INSTITUT NATIONAL DES LANGUES ET CIVILISATIONS ORIENTALES

École doctorale n° 265

Langues, littératures et sociétés du monde

Unité de recherche CRLAO (CNRS UMR 8563, INALCO/EHESS)
UNICOG (INSERM U 992, NeuroSpin CEA/Collège de France)

THÈSE

présentée par

Murielle Fabre

soutenue le 7 décembre 2017
pour obtenir le grade de docteur de l'INALCO
discipline: Linguistique et didactique des langues

The Sentence as a cognitive object

The Neural underpinnings of syntactic complexity in Chinese and French

Thèse dirigée par

M[me] Dan Xu	Professeur des Universités, INALCO
M. Christophe Pallier	Directeur de recherche, CNRS

Rapporteurs

M[me] Naama Friedmann	Professeur des Universités, Université de Tel-Aviv
M. Danqing Liu	Directeur d'études, Chinese Academy of Social Science CASS

Membres du jury

M. Stanislas Dehaene	Professeur du Collège de France
M[me] Naama Friedmann	Professeur des Universités, Université de Tel-Aviv
M. Danqing Liu	Directeur d'études, Chinese Academy of Social Science CASS
M. Christophe Pallier	Directeur de Recherche, CNRS
M. Alain Peyraube	Directeur de recherche émérite
M. Luigi Rizzi	Professeur des Universités, Université de Genève
M[me] Dan Xu	Professeur des Universités, INALCO

Un homme qui cherche la vérité se fait savant ; un homme qui veut laisser sa subjectivité s'épanouir devient peut-être écrivain ; mais que doit faire un homme qui cherche quelque chose situé entre les deux ?

Robert Musil, *Der Mann ohne Eigenschaften*, 1940

The Sentence as a cognitive object.
The Neural underpinnings of syntactic complexity in Chinese and French

Pre-handle a Preamble

The right to write what is right to write with my right that writes

> ABER WARUM SCHREIBST DENN DU?
> —*A.: Ich gehöre nicht zu Denen, welche mit der nassen Feder in der Hand denken; und noch weniger zu Jenen, die sich gar vor dem offenen Tintenfasse ihren Leidenschaften überlassen, auf ihrem Stuhle sitzend und auf's Papier starrend. Ich ärgere oder schäme mich alles Schreibens;* ***Schreiben ist für mich eine Notdurft, —selbst im Gleichnis davon zu reden, ist mir widerlich.***
> ABER WARUM SCHREIBST DU DANN?
> —*A.: Ja, mein Lieber, im Vertrauen gesagt:* ***ich habe bisher noch kein anderes Mittel gefunden, meine Gedanken los zu werden.***
> —*B.: Und warum willst du sie los werden? —A.: Warum ich will? Will ich denn?* ***Ich muss.*** *—B.: Genug! Genug!*
>
> FRIEDRICH NIETZSCHE, *Die fröhliche Wissenschaft, Book II:93, 1882.*

This manuscript will carry the un-concealable and undeniable mark of *not having been written* for the two years where writing and typing have been impossible for medical and surgical reasons. This period of longing to write allowed me to take some distance from the matter that will be dealt in these pages and offered me a perspective on some scientific intellectual automatisms and more broadly on the long-standing issue of pluri-disciplinarity in knowledge. This lapse of time metamorphosed this reflexive latency into the opportunity to think about what it means to do research. The last part of the manuscript – the epilogue – will gather several thoughts I finally decided to share in this

academic writing.

As it can happen in the field of music, the MANUscript you are going to read qualifies as an *oeuvre à plusieurs mains*[1], in that it was concretely written with the help of multiple hands. For this reason, I would like to dedicate these very first pages to all the people that helped me type it, both in a *direct way*, by typing[2] and *in a not so indirect way*, by healing my hands so that I could finally, after so much longing, be liberated as expressed by Nietzsche's epigraph and type it down[3].

My intention is to thank them with a short digression on what it means to write, compared to talking or speaking – what I actually did a lot to them. This could also be considered as an encouragement to make them write down all the interesting things they think and shared with me, which could definitely benefit the two-healthy-hands large population that could not otherwise have the chance to encounter with them.

> Mais pourquoi écrire si la parole existe?
> C'est que **l'immédiat, ce qui jaillit de notre spontanéité**, fait partie des choses dont nous n'assumons pas intégralement la réponsabilité parce que cela ne jaillit pas de la totalité de nous-mêmes, c'est une réaction toujours urgeante, pressante. [...] Par la parole nous nous rendons libres, libres à l'égard du moment, de la circonstance assiégeante et immédiate. Mais la parole ne nous recueille pas, pas plus qu'elle nous crée. [...] Et c'est de cette déroute intime, humaine - pas d'un homme en particulier mais de l'être humain -, que naît l'exigence d'écrire, **on écrit pour regagner le terrain sur la déroute continuelle d'avoir longement parlé.** [...] **écrire à cause de cette intime nécessité de se délivrer des mots, de l'emproter totalement sur la déroute subie** [...]. Les mêmes *mots* auront, dans l'écriture, une fonction différente, ils **ne seront pas au service du moment oppresseur**; ils serviront pas à nous justifier devant l'attaque du momentané [...].
>
> Sauver les mots de leur vanité, de leur vacuité, en les durcissant, en les forgeant durablement, c'est ce but que poursuit, même sans le savoir, celui qui véritablement écrit. [...]
>
> Ecrire ce n'est ni plus ni moins le contraire de parler; on parle dans l'urgence d'une nécessité momentanée, et ne parlant nous nous constituons prisonniers de ce que nous avons énoncé, tandis que **dans l'acte d'écrire résident libération et permanence**.
>
> Sauver les mots de leur instantanéité, de leur être transitoire et les conduire par notre réconciliation vers le perdurable, c'est la tâche de celui qui écrit.
>
> Ce qui se publie l'est pour quelque chose, pour que quelqu'un d'unique ou bien un grand nombre de personnes. Parce qu'ils l'ont su, vivent en le conaissant, pour qu'ils vivent d'une autre façon après l'avoir appris ; **pour libérer quelqu'un de la prison du mensonge, ou du brouillard de l'ennui, qui est un mensonge vital.**
>
> Maria Zambrano , *Pourquoi on écrit. L'inspiration continue. Essai pour les perplexes.* pp. 19-22, 2006.

This manuscript has been *typed* in order to meet three types of requirements and many constraints. It was written in first place for academic reasons, *pour obtenir le grade de Docteur*[4] as stated on the cover page.

Secondly, it was written for the pressing need of being freed from the thoughts that have been dwelling in my mind in these Doctoral years. As magnificently expressed in the text of the philosopher Maria Zambrano (see excerpt on this page and next page footnote for the English version), I had the intimate necessity of *being delivered from the spoken words* that had stated carrying me away.

Thirdly, this manuscript was also written for society – its ultimate and majority financial investor and supporter. However, the common income tax payer represents a type of reader that would unfortunately not have enough background to penetrate the terminological fog and specialization in which science is immersed today, in order to literally enjoy the *thrill and inebriation* of its scientific content. We then thought he would enjoy some other reading connected to my doctoral work: all the long and short quotes and epigraphs are dedi-

1. A work written by more than one hand

2. Bénédicte Tard, Anna Zyw, Christian Popa, Stefania Di Tommaso, Signe Hässler-Andrieux, André Fabre.

3. Special thanks go to Michèle Bivert, Didier Ngo, Philipe Terrade, Gabrielle Vo-Dominé, Georges Nguyen Van Duc, Emilien Vernet, Valérie Vuillemain, Didier Leclerct, Jean-Paul Billot, etc.

4. In order to obtain the Ph.D. degree.

cated to him. Nonetheless, the society he represents deserves, if not being delivered from 'the jail of lies' *prison du mensonge* – who would dare saying something like this can simply be achieved in a Dissertation –, at least being 'delivered from the foggy prison of boredom', *brouillard de l'ennui*[5].

Thus, for the bravest common reader that would feel like going through it – and for the academic reader too – I prepared another kind of enjoyable entertainment: a words treasure hunt which runs through the lines and the linguistic examples. Incongruous lexical items (i.e. linguistic dimension), almost all low-frequency words (i.e. psycho-linguistic dimension) were chosen by acquaintances, colleagues, friends, bar servers and common tax payers to be inserted in this academic writing. This was meant to give a concrete expression to the participation of society to it. So that if the reader is momentously less enthralled by gaining greater knowledge, or enjoying the witty wisdom of quotations (which offer sometimes a way out of some of our epochal problems), he might wander at the end of the manuscript, B at page 821, and find the list of the words to be found in the treasure hunt. We can therefore say that this manuscript meets up several constraints: (i) physical constraints following surgery, (ii) disciplinary constraints, given its pluri-disciplinarity, (iii) "social" lexical-stylistic constraints, and (iv) "epochal" constraint of saving knowledge from instantaneity.

Thereby, we already own the reader an apology for the length of this manuscript. As *saving words and their meaning from instantaneity* is one of the broad vocations of writing – and maybe by excellence the aim of academic and scientific writing, this kind of writing can't be a matter of verbal instantaneity – just try to dictate your Dissertation to a computer or a human to see. Instead reading can easily become an affair of consumerist instantaneity. Hence, to prevent the reader from falling into the post-modern cultural trend for instantaneity, we gave this manuscript the shape of an antidote to instantaneity, with some long initial frame-setting reflections due to a long-standing French academic tradition, this may probably train the readers' patience and prolong the era of 'non-brevity'. The speedy reader will forgive this lengthiness that is the indelible mark this manuscript will carry, for not having been written for two years. This period without hands immersed me in some deep thoughts about the 'fragmentarity' of the academical world, where pluri-disciplinarity is still an ongoing fight.

5. We reproduce here the quotation from Maria Zambrano that is in previous page. " But then why write if speech exists? What is immediate, what springs out of our spontaneity is amongst those things we don't fully take responsibility for, because it does not spring out of the entirety of ourselves, it is a reaction, always urgent, insistent. [···] Through speech, we free ourselves, we are free from the moment, from the circumstance, immediate and besieging. But speech doesn't receive us, no more than it creates us. ··· And from this intimate, human debacle –not of one particular man, but of all human beings – is born the urge of writing, one writes to win back the ground on the continuous defeat of having talked at length. [···] we write because of this intimate necessity of delivering ourselves from words, of totally prevailing over the defeat we endured. [···] The same words will have, in writing, a different function, they won't serve the oppressing moment, they won't be used to justify ourselves in front of the "temporary"'s attack. [···] Saving words from their vanity, from their vacuity, by hardening them, by forging them lastingly, this is the aim pursued by he who truly writes, even though he himself doesn't know it. [···] Writing is more or less the opposite of speaking; we speak in the urgency of a momentary necessity, by speaking we make ourselves prisoners of what we enunciate, while in the act of writing reside both liberation and permanence. Saving words from their instantaneousness, from being transitory, and guiding them through our reconciliation to being perdurable, that is the task of he who writes. What is published is for something, for someone unique or for the greatest number of people. Because they have known it, because they live while knowing it, so that they can live it in a different way after knowing it; in order to free someone from the prison of lie, or from the fog of boredom, which is a vital lie." Maria Zambrano, *Why do we write. The continuous inspiration. Essays for perplex people* (2006:19-22).

All in all, some pages of this work, if not all, would like to be antidotes to these two types of post-modern cultural trends: fragmentarity and instantaneity. In short, *e pluribus, cognoscere unum* will be the manifesto of this *manuscripted fragment of knowledge* you can now hold in your *hands*. It is our greatest pleasure to now share it.

Prologue

Each generation doubtless feels called upon to reform the world. Mine knows that it will not reform it, but its task is perhaps even greater. It consists *in preventing the world from destroying itself*. Heir to a corrupt history, in which are mingled fallen revolutions, technology gone mad, dead gods, and worn-out ideologies, where mediocre powers can destroy all yet no longer know how to convince, where intelligence has debased itself to become the servant of hatred and oppression, this generation starting from its own negations has had to re-establish, both within and without, a little of that which constitutes the dignity of life and death. In a world threatened by disintegration, in which our grand inquisitors run the risk of establishing forever the kingdom of death, it knows that it should, in an insane race against the clock, restore among the nations a peace that is not servitude,

Reconcile anew labour and culture, and remake with all men the Ark of the Covenant. It is not certain that this generation will ever be able to accomplish this immense task, but already it is rising everywhere in the world to the double challenge of truth and liberty and, if necessary, knows how to die for it without hate. Wherever it is found, it deserves to be saluted and encouraged, particularly where it is sacrificing itself.

Albert Camus, *Nobel discourse*, 1957

Ph.D., an answer to our epoch

This PhD manuscript was written in a precise moment in time, and it is to this time, my and our epoch, that I would like to dedicate this prologue. As classically required by a *Prologue*, this digression is dedicated to inscribe this work in what happened before its narrative time, or setting the situation prior to theatrical action. Hence, by a short focus on our epochal background, it will express the *ideal* underlying this manuscript.

Leaning back on the ancient narrative strategy of Greek Drama, in this first scene of the play, we will adopt the voice of the hero's monologue, borrowing the words of one of the most notable cultural hero of the last century, Albert Camus, to depict its fundamental cultural ideal.

In the Medieval ages, the habit of honoring the personality to which the play was dedicated was introduced in the first scene. With this prologue, I would like to dedicate this work to our challenged epoch and to the generation that still has to build it and sacrifice for it.

> Chaque génération, sans doute, se croit vouée à refaire le monde. La mienne sait pourtant qu'elle ne le refera pas. Mais sa tâche est peut-être plus grande. Elle consiste *à empêcher que le monde se défasse*. Héritière d'une histoire corrompue où se mêlent les révolutions déchues, les techniques devenues folles, les dieux morts et les idéologies exténuées, où de médiocres pouvoirs peuvent aujourd'hui tout détruire mais ne savent plus convaincre, où l'intelligence s'est abaissée jusqu'à se faire la servante de la haine et de l'oppression, cette génération a dû, en elle-même et autour d'elle, restaurer, à partir de ses seules négations, un peu de ce qui fait la dignité de vivre et de mourir. Devant un monde menacé de désintégration, où nos grands inquisiteurs risquent d'établir pour toujours les royaumes de la mort, *elle sait qu'elle devrait, dans une sorte de course folle contre la montre*, restaurer entre les nations une paix qui ne soit pas celle de la servitude,
>
> réconcilier à nouveau travail et culture, et refaire avec tous les hommes une arche d'alliance. Il n'est pas sûr qu'elle puisse jamais accomplir cette tâche immense, mais il est sûr que, partout dans le monde, elle tient déjà son double pari de vérité et de liberté, et, à l'occasion, sait mourir sans haine pour lui. C'est elle qui mérite d'être saluée et encouragée partout où elle se trouve, et surtout là où elle se sacrifie.
>
> Albert Camus, *Nobel discourse*, 1957

As the tradition of Greek Ancient Comedy and Theater[6] imposes, the entire world and its vicissitudes are invited to come into play in the Prologue, and this will reveal its cultural, social and political dimension, in a time where Camus' words seem to be prophetically written for our generation. They indeed echo – *mutatis mutandi* – with an incredible vividness the circumstances we are now facing in Western Europe.

These words naturally contribute to inscribe this pluriannual intellectual work in the spirit of sacrifice Camus is addressing here. For my generation, as for many others to come, deciding to pursue an intellectual or cultural work means to commit to more than an individual choice.

Today, the wide scope ideal of "preserving the world from being taken apart" – *empêcher que le monde se défasse* – can find a concrete expression in the intellectual challenge of *reconciling work and culture* through a PhD program, to end-up in reconstructing a transitory disciplinary unity out of the post-modern academic fragmentarity. In this sense, a PhD – and a certain way of doing Research – can qualify as an answer to our epoch and its burdens.

6. Not forgetting that Greek theater was born at the same time as democratic thinking. Its ancient form belonged to the socio-political institutions of the Greek Polys. Greek Theater as a narrative form was also chosen here to underline that PhD can have a social and political dimension.

Universitas scientiarum

When the world experiences the *urgence* it is facing now, the luxury of an intellectual occupation like a PhD – even running against the time dedicated to one's youth –, could indeed seem on the surface a practical way to escape the responsibilities imposed by our historical configuration. However, as expressed by Camus' words, reconciling work and culture is a real intellectual challenge in the realm of *universities*. In other words, the ideals we will try to *embrace* are the original challenges Universities were facing in their medieval dawning: (i) cultural transmission, by limiting terminological fog, (ii) culture of concretely meeting alterity, by analyzing Chinese language, (iii) dialogue and encounter in a ***pluri-disciplinary dialogue***, as the emblem of the whole research project.

Dialogue – in its etymology (σ'υν λόγω) literally meaning 'through the *logos*'–, also signifies an activity or something happening 'through *Reason*', which is indeed the genuine experience of University. Dialogue allows to experience the fact that despite the post-modern hyper-specialization of academic disciplines, which often makes it difficult to communicate with each other, the ***Universitas scientiarum*** – the university of all types of knowledge – can actually constitute a whole. Each discipline working on different aspects of reality on the basis of a single rationality with its various dimensions, and sharing responsibility for the use of reason through the *logos*. Academic work is in the end a theater play where disciplinary dia-logues *dia-logos* take place, allowing the disciplinary alterities to meet up on a research object.

The intellectual adventure of pluri-disciplinary research, its rocky road, its long and treacherous processes in nowadays intellectual fragmented cultural context has been a path full of pitfalls, a bumpy but exciting trip.

The hardship of carrying out a work which over-arches both Humanities and Scientific quantitative disciplines condensates the original spirit of *Universitas scientiarum* and qualifies this work as "going against the wind" of disciplinary fragmentation[7].

Dialogue structure

Following the spirit of academic dialogue culture, we decided to give the whole manuscript the dialogic narrative shape of a theater play. Thus, we shaped the whole manuscript structure around that of Ancient Greek theater play, hoping this choice will not end up being a pluri-disciplinary *tragedy*.

The fundamental reason for choosing the theatrical narrative form is that theater is the realization on a stage of a thread of dialogues, a narrative form embodying DIALOGUE in itself and one of the cornerstones of Occidental culture and democratic thinking[8] which strongly contrasts with our epoch overwhelmed by debates and hardly alimented by dialogues[9].

---

7. We want here to condensate the original spirit of the *Universitas scientiarum*, that of dialogue through the use of reason: would anyone say that literary disciplines use less human reason than scientific ones? Or even dare saying that a qualitative reasoning is less reasonable than a quantitative one? Finally, we will here enjoy the comfort of filling the gap of the dichotomy distinguishing two approaches to science a 'hard' and 'soft', in French *'molles'* and *'dures'*.

8. Cf. previous footnote

9. We rarely assist to Dialogues and to their capacity of making the two parts encounter, listen to each other and tentatively converge; dialogue is often substituted by debates, where a polite or impolite turn-taking is on stage, a kind of exchange where the two parts leave the scene being unchanged and uninformed by the other.

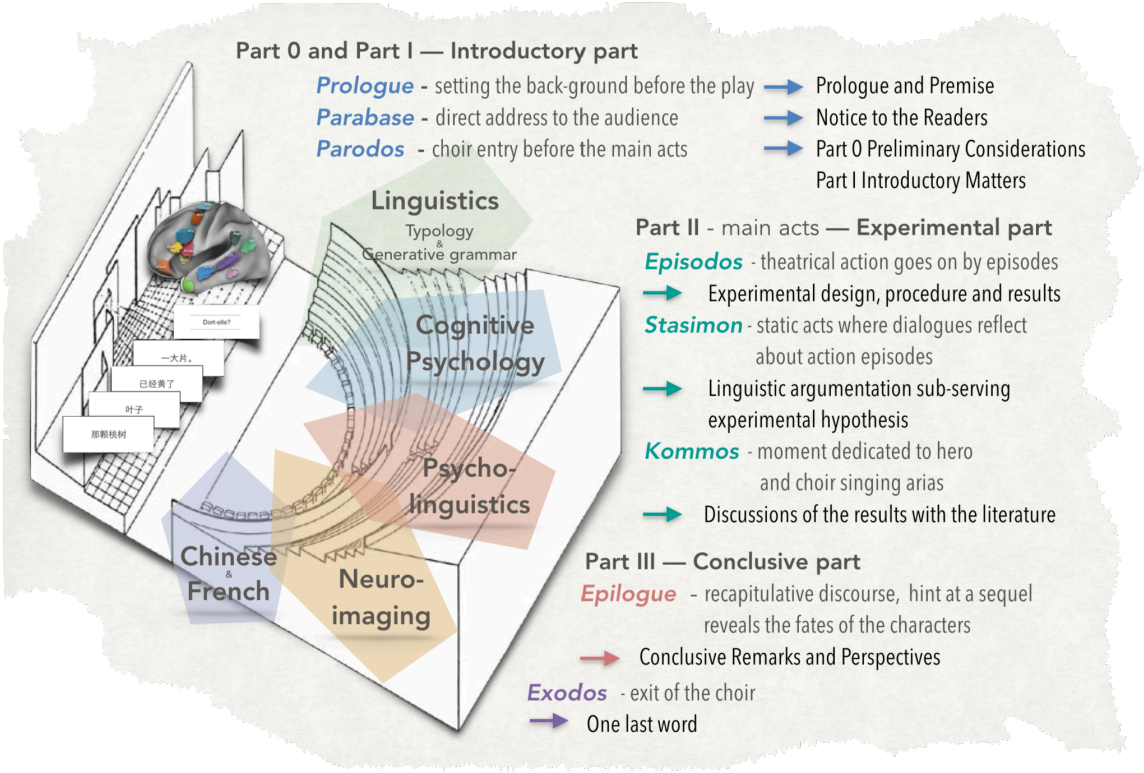


Figure 1 – On the left: A schematic representation of the chorus of disciplines shedding light on the issue of syntactic complexity and its neural underpinnings. on the right: The parallel between Ancient Greek "theatrical structure" and the structure of the thesis manuscript. This configuration will reinforce the Aristotelian rule of Unity of place in theater. PROLOGUE: A monologue presenting the drama's topic. PARABASE: Direct address to the audience. PARADOS: The entry of the chorus, explaining what has happened leading up to this point. EPISODE: The main section of the play, where most of the plot and action occurs. STASIMON: The chorus comments upon the episode to the audience. KOMMOS: The actor and the choir interacts singing arias. EXODOS: The final chorus chant and discusses the moral of the drama..

We hope that such an ancestral narrative form experimentally proven through centuries will be the guarantee of conveying clarity to the reader and constrain the disciplinary tendency to fragmentation. The chapter scenes of this manuscript will hold the type of dialogue that was the basement of academic world: *Disciplinary dialogue*[10].

As in ancient Greek theater, we will do our best to preserve the Aristotelian unity of place. In this enterprise the territory under discussion will be the sentence as a syntactic unit and its neural underpinnings while the different disciplines will be as a Greek choir shedding different lights on the stage, as illustrated in figure 1.

Having taken the firm resolution of avoiding Linguistics-only chapters and Neuroimaging-only chapters, this disciplinary 'chorality' might give an initial impression of untidiness that will gradually disappear as the reader gets acquainted with this dialogic

10. Presenting a scientific work as a dialogue is nothing new, one has just to remember Galilee presenting, for the sake of clarity, its scientific arguments in the form of a *Dialogue* between Simplicio, Sagredo and Salviati in the *Dialogue Concerning the two Chief World Systems* (1632).

shape. However, we should warn the readers that the manuscript will follow this Chestertonian maxim : "If a book is a book to be lived in, it should be – like a house to be lived in – a little untidy"[11]. And I definitely have to say that I have lived in my PhD for quite a few years.

As for structure, figure 1 reproduces the theatrical structure of the manuscript: part II in particular will feature alternations of Episodes, Stasimon and Kommos. This structure appears best suited to achieve a coherent *whole* that would narrowly interweave Linguistics and Cognitive Science to reproduce the real interdependence of the two disciplines at each step of my research activity. We let the house a little untidy, hoping the reader will feel comfortable to hear the weaving loom of a work-in-progress research.

Imaging and images

A second cardinal aspect of this work, inscribing it in our epochal background, will 'unexpectedly' bring us back to the dawning of European culture in its Middle Ages, an epoch where images played a central role in the transmission of knowledge. The colored stained glasses filling the windows of European cathedrals – so-called *Biblia Pauperum* – were telling stories and vulgarizing through images.

Similarly, the advent of neuro-imaging methods and the proliferation of images of brain activations in scientific vulgarization reveals that the medium for Research transmission has changed in nature. Research communicates less and less through writing, discourse or digits, it uses *images* to explain and vulgarize research knowledge.

This apparently banal shift is actually a huge one. It represents a fundamental paradigm shift in that, images, contrary to discourse, text or digits, offer an immediate and intuitive medium summarizing research results, that can be, in some aspects, transparent to anyone. This turn-over or revolution has of course myriads of consequences at different epistemological levels[12] – both on the way research is done by the *"initiés"* and on the way society looks at it or speaks about it. However, this is not the place for a thorough analysis of the science-society interface, and entering the half-religious labyrinth of scientific neo- or post-positivist discourse might be more than perilous here.

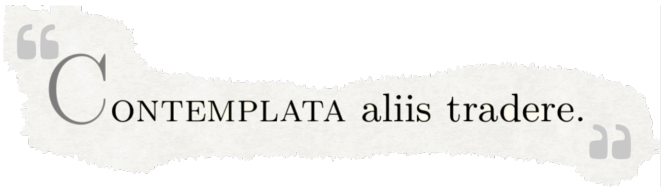
Nonetheless, one aspect should be brought to light concerning this image-producing scientific revolution : it reintroduces a typically medieval activity, fallen into desuetude, that of ***contemplation***. Images, real ones or mental ones, are the object of sight and of a particular intellectual activity called contemplation. Hence, the images that came out from the experiments I run during this PhD were of course first of all the object of 'scientific contemplation'[13].

This recent image-producing scientific paradigm shift offers an additional reason to consider this work as rooted at the dawning of universities, when their founding fathers expressed its vocation as to transmit to others the object that had been contemplated :

11. G.K.Chesterton, in *Fiction as Food*, part 2

12. Roland Barthes summarizes it these terms the ambiguous relationship across ages to images : "Some think that images are a rudimentary system compared to language while others think that signification can not exhaust the ineffable richness of images." In French: "les uns pensent que l'image est un système rudimentaire par rapport à la langue et les autres pensent que la signification ne peut épuiser la richesse ineffable de l'image." from 'Rhétorique de l'image', in *Communications*, n°4 1964. Recherches sémiologiques.

13. For a definition of this concept see the Epilogue §III, p.699.



This motto is reinterpreted here as the necessity to transmit what was generated essentially through images in this neuro-linguistic 'adventure'.

When writing this Prologue and looking back to my academic itinerary, I suddenly realized I might have unconsciously come to neuro-imaging because of its intrinsic image-producing-methodology: images are such extraordinary cognitive objects, like sentences! Images and sentences will be here united for the better or for the worst: what the jury and the reader are essentially invited to see – and I hope to *contemplate* – are indeed brain images and sentences accompanied (i) by the theory and literature reviews that built the hypothesis (and methods) which generated them and (ii) by the interpretations we constructed on them.

Parabase - Notice to the readers

Il n'y a de terrible en nous et sur la terre et le ciel peut-être que ce qui n'a pas encore été dit. On ne sera tranquille que lorsque tout aura été dit, une bonne fois pour toutes, alors enfin on fera silence et on aura plus peur de se taire.
Ça y sera.

*[There's nothing terrible inside us or on earth or possibly in heaven itself except what hasn't been said yet. We will not be easy in our minds until everything has been said once and for all, then we'll fall silent and we'll no longer be afraid of keeping still.
That will be the day.]*

L. F. Céline, *Voyage au bout de la nuit*, 1932.

This manuscript is intended for different types of readers. Thus, following Ancient Greek Comedy rules, a scene with direct address to the audience is allowed: we will abandon the academic fiction and discuss with the different readers in the style of a *Parabase*.

What is very *peculiar* about this common theatrical device is that it stipulates to pronounce ritual satires and *insults* in a procession in front of the audience's first seats. By injuring and mocking, the theater choir used to warms up the audience and guided it – as its etymology testifies '*walk in front*' – to enter and feel involved in the theater piece[14]. This momentarily adoption of the choir direct audience address, will constitute a humorous way for me to let the reader enter in this manuscript by first discharging on its doorstep the load of everyday life to enter the play. These mild injuries shouldn't be taken too seriously. We hope the reader will play the game and enjoy the witty, sometimes cynical, but always humorous person we ended up in becoming during these years.

14. A short historical note is needed here, part of this is a permanence of the Dionysian religious rituals that are at the origins of the Greek Comedy

I will commit to this tradition with humoristic *impertinence* in the next pages dedicated to the notice to the reader. A handful of small scientific but also witty introductory comments to what the reader is going to read are dedicated to the different possible readers of this manuscript, to guide them on how to use, scroll and navigate through this 'letter forest'. Itineraries through the structure have been constructed to answer to the different questions and purposes of a pluri-disciplinary and *multivariate* audience. Feel free to scroll through the serious-funny-off-the-topic-demystifying-sometimes-cynical-scientific small paragraphs.

Thanks to the following paragraph titles, the reader will decide what kind of humoristic, serious, ritual, cynical and scientific pieces of advice and admonitions he wants to read.

The reader should be here informed that there exists an uncensored version of this *Parabase*, and of the manuscript as a whole, that is free of charge under demand.

Hungry Souls

Throughout this manuscript efforts are made to maintain a reasonable balance of disciplinary terminology. However, whatever background a reader would claim to come from, he will find, at the beginning of chapters and sections, short quotations, or 'gobbets'[15] of wit and wisdom. Meaning hunger is what best distinguishes humans, and therefore these literary gobbets are voluntarily addressed to all readers. They contribute to inscribe this manuscript in a long lasting **tradition of epigraphs**[16], a tradition as old as manuscripts themselves (when they were really hand-written). The intellectual exercise is to include quotations about some keywords of the section, chapter, or paragraph they are preceding, in order not to be farragoes. Occasionally, to please and relax the reader from the earnestness of scientific writing, I have dared to give them a freer rein. These quotations may serve a second role to carry on the philosophic part of a PhD, which could also be understood in Foucauldian terms as a "diagnosis of the present epoch", in French terms "*diagnostiquer le présent*[17]", reflected in the way we do research.

Poetry

Poetry will be a leitmotiv throughout the manuscript, not only for occasional digressions to entertain, but mainly because Poetry can be taken as the utmost syntactic capacity of composition of human linguistics capacity, allowing us to make sens of words that actually do not really combine together semantically.

The main idea of this study can be expressed in very few words: the *Syntactic Component* of language faculty is responsible for our capacity of uttering a highly structured linguistic-unit – the sentence – regardless of the the language. This allows to predicate (i.e. say something about something) meaningful information about the world, ourselves,

15. Find here a multilingual translation for the dictionary-allergic reader, gobbets, little meat balls, are boulettes or bouchées, keftadakia, polpettine. In other words, all what we long for after a long day of work

16. Or *Exergue*, from latin *exergum* ("space out of a work" from Ancient Greek ἐκ, *ek* that in front of a vowels becomes ἐξ *ex*, "out of" and ἔργον, *érgon* "work").

17. cf. §Exodos.

and ultimately to be able to combine words to such an extent that humans developed poetry.

Treasure Hunt

If the reader happens to find incongruous words with very low-frequency of occurrence, in the scientific semantic fields concerned by this research, he should not think it is a typing error (although there will be *some*): I asked all the people that witnessed my Doctoral life to contribute with one word that I inserted in the manuscript. In fact their contribution is not only one word: what lies in your hands, dear reader, is indeed not only my work but the work made possible by the existence (or the resilience, if I may say) of human society, institutions, colleagues, friends, physiotherapists, technical manipulators, cleaning clerks, barmen, income tax payers, old teachers, metro voyagers, crazy believers, theologians and family members, poets, linguists, etc. For this reason, this manuscript should honor their lives, as nothing should be taken for granted. In fact, no one could ever dedicate so much time to intellectuality, if this was not supported symbolically and concretely by the shape that was given to our human collective living by civilization. Instead of having a one page list of those words I made the choice of disseminating them across the lines to create a treasure hunt.

The Brain of Brain, a philosophy of the Brain

What is the brain for us living in the 21st century in the Occidental world? This question deserves a whole Dissertation and has been *torturing* me for some years. Here is a short "philosophical" digression about it.

In the immemorial couple of "*same* versus *other*" couple, contemporary philosophy chooses to advantage to the last one. A turnover has been operated in the occidental system of thought starting from the second half of the XX$^{\text{th}}$ century, a shift from the primacy of the *same* to that of the *other*. And the famous French 'philosopher of alterity', Emmanuel Levinas has formulated this point introducing the notion of *"visage"*, face. We can read in *Totalité et Infini* : "La manière dont se présente l'Autre, dépassant l'idée de l'autre en moi, nous l'appelons, en effet, visage. Cette façon ne consiste pas à figurer comme thème sous mon regard, à s'étaler comme un ensemble de qualités formant une image. Le visage d'autrui détruit à tout moment, et déborde l'image plastique qu'il me laisse, l'idée à ma mesure [...] l'idée adéquate."[18] This notion is from my point of view particularly enlightening about the relationship contemporary 'post-modern' man holds with his brain.

One could argue that our contemporary way of thinking is gradually transforming the brain from an organ among others into *"the other of the same"*, or in French terms *"l'autre du même"*. The Brain is gradually acquiring the status of an alterity, a *"visage"*. People indeed train it, feed it, challenge it, sometimes 'change' it, wash it, reset it and even program it. Conceived as the "source-of-the-thoughts-that-can-think-itself", is the brain

18. The way in which the Other presents himself, exceeding the idea of the other in me, we here name face. This way does not consist in appearing as a theme under my gaze, in spreading like a group of qualities that form an image. The face of the Other at each moment destroys and overflows the plastic image it leaves me, the idea existing to my own measure [...] the appropriate image.

nowadays an internally living alterity?[19] Has the brain become a *visage* in Levinasian terms? Probably, as we more and more feel responsible of it. Namely, augmenting one's brain, training one's brain, feeding one's brain, rejuvenating one's brain are among the leading preoccupations of our *paradoxal* contemporary living.

Terminology and notions

As said before, throughout this manuscript efforts are made to maintain a reasonable balance of terminology. Thus, footnotes and an interdisciplinary glossary at the end of the manuscript will lead the readers coming from different disciplinary through the backgrounds taken for granted by other disciplines.

As for specialized and technical vocabulary, this manuscript will try as much as possible to avoid the laziness of abusing of readers' kindness in trying to have intuitions about what is written. We will tentatively do the terminological translation job. As a mark of gratitude, the expert reader of one of the disciplinary satellites gravitating around the research field of this manuscript, will show magnanimity for approximations or infelicitous formulations that are probably to be found in the explanatory diplomacy I had to display in order to bridge disciplines. If the one definition belonging to his field of competence might seem elusive to its specialist eyes, the ones lying beyond his own disciplinary borders might offer him some extra diopter to better focalize what is being talked about.

Lay notions hardly exist, because as notions, they are necessarily the fruit of a first step of abstraction, which, as an abstraction, needs a definition. Lay terminology is a Romantic idea of an idealistic "post-modern pantheism", where everybody would transmit pure meaning without the dirt of language compromise. But if intuitive vocabulary is so perfect and allows such purely direct soul to soul communication, why then study this bodily corrupted machinery that is the brain, and not study the pure uncontaminated deity or even the soul, and shift to Theology.

As for notions we had to say that are many – some may think they are too many – in this manuscript, bu they were essential to interpret experimental data (cf. the importance of theory in experimentation in the Epilogue).

Discovering the narrator's *Weltanschaung* (vision of the world)

Gnōthi seauton or 'know thyself' was carved on the Temple of Apollo at Delphi. The reader will probably witness what it is to be as weird as a no television child of the eighties that had the incredible luck of being given 2.000 years of European and 3.000 years of Chinese culture heritage as her education. I hope readers will be convinced that culture is not a faded living-room wallpaper, but a life. I should probably have become a witty spotter essayist or pamphleteer, but at the time where those were considered real jobs out of which one could make a living, women could not write, or at least could hardly be read. In the language of social media, one would summarize it in "#NoEpochFitsYou". Those interested in my way of looking at the world should read the vast panel of quotes

19. Nietzsche was saying in a more negative scent "Tu vas chercher l'autre uniquement pour confirmer ce que tu es dans son erreur, et tu utilises son soleil pour te chauffer à son erreur." *You're look for the other just to confirm that you're in your own mistake, and you use his sun to warm yourself to his mistake.*

I have been gathering in the last 20 years, particularly Nietzsche ones, as well as the acknowledgments.

Personal Contribution to the fragmentary world

To people willing to know what is my intellectual contribution in this work, I would first answer : Having survived with a brain still at work. This statement might deserve some clarifications and here they are. First achievement: having survived the rational but crazy intellectual challenge of pluri-disciplinarity. This pluri-disciplinary seed was present since the very first burgeoning thought I had about carrying over a PhD research program. Second, the fact of having survived the Humanities vs. Science dichotomy is a second great achievement. This foolish divide lying at the very core of our fragmentary intellectual *Weltanschaung* is still flourishing in the academic world and outside.

In my understanding, this division is a clear mark of decadence, most of all of the intellectual system we live in: no longer capable of generating a unified account of intellectual work, we therefore live in a fragmented world. And, it may be pushing the argument too far, but I would say it has to do with the impossibility to think *truth* as something we can reach for, or as a possible object of human understanding. I will not further develop this intuitive reflection because it would go far beyond the scope of this short foretaste of the *ideal against fragmentarity* that this work embodies.

In sum, this two survival challenges being brought to today's academic defense, are the little but undeniable proof that it is possible today to investigate nature and man under the interdisciplinary lenses, and that disciplines can possibly meet after a great deal of hardship.

Friends

For friends, the footnotes and quotations will be the best miscellaneous of curiosities to make you laugh at or to reason about in my renewed post-PhD social life! Feel free to indulge in the Preliminary Considerations (, page 3), have a look at the Introductory Matters (I, page 77) and do not forget to "contemplate" the figures. Quoting Nietzsche, I'd like to give them an unusual homage: "Me and I are engaged in a dialogue too vehement. How could this be bearable without a friend. The friend is this third that prevents the dialogues between the two to sink in the deepest of the deepest of abysses"[20]. Despite my structural lack of time in these years, my real friends helped me not to sink.

Synthesis is often not the-mother-of-dialogue

Many things could have been synthesized. Who would disagree with this? Personally we subscribe to this remark. However, what the reader does not know is that some of the excerpts, pages, sub-sections, sections in this manuscript were actually added to answer the need of several reviewers. Coming from different disciplinary backgrounds they asked to develop a given point, as they were not catching the ins and outs of a given issue or domain dealt. As for repetitions, we can just advocate that it is the-mother-of-pedagogy.

20. "Je et moi sont engagés dans un dialogue trop véhement, comment serait-il supportable s'il n'y avait pas l'ami, l'ami est le tiers qui empêche le dialogue des deux de sombrer aux abîmes, hélas il y toujours trop d'abîmes pour le solitaire." F. Nietzsche (1844-1900)

Ode to France

If the Anglo-Saxon or European or Chinese readers want to discover a fragment of French intellectual academic spirit, I exhort them to read Part 0, where I commit to the French academic tradition of a preliminary step of notional definitions. Following Antoine de Saint Exupéry last written letter, French civilization appears to be *"un certain agencement des choses"* a "certain layout of things" which we resolutely could not avoid in Part I either.

> Ça m'est égal d'être tué en guerre. De ce que j'ai aimé, que restera-t-il ? Autant que les êtres, je parle des coutumes, des intonations irremplaçables, d'une certaine lumière spirituelle. Du déjeuner dans la ferme provençale sous les oliviers, mais aussi de Haendel. Les choses, je m'en fous, qui subsisteront. Ce qui vaut, c'est un certain arrangement des choses. *La civilisation est un lien invisible puisqu'elle porte non sur les choses, mais sur les invisibles liens qui les nouent l'une à l'autre, ainsi et non autrement*. Nous aurons de parfaits instruments de musique, distribués en grande série, mais où sera le musicien ? Si je suis tué en guerre, je m'en moque bien. Ou si je subis une crise de rage de ces sortes de torpilles volantes qui n'ont plus rien à voir avec le vol et font du pilote parmi ses boutons et ses cadrans une sorte de chef comptable (le vol aussi c'est un certain ordre de liens). Mais si je rentre vivant de ce « job nécessaire et ingrat », il ne se posera pour moi qu'un problème : que peut-on, que faut-il dire aux hommes ?[21]
>
> A. de Saint-Exupéry, "Lettre au général X", 30th july 1944

Hence, even though the manuscript is not written in French, I do commit to this French way of understanding "the invisible links that knot together one thing to another".

Linguists

Those interested in linguistics may enjoy all the linguistic details cognitive scientists don't usually enjoy, about Sentence utterance particles, Topic in Chinese, our study on Scene-setting Topics in French Neapolitan and Chinese, the prosody of Topic constructions, details about Question formation in French, etc. And they will discover how lively is linguistics nowadays, so lively that it is metamorphosing into Bio-Linguistics with all the additional complexity dragged in by the fact that speakers are not only mouths and thoughts, but are *to some extent* cerebrally active too.

21. "I do not care if I am killed in the war. But what will remain of what I have loved? By that I mean not just people but customs, certain irreplaceable intonations, a certain spiritual radiance. What will remain of the farmhouse lunch under the olive trees of Provence, or of Haendel? The things that endure, I don't give a damn. What is valuable is a certain ordering of things. Civilization is an invisible tie, because it has to do not with things but with the invisible ties that join one thing to another in a particular way. Even if we have the most perfect music instruments, mass-produced, where will be the musician? I do not care if I die in the war, or if I get in a rage because of these kind of flying torpedos which have nothing to do with actual flying, and which change the pilot into an accountant by means of indicators and switches (flying is also a certain order of links). But if I come back alive from this "ungrateful but necessary job", there will be only one question for me: What can one say, what does one have to say to mankind?"

Linguists: where are these sentences from?

For the linguists who cherish 'naturality', or for those linguists who will read this manuscript to cherry pick linguistic examples from it: the Chinese Mandarin linguistic data and example sentences from other languages in this manuscript come from three different sources: (i) previous literature, (ii) my own consultations with native speakers, and (iii) purposely tailored sentences by native speakers and myself for our four experimental designs. To the greater degree possible, I have tried to constrain variation by focusing on a specific dialect: the Mandarin spoken in Beijing, the surrounding Hebei province and northern dialects provinces from northern mainland China. My primary informers were two female speakers in their thirties from Beijing and Baoding, Hebei and two male native speaker from Heilongjiang and Shandong. I am extremely grateful to everyone who has contributed with judgments, intuitions, examples and discussion. In particular, I would like to thank them here too: Yao Rongyu, Hu Po, Li Ting and Yan Shaorong.

A special effort has been made to create the 21 back-ground narratives for the Mandarin experiment in chapter 5. We constituted a small group of native speakers getting inspired from various narrative sources. Yan Shaorong, Luo Yingyi, myself, Duan Yuan and Wang Di[22]. Sentences were each time checked against other native speakers from Beijing for naturalness.

Mandarine Chinese(s)

Since Mandarin Chinese is nowadays spoken and written by a wide range of speakers with different linguistic backgrounds, having being raised by speakers with heterogeneous linguistic and dialectal backgrounds, there is an important issue as to the extent of variation in how topic markers like *-ne*, *-a*, *-ya* are used across different dialects and social groups. The previous literature on topic marking in Mandarin Chinese has tended to abstract away from the effects of linguistic variations on Topic markers. One remarkable exception is the deep investigation of Topic markers in Chinese varieties in a book by Xu Liejiong and Liu Danqing on Shanghai dialect's Topics.

Although I will not offer any characterization of the nature of this variation here, it is almost certain that some variation does exist. At the beginning of my research, when I started eliciting judgments about a particular use of *-ne*, *-a*, *-ya* as Topic markers, the first informants would accept the marked sentences although with some resistance, and say that it sounded like something else had been said before the sentence or that they could imagine to say this sentence in a certain particular context or under some precise circumstances. Given that our initial interest for Topic-Comment sentence was related to their non-markedness and 'basicness', we started testing Topic-Comment sentences once the Topic marker was dropped. We did not attempt to categorize these intuitions here, it will constitute an investigation direction for further research.

Regardless of the language of the example, the translations and glosses are generally my own unless specified, this is in order to maintain consistency . In a number of cases, I have added a second translation to the one given in the literature in order to conform to the register of the sentences and make the English more colloquial and highlight the

22. This work must have inspired the last person in the group, because she became six months later a travel journalist.

relation of the sentence to the surrounding discourse-context and the intention of the speaker uttering it.

A Vision of Research

For those interested in my vision of research, the Preamble, the Premise and the Epilogue should be a succinct introductory help to understand the importance I give to theorize the experimental approach. For a more structured outlook, Preliminary Considerations are dedicated to gather everything that nourished conceptually the development of the methodology and the experimental settings presented in Part II. In a nutshell, this research work is a quest against fragmentarity and its broad scope is to reconcile work and culture (cf. the Prologue).

Academic relevance

Those concerned with it academic relevance should read all the dissertation and be free to skip some pages or just look at images, but no footnote can be discarded: the devil is in the details :-). The general take-home message is: the sentence-unit can be studied as a cognitive object considering all the linguistic details that different linguistic frameworks have developed. Different syntactic structures in different languages modulate the cerebral activation of the language network.

Theory-oriented people

Those more theoretical and analytically oriented will find some epistemology in Part 0, and a more formal approach to linguistics in the epilogue. This reader should be delighted to see how much of linguistic theory we can '*trace back*' in the brain, by building crazy complex experimental designs inspired by very simplified versions of formal theories of syntax.

Curious people

Those curious about everything should simply read it all, and skip whatever they want. We would however suggest to this curious reader to concentrate on the first hundred pages and read them as if this was the kind of miscellaneous erudite (but understandable) compilations that were on fashion in the XIX$^{\text{th}}$ century.

Second-language teachers

For those who are responsible for second language teaching, programs or classes, several parts of the dissertation could be viewed as an attempt to bridge the gap between linguistics and second language pedagogy. They will be interested to know what the Chinese natives speakers' brain does when meeting with a common Topic-Comment sentence in chapter 5 and 7 and will probably enjoy the details and quantity of examples present in the manuscript or the 3 experimental corpuses in the Annexes. But, most of all, they will find some raw material to explain in their first, second, third and fourth year classes what is a Topic-Comment sentence articulation in Mandarin Chinese and a

clear minded description and analysis of what is a Hanging Topic by opposition to what is Topicalization.

Linguistic Typologist

Those into Linguistic Typology will enjoy how a typological differences can be brought to neuro-linguistic testing in an fMRI, and be cognitively investigated. Those who try to unite typology and formal approaches to linguistics will find some very initial steps toward a research project uniting their data-driven approach and more theoretically oriented approaches to syntax into experimental projects.

Research's Nostalgic souls

The research world is often described as partitioned into three populations: the nostalgics, the winners and 'normal researchers' in-between the two. The souls nostalgic for a better world of research or for a past Golden Age will realize here - I hope - that nostalgia is only good for *schwärmerisch* German romantics: now has come the moment to work out the best of ourselves to demonstrate that theory is not dead, and that human intellect can do more than big-data auto-revelations and proto-religious emerging epiphanies. In one word, strive for intellect and go back to the broad gaze of Renaissance in a technical postmodern context. To which, cynical Frenchies would answer : *"c'est pas gagné"*, which sounds in English "the task is certainly not going to be easy to achieve". Anyway, drunkard or not, you will find here the cocktail you were longing for : 1/3 theory, 1/3 experiment, 1/3 interpretation. A classical recipe for a good researcher's Campari Spritz.

Scientific distance or the story of the cold souls worshiping neutrality

Yet, on the other side of the research universe lives another anti-nostalgic population: the cold, efficient scientist. To cure his Promethean Post-modern Pride (PPP for those who forgot the definition) we added to this manuscript some ancient traditional remedy: omniscient narration and humor. Recurrent personal comments, omnipresence of the narrator ego, will hold as an aspirin for the classical headaches scientists get by trying to give an impersonal style to their writing. Pretending to be objective and coldly distant with the content is an academic posture, that I would call a reliquary of positivism's waterproof make-up of superficiality. No ode to the religious thinking making quantitative Science as a invariable truth will be found here. This quantified truth is ignoring all the epistemological pitfalls in order for the contemporary scientist to feel Prometheus (at least some minutes per months) or the Guru priest of the purest deity in the world, greater than truth itself, objectivity.

I happen to be of those who think that science is neutral only for the positivist naive ideological thinkers, and that objectivity should go along with epistemological awareness. In other words, intellectual and scientific integrity are better served by the distance of auto-irony and of the salutary philosophical gaze on science, than by the moralistic neo-positivist poison of 'neutrality'.

Take-home message for nice Bourgeois dinners

To bourgeois readers wanting to know what they should say if they had the occasion at a dinner to chat about this research work saying they had read part of this manuscript (nobody will believe this, but you can still try), I would say : "this research work is about how the brain encodes sentence structures in French and Chinese".

Tweet my PhD

For the same social necessity linked to bourgeois dinner, you can find here the 140 characters you need in order to feel a real post-modern hero who promulgates science through social networks - without even having a nice dinner with good wine to accompany your valuable narcissistic mission. Lonely in your room with your telephone, you will perpetrate your communion to the global world by twitting science, disseminating the gospel of post-modern doubt-oriented Reason, feeling the priest of pseudo-positivist religion of science. I will put one additional word so that the tweeter that is in you can express his post-modern arbitration to choose what word to exclude, and feel immensely powerful and free. "Research tells us the brain network for language is modulated by the reading of different grammatical structures, comprising questions and Topic sentences in French and Chinese" or "Come and discover what is your brain like when you listen to French or Chinese different grammatical structures".

Supervisors and an anonymous reviewer

For my supervisors and anonymous *mysterious* reviewers of this manuscript, it would probably be a way to discover more about me, and most of all, a way to discover the hurricane of thoughts I try to master when I talk to them in order not to over-flood them. Probably, I'm mistaken and they already know about this hurricane, and they just never made it appear in our interactions, who knows. Anyway, everything here is probably written thinking about what they would think about it, as it usually is for supervisors, and as it probably should be.

This last consideration about the relation between writing and being read, goes along very well with some philosophy lines concerned with *alterity* written by the French philosopher Emmanuel Levinas (1906-1995), stating that the *Master* – in medieval terms – is a total alterity, in French *"le Maître est alterité totale"*. More specifically, in *Totalité et infini* (1961), he defines in very precise and enlightening terms the kind of alterity an anonymous reviewer played and the position as supervisors gave to my two *Directeurs de thèse*. This quote is dedicated to clarify the cultural and social mission of *Masters* in society:

> "Le visage n'est pas une image, il parle et enseigne et introduit du nouveau dans la pensée, *c'est l'exteriorité qui accompli la liberté au lieu de la blesser, l'extériorité du maitre*. L'autre n'est pas pour la raison un scandale qui la met en mouvement dialectique, mais le premier enseignement."
>
> "The face is not an image, it speaks and teaches and introduces the new into a thought, it is the exteriority that accomplishes freedom instead of hurting it, the exteriority of the master. The Other is not for reason a scandal which launches it into dialectical movement, but the first teaching."

In more personal terms, I would define this intellectual relationship as a helpful distal proximity.

French

Those who like my French writing will be happy to read the English version first, because they should scarily discover that French doubles the length of the manuscript. I will soon write books in French, so, do not worry.

Working language: English or ... the language of Humor

Writing in English obliges to pay respect to the greatest British tradition ever. The most British of all the virtues, so British that hardly any other culture understands it : wit and humor.

As for humor and wit they will be my homage to the language I write in, but for those who like British-English high standards it is more to their understanding I will pay an homage. To mobilize their sympathetic tolerance, I will share part of my language background and my youth idealism on this matter. Before I turned 16, I wanted to speak like the Queen of England, drinking tea with a dash of milk. This dream vanished when I spent a year in a Bostonian High School, where burgers and Wendy's quarter pounds trained my mouth. Those readers will surely notice I decided to express my linguist love for orality in language by keeping abbreviation on auxiliaries in order for the writing to be vivid, and keep the reader awake through these hundreds of boring pages. You may qualify this as a linguist monomania, a reliquary of adolescence, a linguist obsequious or religious pledge payed to language naturality, but it is a choice I assume.

Some academic readers used to reading today's scientific articles will surely find my English writing too colloquial. They should rather spend their time in blaming the English standards of nowadays' scientific articles writing; and maybe feel their responsibility in having given us such a flavorless language in heritage. In fact, the articles written in the 6o's and 7o's were not exhibiting the dullness of today's scientific English.

A light version Chauvinism is glowingly on fashion

Anyway, as I do not write in French, I will be obliged to do the "Frenchy" from time to time, saying for example that this author, researcher or intellectual is French, which does not change anything to his intellectual or scientific contribution, of course, but has the effect of making me feel his or her presence near me and simply making me happy! After all, I heard that, lately, nationalistic chauvinism was on fashion ...

Omnipresent narrator

The most evident thing in this manuscript is that its intellectual and scientific content is the result of the quest of the person I am. My probably obscene presence through these pages is, in first place, a plea for honesty and loyalty to this very first evidence. To understand the "in-between" the lines and the overt claims of this manuscript, one has to bear in mind that here objectivity goes along with incarnation and all the epistemological limits it implies (i.e. cultural background, illness, worries, fears?). Articles, ideas, instead of being just quoted will often be defined as interesting, nice, etc. This will contribute

to avoid banishing the famous gnoseologic ego from science. After all why should we be hiding behind such an objective writing style? Why should we play the game of being impersonal? Is reality UN-personal?

In fact, if all this disqualifies the scientific value of this research work, then science has to be done by machines ruminating data like cows do. And Universities should become unities of productions full of the Londonian smog of "neo-industrial mechanical Cows Utopia", where truths are pornographically automatic, and culture is the cozy bedfellow of some decadent beauties like *production*, *efficiency*, *automation* – as it was the case in the good old times of propaganda and ideology of the 20^{th} century – but definitely *not* the bedfellow of the evergreen freedom.

All in all, a PhD manuscript is the kind of storytelling where the narrator has to be omniscient.

Forefathers and foremothers

For the grandparents who did not really get to know me, they could have discover in this manuscript how society evolved and how the brain became central.

For the grandparents who met me as a child: grandfather could have read it all and added examples in the 7 languages he mastered; and grandmother should have benefited of its results to speak again.

For my parents, I have to thank them for giving me the raw material to become an intellectual.

Intellectual itinerary

Those who are interested in the intellectual itinerary I have been carrying on during this doctoral experience will gradually discover it in Part (chapter 1) and will be interested by the Premise, the Epilogue and Exodos. Astonishingly, these sections are free of the arguments which overcrowded the linguistic and experimental parts of the manuscript and crucially free of Figures. The reader may wonder why such a counter-intuitive piece of advice is given here. The answer is simple: an intellectual itinerary is, as every itinerary, more than the pictures that were taken during the trip. It is best described as an intellectual life and as such only gets inspired by reality (part I and II), without being a total mimesis of it where images take over the role of words.

Omniscient Reader

The author in me cherishes the secret hope of finding an *omniscient reader* of the kind of Pic of the Mirandola (1463-1494). Discussing *"de omni re scibili"*, he will take over the effort materialized in this manuscript and continue it, enriching our contemporary knowledge on syntax and its neural underpinnings across different languages. However, I am aware that this kind of Renaissance *ubris* lies today in the arms of Big-Dataists and this will probably be the kind of reader taking up this challenge, but without discussing it or sharing it, remaining in the shadow of his deep-learning algorithms that sadly do no 'represent' the data they manipulate.

"Terribly long terms"

When the reader will find long agglutinations of words defining a concept, he should be aware that this was purposely done to make the pluri-disciplinary mutual understanding possible. Cases such as 'Chinese style' gapless base-generated in-situ scene setting Topics will be met and each of its term is destinated to a reader coming from a different linguistic or disciplinary background. In this case:

- 'Chinese style': typologists
- Gapless: cognitive scientists
- Base-generated: generativists
- in-situ: naive readers
- scene-setting Topics: general and generativist linguists

Our choice might not be felicitous or aesthetic but it has the tremendous advantage of guaranteeing mutual understanding.

Cognitive Neuro-Science

Those only interested in neuro-science should at least read a few pages with no brain pictures on it. They might feel the fresh breeze that only sophisticated linguistic experimental hypothesis can blow, even under the hot midday sun of statistical zenithal evidence. Doing so they will discover a new definition of man: the homo phraseologicus, a kind of homo sapiens sapiens that is so sophisticated that he can play with languages and do poetry.

Neuro-*blobology* in Neuro-imaging of cognitive functions

If you were happy to read some miscellaneous compendium of the "Blob of this and blob of that" from this incredible scientific new discipline – yes, we can see neuro-blobs (neural cluster of activation) of this and that everywhere –. If you are ever looking for an interesting and insightful proof of the "Segregation of this and that cognitive function somewhere" : remember that it maybe on fashion. Do not get misleaded by the new fancy packaging of brain images... It is since medieval times that blobbology exists (see Figure 2, p.xxxv). Old wine in new bottles.

Humor as *back and forth* thinking device

To dissert about humor, I can't help landing the microphone to G. K. Chesterton (1874-1936), one of my silent and distal intellectual friends:

> Now our modern discussions about everything, Imperialism, Socialism, or Votes for Women, are all entangled in an opposite train of thought, which runs as follows: —A modern intellectual comes in and sees a poker. He is a positivist; he will not begin with any dogmas about the nature of man, or any daydreams about the mystery of fire. He will begin with what he can see, the poker; and the first thing he sees about the poker is that it is crooked. He says, "Poor poker; it's crooked." Then he asks how it came to be crooked; and is told that there is a thing in the world (with which his temperament has hitherto left him

> unacquainted) –a thing called fire. He points out, very kindly and clearly, how silly it is of people, if they want a straight poker, to put it into a chemical combustion which will very probably heat and warp it. "Let us abolish fire," he says, "and then we shall have perfectly straight pokers. Why should you want a fire at all?" They explain to him that a creature called Man wants a fire, because he has no fur or feathers. He gazes dreamily at the embers for a few seconds, and then shakes his head. "I doubt if such an animal is worth preserving," he says. "He must eventually go under in the cosmic struggle when pitted against well armoured and warmly protected species, who have wings and trunks and spires and scales and horns and shaggy hair. If Man cannot live without these luxuries, you had better abolish Man." At this point, as a rule, the crowd is convinced; it heaves up all its clubs and axes, and abolishes him. At least, one of him.
>
> Chesterton, The Man Who Thinks Backwards, p.33.

When I decided to insert the perspective-taking humor offers in this academic production, I soon realized that the language of contemporary scientific articles was not very inspiring for this quest for humor. Hoping that this style will help achieving one of the objectives of this manuscript: the surgical ablation of Positivism Promethean Pride (PPP) from contemporary neuro-cognitive science and its persistent heavy-to-digest American Chantilly of overstating research results. As a Chinese aphorism would joyfully state: "驴子跑不快, 名之为马以枉然" *If your donkey is not fast enough, it is useless to call it horse.* (Taolifu). Humor and perspective-taking will be the surgeon of this monstrously unlikely surgery. Hoping this effort will at least benefit a moment of solace to the reporters and to the occasional reader. Unveiling this occult aim will probably help the reader understand the place held by Nietzsche perspective-taking on science in the main matter first page (page V), a few chapters incipits and in the *Epilogue*.

If the reader finds this frivolous and time consuming he should resolutely enter the fight for a machine world where future PhD manuscripts will be written by report machines, and efficiency will have the last word on art and philosophy.

Politics

If this category of readers still exists at all in our democratic period of decadence, they will enjoy some footnotes on the actual American president and a section on the former French president's favorite syntactic structure. Guess what is was? Topic-Comment constructions in colloquial French!

Too much and not enough

Measure is a hard thing to achieve and this manuscript is a bit too *something...* too extensive, too poetic, too linguistic, too french, too experimental, too eclectic, too cynical, too positive, too vintage, too old-style, too ironic, too existential, too structured (too many titles), too academic, too meta-reflexive, too something. In other words, to read it on should just drop one 'o' TO enjoy it.

Uncensored version

For those who want to see what I can dare writing, I can offer an uncensored version of the manuscript with 50 more pages of post-modernism critique and caustic Confucian name-rectification.

Premise, the necessity of naming

One must name things
with strength and truth,
this strengthens and exalts life.

Thomas Mann, Der Zauberberg, 1924

Following the ideal expressed by Thomas Mann's words, this doctoral work is among others the result of my profoundly fascination for the idea that what makes human life great is the possibility to *name* and know reality [23]. During these years, I happened to understand that one of the central aspects of the scientific investigation is to first *name* things and forge concepts out of observation, to subsequently build on them hypothesis, and then verify the model constructed through this process. A large part of this manuscript is dedicated to this "naming quest" that preceded and followed our experimental approach.

The title of this premise is namely inspired from the work of a contemporary philosopher, Saul Kripke, "*Naming and necessity*" (1980), who developed a theory of identity, where the main thesis is that *names* pick up necessary features of things that are true in every possible world. This philosophical point of view on naming is essential to build the kind of trans-disciplinary approach that this manuscript coveys. The epoch of models and theory blossoming in the experimental disciplines we will be concerned with, has been gradually fading away, and very few models of the implementation of linguistic knowledge in the brain are at reach nowadays. As a consequence, our "naming quest" will have to additionally imply some tentatively scrupulous definition of the different concepts that grounded our approach. The notions called into play in this research will be humoristically presented as *Cooking ingredients* in Part 0 Preliminary considerations. If there had been ready-made neuro-linguistic theories, having already *named* the issues we will address, we would not have had to delineate a rather long (and maybe ram-

23. I could have chosen to rely again on Albert Camus to express the same idea: "Mal nommer les choses c'est ajouter du malheur au monde", *badly naming things adds misery to the world.* But I thought that Thomas Mann's approach was less negative and more centered on a dynamic ideal than the subtle culpabilization lying between the line of this last frenchie quote.

bling) preliminary setting of the field of inquiry and introductory matters, for which we apologize in advance.

We will tell here the double tale of a Linguist in the land Neuro-cognitive Science, and that of a Cognitive Neuro-imager in the land of Linguistics, with as little jargon and technical apparatus as possible, but with a structured theoretical toolbox of notions.

Defining and naming is more about sentences than words

Parsimoniously using words, finding the right words, abolishing empty words, stumbling nonetheless into the wrong words, creating new words, or finally choosing conventional ones will not make a difference in the act of naming and defining. What is needed to name a reality, even scientifically speaking, are certainly good words, but most of all, it is their combination, their arrangement, their relationship, the way we bring them together – which will ultimately determine if we succeeded in defining something. Only word combination can let meaning emerge to describe and *name* things. Hence, the main thrust of this Dissertation, will be that there exists a unit in language, that allows us to say something about things, and this natural unit is the sentence. The human cognitive ability of uttering and understanding sentences will be the object under study across languages presenting a different typology of grammatical relationships among the components of the sentence-unit. This means that the sentence will be here considered as a cognitive object.

Why did I start

The fascinating mystery (it is still one indeed) of linguistic human capacity and its variety of expressions through the languages of the world was the first step for my engagement into this research. The second came after, when I discovered the intriguing possibility that this mystery could have a flesh, a materiality: the **biological actualization of sentences** the brain appeared to be, which allowed such an immaterial entity like language to have an immanence.

The initial idea that sentence-unit and its topic-comment articulations were some basic logical unit of predication came when I entered Chinese Grammar and I had to change my mind on what it meant grammatically to say something about something. All the possible word-orders and combinations that came to my mind when speaking Chinese were initially wrong, my Chinese utterances were impossible sentences; although all the words were right, they systematically needed reformulation. I later learned, I first had to posit a topic and then say something about it.

The main idea of this study can be expressed in very few words: the *Syntactic Component* of language faculty is responsible for our capacity of uttering a highly structured linguistic-unit – the sentence – regardless of the the language. This allows to predicate (i.e. say something about something) meaningful information about the world, ourselves, and ultimately to be able to combine words to such an extent that humans developed poetry.

Poetry and the utmost Syntactic capacity composition

A simple demonstration can show how Poetry can be taken as a typically emblematic example of human syntactic capacity. Poetry is indeed the human habit (or cultural

tradition) where a sentence has to be structurally perfectly correct and semantically absurd. Take a sentence like the French sentence *"Le signal vert indique la voie libre."* "the green signal indicates free way" and replace each lexical word by another of the same category following the original in the french dictionary alphabetical order and you will end up with "Le silence vertébral indispose la voile licite" *the vertebral silence aggravate the licit sail*[24]. Through this lexical shift the sentence's structural scheme stays intact, moreover interpretable meaning or referentiality is blurred or lost like in futurists' or surrealists' French poetic masterpieces. As announced in the Notice to the reader (*Parabase*), poetry will be a leitmotiv throughout the manuscript.

Defining the wider Scientific question

By defining human linguistic capacity in chapter 1 we will sketch the characteristic of what we will call the *Homo phraseologicus* and thereby demonstrate that human linguistic capacity is essentially described by the ability to utter sentences. Although, saying that syntax is responsible for order and structure in the Language *module* is nothing more than a glose of its etymology, we will use this term throughout the manuscript to refer to the hierarchical properties that are attributed to syntactic representations of the sentence. Given the cardinality of syntax in the human language faculty, the simple definition of Man as an *Homo phraseologicus* could be seen by many as a complete truism, if one did not attempt to account for syntax's cerebral underpinnings too.

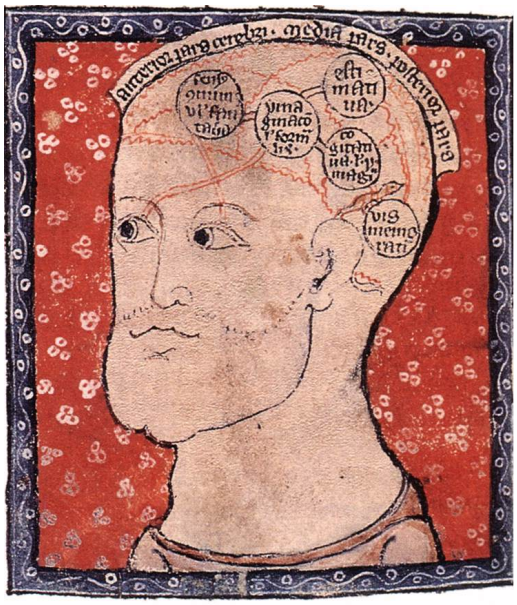


Figure 2 – *Diagram of the brain*, illumination on parchment dating back to around 1345. Dimensions: 21,7 x 14,2 cm (entire page). University Library, Cambridge.

Hence, this complex, and nonetheless universal, linguistic-unit that is the sentence, becomes the object of inquiry of a discipline designated here by the name of neuro-syntax, asking how and where sentence hierarchies are represented and processed by the mind and the brain. This discipline is still at its infancy in terms of the epistemological itinerary that has been accomplished with neuro-image compared to the state of the art depicted in the illumination (Figure 2). One should admit that we are, even nowadays, not far from what the medieval man was doing when he painted a mapping of cognitive functions onto the brain[25]. This drawing offers a glimpse on how

24. Tesnière, in *Elements de syntaxe structurale*, chapter 20, parag. 17, pp. 41-42. An example that astonishingly near to the chomskian one *Colorless green ideas sleep furiously*, except it anticipated it by approximately 20 years.

25. The inscription on the top of the head shows already a clearly tripartite approach to the brain, one can read: anterior, middle and posterior part. The circular modules depict different cognitive functions - faculties at that time - attributed to different parts of the brain in Medieval Times. In the frontal

European medieval men understood the part the brain was playing in human perception of the world. Nowadays, it would be impossible to explain the totality of cognitive processes we map on the brain with such labels as *"sensus communis vel sensatio"*, *"ymagination vel formalis"*, *"estimativa"*, or *"cogitativa vel ymaginativa"* et *"vis memorativa"* (see footnote for translation) because we use finer-grained labels. However Brain Imaging did not change the epistemological paradigm.

These labels reflect the conception Medieval times had of the presumed mental or cognitive processes that could be attributed to certain brain areas. As a matter of fact, this Medieval illumination allegorically summarizes one of the scientific goals of this research: *mapping onto the brain the linguistic faculty for understanding complex* ***sentences***[26].

Moved by the same mapping attempt[27], this manuscript will feature the discovery trip of a linguist in the land of the brain. Thanks to the contemporary technical advancement of fMRI and EEG techniques, this trip benefited from brain maps in order to be able to trace back in the cerebral kingdom some aspect of sentence syntactic complexity in French and Mandarin Chinese.

The way we will dig into this kingdom deserves to be briefly considered. To avoid any sort of equivocation, one has to bear in mind that what is here at stake is not a neo-phrenologist attempt to have a one-to-one correspondence between different brain areas and the way the brain processes, encodes and therefore represents sentence's structures across different languages. The main claim of this work is that the level of description offered by linguistic and its theory oriented fine-grained distinctions is a better guide to investigate the instantiation of syntax in the brain. The sentence will be considered from the point of view of the cognitive representations and processes it yields in the brain[28].

For a Science of Polyphonic Definitions

Taking the risk of being redundant, I continue here a reflection on naming and terminology. Hoping that the reader will not find this a pedantic french way of writing a thesis, it remains that the interdisciplinary aspect of this research deserves an accurate

---

part we see two functions linked to the eyes by the chiasma of optical nerves, respectively *"sensus communis vel sensatio"* (common sens and sensations) and *"ymaginatio vel formalis"* (imagination and shapes). The denomination *"formalis"* indicates here, the shapes and the image patterns. The middle part hosts modules named by *"estimativa"* (estimation) and that of *"cogitativa vel ymaginativa"* (cogitation and imagination). The posterior portion is described by medieval specialists in general as housing the *"vis memorativa"*, in other words the memory force to which they used to add a dynamic element, the cerebellar vermis, whose task was to open and close the channels accordingly to the needs of thinking activity. This dynamic instance is here represented as a small lizard right above the ear. We can note, *en passant*, that, already at that time, the mapping of cognitive functions into the brain was hypothesizing what will be later called a *modular* way of understanding the functional architecture of the mind, featuring mental processes as interactions of different *faculties* (see J. Fodor (1983), *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass.: MIT Press.)

26. Note that one could actually argue that mapping cognitive processes or representations onto distributed neural assembly is fundamentally different from mapping cognitive functions onto brain areas.

27. One has to note here that the largely shared idea that scientific investigation of the material substrate of superior cognitive functions started in the XVIIIth century, or later, is proved wrong by this illumination. Medieval men were not so *'archaic and outmoded'*, they did already think that cognitive functions had a material substrate!

28. Given the title I chose, the question of whether or not a *module* (i.e. a functionally specialized cognitive system), exists for sentences is of course far beyond reach but can still be seen as a silent thread of the discussion carried out between these lines.

presentation. In fact, forging an interdisciplinary gathering around some neuro-syntactic issues participates to the clear delineation of a type of experimental approach and this may actually constitute one of the main academic contributions of this doctoral work.

To do so, we introduce here the concept of *Polyphonic Definition*. Compared to medieval times, today we not only use brain imaging to map linguistic cognitive functions onto the brain, but, we also investigate the brain through a complex interplay of disciplinary backgrounds, encompassing linguistics, psychology, cognitive science, neurology, aphasiology (see Figure 3).

These disciplines are called into play to gain a complementary outlook on such a complex object of research as the brain. However, these different disciplines do not simply gather around the issue of the cognitive mapping of the brain, but they crucially allow three different types of interrogation to be carried out in the field of cognitive sciences: (i) the 'what' question, an interrogation about the nature of the linguistic phenomena under examination, (ii) the 'how' question, corresponding to the question about the nature of the mental representations and operations underlying sentence understanding and language faculty, and last but not least (iii) the 'where' question, which raises the question of cerebral networks, local or distributed, that underpin linguistic behavior for sentences.

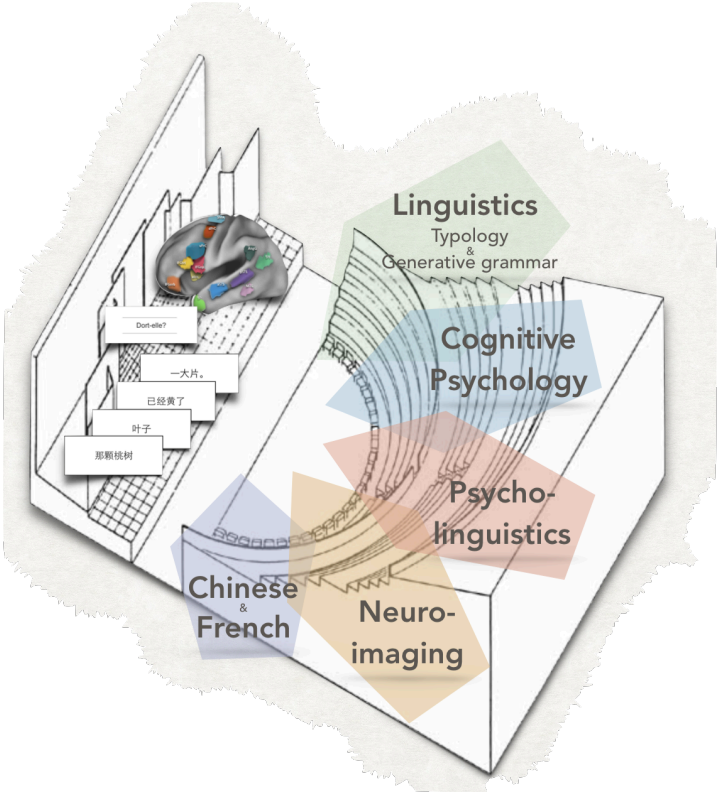


Figure 3 – Disciplinary chorality to have a complete overview on Syntax and the Brain.

Hence, the idea lying behind the word PLURI-disciplinary is that of a multiplicity of voices – in one word that of a *chorality* or *polyphony* – gathering disciplines in a harmonic configuration, where nobody's voice is canceled at any moment, where round and sound harmonics need the contribution of the timber of every particular disciplinary voice to depict reality. This will ideally lead to the unity of the research polyphony by which we tried to depict sentence syntactic complexity in the brain[29]. Hence, chorality through disciplines will also go along with chorality through different descriptive and formal linguistic traditions trying to have them converge toward a reconciled linguistics chorus[30].

I happened to understand that terminology problems were hardly ever linked to

29. This highly risky configuration and this polyphonic aspiration of human tragi-comedy is perhaps wisely relegated by a medieval connoisseur of human condition – Dante Alighieri – to a beyond worldly experience in his Divine Comedy's Paradise. For the moment, we are bodily and cerebrally down here, I will nonetheless try to offer some fragments of *paradise* harmony through these pages.

30. It does sound better in French: *"vers une linguistique réconciliée"*.

words, but came from the difficulty of defining the object under investigation (i.e. level of analysis, biological implication, representations vs. processes etc.). It appeared clearly that a 'straight to the point attitude' is hardly ever a mark (i) of synthetic thinking, (ii) of precision and clear-minded thinking, but rather a disciplinary automatism, on an object that is not questioned, probably because not understood, or even worse because it is to difficult to define. The myth of *simpl-exity* is indeed a myth[31].

Delineation of our research approach

In sum, the intellectual project and challenge of this PhD[32] is to dare avoiding a bi-disciplinary research where Linguistics would be called into play to have some nice and somewhat sophisticated experimental hypothesis and then would be ostracized and relegated to the role of an inspiring muse to adore or to blame when results are not on cards (and to pray again to inspire a nice 'story' out of incomprehensible imaging results). No, the research project lying at the core of the present work is of mutual continuous disciplinary dialogue gathering around syntactic structures as a linguistic phenomenon to be understood, described, analyzed, modelized and tested experimentally. In more methodological terms, rather than pursuing an approach in which linguistically postulated descriptions are chosen to be then validated by neuro-biological data[33] , we will try to build on a position where, taking linguistic descriptions and categories seriously – both from typological and generativist theories –, we will use them to carry on a reflection on what are the abstract categorical representations on which the brain computes the sentence-unit, before going to experimentation.

To engage in this hard enterprise, I had behind and beside me two incredibly open-minded Supervisors who never thought that something was useless or 'stupid' just because it was coming from the other discipline[34]. The acknowledgments will be the place to thank them, but I would still love to say that without their help and example this manuscript and the research approach it embodies would not have seen the light of day.

31. Throughout this manuscript, efforts are made to maintain a reasonable balance of terminology and methods presentation. The glossary at the end of the manuscript (809) will offer explanations to the readers coming from different disciplinary backgrounds. The knowledge presented here was gradually constructed through description, analysis, experiments and theory, with no presumption of being thorough, but with the exigence of knitting together the babel of detailed terminology, that makes every discipline so advanced but in the meantime also terribly out of reach for other disciplines. This effort will probably be vain because I have started to be deeply disillusioned about the possible success of the kind of pluri-disciplinary adventure I started with great enthusiasm during my Master2. Neurolinguistics is still a discipline to come and the academic world is still a highly disciplinary one. The reader will taste in this part how idealistic a young researcher can be. I hope this vain stroll across the fresh green path of intellectual spring time will be nonetheless enjoyable.

32. Some considerations on our philosophical and epistemological research approach are to be found in the epilogue.

33. A position or a role that I ended up calling "playing the linguist jukebox for experimental psychologists".

34. Xu Dan, my first supervisor, continuously supported me in the fragile equilibrium of being 'in between disciplines', while Christophe Pallier never hid his fascination for the ordered and sophisticated world of Linguistics and Syntax. Another person that arrived at the half of my itinerary was Luigi Rizzi whose elegant magnanimous Italian diplomacy resurrected the hope that there could be communication among the different disciplines and Linguistic traditions I was trying to put into resonance. Later came Naama Friedmans's publications (and in flesh and bones) to show me a way to conjugate my theoretically-oriented research question with the linguistic life of real people, thus renewing my interest for aphasiology.

Outline of the thesis: Structure and scientific questions

Delimitation of the field of inquiry - Part 0

As a consequence of the trans-disciplinarity of our object of study, the first two Parts of this manuscript (Part 0, p.3 and Part I, p.77) will are inspired from the Greek theater *Parodos*: the choir of disciplines constituting the foundation of our approach will enter the scene to expose what precedes what will be *happening* in our research work. Its main aim is to help the readers with different disciplinary backgrounds to enter the *play*, by giving them the common background knowledge the other disciplines take for granted. This will participate to *chorally tune* the divers readers to overcome the natural disciplinary *fragmentarity*.

Cooking ingredients: notions and methods at stake - chapter 1

The reader is walked through the essential notions framing our research interrogations, and through some preliminary argumentation that introduce some central aspects and assumptions that grounded our approach. For instance, chapter 1 delineates (i) the cardinality of syntax in the definition of human language faculty and cognition (§1.1 and §1.3), (ii) the essential aspects of what cross-linguistic variation and its limits can tell us about the mind and how the brain represents syntax (§1.2), (iii) some considerations about the consequences of understanding Linguistics as Bio-linguistics (§1.4), and lastly (iv) a few methodological assumptions linked to Neuro-imaging and Cognitive Neuro-science (§1.5).

Hence, chapter 1 offers some framing of our field of inquiry by presenting the notions and methods at stake in this research project. We humoristically called this chapter in a non-conventionally academic way – *Cooking Ingredients* – to express its clear-cut scope: bring to light, in a strictly non-chronological way, the concerns, issues and challenges of our research project examining *the sentence-unit as a cognitive object*. Seemingly too broad, these Preliminary Considerations will actually be essential in unraveling some of the intellectual and notional pillars that are usually taken for granted.

Introductory matters:a polyhedral point of view on the sentence - Part I

While Part 0 consist in general circumscription of the field of inquiry presenting the interrogations raised by considering *the sentence as a cognitive object*, Part I offers an extensive introduction to the experimental hypothesis for the French and Chinese experiments in chapters 4,5, 6 and 7.

The Sentence as a natural universal structured complex syntactic object - chapter 2

Chapter 2 sets the stage for all the research questions that will be dealt in this manuscript by introducing the theoretical context, the scientific goal of our research, and the main hypothesis and claims of the thesis. Note that this will be achieved through a gradual step-by-step argumentation.

Here, the interweaving of the voices of Psycho-linguistics, Neuro-psychology, Brain-imaging and Linguistics will allow us to delineate the argumentations that preceded the choice of our the experimental conditions. The we will depict, in a pluri-disciplinary

and *choral* fashion, the sentence-unit as a (1) natural, (2) universal, (3) structured, (4) complex syntactic object (from §2.1 to §2.4).

The linguistic theoretical background for our study on French syntactic complexity will be given by presenting the different movement types we included in the fMRI design in chapter 6.

By the end of the last two section of chapter 2 (§2.3 and §2.4), a state of the art of different dimensions of syntactic complexity will have been sketched, so that the reader will already be given the possibility to fill in the blank brain (Figure 4 p.xliii to be cut and carried all along reading) with an initial sketch of the cerebral network that underpins sentence comprehension and the representation of syntactic complexity.

Sentence as Topic-comment articulation - chapter 3

While chapter 2 presented the theoretical and broad linguistic grounding of our experimental hypotheses, chapter 3 outlines the detailed syntactic and psycho-linguistic motivations behind our experimental research on Topic-Comment articulation.

Based on the literature, we introduce the reader to the notions of Topichood (§3.1), the issue of Topic-Prominence (§3.2), and their syntactic analysis. The second part of chapter 3 provides both a typological approach and a formal detailed syntactic analysis of the Mandarin Chinese Topics and Left-peripheral phenomena (§3.4). Moreover, this chapter undertakes an initial study of an observed tendency in Mandarin Chinese for the sole arguments of unaccusative and unergatives verbs to surface to the right of the verb in Topic-comment construction. The particular case of Scene-setting Topics in analyzed in parallel with French and Neapolitan relatively similar linguistic phenomena (§3.3).

As it was the case for chapter 2, along this chapter too, we will gradually build the foundations of the two experimental studies concerned with Chinese Topic-comment constructions, presented in chapters 4, 5 and 7.

Experimental approaches to the Sentence-unit

After having theoretically and neuro-linguistically grounded our experimental hypotheses, Part II enacts a change of perspective, giving the go to the experimental development of the current research on *sentence as a cognitive object*. The *Sentence* will be analyzed by focalizing in each chapter one of the aspects that makes it a particular cognitive object: (1) its prosody and the syntactic information it encodes, (2) its interface with discourse when embedded in context, (3) its syntactic complexity ; and finally (4) its syntactic articulation into a Topic-comment construction and its intra-sentential dependency-links.

Sentence as an oral unit modulated by Prosody - Chapter 4

The last 10 years have been registering a increasing use of automatic synthetic announcements to the travelers in the subway and in train stations, making it quite intuitive for everybody that sentences are more than an ordered sequence gluing together isolated words. The amount of linguistic information that is conveyed by prosody in a sentence is considerable, an in chapter 4, the sentence will be considered as a prosodic-unit. A phono-accoustic analysis will allow us to single out the characteristics of the prosodic pattern marking Topic-Comment sentence in Contemporary Mandarin Chinese. Secondly, the

way these prosodic contours influence the understanding of Topic-Comment sentence's articulation is investigated behaviorally by recording grammaticality and truth-value judgments of native speakers during in context sentence processing.

Sentence as a unit immersed in discourse Context - Chapter 5

The sentence-unit is more than an isolated syntactic-unit, it is immersed in linguistic context, and possesses syntactic means to encode its interface with the discourse-level, as it is the case in Topic-Comment articulations. Chapter 5 offers an experimental in-context investigation of the online comprehension of Mandarin Topic-comment sentences that have no selectional relationship with the main verb. The aim will be to answer the question of how happens the temopral unfolding of discourse-interface mechanisms during online auditory comprehension of Mandarin Chinese Scene-setting Topic sentences embedded in discourse context. Event-related potentials (ERPs) will allow us to track the mechanisms of sentence-discourse interface yielded by the presence of the Topic in the sentence, and to observe when discourse context and its informational load is accessed in 'Chinese style' Topic sentences.

Sentence as a syntactically complex unit - Chapter 6

Chapter 6 will consider the sentence-unit as a complex syntactic object having undergone different types of syntactic transformations, respectively linked to question formation, clitic placement and unccusativity in French. This fMRI study contributes to the investigation of syntactic knowledge in the brain by examining the determinants of cerebral activation to (i) different syntactic movement derivations, (ii) intra-sentential dependencies linked to the presence of clitics and wh-words and (iii) to the presence of multiple empty syntactic elements (not phonologically realized) left by syntactic movement by the displacement of the verb or of an argument.These sentence complexity dimensions, linked to different types of syntactic movements, and their modulation of cerebral sentence network will contribute to the functional characterization of the Sentence *distributed* network, though tradition contrast-based approach and unsupervised analyses.

Sentence as a Topic-Comment articulation - Chapter 7

In the last experimental chapter, the sentence will be considered in its possibility of establishing an interface with discourse thanks to its syntactic articulation. In chapter 7 the rich set of syntactic properties characterizing Mandarin Topic-Comment constructions will contribute to answer the question of how the sentence-discourse interface and intrasentential dependecy-links are achieved by the brain. Through a series of minimally differing contrasts we will examine (i) the cerebral representation of sentence structure's domains (i.e. VP, IP, CP), and (ii) the neural substrates of different *overt* and *covert* realization of intra-sentential dependency-links (resumptives, null pronouns and gaps).

Furthermore, in the perspective of offering neural-based arguments for the linguistic controversy on Base-Generated Topics and Moved Topics in Topic-Prominent language like Mandarin Chinese, we will compare Topic-Comment articulations obtained by base-generation versus movement analysis.

Conclusion and Epilogue

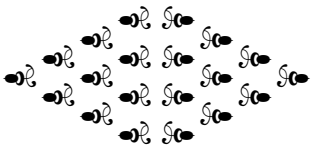
This manuscript ends with a conclusive summary and remarks on the experimental findings of current research. Some research perspectives will also be gathered in this Part III, followed by an Epilogue presenting some epistemological considerations (p.697). This Epilogue will give some emphasis on several aspects we learned during this research work such as: giving priority to experience, what is really implied in experimentation, the cardinality of theory, the importance of harvesting past knowledge, an intellectual reflection on phenomenology, and finally some remarks about this marvelous *chimera* of pluri-disciplinarity that seduced us for several years.

We humbly considered this considerations as a simple or clumsy reliquary of our evolving understanding of what it means to do the kind of research we engaged ourselves into some years ago. It mainly conveys the kind of synergy that arose from every-day-work and the act of continuously questioning our own method and epistemological itinerary. They, therefore, shouldn't be considered off topic, but a perspective for refining the human side of research method: *epistemology*. These will qualify as the "Philosophy pages" for the grade of Philosophiae Doctor they are meant to materialize.

Part IV offers a variety of Indexes and Glossaries to tailor your own trip into the tree forest of this manuscript. This part is accompanied by a short Bibliographic essay on the different articles, books and PhD manuscripts that contributed to the shaping of our research project.

Part V includes Annexes with the totality of my experimental corpora, some short literature review on some aspects of sentences processing and particular brain areas, supplementary materials for each experimental chapter, and the results of a supplementary fMRI study directly comparing Chinese and French.

Last but not least, the reader is given a Brain map to carry along the pages (see figure 4, to be cut along the dotted lines). Its utility is twofold: first, it will tentatively help the non-familiar reader not to get lost across brain areas, and second, it will possibly allow the reader (who would like to play the game) to mark the particular functional mapping of different linguistic phenomena or processes that he would attribute to the brain areas constituting what we call the *Sentence Network*.



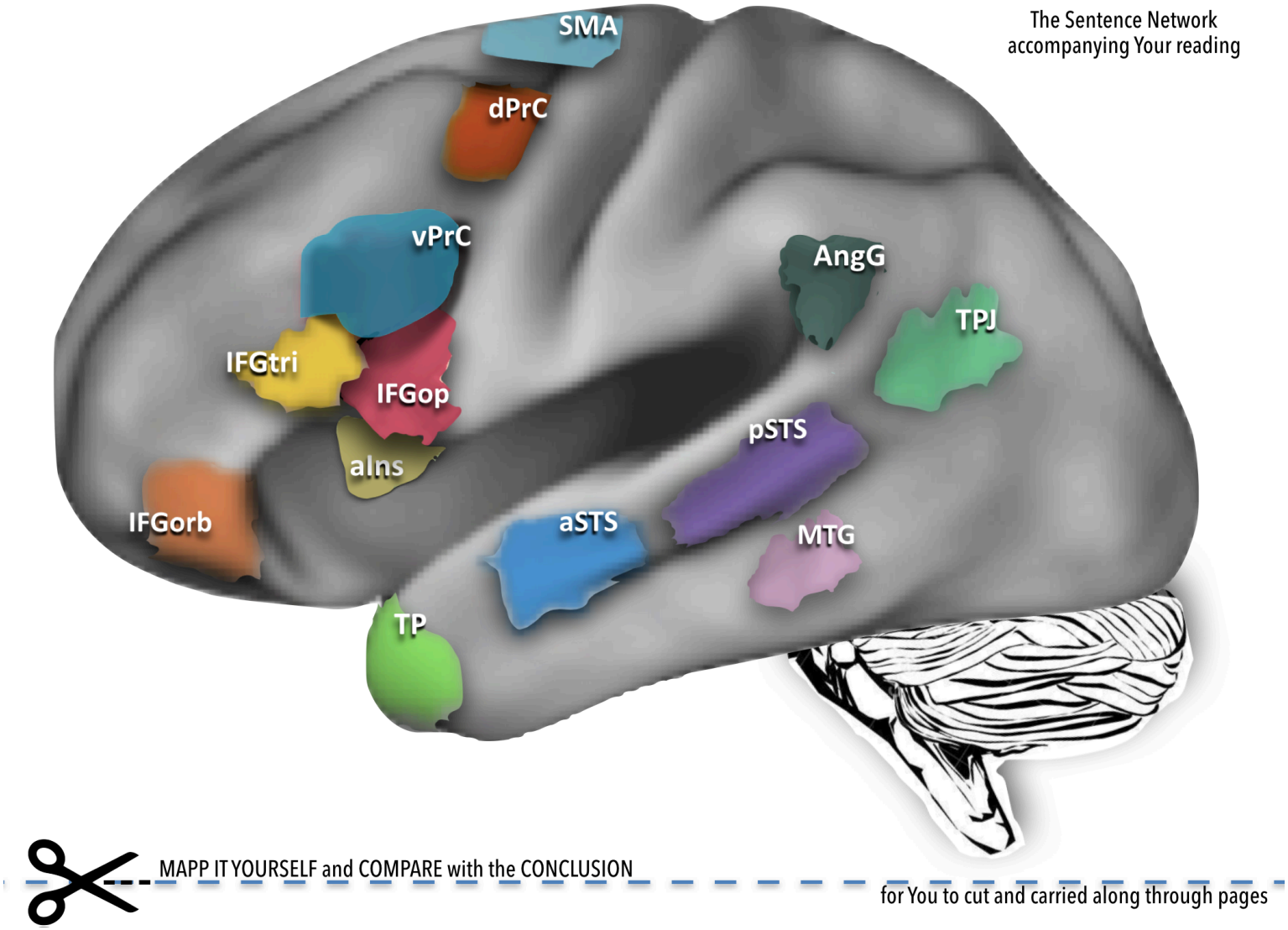


Figure 4 – The characterization offered by this PhD of the sentence hierarchical structures network.

Acknowledgements

"When I was a younger man, art was a lonely thing. No galleries, no collectors, no critics, no money. Yet, it was a golden age, for we all had nothing to lose and a vision to gain. Today it is not quite the same. It is a time of tons of verbiage, activity, consumption. Which condition is better for the world at large I shall not venture to discuss. But I do know, that many of those who are driven to this life are desperately searching for those pockets of silence where we can root and grow. We must all hope we find them." *"Silence is so accurate."*

Mark Rothko (1903–1970)

It is with great pleasure that we bow to the tradition that still gives the opportunity to express personal gratitude to all the human non-scientific contribution of people who have accompanied, understood, encouraged, and supported me throughout my years of doctoral work.

Not only the object of my research couldn't exist if by chance silence disappeared on earth. Noise works against sentences and even words cannot be heard in noise: language needs silence and most of all sentences need silence to be proffered and heard, to have meaning. So that the first acknowledgment I should utter is to silence.

The first thanks goes to a seemingly immaterial element, thus existing only because men and women give it free rein: *Silence*.

As wonderfully expressed by Rothko[35]: Silence is a place to find things! And, I would add that silence is the locus of research. A Ph.D. mainly needs silence as a fuel for thoughts, as a first place to make thoughts live in, be internally expressed, to allow the kind of "*recollection in tranquility*" that was theorized by romantic poets such as William Wordsworth (1770-1850) [36].

35. In Mark Rothko, *The artist's reality : philosophies of art*. Uncompleted manuscript, Christopher Rothko ed., New Haven, Conn.: Yale University Press, 2004.

36. *Recollection in tranquility* is the term used by British Poet William Wordsworth to discuss how

poetry emerges. We can read in his Preface to Lyrical Ballads (1800): “I have said that poetry is the spontaneous overflow of powerful feelings: it takes its origin from emotion recollected in tranquillity: the emotion is contemplated till, by a species of reaction, the tranquillity gradually disappears, and an emotion, kindred to that which was before the subject of contemplation, is gradually produced, and does itself actually exist in the mind.” (Preface to Lyrical Ballads by William Wordsworth, 1800, in The Harvard Classics. 1909–14).

l

Table of Contents

Contents

PART

Preliminary Considerations on the Field of Inquiry

Chapter 1

Delimitation of the Field of Inquiry

> Suppose that in some convulsion of the planets there fell upon this earth from Mars, a creature of a shape totally unfamiliar, a creature about whose actual structure we were of necessity so dark that we could not tell which was creature and which was clothes. We could see that it had, say, six red tufts on its head, but we should not know whether they were a highly respectable head-covering or simply a head. We should see that the tail ended in three yellow stars, but it would be difficult for us to know whether this was part of a ritual or simply a tail. **Well, man has been from the beginning of time this unknown monster.** *People have always differed about what part of him belonged to himself, and what part was merely an accident.* **People have said successively that it was natural to him to do everything and anything that was diverse and mutually contradictory**; that it was natural to him to worship God, and natural to him to be an atheist; natural to him to drink water, and natural to him to drink wine; natural to him to be equal, natural to be unequal; natural to obey kings, natural to kill them. The divergence is quite sufficient to justify us in asking if there are not many things that are really natural, which really appear early and strong in every normal human being, which are not embodied in any of his after affairs.
>
> G. K. Chesterton
> *The Venture Annual*, 1903.

Surprisingly, *Language* is one of those few things "that are really natural, which really appear early and strong in every normal human being" and that has strangely escaped *after affairs* of people's oscillatory opinion between natural and unnatural. This second part of the Preliminary Considerations should clarify the definitions of some core notions like language, man, symbolic reasoning, innateness, and human language faculty, on which this work is built. It will commit to the French academic tradition of *notional definition*[1]. We will touch here to the most challenging part of a pluri-disciplinary research project: the mutual understanding of the different disciplinary a priori. We will try to track down, if not all, at least the essential notional apparatus usually taken for granted once we come from one of the different fields of knowledge complementing each other in this work, principally to give the reader, coming from different disciplines and backgrounds, all the necessary conceptual tools and notions to understand this research's context and goals in a choral description.

The witty demonstration of Chesterton is here to remind the reader that in the realm of human knowledge about Man, what is natural is far from being evident, especially across time and centuries. Our research object, the sentence, is one of those evident units in language, but as soon as we consider the *sentence as a cognitive object* a certain amount of non-evident affairs emerge. The second non-evident issue will be to conceptually bridge the foundational concepts of linguistics and those of neuro-biology and psychology, what is commonly called the linking hypothesis problem (cf. Epilogue, page 709). Thus, in a probably too audacious way, we will attempt to find a **chorality** among the voices that rise on a fair number of controversial issues corresponding to the following five Sections:

1. what is language faculty in the definition of man?
2. what are languages in the definition of language faculty?
3. what is syntax and computation in the definition of language faculty?
4. what role plays the brain in a definition of language faculty?
5. what constraints neuro-imaging techniques introduce in the study of language in the realm of neuro-cognitive science?

Hence, the following chapter will chorally introduce the reader to the different issues at stake in this pluri-disciplinary research universe - with all the satellites and planets of its theoretical approaches and experimental methodologies.

One secret a great chef will never reveal are his cooking ingredients, but even more confidence will be cast on the way the chef composed them. On the contrary, the logic of intellectual work is far from the one of the glory of being selected in Michelin's guide. The tradition in research is to share the one's clearest notional articulation. Hence, this preliminary chapter will offer a presentation that could be called in a more trivial way the *notional "cooking ingredients"* of this Doctoral work – the context and key issues covered by this manuscript. In this chapter (and in the next one too), the reader will not deal with a historical background review, or with a history of the different disciplines and relative scientific questions. By no means this review will be chronological. On the contrary, I will freely retrace the main elements that determined a comprehensive understanding of the issues at stake in this research, in what I would call "a tale of my understanding discovery", a tale of a Linguist in the land of Cognitive Neuro-science and a tale of an experimental Cognitive Psychologist in the land of Linguistics. The reader

1. See the Outline of the manuscript and the Notice to the Reader for a thorough justification on the presence of this Part (p.xxii).

will be walked through the essential notions framing our research, and the essential assumptions that grounded our approach.

We will now invite on the stage the *choir* of the past and present intellectual contributors to my reasoning. As it was traditionally the case in ancient Greek theater, the Parodos, will feature the first great aria of the choir after the Prologue, in a polyphony of authors, disciplines, intellectuals, theories of different epochs.

Overview of the contents of this chapter

Cooking ingredients: notions and methods at stake

> *Although we think we govern our words, [...] certain it is that words, as a Tartar's bow, do shoot back upon the understanding of the wisest, and mightily entangle and pervert judgment. So that it is almost necessary, in all controversies and disputation, to imitate the wisdom of mathematicians, in setting down in the very beginning the definitions of our words and terms, that other may know how we accept and understand them, and whether they occur with us or no. For it cometh to pass, for want of this, that we are sure to end there where we ought to have begun, which is - in questions and differences about words.*
>
> Francis Bacon, *The advancement of Learning*, 1605.

As synthetically stated by Francis Bacon this section contributes to "setting down in the very beginning the definitions of our words and terms". Ironically, we would like to prevent having questions about our *words*, rather than about our *sentences*. The aim is to present or at least enumerate the conceptual tools that are needed to ask what it would mean for a certain linguistic level of analysis -the sentence- to be considered a cognitive object. We will outline the theoretical apparatus related to the issue of cerebral representation and processing of the sentence as a syntactic unit. This will allow to have a re-examination of the presuppositions hiding behind the fact of correlating linguistic phenomena and their theoretical description with brain responses.

The following discussions on language, syntax, cognition, brain and neuro-cognitive methods will suggestively help trace the main lines of the further argumentative developments we will present in the Introductory Matters (Part I) and Experimental development (Part II). Hoping that, the non-linguist readers and the non-neuro-cognitive reader can get acquainted to the key relevant concepts and tacit background assumptions that do not directly belong to their field.

Speaking of intellectual "culinary" association of notions and ideas, we will try in this section to replace a certain number of concepts in the mouth of the intellectuals that – according to our knowledge – first formulated them. As is traditionally said, we will try and give to Caesar what belongs to Caesar, or in our case we will give to Humboldt, Plato, Hobbes, Locke, Leibniz and Jespersen what belongs to them, and, what fashion in research is usually attributing to near past production. In this way, our commitment to finding an antidote to fragmentarity will be declined – in a hopefully successful manner – in an antidote to time fragmentarity. Hoping the reader will enjoy the vicinity of the different voices that contributed with penetrating insights to the vast interrogation about

man and language. We will gradually try to build up chorality by bringing into resonance these contributions in a non-chronological way.

This dialogue through time will allow an exceptional notional harvesting that is far from being just an historical excursion or a nostalgic intellectual digression. It represented in fact a substantial and non-secondary contribution to the development of our experimental hypothesis, in that it allowed us to make important choices regarding both our theoretical approach and methodology[2], and to identify and refine the context and goals of this investigation about sentence structure and its neural substrates.

1.1 Man and language

> "L'homme n'atteint pas le fond de l'homme. Il ne trouve pas son image dans l'étendue des connaissances qu'il acquiert, il trouve une image de lui-même dans les questions qu'il pose.
>
> *["Man can never plumb the depths of his own being; his image is not to be discovered in the extent of the knowledge he acquires but in the questions he asks.]*
>
> André Malraux *Antimémoires*
> 1967

In the watermark of this manuscript are two cardinal interrogations about human condition and language[3]. In fact, what Malraux states in this epigraph is true for this manuscript: it would indeed be difficult to lean on the question of language without understanding it as a one of the most definitory aspects of man. *"The image of man"* one can acquire by reading this research work, is of course tightly linked to the question of language, but more specifically linked to what is most exceptionally human in language: its structures and the possibility across languages to have both recurrent and diverging patterns in sentence's structure. Hence, as our most fundamental question is 'What is a sentence?', one could continue Malraux's statement and ask 'what does this linguistic unit tell about Man?'[4].

While common sense generally acknowledges that the ability to use language is one of the major factors that sets humans apart from other species, the way philosophical

2. This is not only true for the itinerary through history of ideas, but also for the birth of techniques that allow us to observe brain activity, see Annexes on Electroencephalography EEG **??** at page **??**.

3. I here confess in a burst of irony that one of my greatest regrets in life is that Prefix, my beloved 11 yeas old back cat, does not speak, and therefore cannot tell me about the way it sees the world as being a cat, and what would it be like to read cat poetry (because I am sure my cat is a poet).

4. Being rather attached to poetry, the question about syntax and sentences articulations, is, in my understanding, ultimately the question about what permits poetry to exist. Poetry, - thanks to the mastery of these creative "structure tamers" that poets have always been - is an incredible agglutination of lexicon characterized by disparate and non-overlapping semantic fields that are mysteriously and syntactically glued together to allow us to build up new images and unconventional representations. One could, actually say that poetry will stay impossible for or inaccessible to animals.

inquiry and later Cognitive Science orchestrated a number of specific demonstrations for this claim, brings us to set up an intricate conceptual panorama in the next coming Sections.

1.1.1 Language, a definition of man ?

Language faculty is a philosophical, linguistic, cognitive and social touchstone, in that it has always been considered as conferring a unique status to Man, among all living species. The origin of human linguistic capacity, the conditions of its emergence, the nature of its production and perception, together with the investigation of its understanding, its acquisition and pathology, and the exploration of its diversity through world's languages, have all been at the origin of innumerable intellectual works.

Figure 1.1 – Capital from the Basilica of Saint-Benoit-sur-Loire.

Interestingly, the study of observable linguistic facts in humans have frequently led to more existential questioning about human nature too, giving often birth to myths and taboos (cf. Parisian Linguistic Society ban in 1886). In one among the most forgotten civilizations that built the European cultural substrate – the Celts – there existed a deity of Eloquence. This god was so cardinal that it succeeded in surviving successively Roman and Christian iconography, to finally reach early medieval Romanesque architecture. This Celtic Gaulish god of Eloquence *Ogmios* - renamed as Hercules Gallicus in Renaissance times - was not only personifying the power of speech (eloquence), but also considered the mythic inventor of Oghamic alphabet, a writing system of primitive Irish. How was the power of speech understood at that time? A Greek satirist of the 2nd century B.C., Lucian of Samostate, described this deity as covered by a lion's fur, attracting towards him a considerable multitude of beasts by attaching them with chains. But the god of Eloquence was not brutally attaching beasts, it was using a particular type of chain. The fine chains of gold and amber coming out of his mouth were symbolically representing the power of the Word, attracting beasts and lions and transfiguring them into Man.

This iconographic theme depicting the god of Eloquence tiding beasts with the ropes of language can be found all around Celtic Europe on vase pottery, sculpted reliefs, coins and stone tablets in Spain, Italy, France, Ireland, Austria, etc. The picture beside is reporting and example found on a capital from the Basilica of Saint-Benoit-sur-Loire nearby Paris[5]. Its reliquescence in the iconography of the new christian word is a clear

5. This capital has been commented in an article by Régine Pernoud, a famous historian of the Medieval period and Custodian to the National Archives. "Nuit obscure sur les Celtes: un chapitau de Saint-Benoit-sur-Loire". In *Renaissance de Fleury, De la nuit des Celtes à la clarté Romane* n. 249, mars 2014, pp. 20-24.

mark of its importance[6], and it shows how the idea that man is defined by language faculty is constant and long-lasting, across different cultures, religions and ages,

Since Celtic and Medieval times, the delicate question of the difference between the animal and man has been constantly brought back to the scene. Initially, through the prism of religion and philosophy, and only after by linguistics. It is interesting to note that the two fundamental themes of the nature of man and the origins of language have never stopped to be hotly debated. Just consider the intriguing XIXth century's debate about language and its origins, when in 1886 the decision of the Paris Linguistic Society to censor research about the origin of language (in absence of proved evidence) literally inflamed the intellectual world[7].

These examples show how debated has been one of the central issues dealt by this Doctoral work. Thus, resisting the facility of the *argument from authority* to automatically qualify man with Aristotle's words as ζοον λογον εχων *'zoon logon êkhon'*, we will go back in time to refine through successive approximations, the question of man and language. This overview, as it was already noted by Gaston Bachelard, will serve scientific conceptualization, in that, "scientific conceptualization is a series of successive and well-ordered approximations, this is the reason why their periodical and regular assessments and overviews, and, the fact of reflexively contrast them to put them in perspective do not seem to us superfluous" (Bachelard, 1938:61)[8]. Hence, we hope this reflexive overview will help the reader as it helped us to put this central debate in perspective. We will now move to the XVIIth century to identify through Descartes' reasoning what is specifically humand in language and how the way to address it has not evolved much since that time.

What is human in the Language faculty?

Interestingly enough for our contemporary reflection on these matters, Baroque times added to the man versus animal dyptique a third cardinal element of comparison: the *machine*. In a famous argumentation[9], the distinction between animal, machine and the *reasonable* man, has been magisterially treated by the founder of philosophical rational thinking, the french philosopher René Descartes (1596 – 1650). We will now enjoy it, before analyzing its programmatic and sometimes prophetical spin-off.

> "*If there were machines bearing the image of our bodies*, and capable of imitating our actions as far as it is morally possible, *there would still remain two most certain tests whereby to know that they were not therefore really men*.
>
> Of these **the first** is that **they could never use words or other signs arranged in such a manner as is competent to us in order to declare our**

6. According to Georges Dumézil (INALCO) it can be related back to the linking deities like the Scandinavian Odin and the Indian Varuna. It could have been christianized in Saint-Peter-in-Chains (San-Pietro-in-vincoli) in Rome.

7. We will come back to this point in a few pages when addressing the notion of Universal Grammar. And, thanks to the introduction of the brain into linguistic reflection – and the consequent birth of neuro-linguistics as a discipline – we will question the origin of Language faculty in biological terms.

8. As already announced above, this point by Bachelard is the main reason of being of these sections. It will in fact, not only pay a pledge to the french intellectual academic tradition of defining the concepts that are going to be used in order to frame an argumentation.

9. Addenda to the text in squared brackets are taken from a commentary of Étienne Gilson, a philosophy professor at the Collège de France that re-edited the *Discours* in 1939.

thoughts to others; for we may easily conceive a machine to be so constructed that it emits vocables, and even that it emits some correspondent to the action upon it of external objects which cause a change in its organs; for example, if touched in a particular place it may demand what we wish to say to it; if in another it may cry out that it is hurt, and such like; **but not that it should arrange them variously so as appositely to reply to what is said in its presence, as men of the lowest grade of intellect can do.**

The second test is, that although such machines might execute many things with equal or perhaps greater perfection than any of us, they would, without doubt, fail in certain others [suspicamur ea liberum arbitrium no habere] **from which it could be discovered that they did not act from knowledge, but solely from the disposition of their organs**: for while Reason is an universal instrument that is alike available on every occasion, these organs, on the contrary, need a particular arrangement for each particular action; whence it must be morally **impossible that there should exist in any machine a diversity of organs sufficient to enable it to act in all the occurrences of life, in the way in which our reason enables us to act**.[10].

Again, by means of these two tests we may likewise know the difference between men and brutes [*Contra* Montaigne]. **For it is highly deserving of remark, that there are no men so dull and stupid [*Contra* Montaigne], *not even idiots, as to be incapable of joining together different words, and thereby constructing a declaration* [sequence of words with a meaning] by which to make their thoughts understood;** and that on the other hand, there is no other animal, however perfect or happily circumstanced which can do the like. Nor does this inability arise from want of organs: for we observe that magpies and parrots can utter words like ourselves, and are yet unable to speak as we do, **that is, so as to show that they understand what they say;** in place of which men born deaf and dumb, and thus not less, but rather more than the brutes, destitute of the organs which others use in speaking, are in the habit of spontaneously inventing certain signs by which they discover their thoughts to those who, being usually in their company, have leisure to learn their language. And this proves not only that the brutes have less Reason than man, but that they have none at all: for we see that very little is required to enable a person to speak; and since a certain inequality of capacity [*Contra* Montaigne and Charrons : there is more difference between a man and another than between a man and an animal] is observable among animals of the same species, as well as among men, and since some are more capable of being instructed than others, **it is incredible that the most perfect ape or parrot of its species, should not in this be equal to the most stupid infant of its kind, or at least to one that was brain-damaged,** unless the soul of brutes were of a nature wholly different

10. Nowadays the temptation to reduce human language to communication and to strings of words has surfaced again, and academic appetite for complex system oxymoronically too. The Big Data era opened its doors less then 10 years ago, and we will be witnessing a revival of language reductionist approaches, "the machine metaphor will try to access to its embodiment again", and the reality of poetry will be there to show that language is still for those futile humans who think about the questions that death addresses to life (cf. Poetry in the Notice to the reader

from ours."

René Descartes, *Discourse on the Method*, V part., p. 55-59. 1673

If one tries to bind today's research advancement with the boldface parts of this text, it is hard not to have the impression that Descartes was standing on a balcony from which he could enjoy an anticipatory vision of Cognitive Neuro-science. As a matter of facts, the rationalist philosopher was, with Francis Bacon, among the first to turn his interest towards the heyday of machine era, and, in order to identify true man from hypothetical machines or animals[11], he invites us to proceed in two ways. We will first analyze in details the first because of its pertinence for our research topic.

The first, indeed, consists in checking the syntactic ability of the individual, in that "they would never (1) use words or other signs composing them as we do (2) to express to others our thoughts." We can find here the first start of the reflection on language faculty as being (1) eminently a matter of arranging words into linguistic structures, and (2) as being linked to human symbolic thinking. The philosopher further continues the comparison between man and animals developing it in the same direction: "For it is highly deserving of remark, that there are no men so dull and stupid, not even idiots, as to be incapable of joining together different words, and thereby constructing a declaration by which to make their thoughts understood; and that on the other hand, there is no other animal, however perfect or happily circumstanced which can do the like." Here, this capacity of joining together different words and constructing a declaration to express thoughts is taken as a definition of Man and does not seem to be linked to intelligence or other skills. In fact, Descartes systematically uses comparisons between the capacities of animals and those of humans characterized by extreme behaviors, like stupidity, daze, infancy, brain damage: "***It is incredible that the most perfect ape or parrot (1) of its species, should not in this be equal to the most stupid (2) infant of its kind, or at least to one that was (3) brain-damaged, unless the soul of brutes were of a nature wholly different from ours.***" The parallel with today's research advancement is astonishing in these lines. What Descartes could never have imagined is that the research program of the next 400 years, till the advent of Cognitive Neuro-science of language, would have been summarized here. In fact, his sharp and penetrating observations on human language capacities had already pointed the path we are now walking in at least four main directions by studying:

- "the most perfect ape or parrot of its species", including bird songs, monkey and primate proto-linguistics skills[12];
- language acquisition analyzing the linguistic behavior of infants;
- the linguistic behavior of "brain-damaged" individuals in aphasiology;
- the neural underpinnings of the "use words or other signs arranged in such a manner to express thoughts" in neuro-syntax, but even at highest level of linguistic analysis in the study of cerebral correlates of discourse information.

In conclusion, while rational thinking of Descartes times would not further question the limit between man and animal, philosophers and scientist, in the last 150 years,

11. Here again surfaces the need of an alterity to sketch a definition of the object under observation. We will see this all along the manuscript, particularly in the role that Chinese language frequently plays as a linguistic alterity (see Notice to the reader).

12. Interestingly, today's research is developing the study of birdsong in both a neuro-biological manner with intra-cranial recordings and in a computation manner by modeling their songs (see reference articles Abe and Watanabe, 2011 and Berwick at al. 2011)

have tried to ponder the uniqueness of human language, of a particular fascination with the linguistic and cognitive capacities of great apes. We will therefore indulge in some of the differences between man and animal that are still up-to-date in the following Sub-sections.

1.1.2 The filiation of language : thinking and symbolic faculty

The link that is made by Descartes between the faculty of composing words into sentences and the expression of thoughts will be long lasting and it will be further analyzed into a aptitude for symbolic cognitive faculty.

The point that is made here on the linguistic expression of thoughts could seem distant from our primary interrogation about the sentence, but it is indeed cardinal to grasp what makes sentences the medium for thought. And, in the early XVIIth century, Descartes had already pinpointed something that lies at the very core of the interrogation of this research work: namely, the possibility that human language offers, not only to build meaning in groups of words governed by rules that we call sentence units – capsules of structured meaning –, but also to build and articulate discourse out of a set of separated sentences. We can understand this last point in the following manner: the fact that a sentence unit can have an interface with the discourse level can constitute an anchor between each self-standing syntactic unit and build up an articulated discourse[13], to the point of allowing humans to express their thoughts.

In contrast, as we can read in Descartes too, animal signaling calls for automatic reaction, it remains an instinctive behavior and can hardly qualify as conveying thinking or signify a thought. This is why the debate on the origins of language faculty has long focused on the question of the development of *symbolic thinking*.

A cognitive faculty to manipulate symbols

A rapid review of the numerous attempts to teach human language to chimps should be instructive of what it means to manipulate symbols.

The very first experiment started in the 1920's[14] with Robert Yerkes who failed to teach English to a chimpanzee, and first thought at teaching Sign Language in order to overcome the articulatory disadvantage of chimps. This experiment was followed by other attempts to raise chimpanzees in an interactive family context, in order to question the specificity of language in Humans. The Kellog family raised the chimpanzee Gua together with their child Donald in the 30's, and the Hayes raised the famous Viki (1947-1954) in the 50's. Their results encouraged the scientific community, and three different projects were then started in the United States at the beginning the 70's. A couple of scientists (Alan and Beatrice Gardner) and Francine Patterson experimented teaching sign language (ASL) respectively to a chimpanzee (Washoe) and to a gorilla (Koko), while David and Ann Premack experimented teaching lexigrammes - associations between a meaning and a visual element - using plastic shapes on a magnetic blackboard to a female chimpanzee named Sarah. These experiments showed the possibility of teaching chimps a

13. This aspect will be fully developed in chapter 3 and The structure of the sentence with its discourse interface will be introduced in chapter 2.

14. The idea to teach language to chimps was first formulated in the XVIIIth century by de La Mettrie (1748).

communication system that could be seen as near to language in terms of communicative function, but quite distant in terms of form and structures.

The case of a chimpanzee named Nim Chimpsky reported by Herbert Terrace (1979) started relativising the interest for these experiments. A fine grained analysis of its linguistic production revealed, in fact, stereotypical and poor utterances based on a great number of repetitions, and on no more than two symbols combination, that were often not meaningful. If the capacity to compose words was dis-confirmed, the surprising learning capacities of a chimpanzee named Kanzi attracted much attention in the 1980's. Sue and Duane Savage-Rumbaugh reported that its communicative production had reached a thousand words.

The evaluation of what are the communicative capacities in animals, and all the possible considerations on the differences between more gestural communication and linguistic *double articulation* (A. Marinet)[15] shouldn't distract us from the essential fact that these studies reveal: man disposes of a cognitive capacity of manipulating, composing and creating symbols in an effortless manner, which is particularly salient in manipulating symbolic linguistic elements. The french linguist Émile Benveniste has some clarifying statements on how the mastery of the symbolic faculty represents a difference between human and animal language:

> "Using a symbol is the capacity to identify the characteristic structure of an object and to identify it in various contexts. It is that which is peculiar to man and which makes man a rational creature. **The symbolizing faculty**, then, permits the formation of the concept as distinct from the concrete object, which is only one realization of the concept. Here is the basis for abstraction [...]. Now, this **representative capacity, in essence symbolic**[16], which is at the basis of the conceptual functions, appears only in man. [...] Let us first take great care to distinguish between two notions which are very often confused in speaking of "animal language" - the signal and the symbol. A signal is a physical fact bound to another physical fact by a natural or conventional relationship: lightning heralding a storm, a bell announcing a meal, a cry proclaiming danger. An animal perceives the signal and is capable of reacting adequately to it. It can be trained to identify various signals, that is to say, to connect two sensations through the correlation of the signals.[...] But [the symbolic faculty] uses in addition symbols that have been instituted by man; one must learn the meanings of the symbols, one must be able to interpret them in their signifying functions and not simply perceive them as sensory impressions, for symbols have no natural correlation with what they symbolize. Man invents and understands symbols; the animal does not. [...] It is often said that the trained animal understands human speech. In reality, the animal obeys the spoken word because it has been

15. The possibility in human language to build meaning in a first articulation, assembling phonemes without semantic value into semantic units that are called morphemes and a second one assembling morphemes into meaningful structures that can be lexical or syntactical (e.g. words, phrases or discourse). See Sub-section on Hockett's analysis of language faculty, page 16)

16. The faculty of abstract representation allows to consciously take something for something else, as for example playing with a doll or using a word, knowing that what we have in our hands is something else than a baby or that the sound of a word is not the thing we are referring at. As R. Bathes used to say: "Every object can pass from a closed, still existence to an oral existence open to the appropriation society can make of it" (Barthes 1957: 216). As a matter of fact, a sea shell can become a coin to pay at the grocery store as it was in Phoenician time, through convention.

> trained to recognize it as a signal[...] Between the sensory-motor function and the representative function is a threshold which only human beings have been able to cross. [...] The emergence of Homo in animal series may have been helped by his bodily structure or his nervous organization, but it is due above all to his faculty of symbolic representation, the common source of thought, language, and society."
>
> Émile Benveniste, *Problèmes de linguistique générale*, 1966.

To this reflection on human aptitude for manipulating and forging symbols, we can add what Charles Darwin pointed out in his great classic the *Filiation of man* (1871): what fundamentally distinguishes man from animals is the extraordinary development of his mental faculties, allowing him to associate the most different sounds with ideas. Here again, we notice the consideration of the link between mental faculties and the possibility of language[17].

In sum, language not only appears to be distinctive of Man compared to animals, but a link is here established between linguistic abilities and the sort of symbolic mental activities that have been claimed to be specifically human (cf. Tattersall's oft-repeated claim[18], see Tattersall (2012) and Tattersall and DeSalle (2012)).

Assuming this point of view on language faculty, we can say that linguistic capacity reveals the human *faculty of thinking*, precisely in his capacity to compose abstract symbols like words into sentences: Language, Sentences, Reason, Thinking and Man appear finally to be closely linked. Furthermore, this last point can bring us to look at Aristotle's assertion about man and logos as bearing an original lexical ambiguity: the Greek term *logos* means at the same time language, reason and thinking. Thank to this lexical ambivalence, Man could equally be a language being, a rational being, and a thinking being. In other words, the fundamental expression of syntax, embodied by the sentence unit, can be seen as essential to human language faculty, to the point that it ultimately marks the possibility to compose symbols in rational thinking. This last point strongly echoes the issue of the creative aspect of language that we will now address.

1.1.3 Language, Logos and creativity = human sentences

> *"Non lexis magna, sed phrasis."*
>
> Seneca the Elder,
> *Controversiarum*, Lib. 3, p.247

17. "That which distinguishes man from the lower animals is not the understanding of articulate sounds, for, as every one knows, dogs understand many words **and sentences**. [...] It is not the mere articulation which is our distinguishing character, for parrots and other birds possess this power. Nor is it the mere capacity of **connecting definite sounds with definite ideas**; for it is certain that some parrots, which have been taught to speak, connect unerringly words with things, and persons with events. The lower animals differ from man solely in his almost infinitely larger power of associating together the most diversified sounds and ideas; and this obviously depends on the high development of his mental powers." Charles Darwin, The Descent of Man and Selection in Relation to Sex, 1871, p.85, ed. 1889, New York, D. Appleton and Co.

18. Ian Tattersall is a evolutionary paleo-anthropologist who argues for a link between linguistic abilities and the type of symbolic activities that are claimed to be specific to humans. In this, he states that the fossil evidence for these activities linked to symbolic behavior can be used to date the emergence of the modern linguistic mind in human lineage. Ian Tattersall thinks that this "symbolic activity" has a "rather recent origin", and that "as far as can be told, it was only our lineage that achieved symbolic intelligence with all of its (unintended) consequences". It should be noted that some of his positions are debated among specialists of Evolution.

From what we have said so far, it might seem plausible to conclude that this ability for Sentence unit appears as a particular Homo Sapiens Sapiens, that we could therefore re-baptize as Homo Phraseologicus. We see here how much the general attempt to give a biological understanding of language, to introduce linguists to this dimension, is gradually drifting our argumentation towards 'humanizing' and 'sententializing' language. And, we will now see how much of what is specific to human in language resides inside the domain of the sentence as a unit.

The creative aspect of language

In his *Discourse on Method* Descartes addresses a second central aspect of human language capacity, that of its productivity and creativity. By contrasting the capacities of animals, and specifically those of birds, with human linguistic behavior, he concludes: "Nor does this inability arise from want of organs: for we observe that magpies and parrots can utter words like ourselves, and are yet unable to speak as we do, that is, so as to show that they understand what they say; in place of which *men born deaf and dumb,* and thus not less, but rather more than the brutes, destitute of the organs which others use in speaking, *are in the habit of spontaneously inventing certain signs by which they discover their thoughts to those who, being usually in their company, have leisure to learn their language*." Here surfaces another crucial distinction: the creative aspect of language and communication in man strongly contrasts with the fact that animals cannot spontaneously create new ways of communicating.

However, his reflection goes beyond simple observation, and he further interprets this lack of creative linguistic capacity observable in animals and machines in these terms: "although such machines might execute many things with equal or perhaps greater perfection than any of us, they would, without doubt, fail in certain others [suspicamur ea liberum arbitrium no habere] from which it could be discovered that they *did not act from knowledge, but solely from the disposition of their organs*: for while Reason is an universal instrument that is alike available on every occasion, *these organs, on the contrary, need a particular arrangement for each particular action*; whence it must be morally impossible that there should exist in any machine a diversity of organs sufficient to enable it to act in all the occurrences of life, in the way in which our reason enables us to act." In this second remark on this creative aspect of human linguistic behavior, Animals are described as being determined by "the disposition of their organs", while Man, possessing "Reason as a universal instrument", "acts from knowledge" without being deterministically constrained by his organs and/or by situations in his speaking act. We see, here, how the classical issue that brought Descartes to postulate the *res cogitans* as a property distinguishing human beings from the animals or machine, is indeed tightly linked to the "creative aspect of the use of language"[19].

Therefore, given the ability that every man shows to *freely* produce new utterances to expose his thoughts in an appropriate manner according to different situations, without being deterministically constrained by those, the philosopher made Language creativity and the 'Logos' (i.e. *reason*) go hand in hand. In other words, the very fact that a sentence is uttered in any situation according to the speaker's state of mind and his thinking, relies on more than just behavioral dispositions and conditioning.

19. Some argue that this focus on language is at the basis of the duality between mind and body developed in his philosophical system.

We will investigate this aspect by focusing on the properties of the sentence-discourse interface of the sentence-unit.

Language creativity is in the sentence

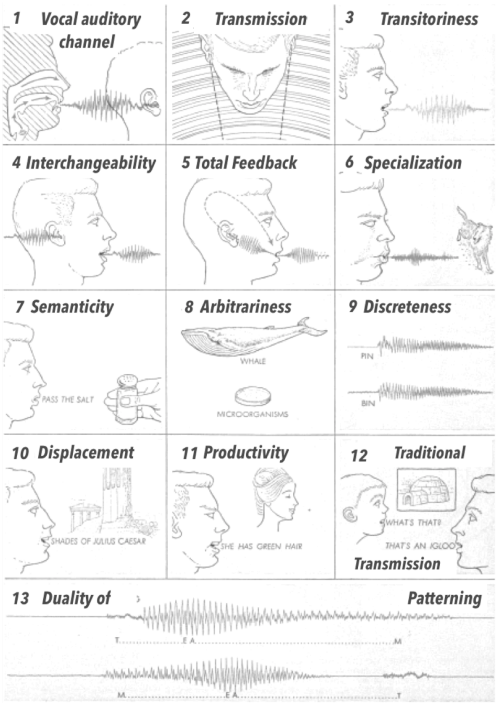


Figure 1.2 – An iconic representation of thirteen characteristics of Human Language, adapted from Hockett (1960).

The sentence unit appears to be part of the domain of this creative aspect of linguistic composition that is performed in a non-deterministic fashion. This last point directly leads us to ask the following experimental question about sentence unit as a domain for this creative productive aspect of human language: can an ape produce a sentence?

Going back to the popular experiments from the 70's reported above a particular study directly asked this question (cf. Terrace et al., Science 1979, experiment performed in Canada). The results revealed that even though the chimpanzee was able after a long training to constitute a corpus of 19.000 different Sign-Language utterances, it nonetheless had never uttered a new sentence (one that it had not already heard). These results are even more significant if contrasted to children's ability to produce new utterances mastering only 128 words, like the one never heard before: "Dad here, hat here" to convey the meaning of "Daddy's hat is here" [20].

These experimental results seemed to conclude a debate dating from the late 50's opposing the two views on human language faculty of the behavioral psychologist B.F. Skinner and of N. Chomsky [21].

While the first attributed to humans a set of behavioral dispositions that enable them to react to the environment and to utterances we hear (Skinner, 1957), the second was addressing the creative aspect of human linguistic behavior by saying that "*man has a species-specific capacity, a unique type of intellectual organization which cannot be attributed to peripheral organs or to general intelligence, and which manifests itself in what we can refer to as the 'creative aspect' of language use*" (Cartesian linguistics (1966), pp. 4-5), that he will further define as grammar.

One could concisely resume these facts and affirmations by saying that while apes or some animal species can have dictionaries of sentences, humans have dictionaries of

20. Note that this fundamental aspect of human language faculty was already formulated by von Humboldt in a concise observation: "language makes use of finite means in an infinite way". This sentence of von Humboldt has been erroneously interpreted, you can find what the author really meant and its context in the next section (section 1.2) about Languages.

21. Chomsky wrote a review refuting most of Skinner's theories (Chomsky, 1959),where the first point of disagreement was, indeed, that language is stimulus independent and historically unbound. The discerning learner will find here neither trace of the debates between Fitch, Hauser and Chomsky against Jackendoff and Pinker that occupied the years 2002 and 2003 on the evolution of language, nor of those that occupied the years 2015 and 2016 around the book "Why Only Us" (Berwick and Chomsky 2016). They are not vintage enough to be in this section. Moreover, they reproduce an old polarization of the scientific debate, in a probably not so constructive way, just by evoking, with new words and new animosity, long-standing debates that the reader will anyway find in this Section.

words they can freely compose according to grammatical structures. We can, therefore, reach the conclusion that human linguistic production proceeds in a *free, creative and non-deterministic manner* [22].

Interestingly, in his critic to Skinner's theories on reinforcement, Chomsky (1959) argued that a theory requires an independent definition of the natural objects it studies, before one can investigate how frequency and reinforcement influence the learning of those objects. In the case of language, he suggested that the natural object is the sentence, and that its definition is to be provided by a grammar (see section 1.3, and Townsend and Bever (2001) for further arguments on this point.).

This creative aspect of language that has never ceased to fascinate thinkers, philosophers, poets and linguists, still remains somehow a mystery that has not found more than a simple description. This being said, linking it to the human grammatical ability for building sentences, transforms it into an experimentally addressable issue, thanks to the advent of modern imaging technologies. The question of human language's uniqueness can now be investigated from a neuro-cognitive perspective, and syntax still lies, as it was already the case in Descartes words, at the core of the debate on 'humanness'. This last point could be seen as the far-reaching and ultimate goal of this research project.

A few definitory properties of Human Languages

Before turning to one of the most debated issue of human language, we will recapitulate here what are the core human language characteristics thanks to the systematic analysis of Charles Hockett in the 60's[23].

In parallel to the debates on language faculty and learning, opposing Skinner's and Chomsky during the 60's, Hockett enumerated 15 definitory properties of human language. Essentially, he stated that these properties were found in all the human languages of the world, and never present all together in an animal communication system. He offered in this way a set of hypotheses that are still in use today to discriminate the aspects that are shared by animal communication from those that are uniquely human. As we can see in the following table, out of these 16 "design-features" of human language represented in the figure above, the first nine are said to be present in the communicative interactions of certain animal species (sometimes taking them in the broad sense), while the four last are hardly observable outside human linguistic system.

While certain low-level properties like vocal-auditory channel are shared by animals, others are more subject to debate, like Interchangeability, Discreteness, Arbitrariness[24]. Double articulation (duality of patterning), productivity or creativity are unarguably uniquely human.

22. One could wittily argue that apes are freer because they do not have rules. It still remains to be proven that not signifying anything when one would want to communicate can be qualified by the adjective 'free'. The grammatical constraints on sentence formation seem to be definitely freeing humans' solitary thoughts into linguistic creativity.

23. Charles Hockett was an American Linguist, teaching linguistics and anthropology since 1946. Note that he will be central in this work for two other reasons: in 1958 he formalized and first introduced the modern usage of the concepts of Topic and Topic-Comment as the most general predicative construction (as we will see in chapters 2), and he introduced the immediate Constituent analysis (Hockett 1960, The Origin of Speech, Scientific American 203, 88–111).

24. Consider the case of the "Kraak Krakoo" expressions of Velvet and Campbell monkeys (results reproduced in Barcelo-Coblijn and Gomila, 2012). For studies on monkeys concatenating vocalizations, see Ouattara et al. (2009).

Some human language properties found in animals' comunication	Crickets	Bee dancing	Western Meadowlark	Gibbon calls	Signing apes	Language in Humans
Vocal-auditory	Auditory only	No	**Yes**	**Yes**	No	**Yes**
Rapid fading	**Yes,** repeated	?	**Yes**	**Yes,** repeated	**Yes**	**Yes**
Interchangeability	Limited	Limited	?	**Yes**	**Yes**	**Yes**
Feedback	**Yes**	?	**Yes**	**Yes**	No	**Yes**
Specialization	**Yes?**	?	**Yes**	**Yes**	**Yes**	**Yes**
Semanticity	No?	**Yes**	In part	**Yes**	**Yes**	**Yes**
Arbitrariness	?	No	If semantic, yes	**Yes**	Largely **yes**	**Yes**
Discreteness	In part	No	?	In part	**Yes**	**Yes**
Displacement		**Yes,** always	?	No	**Yes**	**Yes,** often
Productivity	No	**Yes**	?	No	Debatable	**Yes**
Cultural transmission	No?	Probably not	?	?	Limited	**Yes**
Duality of patterning	?	No	?	[Cotton-top tamarin: **Yes**]	**Yes**	**Yes**
Prevarication					**Yes**	**Yes**
Reflexiveness					Debatable	**Yes**
Learnability					**Yes**	**Yes**

Figure 1.3 – Table of the 15 characteristics of Human Language and the animal species that respectively share some of those properties. Adapted from Hockett 1960 in W-S-Y Wang ed. Human communication: Language and its Psychological bases, Scientific American (1982) (a paper originally published in Scientific American 1960).

Here is a commented list of the 15 features proposed in Hockett's book:

1. Vocal-auditory channel: sounds are emitted by different parts of the vocal tract and perceived by the auditory system.[25]
2. Broadcast transmission and Rapid fading (*transitoriness*): Signal lasts a short period of time, and signal transmission happens in all directions (this is true of all systems involving sound).
3. Interchangeability: All utterances that are understood can be produced by every member of a given species[26].
4. Total feedback: The sender of a message also perceives the message.
5. Specialization: The signal produced is specialized for communication and is not the side effect of some other behavior.
6. Semanticity: There is a fixed and stable relationship between a signal and a meaning. It is often the case in bee dance.
7. Arbitrariness: There is an arbitrary relationship between a signal and its meaning. That is, the signal is related to the meaning by convention or by instinct but has no inherent relationship with the meaning, which is for example not the case for bee communication that is often iconic in that the shape of the dance partially refers to the meaning it conveys, while Velvet monkeys calls do specifically refer to three different types of predators (terrestrial or aerial predators compared to snakes that can be both).

25. The reader might want to refer back to very detailed studies on the vocalization tracts in primate and how the evolution to the human one has occurred, see T. Fitch 2000 'The evolution of speech: a comparative review' in *Trends in cognitive science*.

26. Note that in certain species part of the signaling is reserved to male or female individuals.

8. Discreteness: the message is build up from a relatively small amount of discrete units [27] (e.g., phonemes in human language). Exchanging such discrete units causes a change in the meaning of a signal. This is an abrupt change, rather than a continuous change of meaning.
9. Displacement: Communicating about things or events that are distant in time or space.
10. **Creativity and Productivity:** The capacity to generate and understand messages that are totally new and never heard before. It is the property of a so-called *open* system to potentially build an infinite number of utterances with a finite number of basic elements, by combining the elements differently.
11. Cultural transmission: Each generation needs to learn the system of communication from the preceding generation through cultural learning.
12. Duality of patterning (*double articulation* A.Martinet)[28]: Large numbers of meaningful signals (e.g., morphemes or words) produced from a small number of meaningless units (e.g., phonemes). The duality of patterning is essential to language because it allows to produce from a finite number of phoneme an infinite number of morphemes and from this morphemes to produce potentially an infinite number of words and sentences.
13. Prevarication: Linguistic messages can be false, deceptive, or meaningless.
14. Reflexiveness: In a language, one can communicate about communication.
15. Learnability: A speaker of a language can learn another language.

1.1.4 Is language innate or acquired?

> *In the same sens [in which] we say that in some families generosity is innate, in others certain diseases like gout or grave, not that on this account the babes of these families suffer from these diseases in their mother's womb,* ***but because they are born with a certain disposition of contracting them.***
>
> René Descartes

While the demonstration of the use of symbols or short utterances in a creative and non-situation-specific fashion by great apes failed, the question about the origin of Language in humans remained unanswered. Thus, we will continue on the path opened by the argumentation of Descartes by adding a further element of reflection offered by the work of the French philosopher Jean-Jacques Rousseau, that will allow us to bridge the creative aspect of language and the question of language acquisition.

> "It would further seem from the same observations that the invention of the art of communicating our ideas depends less on the organs we use for that communication than on a faculty that belongs to man, which makes him employ his organs for that use, and which, if he lacked them, would make him employ

27. These discrete signals, when they are in small number, can be repeated as it's the case in bird songs.
28. First articulation is assembling phoneme without semantic value into semantic units that are called morphemes; The second articulation is assembling morphemes into meaning-full structures that can be lexically or syntactically like words mots, phrases or discourse.

> others to that same end. Give man a physical organization as entirely crude as you please: doubtless he will acquire fewer ideas, but provided only that there be some means of communication between him and his fellows by which one might act and the other feel, they will succeed at length in communicating altogether as many ideas as they have to one another. Animals have a physical organization more than sufficient for such communication, and none of them has ever made this use of it. Here, it seems to me, is a most characteristic difference. Those who, among them, work and live in common, such as Beavers, ants, and bees, have some natural language in order to communicate amongst themselves —I raise no doubt about it. There is even reason to believe that the language of Beavers and that of ants are in gesture and speak only to the eyes. Be that as it may, ***precisely because all such languages are natural, they are not acquired; the animals that speak them do so from birth, they all possess them, and everywhere the same one; they do not change them, nor do they make the slightest progress in them.*** *Conventional language belongs only to man*. That is why man makes progress, whether for good or bad, and why the animals do not at all."
>
> Jean-Jacques Rousseau, Essay on the Origin of Language, 1781.

This short quote from Jean-Jacques Rousseau in his *Essay on the Origin of Language* reveals some aspects of human language that brought the philosopher to define it simultaneously as an *acquired and a perfectible* human behavior. The idea expressed here is that the distinction between man and animal is less a matter of thinking, than a question of freedom and of linguistic *perfectibility*. In fact, this statement about the impossibility for animals to have language variation and to modify through evolving conventions the communicative system that is in place since birth, makes our questioning about human language evolve and shift towards a fundamental question: is Language innate or acquired?

The opposition between the natural and acquired character of language has early been addressed by philosophy. In this regard, the two most enduring language models in the history of thoughts can be summarized into two metaphors, that of the wax slate and that of the marble block. The tenants of the first, like John Locke[29], commit to the idea that all knowledge is gained from experience and see infants knowing nothing as the wax patiently awaiting for experience to write on it. While the tenants of the second believe, like Leibniz, that the child, as a marble block, is grained, so that through experience only will some shapes emerge from it. Leibniz expressed the debate with Locke in these terms:

> "The question is to know whether following Aristotle and Locke, the soul in itself is entirely empty, like a *tabula rasa* and whether all that is traced thereon comes solely from the senses and from experience; or, whether like I, with Plato, believe, the soul contains originally the principles of several notions and doctrines which external objects merely awaken on occasions."
>
> Leibniz , *New Essay. Notes directed against Essays concerning Human Understanding*, 1703.

Interestingly, the wax slate versus marble bock debate surfaced again outside the philosophical field in the discussions animating academics concerned with human lan-

29. Book I of the *Essay Concerning Human Understanding*, 1689.

guage behavior in the 50's. As briefly presented above, the behaviorist psychologist B.F. Skinner thought that reinforcement was a big part iofn how children learn a language, while N. Chomsky was arguing that such a statement was absurd, particularly in light of the fact that parents could only teach a small subset of all linguistic information to their children.

Poverty of the stimulus

Once this observation, conjoined to the characteristically effort-less language acquisition of children, was brought into the debate, it offered some central arguments to defend one position compared to the other. This evidence opposed a radical objection to the argument that learning started on a tabula rasa. According to N. Chomsky, the best approach to this longstanding philosophical debate was to begin by asking how a child learns to talk. The question of *innateness* of language could be resumed in an experimental question about child language acquisition.

This last argument will be known under the "poverty of the stimulus" argument. Its main contribution is to point that external linguistic stimulation could not be sufficient to cover all aspects and rules relative to sentences construction in a given language. Humans should therefore have some other form of aid when acquiring the rules of their native language. Most importantly, inserting the child acquisition problem at the core of the issue of language complexity showed that such a learning requires a theory of linguistic structures[30]. In short, what a child can learn is incommensurately more complex than what are his environmental primary linguistic data[31].

In this way, what Chomsky was defining to be innate is a certain knowledge or system of information that every child has on a set of possible universal properties of human languages. It is the interaction of this innate knowledge with the Primary corpus of linguistic data - the observed adults' utterances - that explain the development of linguistic capacities. Thus, one can say that the generative theoretical undertaking considers language as a product of nature and nurture. The particular grammar any individual has is a sort of linguistic phenotype transformed by experience, that is issued from a linguistic genotype. mlkjhgfds Using one of Jacques Mehler's formulations about children learning process, this could qualify as a "learning by unlearning". It would proceed by a gradual unlearning of what is innate, fixating in this way only the particular rules of a given native language (Mehler, 1982). Namely, in the Principle and Parameters framework[32], Chomsky (1993) theorized a number of open parameters, whose values are to be set by the nurturing of primary linguistic data. Hence, each attained grammar – say, for example, French or Chinese grammar – is a linguistic phenotype resulting from the setting of all the parameters in conformity to the Primary Linguistic inputs the child receives. Linguistic variation is therefore to be considered as a function of this parametric variation. Note that the theoretical interest of comparing the brain networks for Mandarin Chinese and French sentences resides precisely here (see Annexes on Chinese and French Localizers' direct comparison §H).

We will now see from a cross-linguistic perspective concretely how languages do not vary at random, but within certain bounds and according to certain patterns. We will

30. This last point will be addressed by focusing on certain syntactic structures in Chapter 2 of Part I.
31. Primary linguistic data (pld): the data children are exposed to while they are learning their native language.
32. see a definition in §1.2.4 (p.31).

harvest the contribution that the study of linguistic diversity and its invariance offers to the question of Human grammar.

1.2 Possible Languages: the 'certain cut' of language faculty

> "Timeo hominem unius linguae."
>
> *["I fear the man having only one language."]*
>
> Lucien Tesnière, 1943

We will devote the next pages to the analysis of the precise limits to the observed cross-linguistic variation, also to see how this 'certain cut' of language has been calling for different principled explanations in Philosophy, Linguistics and in Cognitive Science.

Thanks to the observation of linguistic diversity across the world's languages, we will first address the question of Language universals and linguistic evidence they offer for a universal language architecture.

1.2.1 Language universals and Universal language architecture

> *It must be obvious to anyone [...] that there is such a thing as a basic plan,* ***a certain cut****, to each language. This type or plan or structural 'genius' of the language is something much more fundamental, much more pervasive, than any single feature we can mention. [...] All languages differ from one another but certain one differ far more than others. This is tantamount to saying that it is possible to group them into morphological types.*
>
> Edward Sapir, 1921: 120

The idea that languages that seem very different are in reality variations of a unique theme is antique. The attempts of philosophy to reconstruct a unique grammar of logic relations date probably back the XIII$^{\text{th}}$ century scholar Roger Bacon – the so-called *Doctor Mirabilis*, for those who are acquainted to the logic and optics of medieval scholastic period[33].

His work on linguistic matters was not only substantially heralded as an early exposition of a universal grammar, but in his observations on dialects he showed that one could find properties of the *mother* of the dialectal forms[34].

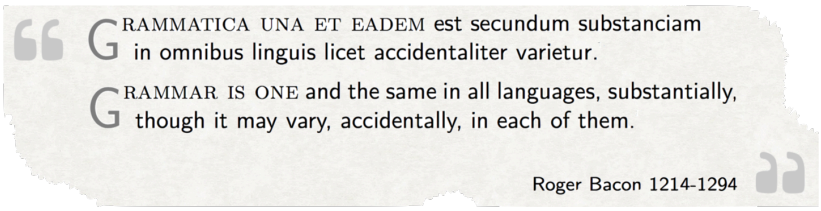
33. We will go back to discuss the issue of grammar of logic in the next Section when making the link between Syntax and Computation.

34. The existence of different *locutiones* in one *lingua* was simultaneously made by Thomas of Aquinas: "*in eadem lingua saepe locutio fit, sicut patet in Francia et Picardia, et Burgundia, et tamen un loquela*

This universal thesis, dear to the Modistae[35], did not prevent him from being interested in all the individual languages he wrote a grammar of. He was so convinced that this kind of universal grammar was underlying all grammars, that he wrote he could teach all that was needed to be known about Hebrew in three days to anyone, who was motivated (and intelligent) enough (*Opus tertium*, pp. 65-66) [36].

This possibility of having a "contrastive" approach to the teaching of different languages is the patent proof of his involvement in fleshing out his thesis about a core central grammar constituting the backbone of all languages. It is this backbone that was kept constant when Bacon undertook to write a Greek grammar starting from the knowledge of Latin the students had, which title could therefore be reworded into a contrastive Greek-Latin Grammar[37].

In two of his linguistics masterpieces, the 'Overview' and the 'Greek grammar', we find an early exposition of the concept of a universal grammar underlying all human languages. The *Greek grammar*[38] contains, in fact, the famous sentence reported in Medieval Latin above.



It is important to say that this is the first definition to be born out of empirical observation[39], and given his rich work on languages and dialects, he can definitely be considered the forefather of comparatist and contastive approaches in Linguistics.

This points out the fact that the difference between languages is not essentially qualitative and the study of this difference will develop across centuries in different disciplines like philosophy, logic, linguistics.

Linguists from the XVIIIth century manifested a certain eagerness towards what will be called Language Universals in the 20th century, and the philosophical tradition described above found a first linguistic formulation in the work of Friedrich Wilhelm Christian Karl Ferdinand von Humboldt (1767–1835) .

est" (in Lectura super Mattheum). But, Bacon's work was more precise on this point because he already differentiated linguistic kinship from variation in usage.

35. A current of thought of the second half of the XIIIth century in Paris University, that had developed a linguistic theory called *grammatica speculativa*. Its aim was to explain the foundation of grammar by concentrating on the construction rules of utterances. In this way, they wanted to state the autonomy of grammar from logic.

36. One can picture the kind of Parisian students of the time if a teacher could write down such a statement.

37. Following Hovdhaugen (1990). Hovdhaugen, Even (1990), "Una et Eadem: Some Observations on Roger Bacon's Greek Grammar", in *De Ortu Grammaticae: Studies in Medieval Grammar and Linguistic Theory in Memory of Jan Pinborg*, Amsterdam: John Benjamins Publishing, p. 117–132

38. The full text reedited in 1902 can be consulted online at the following address `https://archive.org/details/greekgrammarofro00bacouoft`

39. Philosophically speaking, it is necessarily an a posteriori definition (being a philosophical background, and not an apriori and philosophical one as that of the Modistae).

1.2.2 Humboldian Pre-Typology

> *"The terminology in which at present we try to speak of human affairs – 'consciousness', 'mind', 'perception', 'ideas', and so on –will be discarded [...] and will be replaced by terms in linguistics. Non-linguists constantly forget that a speaker is making noise, and credit him, instead, with the possession of impalpable 'ideas'. It remains for the linguist to show, in detail, that the speaker has no 'ideas' and that the noise is sufficient."*
>
> Leonard Bloomfield, 1936

Humboldt's studies on language and his subsequent linguistic philosophical reflections are good company with the traditional Cartesian way of understanding language presented in the Epilogue.

Cartesian linguistics is intimately connected with the notion of Universal or Philosophical Grammar. Given its revival in Chomsky's generative approach to language and the identification of Humboldt as one of his immediate precursors, we should here clarify Humboldt's relation to this tradition. First of all, Humboldt was decidedly critical towards all attempts to construct a system of Philosophical Grammar supposedly underlying all natural languages, because it was patterned after the concepts of Latin and French grammar, and this had resulted in the writing of grammars that violated the nature of Non-European languages by forcing them into the schemes of western system, whose categories were completely alien to their inherent structures (Humboldt, GS Vol 5: 355). However, this does not mean Humboldt rejected a concept near the idea of linguistic universals: that of a "prototype of all language", *Urtypus aller Sprachen*. He theorized the existence of a communicative prototype of human speech, embedded in the structure of language, and manifesting itself in the different languages. One example of its manifestation finds an expression through the system of personal pronouns, namely in the differentiation between the first and the second/third person. He notes that Cartesian grammatical analysis would fail analyzing this linguistic phenomena, because from a logical and grammatical point of view, it makes no difference who the speech is directed at, whatever the pronoun is (first, second or third personal pronoun). In the Cartesian framework, in fact, a given personal pronoun will function anyway as the subject of the sentence. On the contrary, for Humboldt, the first ('I') and third person ('he') are really different entities, and he argues that they exhaust all possibilities, in that they constitute the I and the not-I. For instance, the second person pronoun 'thou' is also a not-I pronoun, but unlike 'he', it does not refer to the "sphere of all beings", it is rather in another sphere of action and interaction. This is the reason why, in his empirical investigations, Humboldt paid special attention to the system of personal pronouns in different languages. He believed that one could reconstruct the specific manifestation of a prototypal speech situation with I and a not-I. Later on, these observations on pronouns will be summarized in Grennberg's Universal 42 :

It should be said that this definitely resembles more the quest of Cognitive Linguistics,

42. All languages have pronominal categories involving at least three persons and two numbers.

than that of Chomskian Universal Grammar. In fact, the importance of linguistic studies lies for Humboldt in the discovery of which part language plays in the formation and transmission of ideas (*Vorstellungen*), not only in a metaphysical sense as conditioning the forging of concepts, but also in the way in which an individual language imparts its imprint on concepts (GS Vol. 6 page 147).

Furthermore, the oft-quoted "infinite use of finite means" refers to a discussion about the role of language in the constitution of thoughts and not mentioning at all the sentence. Humboldt states in this passage (GS Vol. 7 pages 98-99) that language plays the role of constituting thoughts and as thoughts are in principle boundless, language "must therefore make infinite employment of finite means". This being said, this concept can be legitimately seen as introducting the idea of linguistic productivity in general, but not in the field of syntax.

The kind of linguistic universal he was referring to (i.e. I and not-I) was tightly linked to the fact that each language, by its structure and shape, was able to represent a specific *view of the world* - *Weltansicht*-, and this abstract vision of the world constituted the core of his concept of linguistic variety. In Kantian terms, he believed in the universality of the mental structures (i.e. I and not-I), and in Kantian categories representing the rules and the laws of thinking. He further identified these rules of thinking as being ultimately responsible for the system governing linguistic utterances. Hence, he rejected the idea that these structures could already be, by themselves, a kind of logical grammar from which a Philosophical Grammar could directly be born out.

It is important to understand what brought Humboldt to the idea of a 'Universal Grammar' in his concrete linguistic comparative work[40]. Adapting the notion of *Type* – which he first formulated in his work in comparative anatomy at Jena in 1794[41] – to the study of language, he forged a 'Universal Grammar' that would play the role of a *tertium comparationes* for the linguist's comparative study of the languages and language groups. This kind of linguistic prototype was established through a combination of philosophical reflection and concrete methodical linguistic analysis. And, the task of the linguist, then was "to study each language as a fragment of the universal language of the human species" (Humboldt, Essay on the languages of the New Continent)[42]. In this framework, the investigation of a single language with its specific form should be guided by the knowledge of prototype in them, and this work should also contribute to our knowledge

40. One should bear in mind that Humboldt did not work in isolation, but constantly entertained lively and epistolary contacts with leading scholars in Europe and America. Just to cite some of them: Franz Bopp and August Wilhelm von Schlegel, respectively in Berlin and Bonn, Jean-François Champollion and Jean-Pierre Rémusat in Paris, and Peter S. Du Ponceau or John Pickering in Philadelphia and Boston. Humboldt's epistolary exchange about Chinese and other languages with the French linguist Abel-Rémusat is published in French by J. Rousseau and D. Thouard, under the title "Lettres Édifiantes et Curieuses sur la Langue Chinoise. Humboldt-Abel Rémusat (1821 – 1831)". Villeneuve and Paris: Presses Universitaires du Septentrion, 1999.

41. He used the notion of *Type* in his Plan for a Comparative Anthropology (1795) too.

42. In French "*comme un fragment du langage général du genre humain*" (Essai sur les langues du Nouveau Continent, in Stetter 2004 p.238). In this essay he states that linguistic investigation is ultimately philosophical, because languages are not just means for communication, but ways of conceiving the world ('Weltansichten'), sorts of cognitive creations. Languages demonstrate – in the tradition of Leibniz – the "marvelous variety of the human mind".

of the prototype.

As the title of his treatise shows – *Grundzüge des allgemeinen Sprachtypus*, [*Fundamentals of the general Linguistic Prototype*] -, Humboldt vehicles the idea of a prototype of language, which is conceptually very similar to Goethe's idea of a proto-plant *Urpflanze*: a plant not to be confused with a real plant but that embodies the essential features that one can find in all existing plants[43]. However, linguistic form is not something material like a plant, and for Humboldt it mainly pertained to a performance (*Verrichtung*), namely the production of speech. Humboldt finally definied the prototypes as the gathering of elements and rules that can be considered common and essential for speech production in all languages. It is on this last point that Humboldt's philosophical ideas join Chomsky's thought about language: the linguistic prototype is a generative notion rather than a substantive one.

1.2.3 Typology of Cross-Linguistic Patterns

> *"Underlying the endless and fascinating idiosyncrasies of the world's languages there are uniformities of universal scope. Amid infinite diversity, all languages are, as it were, cut from the same pattern."*
>
> – Memorandum to the 1961 Dobbs Ferry Conference on Language Universals (Greenberg, 1963/1966)

During the 20th century two currents ran side by side in the quest for *linguistic invariants* [44].

On one side, the accurate investigation of cross-linguistic properties of natural languages in the so-called Linguistic Typology tradition; and on the other side, the Chomskian approach more closely concerned with the structure of the mind and psycholinguistic evidence[45].

Parallely, the Generative program embodied the need for a theoretical background that would account for impossible natural languages by ruling them out, and in parallel could be empirically falsifiable by making empirical predictions about diachronic language change[46] and language acquisition.

Despite the many differences characterizing these two linguistic traditions, it should be noted that they show astonishing convergences on the fact that not all the possible imaginable grammar and linguistics rules are realized in the natural grammars of the world's languages.

43. Goethe claimed, however, that Humboldt type is of a different nature from his *Urpflanze*, that was something quasi-real perceivable through one's mental eyes.

44. Term chosen in reference to the french linguist Gilbert Lazard, from the Institute of Oriental Languages (INALCO) who published in 2006 an article titled '*La quête des invariants linguistiques*'.

45. Note that a research direction involved with filling the gap between language and the mind can be also identified in the current of Typology in the discipline that is called Cognitive Linguistics. This current proposes to compare the structures and functions of very distant languages to find universals, to then investigate their cognitive bases.

46. This is an often forgotten central evidence for a rule-governed evolution of natural languages.

Inside the first current, one can find differences between the typological approach and the one studying language universals. Superficially speaking, these differences could be expressed in the following terms: studying universals means being concerned with what human languages have in common, while the typologist is more concerned with the ways in which languages differ from each other. Nonetheless, one of the focus of Typology is the fact that the difference between languages is not a random variation, but is subject to limitations.

This quest for linguistic invariants can be found in the American structural linguists Leonard Bloomfield and Edward Sapir, two eminent proposers[47]. The first, Leonard Bloomfield, initiated an approach to linguistics arguing that the study of languages could be scientific. Thanks to his historical work on Indo-European languages, and to field work on Austronesian languages and the Algonquian family, he proposed in 1933 that grammars forged by linguist may be supported or refuted by detailed observation of real languages on the field (L. Bloomfield, 1933). In syntax, he laid the foundations for a theory of constituent structure, including rudiments of Xbar-Theory[48] (Bloomfield, 1933: 194-195). An anecdotal fact could help us understand how significant his contribution has been. One of his eminent students, Charles Hockett, would report years after, that Bloomfield's courses on syntactic analysis made him feel that he was doing syntax for the first time in the history of linguistics (Hockett, 1968:31).

Linguists like Edward Sapir and Leonard Bloomfield started facing the complexity of languages such as Lakhota, Paiute, and Menominee along with French, German, and Latin. According to the limitations on the degree of variation found cross-linguistically, the languages of the world were about to be divided into various types. This is precisely the moment when the American linguist and anthropologist Joseph Greenberg (1915-2001) showed that some patterns, and most of all, some combinations of patterns could be identified among languages. These taxonomic research results were rapidly connected to the works of historical linguistics and its genealogies of language families. Reconstruction studies grew (Greenberg, 1971, 1987, 2002) in such a way that studies on language diversity, linguistic universals and typology experienced a new start (Comrie, 1981; Comrie et al., 2003). This was the case also for a series of Indo-european issues that had been left aside for a period and came back on the stage in the framework of a large-scale genealogical reconstruction of linguistic families (Greenberg, 2002).

This work attracted the attention of other disciplines and with genetic studies of populations (Cavalli-Sforza et al., 1994), the question on the origin and evolution of languages was brought back on the stage (Cavalli-Sforza, 2000), till the latest developments of the so-called genome-linguistics trying to match P-parameters from the formal linguistics framework (Longobardi et al., 2016)[49].

Based on statistic observations of different syntactic properties, Greenberg was able to group up languages according to 45 patterns of typical linguistic properties. These patterns were taken as the evidence that some combinations of linguistic rules do exist while some others do not, giving evidence to what Andrea Moro's has metaphorically

47. This current will be later called American descriptivist school.

48. For an introduction cf. §2.3.1, p.134).

49. Note that Charles Darwin in *The Origin of Species* had already predicted the eventual matching of phylo-genetics with linguistic families to reconstruct human past: "If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man, it would be the best classification of the various languages now spoken throughout the world; and if all extinct languages, and all intermediate and slowly changing dialects, were to be included, such an arrangement would be the only possible one" (Darwin, 1859, ch. 14).

called the 'Boundaries of Babel'. We will now concentrate on the linguistic details of one of these patterns.

Some examples of language universals

Joseph Greenberg published in 1966 a corner-stone book entitled *Language universals*. Starting with about 30 languages from different linguistic families – both historically and geographically distinct –, he proposed a list of 45 *invariants*. This list of statistical regularities bears on multiple aspects of the organization of languages, going from word-order, morphology, casual system, to grammatical gender and number, etc.

As one can concretely see in the first 5 invariants reported below, the majority of Greenberg's proposition bears on word-order considerations: (a) the order of the principal elements of the sentences (Subject, Object and Verb), (b) the order of the nuclear element and determinants of the nominal Phrase, but also (c) on the variation in word-order found in interrogative and subordinate clauses.

> **1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.**
> **2. In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.**
> **3. Languages with dominant VSO order are always prepositional.**
> **4. With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional.**
> **5. If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun.**

Other propositions contained in his list of *Language universals*, focus on the analysis of the presence or absence of certain categories under certain conditions:

> **29. If a language has inflection, it always has derivation.**
> **34. No language has a trial number unless it has a dual. No language has a dual unless it has a plural.**
> **36. If a language has the category of gender, it always has the category of number.**
> **43. If a language has gender categories in the noun, it has gender categories in the pronoun.**

What we see from the above *invariants* are a series of considerations about morphological typology, that is to say the analysis of how the grammatical expression of a certain meaning, like plurality, is found across languages in a number of different ways. For instance, the expression of the difference between singular and plural in nouns can be found in (1) a non-overt expression in languages like Chinese, while (2) the overt presence of a function word can be found in languages like Tagalog, alternatively (3) the option of affixation can be adopted by languages like in French or Swahili, or even (4) sound change can also be a strategy, like in certain well-known lexical elements of English (i.e. woman, women), and finally other languages can opt for (5) reduplication, like it is the case in Malay, where *anak* means 'child' and *anak-anak*, 'children'.

However, what Greenberg did, and what has emerged from work rooted in this first book on Universals (see Dryer (1992a) left vs. right branching; Dryer (2007) and oth-

ers, for 'head-finality' vs. 'head-initiality'), was not only showing the existence of some patterns, but proving that some combinations of patterns can be identified among languages.

Let us now enter in a specific example of this fascinating world of correlations of word-order patterns. For instance, the order of the direct object with respect to the verb was claimed to correlate to varying degrees with the relative order of many other pairs of grammatical elements, as shown in the following table:

VO Word-Order	OV Word-Order
a. P > DP (Prepositional Phrases)	DP > P (Postpositional Phrases)
b. Aux > V	V > Aux
c. copula > predicate	predicate > copula
d. V > manner adverb	manner adverb > V
e. (more) A (than)	'Standard of Comparison' (than) A (more)
f. A > PP	PP > A
g. V > complement/adjunct PP	adjunct/complement PP > V

French happens to confirm the relative word-order pattern illustrated in the above seven points.

In all these examples and in the full-fledged sentence below, the Head precedes the Complement in seven different constructions : Verb before Subject, Preposition before Noun Phrase, etc.

(1) Jean trouve cette lettre sous le lit.
John finds that letter under the bed
'John found that letter under the bed.'

On the contrary, if we observe a classical example, from Baker's book *The Atoms of Language*, the word-order in Lakhota, a language of a Sioux tribe of native Americans, all these ordering relations are inverted inside all phrases types:

(2) John wowapi k'he oyke ki ohlate iyeye.
John letter that bed the under found
'John found that letter under the bed.'

Following, are reported examples[50] of Lakhota post-positional linguistic phenomena[51] where ad-positions occur after head nouns[52]:

(3) Lakhota
a. wicȟa'sâ kiŋ
man the
'the man' (NP > Det)

50. Source: `http://wals.info/refdb/record/Rood-and-Taylor-1996`

51. Rood and Taylor (1996) suggest the following template for basic word order: (interjection) (conjunction) (adverb(s)) (nominal) (nominal) (nominal) (adverb(s)) verb (enclitic(s)) (conjunction), published online at `http://lakxotaiyapi.freecyberzone.com/sk1.htm#83` and `http://lakxotaiyapi.freecyberzone.com/sk1.htm#5`,

52. For further reading, the reference authors on Lakhota grammar are: Rood, David S. and Taylor, Allan R. 1996. Sketch of Lakhota, a Siouan Language. In Goddard, Ives (ed.), Handbook of North American Indians. Volume 17: Languages, 440-482. Washington: Smithsonian Institution.; and Van Valin, Robert D. 1977. Aspects of Lakhota Syntax. University of California at Berkeley.

b. mas'opiye el
store at
'at the store' (DP > PP)

c. típi kiŋ okšaŋ
house the around
'around the house'

d. Othuŋ'wahe etaŋ'haŋ wahi'
town.from arrive I
'I arrive from town.' (DP> PP > VP > Subj)

e. Típi kiŋ okšaŋ e'nazîŋ
house the around stand.they
'They (collectively) stood around the house.'

f. wicȟa'sâ kiŋ o'tapi.
man the many.PL.are
'There are many men

It has to be noted that across the languages of the world, approximately 95%, according to Baker (2002), belong to one of these two word-order classes. Interestingly, Cinque (2013) proposed two potentially new arguments for Universal Grammar that are both rooted in this type of cross-linguistic large-scale observation. Added to the argument of the 'poverty of the stimulus' (see sub-section 1.1.4), the first argument, is based on the fact that, of all the concepts and distinctions that populate our system of thought, only a limited part receives a grammatical encoding across the languages of the world, thus, the different concepts receiving a grammatical encoding should be arguably the same in all languages. The second argument states that quite rigid limits on word-order variations exist, so that as a matter of facts some potential orders are never found. If we observe the statistics reported on the WALS[53], we can discover that of all mathematically possible orders of constituents only a subset is actually attested:

1. SOV: 497 languages
2. SVO: 435 languages
3. VSO: 85 languages
4. VOS: 26 languages
5. OVS: 9 languages
6. OSV: 4 languages.

This statistical distribution confirms the idea previously discussed of possible languages having 'a certain cut', showing that there should be strictly impossible ordering constraints across languages. So "that there is such a thing as a basic plan, **a certain cut**, to each language. This type or plan or structural 'genius' of the language is something much more fundamental, much more pervasive, than any single feature we can mention" (Saphir, 1921).

In conclusion, borrowing the expression from Edward Sapir, this impressive regularity can only bring us to acknowledge that we are here confronted to some large-scale underlying ground-plan, despite the existence of several discrepancies or exceptions. We could also add that, from a cognitive point of view, these patterns are not only highly informative of what can be called "a possible language", but are also urging neuro-linguists

53. Dryer, Matthew S. and Haspelmath, Martin (eds.) 2013. The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at `http://wals.info`, Accessed on 2015-05-14.)

to ask the question relative to the implications that patterns have on the neuro-biological architecture of the human brain: what could be the neural substrates of these linguistics boundaries?

1.2.4 The Cartographic Project and Universal Grammar

> "Trouver ce qu'il y a dans toute langue, et dans le langage général, de régulier, de géométrique, d'architectural"
>
> *["Find what exists in every language, and in language in general, that is regular, geometrical and architectural."]*
>
> Charles Bally 1913, pp.23-24

Universal grammar

When speaking about Universal grammar, it is difficult to avoid mentioning that this concept has not only Medieval, XVIII$^{\text{th}}$ century's and Modern linguistic's roots as we just saw, but it has also modern philosophical ones. Indeed, what Otto Jespersen formulated as *principles* in the language domain – a series of principles underlying the grammar of every language – can easily qualify as the Modern philosophical seed of Universal Grammar (UG), as it was formulated in the Generative framework. This sub-section, first presents the theory that Chomsky developed around the innate factors in human language faculty that determine a class of possible languages.

Shifting from Linguistic Invariants to Chomsky's program, we accomplish more than a shift in linguistics traditions, but a change in perspective. As already presented at the end of the previous Section on the problem of innateness in language acquisition, Languages are considered as having a biological substrate determining part of their underlying principles. Evidently, Chomsky's idea of a deep and universal grammar being the bedrock on which all the grammars particular to each language would develop, can be viewed as joining the linguistic tradition of Language universals. It nonetheless adds a new, external, neuro-biological constraint on language variation: it postulates innate principles representing innate linguistic properties[54].

In his seminal book, on the biological foundations of language, Eric Lenneberg (1967) starts with an important remark on what is a Universal in languages from a *bio-linguistic* point of view:

> "A biological investigation into language must seem paradoxical as it is so widely assumed that languages consist of arbitrary, cultural conventions. Wittgenstein and his followers speak of the word game, thus likening languages to the arbitrary set of rules encountered in parlor games and sports. It is acceptable usage to speak of the psychology of bridge or poker, but a treatise on the biological

54. Chomsky (1964) identified the goal of achieving explanatory adequacy in linguistic theory in accounting for the acquisition of a grammar in relation to Universal Grammar (UG).

> foundation of contract bridge would not seem to be an interesting topic. The rules of natural languages do bear some superficial resemblance to the rules of a game, but I hope to make it obvious in the following chapters that there are major and fundamental differences between rules of languages and rules of games. The former are biologically determined; the latter are arbitrary."

In the Principles and Parameters framework, the syntax of a natural language is described in accordance with general principles – abstract universal properties that are innate –, while language-specific properties are expressed in binary "Parameters" that can be set in various ways (either turned on or off), giving for instance the different patterns we observed in the previous Sub-section[55].

As already noted in a previous section on the innate versus acquired character of language (§1.1.4 p.19), the Principle and Parameters proposition was theoretically and significantly simplifying the learning task, while at the same time providing typological insights in the form of "parametric clusters". We can therefore say that it provided a way of connecting bio-linguistics, the study of language faculty, with language typology results; thus both insights from language typology and language acquisition were gathered to fuel linguistic knowledge on human grammar. This complementary vision was further developed by a new project born among generativists in the 90's (Rizzi, 1997 and Cinque, 1998), the so-called Cartographic Project. We will now present to how this project expresses the potential of cross-fertilization of the descriptive linguistic tradition of Typology and the Generative approach till today[56].

Cartographic Project

The "cartographic project" follows, in fact, from the idea that all languages share common principles in phrase and clause structure building. It developed the central idea of defining a fine-grained map of clause structure (Cinque and Rizzi, 2009) , in the attempt to ultimately draw a map, as precise as possible, of the syntactic configurations underlying the languages of the world.

For instance, an example of this mixed approach could be Cinque's (2005) analysis of Greenberg's Universal number 20 inside the Generative framework. This universal concerns the order of demonstratives, numeral and adjectives with respect to the Nominal element, and among the 24 mathematically possible combinations it is remarkable that only 13 are attested across the languages of the world. Even if variation in the relative ordering of those elements is important, this phenomenon is explained by Cinque in the following manner: only those orders that could be obtained starting from a unique base order [Dem Num A N] and then moving[57] the Noun Phrase NP leftward to higher functional positions are attested.

---

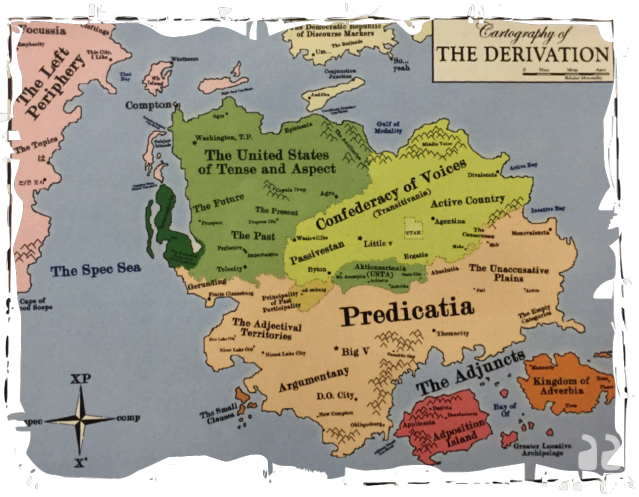
55. For critical readings on Principles and Parameters framework see: Haspelmath, Martin, 2008, "Parametric versus functional explanations of syntactic universals" in Biberauer, Theresa (ed.), *The limits of syntactic variation*. Amsterdam: John Benjamins, Amsterdam, p. 75-107. For recent developments and refinements of Principles and Parameters, see : Biberauer et al., 2014, "Complexity in comparative syntax: the view from modern parametric theory" in Frederick J. Newmeyer and Laurel B. Preston (eds.) *Measuring Grammatical Complexity*, Oxford: Oxford University Press.

56. In this regard several linguists are revisiting Greenberg's (1963) universals, for an interesting approach see Whitman (2008).

57. This movement of the NP being realized in one of the ways interdependently admitted by the syntax of the different natural languages.

Hence, the work on word-order patterns from inside the generative framework (Cinque, 2010) offers some evidence that considering (1) the systematic word-order differences across languages, and (2) the observed relative order of functional morphemes that are overtly realized in the languages of the world, could actually reveal universal hierarchies of sentence functional projections. In fact, the cartography of syntactic structures is a line of research concretely building up an inventory of functional elements to be found inside syntactic structures. These functional elements (heads and specifiers of functional projections) are identified as opposed to the open class lexical ones inside a given language. Taking the open versus close class diagnostic to define Functional elements automatically reveals a huge amount of elements: according to Heine and Kuteva's (2004) evaluation, it should amount to approximately 400 targets of grammaticalization. This gives an idea of the consecutive number of functional projections to be mapped in sentence structures[58].

In the process of identifying functional elements, crucial evidence comes from comparative and typological studies, which provide empirical arguments to test for the cross-linguistic validity of the syntactic configurations brought to light by cartographic studies (i.e. Cartographic Maps) , which could be humoristically summarized in the following sentence map.



As evident in the map presented here, syntactic constructions are complex objects, and as such, a central question in syntactic theory concerns the right structural maps for natural language syntax.

Among the first general observations about sentences across world's languages is a fine-grained analysis of the building blocks of clause structure. It reveals that sentences can be analyzed as bipartite: (1) a nucleus where predicational relations are established and (2) a periphery where properites related to Scope-Discourse semantics are encodes (e.g. clause-typing, sentence-discourse interfacial mechanisms in Topics, etc.).

The different territories of the sentences have been gradually discovered starting from a fundamental assumption from the late 1980s: the structure of phrases and clauses began to be considered as hierarchical sequences of same building blocks according to the fundamental X-bar building scheme[59]. And the main functional elements of the clause were successively split into three main Sentence Domains (functional projections): CP (Complementizer Phrase), IP (Inflectional Phrase) and VP (Verb Phrase)[60]. The need for more articulated phrasal structures in terms of functional projections was argued in Larson (1988) with respect to verb-phrase

58. However one should consider other grammatical categories such as adverbs, adjectives and to some extent verbs as possible candidates to have a functional status, either given their belonging to a close class in certain languages, or because of the very rigid ordering constraints that are applied on one of those categories in the grammar of a given language. The silent character of a certain functional category in a language has been commented by Kaye 2005a and 2006.

59. This will later be replaced by elementary applications of Merge.

60. Chomsky, (1986b) and see Pollock (1989) in order to account for different morphological forms of French Verbs. For a definition of functional projection see §2.3.3.1 (p.140).

domain, and for the Left-Periphery of the sentence in Rizzi (1997).

Converging linguistic evidence showed that the three sentence domains (VP,IP and CP) were better analyzed and conceptualized as sequences of functional projections each one of them following the Xbar structure, and this crucial step of splitting into more articulated hierarchical sequences of functional projections the Sentence Domains. For example, in the inflectional space, a fine splitting corresponded to the identification of a sequence of functional Heads expressing properties of mood, modality, tense, aspect and voice. Belletti (1990) proposed that the higher functional projection of the clause would be the one responsible for subject-verb Agreement (ArgS at the time) and the lower one, the one expressing Tense. Importantly, this ordering is verified and directly observed in languages that express these syntactic properties and relations in the order of prefixes and particles, like Bantu Languages[61].

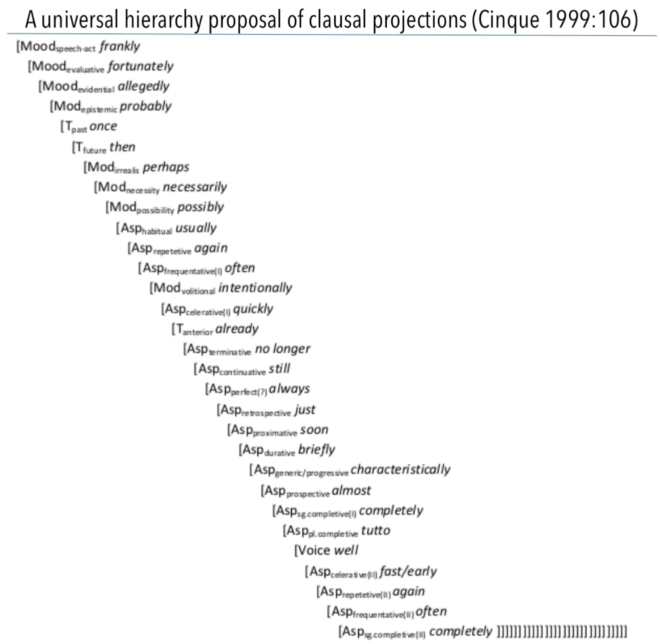


Figure 1.4 – Cartography Project mapping of functional projections of the adverbial domain.

Moving from the core of the sentence toward the periphery, we can observe a tripartite structure of the clauses —argument, predication and illocutory. Cartographic mapping metaphorically places it under a microscope to discover the nature of syntactic structures. The result was that previously ignored linguistic phenomena —such as the syntax of copular sentences, the syntax of finite verbs, or focus/topic constructions —were accounted for in the principle-based framework[62]. Rizzi's (1997) split CP (ForceP > TopP > FocP > TopP > FinP > IP...) is an example, that can be seen as the seminal work having initiated the cartographic enterprise[63].

Echoing Charles Bally's words in the above epigraph, we could say that this project and its formalization started demonstrating the possibility of rendering the entire sentence structure in *geometrical way.* This approach, in fact, builds now on a relatively consistent amount of evidence that there exist distinct hierarchies of functional projections dominating different sentence domains (VP, NP, AP, PP, IP and CP) that could be candidate to be universal not only in the type of heads they involve, but most of all in their number and in their *relative ordering.* This last point remains valid even if languages differ in the type of movements they admit (see Rizzi 1997, Cinque 1990/2003)[64]

61. It is the case in Romance languages too, where the order is the mirror one, featuring : the root, then the tense, and the agreement suffix in last position. According to Baker's Mirror Principle (1985), the verb moves to pick up the closest suffix (the one attached to the stem).

62. Andrea Moro (2001), *The boundaries of Babel* (Chap.6:70)

63. The theoretical question about whether such universal hierarchies of functional projections are to be considered as primitive objects of universal grammar or are to be derived by external conditions of interfaces is presently out of reach. However this will not prevent us from testing for the relevance, for neuro-imaging experiment, of cartographic ordering constraints in the Left-Periphery of Mandarin Chinese, in chapter 7.

64. See chapter 3 and 7 for movement versus Base-Generation in Mandarin Topic-Comment articula-

or in the overt realization of Heads and specifiers[65].

For example, Cinque (1999, 2004) proposes a very rich hierarchy of 30 functional Heads (see figure 1.4). It spans from the upper boundary of vP up to CP, and each functional Head licenses a unique Specifier position (Spec), where a corresponding adverb (in this precise case) may be base-generated. While Indo-European languages pursue an adverbial strategy to encode certain functional projections realizing the functional concepts inherent to these Heads, other languages resort to an inflectional realization for the same functional Heads or a lexicalized Head like a particle. To illustrate this we can take an example from Bocci and Shlonsky about *retrospective aspect* (i.e. "to have just"). While French adopts a periphrastic construction like "venir de", Italian employs an adverb "appena".

Since rarely more than two or three adverbs/inflections ever occur together within a single natural sentence, it has to be noted that the methodology for deriving such a hierarchy proceeds via co-occurrence restrictions.

Importantly, within the cartographic project, the problem of child acquisition is always kept in mind, namely we can read:

> "No particular acquisitional issue arises from considering the structure of the clause provided by Universal Grammar as rich and articulated as that in figure 1.4. The obvious consequence from assuming the universality of [the hierarchy of functional projections in figure 1.4] is that less is left for the child to acquire. He/she will only need to recognize and locate in the appropriate structural places made available by Universal Grammar the morphological and lexical material provided by his/her language" (Cinque 1999:107).

Numerous studies in the last ten years have been increasingly confirming the hypothesis that there should be a universal functional design for the clause and its major Domains that can hold across languages[66].

Macro-parametric variation

Before concluding, it should not be forgotten that this attempt to stay at the confluence between formal linguistic approach and typology is not an isolated one. James C.T. Huang has been lately developing a reflection on the existence of Macro-Parameters, overarching the various parametric variations known to date.

Huang (2005, 2014, 2017) first adopts a contrastive descriptive approach reviewing a large number of distinguishing properties of Modern Chinese syntax, with a special focus on comparison with English. He then shows that the analyzed contrastive properties cluster[67] in a way "that bears witness to the fact that languages differ in ways that allow them to remain in 'a certain cut'." (Huang 2014:3). After having captured these properties syntactically, he inserts them into a more explanatory theory of language

tions.

65. This point will be addressed in chapter 4 and 5, speaking about the overt realization of Topic heads in different variants of Contemporary Mandarin Chinese and in some dialects like Shanghainese.

66. See the series of books on the cartography of syntactic structures directed by L. Rizzi at the Oxford Press, or the following series of individual articles: Rizzi, 2004; Benincà and Poletto, 2004; Frascarelli and Hiterhoeltz, 2007; Cardinaletti, 2004.

67. Meaning that the properties are either jointly present in one language or are jointly absent in another.

variation, inside the Principle and Parameters framework. Interestingly, once pointed out and clustered, these properties provide evidence for the existence of some Macro-Parameters, that is a large-scale correlation of parametric differences among languages.

Concretely, in very stimulating talks and articles (2005a, 2006, 2014) Huang explores the idea that multiple parameters such as Head-parameter[68] (Stowell, 1981), pro-drop parameter[69] (Rizzi, 1982), the wh-parameter[70] (Huang, 1982) and the polysynthesis parameter[71] (Baker, 1996), could actually derive from the traditional distinction between analytic and synthetic languages. Early definitions issued from typology were given by Sapir (1921:127-128): "Languages in which a word tends to consist of only one morpheme are called analytic (or isolating), these languages hardly have any inflection. Generally speaking they are mainly found in three separate parts of the world: East and Southeast Asia, West Africa, and South Africa. While languages in which a word tends to consist of more than one morpheme are called synthetic, they feature inflection, derivation and compounding." Hence, this traditional typological distinction is reinterpreted by Huang in terms of two macro-parameters: *synthetic* vs. *analytic*, that would define a general tendency of languages to '*syntactic concision*' at three different levels:

1. (i) Lexical categories, with examples found in various linguistic phenomenal like the development of light verbs, classifiers, etc.;
2. (ii) Functional categories showing a clustering of distinctive properties (tense, agreement, wh-movement vs. wh-*in-situ*, etc.); and
3. (iii) Argument structure, where composed verbs like 'eat restaurant', 'cut knife', can be more or less productive, or where resultatives are productive, or where unselectiveness of subjects and objects, or even robust pro-drop are observable.

It would be interesting to back to the detailed contrastive account of Huang, when adopting a contrastive approach in neuro-imaging. This is what we did in the Localizers of our two fMRI studies for French and Chinese, we reviewed the salient typological properties of Mandarin Chinese to shed light on the neural underpinning of Chinese versus French sentence and the results are only very briefly presented in the Annexes (H) for the curious reader. This represents a possible further developments of this doctoral work.

In sum, these projects and the framework they build are at the crossroad of disciplinary influences and linguistic traditions. While the cartographic project tries to be as coherent as possible with cross-linguistic and acquisition data, it offers also a rich formal description of the hierarchical articulation of linguistic phenomena, such as Topic, Focus or Clitics. Therefore, it constitutes an interesting theoretical background that allowed us to formulate specific and clear hypotheses to be experimentally tested.

68. Providing an open choice for Phrases (XP) between head-initial or head-final.
69. From which derives a cluster of cross-linguistic differences like null-subject, free inversion, long subject-extraction.
70. Distinguishing *wh*-in-situ languages from wh-movement languages by the occurrence of movement either overtly in syntax or covertly at the level of the Logical Form.
71. Concerning word-order, radical pronoun drop, and characterizing a range of languages on the extent to which arguments are directly linked to verbal morphology

1.3 Language faculty within the Mind and Cognition: a matter of syntax and computation

> *"By studying the properties of natural languages, their structure, organization, and use, we may hope to gain some understanding of the specific characteristics of human intelligence. We may hope to learn something about human nature; something significant, if it is true that human cognitive capacity is the truly distinctive and most remarkable characteristic of the species. Furthermore, it is not unreasonable to suppose that the study of this particular human achievement, the ability to speak and understand a human language, may serve as a suggestive model for inquiry into other domains of human competence and action that are not quite so amenable to direct investigation."*
>
> N. Chomsky *Reflections on Language*, 1975b

From the Linguistic System to Language Faculty

Adopting the perspective of Cognitive Science in linguistics, principally, means shifting the scientific question on language to an interrogation on the specific *knowledge system* mastered by the human mind through the language faculty. Essentially, this means investigating the organization of this knowledge system: a knowledge that is not only the result of a particular form of acquisition, but a knowledge that is at work with its representations and processes during linguistic activity both in production and comprehension. This brings us to ask the question of the structural architecture of this linguistic knowledge, and how its different levels are internalized in the speaker's mind[72].

From Language Faculty to the Mind

The consideration that the study of language is eminently a study of the mind, expressed in the above epigraph, was born in the heart of the XVIIth century philosophy, thanks to Thomas Hobbes (1588-1679) and his interrogations about syntax and the mind. The fact that we can be informed about the architecture of the mind by the investigation of the architecture of language, entails a strong and direct connection between the description and theory of grammar studied by linguists, and the neural architecture

72. Concretely going from phonological, morphological and syntactic to semantic or prosodic representations.

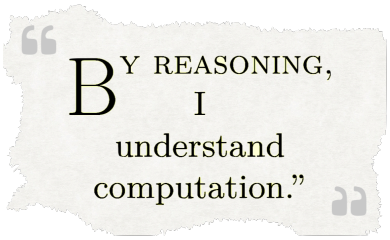
supporting language studied by Cognitive Science. This, in fact, brings one to see the theory of grammar as a theory about brain organization, and, conversely, to take data regarding neural structure and function to bear directly on the theory of grammar. In short, considering the question of why studying language, and specifically sentence unit in the framework of Cognition corresponds, to posit language as an internalized system, and ultimately, to view "language as a mirror of the mind". We will now retrace the successive *shifts* that brought to this assertion.

1.3.1 From cognition to computation

Before entering in the details of the computational aspects of Cognition and their link to language, the following Sub-section focuses on the presentation of one of the first formulations of the idea that reasoning can be equivalent to a computation.

Reasoning as a computation

Diving in these first formulations will help us grasp the deep implications of this point of view (1) for our understanding of language and (2) for the central place grammar and computation should have when considering *sentence as a cognitive object*, as we propose to do in this thesis. The understanding of human reasoning as a computation is nowadays commonly acknowledged, but its roots date back to the XVIIth century (again), in the Hobbesian thinking. In his *De Corpore* (1665, 1.2), we can in fact read the following assertion: "By reasoning, I understand computation."



This idea will have significant connections to later views, both to some of Leibniz's and to more recent approaches that adopt a computational theory of mind. Leibniz will state in his writing that "Thomas Hobbes, everywhere a profound examiner of principles, rightly stated that everything done by our mind is a computation, by which is to be understood either the addition of a sum or the subtraction of a difference" (Leibniz 1666, 3).

He was the first to submit philosophy to a rigorous treatment of logic, thus imposing on it the same criteria that characterize mathematical thinking. The consequences of this apparently simple equivalence are considerable, and make Hobbes thinking incredibly modern and near to us.

This projection of the *mathesis universalis* on language gives an initial start to the logic treatment of language, that Boole will revive later in his treaty *An investigation of the laws of thoughts* (1854). The XIXth century, will namely accomplish the turn to the *algebrisation* of logic, by substituting to propositions their truth-value, namely 0 and 1. In this way, Logic on propositions was transformed into a calculation on values (D. Andler, 1989)[73].

The material substrate to perform these calculations arrived later in the XXth century

73. See among others, Stewart Duncan, 2016"Hobbes on language: propositions, truth, and absurdity". In Martinich and Hoekstra (eds.), *Oxford Handbook of Hobbes*. Oxford: Oxford University Press.

thanks to Turing and Shannon[74], who permitted the link between electric circuit and boolean calculation. With the advent of the computer, being at this point able to materialize propositional logic into electric circuits, the last step towards Cybernetic and Artificial Intelligence was accomplished by bringing on the stage the actual biological substrate of human mental activity – the brain. The parallel study of the mind, the machine and the brain, was inaugurated and later influenced Noam Chomsky's reflection on human grammar (see Chomsky, 1963).

In this rather abrupt transition between the XVIIth century to the 50's, we can understand the extent to which philosophical constructions of Baroque times[75] lie at the core of the computational revolution that took place after the second World War[76].

1.3.2 From computation to sentence-unit

One of the most important elements we are going to synthetically present here is how Hobbes's account of language is critical for his approach of the mind. His philosophical approach to language and mind leads us to consider the *sentence as a cardinal cognitive object*.

As a matter of fact, the ability to speak has been recurrently taken by this philosopher as the *conditio sine qua non* of cognition, and that is why he begins with questions about language and the mechanisms of the mind in three of his masterpieces: *Elements of Law*, *Elements of Philosophy* and the *Leviathan*. In the last of these works, the 3rd chapter is entirely dedicated to language, and as we can read in the following, the equivalence between mental discourse – '*the train of thoughts*'- and the '*train of words*' is established in speech:

> "The general use of speech is to **transfer our mental discourse, into verbal; or the train of our thoughts into a train of words**; and that for two commodities; whereof one is the registering of the consequences of our thoughts [...]. Another is when, many, **use words to signify, by their connection and order one to another, what they conceive or think of each matter**; and also what they desire, fear, or have any other passion for. And for this use they are called *signs*.
>
> Thomas Hobbes, *Leviathan*, 1.4, 1651

Here, thinking (i.e. *ratio*) and discourse (i.e. *oratio*) are joining to the point that Hobbes will state the equation *ratio=oratio*, whose meaning was to be understood in the following way: reason is totally defined in relation to language, and it consists of the faculty to make syllogisms, and therefore to the faculty to operate the calculations they imply. "By their connection and order one to another", words express what man "conceive or think of each matter".

Interestingly, Hobbes goes a step further in the affirmation that discourse (*oratio*) lies at the very heart of predication (i.e. linguistic signification). In *De Corpore* (1655, 2.3),

74. Cf. the formulation of an abstract "universal machine"by Turing (1936) and Shannon's definition of Information Theory in Shannon (1948).

75. As we already saw when analyzing Descartes' programmatic prophecy about the neuro-cognitive investigation of language (cf. 1.1.1), we definitely are in debt with the intellectual heritage of Baroque times.

76. Although Hobbes first defined the various mental processes (compounding ideas, forming propositions, reasoning syllogistically) as being performed by some underlying elementary calculations, like additions and subtractions, it is nonetheless far-fledged to say that Hobbes really initiated Artificial Intelligence.

he states that words alone are not signs: "they are not signs except insofar as *they are arranged* in speech and are its parts". A crucial distinction is made here between words and utterance, while words name things, it is only utterances that have signification, and predication is done at the level of the utterance. One could probably glose his affirmation by saying that language signification emerges only when words become constituents of the sentence, "its parts".

As previously noted in Descartes assertions, Hobbes continues in this direction saying that "Words so connected as that, they become signs of our *thoughts*, are called speech, [...] And words so and so connected, signify the cognitations and motions of our mind" (from *Elements of Philosophy*, Concerning Body, II.ii.2.3, 1656). In this last statement surfaces again the idea that sentences manifest "the cognitations and motions of our mind". In the *Leviathan* (1651) we could already have read that compounding "*signs*" means forming propositions.

How are Propositions formed?

Among all the different shapes that discourse (*oratio*) can assume, Hobbes states that propositions are the level at which the connection between subject and predicate take place by means of a connection or linking sign – *signum connexionis* –, that expresses the power of synthesis and connection that is distinctive of human intellect. Interestingly, Hobbes's use of *continens* to speak about the predicate and *contentum* for the subject[77], is reversed compared to Leibniz's use. Hence, the proposition is defined as being generated by adding the predicate word to the subject word. For example, by adding 'snow' to 'white' we get 'snow +*signum connexionis*+ white'. The temptation to add the copula 'is' would be great, if Hobbes himself had not argued that it was unnecessary, for we could indicate the same thing by word-order rather than having an extra word as the copula. Interestingly, on this point, the difference between Aristotle and Hobbes resides mainly in the fact that the power of synthesis is no more linked to the Being, and all its ontological prerogatives, but the Being is emptied from its ontological content and reduced to a copula, its simple function allowing just to connect a predicate to a subject.

Hobbes states, indeed, that what we need to perform our mental operations of thinking is not the verb "to be", but rather a sign that would allow the processing of the operation of linking together subject and predicate – *contentum* and *continens*. It is difficult not to recognize here an embryo of the description of what will later be Merge operation in the generative program: "a (dyadic) operation that takes two syntactic objects, call them X and Y, and constructs from them a single new syntactic object, call it Z" (Chomsky, 2002).

This introduces us to a second pillar of the dominant view in modern Cognitive Science, where the symbol-processing metaphor has been so pervasive that one could state that thinking means "manipulating" symbols[78]. These symbols are defined as

77. This use is also found in Geulincx's theory of predication. Arnold Geulincx is a Cartesian philosopher that has often been compared to Malebranche and Spinoza, while he developed an original and independent theory of predication and logic that is unfortunately less known that those of Hobbes or Leibniz.

78. Note that in Cognitive Science another trend actually favors an anti-symbolic view – the Parallel Distributed Processing Group (see Rumelhart and McClelland, 1986 ; the PDP conference, 2006). At the end of the 80's, intense debates reexamined in substantial ways the central idea of Cognitive Science that intelligence is the result of the manipulation of structured symbolic expressions, proposing, within connectionist's view, that intelligence had to be understood as the result of a transmission of activation

mental representations, which, in turn, become mental objects on which to perform operations (see Pylyshyn, 1984).

Crucially, regardless of the cognitive domain being investigated, the formal way of manipulating this mental representations is the *rule system that the brain uses* to constitute the syntax of a given cognitive domain. It has, therefore, become widely assumed that cognition is in part computational, and, given this computational basis of cognition, it is not surprising that syntax acquired such a central position in Cognitive Science. Syntax has been understood not only as the unique formal property of linguistic computation, but as a property of cognition (Townsend and Bever, 2001)[79].

Importantly, this third and last shift leads us from a syntax of languages, intended as theory of grammar, to a syntax of cognition, intended as a theory of the computation of the mind. We can therefore schematically represent the shifts we have been retracing in this section in the following manner:

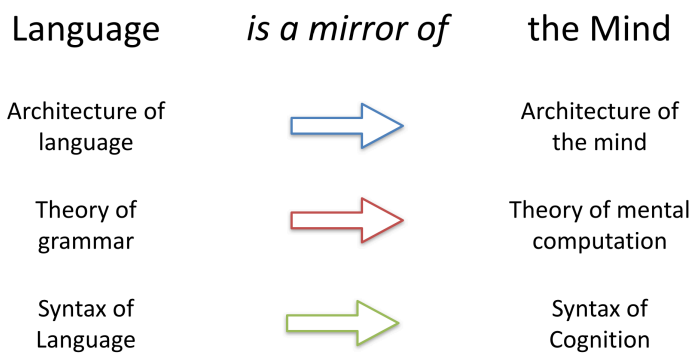


Figure 1.5 – A schematic summary of the different conceptual shifts which are implied in saying the "*Language is a mirror of the Mind*" as reviewed in this section.

1.3.3 A deep structure in languages: Grammar

Having introduced these series of parallels between Language and the Mind, we reach the point where a fundamental question naturally raises: how can we define Grammar in this framework?

The initial intuition that there should exist a hidden structure or form in language is antique, and as briefly mentioned in the previous section §1.2.2 (p. 24), von Humboldt's linguistic studies on many diverse languages[80] brought him to develop the idea that there should exist what he calls an *innere Sprachform*. This being said, a widespread conviction is assigning the first formulation of an internal generative form of grammar to this famous expression. Yet, we should be cautious in taking for granted the direct

levels in very large networks of interconnected basic units. In these debates, Smolensky (1988) opted for a middle way. For an introduction to those debates, see the special issue of *Cognition* co-edited by Steven Pinker and Jacques Mehler (1988:1-2) on "Connections and Symbols."

79. In Townsend and Bever, Sentence Comprehension: The Integration of Habits and Rules Language Speech and Communication, 2001

80. Including Basque, the native languages of Americas, Egyptian, Sanskrit, Chinese, Japanese, and Austronesian like Kawi.

association between the generative program of Noam Chomsky, and the ideas formulated by von Humboldt on Language. Von Humboldt, in fact, is pointing in this expression at the world view – *Weltansicht* – that a given language informs and models in the thinking of its speakers, and not to any types of abstract linguistic grammatical structure. This concept is closely tied to the notion of the spirit of a people (*Volksgeist*) in that it is tightly linked to the idealization of the imagery that occurs in the usage of a language.

It is nonetheless true that we can understand in this formulation a primo manifestation of the concept that a surface language coupled by an underlying level exists, the latter giving a form to the first one. This underlying level concerns abstract ideas, thoughts and concepts for Humboldt, and it has been interpreted by Chomsky with the term "deep structure".[81]

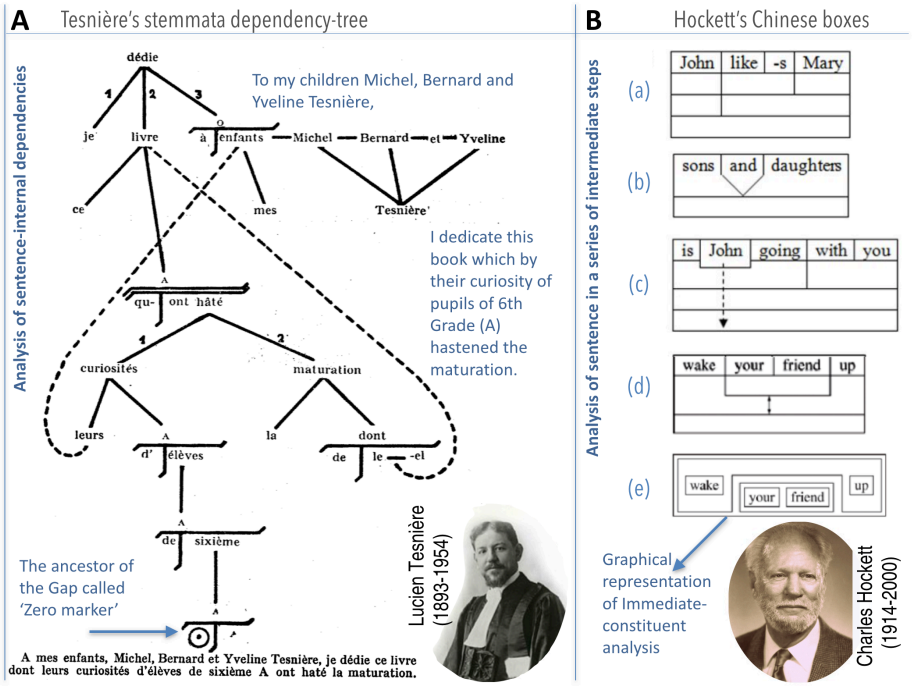


Figure 1.6 – (A) An example of *stemma* by Tesnière (1953). (B) three examples sentences showing the series of intermediate steps that are called *Immediate Constituents* (IC).

Later, another linguist, Lucien Tesnière (1893-1954), reached a somehow similar conclusion from the comparative study of many languages[82]. He deduced that there must be a hidden structure in language that could explain the visible structure of utterances.

81. It is also true that in Humbolt's philosophy of language, a kind of generative view of human language and speech was nonetheless advanced. Namely, he understood linguistic form as a procedural rule having a direction, using his own terms in German: a *Form von Form* (GS Vol 5: 455), rather than as some kind of material shape or fixed objective entity. He contrasted the *forma formans* human language with a *Form von Materie*, in that he believed that the structure and organization of a language could not be gathered from the actual verbal form of its constructions and grammar. It was rather to be obtained from an analysis of the procedures language employs in its speech generation – *Verfahrensweise der Sprache bei der Erzeugung der Rede.*

82. He was a specialist of Slovene and German.

As a matter of fact, he is the first to introduce, in 1933, a sentence analysis consisting of a hierarchical tree-like representation (see Figure 1.6A). The name he coined for this kind of sentence representation was *stemma* (plural *stemmata*). This time the hidden structure was grammatical in nature; in fact such trees (coming from genealogical trees) were graphically representing various sorts of structural dependency relationships, like hierarchical dependency, coordination or anaphoric link (Tesnière, 1953 and 1959). Interestingly, his focus on grammar was paralleled by an interest in mental operations and pedagogic reflection.

Another figure with a great pedagogical talent contributed to an early formalization of syntactic constituents and their derivation. Charles Hockett, from the American linguistic structuralist school, built a method for analyzing sentences of all languages in a series of intermediate steps, called *Immediate Constituents* (IC). In fact, as he needed a tool to represent sentence structure, he proposed a formal graphical representation using embedded boxes or charts displayed in Figure 1.6 (B), the so-called Hockett's Boxes or Chinese Boxes.

Hockett's notation represents the major cuts in the sentence by means of the longest dividing lines in Figure 1.6 (B) allowing the depiction of discontinuous constituents as in (a). For instance, a phrasal verb like '*wake up*' like in '*wake your friend up*' in (d) could be graphically represented as a constituent even if presenting a linearly long-distance relation in the sentence. Figure 1.6 (Bd) represents this discontinuous constituency relation thanks to Hockett's boxes. The idea was to represent the internal organization of a sentence structure as an arrangement of the kind of Chinese boxes without writing anything on them, in order to put in evidence the constituents having the same structure.

What we can note in these two early graphical attempts to help the analysis of syntactic structure is that they contributed to put forward:

(1) Rules and processes that determine basic sentence structure, which is central to the approach Immediate Constituents (IC) theory; and
(2) the Dependency relations within a sentence, like is shown in Tesnière's dependency tree.

Interestingly, these two aspects of sentence structure happen to be the two central focus of the experimental approach adopted in Part II of this manuscript. We will concentrate on the first one by investigating the cerebral underpinnings of Topic-Comment sentences in Mandarin Chinese (chapter 3, and on the second by investigating the cerebral representation and processing of displaced constituents in French (chapter 6) and of resumptive pronouns in Chinese (chapter 7).

Hence, Tesnière's dependency trees, Hockett's Box constituents analysis, and later the X-bar trees, can be understood as representing actual prototypes for Generative trees[83], in that they already show the implementation of a type of linguistic model where one would distinguish between a '*surface structure*', the actual uttered sentences, and a '*deep structure*' underlying the first. This deep structure will be constituting a *Grammar*: a set of simpler kernel structures and rules, from which one can generate the surface structure, as we will see in Figure 2.3.

This basic configuration evolved in the 50's, when the development of the Generative Project would differentiate two kind of mappings in grammar: depth grammar would map on surface grammar, and then the surface grammar would itself be mapped onto the phonological reading, reconstituting the final product of a sentence as it is actually

83. May they be binary branching or not.

uttered[84]. We will go back to these consideration in a few pages when presenting the purely computational aspect of Grammar, but let us before clearly set the place of linguistic Grammar as it was just defined (within the current framework of Cognitive Science) and its understanding of the structure of the Mind.

1.3.4 Language and Grammar as *modules* of the mind

> *One of the hottest questions in the study of natural language comprehension is clearly to know to what extent it is modular and what are the modules.*
>
> David Marr *Vision* 1982:356

From language faculty to Language as a *module* of the mind: some implications

The influence of the embryonic views on the computational aspect of the mind expressed by our now familiar Rationalist thinkers, stretched across centuries to aliment Jerry Fodor and Noam Chomsky's investigation of the functional architecture of the mind.

The central idea of the modern computational theory of the mind is that everything done by our mind is a computation and the mind operates as an information processing device (for example, a computer)[85]. Information theory studies became cardinal in the shaping of ideas about the mind, and this is how the notion of the modularity of mind advanced by Jerry Fodor had a very strong impact, particularly in the domain of psychological studies dedicated to language.

A system is said to be modular whenever its functioning is ensured by a set of subsystems that are functionally distinct. According to information theory, such a system (i.e. a *module*) is called 'encapsulated' when it can only access to the type of data/stimuli that belong to its domain of application (i.e. its database) and when it is blind to that of other processing systems. Some clarification on what is a computational module can be found in a foundational text by another proponent of modularity, David Marr:

> "Computer scientists call the separate pieces of a process its *modules*, and the idea that a large computation can be split up and implemented as a collection of parts that are as nearly independent of oneanother as that overall task allows, is so important that I was moved to elevate it to a principle, the principle of modular design. This principle is important because if a process is not designed in this way, a small change in one place has consequences in many other places. As a result, the process as a whole is extremely difficult to debug or to improve, both by a human designer or in the course of natural evolution, because a small change to improve one part has to be accompanied by many simultaneous, compensatory changes elsewhere. *The principle of modular design does*

84. For practical reasons, I restrain myself to the contribution of Linguists to this issue of logic versus grammatical form. Although it could be interesting to address it in the framework of the modern/contemporary logic of Bertrand Russel, Quine, Frege or even Donald Davidson.

85. See Pylyshyn (1984).

> *not forbid weak interactions between different modules in a task, but it does insist that the overall organization must, to a first approximation, be modular."*
>
> David Marr, 1982, p.102

Hence, modularity as a general principle of organization of cognitive systems has numerous implications. Mental activity in this framework is the result of the coordinated operations of different specialized cognitive processes or modules, where every module is responsible for the processing of a certain type of information. For instance, in the realm of language cognitive functions, it consists in postulating the existence of a system specialized for the processing of linguistic information in the mind.

Fodor's book "Modularity of mind" (1983) has been foundational in this respect. Favoring this kind of modular understanding of psychological faculties in Cognitive Science, it put forward a certain number of properties (i.e. informational encapsulation, domain specificity, innate, rapid, wired, autonomous and non-assembled) that should shape a modular system in order to be considered as such.[86]

Given the pervasiveness of the mind/computer metaphor and the dependence of cognition on computation, it follows that the theoretical foundations of Computer Science, like modularity, were not only placing fundamental limitations on the actions of universal computing devices, as expressed by Marr's quotation, but became of equal significance to any theory viewing cognition from an information processing perspective. This was in fact the case within the framework of a linguistic theory like the generative project which posits complex abstract machinery to capture the intricacies of natural language phenomena. Namely, if one takes Chomsky's (1957) transformational grammar (TG), language is indeed modeled as a multi-strata object in which complex transformations happen between a deep-structure (unseen) representation and a surface form.

In conclusion, ever since Fodor's milestone book, the mental structure has been conceived as a series of psychological mechanisms organized according to a functional architecture of *modules* corresponding to different cognitive systems or shared across them. In this way, a start was also given towards a modular interpretation of language faculty itself, which was further confirmed by psycho-linguistic's experimental results[87] and neuro-imaging results, like the two studies by Fedorenko et al. (2012) reported in the figures here-under, respectively Figure 1.12 and 1.7 p. 46.

For instance, in the study reported in 1.7, brain activity has been recorded during different cognitive tasks listed in (B), among which sentence reading (in black) and non-words lists reading (in gray) constituted critical linguistic tasks. Note that several cognitive tasks encompassing different cognitive systems (e.g. Working Memory, Cognitive control, etc.) with two levels of difficulty (hard and easy) were selected to test for the difference between the modularity of the Language system in contrast with a more general cognitive system that Fedorenko calls the Multi-demand (MD) cognitive system, following Duncan (2010/2012)[88]. The fMRI results reported in (C) show that in both anterior and posterior areas of the left hemisphere, and in two areas of the right hemisphere

86. Although Fodor does not offer a strict definition of modularity, he states that modularity implies graduality, in that a cognitive system can be modular to a certain extent (cf. Fodor 1983:55).

87. See further implications of the modularity of the mind in aphasiology studies. For another point of view on modularity at the level of sentence processing, see Marslen-Wilson (1979)

88. From the results in Figure 1.7 (C), we can note that the cognitive load implied by the more complex level of the Multi-Demand tasks shows activations that go quite systemically in the opposite direction in nearly all ROIs except for Verbal Working Memory hard condition in the Precentral region labeled by the author as Left Middle Frontal Gyrus (LMFG).

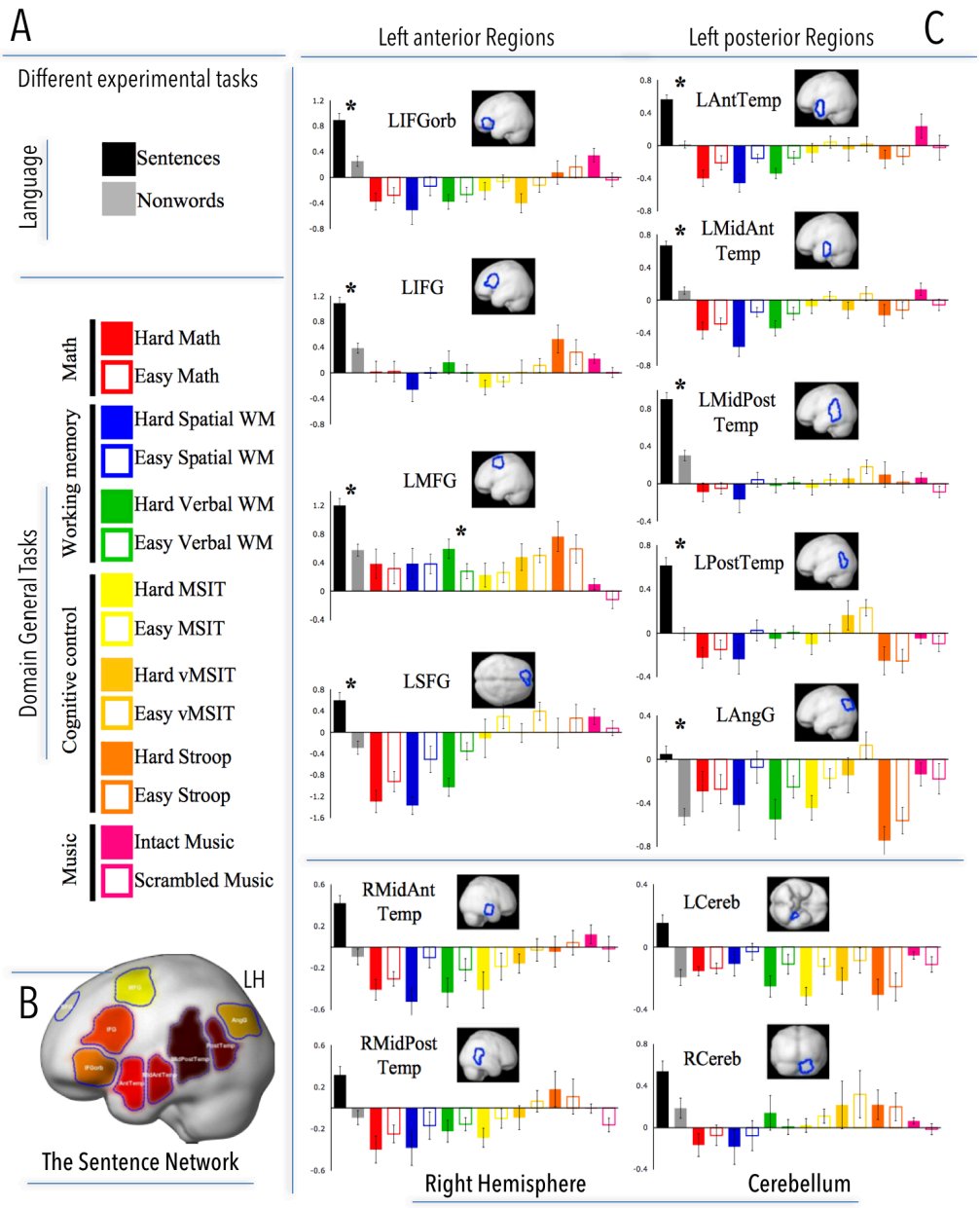


Figure 1.7 – Language modular implementation in the brain. (A) list of the different cognitive tasks included in Fedorenko et al. (2012), note in black the fMRI activation to the sentence reading task. (C) shows the functional activation profiles of Language-Selective and Domain-General Functional Regions of interest, namely activations to different cognitive task in several regions of the language network where sentence reading activation is significantly different from that to the different Multi-demand (MD) cognitive. (B) illustrates a schematic representation of the "Sentence network" in the left hemisphere, encompassing a number of brain regions revealed by the contrast sentences minus Non-words. Some of them showed in this study significant stronger activity to sentence than to other cognitive tasks as detailed in (C). Adapted from Fedorenko et al., 2012.
MSIT = Multi Source Interference Task; WM = Working Memory; Stroop = a psychological test demonstrating interference of processes in reaction times (e.g. naming a color when the name of the color is written in another color.

there is a significant stronger activation for sentences compared to other cognitive tasks. Interestingly, we can see that some Cerebellar regions are also involved in what we are going to call for convenience the ***Sentence network***, in that it encompasses a series of brain regions that have a role to play in the cerebral representation and processing of the sentence linguistic unit.

1.3.5 Linguistic Grammar as a *computation*

> *"The competence/performance distinction is meant to emphasize that a grammar is about the representation and computation of language, not about sentences per se – not directly about the utterances and behaviors of speakers. But one gains the impression from much linguistic writing that grammars in fact are descriptions of data rather than hypotheses about computation and representation."*
>
> Alec Marantz 2005:437

The framework we have sketched until now has been the platform from which the Generative Program took its run-up, to accomplish a groundbreaking bound towards language faculty, by linking *language grammar* and *computation*. However, to fully understand this view on grammar and on computation, it is probably necessary to picture the scientific culture that led to it.

The 60's and 70's were the theater of the cybernetic revolution. American researchers at that time were convinced to be on the verge to solve the linguistic problem by being able to automatize translations from Russian to Anglo-American[89] (cf. historical configuration). In this period, the now familiar metaphor of the machine in Descartes' *Discourse on Method* ceased to be a metaphor, and became a scientific research project. Information theory was now the discipline addressing the problem of the complexity of human, animal and machine communication. In this enthusiastic atmosphere, the attempts of cybernetics to interpret the grammars of human languages using statistics (cf. Shannon theory) were first relativized, and then rapidly failed. It became clear that human minds could not be assessed like machines, and even more clear that human languages could not be represented as linear sequences of symbols regulated only by statistical rules.

In this context, the reflection of Noam Chomsky about human language grammar emerged. To fully understand how the word grammar is used here, we first have to step away from the definition given by the old grammarians and focus on the one given in the previous two sub-sections: grammar is conceived as a *module* of the mind. Chomsky defines it in early published works as a set of rules or an algorithm able to distinguish grammatical sentences from ungrammatical ones. Namely, one could apply this algorithm to a string of words to determine if it is grammatical or not. In his approach, Chomsky

89. Here, we rely on the historical contextualization offered by Andrea Moro in his book *Boundaries of Babel* he had the kindness to offer.

not only tried to show that the structure of human grammar is more complex than a statistically-based model, but he also tried – as previously noted – not to ignore the issue of the inexplicable ease children have to learn the complex system of language.

This idea of an algorithm classifying correct and incorrect utterances emerged directly from the observation that language, and the utterances in one given language, can be potentially infinite, and that people can recognize all sorts of sentences as grammatical, although no one has ever uttered them in front of them before. Taking these facts as foundation, a grammar had to be a model of language that was able to give an account of what enables people to recognize the grammaticality of new sentences they had never heard before. Thus, Chomsky elaborated a system of linguistic generation in which new sentences would be projected from a relatively simple base. In this way, one of the most fundamental properties of human language, its 'creative' aspect – underlined by all the rationalist thinkers we met – had found a formulation: a *generative* finite procedure represented in the brain could potentially produce an infinity of hierarchically structured expressions, having sound and meaning[90].

Grammar being a module of the mind, the generative undertaking considers language as an internal system, similar to what a biological organ could be. Chomsky will call this Internal language "a certain knowledge in the mind of the one knowing a language is acquired by the one that learns it and used by the speaker/listener" (*Knowledge of Language*, 1986). A similitude often used by Andrea Moro can elucidate this definition: like a falling stone does not have the law of gravity written on it, a speaking person when he utters a sentence does not also utter the rules of grammar that governs it. These rules are in fact internal, and the Generative project aimed to give them an explicit formalization, to then propose a mechanism that would organize them (Rouveret and Schlenker, 1998). This understanding of language has, among others, a natural implication for the understanding of the diversity of languages: every different language has to satisfy the basic properties of the human language module in the speaker's mind. And as already pointed out previously, the initial state of this module is uniform for all human beings, and under the influence of the environment, it undergoes changes during the first years of life, resulting in an attained state called the I-language. More precisely, every expression that is internally generated has a particular interpretation on two interfaces:

1. the first for externalization through the sensory-motor modality, and
2. the second as a semantic interface for reasoning, interpretation, inference or planning of other mental processes.

In this way, I-language is viewed as a computational system connecting form (Phonological Form, PF) and meaning (Logical Form, LF) by merging elements selected from the mental lexicon into a structure which is undergoing certain reformulations, certain syntactic transformations. Crucially, the role of these syntactic transformations is to state how the deep and surface structures of sentences can be generated or transformed formally from basic types of sentence structures[91].

In Chomsky's Minimalist Program (Chomsky, 1995/2000) and in the related approaches, the minimal combinatory units of language undergo both the semantic and

90. Note that this is usually the point in Chomsky's argumentation where von Humboldt's observation about human language faculty – finite means could generate infinite linguistics expressions – is given a grammatical computational interpretation.

91. It has to be noted that nothing similar has ever been attested in another organism than man.

phonological interpretation presented above, and combine via a recursive operation, the so-called "merge", that is creating hierarchical constituent structures. Importantly, while each atom of syntactic combination and every licensed combination of elements (be it atomic or already derived) is to be interpreted both in phonology and semantics, the recursive structure-building operations are always that of syntax. In one sentence: the syntax is the sole generative engine of grammar, it is responsible for the recursive hierarchical structure of words and sentences.

Hence, this step forward, made by the early Generative enterprise, was stating that language structures are derivable in a formal system that constitutes a theory – of both representation and computation – in this way, Chomsky (1965) was actually giving a go to a pluri-disciplinary approach to language as competence too.

As stated in the above epithet by Alec Marantz, introducing the competence vs performance distinction within linguistics had, at the beginning of the Generative Project, the corollary of transforming the work of linguists into two main tasks: (1) emitting generalizations about categories and structures of language from distributional regularities within and across utterances (i.e. alimenting a theory of representation), and (2) accounting for what are doing individuals when understanding a sentence. Thus, this theory about the generating mechanisms actually proposed to create linguistic representations, in order to account for the computation mechanisms of sentence[92].

Under these assumptions, we clearly see that this research project transformed Linguistics into an empirical science proposing an explicative hypothesis of the linguistic competence (i.e. the internal linguistic knowledge) of the speaker. To this effect, theory or modeling are not derivable from data alone, experiments are necessary to obtain empirical results. Thus, collecting psycho-linguistic empirical data, like reaction times in priming designs, means focusing on the second task of the linguist, that of describing computational mechanism of the sentence, like we will do in chapter 5.

It should be noted that it is because of these points, and precisely because Generative linguistic represents both a theory of language knowledge and of performance, that it is in use within Cognitive Neuro-science. Moreover, as the categories and operations developed in the framework of Generative Grammar are also hypotheses about the representations and computations in the mind and brain of speakers, that I adopted this framework myself to frame part of the experimental hypotheses in the fMRI studies that will be presented in chapter 6 and 7.

92. An example can be found in the early development of the Learnability theory by Chomsky and Miller in the 1970's. This focus on learnability of language later consisted in modeling real-life language learning incorporating more realistic assumptions about languages and language learners. One of the first study in this direction that had great impact was by Wexler and Culicover (1980). The double scope of their research was showing that a linguistically authentic grammar could be acquired by computational operations of contained complexity, and that this was realizable on the basis of a language sample approaching the language exposition of toddlers. Hence, they started developing a learning model for natural language which could be considered not only as computationally sound, but also psychologically sustainable and nonetheless in tune with current linguistic theory, which at the time was Chomsky's Extended Standard Theory. However, in 1981 Chomsky introduced a new theory with particular assumptions about learnability. Government-Binding was namely defined as having universal principles and a finite collection of parameters codifying cross-language syntactic differences. Principles were fixed and Parameters were seen as learnable. Although Chomsky was not explicit about the mechanism for Parameters setting, he might have regarded it as a relatively simple if not automatic process (that would fit how children acquire a rich target language in just a few years) which was nonetheless centered around the notion of external linguistic input guidance (see Fodor and Sakas (2017), in *Oxford Handbook on Universal Grammar*). For investigation on Parameters setting and its theoretical approach through hierarchies, see 1.2.4, page 35 and Roberts and Holmberg, 2010 or Biberauer et al. 2014.

The distinction between mental representation and mental processing[93] is here introduced for further development. In this research work, we will indeed consider the hypothesis that the representation of syntactic structures is a distinct object from its processing (i.e. from the cerebral activity elicited by its processing). This will constitute a leitmotiv of this work. In fact, distinguishing the manipulation of a given representation from the representation itself appears nowadays to be cardinal in bio-linguistics.

We can say that this evolution paved the way to the advent of bio-linguistics. We will address what change of perspective is implied when considering the brain in the analysis of linguistic sentence structures in the next section.

1.4 Language and brain: bio-linguistics

> "Nicht das Gehirn denkt, sondern wir denken das Gehirn."
>
> *["It is not the brain that thinks, but we think the brain."]*
>
> FRIEDRICH NIETZSCHE

We will now turn to the question of what makes a human brain a 'talking brain'. By reviewing different observations, intellectual reflections, experiments, methods and techniques that concretely enabled us to perform this kind of research project, we will witness the birth of bio-linguistics. In other words, we will observe what happens when the brain is inserted in linguistic analysis. But, before reviewing the steps that thrust the biological investigation of syntax forward, we will stop by to leverage the contribution of two different critical periods preceding the golden age of Bio-sciences and of bio-linguistics we are in.

1.4.1 The unquestionable Question: the taboo of the origins of language

> "ART. 2. – La Société n'admet aucune communication concernant, soit l'origine du langage soit la création d'une langue universelle."
>
> *["ART. 2. – The Society does not accept any communication concerning neither the origins of language nor the creation of a universal language."]*
>
> STATUTE OF *La Société Linguistique de Paris* 1886.

As we can read in the epigraph above, the turn of the 20$^{\text{th}}$ century was the theater of a very astonishing linguistic resilience towards the investigation of the origins of

93. As we saw in Sub-Section 1.3.1, the philosopher Thomas Hobbes had formulated an intuition going in this direction: thinking amounted to performing arithmetic-like operations on internal structures (i.e. mental representations).

language. In our contemporary understanding,this ban has often been interpreted as a conservationist fear or reaction to Darwinian thinking, or even sometimes as the mark of a deeply anchored rejection among linguists of the question of language biological evidence.

Sylvain Auroux (2007: 41-68) has a sightly more sophisticated linguistic analysis of this debated Article 2. He notes that it is not the impossibility to observe the roots that was bothering the linguists of that period. In fact, comparatism was at that time the leading linguistic approach, and it aimed at reconstructing proto-languages. The problem resided in the fact that in the research of the ultimate origin – which concretely meant the search for common linguistic elements to all the families – one would have had to search for the causes and motivations of these roots, for example some onomatopoeia. And this last point would have put into question the totality of the comparativist program and its phonetic laws, which were based on the axiom of the *arbitrariness of its roots* (in that they cannot be considered as iconically motivaated).

Hence, the members of The Parisian Linguistic Society did not think that the question of the origins of languages was not included in the domain of their scientific investigation of languages, but ratherthat this kind linguistic investigation on the origins would have brought to an epistemological impasse, ending up in questioning the core of linguistic comparatism as a discipline: "the axioms of comparative linguistics are incompatible with every research on the origin of languages[94]. We can see, thanks to this explanation, how the idea that languages ultimately consisted in a *convention* was deeply anchored in the XIXth century linguists. It is precisely on this conviction that the bio-linguistic revolution of the 50's will take place.

94. "The point was not to prevent other disciplines to search for the origin of language faculty" (Auroux, 2007:126).

1.4.2 "Bio-everything" science and bio-linguistics

> *Social sciences were envious of natural sciences, their exact methods, stringent results and continuous progress, which have led to think, in the last hundred years, that everything that is in man, in its society and history that was belonging to the physical or biological nature of man was accessible to exact and quantifiable methods. And what was instead out of reach was belonging to the essence of man: self-awareness, moral awareness, values, responsibility, freedom, meaning. Researchers try then either to leave these specifically human phenomena out of their investigations, or to reinterpret and translate them in order to make them reducible to observable and quantifiable matters of facts.*
>
> Jeanne Hersch *L'étonnement philosophique. Chap. 'La Philosophie aujourd'hui'* p.455

On a totally different register, this quote by a contemporary philosopher is very pertinent in that it points to a post-modern tendency to transform every branch of Humanities and Social Sciences into a biologically or cerebrally reinterpreted social science, a "bio-" or even "neuro-" discipline. Linguistics benefited of this vogue of "organic" bio-fashion in science, and we could also replace the neuro-imaging turnover in research (and in popular culture) in this "bio-" trend. As a matter of facts, ever since the end of the nineties even the last 'unquantifiable aspects of man' in the above list have been made to be quantifiable – or at least '*image-able*' – in the brain. One could in fact approximately translate them into some of the research axes that are nowadays found in neuro-imaging: consciousness (self-awareness), reward loop (moral awareness), social cognition (values), decision making (responsibility, freedom), semantics (meaning)[95].

As concrete a demonstration of this "bio-" trend, in the 50's/60's the modern study

95. One could honestly say that the only philosophical interrogation that has not been dissolved in the biology of the brain yet, is probably the quest for the meaning of life, this archetypically philosophical interrogation that every man experiences as the problem of life or human condition. This won't last long, as some happiness studies have lately emerged. In the meantime, Philosophy continue to take as an object "not the man that is a fragment of nature and therefore the object of biology, but the man that is subject and object of history, who contributes to the shaping of society, who has as self-awareness, knowing that he is a mortal, and thinks about his own destiny and experiences with astonishment" (Jeanne Hersch in *L'Étonnement philoophique*).

of language began to take shape within the context of biology. A mile stone for this understanding of language was Eric Lennenberg's book *The Biological Foundations of Language* (1967), literally paving the way to the new discipline we explore in this Dissertation. Interestingly, at the time he wrote his book, Lennenberg had to face the 'unquestionable question' of the XIXth century that we brought up in the previous section: the no-arbitrariness of linguistic conventions. We can read in his preface: "A biological investigation into language must seem paradoxical as it is so widely assumed that languages consist of arbitrary, cultural conventions". To which he adds some clear claim about the difference between the arbitrariness of linguistic convention and the biological motivation of linguistic grammar: "The rules of natural languages do bear some superficial resemblance to the rules of a game, but I hope to make it obvious in the following chapters that there are major and fundamental differences between rules of languages and rules of games. The former are biologically determined; the latter are arbitrary."(E. Lennenberg 1967:2)[96].

His perspective on language remained an inspiring framework until the advent of neuro-imaging techniques. But it should be noted that, before neuro-imaging, a critical period of medical observations and discoveries on the link between linguistic faculty and the brain was instrumental to initiate the thinking, one could say with the words of Lennenberg, that the rules of language were to be biologically determined. The first step towards the hypothesis that the brain was not working as a uniform and homogeneous mass dates back to when a French doctor, Ernest Auburtin (1825-1893), first formulated the possibility that a cognitive function like language could be pointed in a location in the human brain. We will now focus on these foundational observations.

1.4.3 Some famous aphasics

A great leap forward was made in the understanding of the relationship between language and the brain when the 12th of April 1860 a doctor met with Monsieur Leborgne in a Parisian hospital. This patient, despite the meaning of his patronymic, was not short-sighted, but aphasic.

In an epoch where the holistic view of Flourens (1794-1867) – arguing that every brain

96. Full quote: "Wittgenstein and his followers speak of the word game, thus likening languages to the arbitrary set of rules encountered in parlor games and sports. It is acceptable usage to speak of the psychology of bridge or poker, but a treatise on the biological foundation of contract bridge would not seem to be an interesting topic. The rules of natural languages do bear some superficial resemblance to the rules of a game, but I hope to make it obvious in the following chapters that there are major and fundamental differences between rules of languages and rules of games. The former are biologically determined; the latter are arbitrary." (E. Lennenberg 1967:2).

area was participating in an equal manner to mental faculties – was facing the localist and phrenologist view from Joseph Franz Le Gall (1758-1828)[97], Marc Dax (1771-1837), Jean-Baptiste Bouillaud (1796-1881), Pierre-Paul Broca (1824-1880), Gustave Dax (1815-1893) and Pierre Marie (1853-1940), they jointly contributed to bring on the stage the case studies of famous aphasics. These facts and observations definitely challenged the holistic view, making the tenets of the localization by circonvolution definitely triumph.

A week after Monsieur Leborgne's death (the 17th of April) of the same year, Paul Broca described in his autopsy a broad damage to the third circumvolution of the cerebral cortex of the inferior frontal gyrus of the left hemisphere – what will be later called Broca's area (see Figure 1.8).

Still showing a normal level of comprehension and intelligence, and without having any paralysis of the tongue and face, his main symptoms where of linguistic nature: he could only answer questions with the syllables "tan-tan" followed by a renowned french blasphemous curse '*Sacré nom de D...*', as soon he was disappointed of not being understood.

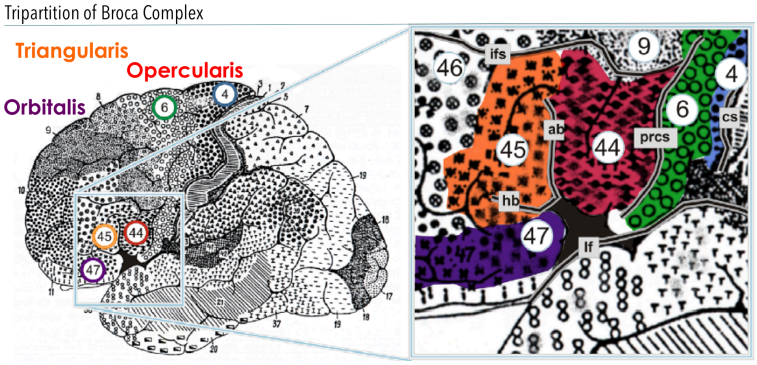


Figure 1.8 – Tripartite parcellation of Broca Complex according to the neuro-cyto-architectonical differentiation between agranular structure in BA6 and dysgranular cortical structure characterizing BA44/45. Cyto-architectonic areas are marked by different hatch marks and are classified according to Brodmann's nomenclature by Arabic numerals in circles. Adapted from Amounts et al. 2010, originally from Brodmann.
BA = Brodman area .

This patient became known to history as "Tan-Tan", and the same year a second patient, monsieur Lelong, who could speak only five words with no grammatical structure, permitted Broca to give a more accurate localization of the region of language production in the posterior third of the second or third circumvolution (Broca, 1861a) (see Figure 1.8 for an actual depiction of this circumvolution). This pathology will be called 'aphasia' accordingly to the name another doctor, Dax, coined, contrary to Broca's nomenclature that was 'aphenia'.

Successive re-analyses of Leborgne's (Tan-tan) brain revealed already in 1906 that the lesion was mainly spreading posterior to Broca's area[98]. And an X-ray study in the 80's demonstrated the presence of lesions to deeper structures (Castaigne et al. 1980). Lately, the white matter underlying the original descriptions was proved to be damaged too (Dronkers et al. 2007). In sum, the lesion of Tan-tan was actually encompassing four cerebral lobes: frontal, temporal, parietal and insular[99]. Generally speaking, the advent of neuro-

97. At the end of the XVIIIth century he emitted the hypothesis that there where cortical localizations for dispositions of the spirit. Excelling in one domain would bring an hypertrophy to the cortical area corresponding to that ability, to the point that the scalp would be lifted up by this local cerebral development. This thesis is commonly called phrenology.

98. The historical configuration for Leborgne's lesion shift in localization is probably to be attributed to the harshness of the debates at the time and to a power struggle between Flourens of the Academy of Sciences and Brouillaud of the Academy of Medicine.

99. The pathology of Tan-Tan would be nowadays classified as "anarthria" (Lebrun 1982, in Moro 2011) and the damage associated with it also includes larger areas, such as the anterior temporal cortex (Déjerine 1914).

imaging comforted this view of a language network stretching around the sylvian fissure, and many more areas were successively brought on the stage. The neural basis of language was revealed to be much more extended than previously thought, however this did not lead to a fundamental relativisation of Broca-centrism, as we will see in the next pages.

In the 1970s, research in aphasiology confirmed that patients with Broca's aphasia had difficulty comprehending syntactically complex sentences. The kind of deficits that fall under this label are: loss of the capacity to produce fluent speech, concurrent omission of both bound or free grammatical morphemes such as prepositions, copulas, auxiliaries, or verbal suffixes (Caramazza and Zurif, 1976). Deepening this understanding became the challenge that neuro-imaging faced in the last decade of the 20th century: to design experimental settings that would allow to obtain experimental evidence of the link between basic and sophisticated properties of human language and neural processes. However, as Lennenberg (1979) had already claimed, the linking hypothesis between language behavior and its cerebral implementation was to be found in linguistic theory. Investigating the biological substrate of grammatical systems needed a theory of how the language is represented in the brain.

As the idea that language faculty and linguistic knowledge were to be biologically supported was largely shared in the scientific community, linguists started to develop theories of linguistic representation that should not only be compatible with the observed linguistic patterns across languages, but also with patterns of language development and aphasiology breakdown.

Since Broca's time, a long way has been walked to achieve projection on brain activity/functioning of a computational theory of the mind and also of linguistic structures and descriptive levels that linguists deal with. And, given the *prima facie* similarities that could be drawn between language capacity and other aspects of human thought (pointed in 1.1), the question was raised about whether a common set of computations was lying at the heart of all hierarchical structure-dependent cognitive processes. This evolution immediately put sentence and its hierarchical structures on the front row of research, and this is where our interrogation on syntactic transformation is rooted.

1.4.4 Linguistics in the brain

Following the aphasiology turn-over undertaken by XIXth century neurologists, psycholinguists of the 70's started to use modern linguistic theory as a tool for experimental approach to investigate aphascis linguistic behavior.

Empirical investigations were carried out to test aphasic's capacities in phonology, morphology, syntax and semantics. Firstly, these studies showed that the human brain was sensitive to this type of linguistic information, in that the examination of the abilities and disabilities of patients presenting the two types of aphasia (i.e. Broca and Wernicke), showed selective impairments of certain levels of linguistic analysis (Caramazza and Zurif, 1976; Caramazza and Berndt, 1983; Warrington, 1975).These findings indicated that the cerebral organization of language could aslo be modular, and that different levels of linguistic analysis could be taken as hypothetical modules of the organization of linguistic knowledge in the brain. There can even be different critical periods to acquire the linguistic knowledge linked to different linguistic domains[100].

100. Further evidence for this modularity of linguistic knowledge can be found in recent studies from

Soon after, psycho-linguists like Blumstein, Zurif and Caramazza designed experiments that were centered around the linguistic distinction following the new theoretical framework given by Generative syntax (Swinney and Zurif, 1995; Blumstein et al., 1998 and much subsequent work), providing early notions about the neural mechanisms of syntactic processing.

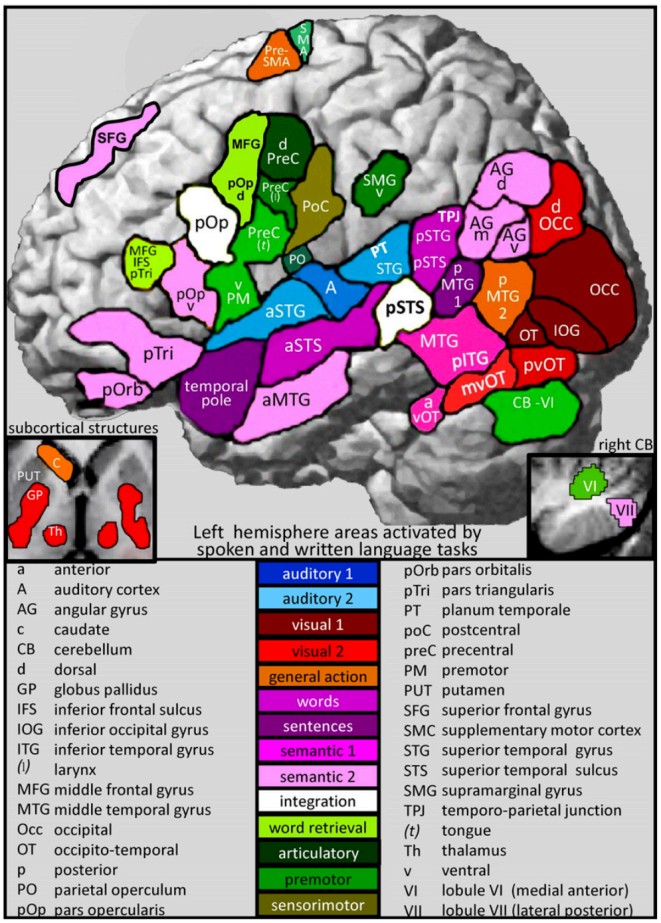


Figure 1.9 – An illustrative sketch of the location of language related activations, based on Price et al. data. The colors indicate the type of task or processing that caused the activation. From Cathy J. Price 2012 *NeuroImage* 62 p. 821.

These first pieces of evidence on the link between linguistic level of analysis and the linguistic organization in the brain, gathered by aphasiology, brought the debate on a new type of questions: what are the units of analysis fitting the relationship between brain and language? Linguistics and neuro-science got closer in order to cartography the linguistic capacities on the cortex, taking linguistic theory as a central tool for the description of linguistic behavior. This was the birth of neuro-linguistics.

More generally speaking, the anatomical organization of a large-scale neural network for language has been characterized since early 2000. Since then, it has become clear that the language network and the part sub-serving sentence comprehension involve an increasingly large number of brain regions compared to the historically defined left posterior superior/middle temporal region (Wernicke's area) and the inferior frontal region (Broca's area). As one can see in Figure 1.9, from a review study by Cathy Price (2012), since the 70's to nowaday, the sentences comprehension network has been extensively characterized in its anatomical distribution mostly in the left hemisphere.

The increasing availability of human electrophysiology and brain imaging not only allowed researchers to start assessing the issue of the syntactic representation of sentence in the brain, but has also made it possible to start questioning the brain with very fine-grained experimental hypothesis on various syntactic phenomena. The last 30 years of neuro-imaging led to the discovery of new data proving that the mechanisms of syntax in the brain are way more complex than the Broca vs. Wernicke dichotomy known in aphasiology.

Naama Friedmann and her collaborators, who focused on the modularity distribution of developmental language impairments in Israel, showing that different language domains can be selectively impaired. For a general presentation of these modular Sub-Types, see a general article by Friedmann and Novogrodsky (2008), while for syntactic deficit see Friedmann and Novogrodsky (2007 and 2006); for Lexical retrieval deficit see Kesselmann et al., 2015 and Fattel et al., 2015; and for Pragmatic deficit see Balaban et al. (2016a/b/c)

Lesion studies were the first to provide a large literature on the neural circuitry underlying sentence comprehension. A very influential one was conducted by Nina Dronkers (2004) combining both precise neuro-anatomical techniques and a very comprehensive behavior testing on different syntactic structures whose Age of Acquisition (AoA) by normally developing children varies from 2 years old to 9 years old, namely:

1. Simple Declarative *'The boy is jumping'* AoA:3
2. Simple Possesive *'The boy has a balloon'* AoA:2
3. Active Voice Word-Order *'The girl is pulling the boy'* AoA:4
4. Double Embedding *'The clown that is big has the balloon that is blue'* AoA:4
5. Agentless Passive *'The girl is being kicked'* AoA:4
6. Agentive Passive *'The girl is being kicked by the boy'* AoA:5
7. Subject Relatives Ending in N-V *'The girl who is pushing the boy is happy'* AoA:7
8. Object Clefting *'It's the boy that the girl kicks'* AoA:8
9. Object (O-S) Relative Clauses *'The girl is chasing the clown who is big'* AoA:8
10. Negative Passive *'The dog is not being outrun by the cat'* AoA:9
11. Object (O-O) Relatives with Relativized Object *'The girl is kissing the boy that the clown is hugging'* AoA:9

The analysis performed on auditory sentence comprehension deficits (in Figure 1.10 A) to define precisely lesion sites in an important group of chronic stroke patients (72 patients with both left and right hemisphere lesions belonging to four neuro-physiological behavior sub-types, including anomic patients, conduction aphasia, Broca aphasia, Wernicke aphasia) revealed that sentence comprehension deficits were reliably linked with damage to the five regions of the left hemisphere depicted in Figure 1.10 (B).

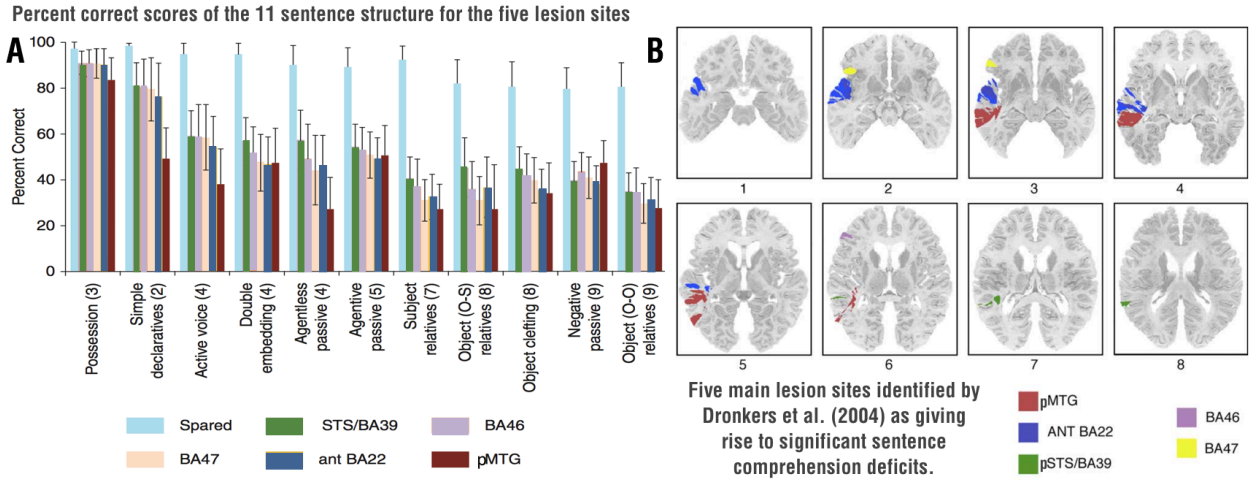


Figure 1.10 – (A) Behavioral results from Dronkers et al. (2004) neuro-psychological study of sentence comprehension. Percent correct scores (and 95% confidence intervals) on each of the 11 subtests are shown for patients with different lesion sites. Modified from Dronkers et al., 2004, p. 163.

Many interesting observations can be drawn from these rich results from aphasic literature, for example it is striking that the 12 patients whose lesions included the posterior Middle Temporal Gyrus (pMTG), only managed to pass the Possession sentence

sub-test, and that their mean scores fell below 50 percent correct for all of the other sentence types sub-tests, thus, indicating a profoundly impaired sentence processing.

More recently, the understanding of language anatomical organization has largely benefited from the increasing use of the so-called *localizer tasks*. These short fMRI paradigms allow to identify subject-specific language network through a very basic sentences versus non-words lists (or consonants strings) contrast[101]. Localizers have received much attention especially since Nancy Kanwisher and Evelina Fedorenko have brought to light their methodological relevance[102] for building probabilistic overlap-maps from subject-specific language network of the 20 people cohort that is usually selected to run fMRI studies (Fedorenko et al. 2009, Fedorenko and Thompson-Schill, 2013).

As one can see in the Figure 1.11 (A), the similarity of individual brain activations patterns to a localizer contrast involving real french sentence versus strings of consonants having the same number of characters, is at first sight quite striking. Across the 20 subjects who participated to the fMRI study that will be presented in chapter 6, the extent of activation and its spatial distribution on each individual anatomy is nonetheless quite different. Hence, the calculation of a Group-specific overlap map[103] showed in (B, red pixels) on the right side of the Figure, allows to overcome the anatomical variability across individual brains, and in a second time to generate functionally-driven cluster regions (B, colored regions with a number), where the voxels that constitute them are shared by the majority of the subjects participating to this study. For comparison, we present in (C), on the extreme right side of the Figure 1.11, the Group-specific overlap map generated through the same procedure and analysis performed on the the data of another localizer paradigm from the fMRI study on Mandarin Chinese that will be presented in chapter 7.

101. The linguist reader should be reminded that, generally speaking, a contrast is the comparison of brain activity during reading of sentences compared to a baseline "condition". It should be noted that these localizers show what is activated in the brain on average for all the sentences presented. This can be taken as representing an activation to all that these sentences have in common.
102. fMRI studies of language classically rely on group analyses, in which the brain activation data from multiple individuals are co-registered and analyzed in a common space of reference coordinates, this method allows to circumvent the problem of the anatomical variability across individual brains.
103. Using the *spm ss* toolbox for Group-constrained subject-specific (GSS) analyses developed by Alfonso Nieto-Castañon.

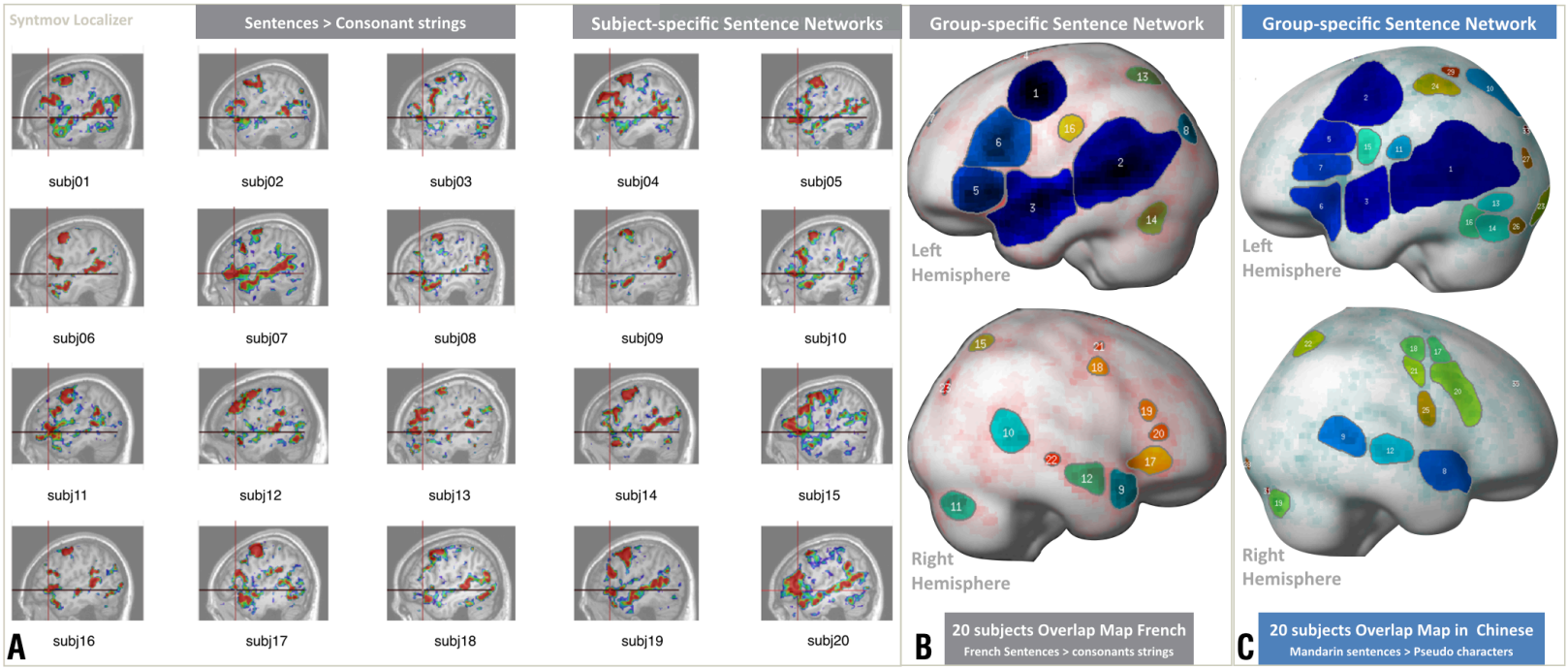


Figure 1.11 – (A) The activations patterns to the contrast of French sentences versus strings of consonants of each of the 20 subjects in Syntmov Project (chapter 6), cross-hair is in Broca's Area, and activation are projected on each subject's own anatomy; (B) Group-specific overlap map for French localizer: red pixels intensity indicates percentage of overlap, colored regions with a number represent Group-specific functional clusters; (C) Group-specific overlap map for Chinese localizer: green pixels intensity indicates percentage of overlap, colored regions with a number represent Group-specific functional clusters.

As shown in Figure 1.12, these tools to define language-sensitive regions in individual subjects, and the functional localisation approach they support has many advantages. First, it shows the robustness of the peri-sylvian organization of the sentence-processing network across individual subjects' brain anatomy. Secondly, among the most practical ones, the analysis of a main experimental task only inside these clusters has the advantage to limit one of the intrinsic problems of neuro-imaging, because its statistical sensitivity is weakened by multiple comparisons performed at each voxel. Thirdly, this method allows to create Regions of Interest (ROIs) that are not only group-specific, but are functionally determined in each individual brain (see an example of the functional segregation they offer in Figure 1.12).

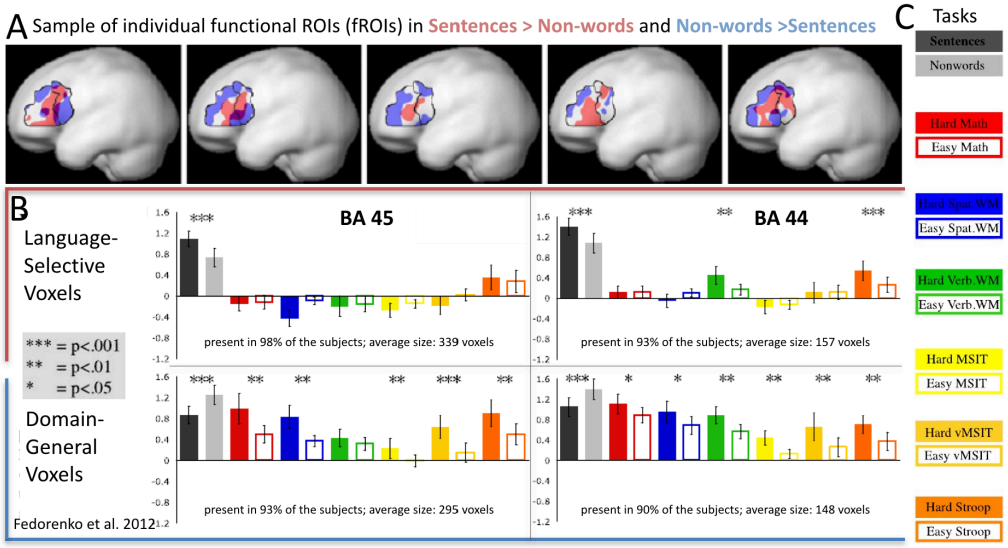


Figure 1.12 – Language is a *module* in the brain. Zoom on Broca internal segregation. (A) shows a sample of individual functional Regions of Interest (fROIs) in [Sentences > Non-words] and [Non-words >Sentences]. Henceforth, fMRI contrasts will be indicated by the symbol '>'. Like in the expression (a>b), literary meaning the brain activation to sentence (a) is superior to the brain activation to sentence (b). All the brain maps presented in this manuscript are the result of the subtraction of the brain map activation to an experimental condition to the one of the second item of a given minimal pair. We can note that even inside Broca's Area fMRI activation to the above contrast do not overlap. (B) Functional Profiles of Language-selective and Domain-General Functional ROIs this shows how inside two parts of Broca's Area (BA44 and BA45) voxels identified by [Sentences > Non-words] are robustly elicited only by sentence stimuli, while voxels identified by [Non-words >Sentences] are not selectively involved in Language, they respond as well to Domain-General tasks listed in (C). Adapted from Fedorenko et al. 2012.

1.4.5 Syntax in the brain: Broca and beyond

> *Fashion, in the feverish sense that it exists today, [...] is as if people began to dig up the foundations of a house before they had finished putting the roof on.*
>
> G. K. Chesterton. *The Illustrated London News*, 26 June 1926.

Syntactic processing and Broca's area are well-known 'cozy bedfellows' ever since Pierre-Paul Broca's time. Despite the fact that further lesion-based evidence severely weakened their relationship – Broca's aphasics are actually good at grammaticality judgments – subsequent PET and fMRI studies on sentence complexity prolonged this happy union: comprehension of complex sentences activates Broca's area more than simple sentences (Stromsward et al., 1996; Rodd et al., 2002 and Ben-Shachar et al., 2003). And a real "Broca-centric fashion" was thereby initiated. Yet, a general review of the literature will reveal that Broca's area and sentence processing is full of controversy. A variety of proposals on its functional characterization have been posited, ranging from a region supporting (1) hierarchical phrase-structure building *à la* Friederici, (2) syntactic movement *à la* Grodzinsky, (3) word-order linearization *à la* Bornkessel-Schlesewsky, (4) linguistic unification modular mechanism *à la* Hagoort, and (5) domain-general functions such as cognitive control *à la* Novick, or (6) working memory *à la* Caplan or *à la* Rogalsky-Hickok.

As proposed by a few researchers, the problem of mapping linguistic distinctions onto the brain, concerns eminently granularity (cf. the Granularity issue presented in the Epilogue p. 709 and Section §1.5.1, p. 68). Syntax, semantics, phonology are not computational tasks in themselves for the brain, they are composed of multitude of subprocessing components and representations. Hence, the quest for the area underpinning the most essential structure building process in the brain - "*the Merge quest*" - could end up into a false convergence between linguistic and neuro-cognitive investigation of language, if the linguistic processes are not decomposed in the right units. We can illustrate this by taking the example of the functional characterization of Broca's Area. It is nowadays clear that a simple and straightforward mapping between this area and Syntax is no longer viable for two main reasons. Firstly, Broca is also activated by non-syntactic stimuli like phonology and lexical tasks (Gandour et al., 2000 for tone perception in tone languages; Poeppel et al., 2004), and secondly, Broca is activated in non-linguisitcs tasks too, like motor imagery or rythmic perception.

However, further evidence in support of the importance of Broca's Areas comes from the following fact: Broca sub-parts BA 44 and 45 are reported as having undergone a considerable enlargement in 'recent' human evolution. Detailed cyto-architectonic analysis of these areas in chimpanzees and humans indicates that BA 44 and 45 are present, in comparable locations, in chimpanzees. However, these regions are among the most greatly expanded cortical areas yet identified in humans: left area 44 is 6.6 times larger in humans than in chimpanzees and left area 45 is 6.0 times larger[104] (Schenker et al., 2010).

104. To give a order of magnitude to the reader we can report the overall increase in brain size: the average human brain is only roughly 3.5 times larger than that of a typical chimpanzee, and V1 (visual

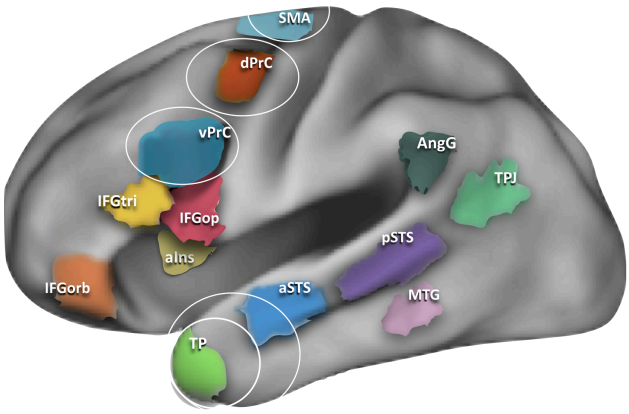


Figure 1.13 – Apriori Region of Interest for Sytnactic complexity retained from the fMRI literature on sentence comprehension. IFGorb: Inferior Frontal Gyrus, Broca's Area *pars orbitalis* (BA47); IFGtri: Inferior Frontal Gyrus, Broca's Area *pars triangularis* (BA45); IFGoper: Inferior Frontal Gyrus, Broca's Area *pars opercularis* (BA44); vPrC: ventral Pre-Central Cortex„ Broadmann Area 6 (BA6) dPrC: dorsal Pre-Central Cortex Broadmann Area 6 (BA6) ; SMA: Sensory Motor Area (/pre-SMA) ; aIns: anterior Insula ; TP: Temporal Pole (or ATL) ; aSTS: Anterior Superior Temporal Sulcus ; pSTS: Posterior Superior Temporal Sulcus ; TPJ: Temporo-Parietal Junction ; AngG: Angular Gyrus.

Nonetheless, up to today an enormous amount of very disparate imaging studies indicate that Broca complex is involved in processes vaguely labeled as 'syntactic'. Without taking into account results from aphasiology, the extreme variety of designs, syntactic manipulation, grammaticality of the stimuli, languages, tasks, modalities, in imaging studies, are generally converging on the fact that Broca complex is somehow involved in syntax and sentence comprehension[105]. This persistent activation enhanced the 'Broca-centric fashion' defined above, to the point that one starts doubting about the validity of a neuro-imaging study on sentence processing when the nice family pictures of Broca and Syntax do not show up[106]. Note that Broca-centrism was already striking right after Broca's death: Pierre Marie, one of the last Interns of Broca, probably lost the possibility to have the Chair of Charcot because he dared call into question the role of Broca's area in Language – and this is precisely how Déjerine obtained this position. Having divided the golden Broca/Language couple costed him dearly[107].

However, in the last decade, a new generation of studies weighs in this question by gradually excentering the debate on new brain regions being involved in syntactic processes or sub-processes. These newly identified brain areas will be central in our fMRI results and their discussion will stretch across different chapters. We will therefore briefly introduce three areas that have recently entered the scene of syntactic complexity and structure building (see circled brain areas in figure 1.13).

1.4.5.1 Temporal Pole (TP or ATL)

Firstly, evidence has been accumulating that the Anterior Temporal Lobe (ATL) may house a network that behaves much more like a syntactic composition system than previously thought in aphasiology, by dealing with sentence-level combinatorics. On one hand, ATL activation seems to be highly correlated with the presence or absence of syntactic information in a sentence (Dronkers, et al., 2004; Friederici, et al., 2000; Stowe et al., 1998). In fact a basic and primary paradigm for examining the cerebral bases of sentence processing has been to simply compare brain activity elicited by sentences to the one in response to un-structured lists of lexical words. These studies commonly find

cortex) is only 1.8 times larger in humans than in chimpanzees.
105. See the very detailed reviews of Embick and Poeppel (2005) and Hagoort (2016) on this point.
106. A second perverse effect of Broca-centrism is that the absence of activation of Broca's area in a syntactic manipulation will probably make it difficult to publish it.
107. For more back-stage stories of the time of Broca, see *Cerveau et Langage*. O. Etard and N. Tzourio-Mazoyer (eds.), 2003.

an effect of increased activity for sentences in the Anterior Temporal Lobe (Humphries, et al., 2005 and 2006; Mazoyer et al., 1993; Rogalsky and Hickok, 2008; Vandenberghe et al. 2002). While on the other hand, a study by Brennan et al. (2012) reported that ATL activity correlated with a measure of syntactic tree node-count providing an estimate of the amount of structure built word-by-word. This result has given the go to a research hypothesis that investigates the mapping of incremental sentence composition in this brain region (Brennan and Pylkkänen, 2016)[108].

1.4.5.2 Precentral Cortex

A second brain area emerges in the skyline of fMRI studies manipulating sentence syntactic structure: the Precentral cortex or ventral Premotor cortex (Broadmann Area 6, BA6). This area is contiguous or right above Broca's complex, it stretches both ventrally and dorsally in the frontal lobe (see Figure 1.13 ventral BA6 in deep blue and dorsal BA6 in dark orange).

Precentral activation has frequently been reported, both in its ventral and dorsal part, and this through a great quantity of studies that focused on the determinants of activation of *syntactic movement* and *word-order manipulations*. However, although the activation of this part of the sentence network is recurrently reported in published brain maps and activation cluster tables, its contribution and sometimes even its presence are hardly discussed. Before discussing it in chapters 6 and 7, we take the occasion of presenting some initial contributions about its functional attribution.

The only occasion where it came on the front of the scene is thanks to a workshop held at Max Planck Institute in Leipzig organized by Christian Fiebach and Richard Schubotz in the year 2003. This event brought together cognitive neuro-scientist from two sub-fields, that of Language processing and of Actions processing[109], on the question of the degree of compatibility between functional model of Broca's area and the premotor/precentral cortex. Interestingly, the idea was put forward that Broca's area (BA44/45) and the left ventral premotor cortex (BA6) together with the frontal operculum support different functions during language processing. Note that the functional differentiation between Broca and the ventral premotor cortex is discussed in the context of the neuro-cyto-architectonical differentiation between agranular structure characterizing BA6 and a dysgranular cortical structure characterizing BA44/45. While Broca's area (BA44/45) can be seen as increasingly activated whenever the internal re-construction of a hierarchical structure from a sequential input is necessary, the Frontal opperculum (FoP) is reported to be involved in the processing of local structural dependencies, following Angela Friederici's interpretation that is differentiating between local syntactic structure building and more long-range sentence structure building (Friederici et al., 2006). Moreover, Fiebach and Schubotz (2006) associate ventrally distributed activa-

108. However, a seemingly competing interpretation of the role of ATL has been advanced by Liina Pylkkänen and her colleagues too. The reproducible and stable activation of ATL during minimal two-word phrases (be they noun-noun compounds Noun-Phrases in Zhang and Pylkkänen 2015, adjective plus Noun-Phrases in Bemis and Pylkkänen 2011 and 2012 or Adjective conjunctions in Poortman and Pylkkänen, 2016), but not in predicate arguments composition, should be understood as mapping to a conceptual combination of linguistic functions (see Del Prato and Pylkkänen, 2014 and Westerlund et al. 2015 in Arabic)

109. A Special Issue of *Cortex* in 2006 reports the different interventions and contributions of this workshop. See for example Friederici, Caplan, Grodzinsky, Koelsch, Amunts and von Cramon Position Papers.

tion of premotor cortex with the *mapping of sequential input onto structural linguistic templates.*

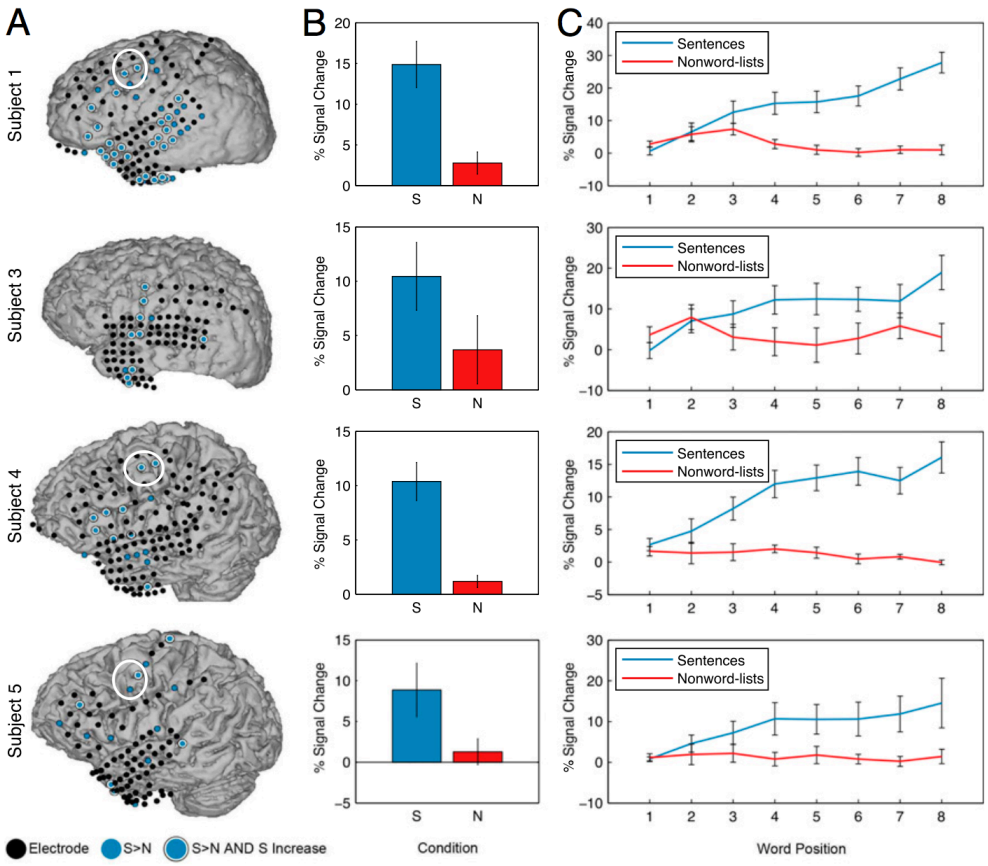


Figure 1.14 – (A) Cortical models of individual subjects showing all electrodes (black), electrodes that show a significant sentences > non-word lists (S > N) effect in odd-numbered runs (blue), and electrodes that show both a significant S > N effect and a monotonic increase across word positions in the sentence condition in odd-numbered runs, [i.e., electrodes of interest, EOIs (blue, circled in white)]. (B) The γ-magnitude for sentences and nonword-lists averaged across word positions estimated in even-numbered runs (i.e., data independent from the data used to select the EOIs). (C) The γ-magnitude for sentences and nonword-lists in each of eight word positions in even-numbered runs. Adapted from Fedorenko et al., 2016.

As for the dorsal part of Premotor cortex is not only systematically present in fMRI results from all kinds sentence localizer tasks and syntactic movement-related complexity manipulations, but also present in the lately developing intra-cranial studies investigating the neural implementation of syntactic information in the sentence (see Nelson et al., 2017 and Fedorenko et al., 2016).

As shown in Figure 1.14, even neuro-imaging results from intra-cranial recording tend to indicate that the "cerebral sentence network" counts an additional area in Precentral gyrus, above Broca's Complex. This study by Fedorenko and colleagues reproduces the now familiar paradigm comparing non-words' lists, jabberwocky (de-lexicalized) sen-

tences and real sentences, and highlights three main hot spots – i.e. specifically showing both a significant Sentence versus > Noun effect and a monotonic increase across word positions in the sentence condition – the sentence network among which the Precentral area we selected as an ROI is present (see blue, circled in white electrodes in Figure 1.14A).

1.4.5.3 Other 'new-comers'

Except on this isolated occasion that was, besides, focusing more on the role of its ventral part, we can say that Pre-central cortex activation in neuro-syntax studies hasn't really received the attention it deserves, especially if we consider its dorsal activation. We will obviate this problem in our literature reviews to be found in our experimental chapters 6 and 7, by discussing (a) its frequent co-activation with the (Pre-) Sensory Motor Area (SMA and Pre-SMA) and (b) how one of its potential role in the language network seems to be different from those argued by Friederici for two other brain areas, namely anterior Insula (aINS also called Frontal Operculum) and SMA. Among the recently debated language areas, another restricted set has started to attach researchers attention, for example Angular Gyrus (whose locus appears to be anatomically debated too) and Tempor-Parietal Junction (TPJ) (see Figure 1.13). However, as their functional characterization seems to be mainly pertaining to sentence process that are less directly syntactic, we will keep them for chapter 6.

In sum, neuro-imaging studies on sentence structure processing are generally consistent with the fact that Broca's area plays a role in syntactic processing, but they nevertheless fail to converge on a single region sub-serving all syntactic processes within the inferior frontal cortex (Newman et al., 2003). One or several computational/syntactic subprocesses involved in syntactic processing are for sure computed in the Inferior Frontal Gyrus (IFG), what is still open to discussion and to experimental evidence is exactly which.

However, this question is often considered as 'THE' cardinal question that neuro-imaging of syntax is actually facing. I do not hide here that this is not my point of view, as the reader might have understood in my critics towards of Broca-centric attitude, I do generally believe that a system (especially a complex one) can be understood and accessed by focusing in its extremes. Moreover, given the way recent neuro-imaging results indicate the existence of a distributed network, the question of a single brain region correlating with the representation and processing of syntactic structures appears 'relativizable' to me.

By contrast, the question of what are the other different brain regions linked to the representation and processing of syntactic structures in the language network received an initial experimental answer in a fMRI study by Pallier and colleagues in 2011. This study constitutes, among others, the initial grounding of part of this research project and offers neuro-psychological support to the linguistic hypothesis that words in a sentence are combined into word groupings (i.e. syntactic constituents).

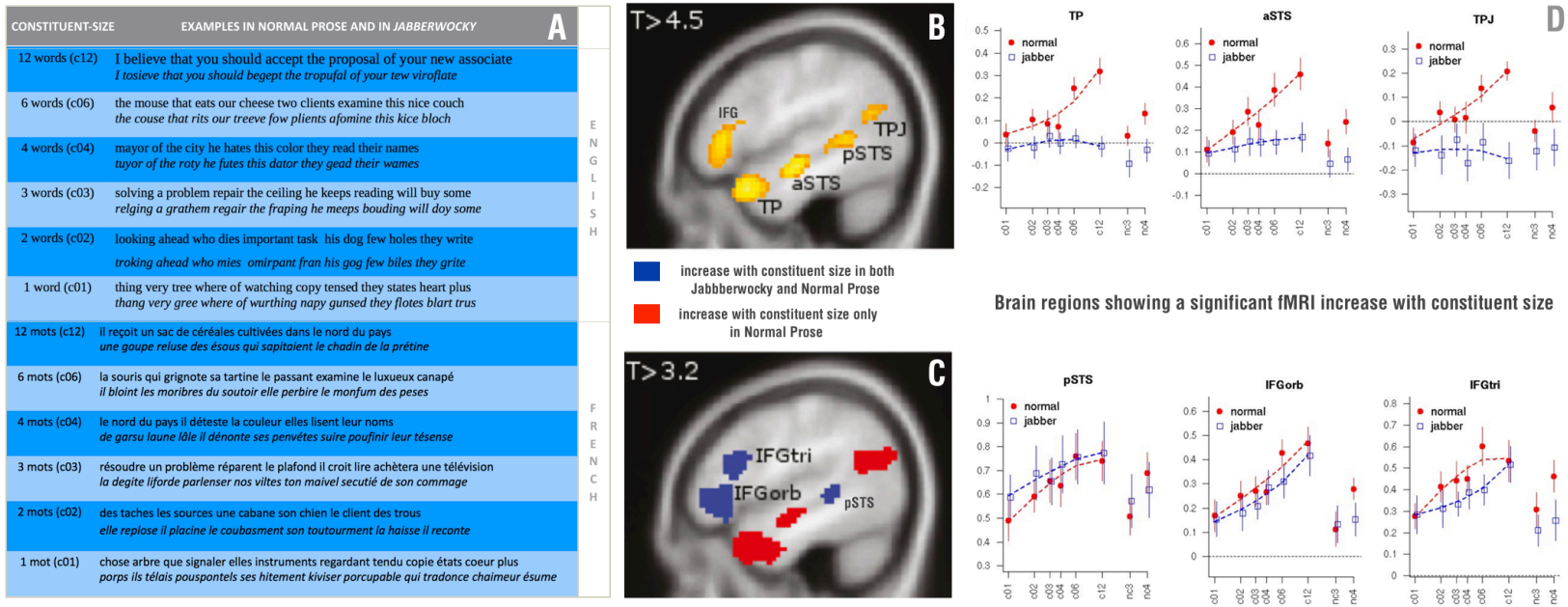


Figure 1.15 – The sentence hierarchical structures network: Brain regions showing a significant fMRI increase with constituent size are shown. (A) Stimuli table in French and comparable sentence conditions translated in English. (B) Results from the normal group who read sequences with actual French words. (C) Results from the jabberwocky group listening to delexicalized stimuli. The regions in red showed increasing activation with constituent size in jabberwocky, while the regions in blue showed a significantly stronger increase in the normal group compared to the jabberwocky group. (D) Amplitude of activations across conditions in the six regions of interest. Conditions c01 to c12 are organized according to a logarithmic scale of constituent size, and the fitting lines are from a regression analysis with linear and logarithmic predictors. Thus, a line on this graph indicates a logarithmic increase of activation with constituent size. 'nc3' and 'nc4' refer to non-constituent control conditions.

In this work, Pallier et al.'s research question could be resumed in the following terms: are syntactic tree structures an actual data structure in which the brain encodes sentence structures? Would increasing the sentence syntactic tree complexity generate an increasing cerebral activation in certain cerebral regions? In other words, if syntactic tree structure is encoded by a distributed cell assembly, its pattern must be increasingly complex as tree complexity increases.

Hence, participants were scanned while reading sequences of 12 words presented sequentially, while the size of the syntactic constituents that could be extracted was systematically manipulated as presented in Table A, Figure 1.15. One of the innovations of this study was to add intermediate levels to the basic contrast between sentences and word lists, which has been the classical contrast for identifying high-level language cortex in production (Indefrey et al., 2001; Golestani et al., 2006) and in comprehension (Kaan and Swaab, 2002; Mazoyer et al., 1993).

Interestingly, results show that areas concerned with constituent structure manipulation have an activation that increases with the size of the largest constituents present in the sentence (see brain map (B) in the Figure 1.15) [110].

An important aspect of this experimental paradigm is that half of the participants read sequences such as the ones listed in Table A (Figure1.15), and the other half read so-called jabberwocky versions, where the content words (but not the grammatical words) had been replaced by pseudo-words.

As one can see in figure 1.15, the fMRI activation of a network of peri-sylvian brain areas increased logarithmically with constituent size, irrespective of the fact that words were presented visually[111]. Furthermore, two regions, the left inferior frontal gyrus (LIFG) and the posterior Superior Temporal Sulcus (pSTS) showed this constituency effect irrespective of the use of actual content words or pseudo-words in the stimuli (see brain map (C) in Figure 1.15). Importantly, this last finding implies that these regions can build abstract syntactic structure from function words and morphology alone, in the absence of lexico-semantic information. In line with studies of natural and artificial grammar learning (Friederici et al., 2006), these areas, which are directly interconnected via the arcuate fasciculus (Catani et al., 2005), appear to play a fundamental role in coding for syntactic-tree structure. By contrast, anterior temporal regions and an area in the temporo-parietal junction exhibited robust complexity effects only when actual words were presented in the linguistic stimuli (cf. brain map (C) in Figure 1.15). One possible interpretation for the activation pattern of these regions is that they at least partly rely on lexical information to construct syntactic constituents.

In sum, the picture of the neural underpinnings of sentence hierarchical structures building and processing emerging from this study is of a distributed peri-sylvian network of regions – encompassing inferior frontal regions and temporal ones. Given this distributed network relevant to constituent-structure building, many questions remain open. Namely, a detailed view of how these different brain regions code for the different aspects of hierarchical sentence linguistic structure is still open. In other words, what are the different sub-processes these different brain regions are sub-serving, and more

110. The authors had also predicted that the peak of this response would be displaced later in time as constituent size increased, which was confirmed by the results and was understood as showing that the deeper constituents can only be computed after a delay, once the lower nodes themselves have been computed.

111. Note that the temporal profile analyses showed that the duration of activation increased in BA45, the posterior and the anterior STS, as predicted from a simple model of integration into constituents.

specifically, which area is intervening at what level?

These issues, linked to the functional characterization of this sentence structure building network – and more precisely how the study of certain syntactic structures and their transformations in Chinese can partly answer them – lie at the core of chapter 7. While chapter 6 addresses the question of how this sentence structure building network "reacts" when constituents in the sentence are moved from their original position.

We will conclude this section addressing an open-to-discussion question linked to this last hypothesis. Although the character of this issue can for the moment only be fairly speculative or theoretical, one could question the above hypothesis that underlies our choice for Pallier et al. ROIs, asking if the cerebral substrates for sentence constituent-structure hierarchy building are the same as those underpinning the representation of syntactic movement operations.

This raises the question of the network dedicated to sentence constituents-structure building mechanisms being hypothetically distinct from the one involving the sentence-structure transformation that is at stake when syntactic movement is theorized.

In other words, would a sentence displaying syntactic transformation elicit the same cerebral network as the one identified for sentence-structure building or should additional regions be recruited to perform the additional hypothesized step of transforming through a syntactic operation the basic canonical sentence into a movement-derived one?

In more concrete and methodological terms, this corresponds to asking whether the fact of choosing the network activated by the parametric increase in constituent-size in Pallier et al. (2011) as Regions of Interest (ROIs) is an adequate choice for investigating the neural underpinnings of syntactic transformations in French and Chinese.

1.5 Experimental methods to study Syntax: a few central assumptions

1.5.1 Theoretical assumptions from Cognitive Neuro-science

Before turning to the introduction of the imaging techniques that are used in the investigation of sentence processing and discussing some methodological assumptions involved in each one of them, we will focus on a few more theoretical presuppositions that Cognitive Neuro-science is bringing into the field of neuro-linguistics. In fact, certain assumptions about how cognitive processes and information processing happen in the brain, bear important consequences on disciplinary dialogue in the field of Cognitive Neuro-imaging, and more specifically in the domain of neuro-imaging of sentence comprehension. As we already saw in Section 1.3, one of the major contributions of syntax theory in Cognitive Neuro-science is that it can offer an articulate system to describe a representation of linguistic information that is computationally implementable. It is exactly on these last elements that we are going to focus next.

A classic approach to link conceptually 'what' (information representation) and 'how' (information processing) in cognition was proposed by David Marr[112] in the 1970s (see for example Marr, 1977). According to his theory on complex information processing systems (1982), a cognitive process should involve several levels of description: (1) the

112. David Courtnay Marr (1945-1980), was one of the originators of the field of Computational Neuroscience, he was Professor of Psychology at MIT.

computational level[113], (2) the *algorithmic level* and (3) the *implementation level*. Thus, Marr theorized the different levels at which an information processing device is to be understood. The top level is the abstract computational theory of the information processing device, a sort of mapping between one type of information to another, defining the performance of the machine. The abstract properties of this mapping should be defined in adequacy with the task, the linguistic one in our case. The second level is the representational level, stating the representation of input and output of algorithmic device implementing the computational theory selected. The third one is linked to the physical realization of the processing system.

In the table below are summarized the questions each of these levels should answer:

Computational theory	Representation and algorithm	Hardware implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

Figure 1.16 – Complex information-Processing System's characteristics. The three levels at which any machine carrying out an information processing task has to be understood. As presented in Marr (1982).

Declining these levels in the realm of our research on language and syntax would give the following:

1. 'Computational level'

As syntactic theory seeks to describe the fundamental property of the sentence – the input of comprehension and output of production – hence, the *computational level* can map onto syntactic theory to uncover the mechanism of the "human sentence processor". In 1963, Chomsky and Miller had already formulated that linguistic behavior, forming the empirical basis of computational theory of language, was rooted in an algorithm implemented in the human brain[114].

2. Algorithmic level

Representational assumptions of the linguistic input constrain algorithmic choices, and one of the major problems at the *algorithmic level* is the choice of the unit of

113. For terminological clarity, we should note that the term computational in Marr refers to the *goal of computation* and not computation itself (see Figure 1.16). Hence, this level of description is in fact usually referred as the *functional* level of description. The term computational in the rest of the manuscript is devoted to Marr's algorithmic level (see Figure 1.5.1).

114. This point will be carried out thoroughly in chapter 6 taking an example from Relativized Minimality (Rizzi, 1990/2001) in syntactic complexity processing.

representation. As already recurrently addressed in this chapter, this problem resides mainly in identifying (a) the fundamental building blocks of the syntactic representation and (b) the combinatorial mechanisms necessary to build the syntactic structure. Just as formal syntactic theory is generally considered as a computational theory of sentences, the theories of sentence processing developed by psycho-linguists can be considered in many ways as algorithmic theories. And why not repeating it again: our research project tries to bring elements of linguistic reflection and theory to answer the fundamental issue of granularity.

3. Implementation level

In neuro-linguistics the *implementation level* has been the focus of many bio-linguistic theories. Developing this level has leveraged results from neuro-psychological patients studies and neuro-imaging technologies (e.g. PET, fMRI, EEG, MEG, ECog, ect.), in order to identify the neuro-circuits involved in sentence processing: what will be called the 'neural substrates' or 'neural bases' of sentence comprehension.

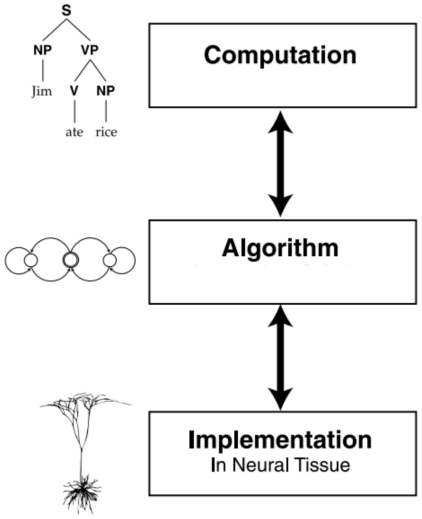


Figure 1.17 – Adaptation of Levels of Marr to Neuro-linguistics, in general terms. Adapted from W.T. Fitch, 2014, *Physics of Life Reviews* 11, p. 331.

In this perspective, the linguistic levels of analysis postulated by linguistics have been assigned the role of *representational primitives*: articulatory features, phonemes, syllables, morphemes, nominals and verb phrases, clauses and constituents, sentence, discourse and narratives have been alimenting the language (domain-specific) representational level of Marr's architecture.

The different levels in Figure are in principle separated, but experimental investigation often forces the researcher to consider the possible interaction between them. Any study of the computational level can only be constructed by observing human linguistic behavior, like acceptability judgment (i.e. the well-formedness of the syntactic structure). However, it is easy to picture that these judgments are closely related to the algorithmic and implementation levels too, for instance in a task implying the processing of complexity.

To these *representational primitives*, a complete toolbox of *computational primitives* are to be added to Marr's algorithmic level. And, given the characteristics of language processing, *computational primitives* have been hypothesized; we can list the most relevant ones for our work: discretisation, sequencing, concatenation, ordering, incrementality, grouping in constituency, hierarchy, establishment of relationship both local and long distance, transformation, prediction[115].

Finally, we can raise the question of the implementation level which is now at the core of Cognitive Neuro-imaging of language faculty: what kind of neural circuits or dynamics may underpin these primitives? Or, choosing a formulation that is more intimately linked

115. Note that the domain specificity is not established for these computational primitives yet, and that they are generally considered as being domain general in humans (see 44, about Fodor theory of the Modularity of Mind).

to our on work sentence comprehension and processing: what is the neuro-biological infrastructure for the representation and processing of sentence's structural complexity?

In conclusion, understanding the relative configuration and interdependence between the three levels in language research is one of the main goals of a complete description of the Cognitive Neuroscience of Language. We can therefore say that this configuration may suggest that any further progress can be achieved only through the closed interaction of syntacticians, psycho-linguists and neuro-biologists (cf. the issue of pluridisciplinarity presented in the Epilogue, p. 705).

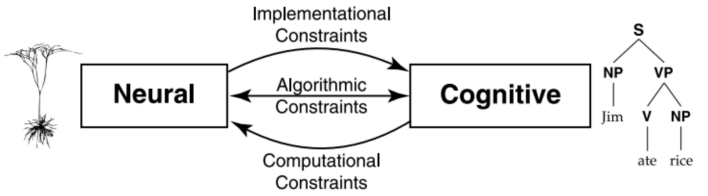


Figure 1.18 – Schematic representation of the goal of understanding the intricacy of today's Neuro-cognitive description of Language at the three levels proposed by Marr. Adapted from W.T. Fitch (2014) Physics of Life Reviews 11, page 331.

1.5.2 Neuro-imaging methods: a few central assumptions

Until approximately four decades ago, the only way to investigate language in the brain was the study of brain lesions and their impact on the language behavior of patients. The selective impairment of specific language aspects would show evidence for mapping this particular aspect on a correspondent lesion site. Neuro-imaging tools have been gradually relativizing this investigation practice, also because brain damage often covers large areas, preventing from constructing a fine-grained mapping between selective impairments and brain areas. Another reason is that sometimes, findings from aphasiology hardly generalize to normal brains compared to damaged ones, this is mainly due to different individual patterns of recovery and plasticity. However, in the last ten years an increasing number of studies with large-sample of aphasics' cohorts have been offering more and more detailed brain damage maps that are accompanied by very detailed linguistic competence test (see Wilson et al., 2012 or Sapolsky et al., 2010).

With the advent of functional neuro-imaging tools, researchers could not only examine the healthy brain at work, but these tools allowed also to study the functional organization of the brain going beyond case studies and dissociations offered by lesion studies. Specifically, ERPs (event-related potentials) and MEG (Magneto-Encephalography), with their high temporal resolution, are informative of the time course of a language process, whereas fMRI (functional magnetic resonance imaging) and PET (positron-emission tomography) provide the necessary spatial resolution to determine the neural *loci* of cerebral language processing.

Overall, these tools and their results have challenged what was known about syntactic processing, essentially showing that the language network as a system is organized into a large number of smaller, interconnected, and tightly clustered (i.e inside Broca complex) brain areas contributing together to language processing. To the extent that the contribution to experimental Cognitive Science of functional Magnetic Resonance Imagining (fMRI), EEG as well as single cell recording is nowadays undeniable.

One of the first central assumptions that Cognitive Neuro-imaging techniques have is that regions engaged in cognitive activities require higher level of activation, which in the case of fMRI would imply higher levels of oxygenation (i.e. Blood-oxygen-level dependent signal – BOLD signal) by an increase in blood-flow in a given region. This variation of

blood-flow is detected by instrument of measure like fMRI that detects the slight changes in a stable magnetic field caused by the blood-flow increase[116]. However, one should be careful because the lack of BOLD signal may not mean the region lacks activation, and moreover, BOLD signal may be the result of both excitation and inhibition.

In PET (Position Emission Tomography) molecules with unstable isotopes are injected in the blood, giving the possibiity to detect their radioactive traces in brain regions where their concentration increases with oxygen demand linked to higher cognitive activity. As these two techniques have high spacial resolution, but a low temporal one, they are mainly used to compare the cerebral activation to linguistic stimuli in terms of localization and higher intensity of activation[117].

Other types of linguistic stimuli need to be investigated with high temporal resolution, especially when one wants to observe cerebral activation in real-time at very precise spots in the sentence, like it is the case in sentences presenting local syntactic or semantic ambiguities or garden-path effects. For instance, Electro-Encephalography measures at the surface of the scalp the electric potentials generated by brain activity through an array of electrodes. The interest of techniques enabling high temporal resolution like ERPs and MEG is linked to the presupposition that language understanding proceeds incrementally and that various sub-components of sentence processing interact in time. Namely, we will use EEG in an Event-related Potential experiment to tag in time the moment in which the information from the linguistic context is processed in sentences having different intonational patterns.

1.5.3 Neuro-imaging: a tool for decomposing Cognitive processes

To these central assumptions linked to the methods in use in the field of neuro-imaging, a few others linked to practice, and as essential as the central ones, should be added. In fact, researchers conducting experiments using these techniques tacitly adopt some shortcuts related to experimental practice and methodology (Bechtel and Richardson, 2010). Here is a concise list of the most relevant and useful ones:

1. The brain contains some regions that are specialized for processing specific types of information. However, regions integrate information from a large number of other regions, which is engaging each region in a complex dynamic network;
2. Functional MRI, by identifying particular areas of increased blood-flow, seems to support the idea that there are highly specialized regions containing modules. However, finding different BOLD activations is not tantamount to identifying cerebral modules;
3. Even though there might be some highly specialized processing regions, these regions might process information generated in other areas and they are also connected to even larger networks of feedback.
4. To obtain a brain map of cerebral activations, one needs to contrast two experimental conditions. This methodological practice has major bearings on the way experimental conditions are designed. So that, articulating the reflection about sentence processing implies systematically decomposing syntactic ability for sentence structure into sub-

116. A side note should be added here. We should say that BOLD signal might not be sensitive enough to neural activity linked to very basic and fundamental processes in that cortical areas might not be right neuro-anatomic granularity at which one should look to find neural cerebral correlate for certain syntactic processes involved in sentence understanding. Other techniques might be more sensitive, and especially temporally sensitive, which could favor the aligning of syntactic theory and cognitive neuro-science.
117. In other words, fMRI and PET studies usually compare linguistics stimuli that are hypothesized as being processed in different parts of the brain or as requiring a greater processing cost.

processes. Therefore, contrasting two sentences would mean subtracting the hypothesized *sub-processes* or *representations* of the second to those of the first.[118]It should be noted that the investigation of the cerebral representation in addition to the classically invoked processes constitutes one of the originality of this work and of multi-variate approaches to neuro-imaging in general. We want to stress here that focusing on the representation means consequently to understand neuro-imaging data as informative about the cerebral *encoding* and not only processing of linguistic structures.

It is important to note that the distinction between mental representation and mental processing[119] will be fundamental in this work. Distinguishing the manipulation of a given representation format from the representation itself reflects the contemporary concern for the investigation of the neural code used by the brain to represent cognitive objects, like the sentence in this research project. This can qualify as a more contemporary methodological understanding of neuro-imaging.

In this research work, we will indeed consider the hypothesis that the representation of syntactic structures with its 'information data format -to say it *à la* Pallier- are distinct from their processing, and consequently distinct from the activity elicited by their processing. This will constitute a leitmotif of this work, although it is still not sure that our experimental data can give results that directly bear on this research issue.

However, given the great amount of technical details and its methodological implications (take for example technicalities linked to fMRI design models), it is beyond the scope of this Section to be exhaustive. Nonetheless, it should be recapitulated that this overall framework urges researchers to engage in constructing experimental designs contributing to the maturation of cognitive theories in order to discover the functional mechanisms that are responsible for the phenomena under observation. Once more, the investigation of the brain implies to identify structures that instantiate a given cognitive process, this is possible only in a given neuro-psychological model that would offer hypothesis to decompose cognitive processes (Logothetis, 2004; Richardson, 2012). More than in other cognitive processes, language studies have the need to be grounded onto a theoretical framework offering detailed account of both the computational – linked to human language capacity –, and of the neural representation of linguistic knowledge. We are going to refine this framework in the next chapters of Part I.

We hope the reader did not just understand these Preliminary Considerations as a respect payed to the gravestones that historically paved the way to the research work and discoveries we are making today. Having walked down this trail will allow, in the next pages, some perspective on the different issues that are at stake when considering: (i) Sentence as fundamental to human linguistic faculty (and therefore a central question for the definition of man); (ii) Sentence as universal across natural (and probably artificial) languages and (iii) Sentence as a fundamental expression of syntax, or as the "kingdom of syntax".

118. Henceforth, fMRI contrasts will be indicated by the symbol '>'. Like in the expression (a>b), literary meaning the brain activation to sentence (a) is superior to the brain activation to sentence (b). All the brain maps presented in this manuscript are the result of the subtraction of the brain map activation to an experimental condition to the one of the second item of a given minimal pair.

119. As we saw in sub-section 1.3.1 page 38, the philosopher Thomas Hobbes had formulated an intuition going in this direction: thinking amounted to performing arithmetic-like operations on internal structures (i.e., mental representations).

PART

Introductory matters

"Novum in vetere latet
et in novo vetus patet."

Augustine of Hippon 354 – 430

"Y así, del poco dormir y del mucho leer, se le secó el cerebro, de manera que vino a perder el juicio."

["Finally, from so little sleeping and so much reading, his brain dried up and he went completely out of his mind."]

Miguel de Cervantes (1547 – 1616), *Don Quixote de la Mancha*, Book I, 1605

A central element of this manuscript are a certain number of states of the art and literature reviews[1] on different cardinal topics addressed by the different disciplinary fields that compose the starting "mosaic floor" for our investigation of the cerebral representation of the sentence.

These two initial chapters draw the background discussion that sets the foundation for the current research, and offer a review of what it means to consider the sentence-unit as a cognitive object, giving it an experimental and linguistic grounding. They outline the linguistic and theoretical motivations behind our experimental research, and provide syntactic analysis of the the linguistic phenomena that were selected to carry out our experimental research (§Part II). In Part II, a series of experimental studies addresses at multiple levels the neuro-linguistic investigation of the sentence-unit as a cognitive unit. More specifically, the sentence-unit will be considered:

1. as a unit having a specific prosodic pattern, in chapter 4;
2. as a unit being linked to context, in chapter 5 ;
3. as a unit transformable through syntactic transformations in a structurally more complex object in chapter 6, and
4. as a unit having multiples representational layers and different syntactic means to achieve co-referential link, in chapter 7.

But before entering in the experimental contributions of this manuscript and in their labyrinths of details, allow me to cast a broader outlook on sentence as a linguistic unit and more specifically as a complex object.

1. This is given by the intrinsic pluri-disciplinary nature of this research work, but also because I personally learned in these years of '*initiation*' to research that reviewing the past helps understanding the present. Which is nothing new, indeed. However, having experienced it makes me really grateful to my two supervisors that both encouraged me in this way. Secondly, this experience makes my grateful towards the Past, with a capital 'P'. Before starting, one could have defined me as a future-oriented mind, and I would not have thought I could have such a passion for the Past. Hence, what Augustine of Hippon used to say about newness and antiquity is actually true in my itinerary: New is hidden in Antique, and Antique is unveiled by New, or in a more literal translation New is latent is Old, and Old is evident in New (Augustinus, Quaestiones in Heptateucum, 2,73: PL 34,623.)

Chapter 2 will be the occasion to articulate a pluri-disciplinary discussion around the sentence-unit and to explain what it means to consider it as a cognitive object. This first chapter will be the theater of a continuous '*holistic*' back-and-forth between syntactic theory and experimental attempts in psycho-linguistics and cognitive neuro-imaging to understand this exceptionally complex level of linguistic description.

Rather than attempt an exhaustive review, our intention is to present the theoretical and notional context of our study. We believe that an approach altering theory and experimentation is necessary for understanding the syntactic faculty of the *homo phraseologicus* defined at the beginning of the previous chapter.

In this way we will introduce the reader to the neuro-linguistic studies that are illustrative of how the theoretical linguistic context we chose to investigate syntax in the human brain is empirically grounded. Namely, the state of the art contained in these two initial chapters will also provide a gradual introduction to the different research questions of the this doctoral work and the experimental hypotheses motivating the five experimental chapters constituting Part II of the manuscript. We will retrace the linguistics, psycho-linguistic, neuro-psychological and neuro-imaging arguments for investigating certain sentence structures compared to others and certain syntactic processes compared to others.

Although the initial chapters gathered in Part I may carry a genuinely informative character -that the different readers might possibly find generic or redundant when it will come to their own specialization-, its value probably does not reside only in the great deal of information reviewed, but in the fact it may permit to find a way to put into resonance the different approaches to sentence-unit that constitute the disciplinary back-ground of this work.

Chapter 2, a polyhedral point of view on the sentence

> "Toute la curiosité, il est vrai, dans le cas d'aujourd'hui, porte sur *l'interpénétration*, mais en parler, impossible sans la confronter au ***concept***."
>
> *[It is true, nowadays, all the curiosity is about interpenetration, but its impossible to speak about it without being confronted with the* ***concept****.]*
>
> MALLARMÉ *Divagations*, CHAP. HAMLET, 1897

Chapter 2 sets the stage for all the research questions that will be dealt in this manuscript by introducing the theoretical context, the scientific goal of our research, and the main hypothesis and claims of the thesis, which will be gradually build up by a paced step-by-step argumentation.

This chapters is the fruit of a meticulous excavation in the literature of at least four disciplines (linguistics, psycho-linguistics, neuro-psychology and neuro-imaging) that enriched my understanding of what the brain could represent of the sentence-unit and of its syntactic structure.

As stated by Mallarmé in the above epigraph the interpenetration cannot avoid being confronted with concept, and the interpenetration of these disciplines and of their respective points of view and approaches on the sentence-unit will offer the occasion of a notional introduction to a series of formal concepts that will be instrumental to ground our experimental work in Part II. This notional work we will be carried on, particularity in chapter 2, it will mainly attempt to put into resonance the different disciplinary analysis and findings to delineate a definition of sentence as a natural (§2.1) and universal unit (§2.2) across languages, having both a basic internal structure (§2.3) and an internal structural complexity (§2.4). All the argumentation on sentence as a natural, universal, structured and complex syntactic unit will be elaborated in the effort to find a new comprehensive way to define Sentence as a representational unit for the mind and the brain.

As our dissertation title states sentence will be considered here as the object of cognition, and not only as the largest syntactic unit of the language system. In other words, the sentence as a linguistic object will be see from the perspective of the mind and of the brain. Therefore, a linguistic and more theoretical approach will run in parallel with a more empirically-oriented state of the art, where sentence's complexity will be seen from the polyhedral angle of theory and experimentation in Linguistics, Psycho-linguistic and Cognitive Neuro-science[2]. This itinerary will lead to show how, for all these different aspects characterizing the sentence-unit, linguistics and psycho-linguistics have provided crucial and testable hypotheses to examine the neural implementation of sentence comprehension through the experimental results that have been gathered since the 1950's.

The unifying focus will be to see what are the sentential properties as a syntactic-unit that can be found across all languages, showing how the fact that all natural languages do structure the utterance in a syntactic-unit having certain characteristics makes the sentence-unit a level of representation that can be defined as natural, thus, begging for a coherent biological/cerebral implementation and description. By offering linguistic examples from a vast cross-linguistic data-set we want to offer the reader an outlook on linguistic facts combining Typology and formal approaches to grammar and syntax.

Hoping to demonstrate the potential of cross-fertilization of these two approaches, on one side, Typology will offer a rich linguistic description of sentential phenomena focusing on their linearity dimension. Given its more data-driven approach, the typological construction-based approach will offer a detailed account of the surface properties of sentence structures, where linguistic regularities linked to sentence linearization (e.g. word-order) emerge from large cross-linguistic corpora and studies. On the other side of *linearity*, the hierarchical sentential dimension, will be theoretically grounded thanks to the rich set of abstract descriptions of the sentence hierarchies and of syntactic operations offered by formal linguistic approach. Given the more abstract level at which these linguistic generalizations take place, they can better fit the neuro-cognitive attempt to investigate the operations of the mind on sentence structure and to identify its neural underpinnings.

Hence, we will try to convince the reader that the neuro-linguistics investigation of the linearity and hierarchy dimensions distinguishing the sentential linguistic phenomena, can benefit of representing syntactic structure in its two dimension (i.e. linearity and

2. For the reader that would approach chapter 2 (and maybe chapter 3) just reading information about his own field of expertise we have to confess that we really made it really difficult for him to do so. The rest of the manuscript will be more easily bite-able as desired.

hierarchy) thanks to a representational tool -the syntactic-tree- that can offer concrete testing hypotheses to investigate the neural implementation of syntax in the brain.

For this reason, the reader will be introduced to a number of different syntactic complexity measures linked to the representation of the sentence structure in a tree-like format.

Not only syntactic complexity will be defined linguistically, but in order to inform the delineation of a neuro-imaging state of the art on how sentence's internal structure can be represented by the brain, psycho-linguistic en neuro-psychological findings will guide the delineation of our experimental hypotheses too.

One issue, that will constantly be in the filigree of this first Part I, is that the bulk of current neuro-syntax research could be seen as beset by a *Granularity Mismatch Problem* (see Poeppel and Embick 2005), as we argued in chapter 1 with our review on syntax and the brain[3].

Specifically, we claim here that this problem could be caused by the coarse-grain level at which linguistic computation is generally posed to be taking place in the brain. We oppose to the the broad conceptual distinctions of sequential word-order, scrambling, embedding and word composition or unification that form the actual basis of the main stream approaches in neuro-imaging of the sentence-unit, a more fine-grained outlook on syntactic complexity and the related processes.

Namely, we will argue for the cognitive and neuro-psychological adequacy of syntactic theoretical fine-grained distinctions concerning phrase structure, movement types, feature-checking, empty syntactic position, sentence domains, syntactic phenomena at the interface with discourse, and ultimately syntactic-tree representation for deriving complexity measures to be correlated with brain activity during sentence's comprehension.

These syntactic details and processes are not only central to linguistic theories, but have initial experimental and neuro-psychological evidence in the literature, as a retrospective account of current and past imaging data will demonstrate.

3. cf. §1.4.5 underlying the excessive focus on Broca's area contribution to syntactic faculty in what could be defined an oxygen-free marriage impeding a salutary oxygenation thank to some close friends brain areas that constitute the sentence cerebral network (see page 61).

Chapter 2

The Sentence: a natural universal structured complex syntactic object

> "Il y a énonciation dès que se réalise l'un de ces actes discrets et chaque fois uniques par lesquels la langue est actualisée en parole par le locuteur."
>
> *["Utterance occurs as soon as one of those discrete and each time unique acts by which language is actualized into speech are realized"]*
>
> Ch. Bally, 1956:251

This chapter, we will address the heart of current research: the sentence-unit, and describe some of the characteristics of the sentence-unit that will appear to be of special relevance throughout the thesis.

As mentioned in the Preliminary Considerations, the sentence as a unit, with its structure, lies at the core of human language faculty. We already went through a number of different view angles to look at the sentence: the sentence as an utterance, manifesting human thoughts (§1.1.2, p. 12), the sentence as a syntactic structure (§1.1.3, p. 14), the sentence as a typological word-order unit varying across languages (§1.2, p. 22), the sentence as a faculty of the mind (§1.3.2, p. 41), the sentence as a cognitive calculation (§1.3, p. 37), and finally as being underpinned by a cerebral processing network (§1.4 and, p. 45).

After the *innere Sprachform* of von Humboldt, Charles Bally[1], one of Saussures precursors, brings back on the stage of our linguistic reflection the topic of the 'internal form' in an new and direct manner by involving once more the *mind*. In his *Précis de stylistique* (1905), he redefines linguistics as[2]:

1. Charles Bally (1865-1947), was the co-editor with Albert Sechehaye of the *Cours de Linguistique Générale* of Ferdinand de Saussure. He is evidently to be considered a disciple of Saussure but he also development his own and independent reflection on language. Starting from the idea of establishing a*psychological linguistics*, he aims at a linguistics that would take into account '*le sujet parlant*' the speaking subject: the study of the linguistics systems should run in parallel by the study of the activity of the speaking subject. It is from this interest for what is happening in the mind of the speaking subject that he will articulate a *theory of enunciation*. For a detailed introduction to his contribution to linguistics see Chiss (1985).

2. In French: "[la linguistique] est basée sur l'observation de ce qui se passe dans l'esprit d'un sujet

> "being based on the observation of what happens in the mind of the speaking individual when he utters what he thinks"

One couldn't find a more perfect definition to ground our investigation of the sentence syntactic unit as a cognitive object. This is the level at which we will situate the observations, arguments and analysis that the reader will find in this Chapter. Given this linguistic *mind-set*, we will now address the issue of what is a sentence from a more linguistic and psycho-linguistic point of view to actually have all the means to finally ask what it means for a sentence to be a natural unit, to be universal, to be syntactic (i.e. a structure), and to be complex from the brain view angle.

Yet, the reader should be reminded that the attempt to draw a relatively complete, and tentatively clear outline of the sentential phenomena with the theory that goes along with the sentence-as-a-unit, is shaped by the need to specify anything we or the brain (minimally) need in order to represent and process sentences units. This will occasionally bring us to review also some of the most basic concepts and processes found in a formal approach to syntax, in that they will greatly help to carry along our chief interrogation about how syntactic structures are instantiated in the human brain. Thus, we will consider a series of syntactic notions and processes, constituting what the mind and the brain should essentially instantiate to represent and process sentences and their structures:

1. different ways to manage reference assignment inside the sentence unit (§2.1.1)
2. a hierarchical structure building operation (§2.3.1 and §2.3.2)
3. a structure building process able to complexify basic sentence structures present in each language by displacing sentence elements (§2.4)
4. and a system to identify structurally and theoretically posited null elements like null objects or other elements that are understood but not phonetically realized (§2.4.3)
5. syntactic means to orchestrate the link between the sentence-unit and the discourse-level (i.e. the sentence discourse-interface (§2.1.2, §2.2.4 and §2.4.4)

Along this neuro-linguistic 'trail', internal and external evidence will be systematically presented for the different properties of the sentence-unit in this chapter. Thereby, we will introduce the notion of *psychological reality* of linguistic structures, which posits that these linguistic representations are not only merely descriptively convenient, but can notably be considered to be psychologically active during linguistic behavior, and therefore 'real'. This will lead to gradually present our central experimental hypotheses for the experiments in Part II, and, to introduce our more general claim that sentence structures posited in linguistics not only can be correlated to processing cost recorded in brain-imaging, but could also identify the representation format the brain uses to build and understand sentences.

We will proceed by exposing first linguistic arguments in favor of each of these definitory properties (i.e. natural, universal, syntactic, complex, unit) and then gather some examples of experimental evidence both from the field of psycho-linguistics and of neuro-imaging. So that, the promise of having in each chapter a theatrical act with its Episodos

---

parlant au moment où il exprime ce qu'il pense" (in ch.I:83-84, Genevra, 1905), or elsewhere we can read: "L'étude de la langue n'est pas seulement l'observation des rapport existant entre des symbols linguistiques, mais aussi des relations qui unissent la parole et al pensée [...] c'est une l'étude en partie psychologique, en tant qu'elle est basée sur l'observation de ce qui se passe dans l'esprit d'un sujets parlant au moment où il exprime ce qu'il pense." in *Traité de stylistique française* (1909:2).

(experimental evidence, Stasimon (linguistic arguments), Kommos (literature review and discussion) will be kept.

Overview of the contents of this chapter

2.1 The Sentence is a natural unit

"Nullum habile membrum est, si corpori par est. Splendor orationis quantus nescio an in ullo alio fuerit : ***non lexis magna, sed phrasis."***

SENECA THE ELDER,
Controversiarum, LIB. 3:247

"Anyone who seriously approaches the study of linguistic behavior, whether linguist, psychologist, or philosopher, must quickly become aware of the enormous difficulty of stating a problem which will define the area of his investigations, and which will not be ***either completely trivial or hopelessly beyond the range of present-day understanding and technique."***

NOAM CHOMSKY, 1959

Asking what is a sentence can appear completely trivial, as this linguistic unit may be considered as a totally self-evident one[3]. However, trying to define some of its properties will help us grasp the cognitive processes that are implied in its understanding and cerebral processing.

There is no doubt that sentences have always been, and will probably always, be paramount for the linguist, but what we want to put forward here is that sentence-as-a-syntactic-unit is a ***natural*** level of linguistic description *and* knowledge, this fact consequently begs for a biological foundation. In other words, the sentence, in virtue of its naturalness *should* be represented in the brain as an object, and should ultimately constitute both a representational and a processing unit.

Non lexis magna, sed phrasis: more then words

An intuitive answer to this question was offered early in the XIXth century by Wilhelm Wundt[4], who was already framing the problem of linguistic description as the description of linguistic knowledge. In his investigation he advanced the idea that:

> "The natural unit of linguistic knowledge is the intuition that a sequence is a sentence."

3. "A Review of B. F. Skinner's Verbal Behavior" in *Language*, 35, No. 1 (1959), 26-58., 2007

4. Wilhelm Wundt (1832-1920) was a German experimental psychologist that deeply influenced L. Bloomfield. He is considered a funding figure of modern psychology and carried out an in-dept reflection on the psychology of ordinary people's linguistic behavior.

His claim was based on some simple linguistic observations showing that words were not the basic and natural unit of linguistic knowledge. His first argument was that even an isolated word like 'stop' or 'stay' are not understood as words, but as a sentence: "*Stop!*". Therefore, as for the definition of what a sentence is, his argument goes as follows: one cannot simply define sentences as the sequence of the words constituting it, because there are single-word sentences. Secondly, he noted that the meaningful relation between words is not sufficient to define what a sentence is, because there are meaningful relations within certain word sequences that, nevertheless, are not sentences ; consider for example the following sequence : "Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday"[5]. This kind of considerations brought Wundt to claim that the sentence must be defined as a sequence that native speakers of a language *intuitively* believe (1) to convey a complete proposition in (2) a linguistically acceptable form.

To these arguments pointing at sentences-level as being the reference-unit of linguistic knowledge, should be contextualized with the advancement of linguistics at the time Wundt was writing. Non-Indoeuropean languages had in fact largely brought linguists to assert that words in a sentence were much more than simply sequences, an that there were patterns and regularities. Moreover, the study of syntax in the XIX$^{\text{th}}$ century was already about the organization of words into abstract syntactic structures expressing grammatical relations between words and phrases independent of their meaning. In his *Essays on Mental Structure* (2007), Ray Jackendoff echos Wundt's assertion by arguing :

> "In language, the sentence provides the ***minimal domain*** into which elementary meanings can be placed and combined."[6].

It is indeed commonly acknowledge that meanings of words are **compositional**[7] at the sentence-level, in that they modify and impact the meaning of that individual words would have in a isolation.

Consider for example the isolated words *hate*, *Sparta*, *Athens*[8]. The first word, *hate*, is an internal state featuring a negative feeling towards an entity; the second is a warrior's city-state of Greek antiquity; and the last is an actual city and the capital of ancient Greek philosophy. Composing these words into a sentence unit - "*Sparta hated Athens.*" - the whole unit actualizes some particular elements of the root meanings of the isolated words into a new interpretational unit where for example the predicated meaning at the sentence-level limits to ancient times the rivalry, and by virtue of the fact that these city-states are form Antiquity, the notion of hate is actualized by our general knowledge that at that time hatred could not result in a suit in the tribunal or in the European court of Human Right, but rather a decision to send an army towards that city. The same is valid for Athens another sentence could have actualized its actual identity of being the capital of a European country of a touristic place. Hence, the sentence-level unit and its arrangement not only isolates certain meanings of the words -like in the case of polysemous words whose meaning is actualized by the sentence context-, but it

5. Lists of words were already chosen as a baseline for sentences as we will see it will be the case in the first 10 years of neuro-imaging of language.

6. Jackendoff, R., Language, Consciousness, Culture: Essays on Mental Structure (Jean Nicod Lectures). Cambridge, MA : MIT Press, 2007. p. 403.

7. The term compositional indicate here the fact that the meaning of words appears to be linked to that of other words in the sentence.

8. We reproduce here an example from Townsend and Bever (2001), in honor of their founding book and to enjoy speaking a bit about antiquity.

provides a structure that can access conceptual knowledge to yield a new and specific sense, of the word *hate* in our example.

Experimental psychologists from the 70's added to these observations on 'compositional' sentence-level meaning, three qualifications of the sentence-unit, stating it is a verifiable, answerable and remember-able unit (Johnson-Laird, 1974). But, before turning to these experimental considerations we will review linguistic internal evidence that there exist natural markers of this sequence, and evidence for certain linguistic phenomena, which are taking place only inside this linguistic sequence. Phenomena like punctuation, sentence final markers and co-reference phenomena will be instrumental in defining sentence as a natural unit.

2.1.1 Co-reference phenomena

One of the first aspects that participate to the intuition native speakers have of the sentence as a unit, consist in showing linguistic arguments that certain syntactic phenomena are uniquely possible within sentential boundaries, like a particular set of sentence-internal relationships and dependencies that are commonly called co-reference. Hence, when an element of the sentence gives its meaning to another noun in the sentence, it is usually called its antecedent[9]. However, the notions of co-reference, co-indexation[10], or antecedence are actually quite general ones, and a closer look at sentence-internal reference between nouns, pronouns, and syntactic positions -the so-called Binding phenomena- reveals that there are dependencies and relationships that can only apply inside certain syntactic boundaries.

Consider the cases in example (4), the reflexive pronoun *itself* in (a) is clause-internally bound to its antecedent *Sparta*, and in (b) the pronoun '*it*' cannot refer to its antecedent NP but has to refer to an exogenous referent to its own clause.

(4) Refexives versus Pronouns

a. When Sparta$_i$ attacked itself$_i$, Athens$_j$ did not survive long.

b. When Sparta$_i$ attacked it$_{j/*i}$, Athens$_j$ did not survive long.

We can see from the distribution of anaphors in (4), there exist a constraint on this type of antecedent-anaphor relation:

1. Reflexives like in (a) can be qualified as anaphore having endogenous reference ([+Anaphor] and [-Pronominal]), *their possible antecedents are strictly limited to the domain of their own clause*,
2. while pronouns like in (b) can be defined by the features [+Pronominal] and [-Anaphor] in that they must refer to some antecedent *not immediately in the domain of their own clause*.

This difference between exogenous and endogenous reference allows syntactic theory to identify the domain of the scope of various kinds of syntactic rules. Importantly, these domains depend more on locality understood in hierarchical terms than on linear proximity, in fact in (4b) *it* and *Athens* are contiguous. Hence, the Binding Principle A formulate the distribution of anaphors as follows: An anaphor must be bound in its

9. It is then marked by subscript letters called indexes

10. Binding counts as a kind of co-indexation, that happens when one of the two NPs c-commands (cf. Glossary) the other, namely co-indexation alone does not constitute binding.

binding domain (i.e.the clause containing it)[11]. We will resume to these issues in §3.2.3.3, where reflexivisation will be used as a test to analyze the distinctive characteristics of Topics versus Subjects[12]

Accordingly we can list the different kinds of antecedent-referential relations that can be found in the sentence unit:

1. Referential expression are NPs that get their meaning by referring to an entity in the world having the feature [-Anaphor] and [-Pronominal];
2. Anaphors are Noun Phrases that obligatorily get their meaning from another Noun Phrase in the sentence ([+Anaphor] and [-Pronominal]) ; and
3. Pronouns are Noun Phrases that can get their meaning from another word in the sentence, carrying the features [-Anaphor] and [+Pronominal].

These basic features of anaphoric relations, reveal that sentences have syntactic devices to cross-refer from one point of the sentence to a phrase in another part, but that this link is bound by certain rules. Here is an initial argument to demonstrate why structural relations are so important in the sentence, and how investigating sentence-internal dependency-links, from the brain point of view, can constitute a window on the cerebral representation and processing of sentential boundaries[13].

We will experimentally ask this question in chapter 7 studying some particular syntactic configurations that Mandarin Chinese offers on these co-reference issues (cf. chapter 3).

Chinese examples offering cross-linguistic validity of the above binding patterns can be found in Xu Liejiong (1996), who presents a cautionary tale in his review of judgment studies on Chinese reflexives. The debate on reflexives in Mandarin started when Battistella and Xu (1990) found that naive Chinese speakers consistently interpreted the reflexive 自己 *zìjǐ oneself* in complement clauses as co-referential with the matrix subject, a pattern that would indeed violate the above cited principle. This experimental result was subsequently explained by arguing that the long-distance binding pattern reported by Battistella and Xu (1990) was due to the use in their experimental stimuli of particular matrix verbs like 告诉 *gàosù* 'tell' that can semantically facilitate the binding pattern with the reflexive '*oneself*'. Later on, Ho (1995) controlled for these verbal biases by using verbs like 劝告 *quàngào* 'advise', and found that speakers preferred local binding like in the above example (4a).

The literature on this topic is extensive and beyond the scope of this section, but Xu Liejiong (1996) concluded that this variation in linguistic judgments was requiring to take into account the structural prominence linked to some pragmatic articulations, a hypothesis that was made more precise later on by Huang and Liu (2000) and Pan Haihua (2001). All in all, these results from Mandarin show that the definition of *binding domain* is far from trivial, and the following examples show that English is not exempt from this type of controversies: "Paul believes that stories about himself are exaggerated." or "Natasha saw a cockroach near to her.".

11. To which we should add Binding Principle B (Chomsky, 1981): A pronoun is free in its governing category. (A: An anaphor is bound within its governing category.)

12. The fact Topics have control on co-reference and pronominal deletion but not on reflexivisation will be a syntactic argument to prove their clause-external position in the sentence-discourse interface domain (the CP domain).

13. The fact that different co-referential links can take place across or inside clausal/sentential boundaries, indicates that the brain might have a mean to represent sentential boundaries in order to manage the right dependency-link rule in endogenous or exogenous co-referential assignment.

Yet, this being said, evidence from children acquisition patters and from selective-linguistic impairments, nonetheless, reveals that a difference in treatment between reflexives anaphors and other binding configurations involving pronouns is real in these populations (see Drijboom, 2010).

Some Dutch populations of agrammatic and of Wernicke's aphasics characterized by a syntactic deficit show an interesting impairment pattern in relation to binding. Ruigendijk and Avrutin (2003) investigated their comprehension of pronouns and reflexives, and found that the comprehension of pronouns was more impaired than that of reflexives. Converging results from Ruigendijk et al. (2006) show that binding rules to establish reference-relations in reflexives appear to be spared in the linguistic behavior of these syntactically impaired aphasic populations, while the process to assign a thematic-role to a resumptive pronoun is actually impaired. From the point of view of acquisition of these binding rules, without any impairment, the acquisition patterns reveal a delay in mastery for pronominal anaphoric relations in both production and comprehension. For instance, Hebrew children allow, until the age of 6, local antecedents for pronoun binding interpretation, while this represent the rule to apply for reflexives (Ruigendijk et al., 2010). The authors explain this asymmetry in acquisition of reflexive and pronoun binding -the so-called "the Delay of Principle B Effect"- following Reinhart (1983), who proposed two types of possible referential dependencies between a pronoun and its antecedent. The first '*variable binding*', governed by the binding principles we just saw, an the second '*co-reference*', governed by independent rules, that could be for instance rules relative to discourse (Reuland, 2001) [14]. This discussion will be further resorted to, when we will compare the impairment patters of empty syntactic element left by syntactic displacement and resumptives in particular populations (§2.4.3, p. 185). Meanwhile, we schematically present in Figure 2.1 (A) the indicative localization of the various complexity effects related to the establishment of anaphoric reference in the sentence as reported in various fMRI studies that have been successively replicated by Hammer and Colleagues (2007/2011) in adult normal populations.

14. See also Szterman and Friedmann (2006) for a language-impaired population (with hearing deficit) that can bind resumptives and not empty syntactic elements like *traces*. We will return to these aspects while discussing gaps and traces in next Section 2.4.3.

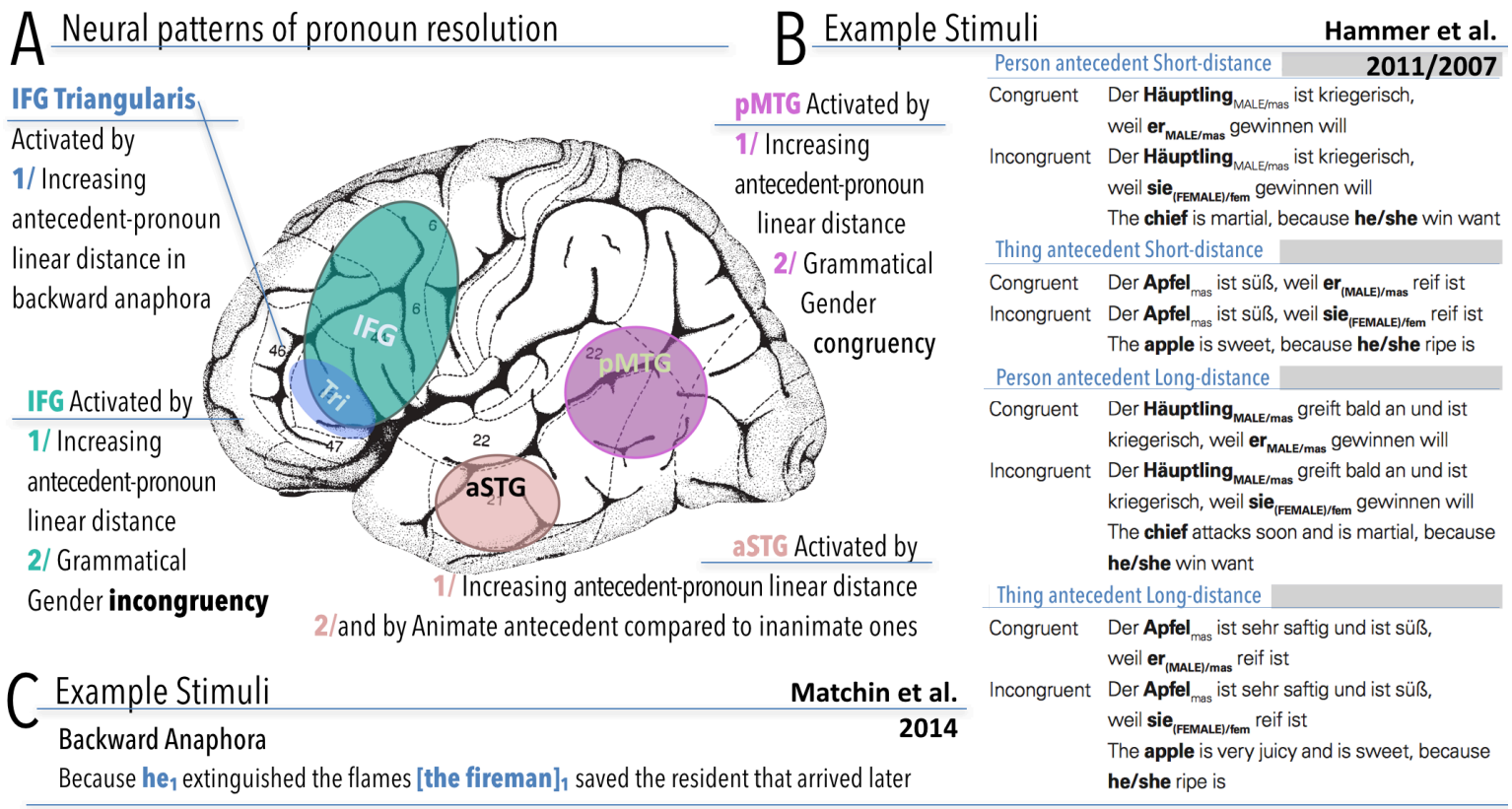


Figure 2.1 – (A) A possible mapping of the processes of semantic and grammatical information retrieval and integration during Long-distance pronoun resolution. Schematic representation of neural pattern of pronoun resolution, summarizing findings from Hammer et al. 2007/2011 and Matchin et al. 2014. (B) Example stimuli from Hammer et al. 2007/2011, manipulating semantic and grammatical features of pronoun binding. (C) Example stimuli from Matchin et al. 2014 on backward anphora. IFG: Inferior Frontal Gyrus, Tri: Pars Triangularis, aSTG: anterior Superior Temporal Gyrus, pMTG: posterior Medial Temporal Gyrus.

In Figure 2.1, fMRI results from forward anaphora (Hammer et al. 2007/2011) and backward anaphora resolution (Matchin et al. 2014) are summarized by putting forward the experimental manipulation of two main linguistic parameters involved anaphora resolution:

1. (1) the first, *linear distance*, is manipulated to try and capture the Working-Memory effects related to the maintenance of the antecedent until the co-referential relation can be established at the place where the pronoun is found in the sentence, and
2. (2) the second is linked to different semantic/grammatical congruency effects, as illustrated in the experimental design in Figure 2.1 (B).

Hammers and colleagues, despite using agrammatical stimuli[15], show a dissociation between processes linked to the incongruency between gender marking of the two elements in co-referential relation (i.e. the antecedent noun and its anaphora), and processes linked to the more semantic aspect of their animacy feature. While grammatical gender incongruity and congruity respectively evoke activation in the Broca complex (IFG, MFG and Precentral Gyrus/BA6) and in posterior Middle Temporal Gyrus (pMTG), animate antecedents compared to inanimate ones elicit a greater activation in the anterior Superior Temporal Gyrus. Thus, an effect of distance between antecedent and pronoun was observed in these four regions[16]. These finding are very interesting in that they tentatively 'localize' the distinct syntactic and semantic processes sub-serving the information 'matching' required to bind pronouns to their antecedents.

In sum, the set of sentential phenomena such that an element determines the "interpretation" of a distant (i.e. non-adjacent) pronominal element be it a reflexive, a reciprocal or a pronoun, reveals how reference-relations among sentential elements are syntactically restricted by clausal boundaries, so that this binding is not only a referential link -a semantic index-, but it has to take into account the sentence's hierarchical structural relations. The definition of the extent of syntactic domain, where different binding relations can occur, represent a crucial knowledge to understand these distal co-referential relationships in the sentence. Hence from a meuro-linguistic point of view, this knowledge -either known, or acquired without explicit training by all the native speakers of a given language- must be somehow represented and operational in language comprehension in terms of a syntactic domain on which a limited set of rules for binding apply.

2.1.2 Sentence-final utterance markers

A second element that can be brought as evidence that sentence is a natural unit is its punctuation in written language or its natural gradual decline of intonation towards the end.

However, punctuation, and in particular the introduction of full-stops, is a relative new entry in the history of writing. Before the medieval ages, Ancient Greek and principally Latin used the so-called *scriptio continua* where no typographical white space was found between words. This being said, identifying sentence-units was less of a problem,

15. A practice that only very indirectly can attest for the functional link between activity of a certain brain area and a given linguistic process.

16. The overall analysis of sentence reading data also revealed bilateral involvement inferior frontal regions interpreted as linked to integration processes, and the anterior Cingulate Gyrus whose contribution was understood as conflict resolution involved in pronoun resolution.

in that the verb could easily signal the end of the sentence, and the inflection the end of a word. The first punctuation mark introduced in the history of writing was not the full-stop, but a point to mark word boundaries, it was then substituted by typographic white, which will be generalized only in the VIIth century in Europe, and is still not adopted in Modern Chinese writing. Medieval copyists adapted the three punctuation marks of Aristophane of Byzance (257 –180), a curator of the Library of Alexandria that found three devices to facilitate the copy of manuscripts of Greek works: (1) the "perfect point" placed at the upper extremity of the last letter of a world to indicate that one could to the next line because the meaning of the sentence was completed (or *perfected* this is why it is called perfect point); (2) the 'middle point' placed at medium height an equivalent of our contemporary semi-colon; and (3) the 'under-point' placed at the bottom of words corresponding to our full-stop. Gasparin of Bergam (1370-1431), will write the first treaty on punctuation, the *Doctrina punctandi*, pointing out for the first time that punctuation's function helps breathing during reading.

2.1.2.1 Utterance markers and the end of the sentence

Although the full-stop usage is now generalized in written language, until recently various languages in the world used, and sometimes still use nowadays, sentence-final markers to signal the end of the sentence or to modify its modality, the way question marks do[17].

For example, Modern Chinese doubled its sentence-final markers to mark interrogation and exclamation with modern typographic question and exclamation marks. A similar syntactic use for total interrogative mode is observed in sentence-final particles of Austronesian languages, like Atayal a Formosan dialect using the marker -gaʔ (Saillard, 2010) to this effect, or in Chinese Hakka dialect (Gan-Hakka subgroup) where a number of markers signifying both modality and the subjectivity (i.e. mood) of the speaker are available (Chappell and Sagart, 2010).

2.1.2.2 Assertion markers

Modality is not the only function of sentence-final maker to be found cross-linguistically, namely simple assertion can be marked in many languages by a specific sentence-final morpheme.

For instance, in ancient Chinese the sentence-final particle 也 *yě* marked simple assertion (the sentence-final particle of assertion 焉 *yān* was playing the same role as *yě*)[18]. For this role modern Mandarin, has been replacing assertive markers by copula 是 *shì*, while other particles are found in sentence-final position to mark other types of enunciation like 矣 *yǐ*, indicating interjective nuances or accomplishment, like in 5 (c) [19].

17. Moreover, as sentence-final markers are strong marker of orality, considering them as evidence for the naturality of sentence-unit is a better linguistic argument than punctuation. Although we should acknowledge that the short history of punctuation reported above clearly indicates that the main motivation of its introduction has indeed to do with the marking of oral information in written modality.

18. It can also be found in sentence construction with a Topic + Noun + *yě*, in which it plays its role of assertive marker.

19. It should be noted that these final particles where already present in Ancient Chinese poetry to give rhythm (之 *zhī*, 也 *yě*, 矣 *yǐ*), and particularly the particle 兮 *xī* was signaling a place where musical background and words could mark a break and balance each other. Specifically, 也 *yě* and the particle 兮 *xī* have been analyzed as pace-holders (pinyin 音律詞) in the Book of Poems 《詩經》*Shijing* (11 –5 century BC), as shown in the following example: 允矣君子，展也大成。《詩 · 小雅 · 車攻》. or 天下之無道久矣。tiān xià zhī wú dào jiǔ yǐ [under the.sky zhi wu till long ASSERT.] 'It has been long since the

Consider the following two examples from Classical Chinese on assertion marking and sentence-final mood marker:

(5) a. 未有仁而遗其亲者也。(孟子)

wèi yǒu rén ér yí qí qīn zhě yě
NEG. there.is goodwill and neglect Poss.3SG. kin SUB. ASSERT.

'There was no man of good will that was neglect his kin.' (*Mengzi*)

b. 原公子忘之也。(孟子)

yuán gōngzǐ wàng zhī yě
wish prince forget 3SG. ASSERT.

'It would be desirable that you forget.' (*Mengzi*) in Djamouri 2010

(6) a. 我生不有命在天? (書 · 西伯戡黎)

wǒ shēng bù yǒu mìng zài tiān
1SG. like NEG. have one's-lot-in-life in heaven

'Doesn't my life has its fortune in heaven?' (Book of Documents Shu, Xībó kān lí)

b. 我生不有命在天乎? (Shǐjì 史記 · 殷本紀) 109–91 B.C.

wǒ shēng bù yǒu mìng zài tiān hū
1SG. like NEG. have one's-lot-in-life in heaven EXCL.

'Doesn't my life has its fortune in heaven?'

Examples in (5) are far from being exhaustive of the type of final particles Classical Chinese features, but following Feng Shengli (in prep.) and some interesting diachronic observation can be made about the sentences in (6), where the Spring and Autumn period (春秋 770–476 B.C.) seems to be boundary where the sentence-final mood marker 乎 hu starts to be in use as attested by the opposition between (a) and (b) and by previous diachronic analysis of mood-markers in the literature (Wang Li, 1980; Pan, 1983:168 and Zhang, 1982)[20].

An interesting spectrum of sentence-final modal particles distinguishes Contemporary Mandarin Chinese[21]:

1. 啊 *a* marks interrogation or interro-exclamatives, whose variant is 呀 *ya* ;
2. 吧 *ba*, final-particle mark imperative ;
3. 啦 *la* being the contraction of LE+A (了 + 啊) marks exclamations and reinforcement of an affirmation; as an assertion marker, it stresses the proposition it marks, expressing the joy, astonishment or anger as in examples (7d) and (7e).
4. 吗 *ma*, final-particle marks interrogative modality ;

world is virtue-less.' or 'The disorder of the world has been there for long.'

20. As for the appearance of sentence final particles Feng Shengli (in prep.) advances that tonal languages could manifest a predisposition to encode intonational inforamtion linked to mood-markers and modality int o sentence final particles. He points to several examples in tonal languages like Thai, Vietnamese and some African Language like Bada-Linda exhibiting both Tones and Final particle.

21. Importantly, following Rizzi's (1997) split-CP hypothesis, the syntactic position of the Sentence Final Particles have been mapped on the Chinese Left-Periphery and see Sybesma and Li (2007), Li B.'s PhD (2006). For a mapping of Cantonese Sentence Final Particles see Wakefield's PhD (2010). Sentence Final Particles are considered by these authors to be discourse morphemes that link the sentence to the discourse, and their syntactic position is therefore assumed to be higher than ForceP. We will resume to this point later in this chapter §3.2.3.2.

5. 呢 *ne*, final-particle marks also interrogative modality, conveys evidence in rhetorical questions and reinforces a statement[22]. This last particle has also been identified as marking Topic in Chinese an will retain our attention in the next chapter (cf. §3.2.2.3, p.298).

Interestingly, young generations communicating on social networks make an abundant use of this sentence-final markers (see statistics from GoogleCh in Yuan Zhongjun's PhD, 2016), that allow to translate emotional prosody in the written form, thus conveying speakers' attitude information. By forging neologism belonging to this functional class, social network users simultaneously avoid to use punctuation with these particles. Question marks and exclamation marks are used alone just to signify a general state of interrogation or of astonishment and surprise in the user (Yuan Zhongjun, 2016; Feng in prep.)[23].

The following examples illustrate the pragmatic richness of these modal and mood-markers. Namely, their translation will show how this type morpho-syntactic marking is expressed in English or in French thorough onomatopoeia or long periphrasis.

Consider the following sentence final mood-markers and their examples:

1. 呸 *pēi* expresses anger, as in (7a)[24]
2. 哟 *yo* in (7b) is used to express imperative tone with a hint of irony, like in French '*hein*!'
3. 呗 *bei* to signify a reluctant acknowledgment or concession as in (7c) and
4. 哦 *o* marking the fact the user just realized what he uttered as shown in the above example (7f).

(7) Mandarin Sentence-final mood markers

a. 沃尔沃对金属漆要另外收费？呸！

wòěrwò duì jīnshǔqī yào lìngwài shōufèi ？ *pēi*
Volvo for metallic-paint wants additional charge ? PART*pei*

'Volvo also charges extra for metallic paint ? bah!'

b. 谷歌能有这样的领导，发表这样的言论，不一般哟！

gǔgē néng yǒu zhè yàng de lǐngdǎo, fābiǎo zhè yàng de yánlùn, bù yībān *yō*!
Google can have this type of/DE CEO, express this kind of/DE discourse, not banal PART*yo*

'The fact Google has such a CEO making such public speeches. It really makes the difference!'

c. 你关了中国的，我就用美国的呗

22. This elusive presentation and this characterization out of context of Mandarin sentence-final particles is mainly due to lack of space this topic is too vast to be reviewed here, for extensive discussion on the semantic and pragmatic properties of final particles and the syntax of Mandarin sentence-final particles see Li B.'s PhD (2006)

23. Interestingly, this pattern of usage is opposite to the one attested in French web and text-message linguistic interactions, see Véronis and Guimier de Neef (2006).

24. Note that some of these particles like *pei* 呸 have also a self-standing use as exclamation particles (free morpheme) that are found at sentence-initial position, and when they are used as such they carry sometimes a different meaning. In the case of 呸 *pēi* its meaningful shifts towards disgust like in English *pah*!, *bah*! or *pooh*! expression or the French word '*beurk*!'. Or the modal particle for surprise 啊 *a* changes tone when used as an exclamatory particle to signify 'what?', shifting from first to third tone.

nǐ guān le zhōngguó de, wǒ jiù yòng měiguó de bei.
you close LE ChineseREL., I then use American REL. PARTbei

'If you close the Chinese (Google), then, of course, I'll (be obliged to) use the American one.'

d. 百度可开心啦

Bǎidù, kě kāixīn *la*
Baidu, veryEMPH. happy PART*la*

'Baidu, what a joy!'

e. 你是去是留，无所谓啦

nǐ shì qù shì liú wúsuǒwèi la
you are go are leave whatever PART*la*

'Stay or leave, it has anyway no relevance!'

f. 它有政治目的的哦。

tā yǒu zhèngzhì mùdì de *o5*
it has political aim DE PART*o*

'Ah, I just realized, he has (his own) political aim behind!'

Comparing some of the Chinese counterpart to English exclamation particles we can observe a distributional asymmetry, while English exclamative particle are sentence initial Chinese ones tend to be sentence final, as shown in Table (2.1) below[25]:

Table 2.1 – Parallel between English sentence-initial exclamative particles and sentence-final Chinese counterparts, English part adapted, Chinese part elicited with our informers..

Particle	Approximatve glose	English use	Conversational use	Chinese counterpart
aha a-ha	'understand'	"*Aha*! So you took the money!"	acknowledgment	原来你拿钱啦
wow	'amazing'	"*Wow*, that's incredible!" (to be off the charts)	astonishment	没治了哎!
boo booh	'That's bad'	"*Boo*, get off the stage!"	disappointment	上来啵

Assertive markers can be found also in the simplest sentence-unit, the so-called minor-sentence type, in which a non-verbal element is the predicate. For instance, some languages distinguish between copulas from so-called *predicators* both marked in bold in the examples reported in (8) form Bambara (Shopen, 1985:170) one of the official languages of Mali. As shown in (a) the Predicator is used to mark a nominal in a one word utterance when there is no overt subject, it function is then to mark that this Noun Phrase is a predicate. In (b) we see indeed that predicator *don* is distinguished from the copulative verb *ye*:

25. We will resume to the issue of interrogative intonation in section §2.4.4.2 by considering the case of French simple yes/no questions that are analyzed theorizing the existence of functional morpheme (Cheng and Rooryck, 2000) for pure interrogative intonation in French.

(8) from Bambara

a. Alamisadon **don**
Thursday PREDICATOR
'It's Thursday'

b. Bi ye Alamisadon **ye**
today PRES Thursday **be**
'Today is Thursday'

2.1.2.3 The peculiarity of Sentence-final position

From a psycho-linguistic point of view, an interesting study on Mandarin Chinese by Chu (1998), shows how punctuating a text reveals that sentence-final utterance marker are systematically taken as functional markers for sentence ending. Chu's experiment asked to punctuate a paragraph to 91 Chinese native speakers, and the author found that the majority of the full stops located by native speakers corresponded to a series of rules including not only some overt linguistic "cues" like sentence-final particles, conjunctions, adverbs, verbal affixes, but also Zero-Anaphoras.

Importantly, the position of this kind of assertion markers indicate they are scoping over the whole proposition, and for this reason they have to be opposed to others morphemes directly linked to the verb that are present in languages like Afar (Afro-Asiatic, Cushitic[26]). This language posses in facts predicate assertive marker that are attached to the verb and are not sentence final,like *-éh*, that allow to oppose (a) from (b) in the following example:

(9) from Bambara

a. ka1 ablé**-h**
him see-ASSERT
'I see him [I affirm I see him]'

b. ka1 ablé
him see
'I see him [him, notanother]' (Morin, 2010)

Some languages even present syntactic means to express negation in sentence-final position, with the use of a specialized morpheme expressing negative existential value like in Gula *tɔʔ'* final particle (Group bongo-Bagirmi, Kara branch in Central Africa Tchad Soudan), scoping of the whole sentence's propositional content.

2.1.2.4 Sentence-final particles: mood and modality

From these cross-linguistic evidence, we can conclude that sentence-final particle are to be analyzed as functional morphemes that apply to the whole sentence. Being present at the frontier between utterances, they are called utterance markers, which constitute a grammatical word class on both functional and distributional grounds. They are divided into two main groups: (a) *modal markers*, modifying the sentence type (i.e. assertion, interrogation, etc.) and (b) *mood markers* or speaker's attitude markers. Hence, they have the function of syntactically encode information about the illocutory value of the

26. Specifically, from the brunch of Lowland East Cushitic, spoken by the Afar people in Djibouti, Eritrea and Ethiopia.

sentence, or to encode the point of view of the speaker on the proposition they appear in[27].

Taking another example of sentence-final particles from Dagara (Burkina Faso), we see that these both carry both standard interrogative function and information about attitude of the speaker in an inseparable fashion. The most neural interrogation mark being *bu* which is far from being the most frequent. While the following type of discursive markers in French or English are marked by adverbs, full expressions and locutions in Dangara are integrated din the grammatical system (Delplanque, 2012:9):

(i) He came + **bii** ? *'He come, or not?'* ;
(ii) He came + **v** ? *'He come, tell-me?'*, *phatic* value with a perceptible lengthening of the vowel ;
(iii) He came + **jàà** ? *'He come, really?'*, expressing reticence ;
(iv) He came + **ke'**? *'He come, thus?'* [If I understood well.] rhetoric interrogation like adding '*in the end*' in English.

Hence, we collect further evidence in support of the fact that there exist a universal and natural sentence-level linguistic unit from the important number of languages in the world that posses sentence-final utterance particles which cover more than the basic interrogative mode function:

1. Basque (Euskara) *omen*: testimonial particle (Oyharçabal, 2010:857);
2. Mandchou (Xibo) *dere*: doubt particle (Beffa, 2010:964);
3. Mongol has interrogative, exclamation particles, and others for the expression of doubt or confirmation (Beffa, 2010:957);
4. Northen Same *dus*: interrogative particle (Fernandez-Vest 2010:820);
5. Mayangna *dai*: for temporal making of past perfect (Benedicto, 2010:1434);

We start here to see the emergence of elements from the discourse-level entering in the sentence. This is gradually unveiling one of our central concern in the neuro-linguistic investigation of the sentence: the sentence discourse-interface, namely the possibility that discourse-level information appears inside as shown in Figure 2.2.

In fact, the mood markers we saw in Chinese are far from being isolated linguistic phenomena, Kwa languages (Niger-Congo family) posses an rich grammatical class of this type of utterance particles marking illocutory force or attitude (Ameka, 2010/1998).

Among those, we report in Figure 2.2 a rich inventory of sentence-final utterance particles found in story-telling in the Ewe (or evhe) dialects of the *gbe* group, to give the non-linguist reader an order of magnitude of all the discursive functions that can be grammaticalized in a language.

2.1.2.5 Sentence-final particles: syntacticizing their semantic and pragmatic richness

The semantic/pragmatic richness of the grammatical coding of subjectivity observed in this class of utterance particles across languages, can reach the syntactic expression of '*frustrative declarative*' thanks to the sentence-final particle *etaop* (Galúcio, 2010), as exemplified by the following sentence in Mekéns (Sakïrabiát, Tuparí), a nearly extinct Tupian-language in the amazonian region of actual Brazil[28]:

27. We will see later in Section 2.4.4 how these apparently distant linguistic functions can be actually represented the same sentence Domain in the realm of the sentence functional skeleton than we will adopt as a measure for sentence syntactic complexity.
28. Note that Tupian is the same branch of Munduruku Languages.

(10) Mekéns

~~isii~~ o-so-a kwat õt i-aka etaop
deer ASG.see.VT exit 1SG.pron MO.follow frustrative.part.
The deer saw me and escaped, I followed him but couldn't get it. in Galúcio, 2010.

What we actually want here to underline is how central and cross-linguistically highly exploited is the discourse interface of the sentence. This point introduces one of the central claims of this work: sentence as a unit can be explored from the point of view of its interface with the discourse, by experimentally addressing the ways in which languages syntactically express this interface where actually syntax, semantics, and pragmatics interact in a complex manner.

There also exist some extreme cases of languages like Standard Thai (*Siamois*) that mark every sentence with a least one final-particle (Gsell, 2010), comprising modal ones determining interrogation (*mǎj*), disjunctive interrogation (*ruǐ:*), mildend order (*rʏ́ʔ*), question tags (*sîʔ sîʔ*) or locution like 'and so' (*náʔ*); or politeness one that distinguish the gender of the speaker, like assertive particles (m.*nîa*) and (f. *kʰráp*), question particle (m. *kʰàʔ*) and (f. *kʰáʔ*)[29].

Interestingly, the kind of discursive function illustrated by these sentence final-particles in Ewe and by the rich inventory we presented, is in sharp contrast with languages that do not encode these discursive functions (mood and modality) syntactically, and that have to express the semantic equivalents in a wide variety of ways, and mainly by intonation, word-order or explicit attitudinal expressions like '*I hope*', '*do you suppose*', 'please' as seen in the translation of examples in (12). The length of our translations and the need of long periphrasis is the demonstration of how the *syntacticization* of the large spectrum of discursive, attitudinal, deferential, politeness markers is a highly efficient procedure to convey rich semantic/pragmatic meaning in one syllable having a definite syntactic role in the sentence[30].

Hence, the ultimate aim of the following rich inventory of languages and sentence final-markers is to convince the reader that leaving outside of the domain of research on language faculty, and more specifically out of neuro-syntactic research such an incredibly sophisticated and efficient set of syntactic phenomena would be a pity[31].

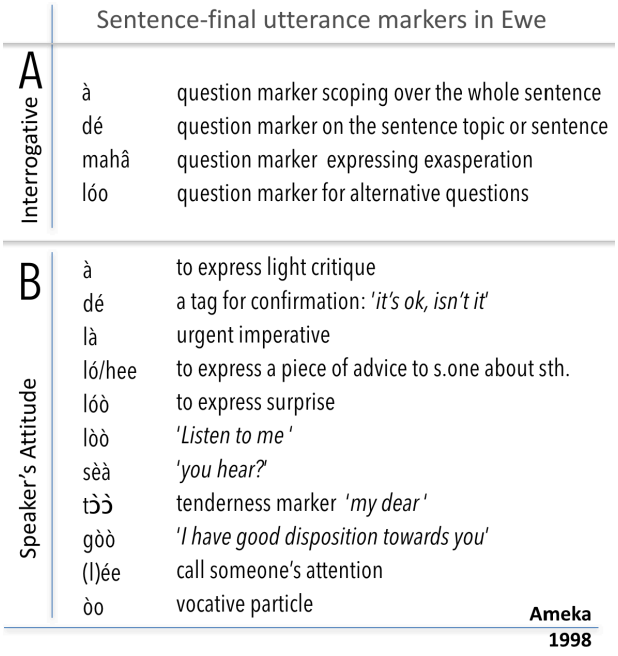


Figure 2.2 – The particles reported by Ameka (1998, p. 183) are naturally used in Ewe story-telling to mark the two classes of particles, in (A) interrogative function and in (B) attitude in the interaction/address mode between participants of a particular communicative situation.

29. This particles' grammatical class is in fact as complex as the pronominal one in Thai (Cooke, 1989).
30. Note that several works inscribed in the Cartographic Project have addressed the issue of sentential particles and their functional projections inside fine structure of the Left-Periphery of the sentential structural skeleton. For a study on sentence-final particles Italian dialects, like Venetian, see Poletto and Zanuttini (2011).
31. We will later show how simple word-order information without even the need of morpho-syntactic

2.1.2.6 Sentence-final particles: A link with Topic-Comment?

Continuing in this direction, some examples from Japanese and Tagalog (Kuno, 1973) illustrate how these mood particles are frequently associated to Topic markers[32]. Interestingly, the functional characteristic of the mood markers in (11) and (12) is to modify what would be a neutral form of utterance and to enrich it with mood/attitude interpretation, by doing this they establish an interface with discourse, that is here complemented by the presence of a Topic-Comment articulation in Japanese and in Tagalog, marked by Topic particles.

By this observation, we anticipate one of the central topic of this research project, Topic-Comment construction that will be thoroughly introduced in next chapter. Examples in (11) and (12) clearly use the final particle to mark the end of the Comment clause.

(11) Japanese

a. Kore wa hon desu **yo**.
this TOP book is STATEMENT
'[I am telling you that] This is a book.'

b. Kore wa hon desu **ka**?
this TOP book is Q
'Is this a book?'

c. John wa baka **sa**.
John TOP foolish EVIDENTIAL STATEMENT
'[It goes without saying that] John is a fool.'(Kuno, 1973)

(12) from Tagalog

a. Mabuti *a* ang ani?
good Q TOP harvest
'Is the harvest good?'

b. Mabuti *kaya* ang ani?
good Q.SPECULATIVE TOP harvest
'Do you suppose the harvest will be good?'

c. Mabuti sana ang ani
good wish TOP harvest
'I hope the harvest is good.' (Kuno, 1973)

In conclusion, the existence of such a crowded population of sentence-final markers can be taken a linguistic internal evidence for sentence-as-a-unit:

1. (1) Firstly their function at the interface with discourse helps in the definition of the sentence-unit in that utterance markers are to be found at the frontier between utterances. In other words the interface helps us to define the sentence.
2. (2) Secondly, the fact that these makers scope over the whole sentence, on one side, to add modality to it and creating "sentence types", and on the other side to modify the pragmatic interpretation of the sentence by adding information about the subjectivity of the speaker.

marking is another highly efficient manner to convey the sentence articulation with discourse in Mandarin Chinese, but also in many other languages.

32. Another important discursive role we will present in the next section 2.2.4.

2.1.2.7 Some psycho-linguistic properties/aspects of sentence-final position

One could link the cross-linguistic pervasiveness of utterance final-markers with psychological studies that explored the processes happening at sentence final position. In the 70s, in the attempt to give experimental evidence to the hypothesis that the proposition is the unit of recording (i.e. memorization) during speech perception (Bever 1970a; Fodor, Bever, and Garrett 1974) sentence-final position was unanimously attributed a special status. Psycho-linguistic studies supporting the so-called '*recording hypothesis*' (i.e. memorization) revealed that the end of each proposition was the locus of the memory fixation of the sentence's memory recording during sentence listening. Their results show that at sentence-final position reaction-times to played click sounds are slow (Abrams and Bever, 1969)[33]. This decrease in attentional capacity of detecting click sounds was taken as the mark of mental load associated with the final stage of sentence processing, and was interpreted as a memory recording step of the proposition into a deep representation (Bever, Garrett, and Hurtig, 1973).

To these findings should be added results from other studies on sentence comprehension, which focus on a broader level of psycho-linguistic analysis (see Johnson-Laird, 1974 among others). Their overall contribution puts forward that sentence-unit is a naturally *answerable* and *verifiable* unit, which means from an intuitive point of view that any individual can decide if a sentence is true, just by verifying that its propositional content is true. More recent investigation of sentence final-psycho-linguistic processes in Eye-movement studies on reading (Bonhage et al., 2005)[34], and in online-ERP studies of sentence processing observed a so-called wrap-up effect at the end of each sentence (van Berkum et al. 2005).

Leaving a host of details aside, the linguistic and psycho-linguistic elements presented until here suggests that the sentence is a unit:

1. A syntactic unit, where only certain co-referential relations can be internally established (§2.1.1) ;
2. A natural orally-marked unit of the linguistic system, possessing a sentence-final position that is cross-linguistic used for markers scoping over it as a whole (§2.1.2) ;
3. A natural level of memorization, in that sentence-final position is a place where particular psycho-linguistic processes are observed (e.g. memorization, wrap-up effects).
4. Next sub-section §2.1.2 will add to this frame additional psycho-linguistic evidence for considering the sentence as a unit of perception.

We hope we could convince the reader that focusing on the sentence interface with discourse is a promising and new way to look at the sentence-unit. Investigating sentence interfacial linguistic phenomena can, in fact, reveal a wider spectrum of syntactic configurations than the usually studied syntactic phenomena in psycho-linguistics or neuro-linguistics. Very interesting issues may arise, right at the heart of questions of how syntax, semantics, and pragmatics interact. This point will be approached in this

33. Note that their actual detection has been reported to be poor (Bever, Hurtig, and Handel, 1975). Even tones stimuli played during sentence comprehension are hard to discriminate at sentence-final position (Holmes and Forster, 1972)

34. Eye-movement and fMRI study on predictive anticipation of sentence final words, where the final target word was displayed with a temporal delay and its screen position was dependent on the syntactic word category (nouns vs verbs). During the delay, statistically significant anticipatory eye-movements (measured by selection time and response time) into the correct target word area were observed and interpreted as indicative of linguistic predictions.

research work thanks to Mandarin and French and through the prism of the cartographic approach presented in chapter 1 (§1.2.4, p. 31)[35].

In our experimental approach to French Question formation and Mandarin sentence-discourse interface we will focus on the two linguistic functions sub-served by this kind of sentence-final markers, namely modality (for French questions) and illocutory value of the sentence in Mandarin Topic-comment sentences and Focalization constructions. More specifically, the ERP-experiment presented in chapter 5 will evaluate on-line brain responses to sentences in context in order to track the point of the sentence structure where context informational load is accessed during on-line sentence comprehension. In this way we will characterize the sentence-discourse interface of Mandarin Topic-Comment articulations and of French question formation from a neuro-linguistic point of view.

2.1.3 The sentence-level defines the fundamental object of language perception

At the turn of the second millennial, a book opens up the heyday of the ***Sentence as a Cognitive Object***. Thanks to Townsend and Bever's book '*sentence comprehension*', the sentence became *a Case Study in Cognitive Science*[36]. The authors develop a number of arguments and present a handful of experimental evidences to argue for treating sentences as a natural linguistic representation (in opposition to a word-level representation).

We will give a brief review of experimentally grounded arguments indicating that the sentential linguistic level has an independent representation during comprehension.

Pioneering experiments in psycho-linguistics were conducted in the 1950's and 1960's, mostly due to the initiative of George Miller and his colleagues, who wanted to determine the behavioral relevance of the sentence-unit. At that time, the question about the perceptual pertinence of sentence was formulated as the search for the unit or level during speech perception where the *acoustic-to-linguistic* transfer was actually taking place.

In a series of studies Miller and colleagues used a simple experimental setting that was based on an ordinary observation: spoken language is extremely resistant to noise and other types acoustic interference. Adjusting the noise of recording of independent words to a level where they could be correctly recognized 50 % of the time in isolation, they made subjects listen to two words sentences composed of the independently recorded words, so that sentences were either correct like (a) "*Horses eat.*" or incorrect (b) "*Horses cry.*" (Miller, 1951). And, the following hypothesis was formulated by the authors: if in this experimental setting the acoustic shape of each word is mapped independently onto a linguistic representation, a sentence composed by stringing together the words that are interdependently understood half of the time should be then understood only 25 % of the time. Results showed that when words are in a sentence, they are recognized much more than 50 % of the time. Moreover, even syntactically well-formed sentences that do not make semantic sense, like '*Horses cry*', enhance the perception of words items. As a humorist side note, could we, then, interpret this result as natural propensity for poetry?[37].

35. Hypothesizing that there exist a Universal order in the functional architecture of the clause, and more specifically in what regards the domain of the sentence that is at the interface with discourse. See §2.4.4 (p. 197.

36. Townsend and Bever, (2001, chapter I) Sentence comprehension. Cambridge, MA: MIT Press.

37. cf. Poetry is a syntactic challenge and not a semantic one, as we said in the Introduction to this

To this experimental evidence one should add an intuitive fact that was further investigated by Miller and Isard (1963): when words are in a sentence they seem acoustically clearer. These authors experimentally examined the memorization of lists of words, showing that the perceptions of words were more resistant to interference and that words were more easily recalled when they were appearing in correct sentences, compared to the same words in lists, or in ungrammatical sentences.

Given these findings, the *word-level*[38] appears to be the one at which the *acoustic* information is mapped onto linguistic representations, and the *sentence-level* appears to be the unit at which "*speech* perception" is happening. Hence, we can end this section with Townsend and Bever's comment (p.10–11; 26–27) on these experimental proofs saying that:

> "Through all this, an essential psycho-physical feature of sentences remains true: words are especially behaviorally compelling when they are arranged in sentences".

2.2 The Sentence is a universal unit across languages

As we saw in previous section, sentences are more than the words that compose them, both from the point of view of their sentence-level meaning interpretation, and from a more structural point of view. For these two claims, two main evidence were brought to the readers' attention: (1) the sentence sequence appears to be a domain inside which certain co-reference rules are to be applied, (2) utterance markers (both modal and discourse-semantics ones) scope of it entirely and (3) sentence-level and not word-level is the unit of speech perception. We turn here to an analysis of the universal characteristics that distinguish sentence-unit and make it a cross-linguistically universal unit.

2.2.1 Word groupings

Moving from the first step remark of Wundt that "*the natural unit of linguistic knowledge is the intuition that a sentence is a sequence*", we will now show that these sequences have universal structural properties: (1) sentences universally have constituents allowing word-groupings, and, (2) sentences universally present grammatically motivated sequential orderings of their internal elements (i.e. word-order constraints and patterns).

These two points will lead us to consider sentence's internal organization both at the lower level of constituents and at the sentence-level (i.e. word-order constraints), highlighting a series of linguistic facts that point to some crucial universal aspects of sentence-internal structural relations. Psycho-linguistic considerations about constituency and prosodical grouping will be giving initial arguments for considering sentence as more than a linear sequence, while the hierarchical aspect of constituency structures will deepen in the next section (§2.3, dedicated to hierarchies in the structure of the sentence and to the universal format offered by X-bar theory, §2.3.1.1).

Let us start by observing the following examples from a famous jazz standard by Cole Porter, marvelously interpreted first by Billie Holliday in 1946, and later popularized by Frank Sinatra:

manuscript (see p.xxxiv).

38. Comprising also phonemes and syllable decoding.

(13) a. What Is This Thing Called Love?
b. What Is This Thing Called, Love?
c. What, Is This Thing Called Love?

The three sentences in (13) shows how the exact same sequences of words can have different meanings accordingly to the grouping of the words that is assigned by oral phrasing and pausing signified by the punctuation marks at different places in the sentence: a manipulation that Billie Holiday actually perform in her singing interpretation. This not only shows that spoken sentences carry much more information that written ones, but that intonational cues (e.g. prosodic pausing) carry a type of information about the syntactic grouping of words, which greatly impact sentence's interpretation. Prosodic phrase-structure information is shown here to perform the same function as syntactic information. We will experimentally address this syntactic function in chapters 4 and 5 by observing behavioral and electro-physiological responses to sentences where this type of prosodic boundaries have been removed, and the listeners can only rely on word-order cues to build and understand the syntactic structure of the experimental sentences.

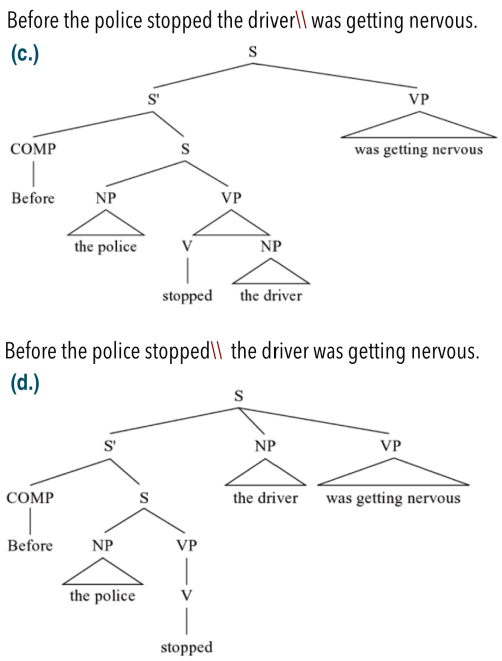


Figure 2.3 – Syntactic-tree diagrams representing the different syntactic chunking yielded by the position of the line break (i.e. \\) of sentences (c) and (d) in the experiment by Kennedy et al. (1989).

This prosodic chuncking of the utterance is universally required to understand a sentence, and its essential role has been confirmed by experimental psychology. Among the numerous experiments bearing on this central issue, we can cite Kennedy and colleges (1989)[39], who investigated how line breaks that separate words belonging to the same constituent hinder reading by augmenting reading times compared to line breaks that are coherent with the grouping of sentence's primary constituents. Compare the following first two examples with the last two in (14):

(14) a. When I've finished running
the class can we all go home?
b. When I've finished running the class
can we all go home?
c. Before the police stopped the driver
was getting nervous.
d. Before the police stopped
the driver was getting nervous.

While (a) and (b) show how line breaks can ease or not sentence parsing, while (c) and (d) show how line brakes, like prosody can cue for a certain syntactic configuration of the sentence. Anticipating section §2.3, we show in Figure 2.3 how representing sentence in syntactic-trees can help grasping the syntactic differences between pairs of sequentially identical sentence, like (c) and (d).

39. It could be particularly informative of why the reader sometimes has troubles reading some of my sentences.

2.2.1.1 The psycho-linguistics of constituents

One last empirical results confirming the organization of the sentence-unit into constituents comes form a pycho-linguistic paradigm based on a click detection task happening during sentence listening. Participant are invited to listen to sentences in one headphone, while clicks are broadcast in the second headphone, they are then asked to report the exact location of the click afterwards. In this kind of tasks, errors in locating the clicks are hypothesized to reflect ongoing processes during sentence understanding, particularly at the time point in the sentence where the click occurs. The several studies that investigated word grouping using this paradigm reported that participants systematically mislocated clicks towards main constituents boundaries, while they objectively had occurred before or after (Abrams and Bever, 1969; Bever, Lackner, and Kirk, 1969). Consider for example the click occurrence in "*That he was (click) happy was evident from the way he smiled.*": its position was systematically reported after the end of the verb phrase. Relying on these observations, Fodor and Bever (1965) proposed that immediate surface phrase structure was assigned directly during comprehension, advocating for the existence of an early stage of access to major phrase segmentation during sentence online processing. These studies and their methodology were long debated, and the methodology was improved in the 70's by adding both windows to report the clicks, and some experimental control sentences where clicks were absent. And finally, the results remained the same (Bever and Garret, 1973 ; Bever and Fodor, 1966): clicks' position was reported as being shifted towards main constituents' boundaries, showing in this way that constituents could be considered as processing units.

This universal aspect of sentence *chunking* into constituents was observed in a recently published neuro-imaging study by Matchin and collaborators (2017) that largely replicated the study by Pallier et al. (2011) (presented in chapter 1, page 66)[40]. Interestingly, these authors give a new alternative interpretation to processes involved in constituent-structure hierarchical relations needed to process a sentence. They interpret (1) syntactic structure building processes as being distinct from a (2) syntactic prediction ones, which are taking place at the level of the sentence. Their experimental design tries in fact to isolate these two processes by manipulating them at two different syntactic levels. They identify the syntactic combinationatorial and building processes as taking place at the level of two-words constituents (i.e. Noun Phrases and Verb phrases like '*the baby*', '*their bill*')[41], and syntactic prediction mechanisms mainly at the sentence-level (e.g. "*the poet will recite a verse*")[42].

Their results show significant activation of IFG (pars triangularis and pars orbitalis) and pSTS uniquely for sentence stimuli when contrasted with both unstructured lists and two-word phrases. This response pattern is interpreted by the authors as showing that Broca complex is here sub-serving an active syntactic prediction process. On the other

40. As a reminder this study showed that cerebral activation in certain areas correlates with the number of constituent that can be built in a sentence, a measure the authors called 'constituent-size'.

41. Note that syntactic combinationatorial and building processes are needed both at the level of constituents and sentence.

42. The authors report as handful of precise and detailed linguistic manipulations that were implemented in order to constrain the morpho-syntactic building operations at sentence-level. Namely, same simple, active sentence structure was selected. Subject and object were always consisted of a simple Determiner-noun sequence. And, "in order to reduce overt morphological complexity that might induce structure- building operations, the content words (nouns and verbs) in the natural condition bore no overt inflectional morphology (i.e., all singular nouns, no tense/agreement on the verb)" (*ibid* p. 109).

hand, two-words phrases showed a retained lenient activation of the Anterior Temporal region (aSTS and TP) and of Pars orbitalis. Only when compared to 'delexicalized' Jabberwoky two-words phrases, like '*our speet*' or'*their subex*', a significant cluster was reported in MTG, suggesting that constituent building process might involve a more Temporal distribution.

This emphasis on prediction processes in sentence processing is a new view angle on syntactic faculty that is growingly supported by a number of studies (see Lau et al. (2006), who showed that ELAN ERP-component amplitude elicited by word category violation was modulated by the strength in expectation for a particular word category in a 'constituent' ; Kamide et al. (2003) for eye-movement evidence for the so-called *top-down* expectations during incremental sentence processing).

2.2.2 Basic Sentences and word-order across languages

Now, that we established that the sentence can be defined as a sequence of constituents, our interrogation can move to what are the universal characteristics of *sequences of constituents* across-languages. Answers come from different linguistic back-grounds that both seek to characterize the language system as an articulation of universal and peculiar elements, beyond language diversity:

1. On one hand, the work of Typology has revealed that sentences are universally showing word-order constraints of the most basic constituents and elements of the sentence. This order is generically understood as the order of subject (S), object (O) and verb (V) in a typical declarative sentence. This approach also describes the grammatical patterns a given word-order yields on the ordering constrains of other functional elements of the grammar of a language. We already had an overview of the regularities that have been discovered through the meticulous field work and analysis of the world's languages by typologists in chapter 1. The fact that the languages of the world universally show language-specific basic word-orders correlating with a number of grammatical aspects of their grammar, will have important consequences on psycho-linguistics investigation of language processing,as we will see.
2. On the other hand, the answer from Generative linguistics has been be to argue that these basic word-order properties shown in natural languages point to an economy-based design of human syntax. In the realm of Government and Binding theory, this argument has been developed by formulating two types of rules that are to be found across languages: Principles and Parameters (cf. chap. 1, p.31). Principles are meant to capture what is universal in language and are therefore invariably shared across languages, while Parameters characterize the space of possible variation in a linguistic system, consequently parametric choices of a language can be seen as the formalization of the variations patterns found in comparative studies.

Actually, what we will try to adopt here is a unifying approach leveraging the descriptive power of both linguistic traditions to obtain a complete and *un-fragmented gaze*[43] on the sentence-unit and its rules.

This comprehensive understanding of the sentence unit has revealed essential to sought for the determinants of brain activity during sentence understanding. The data-driven perspective of Typology, built from the observation of a huge amount of languages offered some descriptive tools for observable surface properties of syntactic units in their

43. cf. fragmentary vision of the world in post-modern Times in the Epilogue §III, p. 707

linearity dimension. And, by carrying out a study of the grammar of languages at a higher level of abstraction, the formal approach offered tools to describe some possible underlying universal mechanisms linked to the fundamental *hierarchical* nature of sentence-unit organization. We argue here, that these two approaches united allow a mapping of syntactic ability and knowledge on cerebral activity, which is keeping at the center the two fundamental aspects of sentence-unit -its linearity and its hierarchy- that actually constitute the two central problems the mind or brain has to solve to understand a sentence[44].

Namely, from these arguments and observations we can move a step further and approach the sentence as a linguistic unit in which ordering constraints on the main constitutive parts are specified in each language particularly for basic sentences, and consider a few Principles that can universally apply to this linguistic unit. Given our focus on what happens in the mind and the brain of the listener, we will add to these linguistic outlook on what is universal in a sentence, another point of view on linguistic phenomena, that of psycho-linguistics, which will contribute to confirm and reveal many universal aspects characterizing the sentence as a unit.

Children's use of canonical word-order structures

First, psycho-linguistic evidence on this point comes from children use of canonical word-order structures. It is generally acknowledge that children acquire the schema of the canonical sentence in their mother-tongue at early stages of the acquisition of syntax (Slobin and Bever, 1982), and this is cross-linguistically true.

Studies in languages where the canonical sentence structure is not always syntactically simple have revealed that children have no difficulties if they were exposed the complex syntactic construction that is central to the grammar of their mother-tough and therefore produced frequently in child directed speech. An example comes from Sesotho (southern Bantu language) that has a complex inflectional system where the verbal complex not only encodes subject/object marking, tense/aspect, but also grammatical functions such as passive, applicative, causative, and reflexive (Demuth, 1992).

Sesotho children are capable of comprehending and producing passives as early as age 2;8 (Demuth, 1990), while English children learn it at an average age of 4 (cf. chapter 1, p. 57). Compared to English learners, they show spontaneous and creative production of passives, relative clauses, and left-dislocated constructions before the age of three, structures that are generally attested to have later acquisition in Indo-European languages. These results provide evidence for considering that the basicness of word-order and of canonical sentence structure can be viewed as language-specific. We will come back on this issue, addressing the issue of acquisition patterns of Topic-Comment articulation in Chinese children in next chapter (§3). In the meanwhile we can cite a longitudinal study by Erbaugh's (1992) on four Chinese children (1;10 to 3;10), where strict adherence

44. This mixed approach can also considered as linked to Backer (2010) formulation of a 'Formal Generative Typology' when this author puts forward that typological variation can be dealt with a reasonable success when it is carried on at a more abstract level within a formal system of description. In other words, he advocates for a typological approach taking advantage of the tools of formal grammar, and seeking to answer the following questions: (i) what properties of natural languages are universal an inherently human?, (ii) what properties of natural languages vary cross-linguistically? and (iii) conversely what are the syntactic aspects of this cross-linguistic variation that are systematic and patterned ? This position reflect partly the spirit of the Principle and Parameters period of Generative grammar where deep concern for the comparative dimension of linguistic variation was developed.

to canonical SVO word order was observed (Erbaugh, 1992:416). The author actually suggest that Chinese children's word order appears even more conservative than that of adults'.

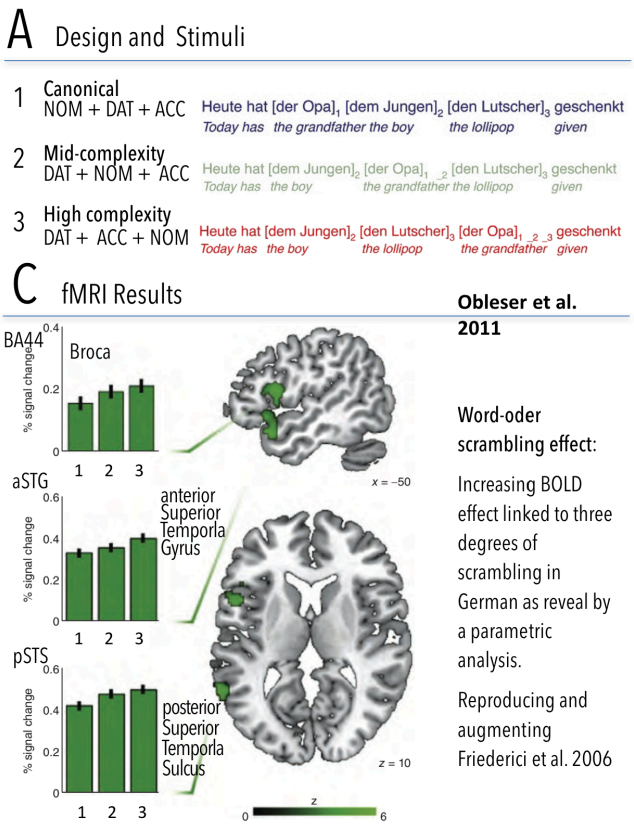


Figure 2.4 – (C) Barplots reporting signal percentage change in brain regions where activity increases with word-order scrambling in German sentences like in (A) from Canonical in (1) to High Complexity in (3) Adapted from Obleser et al. (2011).

She advances that this phenomena is due to their processing capacities and a general strategy of high consistency with word-order rules (Erbaugh, 1992:416-417) in language lacking morphological markers to mark agreement, number, gender or case, and where word-order is consequently considered as the most important syntactic device in Chinese for sentence interpretation (cf. Chang, 1992)[45].

All in all, such results concur with Slobin and Bever (1982) in affirming that canonical word-order schemes are accessible and plays a crucial role in children's early sentence comprehension and production across different languages.

Neuro-imaging of word-order variations

The linguistic and psycho-linguistic evidence for the existence of basic sentence word-order structure is further supported by neuro-imaging studies on word-order scrambling, showing how cerebral activation is proportionally augmenting with the degree of word-order scrambling from initial canonical word-order (Friederichi et al., 2006)[46]. Figure 2.4 illustrate the results of an fMRI study on German by Obleser et al. (2011) that found evidence of how departing from the canonical word-order of a given language elicits complexity effect at the brain level. Figure 2.4 reproduced in (A) the three different level of word-order scrambling[47]. These results are in line with a number of studies on German[48] investigating the processing of this kind of linearization of linguistic dependencies found in free word-order languages that we will

45. As we will see in section §3.2.4 this is could constitute an explanation for the fact that in Chinese children non-canonical word-orders are not as productive as hypothesized by topic-prominence claim, or at least relatively scarce in the early stages of syntax acquisition, although they are available and common in Chinese adult speech.

46. Note that in Friederichi et al. (2006) the complexity effect correlated with the decrease in acceptability of the stimulus sentences

47. Scrambling is a syntactic phenomenon that is observable in non-canonical ordering in free-word order languages like Japanese and German. Not that it involves the same filler-gap dependencies that syntactic movement generates, but it obeys rather different syntactic constraints. We will present movement-related syntactic complexity effects further on, in section 2.4.

48. See Grewe et al., 2005 for rule-governed scrambling of pronominal and non-pronominal objects ; Fiebach et al., 2005 for German scrambling in wh-questions ; Wartenburger et al., 2004 that found only Wernicke area's (pSTS) activations).

present shortly because they show how a certain word-order rules are to be linked to different types of prominence hierarchies among sentence-internal functions.

2.2.2.1 Thematic hierarchy and universal basic Word-Order

Underlying most syntactic accounts of the syntactic complexity is the intuition that somehow the clause has a basic word-order. However, this basic word-order has not only been described as the relative order of Subject, Verb and Object and all the correlated patterns that are studied in Typology, a second important word-order linearization principle has been formulated. A series of thematic hierarchy principles where semantic information types for the different participant roles to the event structure of the verb have been identified as contributing to syntactic complexity of the sentence-unit (Jackendoff, 1972/2002).

Consider for example the very general rule that thematic subjects like *agents* are higher ranked than direct objects (themes) an thus precede them in the linear order of the sentence. The main idea of these thematic rankings is namely to correlate the relevant semantic aspects of the under-lying meaning of participants to an event (i.e. thematic roles) with their observed surface syntactic ordering or grammatical encoding across languages.

The notion a thematic hierarchy or ranking has been derived from the observation that different individual semantic roles like agent, causer or patient and causee, can be grouped into more abstract Proto-roles providing a highly reduced role inventory of all the possible semantic roles a verb can give, such that a Proto-Patient (i.e. *Undergoer*) in an event requires the presence of a Proto-Agent (i.e. Actor), in fact Proto-Agent and Proto-Patient entailments have been formalized as coming in pairs (Dowty, 1991; Croft, 1998; Primus, 1999)[49].

There exists multiple accounts for these thematic hierarchies ranging from the relative semantic salience à la Fillmore (1977) or prominence (à la Jackendoff or Van Valin), or even discourse topicality scale of argument types (à la Givón) for the participants to an event. Two main currents can be identified, the first (1) concerned with the encoding of relations of 'semantic' *structural* prominence among the argument of a verb, and the second (2) formalizing a 'salience' hierarchy reflecting the relative topicality of arguments (i.e. arguments bearing higher semantic roles are more likely to be topics than those bearing lower roles on the topicality hierarchy), namely Givón sees grammatical relations as directly encoding the information structure notion of topic (1984:134).

Different formulations have been successively proposed by Givon (1984)[50], Jackendoff (1990)[51], Dowty (1991)[52], like just to cite some of the most relevant ones, but all hierarchies express the idea that there exist a set of generalized proto-roles that capture valid linguistic generalizations in different languages.[53].

However, abstracting out from these different analyses[54], we can simply consider a

49. See the Role and Reference Grammar (RRG) actor-undergoer hierarchy (Van Valin and LaPolla 1997:126–27) implicitly attributing a prominence relation between the two arguments of transitive verbs.

50. Givon's Hierarchy (1984): Agent > Dat/Benefactive > Patient > Location > Inst

51. Jackendoff's Hierarchy 1972: Agent > Goal/Source/Location > Theme ; Jackendoff 1990: Actor > Patient/Benefactive > Theme > Goal/Source/Location

52. Dowty's Hierarchy.

53. Note that these thematic dependencies encoding prominence relations among a set of semantic notions have been formulated to constitute a language-independent generalizations ranking the possible semantic roles, and being conceptual in nature, they are assumed to be universal.

54. And certain number of contradictory hierarchy scales have been proposed to account for certain syntactic phenomena, but none ended up in providing a comprehensive system to answer the question

series of more general "*natural prominence scales*" that can be seen as expressing, in a more general way, the attempt of these different analyses, by ranking the possible values of some semantic, conceptual, pragmatic, or morphological attributes of a particular linguistic unit. These prominence scales can be appealed to in accounts of a wide range of linguistic phenomena like coding properties of arguments, word-order and morphological case encoding:

a. Person : first, second > third[55]
b. Referentiality : pronoun > proper name > common noun
c. Animacy : human > animate > inanimate[56]
d. Definiteness : definite > specific > non-specific[57]

For instance, a given Noun Phrase in a sentence can be viewed as showing a certain set of values drawn from the above *natural prominence scales*, it could be (1) an agent that is high on the '*thematic scale*', (2) a human, which is high on the '*animacy scale*', and thus being associated with (3) the subject-role, which is high on the '*grammatical relations hierarchy*'. Given this cluster of values on the different natural prominence scales, this Noun Phrase would take either nominative case or would be unmarked because of its basicness[58]. This example clearly shows how the properties on an NP can align in a particular way with the agent semantic-role, thus giving rise to what can be called a thematic hierarchy effect.

Moreover, these scales have been attributed a "cognitive salience", we can see that the semantic roles are ordered according to some notion of cognitive salience that is determining the semantic properties a Noun Phrases in a given sentences. In this regard, for instance taking into account the animacy feature of a Noun Phrase given it centrality in determining morpho-syntactic patterns in a wide range of languages (Comrie, 1989)[59], should be regarded as important factor as we will see shortly.

Interestingly, this conception of thematic hierarchy scales is often coupled with the assumption that the prominence relations between arguments in the event structure is mapped on to syntax (Marantz, 1993), which ends up assigning onto structurally higher syntactic positions, semantically more prominent arguments: "the lowest role on the Thematic Hierarchy is assigned to the lowest argument in constituent structure, the next lowest role to the next lowest argument, and so on." (Larson 1988:382)[60]. Nonetheless, such accounts raise a number of questions, especially if one considered them as generating a structural expectation in the listener during sentence comprehension.

Although, German morpho-syntax does not encode animacy distinctions as the languages that served to build some of these hierarchies, the application of an animacy

of how these semantic features systematically play a role in determining morpho-syntactic patterns in a wide range of languages, for an in depth discussion, see Levin and Rappaport Hovav (2005).

55. cf. Humboldt universal, chapter 1.

56. See Comrie (1989) and Tomlin (1986) who presented a study of the role animacy plays in the ordering of arguments in (dis-)transitive sentences.

57. We will further address this scale analyzing topic NPs in Mandarin Chinese.

58. Certain clusters of values can be considered as basic, as in the reported example, therefore only the divergence from the set of characteristics of a given cluster of semantic and grammatical roles typically yields explicit morpho-syntactic markedness.

59. For example it determines the linear ordering of arguments in certain languages like Fore, a language of Papua New Guinea. (Tomlin, 1986).

60. Already Keenan and Comrie (1977), had formulated the "NP accessibility hierarchy" showing a relative accessibility to relativization of NP positions in simple main clauses.

hierarchy is nonetheless apparent in the linear ordering of arguments in this language: (1) animate arguments should precede inanimate arguments, (2) higher-ranking participant roles should precede lower-ranking participant roles, and (3) pronouns precede non-pronominal arguments in unmarked German clauses and independently from their status as subject or object.

Given this rich set of word-order properties, it is not surprising that several neuro-imaging studies (ERP and fMRI) have been conducted on German word-order to try and clarify (1) the role of these hierarchies in the online cerebral processing of the sentence, and (2) the complexity effect that are observable when these principles are violated and the complexity of the form-to-meaning mapping consequently increases.

An fMRI study by Grewe and colleagues (2006) investigated brain activity responses to word-order deviations from the default order of grammatical functions in German. The authors showed that the permutation of word-order relative to the animacy feature rules in German can modulate cerebral activation at the border between Broca complex (pars opercularis) and the anterior portion of the ventral Premotor Cortex (vPMC) as illustrated in Figure 2.5D. According to the authors' interpretation, this area engages in a crucial aspect of the form-to-meaning mapping during sentence comprehension by specifically reconstructing the interpretive status of sentential arguments from their linear position in the sentence. An earlier study (Grewe et al., 2005) had already demonstrated that this very same area was sensitive to the language-specific principle ruling the linearization of pronouns and non-pronominal noun phrases in German.

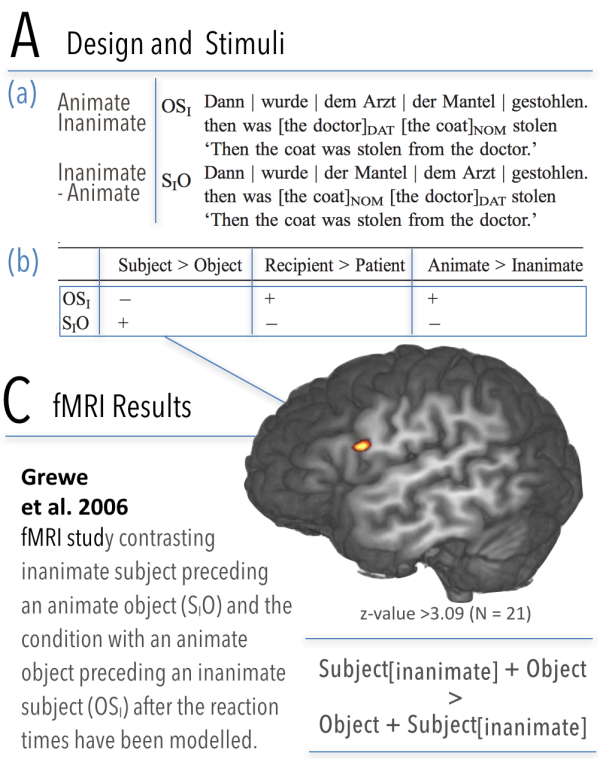


Figure 2.5 – (Aa) Critical experimental stimuli examples altering animacy word-order. (Ab) Experimental design's manipulation relative to three thematic scales dimensions, respectively Subject>Object, Recipient> Patient, and Animacy. (C) Group-average brain map of the animacy effect elicited by german word-order where Inanimate subjects precede the object.

2.2.2.2 Effect of animacy in syntactic structures

Generally speaking, the importance of Animacy has already been demonstrated by neuro-cognitive findings. Several event-related brain potentials results provide converging evidence for its role in disambiguation during sentence comprehension. For example, many studies reported interactions with case-marking in the online computation of a thematic hierarchy (Frisch and Schlesewsky, 2001).

Importantly, animacy manipulations were also reported to invert a well-documented complexity effect between object-relative clauses versus subject-relative clauses. Namely, an fMRI study by Chen and Colleges (2006) found that an increase in activation of Broca complex was observable for object compared with subject relative clauses only when the subject of the relative clause is inanimate and the subject of the main clause is animate.

(15) An Object-relative clause versus Subject-relative clause

a. SO: "The reporter [who the senator attacked ___] admitted the error."object-extracted relative clauses

a'. SS: "The reporter [who attacked the senator] admitted the error." subject-extracted relative clauses

(16) An Object-relative clause versus Subject-relative clause (Chen et al. 2006)

b. SO-Ani-INA: "The golfer$_{\text{ANI}}$ [that the lightning$_{\text{INA}}$ struck ___]survived the incident."

b'. SO-INA-Ani: "The wood$_{\text{INA}}$ [that the man$_{\text{ANI}}$ chopped ___] heated the cabin".

As The classical syntactic complexity effect reported in the very first neuro-imaging studies contrasting an object-relative clause and the subject-relative clause like in (15) was found only when the subject of the relative clause is inanimate and the subject of the main clause is animate like in (16b) and not when opposed to a syntactically identical sentence with a animate subject in the relative clause like in (b').

Despite the fact that syntax and semantic processes are structurally intertwined in a particular manner in these sentences, Chen et al. argue that their finding suggests that activation of Broca's area reflects the relative difficulty of thematic-role assignment: inanimate referents being dispreferred agents for actions, and animate referents being dispreferred under-goers of action[61].

To these arguments should be nonetheless added some experimental evidence showing that animacy is not the only semantic feature that yields complexity effects at the level of brain activations. A more recent study by Bornkessel and Colleagues (2009) revealed a wider network than the one we just saw in the study by Obleser et al. (2011) and Grewe et al. (2006) (compare Figure 2.4 and Figure 2.6C). This study manipulated another semantic feature known to influence word-order linearization patterns across languages: the referentiality of the Noun-Phrases in a sentence. Interestingly, the natural prominence scale for referentiality presented above is actually predicting the complexity effects observed in this study in particular in pars opercularis [-53 11 5], as shown in 2.6 (B).

Animacy, referentiality and definiteness will be relevant aspects for the research conducted in this thesis, crucially these semantic variables will be controlled for in our four experimental designs to be able to isolate effects linked to the syntactic parameters manipulated in each study. Topic-Comment articulations will be particularly subject to these kind of discussion (in chapter 7), because Topicality do carry an 'aboutness' feature that is highly sensitive to some of the "natural prominence scales" introduced in this section. Presenting each experimental design, we will turn to more concrete considerations, and we invite here the reader to keep in mind that the studies that we just presented indicate that a certain sub-set of brain areas are sensitive to diverse principles governing linear order in language as brought to light by linguistic analyses we presented about thematic hierarchy and the semantic features that correlate with it.

In sum, every language has its basic word-order correlating with another sets of rules governing other functional elements present in the language, to which one can add some

61. See also Kuperberg et al. 2008 for neuro-imaging results going in the same direction.

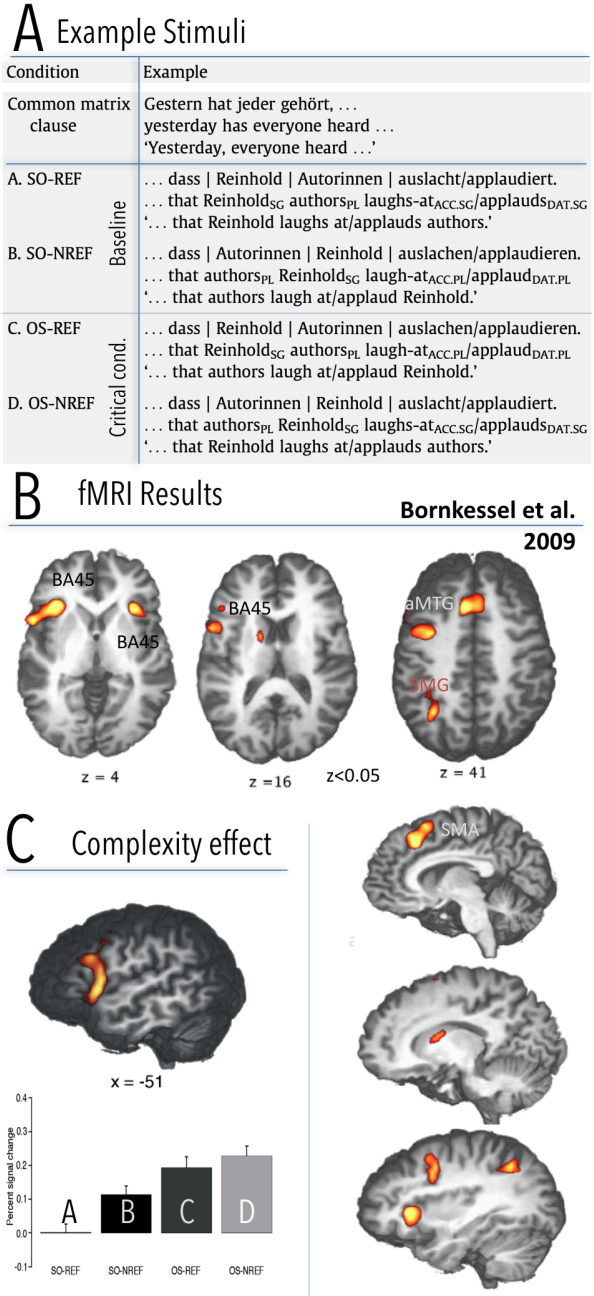


Figure 2.6 – (A) Critical experimental stimuli examples and design alternating subject-object word-order and referentaility feature. (B) Group-average brain map of word-order permutation effect [(A+B) Object-Subject > (C+D) Subject-Object. (C) Group-average brain map and percent signal change at the cluster peak for the complexity effect [A<B<C<D]. Adapter from Bornkessel et al. (2009).

rules called Principles that universally apply on this linguistic unit. The existence of a *basic sentence word-order structure* in every language is psycho-linguistically important, in that it implies that a certain word-order is *expected*.

Continuing further after these initial experimental pieces of evidence for the "*basicness*" of canonical word-order as identified in linguistics and through the neuro-imaging lens, we will ask what is a basic sentence from the point of view of *linguistic behavior*. Yet, if sentences in all languages appear to package information inside the sentence according to basic structures, however the notion of basicness in language is often advocated, but very rarely defined. We will introduce the how experimental psychology arrived to define it.

2.2.2.3 Basic Sentences seen from the psycho-linguistics point of view

In the 1970s, psycho-linguists qualified the notion of basic sentence by defining as the sentence pattern that should presumably be the object of subject's expectations in a given language. This notion was adapted from Fodor, Bever, and Garrett (1974), where the hypothesis was first formulated that a reader initially assumes the sentence as a sequence of constituents, where a noun phrase is followed by a verb phrase, and then followed by a second noun phrase: [NP + VP + NP], and then assigns them respectively the roles typically found in a common clause in the declarative form, namely subject, main verb, and object (in a language with SVO word-order)[62]. The term dedicated to this basic sentence pattern in experimental psycho-linguistics was "canonical" (Baldwin, 1976, 1977). Thereby, a sentence that infirm this expectation for a basic grammatical pattern would be defined as "non-canonical," and if it confirms the expectation, it could be called "canonical".

Sentence-internal punctuation

In a pioneering experiment Baldwin and Coady (1978) explored the relationship between punctuation and grammatical word-order expectations in young (10 years old) and adults readers. Their experimental design, manipulated sentence word-order to vary the criticalness of punctuation cues, ranging from critical to redundant. Sentence-internal punctuation is here again assumed to be an orthographic device signaling syntactic patterns to the reader, and in critical experimental conditions it is, thus, introduced to rule out the default case (i.e. the canonical word-order assumption[63]. Hence, when sentences are non-canonical, punctuation seems essential in arriving at appropriate syntactic analyses. When sentences are canonical, punctuation appears merely to reiterate grammatical information already provided by word-order.

The interaction plot in Figure 2.7 shows that for canonical sentences the presence or absence of punctuation had no relevant effect upon comprehension for the two age group. On the contrary, the comprehension drastically fell for both groups when non-canonical sentences were presented without punctuation. A closer look to the results for punctuated non-canonical sentences, reveal that the performances of adults and children diverged:

62. The "canonical sentoid strategy" advanced in the 1970s (Bever 1970, Fodor et al. 1974) was indeed a strategy that "whenever one encounters a suface sequence NP V (NP), assumes that these items are respectively subject verb and object of a deep sentoid" (Fodor et al. 1974:345 p).

63. Note that here is hiding the implicit notion that grammatical expectations are the products of active reader strategies.

the punctuation had no effect on the comprehension of the non-canonical sentences for the fifth graders, while adult subjects comprehended punctuated non-canonical sentences practically as well as punctuated canonical sentences.

The results indicated that criticalness punctuation varies as a function of preceding word-order in adults, while fifth graders, in contrast to adults, tend to ignore grammatically critical punctuation cues relying more on word-order cues to structure. Interestingly, the fact that the importance of word-order is stronger in children, shows that word-order is taken in this population as a primary source of syntactic information. All in all, word-order and punctuation had a profound effect on the ability of the subjects to understand the stimulus sentences.

In conclusion, these results offer psycholinguistic evidence for the fact that the canonical word-order of the constituents in a give language does matter in sentence comprehension and that it constitute a grammatical expectation in readers, that in adults readers is constrained by punctuation syntactic information. We will now turn to what is cross-linguistically expected to be found in a sentence, by introducing here the basic properties of simple declarative sentence, seen both from the point of view of Typology of Generative Grammar.

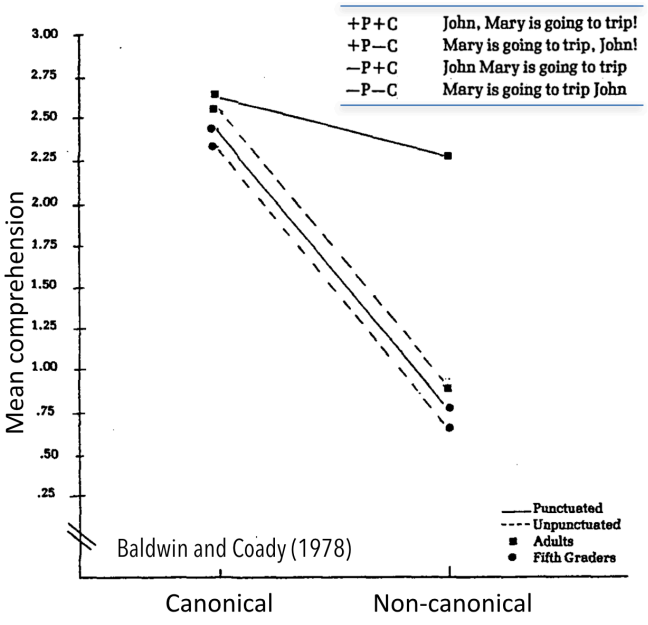


Figure 2.7 – Mean comprehension deceases for non-punctuated and non-canonical sentence. Adapted from Baldwin and Coady (1987).

2.2.2.4 What is cross-linguistically basic in a Sentence

In typology, the issue of variance and invariance is addressed by analyzing universality and variability of word-order constuals across typologically divers languages. As mentioned above, word-order is understood as the order of elements in a typical declarative sentence with a transitive verb.

The vast majority of the languages of the world fall into one of three groups: SOV (Japanese, Tamil, Turkish etc.) ; SVO (Fula, Chinese, English etc.); VSO (Arabic, Tongan, Welsh etc.). SOV and SVO together are found in more than 85 percent of all languages, while VSO is only found in around nine percent. Nothing wrong with the three other possibilities: VOS, OVS and OSV, however, they are exceedingly rare and typically occur in areas that have been relatively isolated (Eifring and Theil, 2005).

In this distribution of word-order patterns across languages, we clearly can see that subjects strongly tend to precede objects. Different hypotheses have been advocated to account for the fact that the subject tends to occur early in the sentence. One of the first argument advocated is linked to agenthood and animacy, subjects are prototypically agents and human, and secondly the thematic role of agent tends to precede the thematic role of patient.The third argument is that information that is more thematic tends to precede information that is less thematic. And, since very often the subject

is also a theme of discourse, it tends to precede the object that is usually expressing new information. When other elements are more thematic than the subject, they are often moved out of their original position and placed before the subject, as in Mandarin Chinese:

(17) 这本书我不喜欢

zhe4 be3n shu1 wo3 bu4 xi3huan1
This CL book I NEG like

'This book I don't like'.

On this very tendency of having a subject in the sentence, it should be noted that among the fundamental innate endowments of our linguistic faculty - a finite set of Principles that are common to all languages - Generative approach identifies as one of the most basic example of linguistics Principle the following:

A sentence must always have a subject, even if it is not overtly pronounced.

According to these facts and the theoretical assertion that that has been made about them, our first interrogation about what is the sentence, is subject to a shift towards another question: what is the subject in a sentence. Thus, we will now approach the notion of basic sentence as being the domain where a subject or/and topic dwell.

2.2.3 The Sentence as the kingdom of the Subject

From Aristoteles to Descartes the notion of sentential *subject* indicated been 'what is told about in a sentence'. The etymological origin of the word 'subject' comes from ancient Greek *hypokeimenon* 'that was thrown under", that was literally translated in the 5th century in Latin into *sub-jectus* "that stays under" [64]. This first hint coming from etymology indicate that the subject is something that undergoes -that is subject to- the verbal predication, in other words that affected by the verb's predication[65].

A fairly counter intuitive fact comes from the cross-linguistic investigation of the syntactic encoding of Subject role. Namely, considering even general properties that Subjects exhibit across languages, we can note that there aren't any that can really be considered as universally shared by all languages[66]. Many languages do mark subject grammatical function by the position with respect to the verb, case-marking or cross-referencing with the verb. But, these coding features are not present in all languages[67]. For instance, Hindi shows no constant cross-referencing for the subject: the subject of transitive verbs may be alternatively nominative or instrumental or nominative or dative

64. A. Moro in 'The Boundaries of Babel'(2008) offers a reflection about this etymology, asking the following question: Why should something that "stays under" be a "*subject*"? He recall that the notion of *substance* itself has the same etymology, advocating this is roughly because, in the ontological model that was largely elaborated in Aristotelian thought, the idea was that under all properties there existed just pure things in the world, entities; thus, a subject is just what is "under a property (or a set of properties).

65. In classical transformation grammar (Chomsky 1965) the subject is described as the Noun Phrase under the the sentence node (S) in opposition to the object being the Noun Phrase under the Verb Phrase.

66. For cross-linguistic properties of subjects see Keenan (1976).

67. For an overview of the preverbal subject field in the realm of the Cartographic approach see Cadinaletti (2010) 'Towards a cartography of Subject Positions', where the author claims its relative homogeneity across languages, compared to th epost verbal subject post-field.

(ref). Another example of inconsistency in marking grammatical relation is found in 'split eragtive systems'[68] frequently found in Austronesian languages that present a marked 'ergative' form expressing Agent function in opposition to an unmarked 'absolutive' for object or subjects of intransitive verbs, like in Warlpiri example in (18) (Shopen, 1985; Comrie, 1978) [69].

(18) from Warlpiri (Shopen, 1985:57 & 55)

a. Ngaju ka-rna-ju nyuntulu-rlu nya-nyi
me(ABS) PRES-you-me you-ERG see-NONPAST
'The man is spearing me.'

b. Ngaju ka-rna nyuntu-kurra parnka-mi
I(ABS) PRES-I you-ALL run-NONPAST
'I am running towards you.'

(19) from Icelandic (Shopen, 1985:107 & 55)

a. mér líkar vel við henni
me(DAT) likes(3SG) well with her
'I like her.'

b. mér lika/líkar þeir
me(DAT) like(PL/SG) them(MASC-NOM-PL)
'I like them.'

Even in languages like Modern Icelandic that do show a grammatical coding of the subject both by pre-verbal position, nominative case marking on Noun Phrase and verb agreement, we can find exceptions of subject presenting oblique case with some particular verbs[70]. Furthermore, as we see in (19) these subject just have pre-verbal position but don't agree with the verb.

Without further developing these observations, we see that the great variety of NP-marking and cross-referencing patterns found cross-linguistically gives a general picture where positional properties are probably the most reliable, but where still grammatical coding feature fail to provide stable syntactic hints to this sentence functional role.

2.2.3.1 Null Subject Parameter

To this should be added that not all languages always have overtly expressed subjects in their basic sentences, particularly pronominal subjects in finite clauses. This observation motivated the postulation in the Principles and Parameters framework for a 'Null Subject Parameter'. More precisely, the languages concerned by this parameter range from languages always requiring the subject function to be expressed as a nominal, to languages where the subject pronoun is redundant because they mark their verbal form with both number and person of the subject, or to cases where subject ellipsis is very common in various type of grammatical constructions, like coordinate clauses, 'while+gerund', relativization strategies, infinitival complements (Avery D. Andrews, 1985).

68. Split ergative systems are shown in languages that partly have an ergative behavior, but employ another syntactic or morphological marking, usually an accusative one, in some contexts (cf. Glossary).
69. Warlpiri is an aboriginal language of the pama-nyungan family, spoken in the central west part of the Northern Territories of Australia.
70. As a side note, this aspect could also be linked to the semantics of the verb 'like'.

A detailed overview of cross-linguistic variation in this respect shows a fairly counter-intuitive picture: Languages that allow null subjects are actually significantly more widespread than those that don't (Newmeyer, 2005)[71].

Since Rizzi (1982, 1989), the syntactic configuration linked to Null Subject Parameter has been described as the possibility for the Inflectional Phrase (Infl) to license a null pronoun (little *pro*) in its Specifier position, as shown in example (23c).

Further observations brought to light that this parameter is often accompanied by other tendencies[72], forming, therefore, what we already called, in Sections 1.1.4 and 1.2.4, a parametric cluster (Rizzi, 1986), where a null (referential) pronominal subject is allowed with no particular formal or contextual restriction in Italian example (a), while this is not the case in English (a')[73]. This *pro* licensing is traditionally assumed to correlate with at least two other properties typically displayed in the two classes of languages identified by this parameter: Verb-Subject inversion, shown in the below example (b) and absence of that–trace effects shown in (c).

(20) The possibility of a silent, referential, definite subject of finite clause

a. 'Sei simpatica.' Lit. 'Are nice', 'You are nice'

a'. * Are nice. vs. You are nice.

(21) Possibility of VS ('(free) subject inversion')

b. Parlava Gianni/un ragazzo.

b'. * Talked John/a boy. vs. John/a boy talked.

b". * Parlait Jean/un garçon vs. Jean parlait.

(22) Subject extraction through an overt complementizer (i.e. the absence of complementizer-trace effect.

c. Chi credi che verrà? Gloss: Who think.2s that come.FUT '*Who do you think will come?*'

c'. * Who do you think that will come?

c". * Qui as-tu dit que ___parlait ? vs. Qui est la personne dont tu a dit qu'elle parlait ? (Chi hai detto che parlava?)

Hence, Italian shows all the characteristic to belong to this cluster of Null Subject pro-drop languages, in that it allows sentences like 20 (a), (b) and (c), by contrast English does not, as shown in (a'), (b') and (c'), and neither does French in (a"), (b") and (c"). Note that in non-Null Subject languages the order VS is lexically constrained and typically possible with some verb classes, like unaccusatives as we will see in chapter 6[74]. Gilligan (1987) have offered a systematic analysis of the cross-linguistic distribution of the properties presented in example 20. He compiled a hundred languages balanced sample, selecting varied geographical distributions and linguistic families. The results showed a

71. Examples of these languages encompass most of older Indo-European and Modern Romance languages, or Celtic languages (except Irish), just to cite a few.

72. Allowing a wh-movement from a finite embedded clause across a complementaized (Perlmutter, 1971).

73. By contrast, English is a non-Null Subject language in that it does not allow the subjects to be omitted in this type of structure, However, take an example from old English in Shakespeare: *'Wilt come?'* meaning '*Will you come?*'(from Stephano in 'The Tempest', III.2). It could be argued to have a null *pro* subject, and hence to have the structure [Wilt *pro* come?], with *pro* having essentially the same interpretation as the second person singular pronoun 'thou'.

74. More recent studies brought to light the fact that VS inversion is not free in null subject languages, but that it is clearly "discourse related", as discussed in detail in Belletti (2001, 2004) with examples of question-answer pairs.

interesting pattern offering a validation of the possible universality of the characteristics linked to this Parameter: out of the sixteen possible combinations of the four properties, the Null Subject parameter was built on, only seven occurred more than once, and two only once[75], thus showing a cross-linguistic correlation between these different properties (see Table 2.3).

Combination of properties of Null Subject Parameter	Distribution across languages			
	yes–yes	yes–no	no–yes	no-no
pro_{+ref} – pro_{-ref}	24	0	15	2
pro_{+ref} – Free Inversion	22	49	11	15
pro_{+ref} – *that*-t[3]	5	3	2	1
pro_{-ref} – Free Inversion	14	25	1	1
pro_{-ref} – *that*-t	7	2	0	1
Free Inversion – *that*-t	**4**	**0**	**3**	**4**

Table 2.3 – A numerical breakdown for six theoretically relevant correlations. Null-subject [+referential] : pro_{+ref}, Null-subject [-referential] : pro_{-ref}; That-trace effect : That-t. Adapted from Gilligan (1987).

2.2.3.2 The Sentential Subject and its *Aboutness* feature

Yet, the questions of the functional marking of subjects or of their overt expression in sentences, shouldn't distract us that from a fundamental consideration: sentence subjects are associated with certain interpretative properties that make the subject be the argument *about which* the event is presented. Going back to what we saw when introducing typological work on Hierarchies of Grammatical Relations, we can argue here that Subjects carry this aboutness feature because they are more thematic that objects, that they out-rank objects in the thematic hierarchy (Jackendoff, 1972).

Moreover, this particular aboutness function is what makes passive and active sentences differ in "*aboutness*": (a) is about a truck and (b) is about a bus.

(23) a. Un camion a tamponato un autobus.
A truck has bumpted a bus
'A truck bumped into a bus.'

b. Un autobus è stato tamponato da un camion.
a bus has been bumpted by a truck
'A bus has been bumped into by a truck.'

c. Poi, *pro* è partito.
Then (it) has left
'Then it left'
(Calabrese, 1986 in Rizzi 2012)

This difference in *aboutness* has important consequences for the overall interpretation of the sentence, and for its discourse articulation. Namely, in a Null Subject languages like Italian, where we just saw that a silent pronoun (*pro*) is left in the syntactic structure in subject position, the referent that this silent pronoun should pick up from discourse context, is the previously established "aboutness" subject. Thus, if (c) is preceded by

75. For debates about how to theoretically interpret these figures see Roberts and Holmberg (2005) or Baker (2001) showing the substantial occurrence of several of the predicted correlations appear to provide important support for the parametric model (*contra* Newmeyer's (2005) proposal).

(a) its interpretation will be that '*the truck* left', while if it is preceded by (b) its interpretation will be that '*the bus* left'.

Conversely, without entering into the considerations about the prominence-mechanisms that are generally associated to word-order changes in the languages of the world (i.e. contrastive, list understanding, focalization, etc.), truth-conditionally equivalent sentences might fundamentally differ if we consider something that has been called *information packaging* by Chafe and Lambrecht in the '80s (see next Chapter (Ch. 3 and 3.4. Consider the following examples :

(24) a. Jacqueline Kennedy married Aristoteles Onassis.
b. Aristoteles Onassis married Jacqueline Kennedy.

Thus, sentences in (24) are true under the same circumstances, but, under normal prosody, they carry different propositional information: while (a) is to be understood as being an utterance about Jacqueline Kennedy, (b) is an utterance about Aristoteles Onassis.

So that the difference is not in what they say about the world but in how they say what they say about the word. In other words this non-truth conditional difference in sentence understanding we observe is a difference in *information packaging*.

Thus, the conclusion which these last observations leads us to, is that sentence unit has the fundamental and universal characteristic of being a predication about something or someone and that this aboutness role can fall to a given constituent for further discourse use as we saw in example 23 (c).

An often advocated characteristic for subjects is *definiteness*. Although many languages do not have a strict requirement on this feature, like English ('*A cat was miaowling.*'), many still show a tendency for subjects to be definite. A corpus study on English written narratives by Givon (1979:51-53) reports that 91% of the subject to be definite, while only 56% of the direct objects were (object are usually understood as new information). These facts among others are explained by a tendency for subjects to embody the aboutness feature, the so-called topicality. These aboutness relations inside the sentence unit, are realized across the world's languages in different ways, and interestingly attributing different grammatical encoding to the constituents that functionally carries aboutness feature.

In the next Sub-section we will briefly present the functional role that embodies this aboutness relationship, the Topic, and the kind of syntactic configuration in which sentences carrying a Topic have.

We will therefore turn to the question of how this sentence aboutness relation is grammatically encoded in languages that do not choose the Subject-predicate relationship for this effect.

2.2.4 The Sentence as the kingdom of the Topic

As we just saw, grammars of natural languages offer various syntactic means to express the same basic informational content conveyed by a sentence-unit. These different syntactic encodings tend to correlate with the aboutness feature that speakers want to establish in the sentence, and with the different informational statuses that speakers assign to components in the sentence-unit.

A fair number of languages present syntactic (and/or prosodic) constructions and explicit morpho-syntactic marking for representing similar propositional content, differing only in the way the information is structured inside the sentences (cf. the notion of *informational packaging*)[76]. Typically, the marking of the 'aboutness' functional role of a constituent in the sentence-unit yields a structuring of the utterance according to the Topic-Comment articulation.

From a general point of view Topics are principally found in sentence-initial position being marked by a prosodic break (25c), a Topic particle (26a) or a specific pronominal form (25b and c), like in following French and Dagara (Burkina Faso) and Gungbe[77] examples:

(25) Dagara (Burkina Faso) and French

a. n bέ ŋuúd a-daán ɛ́
I not drink DF-beer ACT
'I don't drink beer.'

b. màán bέ ŋuúd a-daán ɛ́
me not drink DF-beer ACT
'Me, I don't drink beer.'

c. Moi, j'bois pas de bière.
me I drink NEG partitive beer
'As for me, I don't drink beer.'

d. màán yaa bέ ŋuúd a-daán ɛ́
me not dring DF-beer ACT
'As for me, I don't drink beer.'(Delplanques, 2012)

(26) Gungbe (Gbe)

a. Un sè [do [dan lo yà [Kofi hu ì]]]
I heard that snake the TOP Kofi killed it
'That snake, I heard Kofi killed it.'

b. Un sè [do [dan lo wè [Kofi hu_]]]
I heard that snake the FOC Kofi killed
'THAT SNAKE, I heard Kofi killed.' (Aboh, 2004)

The sentences in the above two examples (25) and (26) are representatives of some of the different syntactic configurations that can characterize Topic-Comment articulations across languages. We see in (25) that the utterance in (a) is given the a Topic-Comment articulation in (b) but without any additional marker except the fact that the pronominal subject playing the role of topic is changed into a particular pronominal to account for this role, as it is the case in french too, when saying "*Moi j'bois pas de la bière.*" [me I.don't drink beer]. Alternatively example in (c) show that in the same language one can find the expression 'as for' to convey in a more explicit way the aboutness feature of the topic in respect to the whole sentence. Example (26) shows that Gungbe linguistic system offer the possibility to mark topics through a dedicated morpho-syntactic marker *yà* that is specialized for Topic and critically distinct from that for marking another discourse feature that is focalization, as shown in (b) where capital letters in the translation indicate emphasis.

76. This phenomenon calls in to play the notion of *informational packaging* that will be later addressed when drawing a thorough overview of the definition of topic notion in chapter 3.

77. The second most spoken language in Benin, belonging to the Gbe cluster of the Niger-Congo Languages.

To these Topic-Comment sentences one should add a more extreme type, where the grammar of a language does not specify any association between the Topic element and the rest of the sentence, and the Topic needs not have a semantic role in the Comment sentence. Consider a situation, where in a close fiends circle, one friend (David) had just being left by his wife (Jannie), and at dinner one of the friends could ask the following question:

(27) "Speaking of David, what has Jannie been up to, lately?".

Although these type of sentences are considered to have a minor role in the system of languages like English or French, they are far more frequent than one could imagine. Let us consider now, another example that relies less on sharing a detailed contextual background, where we can clearly see that the topic introduced by '*as for*' cannot have a role in the comment clause[78]:

(28) a. As for American self-confidence, Columbia gave people a lift.
b. * American self-confidence, Columbia gave people a lift.
c. * It was American self-confidence that Columbia gave people a lift.

Above examples in (b) and (c) represent attempt to force the Comment clause ('Columbia gave people a lift') to assign a semantic role or a grammatical role to the Topic Noun Phrase, but their agammaticality shows that topicality can be irreducible to a function assigned by the clause in the Comment. In other words, only the preposition 'As for' assigns here a role to the sentence initial NP 'American self-confidence'. We can tentatively define the role played by As for in the above sentence by relying on Chafe's definition of Topic (1976): "*the topic constitute the framework within which the main predication holds*".

Cross-linguistic studies in the 70s will reveal that in many languages this kind of sentence articulations are the predominant form of basic sentences in ordinary language use. Since then, languages like Mandarin Chinese have been called 'Topic-Prominent Languages' as defined by Li and Thompson (1976). We will resort to a precise and detailed description of topicality in the next chapter (§3), but before we will clarify the kind of predicative relation between Topic and Comment shown in the above example.

2.2.4.1 The most general predicative construction

The notions of Topic and Comment, in their modern usage, were first introduced in early 1958[79], by Charles Hockett, who stated that Topic-Comment sentence articulation articulation could be considered as the most general predicative construction[80]:

78. One might find it inadequate to speak about politics in a PhD, unless it is dedicated to Political science, which is not our case. But the example here under was offered by the reading of a fieldwork linguistic book, where Avery Andrews discusses topicality with English examples. We could imagine rephrase this example in the light of actual world affairs to give it a fresh new look: "As for American self-confidence, *Trump* gave people a lift."

79. Hockett, Charles F., A course in modern linguistics. New York : Macmillan, 1958. p. 621., Note that the term Topic made its first appearance in this book by Hockett.

80. We already met with Charles Hockett in sections §1.3.3 and 1.1.3. As a recapitulation, he was an American linguist that became known for his first formulation in a milestone article 1960 a list of 13 characteristics of human language compared to animal communication. The first 9 would be shared with other animal species, while the last 4 would be typically and exclusively human. Among those we can of course find the double articulation by Martinet (called "duality of patterning"). Hockett 1960. The Origin of Speech, Scientific American 203, p. 88–111.

> "The most general characterization of ***predicative constructions*** is suggested by the terms "topic"and "comment"[...]: The speaker announces a topic and then says something about it."

This characterization fits very well to the English examples give above, and expresses a way of understanding predication that constitutes a change of paradigm in the way of conceiving sentence as the 'kingdom of subject': the subject ended up being *'thrown under'* a Topic ; if we may afford a pun, it was "*sub-jected*" by Topicality. Such a shift, permitted to start considering the sentence-unit as a syntactic unit having a discourse interface.

My study of the cerebral representation of the sentence is conducted along the lines of this *Copernican shift* where the concept of sentencential subject 'turns around' that of Topic. We will claim that, not only from a linguistic point of view, but also from cognitive point of view, the analysis of clause structure should acknowledge two important factors: first the discourse interface that topics embody in the sentence, and second the particular *predicational* relationship they establish with the rest of the sentence, the Comment. So that, the *Predicative construction* inside the sentences are not uniquely or always governed by a verb, as we saw in sentences like in (28): "As for American self-confidence, Columbia gave people a lift." where we cans say that the predication about 'American self-confidence' is independent of verbal theta-role assignment[81].

Let us now consider a few examples in (29) of this predicative construction excentering the verb from other languages. We will note how difficult it is to give an appropriate translation without resorting to a lengthy paraphrase or to 'as for' expression, or 'speaking of' constructions that really do not reconstitute the naturalness and basicness of this sentence articulation.

(29) a. a. 那场火，心亏消防队来得快。Mandarin Chinese

Nà-cháng huǒ， xīnkuī xiāofángduì láide kuài.
that-CL fire, fourtunate fire-brigade come quick

'That fire fourtunately the fire brigade came quickly'
(Li and Thompson 1978)

b. Lahu

Ho oe ma-qho yi ve yo
elephant TOP nose long PCL DECLARATIVE

'Elephant, noses are long.' (Li and Thompson, 1976)

c. Japanese

Nihon wa, Tokyo-ga sumi-yoi.
Japan TOP TokyoNOM easy-to-live-in

'As for Japan, Tokyo is comfortable to live in.'(Kuno, 1973:65)

A characteristic of Hockett's North American approach, compared to the early European explorations on the structure of utterance, is that, even though in the line with

81. For an interesting point of view on a similar kind of non-theta predication see the notion developed by Caroline Heckoch (1993) as 'syntactic predication' with examples from Japanese Topics. Heycock (1993). Syntactic Predication in Japanese. *Journal of East Asian Linguistics* 2:167-211.

the early functionalism of Firth[82] and the Prague School tradition that we will present in chapter 3, its major claims about Topic and Comment structure are largely based on linguistic observations made on Chinese language. Namely, the encounter with Eastern Asian linguistics led to important new insights on sentence structure, with a new understanding about the relation between discourse and syntax.

2.2.4.2 The Topic and its *Aboutness* feature

Consider the following example from Mandarin Chinese, in parallel with previous Italian examples in (23) presented the preceding sub-section and reported in (d, e, f):

(30) a. 那棵树叶子大，

nà-kē shù yèzi dà
this-CL. tree leaves big

Lit.This tree leaves big 'The leaves of this tree are too big'

b. 那棵树的叶子大，

nà-kē shù de yèzi dà
this-CL. tree DE leaves big

Lit.This tree leaves big 'The leaves of this tree are too big'

c. 所以我不喜欢。

suǒyǐ wǒ bù xǐhuān ___
so 1st.SG NEG like ___

'so that i don't like ___(it).'
a+c -> *I don't like the tree,*
b+c -> *I don't like the leaves*

d. *the truck left*

Un camion a tamponato un autobus. Poi *pro* è partito.
a truck has bumpted a bus then *pro* has left.

'A truck bumped into a bus, then ___left.'

e. *the bus left*

Un autobus è stato tamponato da un camion. Poi *pro* è partito.
a bus has been bumpted by a truck then *pro* has left.

'A bus has been bumped into by a truck, then ___left.'

In the Chinese Mandarin sentences in 30 (a) the presence of a sentence-initial NP '*that tree*' playing the role of the sentence Topic make the sentence subject 'leaves' loose the interpretative properties that define it as the argument *about which* the event is presented. So that, as we saw in the Italian example contrasting active and passive sentence

82. Discourse analysis of the sentence from some linguists disciples of J. R. Firth like M. Halliday will be presented in chapter 3. We can here recall that J. R. Firth used to encourage his students to carry out research on a number of African and Oriental languages. This encouragement made some afterwards well known linguists like Mitchell work on Arabic and Berber, Palmer investigate Ethiopian languages, and Michael Halliday concentrate on Mandarin Chinese.

articulation in (d) and (e), this difference in '*aboutness*' has non only important consequences for the overall interpretation of the two sentences, but also for their discourse articulation in one case the truck left (d), and in the other the bus left (e). Analogously, if the mandarin sentence in (c) is preceded by (a) the interpretation of what is not liked is the tree (i.e. the sentence Topic), while if it is preceded by (b) the object of like are the leaves, namely the sentence subject.

While in the Italian example 23 reported in (d) and (e), this shift in aboutness was given by a change in voice automatically provoking a change in grammatical subject between 'truck' and 'bus', in the Mandarin (a) and (b) this shift in aboutness is given by the presence of a sentence initial Topic NP.

This aboutness shift and the different sentence syntactic configuration that permit it will be tested in the fMRI experiment presented in chapter 7, by contrasting exactly sentence 30 (a) with 30 (b).

2.2.4.3 Null Topic Parameter

Over the last 40 years, the pioneer study of Li and Thompson (1976) demonstrating Chinese as a Topic-Prominent language, contributed to accomplish the shift towards sentence-based accounts of Chinese sentences' articulation with the discourse. In this way, the Topic was included in the so-called basic way of articulating a sentence unit (generally heading in sentence's initial position). Syntactic rules and descriptions of Mandarin were simplified and clarified by taking into account this sentential element having an interface with discourse.

Interestingly, this functional role was soon after inserted in the Principle and Parameter framework and Huang (1984) proposed a Null Topic Parameter. The observation that in Chinese the answer to the question where is Lisi 31 (A) could be followed by the kind of answers reported in 31 (B1) and (B2), brought to state that certain languages allow arguments to drop if they are in Topic position. Namely, in (B1) and (B2) the understood object of the verb is located in Topic position before dropping. Importantly, this argument is not restricted to Chinese examples, but it can generalize to other languages, as shown in the question about the whereabouts of Lisa in 32 (A) can be answered either by 32 (C1) or (C2) (see Huang, 1982 or Ross, 1982), where the pronoun '*ihn*' referring to Lisa in sentence initial position is dropped in virtue of its being the Topic of the sentence. Consider as a proof the un-grammaticality of (C1*) and (C2*) in 32[83].

(31) A: Where is Lisa?

a. B1: 张三说他没看见 ___.

Zhǎngsān shuō tā méi kànjian ___.
Zhangsan say 3SG NEG see ___

John said that he didn't see [her]'

b. B2: 张三看见了.

Zhǎngsān kànjian-le ___
Zhangsan see-PERF. ___

John saw [her]'

83. Consequently, it should be noted here that this shows that the missing argument is not licensed by formal features of the Tense node (T) as it is the case in Null Subject Parameter, see above Sub-section 2.2.3.

(32) A: Wo ist Liza? 'Where is Lisa.'

a. C1:

Heute ___hab' shon gesehen.
today have 1SG already seen

I have already seen [her] today.'

b. C2:

___hab' ich in der Bibliothek gestern gesehen.
___have 1SG in the library yesterday seen.

'I saw her yesterday in the library'

c. *C1: *Heute *ich* hab' ___shon gesehen.

d. *C2: **Ich* hab' ___in der Bibliothek gestern gesehen.

In sum, the corollary to which Huang's observation leads us to is that Null-Subject and Null-Topic Parameters can jointly classify languages into four types:

1. > [+null Subject, -null Topic] : like Italian and Spanish , etc.
2. > [+null Subject, -null Topic] : like Mandarin Chinese , Japanese and European Portuguese , etc.
3. > [-null Subject, +null Topic] : like English and Modern French , etc.
4. > [-null Subject, +null Topic] : like German and Swedish, etc.

Interestingly, Huang (1984) also noted that Null-Topic parameter was correlated with another one featuring the drop of the object - we can call object drop Parameter. We will leverage on this linguistic characteristics of Mandarin Chinese to investigate the cerebral underpinnings of dependency-links and co-reference inside the sentence in chapter 7 and present their linguistic analysis in next chapter (§3.4).

2.2.4.4 Null-objects in Chinese

Under appropriate discourse conditions, Mandarin Chinese allows both subjects and objects to be phonologically null (Huang, 1984, 1989). As shown in (36), all the responses by Speaker B are correct: the subject (i), the object (ii) or both together can be dropped (iii).

(33) Null Objects

a. Speaker A: 张三在写书吗?

Zhāngsān zài xiě shū ma?
Zhang San PROG. write book Q.

'Is Zhang San writing a book?'

b. Speaker B:

i. 在写书。

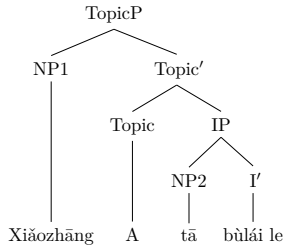
e zài xiě shū.
PROG. write book

'[He] is writing a book.'

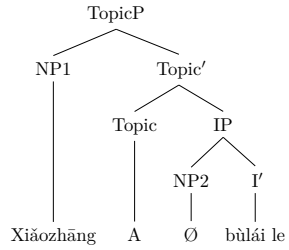
ii. 他在写。

Tā zài xiě *e*.
he PROG. write
'He is writing [a book].'

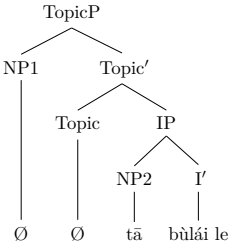
iii. 在写。
e zài xiě *e*.
PROG. write
'[He] is writing [a book].'



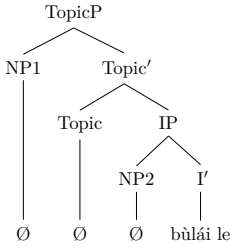
(a) *Xiaozhang, he did not come.*.



(b) *Xiaozhang, did not come.*.



(c) *(Xiaozhang,) he did not come.*.



(d) *(Xiaozhang, he) did not come.*.

Figure 2.8 – The possibility of dropping subject and object in Mandarin Chinese is here illustrated through syntactic-trees diagrams from Xu and Liu (1996): *Xiaozhang did not come.*.

Huang analyses null-objects on a par with topicalized objects, the difference being that when a null object is shown the topicalized element is null as well, as shown in the agrammaticality of (34b):

(34) Null Topics and Null Objects

a. 张三说 [李四不认识]

[Top e_i], [Zhāngsān shuō [Lǐ Sì bù rènshí e_i]].
Zhang San say Li Si not know

'Zhang San say Li Si not know X'

b. * [他], [张三说 [李四不认识]]

[Tā$_i$], [Zhāngsān shuō [Lǐ Sì bù rènshí e_i]
[Him$_i$], Zhang San said Li Si not know e_i.

'*Zhang San say Li Si not know him.' Huang (1984)

From this follows that Huang identifies null objects as variables in that they are bound by A-bar topic-operators. A point that will be criticized by Xu (1986), who argued that null objects cannot be variables as they do not undergo island constraints.

Since Huang (1984) the radical pro-drop observed in several far-eastern oriental languages illustrated by examples in (36) has been associated with the topichood of the missing argument as show in the pattern and the Null-Topic parameter. The main claim of Huang being that null objects must be bound by the discourse topic, in his own words: 'Its reference must be a discourse topic, someone or something that a given discourse is about' (p.541).

As for subject-drop in (36b), we saw in previous section that Italian witnesses subject-drop too (compare with 36i and iii). but has a rich verb inflection as shown in the verbal paradigm of the verb 'to write', e.g. 1SG:*scrivo*, 2SG:*scrivi*, 3SG:*scrive*, 1PL:*scriviamo*, 2PL:*scrivete*, 3PL:*scrivono*. This property is actually licensing null subjects, while Chinese hasn't any inflectional morphology: no verb conjugations, no case-marking and no gender on nouns[84].

The possibility for transitive verbs to easily drop their objects in Chinese yields an interesting contrast with English as for Topicalization is concerned:

(35) self-standing comment clause in Chinese but not in English

a. 那本书，他拿了

Nà-běn shū, tā ná-le
that book he take-ASP.

'That book, he took.' Xu (2001:141)

b. 他拿了

Tā ná-le
he take-ASP.

'He took.'

While the comment clause alone is not grammatical in English (b), it is in Chinese (d) thanks to the pervasive presence of null object in Mandarin.

We can further observe a subject-object asymmetry in the referential assignment of null pronominals in subject position (a) compared to object position (b) in the following examples from Huang (1989:187):

(36) Null Objects

a. Subject position: antecedent or Discourse referent (i.e. topic) 张三说 [(他) 很喜欢李四]

84. Notice that Saito (2007) also interprets missing arguments as Topics and relates this option to the lack of required agreement in East Asian languages as opposed to pro-drop in languages such as Italian and Spanish which have rich agreement. In his formal approach to these linguistic facts, he assumes that a covert Logical Form (LF) copying of elements available in the discourse, including *pro* into argument positions takes places. A unified account for both subject and object null arguments is given by Rizzi (1986) who also analyzes missing arguments as *pro*. He states that *pro* has to be licensed by a governing head, a statement that allows to identify the locus of the cross-linguistic variation of the different instantiations of object and subject drop phenomena in the values' parameterization of the licensing heads. Therefore, languages that have no overt agreement (i.e. do not use - features), like Chinese, allow free interpretation of *pro* where it is licensed.

Zhāngsān shuō [*e*/tā hěn xǐhuan Lǐ Sì]
Zhang San say *e*/he very like Li Si

Interpretation 1: 'Zhang San$_i$ said that [he$_i$] liked Li Si.'
Interpretation 2: Zhang San said that [Jane] liked Li Si.'

b. Object position: Discourse referent 张三说 [李四很喜欢]

Zhāngsān shuō [Lǐ Sì hěn xǐhuan *e*].
Zhang San say Li Si very like

'Zhang San said that Li Si liked [Jane].'

c. Subject position: antecedent or Discourse referent (i.e. topic) 张三说 [他很喜欢李四]

Zhāngsān shuō [tā hěn xǐhuan Lǐ Sì]
Zhang San say he very like Li Si

Interpretation 1: 'Zhang Sani said that hei liked Li Si.'
Interpretation 2: Zhang San said that [Jane] liked Li Si.'

In (36b) we observe that the only possible referent for the null object is a discourse one (e.g. Jane), from this follows the analysis that the empty syntactic element is a variable and it cannot be bound by any antecedent. While (36a) admits both being co-referential with a Discourse Topic and the co-referential link with the matrix subject 'Zhang San'. This asymmetry has been analyzed by Huang as showing that (36a) has a null pronominal in subject position that could be replaced by an overt pronoun 'ta' without modifications.

However, while the clear cut asymmetry shown in (36) and the generalization that has been following form it [85] are generally attested some examples suggest that it can be defeated under certain discourse contexts. Consider (37a):

(37) Object position: antecedent or Discourse referent (i.e. topic)

a. 小偷以为 [没人看见]。

Xiǎotōu yǐwéi [méirén kànjiàn *e*].
thief think nobody see *e*

Interpretation 1: 'The thief$_i$ thought nobody saw [him$_i$].'
Interpretation 2: 'The thief thought nobody saw [Jane].' from Xu (1986:78)

b. 小偷他以为 [没人看见]。

Xiǎotōu tā yǐwéi [méirén kànjiàn *e*].
thief 3SG. think nobody see *e*

Interpretation 1: 'The thief$_i$ he thought nobody saw [him$_i$].'
Interpretation 2: 'The thief thought nobody saw [Jane].'

c. 小偷以为 [没人看见]

Xiǎotōu *e* yǐwéi [méirén kànjiàn *e*].
thief *e* think nobody see *e*

85. On those grounds Huang (1989) proposed that *pro* is limited to the subject position of a finite clause in Chinese. This asymmetrical pattern led him to formulate a rule governing empty pronouns the Generalized Control Rule (GCR): Co-index an empty pronominal with the closest nominal element. (from Huang, 1984:552).

Interpretation 1: 'The thief$_i$ [he] thought nobody saw [him$_i$].'
Interpretation 2: 'The thief [he] thought nobody saw [Jane].'

The natural out-of-context interpretation of the empty element in (37) as referring to the thief '*xiaotou*' suggest that the empty element in object position of a finite clause can actually be analyzed as a null pronoun, which in turn appears then not to be limited to subject position. However, what is peculiar here is that the discourse referent interpretation has in this case to be specified by a particular context to be accepted.

Without entering in the details of the debates opposing Xu (1986) and Huang (1989) on these examples[86], we just want to underline how these examples attest once more the importance of in Chinese of the sentence-discourse interface features[87] which always makes it possible to drop objects and interpret the empty syntactic elements replacing them as referring to discourse topic even when the first natural interpretation indicate a sentence-internal antecedent reference[88].

To conclude, this examples show the fundamental different functions played by apparently similar null syntactic elements (i.e. a *pro* and a variable). Although we will resume to this issue with a full-fledged presentation of the different empty syntactic elements theorized by the formal approach to syntax in section §2.4.3 (see table 2.37, p. 187), it is important for delineating our experimental hypotheses to note that no-overt syntactic elements establish different dependency-links with antecedents or with discourse referents.

Our neuro-linguistic investigation will mainly focus on the cerebral representation of non-overt syntactic elements in French and Chinese: chapter 6 will namely correlate brain activity with the number of empty elements found in movement-derived French sentences. Chapter 7 will investigate the cerebral representation of different empty syntactic elements, offered by Chinese syntactic configurations similar to the above one, namely featuring null pronominals (*pro*), overt pronouns and gaps in minimally differing sentences to observe the brain activations they elicit.

Topic-Comment: a different way of articulating the sentence

This brief excursus into some central characteristics of Mandarin Chinese offers an initial grounding for our interest in Chinese language to investigate syntactic structures in the brain. Thus, given what we have presented so far, it might seem the right time to offer the reader as short explanation of why we decided to focus on Topic-comment sentence articulation in this pluri-disciplinary research project on the *sentence as a cognitive object*.

86. Huang establishes that what distinguishes languages that allow null objects from those that do not is that they license of zero topics to bind variables. While Xu (1986) argued that null objects cannot be variables as they not submitted to island constraints, for example.

87. Sections §3.2.3 and 3.2.2.1 will introduce the property of Chinese system assigning to Topics the control over pronominal deletion and co-referentiality.

88. This being said the above example by Xu (1986) could also be reinterpreted as the contraction of (37b) assigning to The first NP the role of Topic, which in turn would yield the structure in (37c)

Linearity and hierarchy As a matter of facts, besides sitting on the throne of Sinology as the discipline of *alterity* to explain the presence of Chinese in the experimental part of this work - the classical '*Chinese is different*' argument-, this choice was motivated by the wish to investigate the biological basis of a language system where articulating sentence according to the Topic-Comment pattern is the predominant sentential form in every-day ordinary usage, where the Topic role is mainly attributed to a sentence-initial constituent on word-order grounds.

Minimal morpho-syntactic marking The second advantage of choosing Mandarin as a linguistic testing ground is its relatively scarce morpho-syntactic marking, to this on can add that, Topic-Comment articulations are indeed among the most unmarked sentence structures, thus *permitting to study how the brain manages this incredible equilibrium the sentence achieves between linearity and hierarchy.* In Mandarin, Topic linear position and a minimal pause prosodic marking (respectively investigated in chapters 4 and 5) are the minimal cues required to maximally change the hierarchical relations in the sentence. Namely, in this syntactic configuration, the receiver of the linguistic stimulus has to build a syntactic hierarchy and to assign different roles to the Subject and the Topic of the sentence, by essentially relying on the linear word-order cues, to achieve an online structure building.

Topic: a dangling element in sentence-initial position The most interesting property of this syntactic configuration is that it allows the first element (the topic) not have an explicit selectional relationship with the main verb, and even not having a semantic role in the rest of the sentence. This crucial aspect of Topic-Comment articulations shown in example (27) –"*Speaking of David, what has Jannie been up to, lately?*" –, makes its very minimal overt grammatical marking (i.e. word-order) responsible not only for being a syntactic cue of the hierarchy between Topic and Comment clause, but it also guides the on-line building of an interface with the discourse information to achieve the understanding of sentences like "*Speaking of David, what has Jannie been up to, lately?*". This aspect, probably implying the reviving of all the contextual information about David to understand the content of the Comment, will be investigate during on-line sentence comprehension in chapter 5 through online recording of electro-physiological responses to Mandarin Topic-Comment sentence embedded in context.

Topic-comment an essential syntactic construct to "say something about something" Addressing the the issue of sentence as a cognitive object, Topic-comment articulations also provided the occasion to test the most extreme and at the same time essential configuration capable of create propositional meaning in human language, in other words, capable of "say something about something", by partitioning the sentence into a two step predication: first positing a Topic and then uttering a Comment about it.

However, this very basic predicative relationship -the fact of saying something about something- that Topic-Comment sentences embody, requires a well defined structural architecture of the sentence-unit. In facts, this minimal predication configuration implies a relationship between the Topic and the Comment-clause that over-arches the subject-verb agreement one, and establishes a sentence-level hierarchy between the Topic and the whole Comment sentence.

Sentence-discourse interface By addressing the issue of the internal structure of the sentence across languages in§2.2.4, Topic-Comment articulations were identified as carrying interesting discourse properties, which enriched our questioning about the structure of the sentence-unit. Taking into consideration the linguistic phenomena that witness of an interface between sentence-level and discourse-level will reveal the structural link that the sentence can establish with discourse and will allow to uncover the syntactic transformation of the sentence-unit's internal structure that occur through these interfacial phenomena. Far from any hasty pragmatic interpretation of these interfacial linguistic phenomena, what was identified in Topic-comment articulation is namely the *structural* possibility to establish a link between sentence-unit and discourse. The syntactic encoding of sentence discourse interface will be at the heart of our experimental research on the cerebral representation of syntactic complexity.

In conclusion, as the fundamental issue addressed in this manuscript is '*how is represented the structure of sentence-unit and its complexity by a human cognitive system like the human mind and the brain*', Mandarin Chinese, thanks to the above described properties, will allow to narrow down this broad issue about sentence structure asking the following questions:

1. what is the difference of structuring the sentence-unit with or without a Topic[89];
2. how is the syntactic hierarchy of Topic-comment represented by the brain both at the level of sentence's incremental processing (ch. 5) or at the level of functional brain activity (ch. 7);
3. how are the dependency links between topic and comment established in different cases where the topic have or not a direct selectional relationship with the main verb or a semantic role in the rest of the sentence.

Next chapter (ch. 3, p.225) will provide a rich overview of Mandarin Chinese Topic-Comment constructions and of their linguistic analysis, including a thorough introduction to the notion of Topic, and its particular syntactic characteristics.

But before, we will dedicate the next sections to specifying how the internal structure of a sentence-unit can be formally described. This formalization step will demonstrate to be an essential tool to study the sentence-unit from the perspective of cerebral processing all along the rest of the chapter.

89. Or around the concept of Subject or around the concept of Topic, as Li and Thompson (1981) used to formulate it.

2.3 The Sentence and its internal structure

> *"The meaning of each individual sentence is the product of the way the meanings of the words combine, guided by syntactic structure"*
>
> Ray Jackendoff, 2007

Despite the fact that the 'physical shape' of the linguistic signal is linear -featuring words arranged in a sequential order- the syntax of sentences in human languages is based on structural dependencies given by the hierarchical organization of words. The study of *syntax* is about the organization of words into phrases, and subsequently of phrases into sentences[90] As we will deal with sentence tree-structure representations all along this manuscript an introduction to what is a syntactic-tree and the advantages of representing a sentence structure in this format will be addressed. In order to guide the reader towards our broader experimental research question about the cerebral representation of the syntactic complexity, we will be brought to reframe the two research issues about syntactic complexity we already presented namely that implied for instance in canonical versus non-canonical sentences and in Topic-Comment articulations.

Grammar or the deep structure of the sentence

So far we have gathered evidence that the sentence unit is a natural and universal unit, yet, what remain to be accounted is its exact nature. One of the first answer to this interrogation echoes Jackendoff epithet, by asserting that the sentence is a unit where "*the way the meanings of the words combine, guided by syntactic structure*". A second answer, placing sentence-unit at the core of syntactic theory, comes from Chomsky's answer to Skinner's propositions in 1959, we already met discussing about Language Faculty in our preliminary excursus in chapter 1 (§1.1.3, p.14 and §1.2, p.16). Chomsky argued that formulating a theory requires an independent definition of the natural objects it studies, and that in case of language, he suggested that this natural object is the sentence, and that its definition is to be provided by a *grammar*.

These two steps are central in our approach, in that they establish a the fundamental distinction between the linear order given by the spoken chain, the linearity of sentence's phonological form, and sentence structural schemes, by positing the existence of a deep underlying structure -a *Grammar*- that is invisible. This underlying level of representation will be the object of our neuro-imaging investigation, and in practically all the chapters of this manuscript, we will be *chasing* in the brain what is *invisible or inaudible in language* -its structure. We will successively investigate:

1. how simple pauses in the sentence (and their related minimal intonation patterns) in the sentence result in structure building and consequent efficient contextual information integration recorded through ERP-EEG technique.
2. comparing abstract grammatical rules defining the main parameters of typologically distant languages like French and Mandarin Chinese can result in different fMRI activations strategies of the sentence network.

90. Interestingly, when Germans first translated the word Syntax they found the perfect fit building the word *Satzlehre*, namely the "science of sentence".

3. addressing the question of the representational complexity in the brain of empty abstract syntactic positions, that are nevertheless necessary for sentence comprehension of French wh-Questions and other complex sentence structures
4. and finally addressing the difference between overt and covert co-reference marking in Mandarin Topic-Comment sentences.

Sentence: a deep virtual form and 'zero markers'

As we have seen in chapter 1 (§1.3.3, p. 41), the development of a linguistic approach to sentence structure around the idea that there exist a deep invisible underlying structure for sentences isn't new. Lucien Tesnière already built a system to represent sentencial dependencies and structures, that already distinguished between structural-order of the sentence -its deep structure- from its surface linear-order, as the transformationists later did in the framework of Generative Grammar.

Not only did Tesnière achieve the first tree-like representation of the sentence's structure, but he started theorizing a non-material structure underpinning the visible structure of the utterance[91]. For him this non-material structure was having an observable and mono-dimensional linear-order, and a hidden and pluri-dimensional structural-order. Interestingly, he also theorized some general binding mechanisms between words -called "*Connexion*", "*Fonction*" and '*Translation*"[92], that can be roughly seen as corresponding to the syntactic-tree shapes and configurations we will observe in generative syntactic-trees in the next sections.

We should note that this way of considering Grammar and syntactic theory, was influenced by the so-called "grammaire logique psychologique"[93] a French linguistic trend of the period arguing that "Linguistic facts are essentially psychological facts". One of the main tents of the psychological grammar, George Galichet, will say about Tesnière's 'Elements of Syntax' that it was extracting "*the deep structure of our language*" (Galichet, 1949)[94].

As shown in Figure 2.9 (A), his structural syntax is built on the relations that exist between the structural-order and the linear one. Thus building the *stemma* of a sentence is equivalent to transform the linear-order into a structural one, translating the "*connections*" of the structural-order into sequences. Hence, we can note that this transformation was already understood as a mental operation and a multiple steps process: "*From this point of view, we can say that speaking a language means transforming the structural order in the linear one, and conversely, understanding a language is translating*

91. It is astonishing to remark that since the very first pages of its 'Elements of structural syntax', Tesnière is citing the notion of *innere Sprachform* of Wilhelm von Humboldt too. One should be non the less reminded that the concept of inner linguistic form was introduced by Wilhelm von Humboldt in order to designate the total worldview (*Weltanschauung*) of people who speak a given language. This concept is closely tied to the notion of the spirit of a people (Volksgeist) in that it is tightly linked to the idealization of the imagery that occurs in the usage of a language.

92. These connections between words (i.e. *Connexion*", "*Fonction*" and '*Translation* are described in the three parts of his '*Elements*' and depict the different and multiple connections that the terms in the sentence can have.

93. Namely, at that time in France an ongoing trend was to consider grammar as a form of Logic. The two principal representatives of this current were Albert Sechehaye, with his *Essai de structure logique de la phrase [Essay on the logical structure of the sentence]* (1929, 1950), and, Georges Galichet who published in 1970 a *Grammaire structurale du français moderne [Structural Grammar of French]*

94. In French "*la structure profonde de notre langue.*"(George Galichet, 1949).

the linear order into the structural one." in Elements (1959:19)[95]. We can further read that "and once disposed in the linear order in the spoken chain the sentence is ready to receive the phonological clothing that will give it an exterior form." (p. 34). These short extracts on the conversion between linear order and structural one from Tesnière's *Elements* offer an interesting reflection on some theoretical implications linked to representing the sentence-unit in a syntactic-tree that is hypothesized to represent this virtual deep structure, and this not only within the generative framework.

Furthermore, Tesnière developed in his linguistic theorization a system of symbols capable of "representing each type of lexical word to express their deep nature without retaining the accidental contingencies" (*Elements*:63). Substituting these symbolic representation to the real words in a stemma will generate a *virtual stemma* like in Figure 2.9 (B). Note that in explaining this substitution to generate virtual trees, Tesnière will refer to algebraic calculus: "*As algebraic method allows to generalize the solution to quantity problems, giving general formulas of solution's types, using symbols will permit us to generalize grammatical questions substituting general formulas of sentence types to the infinite multitude of particular sentences.*" (*Elements*:65). Hence, this 'virtual' underlying structure is already understood as having a computational relevance and as possibly generating expectations (cf. Top-down expectations, p. 104). Moreover, his conception of Grammar and of its underlying rules, led him to theorized the "*Zero marker*" indicating places in the sentence where syntactic facts are not associated to overt morphological marks (see an example of Zero marker in Figure 1.6,p. 42).

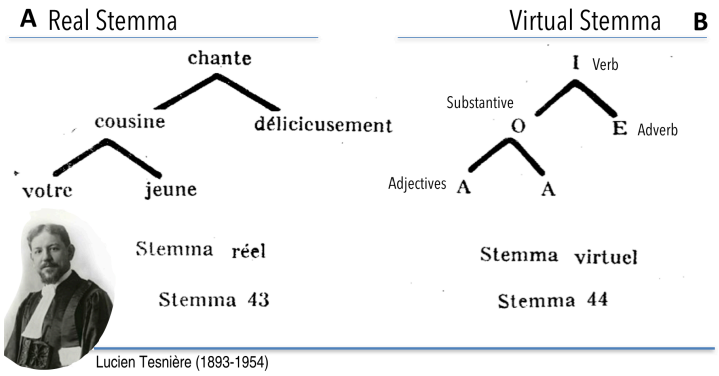


Figure 2.9 – (A) Real stemma of the sentence "*Your young cousin sings marvelously*" (B) Virtual stemma. 0: Substantive, A: Adjective, I: Verb, E: Adverb. Adapted from Tesnière *Eléments de syntaxe structurale* 1976, p. 64.

In conclusion, we see that his approach to the study of Syntax was already gathering two fundamental ingredients for our experimental work on syntax: the focus on language as a system, and the mentalist and cognitive focus.

It is important to note for our theoretical positioning, that while Tesnière trees focused on representing sentence-internal dependency relations and Hockett's boxes focused on the intermediate building steps of constituents structures of the sentence, linguistic trees from the generative tradition have the advantage to incorporate these two main aspect that syntactic theories to to explain how sentences are formed. These two aspects lie at the core of neuro-linguistic interrogation today and in our research project, namely:

1. the Rules and processes that determine basic sentence structure, and of

95. "*De ce point de vue nous pouvons dire que parler une languge, c'est transformer l'ordre structural en ordre lineaire, et inversement que comprendre une langue, c'est transformer l'ordre linéaire en ordre structural.*" L. Tesnière in Elements (1959:19, chapt.6.4).

2. the Dependency relations within a sentence.

As we already had the occasion to point out in our Preliminary Considerations in Part , Generative studies not only provide a formalized system for dependency relations within a sentence, but provide some computational principles for such a formalized system[96], by positing very essential building processes, like Merge and Move [97].In what follows, our aim is less to give a complete survey of a particular linguistic theory, than to inform the reader about how this given theory actually provided us with useful descriptive tools (i.e. syntactic-trees) and a set of possible mechanisms determining sentence structure (Merge, Move, Binding etc.), which in turn will represent testing hypotheses for the investigation of the cerebral encoding of sentence structures.

We can here state that the central methodological issue in this first chapter is to familiarize the different readers to the hierarchically built linguistic structures to be able to ask a now classical question: Do the hierarchical structures postulated by Linguistics correspond to actual representations used in the brain during real-time language production or comprehension?

The representations and experimentation on the concept of sentence carried over the last century will be now presented with special attention to the (1) internal structure of the sentence and to (2) its Building process, the *'famous'* Merge and Move.

2.3.1 Representing Sentence in a tree: Hierarchy and Linearity

We concentrate here on Sentence as a hierarchical structure and on how our understanding on the neural implementation of syntactic structure can benefit from representing sentence structure with the tool we just we just mentioned above: syntactic-trees.

This syntactic representation together with its cognitive and computational assumptions have been considered so fundamental in defining human linguistic ability, that recently Techumseh Fitch – a researcher working at the interface between Cognitive Neuro-science, Primatology and (Computational) Linguistics – has argued that humans have a species-typical inclination to infer tree structures in sets of strings in multiple cognitive domains, that he calls *dendrophilia* (Fitch, 2014). He seeks to characterize the cognitive difference between humans and some other species in terms of the class of models that man can build to make sense of stimuli, and proposes the Dendrophilia Hypothesis in the following terms: "*Humans have a multi-domain capacity and proclivity to infer tree structures from strings, to a degree that is difficult or impossible for most non-human animal species*"[98]. We can note here how the previously defined Homo

96. For an articulated discussion see Rizzi (2013) in *Lingua*, where the author speaks both to linguists and cognitive scientists.

Focusing on the hierarchical nature of syntactic structures, Rizzi illustrates some formal properties and interpretive consequences of structural representations, and highlights the importance of the two fundamental computational mechanisms that natural language syntax requires: (1) a recursive structure-building operation, generating sets of hierarchical representations, and (2) an operation establishing dependencies between more or less distant syntactic positions,(i.e. Move), that is necessary to express several form-to-meaning characteristics of the sentence-unit.

97. Or, one could also imagine to focus on the minimalist Agree and Match mechanisms. We will leave this idea for further investigation. We can anticipate that the minimalist linguists reading chapter 7 will recognize that a certain condition contrast could be interpreted in terms of Match and Agree.

98. We can further read the following: "*To the extent that this hypothesis is correct, it provides a unified computational account of the difference that allows our species, and not others, to acquire language: that we infer trees over linguistic or musical strings, and that this allows us to compute probabilities and infer higher-order rules that would be essentially invisible to a species that does not do so.*" (2014:352).

Phraseologicus emerges again (cf. §section 1.1.3, chapter 1) showing a deeply rooted tendency to automatically interpret sequences of linguistic stimuli in hierarchical ways compared to other species[99].

2.3.1.1 Sentence is hierarchical: X-bar Structure vs. linear strings

Before entering into the details of what the modern versions of syntactic-trees, the problem of representing hierarchical, embedded syntactic structures is not new, and needs to be first addressed at a level that is inferior to sentence-level constructions we will investigate in this research project, that of constituency. Namely, the issue of sentence structure started to be investigated in the 50's and before by first claiming that grammatical phenomena work in terms of constituents as we saw with early attempts to represent sentence constituents (cf. Hockett's Box §42), and later evolved into a syntactic theory whose aim was to elaborate a phrase structure schema that would be common to all phrases in in the 1980's: the X-bar theory[100]. This step forward that this theory brought in the linguistic field is very important for our concern for syntactic structures and their cerebral implementation. Namely, X-bar theory first offered a system of structure projecting common to more than one Lexical category[101], and, in this way it paved the way towards the formalization of a common structural model that could underlay the building of syntactic hierarchical structures (cf. §2.3.3.1, p. 140).

Let us consider the structural scheme in Figure 2.10. Each phrase of category XP is assumed to be structured around a Head (X) that projects a phrases from the lexical entries of category X. Such phrases contain intermediate constituents that are projected from a head X, namely an X-bar (X'), and one optional specifier and a complement.

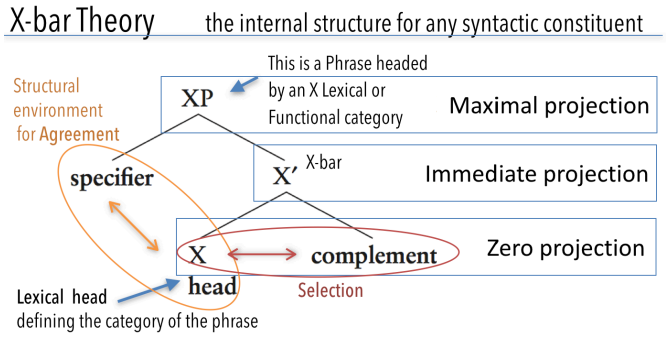


Figure 2.10 – The basic X-bar schemata commented.

This structure call for a hierarchical organization that allows to identify relations in the structure by using the notions of sister and daughter, defined in terms of *immediate domination*. Hence, we can say (1) that A is a sister of B if A and B are both immediately dominated by the same node and that (2) A is a daughter of B if B immediately dominates A. This structural position allow to define the following three grammatical relations Figure 2.10:

(i) A is a specifier if A is a daughter of X";
(ii) A is a complement if A is a daughter, but not a sister, to X'; and,
(iii) A is an adjunct if A is a sister and daughter of X".

99. As a side note this idea was already present in George Miller's (1920-2012) investigation of language (cf. Grammarama project) and T. Fitch extends it to other cognitive domains like music.
100. "X-bar theory evolved from Chomsky's (1970) Remarks on Nominalization but was restricted initially to apply only to the lexical categories (which, for Chomsky, were the four combinations of the features [±N] and [±V]). Inflectional categories such as tense and complementizers remained outside the X-bar system. Things changed with the publication of Barriers, (Chomsky 1986), which opens with an explicit proposal to generalize the X-bar system to Infl(ection) and C(omplementizer)." From Shlonsky (2010) "The Cartographic Enterprise in Syntax". In Language and Linguistics Compass.
101. Next session retraces how this applies for Functional heads as well.

For all lexical categories, Nouns (N), Verbs (V) Prepositions (P) or Adjectives (A), the structure of the phrase consists of an X' (X-bar) syntactic projection (X' or X-bar) that is represented by means of a layered representation, in which the intermediate projection X' between the head and the maximal projection can be iterated X, X', X". We refer to these different levels as Phrase-level (XP), Bar-level (X'), and Head-level (X)[102]

We present in Figure 2.10 the basic X-bar schemata, where X' and XP (or X", X double bar) are called projections of the head X, which is consequently called the 'zero projection', also written X°. The sister of X is called the complements of the head X, and the sister of X' is the specifier of the phrase.

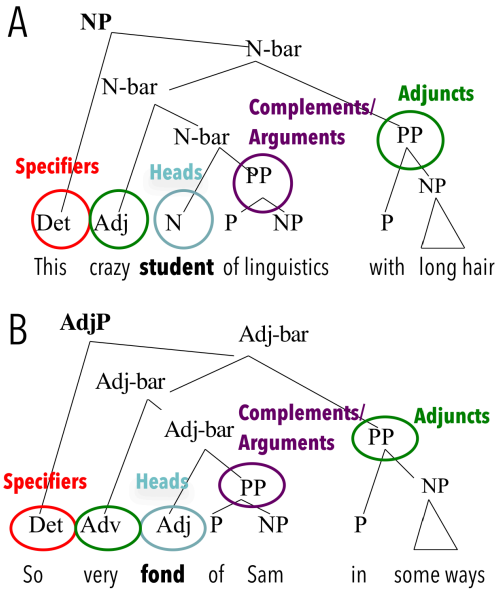


Figure 2.11 – Xbar representation of two different Phrases, (A) Noun Phrase NP; (B) Adjective phrase, which share nonetheless the same tree-like format structure. This representation is a first approximation to show the analogy of structure, we use these label just to draw an example and will not use this notation for the following.

According to the principles of Xbar-theory, all other elements within the XP besides the head are structurally optional, their presence being determined by principles of licensing, like theta-theory or case-theory. Oversymplifing for expository reasons, wa may say that the main relations identified inside this structure are the *head-complement* relation, as the locus of Phrase selection, and the *specifier-head* relation as the structural environment for agreement. Moreover it should be noted that head complement order can be parametrizable according to cross-linguistic variation along head-directionality (head-final versus head first languages). All in all, the X-bar schema achieves a productive and simple generalization of the internal structure of any constituent. According to this theory all lexical categories project the very same structural skeleton).

All these assumptions about the internal structure of constituents, clearly show that this syntactic theory obeys to a general principle. This general principles requires that (1) a complex expression is built up from the meanings of its constituent expressions and from the way they are combined. Moreover, (2) this theory determining constituency, specifies only immediate dominance, and by this introduced configurations where it is possible to distinguish between structural relations that are more or less local. In sum, Dominance and Sisterhood of the different nodes reflect a fundamental hierarchical structural organization.

To illustrate the fundamental assumption of uniformity of syntactic structures represented in a tree-like format, we can observe in Figure 2.11, that Phrases in (A) and (B) share the same pattern structure despite their being the projection of different Lexical Heads, receptively Adjectival head for (B) and Nominal Head for (A) (in light green). If we consider the grammatical relations between their constituent, their internal structural 'deep' organization is the same. We can recognize *Specifier* role is occupied by two distinct determiner 'So' and 'This' (in red), while the adjective 'cray' and the adverb 'very' are in the same structural position as Adjuncts (being the sister and daughter of X"); and the two prepositional

102. Note that in this framework, the value of X ranges over at least the four categories mentioned above, the so-called lexical categories, that can therefore be characterized in terms of [±N] (substantive) and [±V] (predicative) features. In this way it becomes possible to characterize natural classes of syntactic categories.: N [+N,-V], A [+N,+V], V [-N,+V], P[-N,-V].

phrases in purple are in Complement/Argument position of their Head in both structures. From the example in Figure 2.11 we can realize how representing the internal structure of these two phrase helped in grasping the fundamental structural similarity between (A) and (B)[103].

We will retain, then, that such a tree-like representation helps capturing the purely abstract syntactic relations of the phrase, that wouldn't be perceived just by focusing on grammatical categories of the function words present in the sentence.

This argumentation is essential to ground theoretically the several attempts that have been undertaken in neuro-imaging to correlate syntactic-tree complexity measures with brain activity, as we will see in the following sectio, §2.3.3.2 on the Neuro-imaging of syntactic-trees.

2.3.2 Minimal syntactic structure combinatorial operation

2.3.2.1 All we need is MERGE ?

The way of understanding the relation between form and meaning, already present in X-bar theory, lead the Minimalist program (Chomsky, 1995)[104] to introduce an extremely simple sentence building mechanism: *Merge*.

Merge is assumed to be the basic structure building operation that takes a pair of computationally well formed objects, A and B, and forms a new object C:{ A,B} (Chomsky, 1992:2; 2001:2).

Elements are merged in a Spec-Head agreement relation in which the nature of specifiers is determined by the nature of the Head (Chomsky 1993). Depending on the nature of the two merged elements, we can distinguish three main sub-cases of Merge:

1. Head-Head Merge [X - Y] where the two elements are drawn from the lexicon (configuration of mutual c-command)
2. Head-Phrase Merge [X - YP] where a phrase already formed by a previous application of Merge is subsequently merged with a head drawn from the lexicon (the label of the head is given)
3. Phrase-Phrase Merge [XP - YP] where a phrase already formed by a previous application of Merge is subsequently merged with a head drawn from the lexicon (the label of the head is given)

Importantly, the merged elements X and Y can be either two elements taken from the lexicon or complex expressions already formed by previous application of Merge. This principled way to describe the emergence of linguistic structures is a computational operation (assumed to be a binary function) that constructs new syntactic objects by recursive merging of words and phrases together, creates the hierarchical structure of the sentence giving rise to the tree-like representation of syntactic binary-branching trees. This central property of the generative procedure is called *recursivity*, it allows namely an operation to reapply to the result of an earlier application of the same operation (i.e.

103. The notation used here is an old one,(i.e. before Abney's notation), however the point we want to make here is strictly about the structural tree-like configuration that the X-bar schema offers and represents.

104. This shouldn't be taken as a historical remark, in fact already in 1957 the book by Chomsky 'Syntactic Structures' contained what we could call his own version of the structuralist Immediate Constituents analysis, the Phrase structure Grammar that will lead to 'Remarks on Nominalization (1970) a corner stone for X-bar theory.

its own output). Compositionality and generativity being among the first characters of natural language that concerned theoreticians of human syntax, Merge operation became central because of its essential simplicity, determining nevertheless a hierarchical structure [105].

Turning now to the structure of the resulting linguistic object. While in classical X-bar theory labels categories were determined by the X-bar schema where a Lexical head entering syntax would automatically generate its own projection, in a Merge-based approach the standard assumptions about how labels are assigned to the new node created by Merge is that this new entity C should have, minimally, the form A, B where A and B are the merged constituents and is an identifier, called *label*, that would express the category which C belongs to (Chomsky, 1995:243)[106]. Importantly, each element and each merger involves an interpretation both phonologically and semantically.

2.3.2.2 Mergeability and Feature-matching

In the simplest terms, Merge forms a set out of two objects: it basically takes a pair of computationally well formed objects, A and B, and replaces them by a new object C[107].

The basic syntactic structure building operation is not only a procedure that takes two units (words, phrases, clauses, ...) and form them into a single unit, but most importantly it is subject to 'feature-matching', which could be expressed as a matching with what "*the word is seeking to combine with*".

Given this, linguistic symbols are considered as feature sets composed by phonetic, semantic and formal (selectional, categorial, case etc.) features (Chomsky 1995:21-22), as illustrated by Figure 2.12. Most of the current generative frameworks express crucial properties of the linguistic objects by using features, like shown in Figure 2.12(B) for the word 'dog': features determine the position that an object can occupy by triggering its lexical insertion. In Figure 2.12(A), sentence (a) vs. (a') show how a transitive verb selects an argument, while (b) shows a displacement, where the focalized elements 'MARY' stands at the beginning of the sentence rather than right after the verb that requires it as an argument. While the agrammaticality of (d), (e) and (f) shows examples where feature-matching failed for different reasons. Features in 2.12 (c) provide essential instructions for the performance systems about the word '*Mary*', which in turn shows the importance of the two different way of representing the features of the determiner '*the*' that are shown in Figure (C)[108].It is also important to stress that Merge, as a core building operation, not only associate element after feature-matching, but also determines the feature structure of the resulting object.

As an example of the series of combinatorial processes allowed by Merge consider the following steps to build the simple sentence '*She likes Christian*':

105. For more on the concept of recursion see Watumull et al. (2014).
106. For discussions about the importance of labels in syntactic structure-building computations from the point of view of the Cognitive Neuro-science of Syntactic Structure Building, see Sprouse and Hornstein (2015), and labeling approach presented in Chomsky (2013) with Rizzi's (2015) 'Notes on labeling and subject positions' (already EALING course 2012) from the point of view of Linguistics. This is a hotly debated topic even from the computational point of view. For an overview of the problem, see Stabler (2013).
107. Note that merge, in language, is assumed to be a binary function.
108. For more details about this fundamental difference between a representation where only the lexical entry is specified from a representation where both lexical entry and Principle are represented see Cristiano Chesi PhD (2004) and following articles.

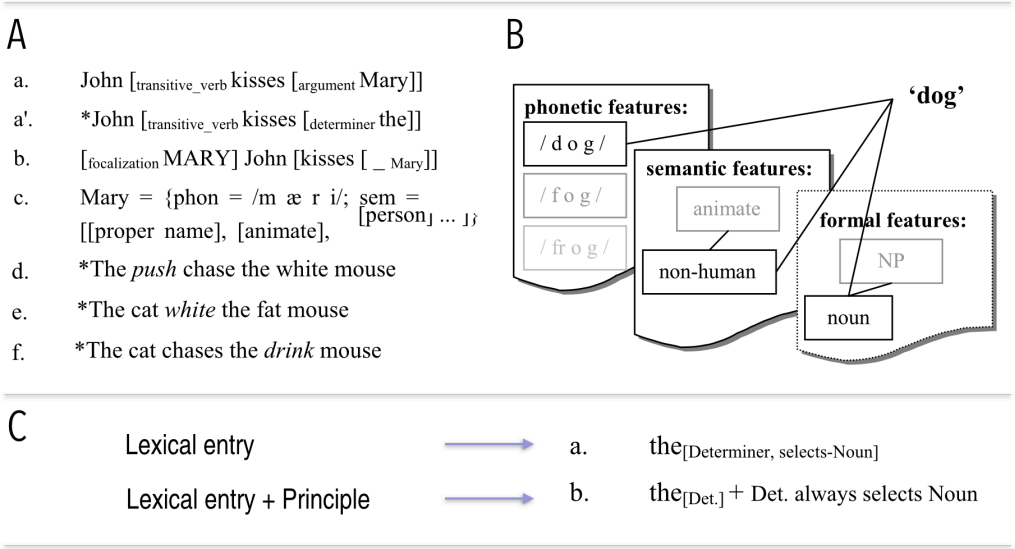


Figure 2.12 – Feature matching and merge-ability. (A): (a) vs.(a') show transitive verb's argument selection; (b) shows the feature of the focalized elements 'MARY' standing in sentence initial position rather in post-verb object-position; (c) provide essential instructions (i.e. features) for the performance systems about the word '*Mary*'. While (d), (e) and (f) shows examples where feature-matching failed for different reasons and their consequent agrammaticality. (B): Features for the word 'dog'. (C): the difference between representing the features of the determiner '*the*' in the lexical entry (a) or as a lexical entry and an associated Principle. Adapted from Chesi PhD (2004).

1. The pair [likes + transitive] is satisfied by a Theme argument in object position
2. The noun [Christian] looks for a verb that requires an argument
3. They match and [like + Christian] is built, after this
4. The new unit [like***s*** Christian] requires a subject that is a third person singular
5. the pronoun [she] supplies this
6. And [she] is merged with [likes Christian]
7. to give -> [she [likes Christian]]

From this process we can see that merge happens if and only if:

1. a relevant local configuration is met among elements that are combined;
2. the feature structures of these elements is compatible.

This second point in (2) is essential for our understanding of a great number of experimental designs in Neuro-linguistics manipulate this feature-matching aspect of Merge, among others Pallier et al. experimental design can be viewed as a direct example, it manipulated the mergeabilty of words to parametrically manipulate the size of the constituents of experimental stimuli see Material reproduced in Figure 1.15 (p.66).

This minimal structure building has been attracting much attention in neuro-linguistics in the last fifteen years, because it has been identified as the generative engine or operation enabling composition (cf. "the Merge Quest" in Broca chapter 1 §1.4.5, p.61)[109],

109. We wont enter here in the debate arguing to ultimately reduce language faculty in the narrow sens to a recursive computational mechanisms as the only component characterizing human species. For more on this topic see Hauser et al. 2002 (Science).

although it wasn't originally designed for this purpose, sentence construction via Merge is a plausible and 'economical' process that could model how the brain may represent the process to build sentences[110].

2.3.3 Syntactic trees and cognitive processes

2.3.3.1 Expanding X-bar system to the sentence

So far, and assuming a rather general approach, we have considered how X-bar theory formulated the part of grammar regulating the structure of phrases, where a head (X°) could belong to lexical categories such as N (noun), V (verb), A (adjective), P (preposition), developing in this way the central hypothesis that all the phrasal categories are structured following the X-bar format. However, nothing has been said about the larger unit of syntactic analysis, the sentence.

A further step in this direction, will lead us to integrate (S) into the X-bar system, by stipulating the projection of functional heads too. In this way, the idea that grammatical categories are constructed according to a fixed template and structural configuration is to be extended to sentence structure.

Hence, a whole array of Functional lexicon consisting of grammatical words, morphemes, such as determiners, complemetizers, auxiliaries, copulas and expressions of tense and aspect, etc. -a range of elements having a more abstract semantic content- entered the X-bar system to define a *configurational structure* in which the contentive elements could be inserted[111]. The underlying assumption being that, like lexical heads (N, A, V and P), these functional heads would have a syntactic projection too.

Hence, functional categories like I (Infl/IP) and C (COMP/CP) could now also to be values of the head X. And, the sentence (S) is therefore reinterpreted as a projection of Infl/Inflectional Phrase (IP), being the functional node that dominates the inflectional morphology of the verb, affixes, infinitival *to*, and tense and/or agreement features, which are not independent lexical categories or even 'words'.[112].

Similarly, S' is also reinterpreted as a projection of C(OMP), the Complementizer Phrase (CP) to host complementizer of embedded clauses, wh-phrases[113]. Analogously, the functional category Determiner (Det) can be incorporated into the X-bar format system as Determiner phrase (DP)[114].

The role of the Complementizer Phrase (CP) Before moving forward, we briefly illustrate the role of the Complementizer Phrase, because it will be our loyal companion until the

110. As a side remark, here my supervisor C. Pallier would say: 'Remember that this could be also be implemented by *reduce* in parsing', but discuss this is will carry us to far from our purpose and I let him defend his purpose in a future article.
111. Contentive lexicon, consisting of nouns, verbs adjectives, etc.that are endowed with descriptive content characterizing events, arguments, qualities, etc.
112. Infl: a functional head containing (in English) auxiliary verbs and/or tense and/or agreement features. Infl was then reinterpreted as a conflation of two separate heads AGR (Agreement) and T (Tense). It is also written as I (I°). Chomsky (1981), Pollock (1989).
113. As for COMP it was originally assumed that wh-phrases such as who in 'who did you see?', would be in the COMP-position in order to explain why, in English, the presence of a wh-phrase usually excludes the presence of a complementizer. The standard assumption now is that COMP, or rather C°, heads its own syntactic projection -CP- and that wh-phrases, relative pronouns and other preposed material are in the specifier position of CP.
114. In this way, NP is reinterpreted as being part of a DP.

last pages of this manuscript. Consider the two sentences:

(38) Role of the Complementizer Phrase

a. a. I will ask [***if*** [Tristan will change his mind]].
b. b. I will say [***that*** [Tristan will change his mind]].
c. c. Which point will Tristan will change his mind on ___?

In (38), we can see that the nature of the sentence unit as a whole -*its sentence type*- is determined by the nature of its complementizers: (a) is interrogative while (b) is declarative. This difference is given by complementizers introducing the bracketed complement clause *if* and *that*, which have different features. The notion of features, as we already saw, refers to the properties of lexical elements, and one of the properties of the item '*if*' or '*which*' in example (38a) and (c) is that they contain wh-features. It is further assumed that they have a complemetizer head C that bears a [+WH] feature, marking the clause as a wh-question in the case of '*if*', and a [-WH] feature in the case of '*that*'.

In brief, the Complementizer phrase is the Functional Projection hosting information about what is the *sentence type*. In theoretical linguistics, it is indeed generally assumed that every sentence has a sentence type an therefore a CP. Every sentence thus contains the C-position, even though this position is not always filled by a lexical item. We will come back to this central aspect of the 'sentence skeleton' at the end of this chapter from both a linguistic and neuro-linguistic point of view.

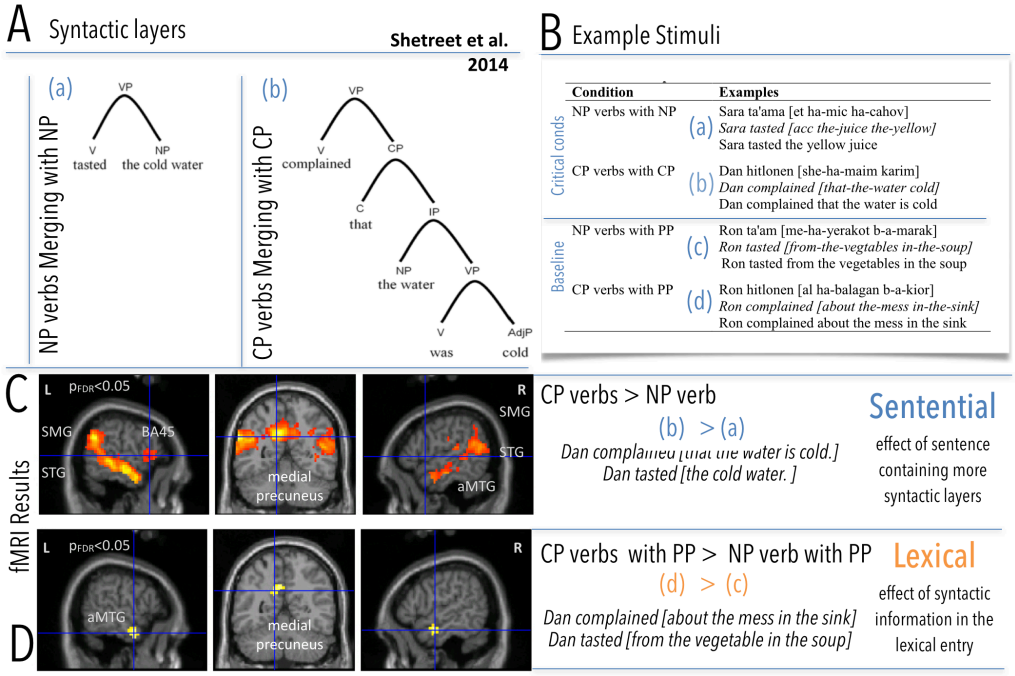


Figure 2.13 – Adapted from Shetreet et al. (2014).

As for the neural underpinnings of this sentential layer, a recent neuro-imaging study by Shetreet and colleagues lends experimental support to the theoretical distinction be-

tween functional projections and phrasal projections we are now considering. From Figure 2.13(B) reports the experimental stimuli the authors selected. Two types of sentences where verbs where either selecting a Noun Phrase (a) or a Complementizer Phrase (b) were contrasted, and the effect of these two different sub-categorization was obtained by contrasting them against a baseline sentence where the same verbs select Prepositional Phrases as illustrated in (c) (cf. example stimuli Figure 2.13(B). fMRI results in 2.13 indicate a radical difference in the extent of activation between the two experimental conditions: verbs selecting complementizer phrases compared to those selecting NPs activate the quasi totality of the sentence processing network in a very bilateral fashion, as shown in Figure 2.13(C).

The portion of brain areas recruited by this broad sentential effect can be interpreted as revealing the complexity of representing the extra sentential layers shown in Figure 2.13 (Ab), that a verb selecting a whole sentence as its argument could elicit compared to the syntactic layers in (a). In (D) we observe a spatially restrained activation in the temporal region (aMTG) bilaterally for the contrast opposing the two verb types when both were selecting a Prepositional Phrase [(d)>(c)]. This is understood by the authors as linked to the richer syntactic information contained in the lexical entry of verbs sub-categorizing for a complement sentence. One could indeed speculate that this area (aMTG) stores the verbal information about the possible sub-categorization frames for Verbs that can both select CPs and PPs frames. In sum, this difference in the extent of recruitment of the sentence network is a first evidence for an increase in activity related to the presence of an extra Complementizer syntactic layer adding complexity to sentence.

Going back to our presentation of the extension of X-bar Theory to Functional Categories, we can state that the main justification for taking these functional elements as heads is that they determine the syntactic distribution of the sentence. This shift from emphasis given to the contentive lexicon (N, V, A,...) to the functional lexicon (D, Aux, C, T, Asp,...) has in fact the advantage of create *a default configurational skeleta* for the insertion of contentive elements and trigger the fundamental computational processes of the sentence, while being also the locus for expressing basic parameters of variation across languages. As we will present in the next sub-section these functional elements give rise to complex configurations, studied in "cartographic" projects (Rizzi, 1997; Cinque, 1999).

Yet, if one should ask how to look at a syntactic-tree, we could first intuitively answer that the position in the syntactic-tree gives a syntactic 'relational closeness' measure of words. While if one considers only linear-order, in (b) compared to (a) [*Murielle* and *says*] are adjacent, which could give evidence of subject-verb agreement, which is true only for (a) but not for (b). In fact, in the hierarchical structure of sentence (b) 'Murielle' and 'says' are far apart and not in a subject-verb relationship (in blue). Two words exclusively dominated by a single node are in a sisterhood relationship (see orange square), but '*Murielle says*' linear sequence is structurally a remote relationship, since the only node that dominates both is the node that dominates the entire sentence (S) (in blue).

Some empirical results confirm that the relational distance represented in the above syntactic tree and by the boundary brackets between the linear sequence "Murielle]]] [says", has a psychological reality in linguistic behavior and sentence processing. We will later introduce them. Meanwhile, we can cite Friedmann and Grodzinsky (1997), who

observed that agrammatic patients have difficulties with embedded sentences even in a basic repetition task.

We anticipate here some empirical results from two classical psycho-linguistic methods used to investigate on-line sentence comprehension: one measuring the reading time (RTs) associated to each word, using either eye-tracking devices during natural reading (Just and Carpenter, 1980; Rayner, 1998), and the other using reading paradigms where the reader can control though a button press the rhythm of words on the screen (Mitchell, 2004). For instance, Holmes et al. (1987) observed a slowing down in reading times at the opening of a Complementizer Phrase compared to that of a Noun-Phrase, which would correspond in the Figure 2.14 to the place where '*who*' is found, this place namely represents the moment in the tree where the embedding of the relative clause starts. Other works using these methods confirmed that the processing time of a given word is influenced by factors linked to the syntactic-tree structural complexity of the sentence (Gibson, 1998, Dependency Locality Theory (DLT) in terms of intervening discourse referents; Roark et al., 2009; Staub and Rayner, 2007).

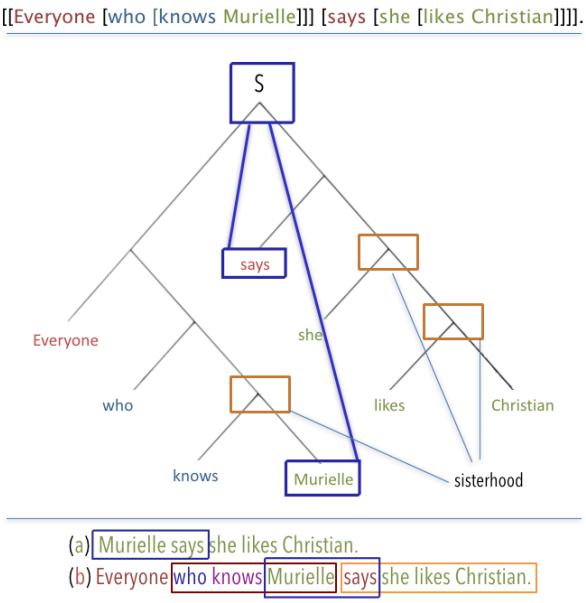


Figure 2.14 – An example showing how to look at a syntactic-tree and the syntactic relationships it represents. In sentences the words *Murielle* and *says* (a) and (b) do not differ in linear distance but do differ in hierarchical distance represented by the syntactic-tree diagram (see blue circled tree-node dominating them), while words exclusively dominated by a single node are in a sisterhood relationship (see orange square).

The advantages of considering the sentence structure in a tree

Before moving on to more complex syntactic-tree representations, a distinctive characteristic of this way of representing sentence structure should be noted. Tree-like representation characteristically distinguishes grammatical categories from the *grammatical relations* (e.g. subject, object) that are defined by their position in the syntactic-tree. An element is assigned a given category depending on the nature of its Head, while its grammatical relation is determined by the tree configurations, that is to say depending on its structural position in the X-bar schema. Thus, an NP will be a subject if it is immediately dominated by a S (I) node, and an object if it is immediately dominated by VP node.

Hence, *syntactic closeness* is expressed in hierarchical terms, simply by using the notions sister and daughter of a tree-like representation: the syntactic relations can be defined in terms of immediate domination, as we did in our example to show how 'Murielle and 'says' were not in a verb-agreement relation although being adjacent. This aspect was clearly represented in the position of non-immediate dominance they occupied in the syntactic-tree in Figure 2.14 (see the blue circled items).

We can conclude that two fundamental types of information about sentence syntactic structure can be extracted from a tree-like representation: the first is *dominance*, and the second is the *precedence* among the lexical items forming a sentence. As illustrated in Figure 2.15:

1. *precedence* represent the spoken linearity aspect of the sentence and is defined on the terminal nodes of the tree (i.e. tree leaves); and
2. *dominance* expresses dependency and constituency relations among both the pronounced words and the empty abstract elements we are soon going to introduce. This dimension fundamentally expresses the hierarchy in the sentence, where dominated elements are constituent parts of the dominating one.

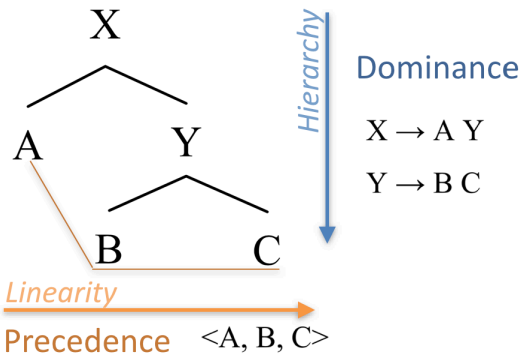


Figure 2.15 – Schematic representation of how the syntactic-tree representational format formalizes the Linearity and Hierarchy dimension characterizing the sentence-unit. Adapted from Chesi (2004).

These two relations represented in tree-structures synthetically constitute the most relevant relations among elements of a sentence[115], to which one should probably add a more complex relation called C-command (i.e. Constituency-command) that has a pervasive use in many syntactic relations (e.g. binding conditions, government relations, scope constraints)[116].

Given this background, we can reformulate the linearly long-distance relations in the sentence, forming the so-called discontinuous constituency relations à la Hockett[117] as occurring when a dominance relation, but no precedence relation is defined between two elements.

From a more general point of view, one could argue that such elementary relations like 'linearization' (implied in precedence dimension), and dominance could have interesting analogies in many cognitive domains, which could justify the interest from the point of view of Cognitive Neuroscience to use tree-like representations to investigate such a high-level cognitive faculty like syntax in human language.

In conclusion, tree-like representations crucially contribute to determining properties of form and meaning of linguistics expressions (e.g. what word is related to other words, what constituent units are formed, etc.), capturing both their *linearity* and *hierarchic organization*.

From what we have said so far, it might seem plausible to conclude that syntactic-trees appear to offer interesting tools to reveal and quantify syntactic relations in hierarchical terms, like we just saw for the metric of *syntactic closeness*, defining a constituent as a string of words such that there is one node that dominates those words and no other words.

115. They are considered in Computational terms as being Primitive in nature.
116. Constituency-command has received different subsequent definitions, and the literature on this is extensive and beyond the scope of this section. C-command is a particularly important syntactic relation, which provides a useful way of determining the relative position of two different constituents within the same syntactic-tree: whether one is lower in the tree than the other or not. This relation can be informally defined as follows (where X, Y and Z are three different nodes). C-command A constituent X c-commands its sister constituent Y and any constituent Z which is contained within Y. From Radford (2009) in *Minimalist Syntax Revisited*.
117. Remember the example of Hockett's boxes featuring such constituent discontinuities by longer lines cutting down some layers of its graphical representation, see §1.3.3, p. 42.

2.3.3.2 Neuro-imaging of syntactic trees

Before turning to computational considerations on syntactic trees, we will briefly present two neuro-imaging studies conducted in Japanese and English that were among the first to explicitly address the question of the neural underpinning of sentence featuring two different sentence-internal dependency configurations, represented by two different syntactic tree-shapes at the end of the '90s.

Seminal work from Stromswold and colleagues (1996) showed in a PET study that sentence structural complexity had an effect in the Broca's complex (Pars opercularis, see Figure 2.16 (C) to the left). In this study Right-branching and Center-embedded structures were selected for comparison, as illustrated in Figure 2.16 (A), the shape of the syntactic-tree of the two conditions differed in the two tested sentence structures.

Soon after, Inui et al. (1998) tried to identify the mechanisms directly related to the processing of sentence whose syntactic tree-shape and consequent levels of embedding differed in a much more controlled manner. The authors constructed in Japanese two experimental conditions using the exact same sets of words to obtain a center-embedded and a left-branching syntactic-tree that were also sharing the same matrix clause, which was not the case in the English PET study where "The child spilled the juice" and "The juice stained the rug." were compared.

Namely, one clear advantage in using Japanese for the kind of contrasts that Stromswold et al. used is that this language allows the matrix clause and the embedded one to be the same in both syntactic configurations. Matrix clause shown in Example (39) are indeed identical: "Taro-ga oshita".

This was not the case in the English study, where the content of relative clause and matrix clause is inverted across the two experimental conditions, resulting in a difference in assertion and presupposition -a parameter that was confounded with the manipulated variable of sentence structure by Stromswold et al.[118].

(39) Japanese

a. left-branching sentence

[Masaru-o taoshita Hanako]-o] Taro-ga oshita.
Masaru-ACC. knocked.down Hanako-ACC. Taro-NOM pushed

'Taro pushed Hanako, who knocked down Masaru.'

b. center-embedded sentence

Taro-ga [[Masaru-o taoshita Hanako]-o] oshita.
Taro-ga Masaru-ACC. knocked.down Hanako-ACC. pushed

'Taro pushed Hanako, who knocked down Masaru.'

Hence, as can be seen in (39) the two experimental conditions (a) and (b) differed only in the word-order (reflecting their syntactic structure) with no semantic overall interpretative 'inversion'[119]. fMRI Results show among other things[120] a difference in

118. It has to be noted that these two studies used very different tasks Stromswold et al. used semantic plausibility judgments, while Inui et al. subjects' task was to understand the relationship among three characters mentioned in the sentences.

119. Nonetheless it has to be noted that the left-branching condition (a) presents a word-order putting the subject in a position that is interpreted as a focalization by Japanese speakers. It could therefore be translated into English as follows: '[It is] TARO [that] pushed Hanako who knocked down Masaro', Thomas Pellard (p.c), which, given the task used in this study would lead nevertheless to an easier assignment of thematic roles despite the understood focalization.

120. Significant activation of Posterior part of the frontal lobe (BA6/9) and the inferior parietal area

activation of Broca's area (BA44 and 45) between the two conditions in favor of center-embedded condition. This finding suggests that the different shape of the syntactic-tree and the correlated hierarchical relations (i.e. embedding) elicit syntax-related processes in these areas.

By directly comparing these two studies, we can observe something we already noted in Figure 2.11 (p. 136), where the same shape/configuration of two syntactic-trees was representing that two different Phrases (i.e. an Adjectival phrase and a Noun Phrase)[121] shared the same structural relations. Similarly, in the manipulation of these two brain imaging studies, it is the structural relation reflected by the similar syntactic-tree shape that matters. The graphical comparison between the syntactic-trees tested in the two studies in Figure 2.16 (C) (blue for English and orange for Japanese) show that Japanese trees are the mirror image of English ones. This 'mirror image relation' between the syntactic structure of the two languages is accounted by a very well-known Parameter defining head directionality, opposing them along the fact that inHead-final languages complements precede their head, while in head-initial languages the head precedes complements.

Given these remarks, we can advance the argument that actually these two experiments tested tree-shapes that share the same relational difference in Japanese and English as they are one the mirror image of the other. By saying this we actually speculate that this is what can account for the fact the two brain imaging results similarly indicate the involvement of Broca complex, and that this brain region might be involved in the representation of the syntactic-tree configuration of the sentence structure.

So all in all, these two early neuro-imaging studies offered a cross-linguistic confirmation that the hierarchical sentence-internal relations that the tree-representation offers are interesting tools to investigate the neural correlates of syntactic processes in different sentence structures and hierarchical relations.

Moreover, we hope by this example to have shown how using a specific language to tease apart confounded experimental variables to investigate sentence complexity can be a clear methodological advantage in a neuro-imaging study. This kind of methodological approach to cross-linguistic diversity will indeed be put to practice in Part II.

We conclude by saying that these linguistic and experimental examples lead us to the following consideration: the most significant fact about syntax is its hierarchical structure, and syntactic-trees reflect how sentences are constructed in an "experimentally handy manner".

2.3.3.3 Sentence's trees and computational process

Beyond the diversity of theories of syntax in formal linguistics, and inside the generative framework, there is a common assumption central to the different approaches (e.g. the derivational theory, the optimality theory of syntax, unification grammars, etc.): explicit models of natural languages should go beyond description of linguistic behavior and *theorize the mechanism that organizes these data*[122].

This common bases can be traced back to the famous distinction between linguistic Competence and Performance in language: Competence corresponding to the knowledge

(BA39/40) were also reported.

121. Namely, Noun Phrase: This crazy student of linguistics with long hair. Adjectival phrase: So very fond of Sam in some ways.

122. see Miller and Chomsky (1962)

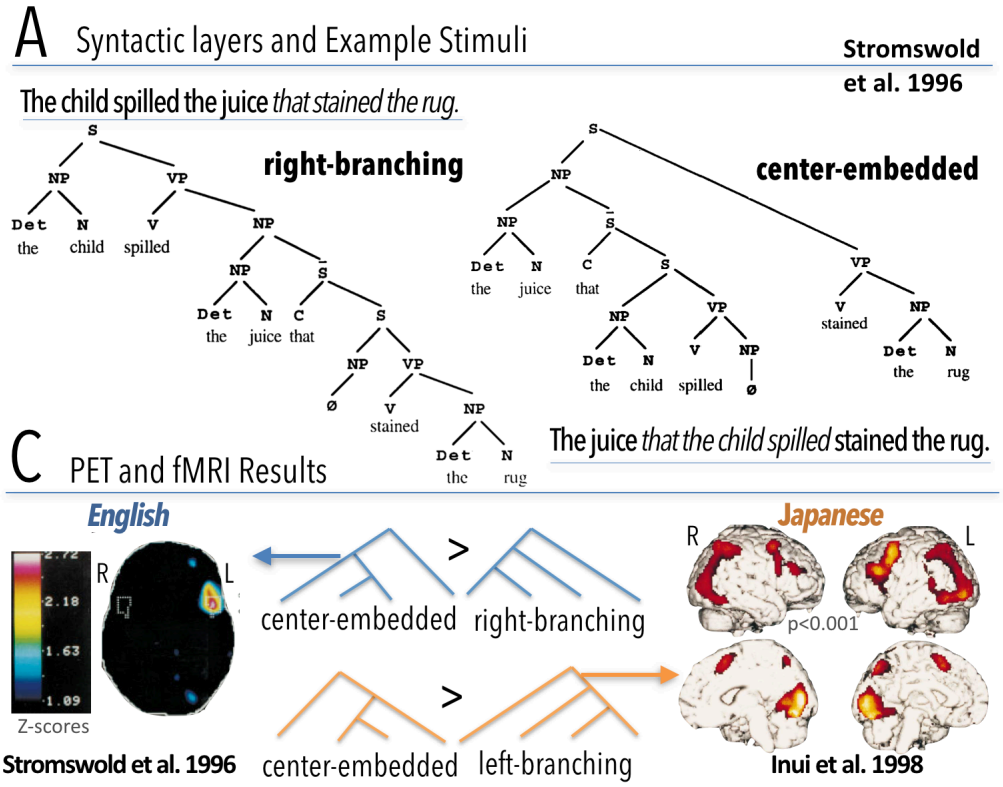


Figure 2.16 – Different Syntactic relations inside the syntactic tree are reflected in more activation the Broca complex irrespective of Head-directionality parameter, as observed in English and Japanese center-embedded sentences compared to right-/left-branching relative syntactic configurations. Adapted from Inui et al. (1998) and Stromswold et al. (1996).

of language, and performance to the mechanisms by which our knowledge of language is put to use[123]. So that linguistic utterances are seen as giving the evidence of a certain set of mental capacities, and it is precisely this shift that can define linguistic as a cognitive science.

However, linguists face a recurrent problem when trying to determine what constitutes the knowledge of a given language: the raw data that linguists start with are informer's intuitions about their own mother-tongue. It is, indeed, very difficult to discriminate or assign informer's intuitions to properties of that language (i.e. Competence) or to other cognitive factors that one could interpret as part of Performance[124]. This is especially true when informers answer to test sentences with the answer: "*sounds funny* or "*sounds awkward*".

This observation leads us to outline a linguistics/psycho-linguistics task-sharing that we already clarified in chapter 1 §1.5.1 (p.68). Grammatical theories have typically aimed

123. Chomsky (1965), Aspects of the Theory of Syntax, MIT Press.
124. See on this topic a classic paper by the psychologist George Miller: "The Magic Number Seven, Plus or Minus Two: Some Limits On Our Capacity for Processing Information" in the Psychological Review (1957), proposing a specific bound on short-term memory across different perceptual domains.

at accounting for patterns of acceptability judgments, methodically founded on contrasts between grammatical and ungrammatical sentences. As a result of this long and collective work, the grammatical constraints of a given language have emerged as a coherent linguistic/grammatical system. However, these theories have remained relatively distant from real-time processes involved in comprehension, speaking and in making acceptability judgments. In parallel, psycho-linguistic theories of language processing, have been, focusing on the role of grammatical constraints in real-time language processes, thinking it was also important to understand whether some constraints derive from limitations on language processing. This focus on the temporal dimension of sentence understanding has been first addressed in language neuro-imaging though EEG technique which constitute a very important tool to investigate to what extent real-time sentence processing mechanisms reflect the details of grammar linguists have been forging.

This being said, as a matter of fact, among all branches of contemporary linguistics, Generative Grammar has to be acknowledged for having addressed the question of the cognitive plausibility of syntactic models, and this for approximately the last half century. Therefore, it is not surprising that studies on neuro-imaging of language have borrowed the formal approach it has developed. Consequently, this is the framework we will adopt to investigate the neural bases of Sentence's syntactic structures in Chinese and French.

One central aspect for the adoption by the Cognitive Science pluri-disciplinary project of the generative linguistic framework, is that the issues of computability and "cognitive plausibility"[125] have been standing at the core of it since the very first moves (Miller and Chomsky, 1962). According to Chomsky (1986a, pp.19-56), the ultimate goal of linguistics in studying *competence* is to characterize the nature of the internalized linguistic system (or I-language) which makes native speakers proficient in their mother-tong. We can read that the a grammar of a language is "a theory of the I-language ... under investigation" (1986a:22), which means that in devising a grammar of a give language, we are attempting to uncover the internalized linguistic system (I-language) possessed by native speakers of that language, that is ultimately an attempt to characterize a mental state -a state of competence, and thus linguistic knowledge. We see here that Chomsky takes a cognitive approach to the study of grammar, investigating (1) what the native speakers *know* about their native language that makes them capable them to understand sentences and speak, and (2) how that linguistic knowledge might be represented in the mind/brain. Since that epoch, Cognitive Science has largely benefited from discussions on the syntactic models offered by "Principles and Parameters" and Minimalist programs. And, syntactic rules have been gradually considered as being at the interface between linguistic competence and performance[126]. Yet, it should be clarified that more strictly speaking from the point of view of computational procedures, the processing system of syntax - the parser[127] - is generally considered as a set of algorithms applying

125. cf. learnability. Generative linguistics has aimed since the beginning to precisely define the representation of linguistic knowledge, trying to take into account as much as possible the kind of properties that make the linguistic representation cognitively plausible, like children language acquisition criteria, psycho-linguistic and cross-linguistic validity, etc.

126. To clarify this assertion we forward the reader to to Embick and David Poeppel (2015) and to Marantz (2005), where we can read: "generative linguistic theory serves as a theory of language within Cognitive Neuroscience. The categories and operations of generative grammar are hypotheses about the representations and computations in the minds and brains of speakers".

127. Definition of *Parser*: a natural language parser is a program for analyzing a string of words (sentence) and assigning it syntactic structure in accordance with the rules of grammar. From J. Hale, "Automaton Theories of Human Sentence Comprehension" (2014). Ideally, the relation between basic

the syntactic rules that are defined independently. And, generally speaking this module (i.e. computational system) can build bottom-up syntactic trees, while a parser reads incrementally from left to right the sentence.

Such computational distinctions led researchers to consider Performance as a set of behaviors, while Competence as a mental system that organizes them. The schema presented in Figure 2.17 gives a more concrete idea of a proposal embodying the competence/performance dualism by Collin Phillips (1996)[128].

Yet, having to look at brain activity elicited by the comprehension of sentence stimuli, however, is compelling us to adopt a unitary approach where both sentence structure's representation and its processing can be the determinants of the observed cerebral activity. Thus, wanting to focus on the representation of the syntactic structures in the brain will not only imply to concentrate on the elements listed in the upper orange box in Figure 2.17, but also to strenuously try to isolate through experimental designs[129] metrics that can correlate with brain activity in areas that should be linked to the representation of the sentence structure. All this, of course, assuming that the language system instantiated in the brain needs to build up and therefore represent sentence structures during sentence comprehension.

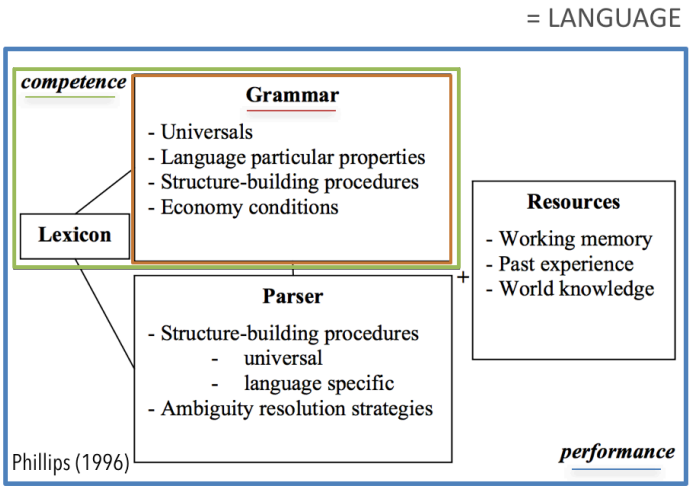


Figure 2.17 – Standard model, representative of the competence/performance dualism, adapted from Phillips (1996:16).

Going into further details about the characteristics these computational models should have to be neuro-physiological grounded is beyond the scope of this Section, however, we could just say, with Poeppel and Embick (2005), that, by hypothesis, computational models should feature *primitives and operations* that (i) should be of the type to be plausibly executed by assemblies of neurons, and than could be (ii) reasonably constitutive of recurrent subroutines of linguistic computation, providing in this way a theoretical foundation.

Hence, one of the fundamental reasons for considering syntactic-trees in a Neurocognitve investigation of the representations of syntactic structures, lies essentially in

parsing operations and basic operations of grammar approximates the identity function. Probabilistic parsers use statistical information to provide the most likely grammatical analyses of new sentences (Everaert et al., 2015).

128. Note that the very basic model illustrated by Colin Phillips' early work is here just to illustrate an approach where Competence and Performance are accounted dualistically. The relation between parser's structure building and Gammar Competence are discussed in depth by more recent work by C. Phillips (e.g. Phillips, 2013 "Parse-grammar relations: we don't understand everything twice"). More recent approach to how language performance is shaped by cognitive constraints, and specifically memory constraints is developed by this author. The relation between linguistic representation and memory architecture is addressed by several article (see Chacon et al., 2015), but a very interesting article 'Encoding and navigating linguistic representation in memory" is found in a special edition of Frontiers in Psychology (2017).

129. See 'The work of experimentation' in chapter 1, §III and the discussion on the decomposition of cognitive processes in neuro-imaging in section §1.5.3.

the the fact that they offer quantifiable measures of syntactic complexity that can be correlate with brain activation data. to this it should be added that although this wasn't the initial idea in generative grammar, under a computational account they could provide a model of the sentence as a cognitive performance or a computation, which in turn would imply understanding sentences as having a history of computational derivations underlying them.

It should be remembered that the very first formulation of the idea that counting nodes in a syntactic-tree could be an index complexity -be it processing or representational complexity- dates back at least to Chomsky and Miller (1963)[130], a measure that is close to what is called in computational linguistics Yngve-dept (Yngve, 1960). The structural complexity of a sentence as expressed by Yngve (1960)[131], and in particular the syntactic tree-depth (Brennan et al., 2012), has traditionally been taken to be one of the central determiners of the parsing difficulty of a sentence. Interestingly, these measures can represent an insight on the incremental and dynamic operations that happen during sentence structure comprehension. Namely, we can say that these parsing-oriented neuro-imaging studies treat the evolving (word-by-word) internal state of a given parser as a possible proxy of the human sentence parser.

At this point, an important distinction is worth noting before continuing. Two main approaches to sentence structural complexity have been correlated in the literature with brain activity: measures of complexity linked to the structural representation of the sentence and measures of complexity linked to the calculation of the structure of the sentence. Namely, on one side we find construction-based manipulations of sentence structure, treating complexity as a static property of the whole sentence without any temporal or incremental dimension[132]. It has to be noted that this approach has a reasonable methodological justification. It was taken in fMRI studies also because of a fundamental discrepancy between the times scales at which sentence comprehension happens – an order of magnitude of milliseconds – compared to the temporal resolution of the BOLD-signal recorded by this brain imaging technique that is on the order of several seconds. Hence, we can synthetically say that these studies quantified the syntactic complexity effect at the temporal resolution of the entire sentence, a methodological choice that has the consequence of observing cerebral correlates of sentence structure's representation instead of those linked to its calculation with on-line complexity measures.

Yet, on the other hand, an increasing number of fMRI studies have been taking into account the incremental and dynamic nature of human sentence comprehension. And, building on Yngve's proposal, they have been modeling the incremental sentence-structure formation through a complexity measure representing the number of syntactic steps -the nodes added to the syntactic-tree of a sentence- that are required for the integration of the incoming word into the already existing tree structure created by the parser based on the previously encountered words.

We will conclude on the topic of the computational interpretation of syntactic-trees by presenting two neuro-imaging results where we can clearly see the enactment of a shift from more concrete and construction-oriented syntactic tree-complexity testing -

130. Some argue that it is also the essential intuition lying behind Minimal Attachment (Frazier, 1985).
131. Note that the author proposed an account of complexity based on the depth of the computational pile.
132. These measures can be therefore associated to the kind of off-line complexity measures offered by early psycho-linguistic attempts to investigate the psychological dimension of sentence structure (e.g. reaction times)

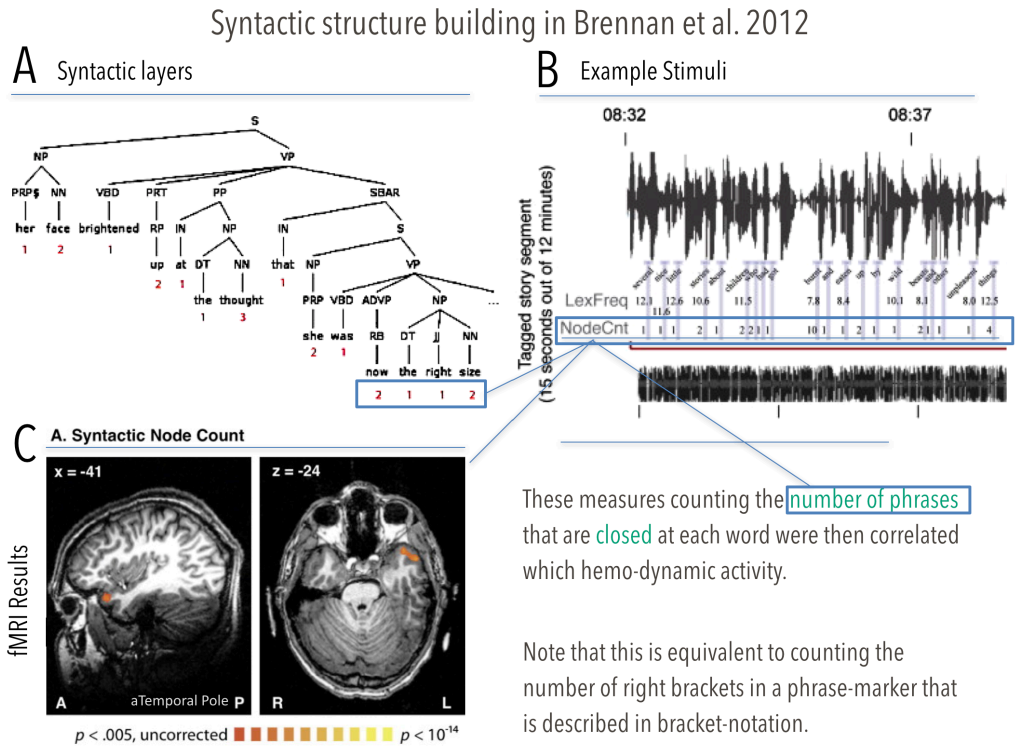


Figure 2.18 – (A) A tree-like representation of a sentence from Alice in the Wonderland. Orange numbers represent incremental syntactic structure building measure counting the number of phrases that are closed at each word. (B) Word boundaries tagging of the story sound signal with Lexical frequency measures and Node count measure. (C) Group fMRI results for Syntactic Node Count, indicating a single cluster region where Bold signal correlates with the syntactic structure building measure selected (i.e. Syntactic node count), adapted from Brennan et al. (2012).

focusing on relatives, question formation, passives, etc.- to a more abstract computational approach. Brennan and Colleagues (2012) quantified the rules and relationships that determine basic sentence structure building using a word-by-word measure of the amount of syntactic structure analyzed at a given time point. This particular type of node-count measure is represented in Figure 2.18 (A) by the red numbers at the bottom of each leave of the syntactic-tree. As we can see what was precisely counted is the number of phrases that became closed at each word[133].

Once calculated for an approximately 12 minutes excerpt of the book "Alice in the Wonderland", these node-count measures were then correlated with the hemo-dynamic activity of participants listening to the recording. Results indicate the anterior Temporal Lobe as the locus of this type of syntactic structure building in a highly ecological story-listening situation, where natural language comprehension is indeed free from artificial task demands. Not only the region involved in what the authors hypothesize to be a process of syntactic structure building is more and more prominently linked to syntactic

133. Note that, as the authors remark, this is equivalent to counting the number of right brackets in a phrase-marker that is described in bracketing notation.

processing in the literature (as we saw in in chapter 1), but the kind of measures of syntactic-node counting that were used represent a good example of one of the possible approximations of the structural complexity that a tree-like representation can offer to investigate sentence parsing.

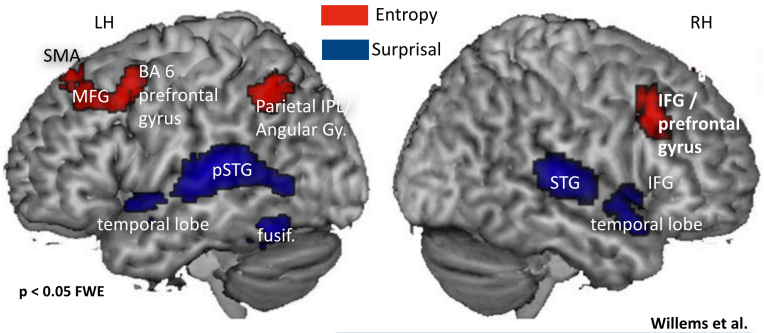
Another study by Willems and colleagues (2016) embodies a more recent attempt to leverage the computational measures linked to sentence-internal structure in order to characterize syntactic predictive dimensions of sentence understanding (see Figure 2.19). Two more abstract computational measures linked to sentence structure predictive processed were correlated with brain activity during story-listening. In this fMRI study, the authors estimated in a word-by-word fashion the entropy of next-word probability as well as '*surprisal*', a measure expressing how unexpected a current given word is given the previously encountered words. This syntactic measure of surprise is meant to represent the difficulty of integrating the current word compared to previously built expectations. It has been in fact hypothesized to be proportional to the cognitive processing effort of integrating a word into the current context by Hale (2001), and it has already proven to be cognitively pertinent in that experimental measures of comprehension difficulty like sentence Reading Times (Smith and Levy, 2013) and with the amplitude of the N400 ERP-component (Frank et al., 2015) have been reported to correlate with it.

Simply put, Entropy measure (in Figure 2.19 in red) is a forward looking measure expressing the strength of expectations about what the next coming word will be: thus, giving a high score (in red) to highly predicted words and low score to indicate how uncertain the language system is uncertain about what is coming next, accordingly red areas should house predictive processes linked sub-serving strong expectation. Conversely, Surprisal is a backward looking measure giving the extent of effort it caused to the system to integrate the present word because unexpected given the previously encountered words: it has thus high scores (in bleu) when the word is unexpected, hence blue areas should host integrative mechanisms that are necessary when exception are not meat or when prediction is low.

The approach and methodology of this study participate to an actual trend of combining computational linguistics with cognitive neuro-science keeping at the core of research approach to language (1) the worry for the naturalness of linguistic stimuli, on one side, and on the other (2) the interest for predictive mechanisms and complexity measures in sentence structure processing.

It is interesting to note at this point. We can remark that this perspective concentrating on the effect of surprise (in different surprisal measures) to look at the cognitive effects of predictions and expectations in language, is currently being revived in a number of studies (see Kuperberg and Jaeger, 2015) for an interesting discussion on predictive mechanisms in language comprehension). Hypothesizing a correlation between the extent of compatibility with prior expectations of a linguistic input and the extent of cognitive effort to integrate a stimulus, fundamentally holds on an apriori, that is: human incremental parsing of sentences relies on the formation of predictions about down-stream input of the sentence and about its overall structure.

Pushing these arguments a bit further, we can bring back on the stage our research *leitmotiv* about the cerebral representation of sentence structure, by saying that the question of whether syntactic prediction has a neuro-cognitive impact during sentence comprehension is also directly connected to the fact that sentence structure should be represented in the brain. In other words, a sentence processing architecture that actively engages in prediction needs to have a sentence representation to do predictions on, which



fMRI results of the whole-brain analysis. Brain areas that were significantly activated in response to the regressors modelling entropy (red) and surprisal .

Entropy expresses the strength of expectations about what the next coming word will be.

Surprisal expresses how unexpected was a word

Entropy results for whole-brain analysis

Region	MNI coordinates (X Y Z)	Cluster extent (voxels)	Maximum t-value
Right inferior frontal gyrus	54 18 40 48 18 48	180	−3.99
Left middle frontal gyrus/ventral precentral sulcus	−46 12 50 −42 8 56 −36 30 46	353	−3.53
Left SMA	−4 42 44		−3.14
Left inferior parietal lobule	−44 −56 50 −52 −56 48 −46 −48 48	235	−3.50

Figure 2.19 – Top: Group Brain maps with significant activations in response to the regressors modeling Entropy measures (in red) and Surprisal measures (in blue). Bottom left: a detained definition of the computational measures correlated with brain activity. Bottom right: Table reporting the coordinates and voxels' cluster extent for the whole brain analysis results for Entropy measure. Adapted from Willems et al. (2016).

indeed turns out to be relevant to our recurrent leitmotif about distinguishing the cerebral representation of the sentence (i.e. its encoding format) from its processing. Put differently, having predictions about the upcoming sentence input, or about the overall structure could be interpreted as ultimately needing an internal representation of the sentence to build predictions for the incremental comprehension of the sentence. Thus, these surprise measures and prediction measures could actually point to brain areas that actually represent sentence structure, to subsequently be able to predictions on (in red), or to locally modify it when a word is unexpectedly met (blue). This is mainly the reason why we consider this findings as being informative for our research direction, which can be defined as ultimately trying to capture sentence structures representations in the brain.

2.3.4 Evidence for sentence structures in psycho-linguistics and neuro-imaging

In this section, we have reviewed a certain amount of phenomena and experimental evidence for what is usually called the *Psychological Reality* of sentence linguistic structures: we went through (1) punctuation and prosody cues to sentence structure, (2) experiments like click-detection and their shifts indicating the processing load of sentence constituent-structure, and (3) we also saw how reading times and Eye-movement paradigms have confirmed that the processing time is influence by factors like syntactic structural boundaries.

Now, before turning to the analysis of the complexity of sentence-internal structure, we should add that support for the psychological reality of syntactic structures comes also from psycho-linguistic studies showing that speakers tend to reuse the syntactic structure of recently heard sentences, a phenomenon known as syntactic priming (Branigan et al., 2000 and Bock et al., 2007)[134]. This brings us to another important type of syntactic priming effects, those observed in neuro-imaging studies.

In fact, at the neuronal level, repetition suppression paradigms[135] has the advantage of being able to detect neuronal populations that are sensitive to properties that are shared by consecutive sentences. A certain number of fMRI studies manipulated the syntactic structure between prime and target sentences looking for brain regions sensitive to the syntactic information sentence-internal constructions are.

Evidence from neuro-imaging: sentence structure can be primed

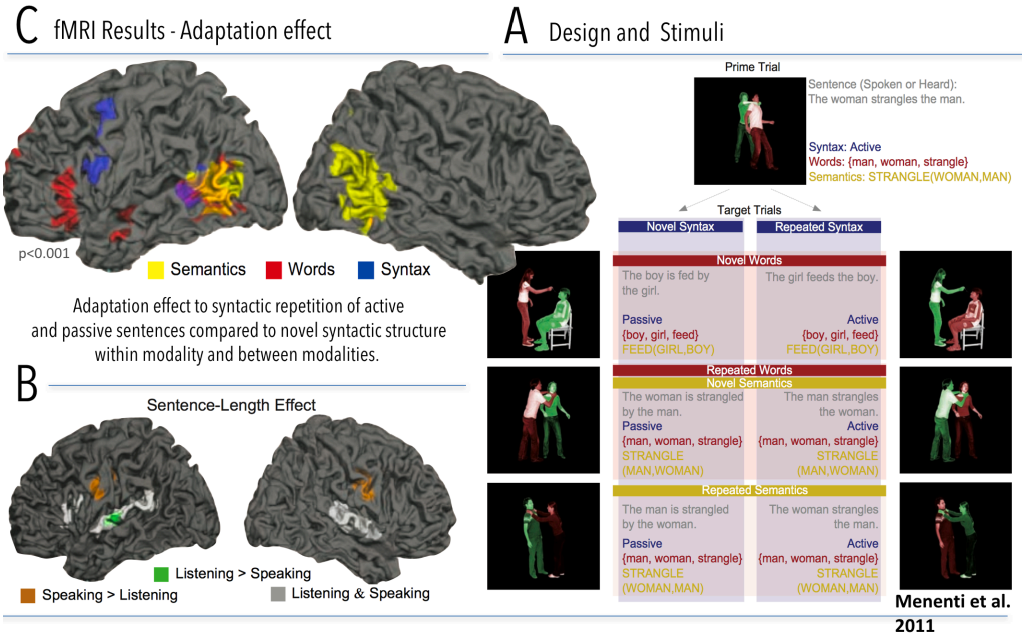


Figure 2.20 – (A) Experimental stimuli examples and design. (C) Group-average brain map for adaptation effects to syntactic (blue), semantic (yellow) and word (red) repetition. Syntactic adaptation of active and passive sentences compared to novel syntactic structure within modality and between modalities. (B) Group-average brain maps for sentence-length effect, contrasting listening and speaking. Adapted from Menenti et al. (2011).

One of the first syntactic priming studies reported repetition suppression effects in the anterior left Temporal Lobe (aTL, Noppeney and Price, 2004), making this region

134. For two reviews see Ferreira and Bock (2006) and Pickering and Ferreira (2008).
135. While priming paradigm in behavioral studies recording accuracy or reaction time defines the observed improvement in behavioral response when stimuli are repeatedly presented, in fMRI studies the phenomenon called of repetition suppression is an observed reduction in neural activity when stimuli are repeated, mainly depending on the functional processing overlaps between the repeated items (see Grill-Spector et al., 2006).

a potential candidate region for the encoding syntactic structure in the brain. Later, another syntactic priming study (Devauchelle PhD, 2008) found in a neighboring region (Temporal Pole, TP) only a priming effect of the whole sentence verbatim (i.e. lexicon and syntax) and the the hypothesis was emitted that left aTL complex (broadly speaking) could be considered as encoding sentence-level global propositional interpretation. While another fMRI syntactic priming study by Devauchelle and colleagues (2009)[136] did not find any syntactic priming effect but only a lexico-semantic priming effect in another close neighboring region (i.e. aSTS)[137]. To this already heterogeneous panorama of findings an interesting study testing for the cerebral correlates of sentence prosodic and syntactic dimensions (Humphries et al., 2005) found in aTL complex a priming effect of the prosodic structure of the sentence, which is indeed tightly linked to sentence structure.

Some more recent syntactic priming studies were able through sophisticated priming designs (see Figure 2.20 A) to reveal distinct neural networks adapting to the three main linguistic processes involving semantic, lexical, and syntactic information in both production and comprehension modality. Menenti et al. (2011) run a speech comprehension and production design where subject had either to produce a sentence according to a transitive verb and a subsequent picture presented on a screen (see Figure 2.20 A for an example). Pictures presented a color coding for the participants of the action (green, for grammatical subject, and red for grammatical object) that was cuing for the production of passive or active syntactic structure. Alternatively, during comprehension trials, a sentence-picture matching paradigm was used and participants were presented with a photograph in gray color scale and an auditory sentence describing the picture.

From the results in Figure 2.20 (C), we can see that only three areas in the left hemisphere adapted to syntactic construction across modalities, namely Pre-central cortex (BA6), Inferior frontal cortex (Broca/IFG) and Middle Temporal Gyrus (MTG). Importantly, this study was also able to bring to light a network adapting to peripheral low-level processes (i.e. number of syllables) involved in sentence comprehension contrasting them between modalities, as illustrated in 2.20 (B).

Thanks to these results we start to see an important division of labor in the cerebral organization of linguistic information for semantics-related (in yellow) and syntax-related processes (in blue), and words activation (in red) lying in between[138].

136. Note that in this study, syntactic construction varied.

137. The authors retrospectively speculate that the absence of syntactic priming could be due to the absence of an active task. A speculation to put in parallel with the fact that syntactic complexity effects in Pattamadilok et al. (2016) were present only when a task was performed on sentences, inorder to explain the contradiction with more recent results like Menenti et al. (2011). Pallier (p.c.).

138. However, given that our research is about syntax, we do not address this issue further, since this would take the discussion too far afield, the reader will find in the annexes a dedicated entry on this important issue D.2 (page 911).

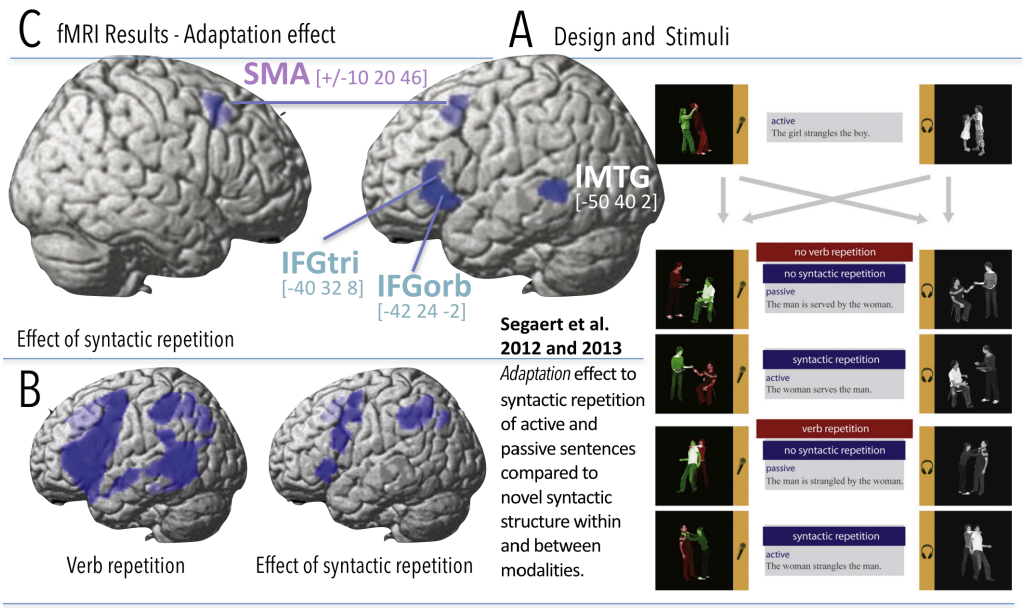


Figure 2.21 – (A) Experimental stimuli examples and design. (C) Group-average brain map of the adaptation effect to syntactic repetition of active and passive sentences compared to novel syntactic structure within and between modalities. (B) Separate Group-average brain maps for Verb repetition on the left and for syntactic repetition on the right. Adapted from Segaert et al. (2013).

These results on syntactic priming where later replicated by Segaert et al. (2013) as shown in Figure 2.21. Interestingly, this study was able to separate effects of verb priming from those of syntactic priming, but found syntactic priming effect only for passive sentences across modalities. Some areas adapting for the syntactic structure of sentences showed in 2.21 (B) and (C) belong to the set of areas we presented in chapter 1, that we will repeatedly 'break through' the sentence comprehension network.

One last example from the fMRI literature on priming of syntactic constructions needs to be reported to add a final confirmation of the existence in the brain of dedicated sub-components in the sentence network for the representation and processing of sentence structures. In fact, syntactic priming studies, compared to conventional subtraction analyses, allow some times complex configurations to be tested. A particularly interesting study for our experimental question on French wh-question, is a perfect example of the complex configurations that can be addressed by syntactic priming experiments. Testing German–English bilinguals Weber and colleagues (2009) tried to answer to a long-standing question in bilingualism: Is syntactic information shared between the two languages processing systems?

The authors investigated syntactic priming in reading comprehension in German–English late-acquisition bilinguals, by priming either with a passive or by an active sentence in English or German, a sentence with a passive sentence structure in English. Repetition suppression effects linked to syntactic structure were found across languages in left inferior frontal, left Precentral and left middle temporal regions of interest as shown in Figure 2.22 (C).

In conclusion, this additional evidence in support of the claim that sentence linguistic

structures are really represented and processed by the brain leads us to now go a step further and ask more cardinal question for our research project that of sentence internal structural complexity.

At this point it will be useful to summarize the discussion on tree representation and sentence structure building so far, by drawing an intermediate summary of what can be considered the complexity metrics that we already can obtain in syntactic trees despite the minimality of the compositional mechanisms that merge is.

Despite its minimality, from what we already saw in this section we a can indeed draw an intermediate summary of the metrics we already presented to evaluate sentence complexity, and list down the advantages for representing sentences structures in tree-like format, by briefly and simply enumerating the different complexity metrics they offer to further build-up experimental hypothesis in neuro-imaging, that we have identified to this point:

1. Syntactic-tree total node counts (total depth);
2. Syntactic-tree depth at a given word (local depth);
3. Structural configuration of the syntactic-tree (shape);
4. Syntactic construction reflected in syntactic-tree embedding (cf. Left-branching versus center-embedded sentences)
5. Relations between syntactic positions in the syntactic-tree

These aspects linked to the sentence representation through tree-like representation will be greatly enriched by enriched by introducing what we are going to call syntactic transformations.

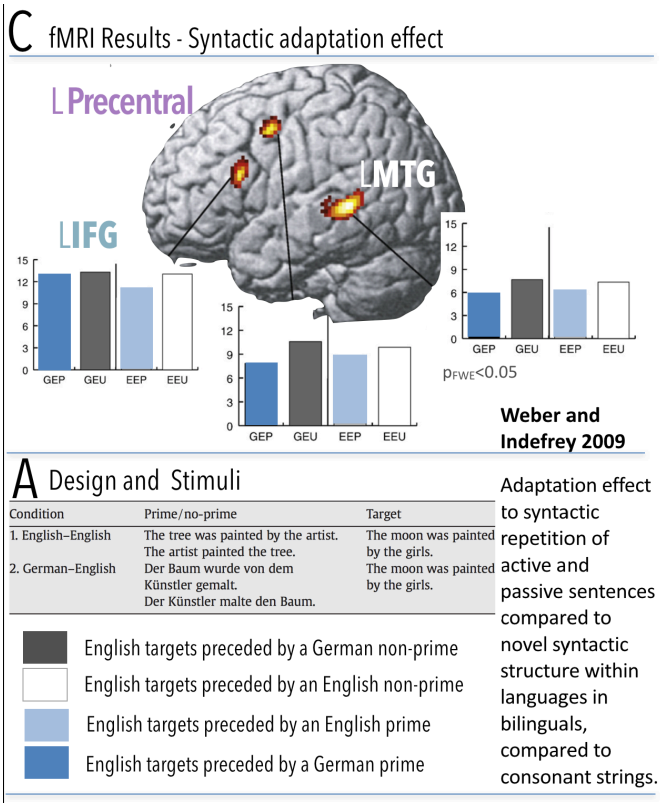


Figure 2.22 – (A) Experimental stimuli examples and design. (C) Group-average brain map of the syntactic adaptation effect to passive and active sentences in French-German bilinguals compared to consonant strings. Barplots of the contrast estimates for the three ROIs for the four different conditions (GEP: English targets preceded by a German prime; GEU: English targets preceded by a German non-prime; EEP: English targets preceded by an English prime; EEU: English targets preceded by an English non-prime). Adapted from Weber and Indefrey (2009).

2.4 Syntactic complexity and transformations

In our step-by-step definition of the sentence unit, we arrive now to the most central aspect for our research -the sentence's syntactic complexity. This section will be dedicated to understanding and analyzing the sentence-unit as a complex representation.

In our review of the possible syntactic processes to represent and process the sentence-unit that could be instantiated in the human brain, we already came across (a) *Merge*, a hierarchical structure building operation, and (b) some ways to manage reference assignment inside a sentence (i.e. *Bind*).

Given the extreme simplicity of the syntactic generative devices introduced until now, one might wonder how the very simple combinatorial operations of Merge, even when complemented by a Search operation identifying the candidate for Merge, can account for all the aspects of sentence structural complexity in one language or more generally in all languages.

The structural minimality of tree-like syntactic representation will be here enriched by an addition syntactic operation that will allow to account for the attested sentence complexity and the wide range of sentence constructions that languages present and will be the focus of a great part of our experimental approach.

We will now turn to describe the sentence as the result of syntactic transformations, considering sentences as having a 'history' of structural syntactic derivations underlying them. Therefore, we will focus on a new structure building process operating on syntactic trees, that derives new structures from basic ones – *Movement*, while next section will be dedicated to introducing the representation of syntactic elements that are structurally needed and semantically interpreted, but not phonetically realized – the so-called *Empty Categories* or Gaps[139].

Wundt's Umwaldungen

The idea that surface word order is the results of a transformations dates back to Wundt's formal analyses[140]. After his initial characterization of the sentence as the natural unit of linguistic knowledge, he assignment of purely abstract syntactic structures to sentences, independent of their meaning, observing that the surface grammatical relations could not capture what were called at that time the propositional relations between phrases of a sentence. Taking sentences like (a) '*Caesar was crossing the Rubicon.*' and (b) '*Cross the Rubicon was what Caesar did.*', he noted that Caesar is the actor in both of sentences, despite its different surface position.

At the level of propositional relations (a) and (b) share the same relations between [agent] 'Caesar', [action] 'cross' and [patient] 'the Rubicon'. This level of representation would be called by Wundt the inner form of the sentence, and the differences at the grammatical would be theorized as being the result of mapping processes that he called *Umwandlungen*, literally meaning "transformations". Importantly, such mapping between inner form (propositional level) and actual grammatical realization of surface sentence were taking into account the actual surface patterns allowed by each particular language.

Starting from this initial intuition dating back to the beginning of last century, next section will deepen our understanding of syntactic transformations and we will dive into what has been called "derivational theory of complexity" (DTC) in psycho-linguistics, an hypothesis consisting in correlating the number of operations that the grammar uses to generate a sentence with behavioral measures recorder by psycho-linguistic tasks, like the time for speakers to process the sentence.

139. The notion of trace was used in Chomsky's Principles and Parameters (i.e., Government Binding theory; Chomsky, 1986), Within Minimalist Program a different approach was developed. While the details of the two theories differ, the basic function of *trace* and *copy* are the same, and we will generically call this empty phonological syntactic elements *gaps*.

140. Wundt (1874) is considered as one of the forefathers of Experimental cognitive psychology, see Wundt (1911) for a vintage look on one century of research on the natural units and structure of language.

2.4.1 Derivational theory of complexity

In order to account for the wide range of possible and complex sentences in human language, seminal works of Chomsky (1955 and 1957) proposed two classes of rules, namely, phrase-structure rules that combine two constituents to create phrases, and transformation rules that map one phrase to another. The transformations type of syntactic rule, namely allowed to add, move, or delete sentence elements from a syntactic position.

Importantly, syntactic transformations are to be understood as a mean to relate superficial properties of the sentence – like the ordering of the elements as they are realized in the surface sentential sequence – and a Deep structure (D-structure)[141] representing the basic argument relations in the sentence, a level of the sentence that encodes the lexical properties of the constituents of the sentence[142]. Movement allows to establish at once several local relations by forming a dependency chain. For instance, a wh-object is inserted in complement position of the verb within the VP area to satisfy the thematic role in its original position, it is then re-merged in the CP so that it can establish a local configuration with the wh-interrogative Head that license clausal-type.

Some of these transformational operations, converting a sentence into other sentence, imply what we will define in the next section a Movement operation[143]. Transformational operations were fist linked to different construction like passivization – "*The Dissertation was read ___(by many people).*"– or subject-auxiliary inversion – "*Will someone ___ read the Dissertation?*" – and will be later subsumed under a single general movement rule (Chomsky, 1981, and subsequent works 1995, 2001).

This understanding of sentence syntactic complexity as the result of a series of transformations of a basic (base-generated) sentence pattern immediately attracted the attention of psychologist, who tried to experimentally ground this view on sentence complexity.

2.4.1.1 Psychological reality of syntactic transformation

In the early 1963, Jacques Mehler investigated how a manipulation of syntactic complexity could impact peoples capacity to remember sentences. Basic sentences, their correspondent negative, passive and interrogative versions, and the corresponding sentences modified by all these transformations at once, were presented to participants who performed a memory task. The results of his study showed that the probability to remember a sentence is inversely proportional to the number of syntactic transformations it carries. This study had a broad impact, it inspired and stimulated, with other studies, the first hypotheses leading towards the research direction that is nowadays looking for the neural code of sentence's deep structure and its syntactic transformations, a direction we are in a way pursuing and expanding today, testing a similar hypothesis by looking at neural correlates of syntactic movement transformations.

At this point a short clarification on the cognitive processes linked to sentence's complexity is necessary to better understand the position we will be taking on syntactic com-

141. Note that the notion of Deep/Superficial Structure were current in the Government and Binding framework, and are realized derivationally in the Minimalist approach so that these two levels are unified. This notion is now replaced by the one of External Merge (see Figure 2.24, p. 164)

142. At this Deep structure level, external arguments are base-generated, while internal ones are governed by the predicate in their base position.

143. See Types of Movement Operations, later in section §2.4.2.1

plexity by investigating syntactic movement-related complexity effect in neuro-imaging.

As we already saw in previous sections, before the advent of neuro-imaging the processing complexity was initially measured in psycho-linguistics thorough a quite simplistic behavioral measure like Reaction-Times (see for example psycho-linguistic experiments reported earlier in section §2.1.2.7), which had the intrinsic limitation of begin broadly capturing some complexity effects happening off-line (i.e. after listening to the sentence, and not during sentence's comprehension) that were difficulty identifiable with a very precise cognitive linguistic process. Alternatively, in another paradigm, as we saw in effects reported in clicks-detection tasks, reaction-times could offer a on-line measure of complexity during sentence listening that can be interpreted as a parsing/processing complexity Through the first neuro-imaging attempts of the 90's the understanding of sentence's complexity was still associated to a processing and parsing complexity measure implying time, and cerebral activations were understood as a proxy of the Reaction-Times measured by psycho-linguists.

However, the introduction of a computational view of sentence comprehension operated a shift towards another understanding sentence's complexity as a syntactic computation, essentially implying two types of "complexities": (1) *temporal* complexity linked to sentence *parsing*, and (2) a complexity dimension linked to the amount of *memory* resources recruited to *represent* the sentence's structure. In neuro-imaging a turn over can be seen in the early 2000 when the view was advanced that cerebral response patterns reflect the representational complexity dimension of visual objects[144]. Following Haxby et al. (2001), this understanding of cerebral activations as reflecting representational dimension of a given stimulus was later adopted also in the neuro-imaging of sentence structural complexity. Namely, Pallier and Colleagues (2011) (see Figure 1.15, p. 66), who interpreted their fMRI results, understanding sentence structural complexity à la Smolensky: the more the sentence is structurally complex the more it recruits memory to be represented. While it is no secret for the reader that we will adopt the second position (cf. chapter 1, §1.3.5), it should be noted that investigating the neural correlate of syntactic-movement we will by no means consider this syntactic operation as having a temporal dimension during sentence comprehension. We will in fact adopt a fully representational understanding of syntactic-movement, for example, by investigating in chapter 6 and 7 the cerebral representation of the empty syntactic positions left by the dislocation of sentence elements.These two fMRI study are namely uniquely conscerned with coprehension ofthes sentences, on the contrary considering production would imply a totally different point of view.

Given this background, if we consider the study by Stromworth et al. (1996) presented in previous section (§2.3.3.2, Figure 2.16, p. 147), it is difficult to attribute to these results a clear-cut interpretation in terms of representational complexity vs. parsing

144. This view was advanced by fMRI studies on the functional architecture of the objects vision in human visual cortex (i.e. ventral temporal cortex) like Haxby (2001) (see also Hanson et al. (2004) for a presentation of the concept of "combinatorial codes" in visual cortex.), who put forward that the visual perception of faces, for example, is mediated by a distributed processing, thus promoting a perspective on the neural systems that underlies object vision, that understood cerebral responses as a distributed system "both in terms of the involvement of multiple brain areas, and in terms of locally distributed population codes within these areas" (Haxby and Gobbini, 2010). This perspective introduced a data analysis method called *decoding*, where fMRI responses are considered as *cerebral maps* (i.e. pattern of responses in the brain) that are not only corresponding to different stimulus categories (e.g. faces, cats, five categories of man-made objects, and nonsense pictures), and that discriminated among all categories.

complexity to these fMRI results[145]. Namely, the fact the authors selected to test for syntactic complexity choosing a sentence structure with an embedding make it even more difficult to disentangle (1) the complexity effects linked to representational complexity of the syntactic-tree, (2) the complexity of the computation of such structure, from an additional complexity parameter that is linked the on-line memory dimension of keeping in working memory the matrix clause during the parsing of the sentential embedding[146].

In sum, we can say that these two positions fundamentally differ in the interpretation they give to cerebral activations in response to complexity. While the first favors on-line complexity measures linked to incremental parsing of the sentence (likewise the computational time needed to parse the sentence), the second favors offline processes that can be view as more tightly linked to representational complexity of the sentence.

A recent regain in interest for the idea that sentences can be represented by the history of their transformations or successive derivational-steps was inaugurated by a pioneering attempt to correlate brain activity with measures linked to *derivational-steps* computed directly from a Parser[147] (Bachrach PhD, 2008). In this approach the priority is given to on-line measures, like Syntactic surprisal and the number of derivational steps, which are incremental measures reflecting a complexity in parsing and computing sentence structure.

Figure 2.23 shows fMRI activations to the two central complexity measures linked with sentence structure incremental understanding: Syntactic surprisal and number of derivational steps needed to achieve a give sentence structure in short narratives created for the purposes of this study. As for these two syntactically-oriented computational measures, on one side the number of derivational steps, is a measure reflecting the local structural complexity of the sentence, in that it represents the number of steps (i.e. Rorack parser's steps = rules) used during integration of a given word into incrementally constructed syntactic structure. On the other side Syntactic surprisal measures the conditional probability of the grammatical part of speech given the preceding context. fMRI activation patterns show a prominent and bilateral involvement of temporal regions for these two syntactically-oriented computational measures.

To conclude, about the derivational theory of complexity, the general underlying assumption of this research direction investigating syntactic complexity, could be expressed in the following terms: the more complex a representation is -the longer and more complex are the linguistic computations necessary to generate a representation- the longer a linguistic task should take, or the more people's brain should activate resources, or put to work brain areas to build or access to this complex representation, while performing sentence understanding (for further details see Phillips et al., 2005; Phillips and Lewis, 2013).

In this approach it is possible to test linguistic theory predictions about complexity (and about the nature of linguistic representation and computations) to correlate them with experimental neuro-imaging techniques, to approach the ultimate goal of a cognitive representation of the sentence unit. Moreover, linguistics also offers a rich set of hypothesis about the similarity of or identity between linguistic representations and elements that are generally investigated by priming experiment as we just saw in the

145. An important aspect in this decision is also linked to the task that different experimental paradigms select.

146. Note that this complexity (working memory) is commonly acknowledged to be present in all kinds of long distance dependency.

147. Roark's Parser, a context-free Top-Dow Parser.

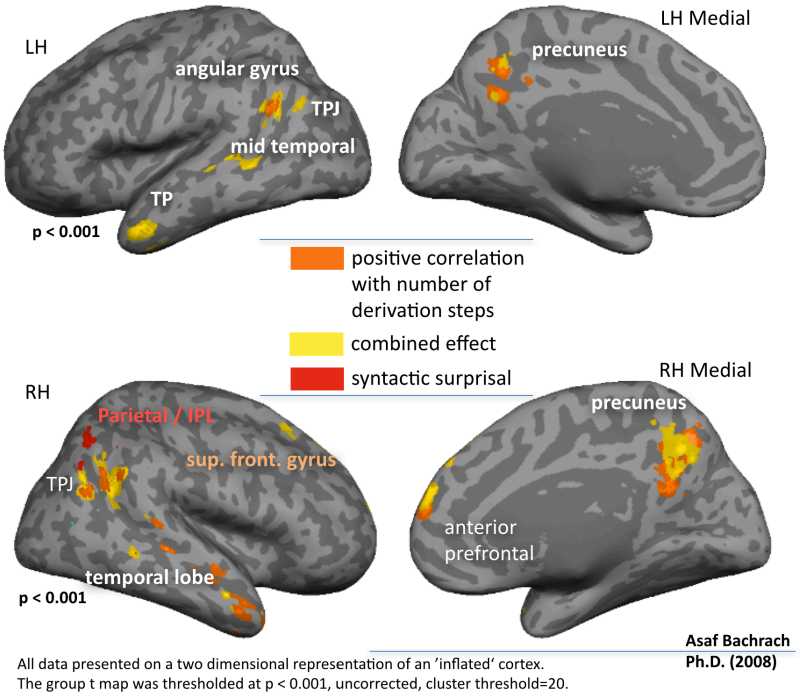


Figure 2.23 – Group-level brain maps showing significant BOLD activity (p<.001, uncorr, cluster threshold =20) for the increasing number of derivational steps (in orange), for syntactic surprisal measure (in red), and for their combined effect (in yellow). Adapted from A. Bachrach's PhD (2008).

above section (§2.3.4, p. 153).

The present work can be seen as inscribed in this research direction, to investigate the neural underpinning of sentence complexity in its transformational aspects thanks to syntactic movement.

2.4.2 Movement, syntactic transformation

Form the very basic observation of complex sentences, we can say that sentence elements are often pronounced in a position different from the position in which they are interpreted. This is typically the case in wh-questions, like "*Which Dissertation should I read?*" which can also be represented "Which Dissertation should I read ___?" But, where in the syntactic-tree does the object of 'read' moves?

As we saw in previous section presenting the different functional projections, the type of the sentence-unit is determined elements found in the Complementizer Phrase, therefore to get an interrogative mode the object of read needs to move high to the Specifier position of the Complementizer Phrase.

These constructions share an interesting property: the constituents at the beginning of these sentences are "displaced" from their original position and this position is indicated by "___" a so-called *gap*[148]. Even if the object appears in a displaced position,

148. The notion of trace was used in Chomsky's Principles and Parameters (i.e., Government Binding

before the verb and the subject, the thematic relations (e.g., who did what to whom) remain unmodified, the theme role, for example, is always assigned post-verbally at the place of the gap.

Movement operation was introduced by transformational theories of generative grammar to characterize these displacement phenomena, as seen in the above wh-questions.

Constraint on movement operations were early identified, and since Ross (1967) the so-called syntactic "islands" defines the structural sentence environments or configurations out of which certain transformation operations cannot apply[149].

Under the *copy theory of movement*, positing that all the relevant information for interpretation is expressed at each site of the syntactic structure, the full representation of the above sentence would be having the a silent copy of the object notated with angled brackets: "Which Dissertation should I read <which Dissertation> ?" it is in fact at post-verbal position that the constituent *which Dissertation* receives from the verb 'read' the participant role that would have been assigned directly to it it if it hadn't moved. The constituent ultimately gets the participant role by being linked with its *gap* via what is called a "chain"[150].

Copy theory linked to movement operations appears in this way useful to account for the interface with semantic interpretation, in that it states that different positions of the sentence sub-serve different functions, one linked to the syntactic position where the syntactic element lands after displacement and one in the extraction site from which it was removed (e.g. post-verbal object position). We will resume this issue in in next section §2.4.3.

What is movement, exactly?

Given this last consideration, Movement appears than to be a particular type of Merge that instead of taking a external linguistic material (e.g. a lexical array) would move a constituent within the existing structure to a new position. As we saw in previous Section, within the Minimalist Program, hierarchical sentence structure is generated via a general Merge operation (see Figure 2.24), where two elements, A and B, are merged via feature-checking to create a constituent C.

The elements in Figure 2.24(a) may be 'atomic' or already the output of a previous merger, or even an internal constituent, this last case is called "internal merge" as illustrated in 2.24(b). This possibility of re-merge allows the operation of syntactic "movement" or displacement, where importantly a single constituent acts within the sentence as if it is occupying two structural positions. Therefore the gap position in the sentence can be see as the place where the displaced constituent is re-merged, this last observation will have important impact on psycho-linguistic study of the processes

theory; Chomsky, 1986), Within Minimalist Program a different approach was developed. While the details of the two theories differ, the basic function of *trace* and *copy* are the same, and we will generically call this empty phonological syntactic elements *gaps*.

149. See Constraints on Movement, later in §2.4.3, and subsequent works from Chomsky (1973/1977/1986) and Rizzi (Relativized Minimality, 1990) provided further and unifying explanations for the island constraints.

150. We can note that for copy theory of movement, all the copies are the same: all the instances contain phonological features. It is only at spell out, the interface with Phonological Form (PF), that an algorithm identifies which copy must be pronounced and which one left must stay silent. Generally the the highest copy is pronounced, but according to some analyses, there are cases where a lower one is preferred (Bocci, p.c.).

happening at *gap* position that we are going to review in the next sub-section.

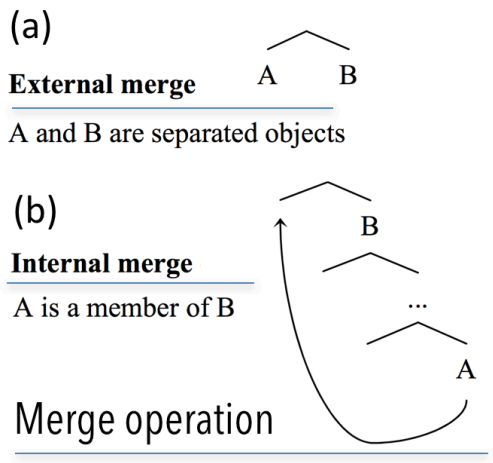


Figure 2.24 – Internal and external Merge schematic representation.

We can thus resume the different possible Merge operation in (1) Primary Merge [**the** cat], with two elements from the Lexicon; (2) Recursive Merge [cuddled [**the cat**]] with one element form the Lexicon and a complex syntactic object (i.e. the output of preceding application of Merge); and (3) [my sister [**cuddled** [**the cat**]] Phrasal Merge where two complex objects, resulting both from previous application of Merge, are merged together.

In this three cases the role of selector (in bold) and selectee is assigned, while in Phrasal Movement a *Search* operation is added within this primary buffer to identify the candidate for Phrasal Merge, and only after this Search step the suitable candidate is internally merged with the whole structure. We can then conclude that Movement appears to be a composite operation combining a *Search procedure* and a consequent possible Phrasal *Merge operation*.

Before moving to the characterization of the different types of movements this minimal External Merge operation allows, we will consider some neuro-imaging results that give initial evidence for the fact that syntactic movement constitute a possible complexity metrics that can find correlates in the brain.

Embedding versus movement

So far we have described neuro-imaging, neuro-psychological and psycho-linguistic findings that confirm the fact that the structural complexity of a syntactic-tree structure (and the internal relation it represents) are a possible account for the complexity effects observe at the level of the brain (e.g. §2.3.3.2, p. 145). Now that we introduced a new syntactic operation to be performed on the tree-like representation of the sentence, we can introduce a study that succeed in decomposing these two complexity parameter (i.e. shape of the syntactic-tree and movement operation) through a sophisticated adaptation experimental design.

The core idea of this fMRI priming paradigm was to enable the adaptation to two different types of syntactic properties of the experimental sentences in Figure 2.25 (A): namely, to the tree-shape (i.e. right branching or center-embedded in blue) in conditions (1) and (2), or to the movement operation of the subject or of the object like in condition (3) (in red). By this experimental design the authors were able to observe different processing difficulties, respectively linked to (1) the establishment of filler-gap dependency generated by wh-movement (subject vs. object filler-gap linking) in subject vs. object-relative clauses, and (2) the complexity of the syntactic-tree configuration opposing right-branching to center embedded relative clauses.

Their findings presented in 2.25 (B) indicate that movement in its filler-gap dependency complexity (see activation clusters in red) shows a adaptation effect in BA44 and BA45 within Broca's complex irrespective of the syntactic-tree configuration of the sentence. While the two complexity parameters, syntactic-tree shape and movement type, involved in the type of relative clause complexity activates BA44 and Left Precentral Gyrus (ventral BA6).

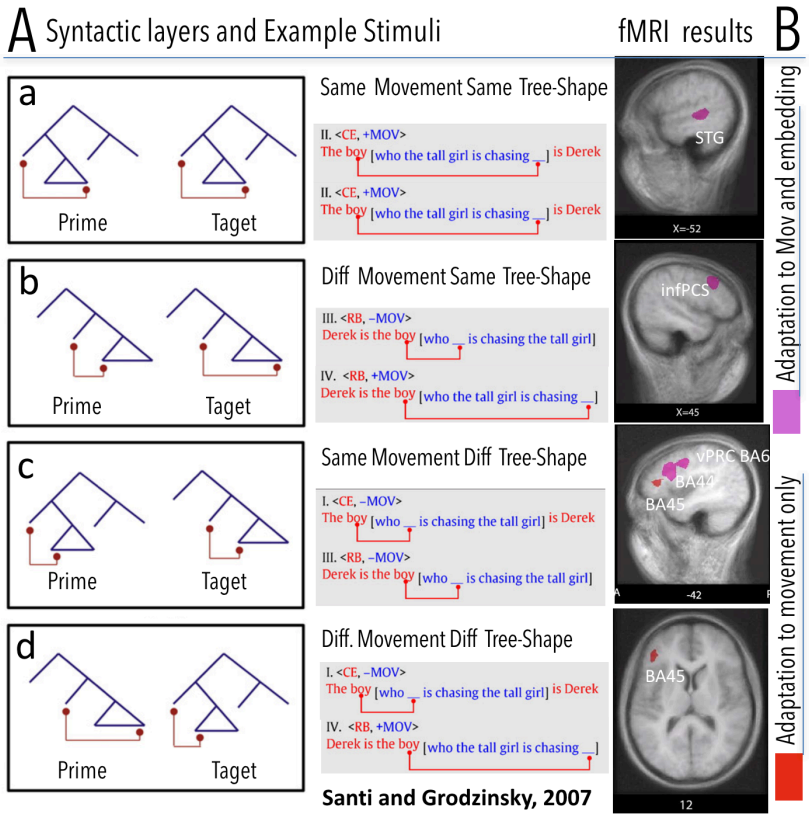


Figure 2.25 – (A) Experimental design and stimuli examples with their tree-like representation on the left. (B) Group-average brain maps of the syntactic adaptation effect to movement operation in red and to both movement and syntactic-tree shape in pink. Adapted from Santi and Grodzinsky (2010).

In sum, results reveal that four distinct regions adapted to both Movement and Embedding in a conjunction of the identical condition in (a), in which both embedding type and movement type is the same, is subtracted from those conditions in which one or both factors change: posterior Left Inferior frontal gyrus (LIFG/BA44), the right inferior Precentral sulcus (RiPS), and the left superior temporal gyrus (STG) inferior Left Precentral Sulcus.

Figure 2.26, illustrates another fMRI study that went in the same direction by differentiating in the brain syntactic embedding complexity metrics from movement operation, shows a interesting contrast between movement-derived Object relatives (a) and embedded declarative featuring a sentential complement (b) in Hebrew. Participants performed a grammaticality judgment task while presented with relative clauses like :

(40) Hebrew

a. relative clause : azarti la-yalda še-[Mary ra'ata t ba-park]
I helped the girl who(m) [Mary saw t in the park]'

b. embedded sentential complements: amarti le-Mary še-[ha-yalda raca ba-park]
'I told Mary that [the girl ran in the park]'

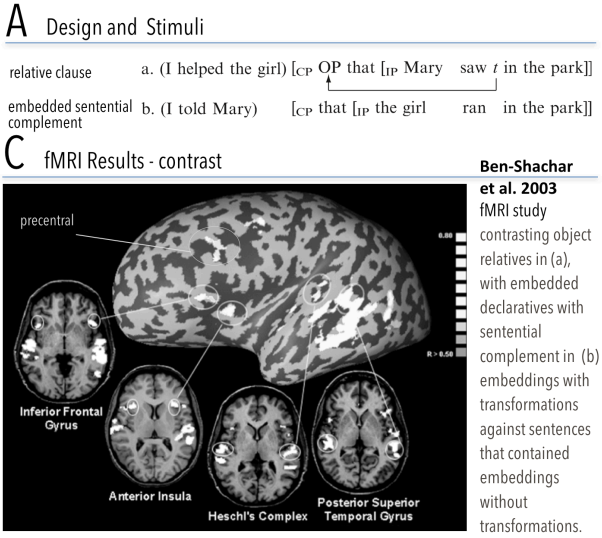


Figure 2.26 – (A) Experimental stimuli examples with schematic movement representation with arrows. (C) Group-average brain maps for critical contrast between Relative clause (a) vs. Embedded sentential complement (b) [a>b]. Adapted from Ben Shachar et al. (2003).

Results in Figure 2.26 (C) show that embeddings with transformations against sentences that contained embeddings without transformations activate Broca's area (BA44/45) and the posterior superior temporal gyrus (BA22/39) bilaterally.

The authors put in parallel the bilateral activation in pSTS for syntactic transformation with previous studies reporting bilateral activation in Wernicke's region as increasing with the distance between the sentence positions linked by transformation (i.e. antecedent-gap distance, see Caplan et al., 2002; Cooke et al., 2002).

They argue that Broca's (lIFG) region involvement in the critical contrast illustrated by (C) should be interpreted as transformation-related activation linked to the structural analysis that is needed in sentences containing syntactic movement compared with simple embedding involved in sentential complements. Moreover, the authors report an distinct effect in the left posterior superior temporal sulcus (pSTS) for an orthogonally manipulated variable liked to the argumental complexity of the verbs used in the different structures.

2.4.2.1 Movement Types

Another important aspect of Movement is that there exist different movement types in that do not operate on the same syntactic element and do not target the same positions in the syntactic-tree (i.e. the landing-site). This section briefly presents the several types of syntactic movements within the framework of theoretical linguistics and discuss four of them, which are relevant this thesis.

Two main types of syntactic position are distinguished within this framework: Argumental positions, A-positions, are positions to which a theta-role can potentially be assigned, like the subject position, or more generally positions that are associated to grammatical functions. Alternatively, A-bar positions, are defined negatively as non-argumental, like the specifier position of the Complementizer Phrase. Given this, we can already distinguish two types of movement, according to the position the moved element targets in the sentence tree-structure: *A-movement* as in NP-movement is distinct from movement to a *non-argument A-bar* position that is found in wh-movement. Let us briefly consider some linguistic phenomena which can be analyzed in terms of movement.

Wh-Movement

Consider first the following sentence constructions, all showing a long distance dependence between a gap and a filler:

1. Topicalization: That girl, I think he likes ___.
2. Wh-questions: Who do you think he likes ___?

3. Relative Clauses: The woman who you think he saw ___is tall.
4. Embedded wh-questions: I wonder who he thinks he saw ___.
5. Clefts: It's Anne that he thinks he likes ___.
6. Tough Movement: These problems are tough to even try to solve ___.
7. Correllatives: The more people I try to pretend I know ___, the more confused I get ___.

All of these constructions can be described with a single rule, instead of being considered individually, as isolated grammatical constructions (Chomsky, 1977, "On Wh-Movement"). Namely, some general principals hold for each one of those:

1. (1) they present the Displacement of a certain type of constituent *wh*-word to a non-argumental position, (i.e. A-bar, where no fixed grammatical function is assigned, as in the case of movement of a subject or object NP to a pre-verbal position in interrogatives);
2. (2) they have a long-distance dependency, whose range is constrained according to some *Islands effects* that would for example forbid the extraction of a constituent from a complex NP or from an indirect question (see *Subjacency condition*, p.186).

This can be exemplified in the contrast between the following two sentences:

(41) a. [I wonder [which problem [Baptist could solve *t* this way]]].
b. * [How do [you wonder [which problem C° [Baptiste could solve *t*]]]]?

Comparing (41a) and (b), note how in (41 b) the embedded Complementizer phrases introduced by a wh-constituent represent a syntactic *barrier*/island for extracting '*in this way*' and move it to sentence-initial position to generate a question '*How*'. We will further specify these issues generically called wh-Islands in the next sub-section (§2.4.3).

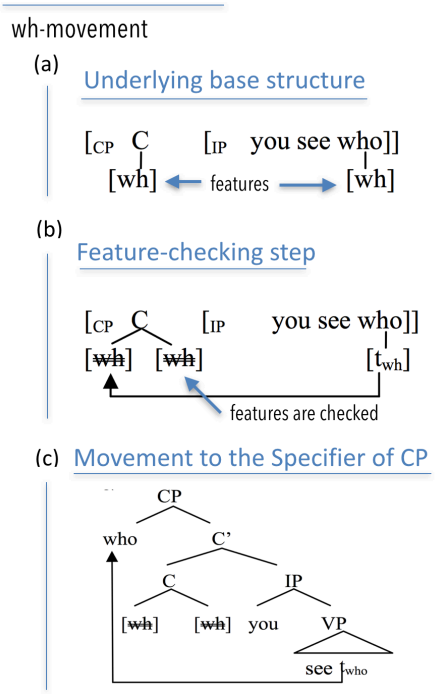


Figure 2.27 – Diagrammatic representation of the different transformation steps implied in the movement of the wh-word 'who' to sentence-initial position in question formation: from base-generated structure in (a), through feature checking step in (b) to movement of the wh-word to the Specifier position int he CP layer.

Another important characteristic is linked to what we previously exposed about the features of Complementizers in example (38). It is assumed that a sentence with an object Wh-word like 'which' has a Complemetizer head (C) that bears a [+WH] feature, marking the clause interrogative, and specifically a wh-question. The assumption is that movement is triggered by the need for feature-checking[151]. A head whose features have to be checked 'searches' into the sentence to find an element bearing its same features. When such an element is found, its features are attracted to the head for feature-checking. We might, then, reformulate the definition we retained until now for wh-movement in order to add the requirement that the Complementizer Head C to whose Specifier position (Spec-CP) the wh-phrase moves to, should have a C-head carrying the [+WH] feature.

151. Note that Starke (2001) approach introduces a difference between bare wh-phrase and the first are attracted by a simple attractor with a [+Q] interrogative feature, while the second is attracted to the CP by a a complex one carrying both [+Q, +NP] features.

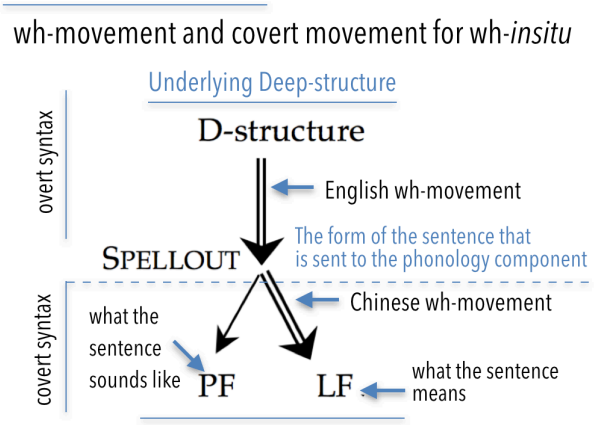


Figure 2.28 – Schematic multi-level representation of Extended Standard Theory (EST Chomsky): Deep structures is the domain that is intended to isolate the thematic relations and the insertion of lexical items in the process. Note that overt syntax is the domain where the movement operations, we will experimentally investigate, take place. For an detailed description of the diagram refer to the text and footnote.

Given this, we can describe 2.27 (c) interrogative clause as follows: the position in CP where the [+WH] feature is, has an agreement relation with the head of the clause C that is characterized by a [+/-WH] feature that was just checked as signified by the three bars in (b) and (c)[152].

(42) Mandarin Chinese Wh-insitu

a. 你看见了谁?

nǐ kànjian-le shéi
you see-ASP. who

'Who did you see?'

b. French Wh-insitu

Tu as vu qui?
you have seen who

'Who did you see?'

c. French movement derived wh- question

Qui tu as vu $t_{<qui>?}$
who you have seen *t*

'Who did you see?'

However, this dislocation pattern might puzzle the reader acquainted with Mandarin Chinese grammar as well as the reader who is acquainted with colloquial French or with all languages that practice *in-situ* question[153]. In fact, French allows the wh-word to remain *in-situ* (i.e. in its base-position), or to move to the CP-layer (in this case the beginning of the sentence). These two possible wh-questions are illustrated in (42) accompanied by a simple Mandarin Chinese example. Yet, how can the framework argued until now account for these phenomena?

There exist parameters for movement across languages that assign an overt versus covert value to a given movement. Such parameters determine, for instance, whether syntactic movement applies before or or after the spell-out of the surface structure (S-structure) as shown in Figure 2.28. It has been assumed that the difference between overt and covert movement lies in the syntactic level at which movement takes place, as we already saw for Verb rising in Head-movement.

Figure 2.28 illustrates the architecture that was retained for the grammar (1993)[154]. Since a covert movement can happen after the branching off to the phonology component

152. There exist theoretical alternatives to this version of Movement triggering. We will not discuss theoretical options like 'feature-strength' or EPP-feature here for reason of concision and because they are less relevant for our introductory discussion as we do not know how these different theories could be experimentally addressed.

153. A constituent is said to remain in-situ, if it does not undergo any kind of movement operation, thus *in-situ* question are formed simply by substituting the constituent being the scope of the question by a interrogative wh-word.

154. With the Minimalist turn Chomsky (1993, 1995) has suggested that language can be defined as being a perfect system of optimal design, that is to say that natural grammars generate structures which are designed to parfectly interface with the other components of the mind like the speech and thought systems. Namely, we can read in Chomsky (2005b:2) that '*Language is an optimal way to link sound and meaning*'. This assertion supposes that the kind of grammar of a language is organized as schematically

(PF), it is un-visible or heardable. Hence, the differences between English and Chinese resides then in when happens the operation 'Move' before Spell-out for English and after Spell out for Mandarin Chinese.

Unfortunately, we do not have in our experimental contrasts a set of stimuli allowing us to observe such parametric differences at the brain level. In the localizer stimuli (cf. §H) we actually selected sentences to oppose these over/covert patterns, but the experimental paradigm (i.e. mini-block design) is not meant to have direct contrasts between individual sentences. French and Chinese sentences featuring this parametric difference were inserted to maximize the typological and formal differences between the parametric settings of the two grammars at a more global level. Hence, we leave this interesting aspect of grammar for future investigations.

NP-Movement

Until now we have discussed syntactic transformations that moves phrasal elements having properties that distinguish them from regular Noun Phrases, like wh-elements, we will now turn to a movement type that displaces a regular phrase into a position where a fixed grammatical function is assigned. For instance, the kind of movement affecting the object and moving it to the subject position, an instance of Argumental movement (A-movement).

This movement type is also called NP-movement and distinguishes another set of syntactic phenomena, like passive (e.g. "The artist was chased by the thief."); cases of subject raising (e.g. "The thief seems to chase the artist."), and unaccusative verbs (e.g. "The pizza fell.". Consider the sentences represented in tree-diagram in (2.29), where NP-movement occurs because in their underlying form.

All sentence types have an empty subject position, because passive verbs, like '*be invited*', and raising verbs, like '*seems*' or Unaccusative verbs like '*fall*', take only one internal argument and, therefore, do not assign an external thematic-role to the subject position (Haegeman, 1994).

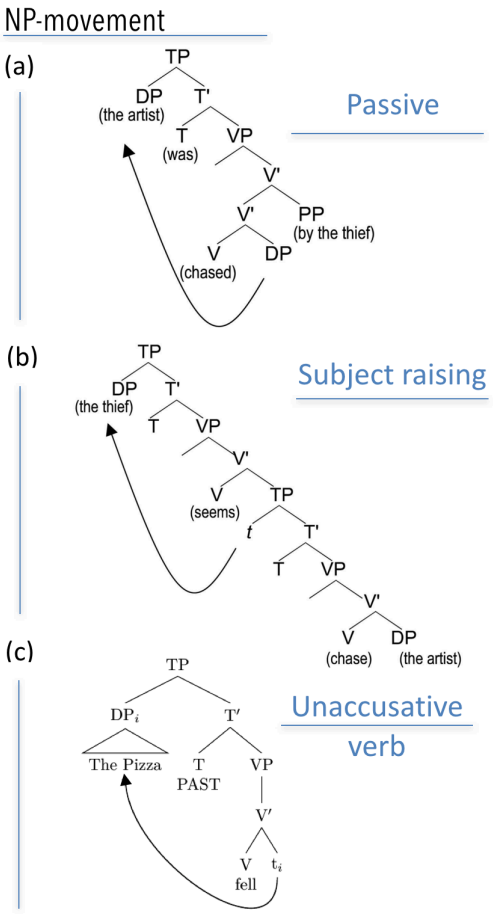


Figure 2.29 – Tree-diagrams of NP-Movement in three different syntactic configuration in English, respectively (a) passive (b) raising verbs like '*seem*', and (c) Unaccusatives verbs.

illustrated by the diagram in Figure 2.28. "The words taken out of the Lexicon are combined together by a series of syntactic computations in the syntax (i.e. in the syntactic/computational component of the grammar), thereby forming a syntactic structure. This syntactic structure serves as input into two other components of the grammar. One is the semantic component which maps (i.e. 'converts') the syntactic structure into a corresponding semantic representation (i.e. to a representation of linguistic aspects of its meaning): the other is a PF component, so called because it maps the syntactic structure into a PF representation (i.e. a representation of its Phonetic Form, giving us a phonetic spell-out for each word, telling us how it is pronounced). The semantic representation interfaces with systems of thought, and the PF representation with systems of speech." Taken from Radford (2009:14) in *Minimalist Syntax Revisited*.

Hence, because all grammatical English sentences must have subjects, the internal argument moves to the subject position.

There are, of course, other crucial differences between the two NP-movement structures. In passives (44 a), an object NP is moved to the subject position, whereas in raising structures (44 b), the subject NP is raised from a lower clause to a higher clause, resulting in an embedded sentence (Figure 2.29b).

We can see from these examples, that the two movements reviewed to this point, despite the distinction in landing sites, do share the characteristic of creating a chain with their antecedent and to leave a *gap* that is co-indexed with its antecedent[155].

(43) wh-movement
   i. Who/*Whom$_i$ do you think ___$_i$ will arrive first?
   ii. Who/Whom$_i$ do believe Anna will invite ___$_i$?

(44) NP-movement
   a. Anna$_i$ will be invited ___$_i$.
   b. Anna$_i$ seems ___$_i$ to be the best specialist.
   c. Anna$_i$ fell ___$_i$.

However, a fundamental difference exists on this last point, particularly in relation to case assignment patterns in the gap-antecedent relation, as illustrated in (43).

A closer look at these examples reveals that in (ai) the moved wh-phrase is nominative, while in (ii) it is accusative, it is the gaps of *who* and *whom* that are case-marked, and it is the case on the gap that will make this theta-role visible for the verb to assign it its object theta-role. Comparing this configuration to (i) and (ii) in 43(b), we see immediately see that gaps are not case-marked, the verb 'invite' does not assign accusative case to 'Anna' and so does the Unaccusative verbs 'seems' or 'fall'.

Thus, we can say that the *antecedent-gap configuration* is reversed between NP-gaps and wh-gaps: passive and unaccusatives in (b) it is the antecedent that is assigned a case, namely the subject role; while in the case of wh-movement in (a), the antecedent is not in a position where case is assigned, a non-argumental position (A-bar). It builds, namely, an A-bar chain where the head of the chain (the antecedent) is case marked and the foot of the chain (the gap) is not. NP-movement is the opposite it case-marks the antecedent and not the gap. This important property distinguishing these two movement types will be investigated in psycho-linguistic studies that we reserve to present in short.

Unaccusative verbs This type of movement can be found also in another type of verb category behaving like the verb 'seem' that we just saw. Linguistic configurations like in "The pate broke" are interpreted in the following manner: the NP argument of Unaccusatives starts in object position and moves to subject position, giving the following analysis [S [NP The plate]$_i$ [VP broke t$_i$]]. Under certain syntactic configurations, Unaccusatives also allow an alternative word-order (i.e. *there*-inversion) where the underlying object remains in object position like in example (45). Since Unergative subjects aren't generated in object position, they aren't allowed to appear there with *there*-inversion as shown by the ungrammaticality of (45a).

(45) a. *There danced three men at the palace.

155. Both movements c-command the co-indexed gap in the extraction site.

b. There arrived three men at the palace.

There is a great amount of cross-linguistic evidence that languages treat intransitive verbs assigning different syntactic position to the one argument that occurs with the verb. The sole argument is generally playing the role of a surface subject, but some verbs, there is evidence that the surface subject is an underlying object, in that it takes non-agents subjects. Namely, the notion of unaccusativity or split-intransitivity resides in a distinction between two types of intransitive verbs:

1. Unergative verbs: where the argument NP is semantically an agent or causer like *run*, *talk*, *shout*, *lie* or *smile*, these verbs can therefore be nominalized. And,
2. Unaccusative verbs, where the sole argument NP is a theme or patient, and most of all, where it can have some of the characteristics of the object of a transitive verb, like die, fall and sit. They are typically involuntary verbs, existential verb, movement verbs.

Interestingly, these difference between Unergative and Unaccusative verbs correlate with many grammatical phenomena that appear to be only possible with one of the classes of intransitives. However, different tests for the Unaccusative/Unergative distinction tend to apply in different languages, for instance, one of the first Unaccusativity tests, proposing auxiliary selection of tensed-verbs to distinguish between these two verb classes, works perfectly for Italian but is less stable for French. We will go into further details about the different Unaccusativity tests that can be used, in the description of our experimental material in chapter 6.

Head-Movement

Continuing our presentation on movement types, it remains to be said that not only phrasal constituents (complete XPs) can move like we saw in wh-movement or NP-movement. In fact, according to the nature of the linguistic element that is moved, we can further distinguish between Head-movements and Phrasal-movements. This shift in nature of the moved constituent comes with certain restrictions. Phrasal constituents can only move to phrasal positions, typically the Specifier position is the "landing site" for a displaced phrasal constituents. Symmetrically, heads can only move to head-positions, this is why this movement is also called *head-to-head movement*. This kind of movement can cross different syntactic layers (e.g. I-to-C movement, V-to-I movement), but it cannot 'skip' over another intervening heads in the structure (i.e. Head Movement Constraint).

The English and French examples in Figure 2.30 show two distinct patterns of movement for inflection which accounts for the different position of Verb in this two languages. In English, the lexical verb remains under V, and Inflection and Verb meet by lowering the infection to the Verb, thus giving the order Adjunct 'always'-Verb in (a). While in French, where the order Verb-Adjunct is observed, the Verb raises to I to unite with its inflection as in (c). Note that in both languages the revers pattern is not available, we can say that they are in a complementary distribution across languages, which has brought to establish a Parameter for Verb raising which is set as overt for French and covert for English[156].

156. It has been argued that French has a strong Inflection, given the richness of its inflectional paradigm compared to English, and that this property has the corollary of attracting the verb to rise it to the Inflectional syntactic layer.

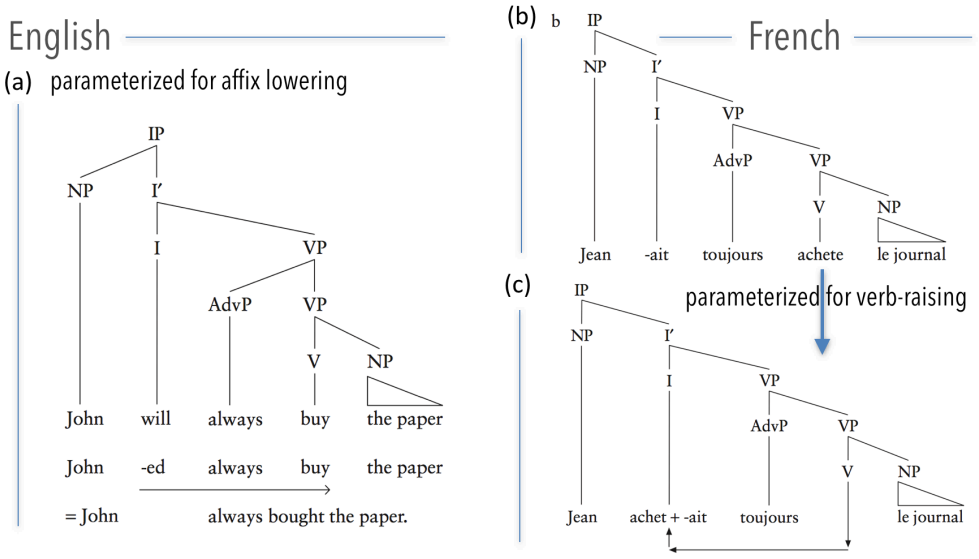


Figure 2.30 – Schematic representation of English (left) and French (right) syntactic transformation patterns in relation to syntactic movement parametric variation.While English features a parametrization for affix (e.g. *-ed*) lowering, French is parametrized for Verb-raising, for instance*achet-* moves higher in the syntactic-tree where the affix for future (*-ait*) is found.

English and French appear here to be distinct on the 'mobility' of their verbs, which we will confirmed hereafter by observing the different patterns the two languages have for yes-no Question formation. Specifically, this verb raising parameter characterizing French lies one of the interest of experimentally focusing on French Verb-movement as we will do in chapter 6 by investigating the neural underpinnings of French yes/no questions.

Before having a closer look to question formation in the following paragraphs, we can note that among the syntactic processes we have been characterizing so far, a typology of parameters can be sketched by distinguishing:

1. *Merge parameters*: ruling whether the Head selects to the left or to the right, does the Head precedes or follows the complement distinguishing Head-initial and Head-final languages as we saw in the Japanese fMRI study on embedded relatives (page 147[157].
2. *Move parameters*: As we saw in the example just illustrated in Figure 2.30 comparing French (V-to-I) and English, showing how a certain Head attract a lower head in I-to-C movement.
3. *Spell-out parameters*[158]: ruling whether a certain Head overt or null, as we saw for Topic heads (§2.2.4), certain languages posses over marking of Topic functional heads while other don't. Or if a Head licenses a null Specifier like in Topic Drop or Null Subject configurations (respectively p.114 and p.118).

This paves the way for discussing the contrastive approach to French and Chinese sentence neural correlate we adopted in developing language-specific localizers (§H) to allow

157. Also ruling what category does a given Head select.
158. See Figure 2.28 page 168 to be reminded of what is *spell-out*: the step where the sentence is sent to the phonology component.

a finer-grain and language-specific localization the Sentence cerebral network. Namely, designing these localizers we focused on tree different ways of accounting for the syntactic differences between French and Chinese: (1) a traditional construction-base account of sentence structure, (2) one based on typological difference like Topic-prominence, and (3) a more abstract account along the above-listed Parameters typology in relation to Syntactic operations.

Verb-Movement

In English subject auxiliary inversion, the sentence '[[Anna will eat the snails.]]' transforms into Will Anna eat the snails?' having an interrogative Force/clause-typing. This way of building yes/no questions in English, where auxiliary verbs invert with their subject is analyzed as the upward movement of the Tense head, triggered by the presence of a null question Complementizer (Ø[+Q]) in the CP-domain. This head-Head moment represented in Figure 2.31 is thus, giving a phonological content to the phonologically null Ø [+Q] complementizer by moving T to C, bypassing the subject.

While in English, only auxiliaries ever occupy the T head as free-standing entities. Main verbs do not raise to T in English an thus cannot undergo T-to-C, in French main verbs undergo V-to-T movement, and yes-no questions are formed by inversion, as an instance of T-to-C movement. Compare the following sentences:

(46) V-to-C head-head movement

a. Mangez-vous les escargots?

b. *Eat you snails?

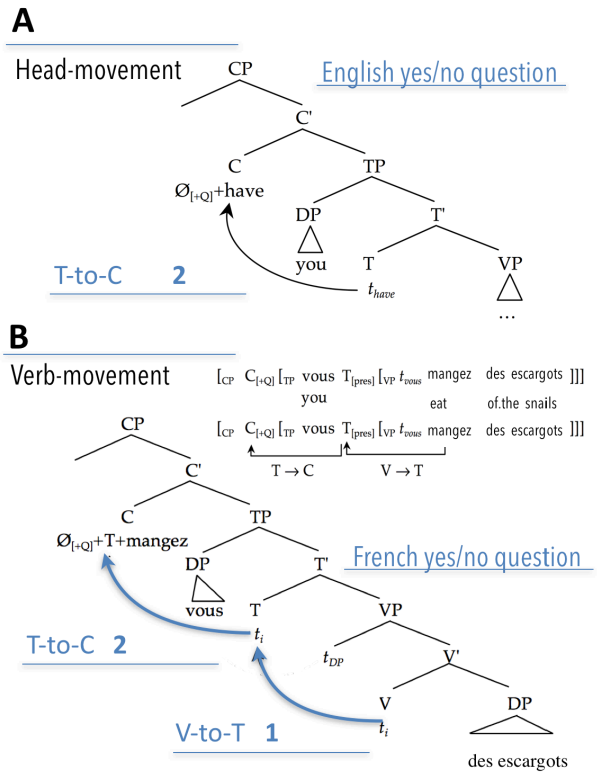


Figure 2.31 – (A) Tree-diagrams of Head-Movement in English yes/no questions. (B) Tree diagrams of Head-Movement in French yes/no questions, where the two steps of Verb movement are illustrated. In (B) the verb moves first from V position to T position in (1, V-to-T), and then to C position in (2, T-to-C), while in (A) only T-to-C is represented.

As we can see in the above tree-diagramm Figure 2.31 Head movement has another important characteristics, it is cyclic, that is to say that it occurs step-wise; like in our example the fist step (1) moves the verb from the V-head to the T node and successively moves to the C domain to receive interrogative Force[159]. We will see in chapter 6 that French partial interrogation like "Où dort-il?" will involve both wh-movement and Head-movement, leading to the following derivational analysis: [CP Où [+wh] [c° dort[+wh]] [IP il [I° *t*verbe]...]].

159. A chain is a sequence of co-indexed positions in a syntactic tree where each position locally binds the next position down

Clitic-Movement

Before we go any further and look at evidence for movement related phenomena in the field of experimental psychology, we will consider one last type of movement, an 'hybrid' one consisting in partly an NP-movement (DP-movement) and partly a Head-movement (X° movement), it will occupy us in chapter 6.

In French, clitics are a particular form of weak pronouns that always precede the finite verb. This tight relationship of the clitic pronoun with the verb is a characteristic of Romance languages[160]. The fact clitics cannot occupy the position in the clause which is usually occupied by full Noun Phrase they replace, has been taken as a starting point for a movement analysis displacing them to a pre-verbal position and assuming they are base-generated in the same position as the nominals they replace and leaving a trace at the base position (Kayne, 1975 and 2000).

Consider in (47) French examples where a weak object pronoun precedes the verb. All verbs selects a direct object, as evident in the agrammaticality of (g), thus for all grammatical examples it is postulated that a displacement of the object from post-verbal to pre-verbal position took place.

(47) Some clitic characteristics

a. "Anna regarde seulement lui." and not

Anna le (*seulement) regarde.
Anna 3SG.ClPronACC *only see

I him *only see.

b. "Ne viendra-t-il pas?" and not

*Ne viendra pas il.
NEG come.FUT not he

'Won't he come?'

c. Possible multiple clitics:

Je le lui donne.
I 3SG.ClPronACC 3SG.ClPronDAT give

'I give it to him'

d. Some written style exceptions

[...] en *fort bien* parler.
Clitic.of.it strong well [to] speak

'(I head someone) speaking very well of it'

e. Clitics are non-accented:

*Anna LE regarde, pas elle.
Anna 3SG.ClPronACC. look not her

'I loook HIM not her.'

f. To form a question "Tu l'a vu.' is transformed into:

160. For instance, in Slavic languages, like Croatian, the clitic pronoun always appears in sentence second position, independently of the position of the verb. See for further arguments H. van Riemsdijk (ed.), Clitics in the Languages of Europe. Berlin: De Gruyter, 543-579.

L'as- tu vu?
You him have seen > Him.have seen?
'I look HIM not her.'

Using the term clitic pronoun[161], we will assume that these are not only placed in front of the verb, but are attached to it. Namely linguistic analyses speak of a last step of incorporation to the verb, as illustrated in Figure 2.32B. Clitics can in facts not be separated from the verb, as illustrated by (47a) even when they are subject clitics like in (47b), the only possible intervening elements are other clitics like in (47c), or for written literary French style a clitic pronoun can be separated from the infinitival verb like in (47d). Among the numerous characteristics, we will later present in chapter 6, they have an important one that is also liked to their topicality, they cannot be accentuated like in (47e). Namely, clitics are used instead of full-Noun Phrases, when the intended referent is sufficiently identifiable by the listener.

From the point of view of syntactic derivation, the syntactic configuration is way more complex. And for the sake of our research focus on the neural underpinning of movement types we do not commit to any specific analysis of cliticization, a hotly debated topic in generative linguistics, since this would take the discussion too far afield. We will simply assume some general and stable proposals (Belletti, 1999; Sportiche, 1996, among others) that pronominal clitics may be considered the weakest form of pronouns, in that at the end of the cliticization process: the clitic ultimately is the head of a pronominal DP (D) and forms a word with the verb that hosts it (cf. a step of incorporation to the verb).

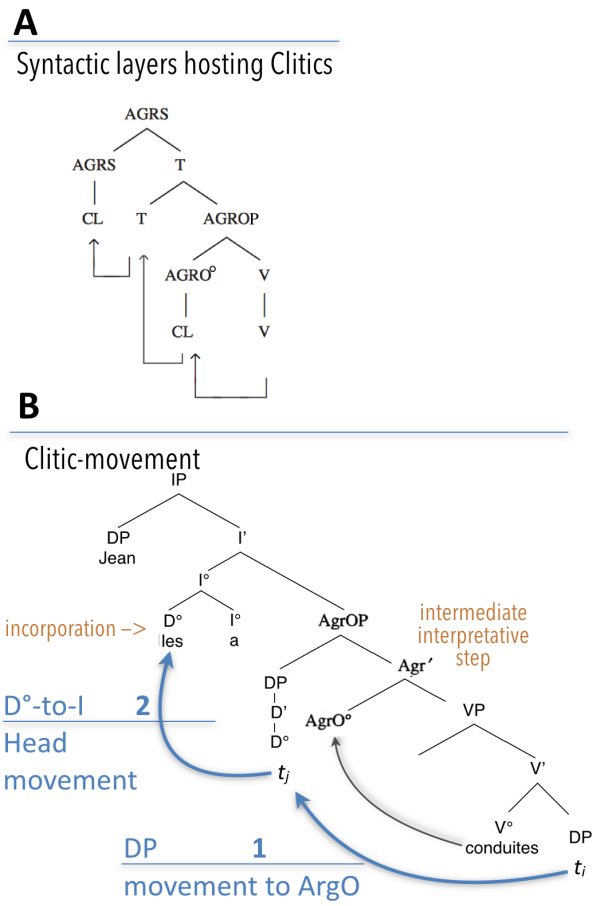


Figure 2.32 – (A) The two movement steps for clitic-movement with an intermediate interpretative step, assuming the detailed derivation of Belletti (1990) and Sportiche (1996). (B) Simplified tree-diagram of the sentence "*Jean les a conduites*" John [them] has driven[3pl.fem] 'John has driven them'.

Without going into further details about how Clitic-movement displaces pronouns, we can say it is another type of movement that shares the step-wise procedure we saw in Verb-movement. In the in sentence tree-structure, clitics are hosted by a functional head Infl (or AgrS, cf. Belletti, 1990) containing the verb as schematically indicated in Figure 2.32 (A). The fact that clitics are analyzed as being heads (cf. Kayne, 1975; Baltin, 1982) is supported by the observation that the clitic and its host undergo movement together as a unit in French interrogatives as in 47 (f), the verb moves to a head in the Complementizer domain, higher than the position of the

161. The term '*clitic*' comes from the Greek word '*to lean on*', convey the idea of the attachment to another word already in its etymology.

subject as we just saw in 2.31 (B) and in the presence of a clitic pronoun, both the clitic and the verb will precede the subject. Thus, given these considerations, we will assume that cliticization involves a further step after moving the object of the verb through NP-movement (DP): movement of the clitic pronoun as a head (X°), into a dedicated head position in the clause; the same ultimately also containing the finite verb and see in 2.32 (B), step 2. Figure (B) illustrates in a simplified manner the derivational steps for the clitic movement for the sentence "*Jean les a conduites*", John [them] has driven[3pl.fem], 'John has driven them':

1. NP-movement in Step 1, where the agreement with the past participle verb takes palace under AgroP, in 2.32 (B) the verb 'conduites', and the clitic is is indeed marked for number and gender.
2. And then Head movement in Step 2, from this position[162] further takes the clitic in adjunct position to the left of the head I° giving the surface SOV word-order.

In conclusion, the idea is that the hybrid nature of the clitic (as both a head and a maximal projection), is responsible for clitic-movement to be partly NP-movement and partly Head-movement. We will come back on some characteristics of clitics and clitic movement in next chapter and in chapter 6.

To summarize, in this discussion we have distinguished three types of movements:

1. wh-movement associated to wh-elements that move the the specifier position of the Complementizer phrase to give clause-type or assigning discourse properties to the sentence like in case of Topic or Focus (the so-called Scope-Discourse properties)[163],
2. NP-movement is the type of movement that takes place in passives and Unaccusative verbs, where an Noun Phrase is moved to an empty subject position;
3. Head-movement, the movement of auxiliaries from I-to-C to form questions in English or the movement of the tensed verb-to-C in French to form yes/no questions.

The table here under summarizes the different characteristics of the three syntactic movement that we have presented.

Table 2.4 – Comparing the characteristics of the different Movements types.

	A-bar Movement	Argumental movement	Head-Movement
Extraction position	Argument or Adjunct	Argument	Verbal Head or Determiner Head (V° or D°)
Displaced element	wh-element or NP	DP no case[-Case]	Verb or clitic
Landing-site	Specifier position of CP-domain (Spec-CP)	Specifier position of IP-domain (Spec-IP)	Head position in I or V
Trigger	wh-Criterion among others	Case theory	Flexionally rich
Derivation path	Cyclic through Spec-CP	Specifier position of VP-domain (Spec-VP)	Cyclic from V to I and to C
Locality constraints	Islands	DP-trace	Constraints

162. accoding to Belletti (1999), the movement is both to Agr(O) and from Agr(O).
163. More precisely, this movement is related to clause typing and to the need to establish the needed operator-variable configuration for the semantic interpretation of the *wh*-element.

This brief and partial outline obviously only offers a small fragment of the deep analyses that have been proposed for movement phenomena, but it will enable us to illustrate in the coming pages, how neuro-imaging studies and neuro-psychology have been offering initial experimental evidence for some of the formal linguistic distinctions we presented. In this way, the different readers have been given an elementary insight on the linguistic phenomena that have been selected in the different paradigms of Part II, and a background to understand their experimental hypothesis.

2.4.2.2 Neuro-psychological investigation of Movement

In psycho-linguistics and aphasiology, movement has soon been identified as a major contributor to the perceptual complexity of sentences in healthy subjects performance (Fodor et al., 1974 engaged in empirical work in psycho-linguistics within the Chomskian framework; see also Neville et al., 1991), but it very soon became a central tool in the evaluation of aphasic linguistic performance. Thanks to the work of Josef Godzinsky, among others, the hypothesis that the impairment of the Broca's aphasics could be given syntactically-based definition involving Movement was early introduced in the field. His 'Trace-deletion Hypothesis' have been claiming that the agrammatic syntactic impairment precisely involved the impairment of movement-related syntactic representations like *traces* (see a detailed discussion in Grodzinsky, 2006).

Agrammatic remediation through training on Movement

A second domain where syntactic movement has proven to be neuro-psychologically relevant is in remediation studies of agrammatism. While therapy for agrammatism used to target surface grammar, recent developments, in the years 2000 (Treatment of Underlying Forms in Thompson et al., 2003 and before that Thompson et al., 1997), have attracted researchers' attention on more theoretically motivated therapy, moving toward a linguistically-based approach for treating aphasic sentence deficits, namely targeting the "*underlying syntax*" of different sentential constructions.

Note that this approach focusing on the underlying representation of syntactic structures has begun to demonstrate a benefit of considering details syntactic representation like syntactic-trees and formal linguistic theory to investigate the cerebral representation of sentence's structure in the human brain. Given the evident scientific relevance of this approach for our research work, we will harvest its experimental and neuro-psychological findings by presenting Cynthia K. Thompson's work on Movement types, and Naama Friedmann's findings and theory on agrammatic and impaired sentence comprehension and production in the next sub-section (§2.4.3, p.189 and §2.4.4.5).

Treatment of Underlying Forms for aphasics

The methodology of the Treatment of Underlying Forms (TUF) consists first in selecting a non-canonical (complex) sentence structure derived by movement, like wh-movement derived object relative clause in English (Thompson et al., 2007), and to perform a training on it until the patients' comprehension or production performance on the trained construction is increased. As illustrated in Figure 2.33C, the assessment of the treatment is then carried on by testing for the generalization of the increased mastery (i.e. comprehension or production performance) on other un-trained sentence

structures that present the same type of movement-derivation, like it is the case for object wh-questions or object-clefts. One structure is usually trained at a time, while untrained sentences are only tested for generalization. Hence, this method operates on the hypothesis that training underlying abstract syntactic properties of the sentence will provoke generalization to untrained structures derived by the same movement operation.

In the last decades this method has repeatedly benefited the comprehension and production of the target and trained syntactic structure in individuals with mild to moderately severe agrammatism (e.g. Broca's aphasia with characteristic deficits patterns; see for a review Thompson and Shapiro, 2007; Thompson et al., 2010). It has been shown that sentences included in the treatment are acquired and remain significantly above baseline performance levels through follow-up phases, as shown in the diagrams reported in Figure 2.33 A and B for both object wh-questions and object-cleft. Moreover interesting generalization patterns are to be observed as far as movement types and syntactic complexity are concerned.

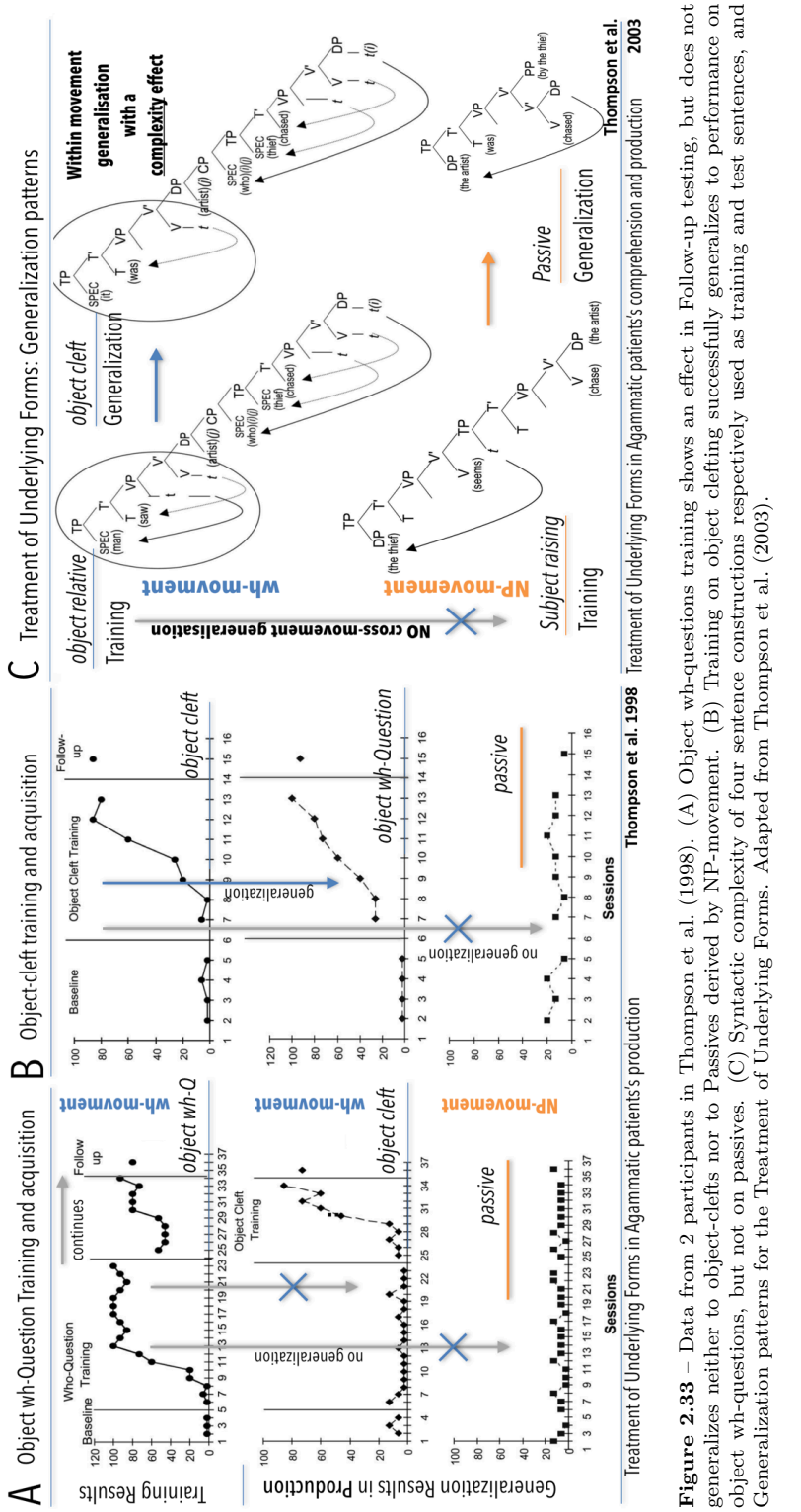


Figure 2.33 – Data from 2 participants in Thompson et al. (1998). (A) Object wh-questions training shows an effect in Follow-up testing, but does not generalizes neither to object-clefts nor to Passives derived by NP-movement. (B) Training on object clefting successfully generalizes to performance on object wh-questions, but not on passives. (C) Syntactic complexity of four sentence constructions respectively used as training and test sentences, and Generalization patterns for the Treatment of Underlying Forms. Adapted from Thompson et al. (2003).

Complexity effect in generalizations patterns A central assumption since Thompson et al. (1998) is that training has to be done on more complex wh-movement configurations in order to generalize to less complex sentence structures. Figure 2.33 (B) shows that when treatment was applied to object-clefts (1) wh-question production spontaneously emerged with no treatment provided, and (2) similar learning curves were observed for both constructions. On the contrary, comparing this generalization patterns with learning curves in 2.33 (A), we see that participants who received initial treatment focused on wh-questions showed no generalization to object clefts. Concretely, training object-clefts improved wh-question production, while the contrary did not yield any generalization effect.

Interestingly, this complexity effect was replicated in the learning and generalization patterns reported by other studies, where training sentence was a object relative clause (more complex) and test sentence structure was an object-cleft, as shown in 2.33 (C).

A closer look at syntactic complexity reveals that both object-clefts and object relatives can be considered more complex than wh-questions based on (1) the complexity parameter of the depth of embedding, and (2) on the presence of additional movement of the subject NP 'thief' from within the Verb Phrase as shown in the syntactic-tree representations in 2.33 (C).

Figure 2.33 (C), concretely shows how object relatives display additional syntactic complexity compared to object-cleft, in that object relatives also require a lexical verb 'saw' in the main clause to assign thematic-roles, while the presence of the copula in the object-cleft main clause does operate a thematic-role assignment.

As for NP-movement derived structures, the complexity parameter that distinguish subject raising constructions from passives can be summarized as follows: in passive an object NP is moved to the subject position within the same clause, whereas in subject raisin, we can see that the subject is moved from a lower clause to a higher position in the tree, thus resulting in an embedded configuration. Moreover, as noted for wh-movement derived structures, in subject raising there are two verbs, the raising verb 'seems' and the two-argument verb 'chased'.

No cross-movement type generalizations Interestingly, an early work by Thompson et al. (1998) – later replicated in Thompson et al. (2003) – tested in two agrammatic subjects the recovery of both wh- and NP- movement derived sentences. The target sentence involved movement of noun phrases (NPs), derived by wh-movement (Object-cleft) or NP-movement (passives), and patients were sequentially trained to produce either object wh-Questions and object clefts (wh-movement) or passives and subject-raising structures (NP-movement). Behavioral results reported in Figure 2.33A, B and C showed generalization patterns that are constrained to movement types, that is to say, training wh-movement structures resulted into generalized production of untrained wh-movement structures without influencing production of NP-movement structures and vice versa.

We find in the absence of an across-movement generalization some initial evidence in favor of considering NP-movement as neuro-psychologically distinct from wh-movement, which will represent one of our experimental hypothesis on French syntactic transformation (in chapter 6). The non-generalization from syntactically more complex object relative to less complex NP movement derived sentence structures[164] in facts indicate

164. For instance, the distance between the moved antecedent and the gap site is greater for wh-

that movement type and not general complexity is trained in this agrammatic population. The understanding and production of Movement to an argument (A-) position (as in NP-movement) appear to be distinct from that of sentence presenting movement to a non-argument (A-bar) position, as required in wh-movement. The 'landing site' where movement terminates in the tree-structure of these non-canonical sentences appears to influence sentence production. These findings show that linguistic properties of sentences syntactic-tree influence sentence production breakdown and recovery in aphasia.

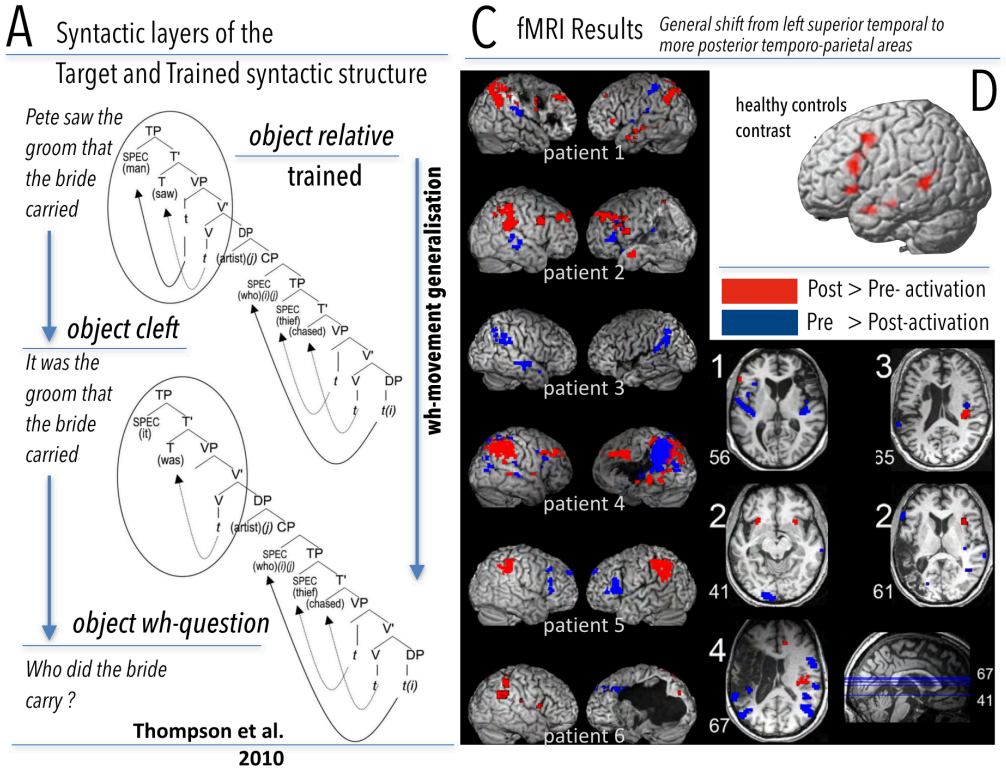


Figure 2.34 – (A) Syntactic-tree structure of the trained (object relatives) and target sentence structures (object-cleft and object wh-question) in Thompson et al. (2003). (C) fMRI activation patterns in 6 patients to pre- and post-treatment contrast from the auditory verification task on Subject cleft, Object clefts and simple past tense declarative sentences in Thompson et al. (2010). (D) Brain map of the complexity contrast [Obj. cleft> Subj. cleft] in the healthy control group.

In Thompson et al. (2010) aphasic participants performed an auditory verification task in the scanner, on object cleft, subject cleft, and simple active sentences, prior to and following the Treatment for Underlying Forms. The treatment resulted in improved production and comprehension of these structures and generalization to less complex wh-movement constructions, as has been seen in previous studies with agrammatic aphasic patients (for a review see Thompson and Shapiro, 2007)[165].

As for the pre- and post-treatment fMRI data, the treatment gains were mapped onto

movement structures than for NP-movement.
165. Note that this work also showed that the treatment appears to affect processing of trained sentences in real-time.

the brain activity for six individuals with stroke-induced agrammatic aphasia, as shown in Figure 2.34 (C).

The differences associated with pre- to post-treatment processing of the targeted complex sentences in 2.34A, reveal that the recovery of target syntactic structure (i.e. object-cleft) operate a general shift from left superior temporal activation to more posterior temporo-parietal areas, bilaterally. Despite individual variation in activation differences from pre- to post-treatment scans, the authors interpret the fact that angular gyrus showed a post-treatment increased activation in the lesioned hemisphere (even for patients whose lesion did not involve adjacent tissue) as the mark of a recovery mechanism, that recruits cerebral tissue outside of the network primarily activated by healthy controls (shown in Figure 2.34D) for processing complex syntactic material, such as [Obj. cleft versus Subj. cleft] (Den Ouden et al., 2012). In the case of Angular Gyrus, an area that plays a more peripheral role in syntactic computation like verb argument structure processing (Den Ouden et al., 2012 and Thompson et al., 2007/2010).

Although these findings should be taken cautiously because of MRI methodological difficulties linked to hypo-perfusion of tissues in post-stroke brain and the consequent time-to-peak delay of the hemo-dynamic response function (HRF), it remains that being able to investigate brain activation patterns after remediation studies is a promising perspective for the understanding of the encoding of syntax in the brain .

2.4.2.3 Neuro-imaging on movement types

Additional evidence for hypothesizing a neuro-psychological distinction of movement types can be found in recent neuro-imaging literature. While in the majority of the neuro-imaging studies on syntactic movement special attention has been given to the internal architecture of a given syntactic construction, that of relative clauses,as we saw for English and Japanese (Fig. 2.16, p.147), very few studies have investigated the cerebral underpinnings to wh-movement with respect to other movement operations, and only a few investigated it in other sentence constructions.

Shetreet and Friedmann (2014) investigated the distinction between wh-movement in Topicalization and Verb-movement in a Verb-second language like Hebrew. fMRI results reveal separate neural activation patterns, as illustrated by Figure 2.35. As we can see from the experimental conditions reported in example (48), the word-order variations from the SVO order in (a), feature two different types of syntactic movement: in (b) the object moves to the beginning of the sentence through wh-movement, and in (c) the verb moves to a pre-subject position through verb-movement.

(48) Experimental stimuli from Shetreet and Friedmann (2014)

a. Hebrew [SVO] declarative

Ha-yalda nishka et ha-doda etmol
The-girl kissed ACC the-aunt yesterday

'The girl kissed the aunt yesterday'

b. Hebrew [OSV] topicalization

Et ha-doda ha-yalda nishka etmol
ACC the-aunt the-girl kissed yesterday

'the aunt, the girl kissed [her] yesterday'

c. Hebrew [VSO] Verb movement - Verb second

Etmol nishka ha-yalda et ha-doda
Yesterday kissed the-girl ACC the-aunt

'Yesterday the girl kissed the aunt.'

For wh-movement in Hebrew Topicalization, the authors report an increased activation in the left Inferior Frontal Gyrus (Broca BA44/45) and in bilateral temporal regions. This replicates what was found for other syntactic structures, like relative clauses and wh-questions, even though topicalization does not share with them the same semantic/pragmatic features. This could tentatively be interpreted as showing that Broca's area and temporal regions activations in wh-movement are most likely linked to syntactic processing, rather than to pragmatic calculations implied by topicalization of wh-questions.

Additional activations for Topicalization were found in right Inferior Frontal Gyrus (BA47), left Precentral gyrus, and the left cerebellum. While verb-movement effect, obtained comparing [AVSO > ASVO], elicited a single activation cluster in the left inferior occipital gyrus (IOG/BA17, lingual gyrus). The comparison between these two movement effect ([OSVA > SVOA] > [AVSO > ASVO]), yielded Broca's area activation, with left posterior temporal regions and in the medial Superior Frontal Gyrus. This last brain region could be interpreted as reflecting the pragmatic discourse calculation linked to topicalization that is absent in Verb-movement. As for Verb-movement related activation, the authors interpret it as being linked to the phonological role of IOG by advocating the syntactic analysis that the displacement in sentence's second position a covert one happening at the level of the phonological component, as it has been argued above for the wh-movement parameter differentiating English and Chinese. Interestingly, activation in the right Broca complex (BA47) is also interpreted as being associated to theory of mind and discourse functions that characterizes Topicalization.

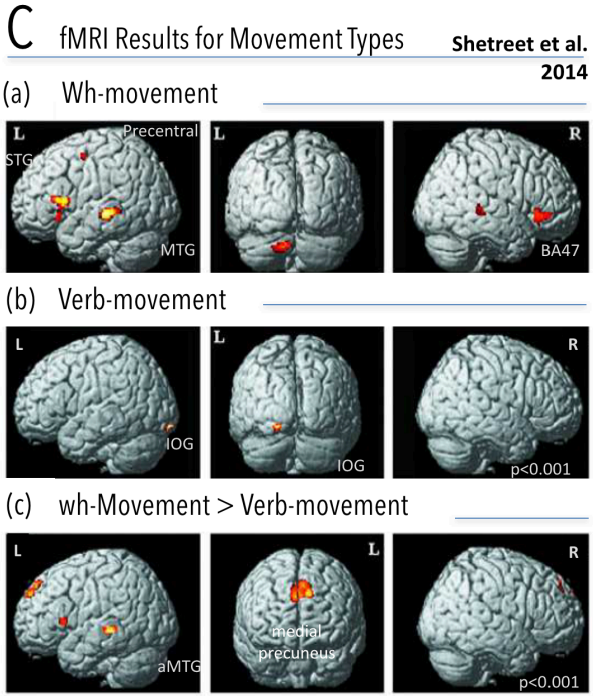


Figure 2.35 – (C) Group-average brain maps for critical contrasts showing (a) wh-movement and (b) V-movement effects against SV0 baseline, and in (c) wh-movemnt versus V-movement contrast mediated by two different baselines, i.e. [OSVA > SVOA] > [AVSO > ASVO]. Adapted from Shetreet and Friedmann (2014).

Importantly, as discussed in previous section (§2.4.2.1, the difference between these two movements is theoretically relevant (see Table 2.4, p.176), and these imaging results appear to support the neuro-cognitive relevance of this theoretical linguistic distinction, we will build on this initial experimental evidence to further investigate the neural underpinnings of different types of syntactic movements in French.

Another very recent fMRI study by Europa and Thompson (unpublished Poster), echoes the same experimental question about the existence of regions of the brain preferentially activated for different movement types using different syntactic structures in English, namely Object clefts contrasted with Subject clefts to observe effects of wh-

movement, and passive opposed to active sentences for NP-movement effect (see Figure 2.36A and B). The experimental design accompanied sentence comprehension with a

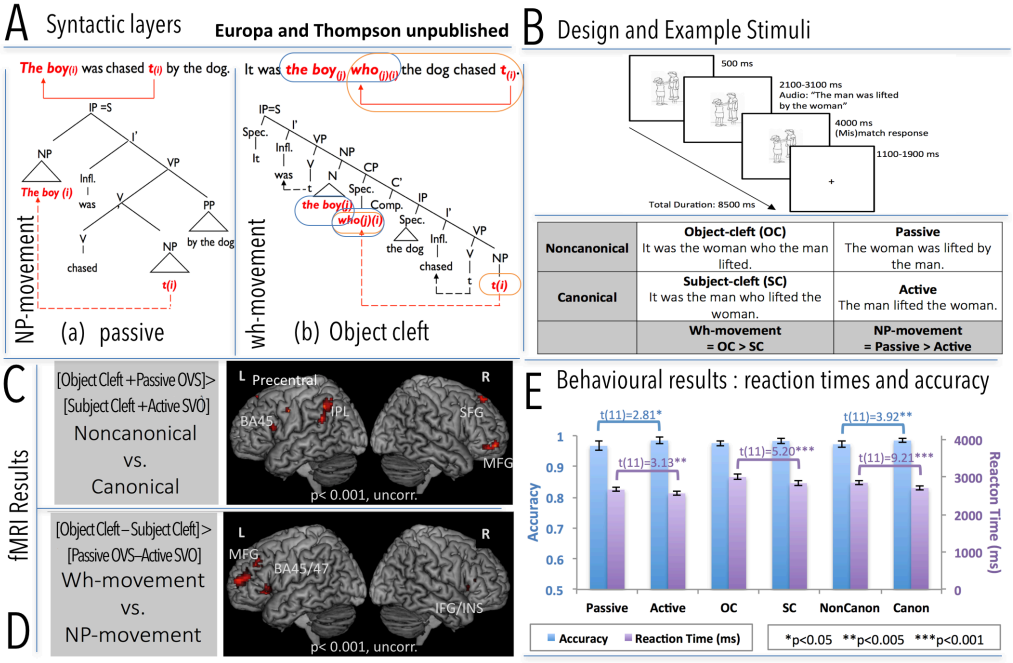


Figure 2.36 – (A) Syntactic-tree representation of the two experimental condition involving NP-movement (a, passive) and wh-movement (b, Object-cleft). (B) Experimental design and sentence stimuli examples. (C) Group-average brain map for the contrast [no-canonical (passives + object-clefts)] > [Canonical (active declaratives + subject-clefts)]. (D) Group-average brain map for the contrast [wh-movement (Object-clefts s. Subject clefts)] > [NP-movement (passive vs. active declaratives)]. (E) Behavioral results for the sentence-picture matching task: reaction times and accuracy by sentence types.

picture verification (see 2.36 B) task that revealed a generalized drop in accuracy and increase of Reaction times in movement derived-sentence structures (i.e. passive and OC) compared to canonical word-order conditions, as shown in Figure 2.36 (E).

The direct comparison of Wh-movement and NP-movement yielded (2.36 D) left frontal activations in Middle Frontal Gyrus and Broca complex (BA45/47), with a small right lateralized activation in Broca adjacent to the Insula, but no significant clusters were found for NP-movement compared to Wh-movement.

The authors decided also to test for a more general complexity measure that is linked to the canonicity of agent-before-patient word-order, and compared (Passive + Object-Clefts) versus (Active + Subject-Clefts) and found bilateral frontal and left inferior parietal cortices, and nothing was found for the opposite contrast as illustrate in Figure 2.36 (C) . Notably, inferior parietal lobe activation in this contrast may reflect thematic reanalysis processes involved in understanding object-before-subject linearization. Broadly speaking, these findings suggest that processing wh-movement requires greater neural resources than NP-movement especially in Broca complex. This pattern that is interpreted by the authors as reflecting the increased cognitive resources needed by the processing of

two different co-referential relations in wh-movement (see Figure 2.36Ab): (1) the first between the *g* and the wh-element 'who', circled in orange, and (2) the second between the wh-element 'who' and the direct object 'the boy', circled in blue.

Although not easy to interpret, these two studies are the first neuro-imaging findings on which we will lean for our own study on French movement types. Anticipating on chapter 6, our experimental approach to syntactic movement transformations will namely involve, wh-movment in object wh-questions, Verb-movement (V-to-C) in typically French yes-no Questions (typical in that the main verb moves), NP-movement in Unaccusative verbs and Clitic-movement of the object an of the locative argument.

2.4.3 Gaps, traces and resumptives

After discussing the nature of the moved constituents and introducing the notion of *landing site*, we now move to the arguments for positing *gaps* at the *extraction site*.

For example in wh-movement, positing these silent copies ensures that the displaced phrase is interpreted as the object of the verb, just as a non-displaced phrase would be. We will now concentrate on the abstract syntactic position vacated by movement: the *gap*. As briefly mentioned when we discussed the neuro-psychological reality of Movement syntactic transformation (§2.4.1.1, p. 159 and 2.4.2.2, p. 177), this operation causes a phrase to occurs in a syntactic position that is different from the basic position which determines its basic semantic role, and a *gap*, an abstract marker needs to be posited in the sentence in order to maintain the structural relation between the surface form of a sentence and its underlying form[166].

Importantly, assuming null elements as a syntactic element, a component of grammar of all natural languages, parallelly goes with the assumption that language learners have the ability of postulating these silent null elements and to place them in the representation of the sentence unit that the learner builds.

(49) a. We *all* could have been being punished for that.
b. We could *all* have been being punished for that.
c. We could have *all* been being punished for that.
d. We could have been *all* being punished for that.
e. *We could have been being *all* punished for that.

Exactly as he learns the possible and 'legal' distribution of the word 'all' in the sentences in (49 from Harwood 2012:7)[167], the learner needs to learn the constraints on Long-Distance Dependencies for the placement of these silent gaps.

166. We should acknowledge here the existence of other syntactic theories that do not theorize abstract syntactic position to account for the same syntactic phenomena. However,given the amount of converging evidence confirming the psychological and neuro-psychological reality of empty syntactic positions (i.e. gaps), we decided not to consider such alternative theories. Yet, it could be interesting to in the future to consider drawing direct comparisons between the predictions that such theories for the experimental contrasts and paradigms we presented tin this chapter and for our two fMRI studies that will address this issue in chapters 6 and 7. Note that in Minimalism the distinction between Surface and Deep structure has been replaced by a unified account we presented introducing the mechanisms of External Merge.
167. The reader might remind the presentation of the Cartographic project that was done in chapter 1 (p. 34) and the elaborate functional hierarchy for the Cinque (1990) proposed. This example echoes the evidence found for clausal (TP/IP) hierarchy in the ordering constraints of adverbs across different languages. But further considerations on this issue an on this example would take us too far astray.

Although the so-called filler[168] can be found at a long-distance from the gap[169], like in "[[This colleague,] [I think [he told you] [he tried to like ___.]]]]", there exist constraints on gap placement, and the filler-gap distance is to be measured in how many clauses the filler is from the gap (not in number of words). Ross (1967) formulated a few of these constrains in wh-movement:

1. the gap cannot be inside a coordinate structure. I saw [the boy and the girl]. > *Who did you see the boy and ___.
2. the gap cannot be inside a sentence that is inside a noun phrase: I like [the fact that he reads books every day]. > *What do you like the fact that he reads ___every day?

3a. the gap cannot be inside the subject: [Pictures of Anna] were available. > *Who were [pictures of ___] available? [Books about linguistics] were on sale. > *What were [books about ___] on sale?

3b. But, the gap can be inside the direct object: You saw [pictures of Anna]. > Who did you see [pictures of ___] ? You read [books about linguistics]. > ?What did you read books about?

4a. The gap cannot be inside an embedded question: They wondered [who talked to Anna]. *Who did they wonder [who talked to ___]?

4b. While the gap can be inside of a plain embedded clause: 'They thought [(that) we talked to Anna]'. or 'Who did they think [(that) we talked to ___] ?"

5. The gap cannot be inside a relative clause or any another long distance dependency: 'I like [the boy that Anna speaks to ___.]'
   *'Who do you like [the boy that ___plays with ___].'

In the above examples, the syntactic units where gaps cannot be found are generically called *Islands*, following the term introduced by Ross. These facts have been given a unified account by advocating a *Subjacency Condition* on movement, which defines the boundaries for syntactic movement, thus determining how far a dislocated sentential element can go through syntactic movement. To explain the detailed analysis that preempted its proposal would lead us too far astray, so we restrict ourselves to its formulation: Movement cannot cross more than one bounding node, where bounding nodes are IP and NP[170].

Having posited these boundaries and rule, what remains to be accounted for is how is the antecedent-gap relation established at the trace site in absence of phonological input. Using again the metaphor of language leaning, we could reformulate the question saying that the learner has not only to know where to posit gaps -knowing how non-overt syntactic elements are licensed in the sentence structure-, but also the way to give content to non-overt elements like gaps. We already met some silent position at the beginning of this chapter (§2.1.1, p. 86 ex.4) presenting binding rules for co-reference. Yet, consider the following examples:

(50) a. Christian$_i$ promised his boss$_k$ [PRO$_i$ to shave himself$_i$]
     b. Christian$_i$ persuaded his boss$_k$ [PRO$_k$ to shave himself$_k$]

168. Definition: the preposed sentential element filling the empty syntactic position in a movement-derived construction.
169. We will later introduce another way of understanding the distance between a filler and its gap, by counting the number of discourse referents intervening between the gap and the antecedent, to account for some psycho-linguistic aspect of maintenance in working memory.
170. Our insistence on Subjacency is mainly due to the argumentative tone, as it offers an empirical argument to show the king of "rules" or constraints the brain should represent and follow to place gaps.

(51) Which picture of himself$_i$ did Christian$_i$ choose ___?

(52) Which picture of himself$_i$ did Christian$_i$ choose <which picture of himself$_i$> ?

Apparently subjectless infinitive clauses contain an "understood" null subject, which qualifiers as a valid antecedent for the anaphora in 50 (a) and (b). Since the null subject found in the two infinitive clauses in 50 has mainly the same grammatical properties as a pronouns it is conventionally designated as PRO (big-PRO). Its interpretation is "controlled" by constituents of higher clause. If the higher verb has an object, it is the object that usually supplies the subject of the lower verb like in (b), with the exception of the verb 'promise' as attested by (a).

We can note that compared to Binding, the previously encountered rule governing overt elements of referential relation in the sentence, the Control rule governing empty pronouns like in (a) and (b) seems to involve a straightforward notion of "locality": select the closest -in terms of tree-structure- possible antecedent for the verb in question.

Another important aspect shown in the referential attribution of the two reflexive pronouns 'himself' in 50, is that the Binding of gaps, occurs following Binding principle applied to lexical constituents. In (a) and (b) empty pronouns co-refers with the reflexive "himself", thus establishing a different referential link with the Nominal elements in the main clause: the empty pronoun PRO is co-indexed once with 'Christian' for (a) and once with 'the boss' in (b).

	Lexical	Empty
[+ anaphora] [– pronominal]	Anaphor	NP-*trace*
[– anaphora] [+ pronominal]	Pronoun	*pro*
[– anaphora] [– pronominal]	Referential expression	wh-*trace*
[+ anaphora] [+ pronominal]	---	PRO

Figure 2.37 – Typology of empty categories. Inventory of the different types of gaps presented, according to the features [-/+anaphoric] and [+/-pronominal].

Long-distance dependencies not only constitute one of the most distinctive properties of natural language, but also provide a window to observe some broader syntactic architectural constraints, as we saw presenting the constraints on distribution of the different types of NP-like anaphors (2.1.1, p.86). Analogously, there exist constraints on where the gaps can be, and on what kind of empty element can be found at the original position of the displaced constituent. Summarizing the empty elements seen till now, we can say that they are mainly divided into two groups on one side traces, NP-gaps/(DP) and wh-gaps, and on the other side null pronouns, comprising pro and PRO. Their different properties lead to establish a typology for Empty categories, that will complement the one established for NPs along the features [-/+anaphoric] and [+/-pronominal] (see Table 2.37)[171].

A recently published fMRI study offers a neuro-imaging insight on these distinctions between different empty and lexical categories in the sentence. Matchin et al. (2014) contrasted English sentences featuring a gap in a wh-question, and a pronoun in Backward Anaphora, as shown in Figure 2.38(A). They report an interesting temporal/frontal

171. We present here a typology of Empty Categories to introduce the reader to the fact that empty elements can be declined into several types. This will be further explained in chapter 3 through several syntactic tests on Chinese Topic-Comment sentences. Chapter 7 will test for the neural underpinnings of the different empty elements and their dependency-links. However, it has to be noted that a crucial difference lies in the adoption of the Theory of traces compared to Copy theory of movement. These two different accounts of empty syntactic elements differ in the their understanding of what is the gap (silent element of the dependency) representing. Namely, assuming the early theory of traces, gaps are considered as simple place-holders for displaced elements where the tail of the is reactivated. While within Copy theory of movement what is "reactivated" at the gap is a whole articulated chunck of syntactic structure with all its derivational complexity (G. Bocci p.c.).

division of labor between gap strategy and anaphoric pronoun. Figure 2.38(B) shows a bilateral temporal activation with an anterior and a posterior cluster for Backward Anaphora compared to *wh*-gap, while the reverse contrast reveals a unique cluster in Precentral Gyrus for wh-questions gaps. In sum, the establishment of a dependency link between to overt syntactic elements in an antecedent-pronoun relation elicit antero-/posterio-temporal activation, while the establishment of a similar co-referential link in a syntactic configuration where a silent gap has to be posited elicit a stronger activation in an area adjacent to Broca complex, the Precentral Gyrus. These results are coherent

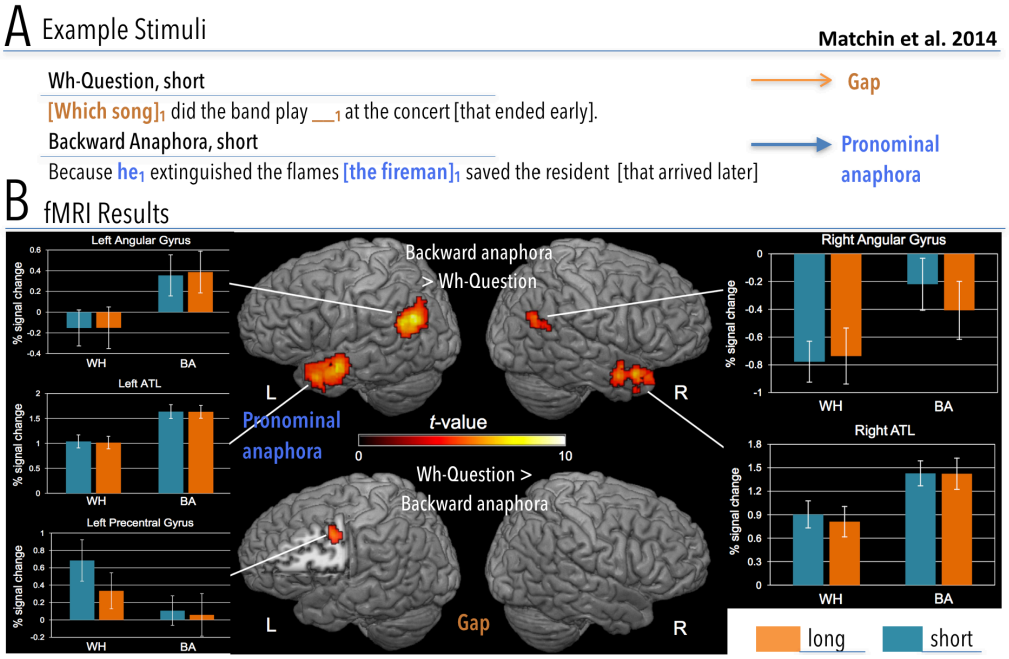


Figure 2.38 – (A) Experimental stimuli examples. (B) Group averaged brain maps results for the contrast between Backward anaphora > Wh-Question in (a), and Wh-Question > Backward anaphora in (b). Barplots report average percent signal change for each condition within selected clusters of activation. Adapted from Matchin et al. (2014).

with other findings: (1) Santi and Grodzinsky (2007a) reported activation in Broca's area (pars triangularis/BA45) for filler-gap dependencies generated by movement and not for canonical anaphoric dependencies; and (2) Santi and Grodzinsky (2010) showed the involvement of precentral Gyrus (BA6) and Broca BA44 and in presence of syntactic gaps (as the result of adaptation to both movement and embedding, cf. page 165). These fMRI findings constitute a further empirical grounding to one of the central experimental hypothesis, we will build on these findings to investigate the cerebral representation of the empty categories left by syntactic-movement transformations in French and Chinese.

We will now continue in this direction by reviewing a few examples giving further neuro-psychological grounding to our movement-base approach of syntactic complexity in the brain.

2.4.3.1 Resumptives: Another kind of Pronoun for a long-distance dependency

Resuming to the question of the neural implementation of the different types of syntactic elements among which a long-distance can be established in a sentence, another co-referential configuration should be added. In fact, dummy Resumptive pronouns can be found were a gap could also be found, like in "This PhD, I think he likes it." One of the reasons for focusing on this type of pronouns is that contrasting them with empty syntactic objects like gaps will deepen our understanding of the cerebral representation of the sentence.

In fact, some languages employ a different syntactic strategy, called resumption, which does not involve a movement-derived gap.Contrary to the movement-based strategy, a pronominal element (i.e. resumptive pronoun), is found at the original clause-internal thematic position where it receives the object role[172].

Concerning resumption and gaps, relativization pattern of Hebrew show in example (53) a combination of strategies: Gap or pronoun for the object in (a), (b), gap for subject (c), and pronouns (i.e. pronoun retention) for other grammatical roles like Dative in (d):

(53) Hebrew main clause relativization

a. accusative : Gap or Resumptive

Ra'iti et ha-barvaz$_1$ **she**-Dudu ciyer oto$_1$.
I.saw ACC the.duck$_1$ **that**.Dudu drew him$_1$

I saw the duck **that** Dudu drew [him].

b. Gap or Resumptive (Givon 1973)

Ha-ish$_1$ **she**- [Yoav raa Ø$_1$/oto$_1$ etmol].
the.man$_1$ REL Yoav saw ___Ø$_1$ /him$_1$(ACC) yesterday

'The man that Yoav saw [Ø/him] yesterday.' (Givon, 1973)

c. Nominative: Gap

Ha-ish$_1$ she- [Ø$_1$ /*hu$_1$ hika et ha-kelev].
the.man$_1$ REL ___Ø$_1$ /*he(NOM) hiy ACC the.dog

'The man that [*he] hit the dog.' (Givon, 1973)

d. Dative: Resumptive

Ha-ish$_1$ **she**-[Yoav natan *Ø$_1$/lo$_1$ et ha-sefer]. (Givon 1973)
the.man$_1$ REL Yoav gave *Ø/him(DAT) ACC the book

'The man that Yoav gave [to him] the book.' (Givon, 1973)

Analogously, Mandarin Chinese alternates resumption and gap strategy in Topic-Comment articulations, which makes these structures the prefect testing ground to investigate long-distance dependencies in absence of embedding and movement. We will namely leverage this syntactic configuration to investigate Movement-based strategy versus Base-generation in Chinese topicalization in chapter 3 form a more descriptive and linguistic point of view, and in chapter 7 as one of our critical fMRI experimental manipulation.

172. In fact, Resumptives are bound by a co-indexed operator that is base-generated in Specifier position of CP as shown by the Hebrew relative clauses in (53), which shows that it didn't move.

The complexity of filler-gap dependencies

Since the early moves of experimental psycho-linguistics, two processes have been posited in filler-gap dependencies: (1) active gap search (see Fodor, 1978) and (2) reactivation and integration of fillers (McElree and Bever, 1989). In active gap search one can say that the human sentence parser actively initiates a search for a potential gap site as soon as it encounters a filler, like wh-words or morphological marking for relative clause, etc. A consequence gap-filling effect has been observed in a behavioral slowing down when a potential host for the gap is filled by overt elements (Crain and Fodor, 1985; Stowe, 1986)[173]. We will resume to the psychological reality of syntactic gaps in the next session (§2.4.3.2), meanwhile we can say that these psycho-linguistic findings suggests an active online prediction and creation of a gap representation as soon as the human parser it encounters a filler.

The long-distance dependencies formed by movement operations are not only of broad interest for linguistic theories and psycho-linguistics, but and have been extensively investigated in their temporal processing in EEG/ERP studies, which were able to track the online syntactic chain between antecedent and gap thank to their a very high temporal resolution (see for example Hestvik et al., 2007).

Copy-theory of movement and syntactic movement became prominent even outside linguistics and inspired in the domain of neuro-psychology, the idea that Broca's area supports the interpretation of displacement. The so-called Trace-Deletion Hypothesis (Grodzinsky 1986, 1995) states that syntactic traces are deleted in agrammatism, and a default sentence understanding strategy is engaged: "assign the role of Agent to the first NP". Broca's aphasics know how to assign semantic roles to NPs when the NPs are in the original place (local to the verb), while when they are moved their interpretation depends on traces and patients performance is at chance. Yet, in structures where there is a second NP that receives the Agent role, agrammatic sentence representation present two NPs conflicting for Agents active role of the verb[174]. Grodzinsky argues that this is the mechanism that induces guessing and chance/random performance on comprehension tests that has often been observed for passives and object cleft and object relatives. However while certain studies (e.g. Beretta and Munn, 1998)[175], were able to confirm the default Agent-first strategy of aphasics implied by the Trace-Deletion Hypothesis, while a number of other studies (e.g. Caramazza et al. 2004 in Italian, among others) fail to prove the double agents competition in agrammatic sentence representation and its predictions.

As picture-matching data do not support the agent-first strategy, another way to test for the representation of the gap would be to focus on languages that alternate the presence of a resumptive pronoun at the gap site. Hebrew, as we just saw in the above example (53), represents the perfect linguistic testing ground to asses in the same syntactic structure the agrammatic comprehension deficit, in presence or absence of gap. Interestingly, the results from a sentence-picture matching paradigm (see Figure 2.39

173. A recent article by Omaki et al. (2015) reports what the authors call an *hyper-active gap-filling* effect when, for instance, the first-encountered verb is an intransitive verb that cannot host a gap.
174. Note that the prediction for such an hypothesis would imply normal patients' comprehension performance in active reversible sentences, but at chance level for passive reversible sentences. A prediction that strongly contrast with initial results by Caramazza and Zurif (1976) who showed Broca's aphasics performed poorly in a sentence comprehension task with semantically reversible sentences.
175. Using a sentence-picture matching task in which one of the pictures matches the meaning arrived by the default strategy.

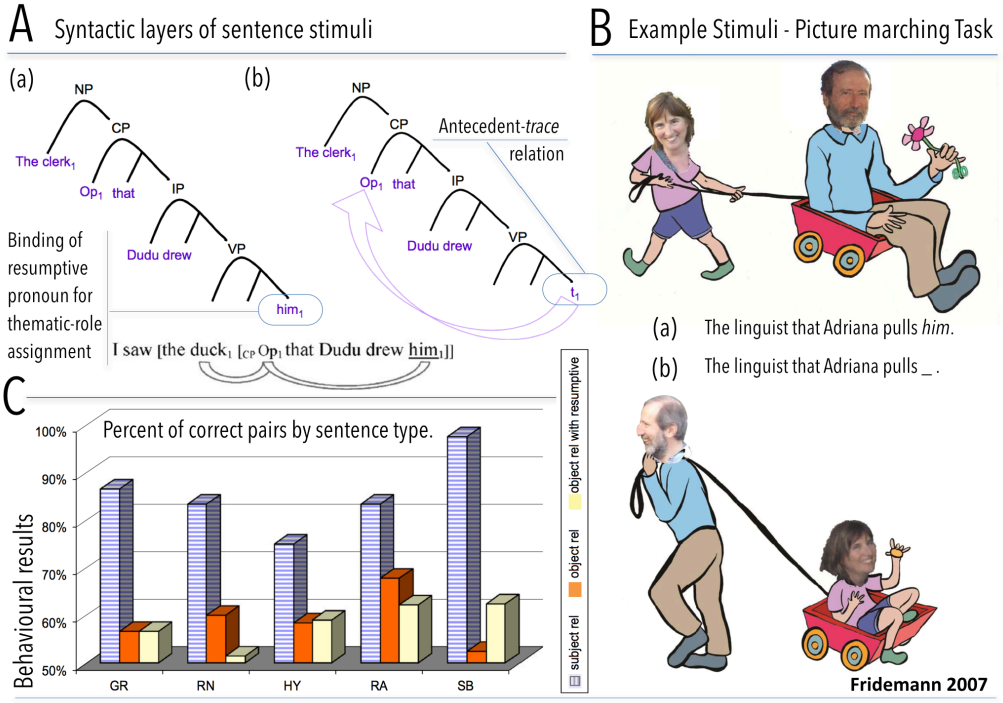


Figure 2.39 – (A) Syntactic-tree of object relative sentences with an without *trace* (B) Binary picture pairs, actually, adapted from Friedemann and Givon (2016), in *Intervention and locality in Agrammatica aphasia.* (C) Experimental results of comprehension of subject and object relative sentences with an without *trace*: Percent of correct pairs by sentence type. Figure Adapted from Friedemann (2006).

(B) (C)) with Hebrew agrammatic patients showed that the presence of a resumptive at gap site has no fundamental impact on processing long-distance dependence between the gap and the filler in object relatives (Friedmann, 2006). Thus, what seems to be impaired is not the gap-antecedent chain, but the more global binding process to assign to a pronoun (having a antecedent) the thematic-role of the verb.

However, the precise role of an overt marking of the long-distance dependency link with its antecedent needs still to be totally elucidate, in that another syntactically impaired population of hearing impaired children seem actually to benefit of the presence of resumptive pronouns in both production and comprehension or A-bar derived sentences. These population shows indeed a Syntactic Selective Impairment (S-SLI) that was consistently reported to selectively target the comprehension and production of movement-derived constructions comprising object relatives, topicalized structures, passive and object wh-Questions (de Villiers et al. 1994; Szterman and Friedmann, 2006)[176].

For instance, Friedmann et al. (2008) report behavioral results indicating that hearing impaired children that show a deficit in A-bar movement derived syntactic structures, like Relative clauses and Topicalization, typically produce grammatical object relative

176. This particular population of S-SLI children weren't exposed to language during the first 8 months of life -a possibly critical period for syntax- because of a non-detected hearing deficit (Szterman and Friedmann, 2006)

clause with a resumptive pronoun, unlike their age-matched controls as shown in Figure 2.40 (D) and (C) who tend to produce them with a gap.

Yet, this second syntactically impaired population (with syntactic Selective impairment, S-SLI) has been shown to have a significant improvement in comprehending Hebrew object relative when a resumptive is added at the gap position. Contrary to the comprehension impairment pattern revealed by the previous picture-sentence matching experiment (Friedmann, 2006) in agrammatic patients, the results for the same task in this hearing-impaired population show that gapped A-bar moved structures (i.e. topicalization and object relatives) are better understood when the resumption strategy is adopter as clearly showed in Figure 2.40 (B).

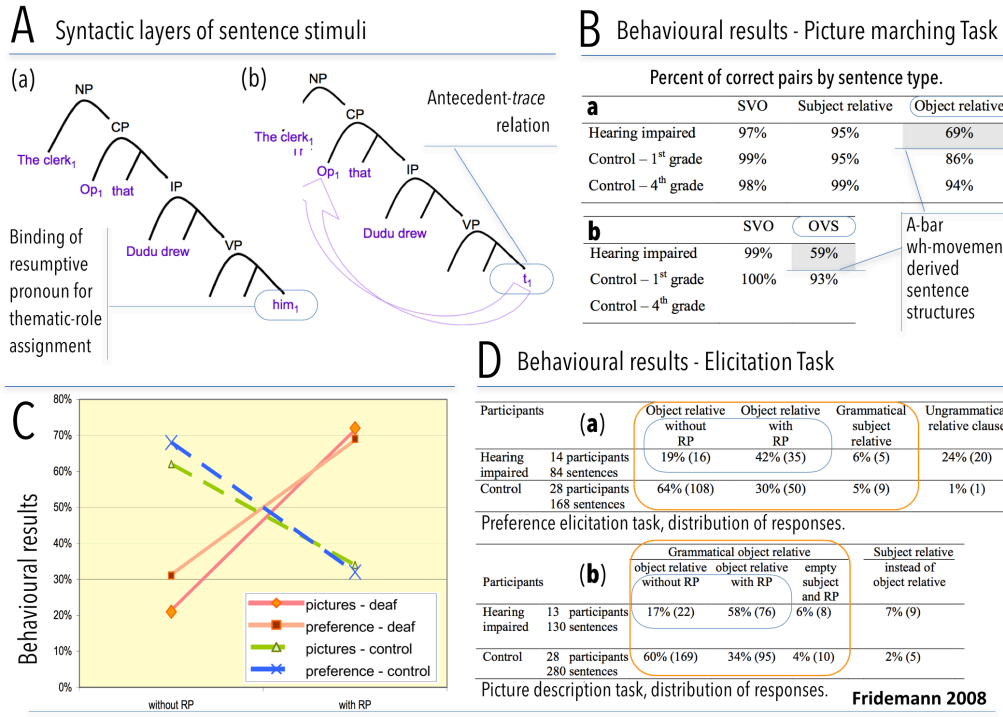


Figure 2.40 – (A) Syntactic tree of object relative sentences with (b) an without *gap* (a). (B) Percent of correct pairs by sentence type. (C) Table showing the performance in elicitation and comprehension tasks of hearing-impaired syntactic SLI children against control groups. (D) (a) Distribution of responses for target object relatives in the preference elicitation task; (b) Distribution of responses in the picture description task, target object relatives. Adapted from Friedmann et al. (2008).

Moreover, the impaired participants also produce ungrammatical sentences were they doubled the relative head in both subject- and object-relatives. overuse of resumptive pronouns instead of gaps both in licit and illicit contexts of relative clauses

While the first population showed an impairment of the binding process to co-index both the gap and the resumptive pronoun with the antecedent. This second population importantly shows a selective impairments for movement-related syntactic gaps compared to the presence of a resumptive. These results were interpreted by the authors as supporting the claim that resumptive pronouns are a last resort when movement is

impaired.

Such a distinct impairment pattern seems to indicate that the representation of the dependency-link between two elements in the sentence, and the capacity of representing the empty syntactic position itself are dissociable. Hence, we could see in this dissociation findings a initial evidence indicating that the human syntactic system has different representations for co-indexation function, that we could define as a mechanism of pointing to the right referent for co-reference assignment, and for the creation or localization of syntactic gaps in the sentence.

These findings are instrumental in introducing one of our main experimental questions. We will investigate in an fMRI experiment with normal adults (cf. chapter 7) a similar syntactic configuration where the very same sentence construction with a [gap] will be opposed to one with a [resumptive]: Mandarin Topic-Comment constructions presenting a gap or a trace will be selected to uncover the neural substrate of this resumption strategy, presenting an overt element for establishing the long-distance dependency, compared to gap (silent element) one. Note that as it is the case in the Hebrew stimuli of Friedmann et al. (2008) our two conditions will also differ in their movement analysis.

Anticipating next sub-section, from Friedmann's (2006) results emerges another possible account for this deficit pattern in agrammatic population: both have to establish a long-distance dependency over another argument of the verb. Both sentences (with and without gap), do have a constituent in the high nodes of the syntactic tree (specifically, in Spec-CP) as we can see in Figure 2.39. It is therefore also possible that the deficit is related to an inability to construct the syntactic tree up to its treetop, the CP node, that would results in an inability to connect the antecedent and the gap. It is along this lines that our experimental design on Mandarin Topic-Comment articulation will be built, comparing movement-derived and non-moved derived sentences both presenting a NP in the high nodes of the syntactic tree with sentences where a constituent has been moved to sentence internal position.

2.4.3.2 Experimental Evidence for the gaps/traces in the sentence

In the early 21st century, work on the topic of syntactic movement to account for sentence complexity had broader impacts on psycho-linguistics and has helped shedding light on mechanisms of sentence processing.

Psycho-linguistic studies have been focusing on this transformational dimension of sentence complexity, positing that successful understanding of displaced constituents would require representing the *gap*, and, being able to reactivate the moved constituent in its original position. This is why, displaced constituents have been investigated with a real-time behavioral paradigm called cross-modal priming (Swinney et al., 1979), that allows one to observe priming effect of semantically related words to the moved constituent. As the prediction from the copy theory of movement is that displacement involves pronunciation at one site of the sentence an interpretation at another, the interpretation site in gap position should be able to prime the semantic content of the displaced word and therefore show a facilitation effect of semantically related words in a parallel lexical decision task, as schematized in Figure 2.41A.

To summarize, this methodology manipulates two factors: (i) Whether or not the visually presented target is semantically related to the displaced element in the auditory stimulus, and (ii) at what point during the auditory stimulus the visual target is presented.

Among the first structure investigated were wh-questions and relative clauses, two occurrences of wh-movement (Frazier et al., 1983; Nicol and Swinney, 1989).

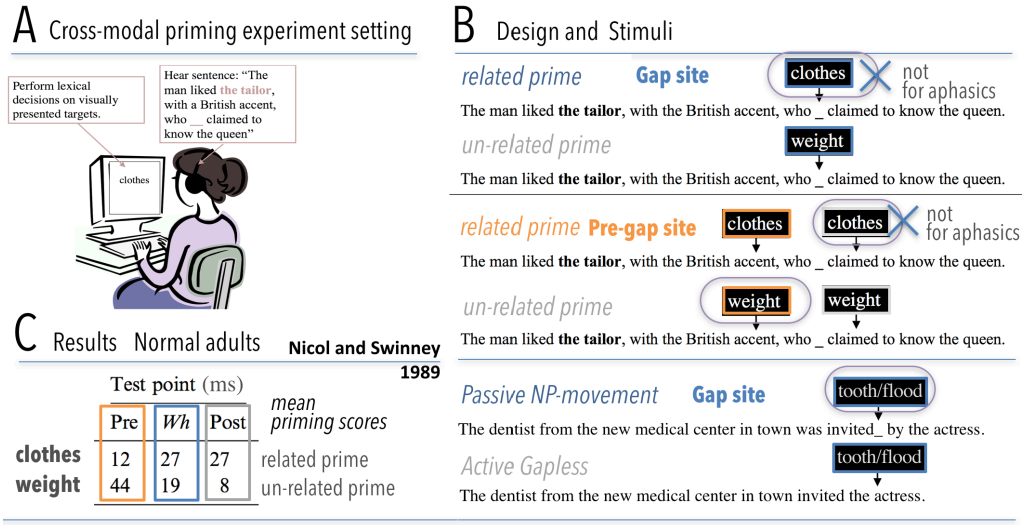


Figure 2.41 – (A) Experimental setting for Cross-modal Priming Paradigm (B) Sample stimuli from the experimental design, circled items are the ones that show a priming effect. (C) Behavioral results, mean priming scores in milliseconds, from Nicol and Swinney (1989).

From the point of view of incremental sentence processing, wh-movement posses two main relevant characteristics. Firstly, it departs from canonical structure and creates types of phrase sequences that do not correspond to any simple un-transformed sentence. And, secondly, when wh-movement occurs in a sentence, there is often an overt marker in the sentence like a wh-word (e.g. that, wh-pronoun, etc.). In the examples reported in Figure 2.41 the marker is the relative pronoun'*who*'. One could, therefore expect that the comprehension system uses such overt markers to immediately engage in establishing a gap for the moved marked element.

Results from Nicol and Swinney (1989) seem in fact to go in this direction. Subjects heard sentences similar to the ones in Figure 2.41B and saw a test probe at a pre-gap point, at the wh-gap, or at a post-gap point on which they were asked to make a lexical decision on a word that was either related or un-related to the gap antecedent: 'clothes' (related) or 'weight' (unrelated). The authors, found a significant facilitation in lexical decision for 'cloths' at wh-gap position and at the post-gap point, but not at the pre-gap-site (see Table C in Figure 2.41 circled probes are the primed ones). This result pattern was taken to indicate that a displaced constituent is re-activated at its gap, a mechanism that was not observed in Broca's aphasic patients, as we will now see.

Classical Cross-Modal-Lexical-Priming studies that were conducted with adults aphasic. (Zurif et al., 1993) show that priming effect is abnormal both in subject and object relatives. The facilitation effect preserved in normal adult population is either absent or shifted in time. This result has been replicated a number of time, and Love et al. (2008) showed a general slowing of lexical activation and a concomitant delay in the formation of syntactic dependencies involving moved constituents and empty elements. Interestingly, at a slower speech rate, Broca's patients appear to form syntactic depen-

dencies like normal adults. The replication of this difference slowing the presentation rate in a sentence picture matching task made the authors conclude that Broca region is to be attributed the formation of syntactically-governed dependency relations more than the role of supporting syntactic dependencies knowledge. According to their discussion the results indicate that Broca supports the real-time implementation of syntactic movement's specific representations.

NP-movement gaps

In the last fifteen years Cross-Modal Lexical Priming studies have found stable evidence for the reactivation of the antecedent NP at the gap position of wh-chains and mainly within relative clauses (Love and Swinney, 1996; Nicol and Swinney, 1989; Swinney et al., 1988; Swinney et al., 1989; Zurif et al., 1993, 1995). However, gaps left by other types of movements have also been investigated, for instance the gaps left by NP-movement in passive sentences and unaccusative verbs have luckily been a well-studied object too.

The comparison between the psycho-linguistic behavioral measures linked to these two gap types, leads us to briefly compare the main characteristics of the movement types that generate them in the following Table 2.5. Bridging together the formal definitions of NP-movement and of wh-movement that were introduced in the previous section, we see that a real typology for these two types of empty categories can be drawn.

Table 2.5 – Comparing NP- and wh-movement gaps.

	NP-gap	**wh-gap**
Moved category	Noun Phrase	XP (NP,PP, etc.)
Landing site	Argumental position NP-position	Non-Argumental position (A-bar) Specifier of Complementizer Phrase [Spec, CP] or Adjoined position
Antecedent Properties		
Case assignment	Yes	No
Chain type	Argumental (A-)Chain	Non-Argumental (A-bar)Chain
trace Properties		
Features	[+Anaphor] [−Pronominal]	[−Anaphor] [−Pronominal]
Binding principle	Principle A	Principle C
Theta-role	Yes	Yes
Case	No	Yes

Given these characteristics, we can add that from the point of view incremental processing of the sentence, the main difference between the traces left by these two movement types resides in the presence of overt markers for wh-movement, while NP-movement has none. For instance, the class of intransitive verbs called Unaccusatives, present only one theme-marked argument, which therefore base-generated in the verb object position. In these verb class the theme-marked argument moves to the subject position to satisfy the requirement that English and other languages have that all sentences must have subjects. Namely, NP-movement preserves canonical word-order. This characteristic aspect of NP-movement is what made psycho-linguists hypothesize and test that the comprehension system engaged into incremental processing of NP-moved sentences would establish

NP-*traces* in their original location only after having assigned meaning (McElree and Bever, 1989, probe recognition task), rather than in a predictive manner as reported for wh-movement (cf. pre-*trace* priming effect in Nicol and Swinney, 1898).

Using the cross-modal priming technique, Osterhout and Swinney (1993) tested reactivation in passives sentences at the *trace* spot, and found a delayed reactivation of the grammatical subject, with effects reaching significance only 1000 ms after the verb. This result further support the idea that NP-traces do reactivate (and prime) their antecedent, but later in time.

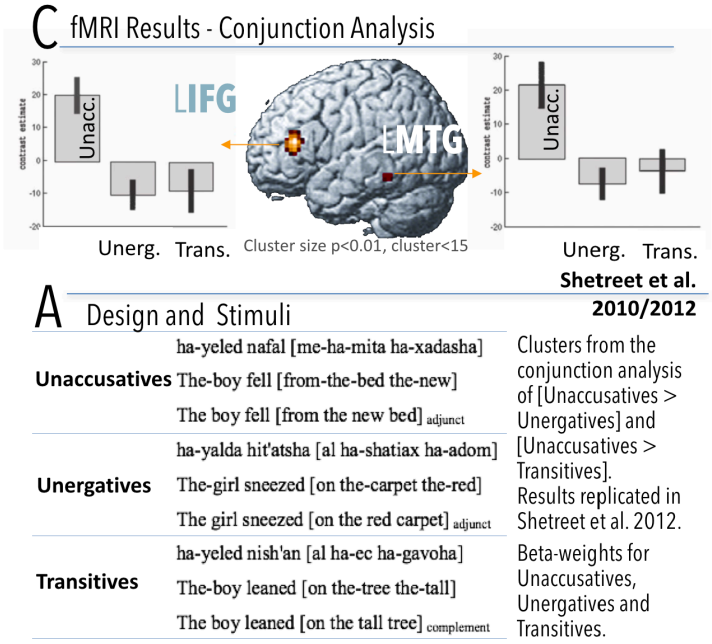


Figure 2.42 – (A) Design and sentence stimuli examples. (C) Group average brain map of the conjunction analysis [Unaccusative > Unergantive and Unaccusative > Transitives]. Barplots report the Beta-weights in the two activation clusters, respectively Left Inferior Frontal Gyrus (BA45/46) and Middle Temporal Gyrus (MTG).

Similarly, delayed reactivation effects have been reported for unaccusative verb's structures, like "The leaf$_i$ fell t_i" that are also known to involve NP-movement in Hebrew (Friedmann et al. 2008) and Spanish (Bever and Sanz, 1997). Friedmann and colleges compared sentences with unergative verbs to sentences with unaccusatives, both alternating and non-alternating unaccusatives[177], and confirmed the linguisitc hypothesis that unaccusative subjects are base-generated in object position, and move to the subject position (cf. Unaccusativity Hypothesis, Bruzio, 1986): the subjects of unaccusatives reactivated after the verb, while subjects of unergatives did not, and alternating unaccusatives showed a mixed pattern of reactivation.

Interestingly, the same comparison was brought to the scanner in a study by Shetreet and collegues (2010, later replicated in Shetreet et al. 2012), with the difference that transitive verbs included in the experimental design were sub-categorized for prepositional phrase and not altering accusatives[178]. The authors reported a conjunction analysis of the two contrasts [Unaccusative > Unergantive and Unaccusative > Transitives] showing activation of Broca complex (BA45/46) and in the posterior part of the Middle Temporal Gyrus (MTG) for unaccusatives, as shown in Figure 2.42. All in all, these findings confirm the hypothesis of the involvement of Broca complex in the syntactic complexity linked to movement-derived linguistic phenomena, and particularly in NP-traces fond in Unaccusatives.

177. Alternating unaccusatives are intransitive verbs whose subject may also appear as the direct object of a morphologically identical transitive verb, like the verb '*to break*'. On the contrary non–alternating unaccusatives are like the verb '*vanish*' that does not have any other form in which it can take an object.
178. It has to be noted, however that Hebrew posses verb classes that are morphologically marked, and certain morphological markers can also derive transitive verbs into unaccusatives. Among the 14 verbs present in the stimuli 10 were carrying this mark (*hit-*). Notably, the authors interpret the activation of MTG in this direction, saying that it may be responsible for the lexical operation that derives unaccusative verbs.

Traces were also investigated using other psycho-linguistic paradigms and methodologies. A visual world paradigm on wh-movement structures, including object-extracted wh-questions and object cleft structures, showed automatic eye-movements to the filler at the gap site in both healthy adults, and listeners with aphasia (Dickey et al., 2007, 2009; Koring et al., 2012).

Additional evidence also comes from event-related potential (ERP) studies. Fiebach et al. (2002) report that wh-embedded questions compared to *whether*-questions that do not entail gap-filling, elicit a positivity at the gap position around 400–700 ms. We reserve for next chapter a detailed account of the different psycho-linguistic and neuro-imaging results (Chapter 3, §3.4.2.2) that can be found for gaps in Chinese topicalization (p.370).

In conclusion, given this rich range of experimental results, we hope the reader will look forward to addressing the issue of the neural underpinnings of movement transformations in chapter 6, where we will investigate the different representations and processes associated to the two types of gaps reviewed here (NP-and wh-movement ones) at the level of brain activity. While chapter 7 will experimentally address the resupmtion versus gap issue by comparing the fMRI activation patterns to gapped Topicalized sentence structures derived by movement to Topic-comment sentences that include a resumptive pronoun and do not involve movement.

As a more general remark, we can say that given this rich range of experimental data, the linguistic constructs described above do have important implications for human sentence processing (as well as production). Furthermore, these results demonstrate that taking into account as a complexity parameter the abstract syntactic object a '*gap*' is, might reveal important aspects of the organization of syntactic representations in the brain. This is what grounded our choice to try and correlate with brain activity the number of syntactic positions (including the number of gaps) present in each sentence of our French fMRI study (chapter 6). The main rational of this resides in the hypothesis that the set of brain areas underpinning the kind of representation linked to *gaps*, we just reviewed, should show an increased activity proportional to the number of empty silent abstract elements present in the syntactic-tree of our complex set of movement-derived we included in the design (Fig. 6.2, p.6.2).

2.4.4 Discourse-Related Features: the Complexity of Sentence Domains

Generally speaking linguistic research, and consequently neuro-linguistics, have maximally developed considerations about the organization of units at sentence-level, focusing on constituency organization, word-order generalizations, rules that bind constituents together, and even agrammaticality-related processing. These sentence-internal linguistic phenomena, however, do not account for another structural link that the sentence-unit can establish with the above level of linguistic analysis: the discourse.

Understanding a sentence is more than a simple matter of decoding words and the hierarchical relations structuring it, it also implies the processing of sentential contextual information, which typically requires inferential processes, like bridging successive utterances, using some background knowledge to disambiguate among different options. All these processes are linked to the higher linguistic level of discourse-context where pragmatic information and interpretation are required as we saw for Utterance particles in section 2.1.2 (p.90). These varied linguistic and psycho-linguistic processes are the expression that some discourse-level semantic interpretation is possible already at the

level of the sentence-unit, and the way the human brain achieves meaning assignment to strings of words in a sentence has to be investigated at different levels.

This approach to sentence is relatively new in syntax and implies looking at the sentence as a way to package information, giving a syntactic interpretation to the possible relation between sentence-level propositional content and the discourse-level. This will lead us to identify in the functional skeleton of the sentence syntactic structure a domain dedicated to the interface with discourse, and in our neuro-linguistic investigation we will consider this interfacial sentence domain as an additional complexity measure of sentence complexity to explore Topic-Comment articulations in chapters 5 and 7, and French question formation (yes/no questions and wh-questions) in chapter 6.

But before entering in this section, and before all the emphasis that it gives to the linguistic phenomena linked to sentence/discourse interface, we start by presenting a psycho-linguistic and neuro-imaging study offering evidence for these kind of phenomena, in order to first give some empirical grounding to the more theoretical claims this section carries. An fMRI and behavioral study in Danish, an SVO language, investigated the effect of appropriate linguistic context on the processing of Danish main clauses with either an initial subject or a topicalized object.

Importantly, Danish is a language which marks pronouns for case but not full NPs, as shown in the example stimuli in Figure 2.43A, so that the participants cannot know from morphological case marking the grammatical role of the sentence initial NP before getting at the non-finite verb position 'invitere' *invite* for subject-initial sentence at the second NP position for Topicalized sentences. This syntactic configuration offers a good testing ground to investigate word-order cues t sentence processing in absence of morphological markers. Moreover, the authors suggest that unless the context gives rise to other expectations, the listeners should expect that sentences have subject-before-object object, in fact they cite a study by Thomsen (2008) showing that in written Danish subject-initial sentences are much more frequent than Topicalized objects, with a ratio of approximately 1:14 in transitive main clauses, and of 3:4 in spoken Danish transitive main clauses.

Participant were invited to read a short context of three sentences and after each target sentence, a comprehension question of the type "Ville Peter invitere Anne?" '*Would Peter invite Anne?*' was presented and responses were collected by button press. Behavioral results in Figure 2.43B show that word-order did not have a significant effect on Response Time, but context had, revealing faster responses to sentences with a supportive context. As for accuracy Kristensen et al. (2014a) demonstrated that topicalized object-initial sentences are more context-sensitive than subject-initials and that context had a important facilitating effect on the comprehension of object-initial clauses given the accuracy patterns illustrated in 2.43B.

The fMRI results in Figure 2.43C show that the nature of contextual information provokes a reduced activity in BA47 for both subject initial and topicalized sentences (occurring after a supportive linguistic information). This effect restricted to one of the apriori Region of Interest, was not isolated, namely the lack of coherent and supportive context elicited an increased BOLD signal in a wide network encompassing several frontal, temporal and parietal regions.

Hence, we can conclude that adding a discourse pragmatic dimension to the investigation of word-order changes in the sentence reveals that sentence processing is highly responsive to discourse information both at the level of behavioral responses and at the cerebral level give the wide activation of the quasi totality of the sentence network.

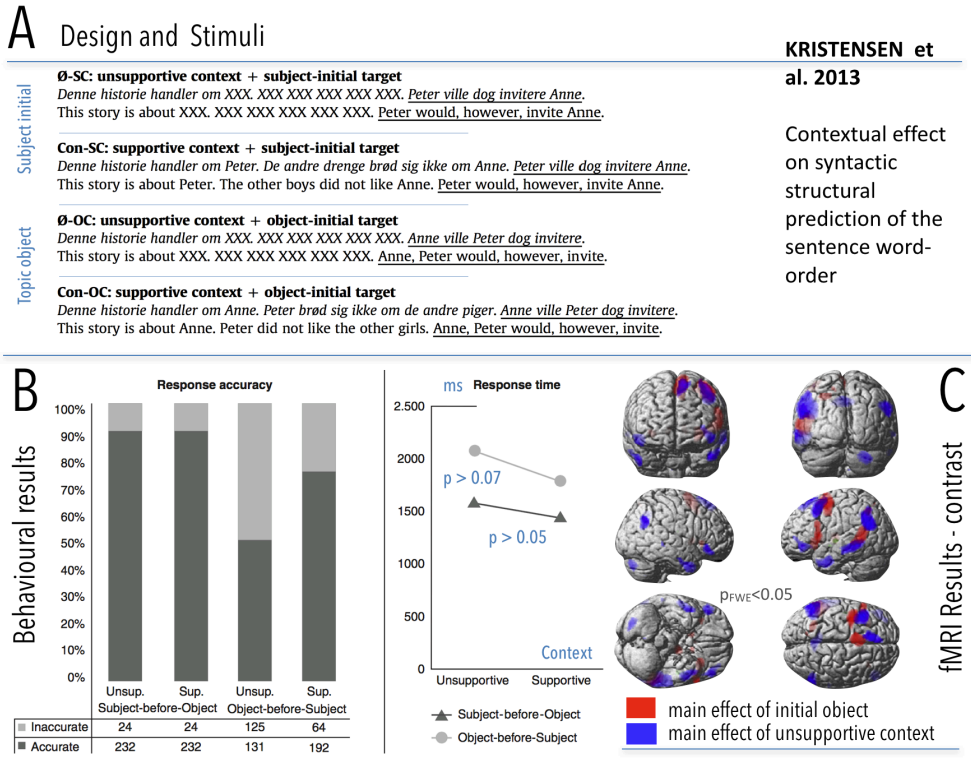


Figure 2.43 – (A) Experimental design and sentence stimuli examples. (B) Behavioral results for the comprehension task, showing Accuracy (left) and Response Time (right) by sentence types and context. (C) Group-average brain map for the main effect to object-initial sentence in red, and for the main effect of unsupportive context in blue.

Thus, having experimentally established through this Danish fMRI result that contextual appropriateness has actually a broad impact on sentence processing, and more specifically syntactic processing, we can continue in this section by presenting three main points:

1. (1) the details of the discourse-related properties represented in the functional architecture of the sentence-discourse interface focusing on those functional projections that are relevant to the research in this thesis (§2.4.4.1).
2. (2) the linguistic analyses of the syntactic processes happening when a Noun-Phrase is moved to the Left-Periphery of the sentence, will reveal crucial steps that will be further correlated with brain activity in the experimental part (§2.4.4.2 and 2.4.4.3).
3. (3) And to conclude, the question of the full/partial representation of the rich functional skeleton of discourse-related properties and its ordering constraints will be linguistically addressed to engage in a reflection of the possible cerebral representation of the ordering constraints put forward by the Cartographic approach (§2.4.4.4 and 2.4.4.5).

2.4.4.1 Overview of the Left-Periphery

Until now, we have assumed a very simplified structural representation of the sentence syntactic skeleton, in which canonical clauses are expressed in terms of three domains [CP + IP + VP] structures. However, we will now introduce some elements of complexity in this functional structure that are hiding behind these abbreviations to show how a richer clausal configurations can account for the above cited linguistic phenomena. This will imply an essential step of "*syntacticization of scope-discourse semantics*" (Cinque and Rizzi, 2010), by reducing some discourse articulations to syntactic configurations in which every specific interpretative property posses a particular functional head in the sentence tree structure localized in a given sentence domain (CP in our case). The role of this functional head will be to trigger the appropriate discourse semantics interpretation[179].

Using the terms of Rizzi (1997), and much related work, this means that core syntax must involve information-sensitive functional projections (e.g., TopP, FocP, InterP)[180].

Hence, following the working hypothesis that "*one morpho-syntactic property = one feature = one head*" the cartographic study of sentence structural articulation establishes a principled typology of syntactic positions dedicated scope-discourse features as illustrated in Figure 2.44. Note that these functional Heads have the characteristic of being interpretable by providing information to the syntactic/pragmatic system in a simple and transparent manner, while empty functional Heads we already encountered (e.g. Infl.) receive a default interpretation. The initial periphery of the sentence (the CP-layer, henceforth, Left-Periphery of the sentence) is thus assumed to be populated by a sequentially ordered series of functional heads such as Question markers (Q), relative markers (R), markers of topicality (Top) and Focus (Foc),etc. But also different subtypes of Topics like Hanging Topic, List reading Topic, Contrastive Topic, to which we will come back in chapter 3 presenting the relative order of Topic and Focus projections in the Mandarin Chinese Left-Periphery (§3.4.5).

A clear advantage of this approach for our investigation about the cerebral representation of the syntactic skeleton of the sentence-unit, is that it offers a unified syntactic account of different linguistic phenomena whose semantic interpretations sometimes take different morpho-syntactic markings[181].It should be noted that this research project on the Cartography of sentence domains, briefly introduced in chapter1 (§1.2.4), gathers an increasing amount of cross-linguistic evidence indicating that the hypothesis of universal functional design for the clause and its major phrases that can hold across a growing sub-set of languages like Romance languages (Benincà and Poletto, 2004); Germanic languages (Haegeman, 2006), West-African Languages (Aboh, 2004 and 2007), Creole Languages (Durrelman, 2007), and East-Asian Languages (Tsai 2007 and Badan 2008).

179. Note that from a structural point of view this puts forth that scope-discourse properties are structurally determined by local head-dependent relations.

180. Note that Interrogative, Topic, and Focus, as typical cross-linguistic Information Structure cornerstones, often involve non-ambiguous morphological expressions correlating with their distinct semantic interpretative properties, and are therefore less controversial compared to other abstractly postulated heads. It should also be noted here that in traditional debates on the interaction between Information Structure and syntax, one can observe a bias in the literature in considering Topic and Focus as purely pragmatic notions, while Interrogatives for instance are considered as having proper formal features in that they more easily present overt morpho-syntactic marking (functional heads).

181. To this should be added that it also accounts for the phonological processes of pitch contour assignment that are yielded by specific discourse articulations (like Topic-Comment or Focus-Presupposition, cf. chapter 4), because as we just saw these *heads* are interpreted (like their scope-discourse properties) at the interface with the sound system (Bocci, 2012).

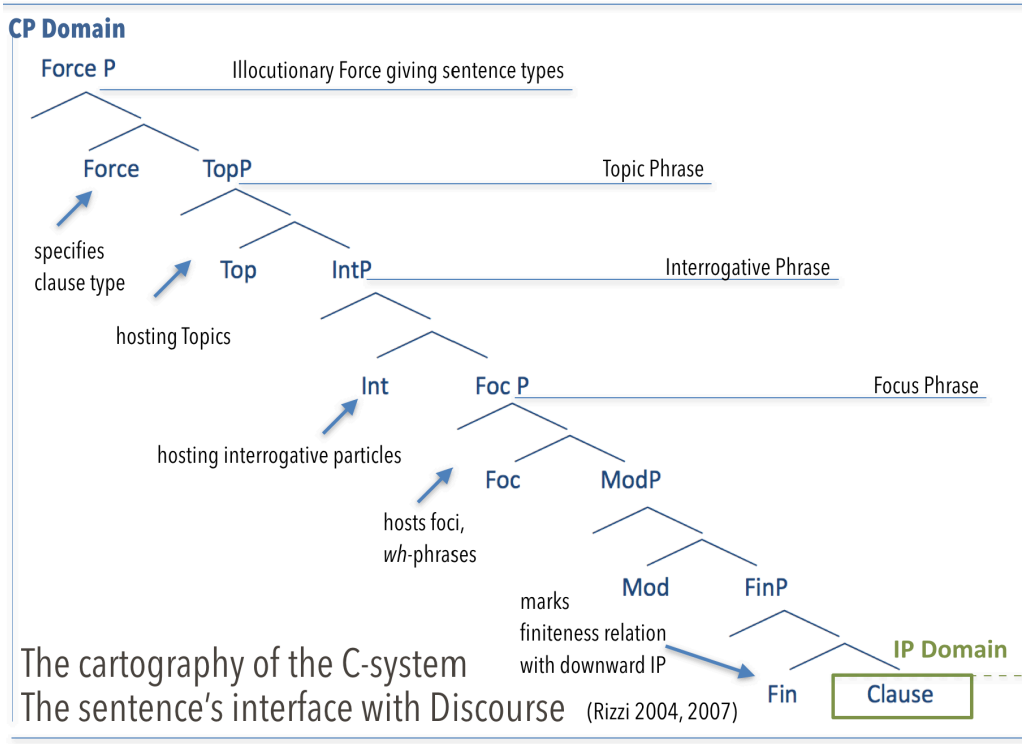


Figure 2.44 – The Cartography of the Complementaizer System, the sentence's interface with discourse. Adapter from Rizzi (2004 and 2007).

The sharp contrast between the richness of the sentence configurations presented in Figure 2.44 for the CP-domain or for the clausal hierarchy in IP by Cinque in Figure 1.4 (§1.2.4, p. 34), shows how a minimal generative device like Merge can generate a clausal structural backbone organized according to several Structural layers (commonly called Split Projections), where each head correspond to a morpho-syntactic property with its interpretative value. This is a crucial aspect of this approach that will allow us to account for a rich range of complex sentential phenomena keeping in the realm of tree-like representation of the clausal structures with their linear and hierarchical mapping, as we already had he occasion to underline in the previous Section 2.3.3 (see Figure 2.15, 144).

Force and FinP

Having a closer look at the CP clausal hierarchy presented in Figure 2.44), we find at the two extremes Force Phrase (ForceP) and Finitness Phrase (FinP), the two parts of the C-domain respectively facing the out side of the sentence and facing inward (Branigan, 1996 and Rizzi, 1997). Force Phrase encodes the relation between propositional content (expressed by IP and VP) and super-ordinate discourse, indicating Clausal Types

through the specification of the Illocutory Force[182] (interrogative, declarative, imperative, optative etc.). While the Finitness Phrase, encodes relations with IP layer that are yielded by certain particular requirements of the verbal system, like in cases where the sub-categorization frame of a verb has an argument introduced by 'for' requiring an infinitival form.Such phenomena illustrate the need of an "agreement" between C domain and the I domain and this Functional Projection is indeed called like this because it has a head with [+/-finite] feature licensing Tense and Mood in IP.

2.4.4.2 Questions in the CP

Sandwiched in between these extreme positions, we find an extra layer called InterP which is commonly assumed to be realized by Question particles as the one we met in Section 2.1.2, and the projections of Topic and Focus, that we will be the concern of much of our discussions in the following chapter 3.

As for sentence interrogative interpretation, an asymmetry between yes/no questions and *wh*-questions has been cross-linguistically attested (Aboh and Pfau, 2011)[183] and cartographic analysis has established that these two forms of interrogatives relate to different portions of the clause skeleton/structure although sharing the same discourse-information property of encoding interrogative force. More precisely, we will briefly present what is their difference and therefore interrogative particles have been located between the complementizer corresponding to 'that' and topic/focus articulation. Namely, this fact that yes/no operators (or particles) and wh-operators (moving to the specifier of the focus projection) activate different articulations within the C-system, respectively InterP and FocP, will be investigated in chapter 6 by comparing brain activation to French yes/no questions in French that are minimally marked by a question mark at the end of the regular declarative word-order, like in example 54a (a and b) and object wh-Question in 54d (d). Hence, let us briefly detail this attested difference between yes/no questions and wh-questions.

As we can see in the following examples French show a quite wide variety of interrogative syntactic configurations, ranging from simple intonational marking, wh-in-situ to complex inversion[184].

(54) French interrogation

a. yes/no question

Tu vois ça ?
you see this Q-INTONATION [Q :]

'Do you see this?'

b. yes/no question

Tu mets ça là ?
you put this here Q-INTONATION [Q :]

'Do you see this?'

c. wh-question

182. Illocutionary force reflect the way in which a speaker uses a proposition in a discourse context, it gives the status of an utterance as a question, promise, threat, etc. (not be confused with illocutionary speech act, expressing the intention of the speaker).
183. Note that these authors analyze this distinction also considering examples from Sign Language utterances (in use in the Netherlands).
184. See Rizzi and Roberts (1989).

Tu as vu qui ce soir?
you have see who this evening Q-INTONATION

'Who did you see?'

d. wh-question

Qui tu as vu t_{qui}?
who you have seen ___

'Who did you see?'

e. 胡匪买了什么呢 Chinese wh-questions + interogative wh-particle

Húfěi mǎi-le shénme (ne))
Hufei buy-PERF. what PART.

'What did Hufei buy?' Cheng 1991, see also Tsai 1994.

We adopt here Cheng and Rooryck's (2000). In (c), French allows the wh-word to remain in-situ, that is to say in its base-generated position, in (d) the wh-pronoun moves to the beginning of the sentence. While (a) and (b) show how yes/no questions can be formulated just by adding an interrogative intonation, in (c) we can see how intonation, plays a central role in the licensing of wh-in-situ, just like wh-particles or real wh-words.

These characteristics of French interrogatives showing that wh-movement can be optional[185] made it a suitable language for the examination of the effect of wh-movement in Broca patients' comprehension. van der Meulen's PhD Thesis investigated the comprehension patterns of French aphasic[186], showing that the comprehension of object questions with wh-in-situ in example 55a (a) is significantly better than that of subject (c and d) and object wh-questions (b) with movement[187]:

(55) a. Object in-situ

[$_{CP}$ [Le garçon arrose qui]]?
The boy splashes who

'Who does the boy splash?'

b. Object moved

[$_{CP}$ [Qui] est-ce que [le garçon arrose t_{qui}]]?
Who Q the boy splashes

'Who does the boy splash?'

c. Short subject

[$_{CP}$ [Qui] [t_{qui} arrose le garçon]]?
Who splashes the boy

'Who splashes the boy?'

d. Long subject

185. French is not the only language displaying this kind of syntactic phenomena where wh-movement is shown to be optional. Ancash Quechua and Malay (Cole and Hermon, 1994/1998), or European Portuguese (Cheng and Rooryck, 2000) can be cited as other examples.

186. As a side note, Ineke van der Meulen argues that the comprehension deficit in Broca's aphasia is related to the type of movement through which sentences are derived. A global assessment of his proposal is beyond the scope of this section, but his analysis represent an additional element that supported to focus par of our experimental efforts on the issue of the cerebral representation of movement types. See van der Meulen et al. (2002/2005).

187. '*est-ce que*', that is here glossed by van der Meulen as 'Q', is argued to be an interrogative marker occupying C. it has the peculiarity to be obligatory in object questions and optional in subject questions.

[$_{CP}$[Qui] (est-ce qui) [t_{qui} arrose le garçon]]?
Who Q splashes the boy

'Who splashes the boy?'
examples based on van der Meulen (2004:81, ex.5)

An interesting linguistic analysis of French wh-questions can shed light on these linguistic and neuro- psychological data. Cheng and Rooryck's (2000) propose, following Cheng (1991), that wh-questions compared to yes/no question involve essentially two operations:

1. (1) *clause typing*, identifying in this precise case interrogative force, identified by [Q:] in the Table 2.45 that can be found in both yes/no questions and wh-question,[188] and
2. (2) interpretation of the *wh-phrase* for the identification of the content of the question expressed by nomenclature [Q: wh] that is restricted to wh-questions [189].

Cheng and Rooryck 2000/2003	French interrogative intonation	Chinese *wh-* Particle 呢
Checks Q- feature in C^0	√	√
Instantiates Q- feature in C^0		
[Q :] *wh- and yes/no-q*	√	
[Q : wh-] *wh-question*		√
[Q : y/n] *yes/no-q*		
root scope	√	√
embedded scope		√
island sensitive	√	
wh- word moves		
Feature moves	√	

Figure 2.45 – Table summarizing the syntactic characterization of French interrogative intonation versus Chinese wh-particle *-ne* 呢. Adapted from Cheng and Rooryck's (2000:18).

By splitting into two the properties of wh-questions, these authors argue that wh-in-situ in French is licensed by what Cheng and Rooryck's (2000) call an *intonation morpheme* whose function is clause-typing, allowing by its presence a yes/no questions interpretation of the SVO basic word-order, without Verb-subject inversion (I-to-C movement). Thus, these authors further argue that three types of question morphemes can be distinguished: [Q: wh] marking wh-questions, [Q: y/n] marking yes/no question and a third [Q:] type under-specified that can be found in both types of interrogatives.

These functional distinctions in offers the possibility to have an interesting comparative overview of french intonational phenomena (or morpheme [Q:]) with Chinese sentence-final interrogative particle *-ne* 呢, that we reproduce in Figure 2.45, showing some syntactic characteristics of these two interrogative linguistic phenomena. We can note how particular French interrogative intonation is: it can mark both plain yes/no questions and wh-question while the interrogative *-ne* particle in Mandarin can only mark questions where an in-situ wh-word is already present like in (54e).

Hence, the comparison between French plain yes/no questions where only the interrogative intonational morpheme is present ([Q:]) and wh-questions will allow to observe at the level of cerebral activation the cerebral correlate of these different interpretative values, separating through fMRI contrasts strategy the interrogative force of [Q: wh] marking Wh-questions from the second role of wh-interrogative pronouns that is the identification of the content. Importantly, this configuration will allow us to separate

188. Authors argue the default value of [Q :] is yes/no question ([Q : y/n]).
189. In facts, i goes in facts without saying that all wh-phrases are not inherently interrogatives an that wh-movement is not triggered always by the necessity of interrogative clause typing as we saw indeed for Topic or relative clauses.

wh-movement related cerebral activations from interrogative force interpretation ones in chapter 6.

Basically, the argumentation presented so far, is what led us to posit that the two operations implied in wh-questions happen to be the properties of two distinct heads in the Left-periphery: Clause-typing being a property of Inter, while identification of the wh-phrase, a property of Foc.

Interestingly, these two operations involved in wh-questions can be further specified following Aboh and Pfau (2011): While interrogative force could be generally characterized as the search for "new information", Focus position in the left periphery would house a process that could be called as q-identification, being an supplementary device stemming from information-structure allowing the identification of the target about which new information is sought, a necessary step to interpret what the question is about.

In short, this would mean that while Int functional head clause-types the sentence, Foc one hosts an operator that assigns a range to the variable that represents the target. This process could be schematically summarized as follows[190]:

1. *Clause-typing*: InterP sets interrogative force
2. -> search for new information
3. *q-Identification*: FocP ranges over the propositional content and its variables (e.g., argument, event, adjunct), and
4. -> provides a value to new information.

Further confirmation for the difference between yes/no questions and wh-questions comes from their observed distinct cerebral activation patterns. An fMRI study by Ben-Shachar and colleagues (2004, Exp.2) compared the effect ofembedded wh-questions and yes/no questions in an embedded context as shown from the sentence stimuli in Figure 2.46A. The participants were presented with the critical sentences in an auditory comprehension task in which comprehension questions referred either to the adjective, the verb, the embedded subject, or the object.

Results in Figure 2.46C show that wh-questions compared to yes/no questions yielded stronger fMRI activation in left Broca area (IFG), left ventral Precentral Sulcus (vPCS) and left posterior Superior Temporal Sulcus (pSTS), accompanied by a marginally significant effect in right pSTS. Interestingly and in contrast with the literature generally showing different effects for the movement of the object compared with that of the subject -like in object or subject clefts (Den

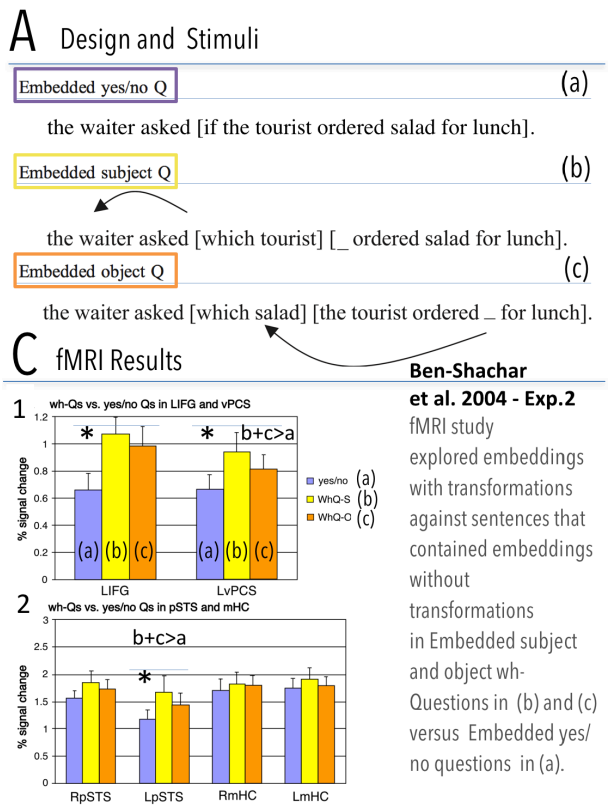


Figure 2.46 – (A) Experimental design and Sentence stimuli examples. (C) Percentage of BOLD signal change for the three conditions in frontal Broca's complex (LIFG) and Precentral Cortex (vPRC) in (1), and in posterior Left and Right ROIs (LpSTS and mHC) in (2). Adapted from BenShachar et al. (2004).

190. This last analysis instrumental in building our experimental hypothesis and materials for investigating French Question formation form the cerebral point of view in chapter 6.

Ouden, 2009)- object wh-questions did not show significantly stronger activation than subject wh-questions even when restricting the contrasts in the analysis by Regions of Interest (ROIs).

We can note that in this study again wh-movement yields an activation pattern that, compared to a non-movement derived baseline, seems to converge with the one observed in the study presented in previous section (cf. Fig. 2.35, p.183), where wh-movement was not associated with interrogative interpretation, but with topicalization in Hebrew. In Ben Shachar et al. (2004), wh-movement in questions increased activation in left inferior frontal gyrus (LIFG) and, as with topicalization, left ventral Precentral Gyrus (L-vPrCG) and bilateral Posterior Superior Temporal Gyrus (pSTG). These areas will therefore constitute some of our apriori Regions of Interest to try and replicate these wh-movement related effects with French wh-questions and Chinese topicalization, in the last two chapters of the manuscript.

2.4.4.3 The Sentence-Discourse interface

Continuing the characterization of the sentence Left-Periphery, it should be said that questions and their interrogative interpretation are not the only elements in the sentence that yield a interface with discourse. We will now turn to giving some essential elements about a syntactic account of the sentence's interface with discourse, by introducing a first broad analysis of Topic and Focus linguistic phenomena in this framework. This is meant to give a clear idea of how Topic phenomena can be given a definition in syntactic terms and to pinpoint the complexity parameter they add to the sentence syntactic structure. This will pave the way to question its syntactic characteristics in Contemporary Mandarin Chinese in next chapter (3) to theoretically and linguistically ground our experimental approach of chapters 4, 5 and 7.

We will now address here the question of what happens when a Noun-Phrase ends up in the Left-Periphery of the sentence. Let us start by considering, the examples in (56):

(56) a. Which Dissertation **Q** should you read $t_{<whichDissertation>}$ tomorrow ?

b. Your Dissertation **Top**, I should read $t_{<yourDissertation>}$ tomorrow.
b'. [“Topic”] Top° [“Comment”]

c. Your Dissertation **Foc** , I should read $t_{<yourDissertation>}$ tomorrow.
c'. [“Focus”] Foc° [“Presupposition”]

Examples (a), (b) and (c) all show a configuration where two kinds of interpretative properties are associated to the every same expression ‘your Dissertation’. For instance, in (b), it must be interpreted as the argument of the verb ‘read’ and as the Topic of the sentence. However, these type of constructions (i.e. A-bar movement chains), where a syntactic element typically occurs in two positions dedicated to two kinds of linguistic interpretive properties, raise the important question about where the scope-discourse property of ‘topicality’ should be interpreted in (b). In facts, while theta-roles (i.e. properties of argumental semantics) are determined with the verb assigning a thematic role to it immediate dependent[191], how is the Topic role assigned?

191. As we already saw in local relations of head-dependent configuration.

The answer coming from Cartographic approach is the following: A Topic head (Top°), occurring in the Left-Periphery of the sentence, activates an interpretative instruction at the interface with semantics-pragmatics for building the sentence into a Topic-Comment articulation as shown in (b'), that is to be read as "please *interpret my specifier as the Topic, and my complement as the Comment*", an similarly in (c') a focus (Foc) head guides the interpretation of the sentence according to the Focus – Presupposition articulation[192].

Importantly, this procedure establishing to have only one single interpretative property per head, allows to keep a certain simplicity in the functional structure of the sentence, respecting in facts the linearity and hierarchical dimensions of tree-like representation. So that, no head has to assign to its dependent the complex twofold property of being both [the patient-object of the verb] and [the Topic of the sentence]. On the other hand, this procedure increases the complexity of the sentence functional skeleton, and makes abundant use of movement to different structural position to pick up various syntactic specifications.

Thus, this procedure can be summarize with an example where it appears clearly that the two functions of the Topic Noun Phrase 'your PhD' are assigned at different steps:

1. (1) [I [will [read [your Dissertation] tomorrow]]]
2. **->External Merge**
3. (2) **Top** [I [will [read [**your Dissertation**] tomorrow]]]
4. **->Search operation**
5. (3) Top [I [will [read [your Dissertation] tomorrow]]]
6. **->Internal Merge**
7. (4) [**your PhD**] Top [I [will [read <[*your Dissertation*]> tomorrow]]][193]

These steps retrace a clear-cut division of labor between External and Internal Merge, when there are two kinds of interpretive properties at the interfaces with semantics and pragmatics: The expression 'your PhD' is merged in two positions, where it picks up the interpretive properties of Argumental "patient of read" by Internal Merge, and Discourse "Topic" by External Merge, respectively. Consequently, External and Internal Merge are not just extra rules to aliment the formal system, but play a critical role at the interfaces with semantics and pragmatics. They appear to be dedicated to the expression of two types of semantic properties: (1) External Merge expresses (among other things) argumental semantics (who does what to whom), while (2) Internal Merge expresses here (among other things) Scope-Discourse semantics (Scope of operators and Discourse related properties like topicality, focus, etc.)[194].

Last but not least, we have to note that in this process a fundamental role is played by the *search operation*, as previously described, this operation operation is conducted

192. This approach to scope-discourse semantics puts forth the hypothesis that scope-discourse properties (e.g. the scope of operators, topicality, focus, etc). also are structurally determined by local head-dependent relations in the left-periphery, being reduced to the syntactic schema of Spec-Head-Complement that we saw approaching X-bar theory.

193. As a side note: we can see in these step by step syntactic structure building operations, the order of the Merge operation is crucial to derive the right structure.

194. Note that as observed by Marantz (2005) a general minimalist guideline is to consider movement as a device to express an interface effect, which is, for instance, particularly relevant in the case of Topics and Focus, as we can see in this example.

within a primary buffer to identify a candidate for Merge. The identified candidate is then merged with the whole structure.

Put differently, looking into the sentence for candidates to find an element bearing its same features (as presented in Fig. 2.27) could be understood as playing the role of a "big brother eye" screening over the whole trees structure constructed so far to find a suitable candidate to perform the internal move step. We could therefore speculate that a brain area sub-serving this search operation would be possibly representing the sentence tree-structure wholly to be able to search though it screening for possible candidate having the right features for internal merge to happen. This speculative reflection of mine will be further discussed in front of experimental data in chapter 6, it will in fact interesting to ask the question of its neural implementation to our fMRI study on French syntactic transforamtions, where we will correlate brain activity with the number of different movement-derived sentence transformations.

The gist of splitting the CP layer

Before concluding, we should address some of the conceptual arguments for the introduction of an articulated structure in Complementizer domain as illustrated in Figure 2.44 (p.201) and for the relative ordering constraints that are found among its functional projections.

Reduced to its essential, the idea of splitting the CP-layer stems from the fact that not only wh-constituents undergo movement to the clausal Left-Periphery, as we just saw for wh-questions, but also Topics. Moreover, when we deal with sentence structures where the verb selects a complement phrase as its argument[195], also this phrasal complement is represented as a CP in that it needs to be '*clause typed*'. Consider for example, verbs like '*believe*', '*wonder*' or '*know*' that select as arguments different *clause types*: [*believe* + declarative clause], [*wonder* + interrogative clause] or [*know* + declarative or interrogative clause].

As for the relative ordering constraints analyzed by the cartographic approach, the example in (57) shows that the relative operator, as suggested by Rizzi (1997, p.289), occupies the highest specifier position, the Specifier (Spec) of Force above the Topic position, while the inverse in (b) is ungrammatical. Consider the syntax of the bracketed relative clauses in the simplified structures shown in (57a') below. As (a) shows a pre-posed wh-expression '*in which*' precedes the pre-posed Topic '*that kind of behavior*', then in (a') the pre-posed relative operator expression occupies specifier position within the Force Phrase.

(57) a. A university is the kind of place [**in which**, *that kind of behavior*, we cannot tolerate]

a'. [$_{\text{ForceP}}$ in which [$_{\text{Force}}$ ___[$_{\text{TopP}}$ that kind of behavior[$_{\text{Top}}$ ___] [$_{\text{TP}}$ we cannot tolerate *t* *t*]]]%TODO muri: deux "t" ?

b. *A university is the kind of place that kind of behavior, [in which we cannot tolerate]

Similarly, the contrast between (58a) and (b) indicated that Topics precedes interrogative pronouns in the sentence's skeleton (Lahousse, 2003:111)

(58) a. Ce livre, qui l'a lu ?

b. * Qui, ce livre, l'a lu ?

195. Because of its lexical restrictions.

These examples in English and French show how the ordering constraints can be gradually built by minimal pairs. For those characterizing the Left-Periphery of Mandarin we will rely on the work of Badan (2008, and her subsequent related work).

A more general question may arise looking at the cartographic map of the sentence-discourse interface when looking for metrics for sentence complexity representation in the brain: are of all the functional heads contained in this fine description actually realized as soon as an element is moved or generate in this sentence domain? The answer from linguistic theory comes by briefly summarizing the essential gist Rizzi's split CP hypothesis: in structures containing a topicalized and/or a focalized constituent, the CP layer splits into a number of different projections- namely a (1) Force Phrase, (2) Topic Phrase and/or Focus Phrase, and (3) Finiteness Phrase, as presented by Radford (2006).

Interestingly, this means in the position taken in Rizzi (1997)[196] that the C-domain only splits into multiple projections in sentence structures containing a topicalized and/or focalized constituent. More generally speaking, this syncretization may be due to Economy principles determining that a head is only projected as an independent head if it has semantic content that requires this, or if it is overt and many other possibilities. Although, on this last issue the position is nowadays different and debated in this theoretical framework, we can nonetheless speculate that a neuro-imaging design pertaining on this issue could be rather informative on the complexity effect that syncretizing or not split-CP projections might engender at the level of brain activity[197]. This question about the actual representation of all the syntactic layers is very important in the perspective of using the Cartographic sentence representation as a metric for syntactic complexity in the brain. The existence of (i) ordering restrictions on the different functional elements populating the Left-Periphery – crucially yielding agrammaticality when not respected —, and (ii) the existence of overt morpho-syntactic heads encoding for them, are already two important pieces of evidence to consider that these ordering rules and morphemes are to be encoded somewhere in the brain. But, another experimental question on the portion of structural skeleton to be built could be namely asked in the following terms: Is the height of a syntactic position within a given syntactic domain encoded somewhere in the brain. Or put differently, where are represented all the branches of the sentence functional skeleton that have been 'activated' by the presence of a Topic, during the building of the cerebral representation of a topic-comment articulation? These question will be asked by investigating the cerebral representation of the sentence structure in the two fMRI designs in the last two chapters.

These remarks pave the way to next sub-section where we will present strong empirical evidence for the neuro-linguistic relevance in aphasiology and neuro-imaging research of the theoretically introduced different Sentence Domains.

2.4.4.4 An apparently *vulnerable* C-Domain

The interest for considering sentence domains in neuro-linguistics can be show in a series of concrete examples, both from experimental studies and from the study of populations

196. The authors takes nowadays a different position on this. Rizzi (1997) posited that in a structure containing no focalised or topicalised constituents, the Finiteness head is *syncretised*, in other words 'conflated', with the Force head immediately above it, so that rather than being realized on two different heads, the relevant force and finiteness features are realized on a single head corresponding to the classical C constituent (i.e. a composite of force and finiteness head).

197. This is an idea that will be pursued in chapter 7 though a linear contrast.

having impairments in language behaviors, encompassing aphasic Patients, children with Specific Language Impairments (SLI).

In this domain of research, compared to neuro-imaging data, evidence comes from a more varied linguistic sample[198]. Experimentally asking the question about the relevance of sentence domains in patient or acquisition studies implies selecting linguistic material distinguishing syntactic operations involving different Sentence Domains, like the C-domain only versus and the I-domain only[199].

A study by Platzack (2001) present evidence from SLI, L2 learners and Broca aphasics of Swedish, showing a different syntactic behavior in production of sentence's lower structural levels compared to higher ones[200]. The three populations were test on three different Swedish sentence structures shown in Figure 2.47, namely subject-initial clauses (with transitive verb) in (a) , non-subject-initial clauses (Verb second, V2) in (b) and embedded clauses in (c), and showed a performance pattern confirming the "vulnerability" of the CP domain even in highly frequent sentence articulations and word-orders.

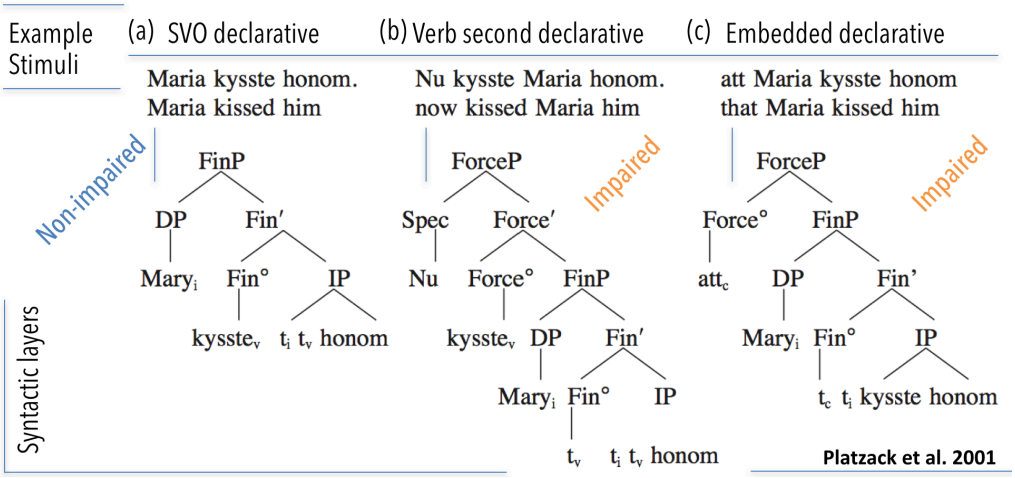


Figure 2.47 – The tree sentence structures tested by Platzack (2001) across three populations (i.e. Children with SLI -Selective linguistic impairment-, L2 learners of Swedish, and Broca aphasics). While (a) sentence structure shows target-like production, (b) and (c) are impaired or mastered later in the three populations.

The main interest of this study lies, in fact, in the advantage of using very frequent and basic declarative main clauses to test the difficulty of C-domain processing. Swedish offers the possibility to de-correlate the presence of a CP syntactic layer from parameters linked to frequency in language use and from other complexity factors that usually associated to syntactic embedding, like semantic reversibility, working memory, etc. Swedish is a Verb-second language (V2 with a VO order), in which main clauses present a finite inflected verb (for tense, but not for agreement). The finite verbs are generally preceded

198. The reader couldn't really realize this because I purposely selected studies coming form different languages to enhance the strengths of the assertions done though cross-linguistic validation.
199. In chapter 6, we will review how different movement types targeting different Sentence Domains show different acquisition patterns in Hebrew children §6.1.1.2 (cf. Fig. 6.5, e.g. Fridemann and Lavi, 2006), p.516).
200. Note that true performance errors in normal/control populations are of the order of magnitude of 2%.

by one constituent, usually the subject, but in about 30–40% of the sentences it is an object (both in written and in oral register), a predicative element or an adverbial that precedes the finite verb. Importantly, as shown in Figure 2.47 in example (b), when the finite verb is preceded by a non-subject constituent (in 30 to 40% of the cases) it moves to the CP layer, as shown in the syntactic-tree in 2.47b. While showing no impairment in the production of sentence's lower structural levels, the three populations all showed a consistent impairment or difficulty for sentence structures involving the Complementizer (CP) sentence Domain even in sentence structure that is a highly frequent one in Swedish like (b). This initial evidence for an increased complexity in mastering the highest layers of the syntactic-tree in these tree populations is further confirmed by complexity effects observed in fMRI studies, and was formulated by Christensen (2008) as the '*Sentence Domain Hypothesis*'.

Sentence Domain Hypothesis and neuro-imaging evidence

More precisely this hypothesis posits that differential activations in the sub-components of the sentence cortical network could reflects the computation of different syntactic domains -the interface levels between syntax, semantics, and pragmatics.

Roughly speaking the division of labor proposed by such an hypothesis attributes different activation patterns or sub-components of the language network to the Complemetizer Phrase, this particular sentence domain that we previously presented closing the IP-domain and linking the propositional information of IP and VP to discourse.

One could speculate -and this is a position I could intuitively subscribe to- that this hypothesis is to be seen as a way to 'modularize' the sentence structural layers into three modules being respectively specialized in (1) Dealing with the syntax/semantic interface for theta-role assignment in VP ; (2) connecting the verb with its Subject and Arguments in IP; and (3) Orchestrating the interface between propositional content of IP and sentence discourse-related interpretations[201].

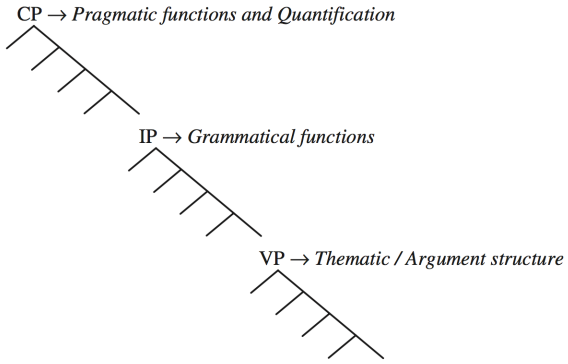


Figure 2.48 – The division of labor proposed by the Sentence Domains Hypothesis from Christensen (2008).

201. Namely, we could speculate that in case of movement across the main sentence's projections, this modular configuration would imply a mechanisms dedicated to search for target position in the target domain, a kind of mechanism like the Search process that is prior to Internal Merge (cf. previous description). This overarching mechanism (or module) that would be in control of the search process in cross-domain movements, could actually be the locus where the different rules and constraints on Movement, like the one specified by Relativized Minimality, are represented. Using a metaphor this overarching syntactic module could be the Big Brother of the sentence's syntactic skeleton, knowing about the serveral constraints that exist in the fine-grained mapping of the sentence's hierarchical structure. However, the problem with this modular view of sentence domains is that by definition modules should be informationally encapsulated, meaning that in order to run its operations a module would only use its own information being blind to other sources of information, which actually should be a problem for an interface domain like the CP. This point not only rises a problem, that was already addressed by Fodor ("Thus, we expert that modularity could be a matter of degrees." 1983:55). But, it also poses a series of problems regarding at what level the Search operation is performed. Namely, if it ought to

For instance, (1)***Inflectional nodes*** would include an agreement phrase (AgrsP, the lowest functional node) which would represents agreement between the subject and the verb in person, gender and number; (2)***Tense Phrase*** (TP) above it, tense inflection of the verb would be represented, verbs would move there in order to collect their inflection like finite verbs; while (3) the ***Complementizer Phrase*** (CP), highest sentence node in the tree is t, would hosts complementizers such as "that" or wh-words like "who" and "what" that moved from the base-generated position within the VP, and the auxiliary in yes/no questions or the verb as we saw in this chapter, or Topic and Focus heads determine the sentence-discourse interface.

Reduced to its essential, what Christensen (2008) calls an *'interface' approach* to the neuro-syntax imaging results present in the literature, builds on the hypothesis that left IFG (Broca) is involved in the interfacing between the (computational) system of syntax and other cognitive systems, including information structure. It is argued that a correlation between left IFG activation and syntactic movement operations involving the top-most node of the clause – the Complementizer Phrase (CP) – is observable in several studies (see Table 2.6, p.216). As we can see in Table 2.6, the CP is involved in topicalization, questions and sentential embedding like relative clauses.

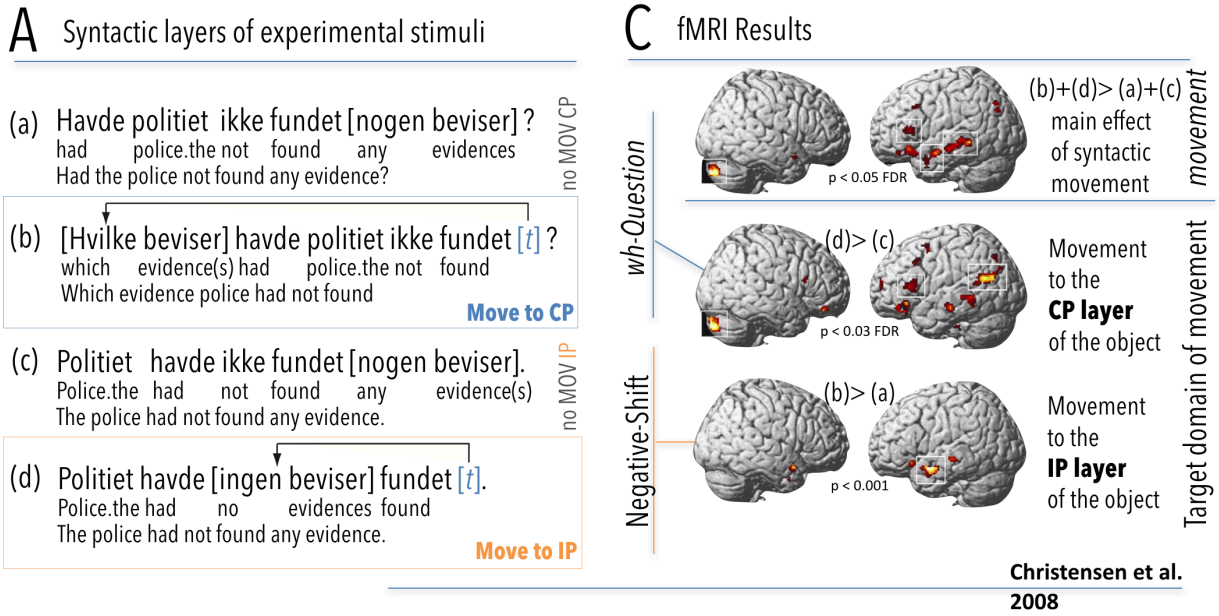


Figure 2.49 – Sentence Domains in three Danish sentence structures where IP and CP are present. (A) Experimental design and Sentence stimuli examples. (C) fMRI brain maps for syntactic effects where constituents target or not the CP Sentence Domain. Adapted from Christensen (2008).

The first experimental evidence for these claims comes from a study by Christensen (2008) where different movements targeting different sentence domains where compared. As one can see from the experimental sentences 2.49A the canonical position, or the 'base-position', of the object is the same throughout the conditions, whereas the position

be performed at the level of the output of each sentence domain module, then it wouldn't be a problem indeed. Hope these speculations could find some lively discussants in the near future.

the object moves to a target positions that are different in the two types of constructions selected : (a) Negative-shift, motivated by sentential negation and (b) wh-movement triggered interrogative force. While Negative shift moves a negative indefinite quantified object to the position of negation in the middle of the clause, i.e. in the IP-domain, in order to license sentential negation. Wh-movement target the CP layer. And as we can see from the fMRI results in 2.49C only wh-question activate frontal areas like Precentral Gyrus and Broca's area (BA45).

This division of labor across syntactic layers can be found in other studies, whose results despite not having been interpreted according to this working hypothesis constitute nonetheless a support for this 'modularized' understanding of sentence syntactic layers.

For instance, Roder et al. (2002) fond a similar pattern of results contrasting word-order variations in German, of the kind *Jetzt wird der Astronaut dem Forscher den Mond beschreiben* [canonical W-O: S-IO-DO] and "*Jetzt wird den Mond dem Forscherder Astronaut beschreiben*" [non-canonical W-O: DO-IO-S] (*Now will the astronaut to the researcher the moon describe.*). Conditions (a) and (b) (see Figure 2.50A), presenting Movement in the VP-domain, showed activation in the right hemisphere in Broca's BA44/45, while the syntactic effect of (c) and (d) was observable in left Broca and precentral gyrus, with additional implication of the SMA as shown in Figure 2.50C.

Likewise, the pronominal scrambling inside IP domain in Grewe et al. (2005) did not activate Broca's area, and the lack of difference in activation between subject and object embedded questions (Ben-Shachar et al., 2004) that we noted in Figure 2.46 (see page 205) can be attributed to the fact that both target spec-CP.

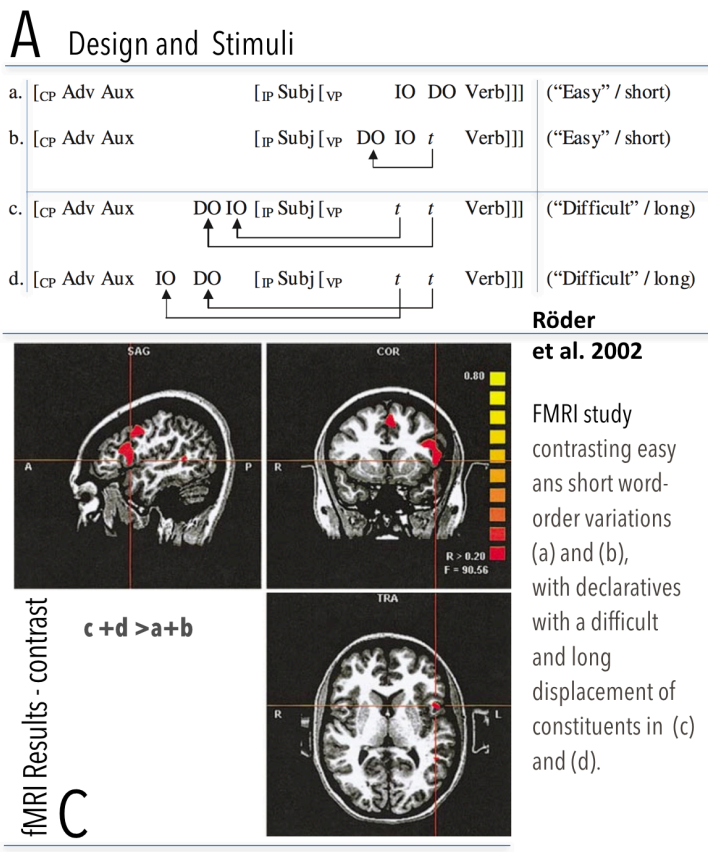


Figure 2.50 – (A) Experimental design and Sentence stimuli examples. (C) fMRI brain maps of syntactic effect (c+d>a+b). Adapted from Roder et al.2002.

Analogously, Ben-Shachar et al. (2004 - Exp. 1, see Figure 2.51) contrasted topicalization in (c) and (d) against Dative-shift in Hebrew in (c), and reported that IP internal movement deriving Dative-Shift sentence [S+V+IO+DO] does not yield Broca's activation. Only wh-Movement to the CP domain appears to activate Broca Complex and Precnentral Gyrus, as shown in the the brain map and bar-plots in 2.51B.

Note that these results are not isolated increased activation for object- vs. subject-initial sentences (with a maximum in BA45) has also been reported for Japanese. Kinno's

et al. (2008) contrasted scrambled sentence[202] (SS) condition *X-o Y-ga hiiteru* 'As for X, Y pulls [it]' in a picture-sentence matching task the contrast between reveals three clusters of activation: left Precentral Gyrus (BA6/8/9), Broca's area (dF3t/BA45) and posterior Superior Temporal Gyrus (pSTG/MTG BA22/21/37). Not only these results

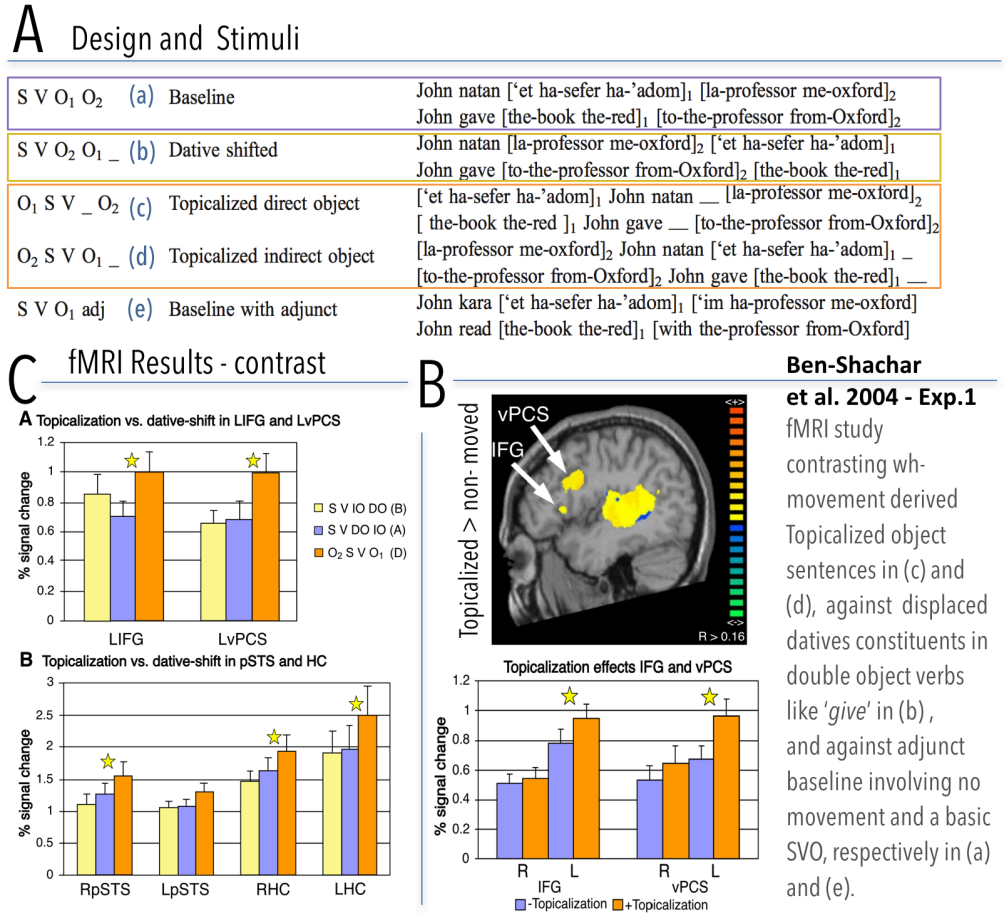


Figure 2.51 – (A) Experimental design and Sentence stimuli examples. (B) Brain maps for contrast Topicalization > non-Topicalized sentences, and signal percentage change in frontal ROIs in left and right IFG [+ Topicalization conditions] (in orange) vs. [- Topicalization conditions] (in blue) and signal percentage change in left and right IFG. (C) Percentage of BOLD signal change for the tree conditions (a) (b) and (d) in frontal Broca's complex (LIFG) and Precentral Cortex (vPRC) in (1), and in posterior Left and Right ROIs (LpSTS and HC) in (2). Adapted from BenShachar et al. (2004 Exp.1).

from Hebrew and German sentence processing land support to the Sentence Domain Hypothesis through different movement types and syntactic constructions, but a retrospective examination of the literature on movement is summarize in the following Table 2.6 showing that a series of findings can be taken to corroborate the hypothesis that the different neural pattern are observed when processing movement-derived sentences

202. Nota bene: the stimuli didn't imply a *-wa* marked Noun Phrase, *-wa* being Topic marker in Japanese, hence, no Topic analysis can be driven for this results, it is as the authors declare it a Scrambling syntactic configuration.

whose constituent target different syntactic domains. We will further continue on this topic, taking into account results from aphasiology that bear on a similar issue. But before, as announced in chapter 1 (§1.4.5.2, p.61), it is noteworthy that a generally under-considered brain area like Precentral Gyrus (BA6 in its ventral and dorsal part) is here stably reported, and more broadly present in the kind of syntactic manipulations of the sentence structure reported in this chapter as a whole – we will come back to it in the experimental Part of the manuscript.

Table 2.6 – Sentence Domains Hypothesis. IFG, inferior frontal gyrus; aIns, anterior insula; vPrCG, ventral precentral prefrontal gyrus; ant. Cing., anterior cingulate gyrus; Heschl's gyrus; pSTG, posterior superior temporal gyrus; PrFG, prefrontal gyrus; L, left, R, right; M, medial. Adapted from Christensen (2008).

Target domain	Movement contrast	Study	Activation clusters					
			Broca/aIns	Precentral (VPrCG/BA6)	Heschl's gyrus (BA41/42)	p.Superior temporal gyrus	Dorsal middle frontal gyrus (BA6/8)	Anterior cingulate (BA24/32)
CP	Obj. rel>emb. decl.	Ben-Shachar et al. 2003	L			L+R		
	Wh>yes/no	Ben-Shachar et al. 2004	L	L		L+R		
	Wh>subj. Initial	Ben-Shachar et al. 2004	L	L	L	L+R		
	Topic>re-serial.	Dogil et al. 2002	L	L		L	M	M
	Long>short subj. (scrambling above subj.)	Fiebach et al. 2005	L+R			L+R		
	Subj–IO>IO–Subj (scrambling above subj.)	Grewe et al. 2005	L+R					
	Double obj. scrambling	Grewe et al. 2005	L oper					
	Easy>difficult (scrambling above subj.)	Röder et al. 2002	L+R	L		L		M
IP	Long>short subj. (scrambling below subj.)	Fiebach et al. 2005	-	-	-	-	-	
	Pronominal scrambling (scrambling below subj.)	Grewe et al. 2005	-	-	-	-	-	
VP	Dative shift	Ben-Shachar et al. 2004	R	R				
	Semantic>non-semantic	Röder et al. 2002	R					

2.4.4.5 Recovery up the trees in aphasics

Yet, another set of neuro-psychological evidence has to be considered in developing our arguments for sentence complexity and hierarchies across the syntactic-tree skeleton. A growing amount of observations from agramamatic sentence production studies and from developmental ones, namely, concerns the correlation between the position of sentential elements in the syntactic-tree and the syntactic impairment patterns observed in these populations.

In this regard, Hagiwara (1995) first suggested that in individuals with agrammatic aphasia higher nodes in the tree are more 'vulnerable' and impaired than lower nodes. Studying the production and grammaticality judgments of a group of Japanese patients, and contrastively considering data from French and Italian aphasic speakers, Hagiwara showed that the impairment of syntactic elements whose position is within the Complementizer domain (CP)had no impact on the impairment of elements represented in the lower nodes, like it is the case for Tense marking and Negation. Importantly, no case was observed in which patients showed an opposite pattern of impairment, that is to say spearing higher nodes syntactic, and affecting syntactic phenomena taking place in the lower functional heads.

This vision of the hierarchical structure of a sentence, establishing the lower the position of a functional head (and its projection) the more accessible it is for agrammatic aphasics, was further confirmed by Friedmann and Grodzinsky (1997), who reported the case of a Hebrew-speaking patient showing an even finer dissociation in his ability for verbal inflectional morphology: tense inflection (and use of copulas) was impaired, but not agreement inflection.

This impairment pattern was then correlated to different position in the syntactic-tree respecting the high/low complexity characterization: agreement takes a lower position in the tree than tense as see in Figure 2.52. Confirming in this way that lower functional projections in the tree are more accessible than those in higher positions[203].

Moreover, embedded structures were severely im-

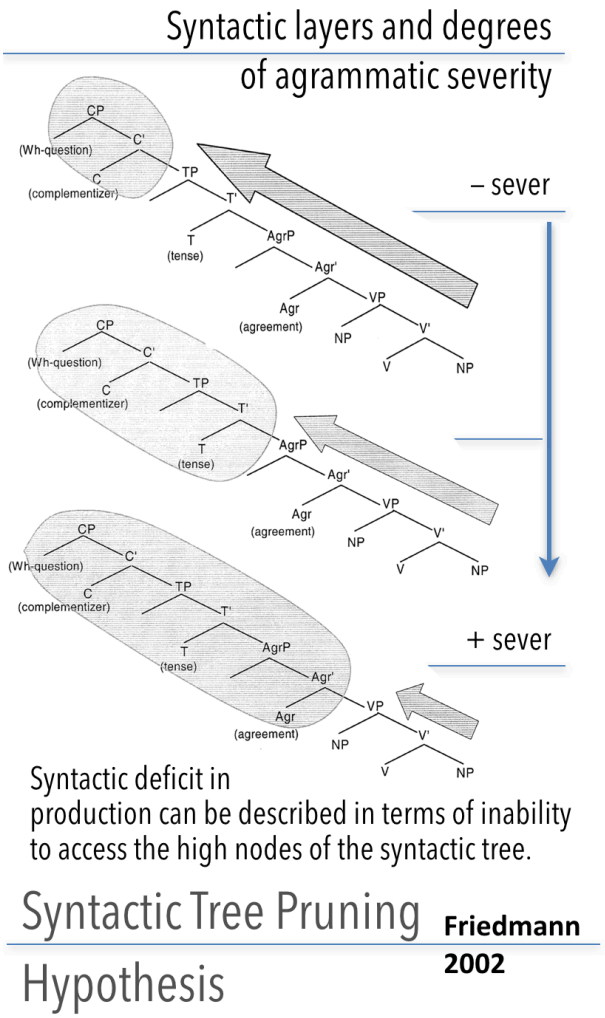


Figure 2.52 – Tree-Pruning Hypothesis. Schematic representation of how the production syntactic deficit in agrammatic patients can be described in terms of inability to access to the higher nodes of the sentence's syntactic-tree skeleton. Syntactic domains circled in grey represent the extent of inaccessible tree-nodes according to the degree of severity of the syntactic impairment. Adapted from Friedmann (2002)..

203. These findings showing selective impairment inside the inflectional Phrase can also be seen as bearing upon the central theoretical issues of splitting functional projections of the sentence functional skeleton, as we saw in this section for the CP-layer (Rizzi, 1997, p.201) and in chapter 1 with the split of the inflectional domain (Pollock, 1989).

paired in this patient which led the authors to formulate the Tree Pruning Hypothesis (TPH), positing that when the lower nodes are impaired, projecting higher nodes in the tree is impossible (also see Friedmannn, 2001, 2002), as if pruning shares had cut down higher tree branches, impairing agrammmatic performance for the syntactic functions housed by cut node and higher up, see Figure 2.52.

Accordingly, higher nodes are more complex than lower nodes. This fits well with our conceptualization that wh-movement is more complex than NP-movement as developed in previous section (§2.4.3.2, see Table 2.5, p. 195). The Tree Pruning Hypothesis was also developed also to account for the patterns of agrammatic sentence deficit in relation to wh-Question production (Friedmann, 2002).

In a group of 19 Hebrew- and Palestinian Arabic-speaking individuals with agrammatism, three distinct patterns of performance were found for production as illustrated in Figure 2.53B.

Friedmannn (2006) suggested that these differences in production can be accounted for by the height on the syntactic-tree a patient can access. Using a metaphor: the higher the patient can climb on the tree, the milder the impairment will be as we can see from the schematic representation presented in Figure 2.53A and B. Syntactic structures that relate to high nodes of the tree are impaired in agrammatism, whereas lower structures result being unimpaired. Moreover this can partly reveal the role of the impaired brain area (Broca's area) in syntactic production.

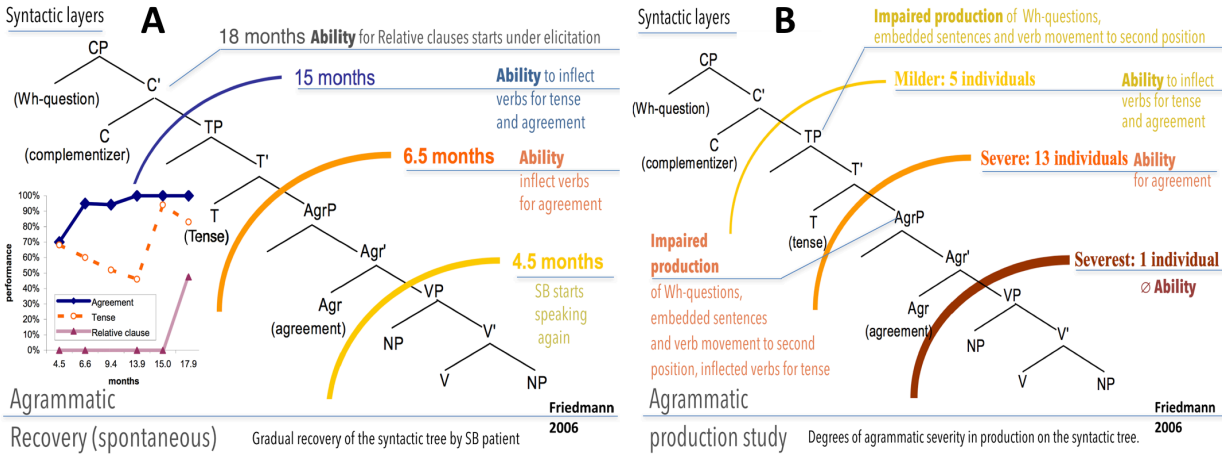


Figure 2.53 – (A) Spontaneous Recovery of SB patient Agrammatic Patients across syntactic layers; (B) Tree Pruning Sentence production patterns ans severity of 19 Agrammatic Patients across syntactic layers. Adapted from Friedmann (2006a/b).

Another study by Friedmann (2006) repeatedly tested the production of functional elements and syntactic structures of a young woman with agrammatic aphasia (SB, 20 years old), whose syntactic impairment was consequent to a traumatic brain injury followed left craniotomy (in left hemisphere including temporal, parietal and frontal lobes, lateral ventricle enlarged, and right hemiplegia). The gradual recovery of this patient, illustrated in Figure 2.53A, shows that the relative order of functional elements on the tree-representation can explain the order of gradual spontaneous recovery of SB's syntactic abilities. Thus, 'the recovery of SB can be described as gradual climbing on the

syntactic tree, at each stage obtaining access to a higher node of the tree' (see Fig. 2.53A). Once more the integration of neuro-physiology/psychology and linguistic theory appears to open interesting perspectives, and syntactic-trees, as a representational format, turn up to be a useful tool to describe the deficit in production of the agrammatic cases.

In sum, if we compare these findings with the ones from treatment studies examining recovery of sentence production and comprehension in agrammatic aphasia (§2.4.2.2, p. 177), we can note that the two approaches converge in identifying the Complementizer Phrase (CP-layer) as a relevant complexity metrics to be considered both in understanding syntactic impairment and in designing treatment. This view angle on sentence's syntactic complexity provides strong empirical grounding for our interest in the CP syntactic layer and in the sentence-discourse interface linguistic phenomena it 'houses'. The overall findings presented in this last section show how investigate the neural underpinnings the syntactic-tree representation could be essential for understanding sentential representation and mechanisms at the level of the brain.

Not only the syntactic-tree ordinality can describe the neuro-psychological recovery and deficit severity patterns as climbing up the syntactic-tree structure, but the neuro-imaging results, reviewed when addressing the 'Sentence Domains Hypothesis', seem also to confirm that the syntactic tree-skeleton can be considered as a neuro-psychologically grounder representational format, in that it can be hypothesized that the different layers of the syntactic-tree have different cerebral underpinnings (cf. Table 2.6, p.216).

It is in these last findings that resides the ultimate interest of selecting such a representation format for sentence's structural complexity.

Cerebral representation of the sentence-tree skeleton

In conclusion, a few words about this neuro-linguistic vision of the syntactic tree are necessary before we proceed. The general idea resides in systematically mapping the functional nodes present in a syntactic tree onto the ability to understand or produce sentences, whose structures require the representation or the movement of one of its constituents to a certain site (or height) in the syntactic tree. The proposal of understanding syntactic comprehension deficit, as deficits in the construction of the syntactic-tree of an input sentence, constitute a shift towards taking into account the representation of the stimulus, more than focusing on the processing or the manipulations and operations that can be done on syntactic-trees. Under the Tree Pruning Hypothesis what is impaired is the building of one or more syntactic layers, that therefore will not possibly host the displaced constituent.

Similarly, what has been described about the 'Sentence Domains Hypothesis' seems to go in the same direction, namely advocating that once the syntactic-tree structure is taken as an interiorized representation, it will cause the brain activity in some areas (according to the authors, Broca's area) to depend on the height in the tree-hierarchy of functional nodes where the elements of a given sentence are hosted, or specifically where the elements having undergone movement land.

These frameworks will be essential in our experimental approach to Topic-Comment articulation and other Left-Periphery phenomena like Topicalization and Focalization in Mandarin Chinese (cf. chapter 7). Thus, investigating different ordinal positions in the left-periphery will allow us to compare the fMRI activation patterns related to syntactic position differing in syntactic-tree height, and to search for a brain area possibly encoding for the relative height in the cartography representing the Mandarin Left-Periphery.

Moreover, we will leverage on Topic-prominence parameter in Chinese to compare fMRI activation patterns for Topics that are Base-generated in the Complementizer Phrase, to those for moved ones, and to moment-derived sentence structures that target lower IP positions.

2.5 Interim summary

In this chapter we defined the sentence as a cognitive object by presenting a neuro-linguistic account of this syntactic unit and its internal structural complexity.

By arguing that the sentence is a natural, universal and complex syntactic-unit we introduced some basic concepts and formal tools that the mind or the brain might need in order to represent and understand sentences.

Carrying along the question of how these syntactic processes are instantiated in the human brain we depicted an original recollection of the neuro-imaging data, that gives evidence for the cognitive pertinence of some of the descriptive formal tools and processes that were formalized by linguistic theories, like :

1. the hierarchical structure building operation called *Merge* ;
2. *Binding phenomena,* as a way to manage reference assignment inside a sentence ;
3. word-order properties and interpretative value inside the sentence
4. a structure building process called *Movement*, operating on syntactic trees, that derive new structures from basic ones ; and,
5. of the so-called *Empty Categories*, as structurally needed elements that are understood but not phonetically realized.

While introducing these syntactic processes, we delineated the linguistic and neuro-psychological interest of the four major axes of our research. This state of the art served namely as a gradual introduction to the experimental questions we will be dealing with in the experimental chapters 4, 5, 6, and 7.

The discovery of the neural underpinning of these more abstract syntactic processes remains the ultimate goal of many experimental attempts nowadays, and we will adopt in the experimental part of this manuscript a concretely linguistic-oriented approach investigating particular syntactic configurations where these elementary syntactic processes can be de-correlated.

As an interim summary, we can list these research questions that will be addressed in our experimental approach:

2.5.0.1 Building sentence's articulation and hierarchy by Minimal marking

There is no doubt spoken sentences carry much more information than written ones, but more specifically we saw in section §2.2.1 (p.101) that they bring cues to syntactic grouping of words that can greatly impact sentence interpretations. We saw how different meanings accordingly to the grouping of the words that is assigned by oral phrasing and pausing signified by the punctuation marks at different places in the sentence.

This prosodic chunking of the utterance is universally required to understand a sentence, and its essential role has been confirmed by experimental psychology, showing how syntax and sentence-level intonational patterns (pauses and syllables lengthening, etc.) are tightly linked.

Given that Prosodic information can perform the same function as syntactic markers, we will experimentally address this syntactic function in chapters 4 and 5 by observing behavioral and Electro-physiological responses to sentences where this type of prosodic boundaries have been removed and the listeners can only rely on word-order cue to build and understand the syntactic structure of the experimental sentences.

An advantage of focusing on Chinese is given by its relatively scarce morpho-syntactic marking, and among Chinese syntactic structures, Topic-Comment articulations are indeed among the most unmarked structure permitting to study how the brain manages this incredible equilibrium the sentence achieves between linearity and hierarchy (cf. §2.3.3.1, page 143, Fig. 2.15).

Namely, in Mandarin Topic linear position and a minimal pause prosodic marking are the minimal cues required to maximally change the hierarchical relations in the sentence.

Hence, chapters 4 and 5 will offer a psycho-linguistic and neuro-linguistic account of two minimal hierarchical structure marking devices of the sentence-unit: (1) its prosodic hierarchical structure marking and (2) the syntactic information linked to its word-order.

2.5.0.2 The representation and processes linked to the Sentence's Interface with Discourse

The importance to describe the sentence as having an interface with the discourse and the effect this has on its internal articulation was thoroughly addressed in section §2.1.2 by showing the richness of pragmatic interpretation linked to utterance and discourse particles, in Section /S 2.4 and more specifically in 2.4.4 (page 197) by highlighting how a syntactic process like Movement can generate word-order changes where constituents target the sentence domain dedicated to the Sentence-Discourse interface.

Scope-Discourse semantics interpretation of sentences will be investigated in chapter 6 by addressing the difference between interrogatives and declaratives that have the exact same surface structure and words in French (cf. §2.4.4.2 and in Mandarin Interpretative features like contrastiveness or Focus discourse semantics. Contrastive Topic [SOV] or even-focus (*lian* Focus) will allow to test for the brain activation of the different sentence articulations and for the pragmatic interpretation linked to Focus-presupposition and Topic-Comment articulations.Importantly, the difference between simple word-order changes and the presence of overt morpho-syntactic marking for pragmatic contrastive interpretation will be observable in chapter 7.

Sentence as a Topic-Comment articulation

Section §2.2.4 introduced how the sentence-unit predication can be generally viewed as a matter of Topic-Comment articulation, once taken into account the aboutness role that the subject or the Topic can play in the sentence. This sentence level-articulation relativizes the otherwise central role of the subject and presents a clear-cut hierarchical structure where the topic and the comment occupy two different syntactic layers. The cerebral underpinnings of what it meant to build a sentence according to the Topic-Comment articulation will be addressed in an fMRI experiment on Mandarin in chapter 7.

Discourse contextual effects in Topic sentences will be investigated in order to offer a psycho-linguistic and behavioral description of Topic-Comment sentences in context in chapter 4. To study the two dimensions come into play during the online comprehension of Topic-Comment structures, the syntactic dimension linked to the hierarchy between

the Topic and the Comment clause, and (2) the informational status in the discursive context of the referent of Topic constituent.

This point will be explored through Electro-Encephalography of Mandarin Topic-Comment sentences *in context* to understand when and how discourse information is activated during the comprehension of sentences begin prosodically marked for a sentence discourse interface in chapter 5.

2.5.0.3 Representing Syntactic-tree complexity metrics

The different complexity metrics linked to the sentence's tree-structure representation were presented in Section 2.3, and one of our experimental focus will be to tackle the neural representation of abstract syntactic categories left by syntactic Movement.

Syntactic Movement

The question of the cerebral representation of syntactic Movement transformations will be addressed in chapter 7 through an fMRI study investigating the difference between moved-constituents and Base-Generated constituents to the Complementizer Phrase (CP). The Topic-Prominence of Mandarin Chinese offers indeed a perfect testing ground for syntactic movement effects thanks to the availability of a syntactic configuration where the complexity dimension linked to sentence initial Topic can be observed both with and without movement (and without embedding).

Differences between Movement types

The fundamental differences between the syntactic properties and configuration yielded by different types of movements reviewed in section §2.4.2 will be experimentally addressed in a dedicated fMRI design in Chapter 6. Question formation in French thought V-movement, wh-movement and their combination will be used to question their linguistics differences and possible additive effects. Moreover, NP-movement and clitic-movement will be used to investigate the neural underpinnings of more local movement types.

A modular syntactic representation of Sentence Domains

Abundant Neuro-psychological evidence and some neuro-imaging data were presented in sections §2.4.4 indicating that a difference in representation or processing of the three main syntactic domains that constitute the sentence structural skeleton has an experimental grounding.

As developed in section 2.4, Sentence domains and landing site of movement transformations are taken here as a complexity parameter, and the difference in targeted sentence domain for moved constituents will be addressed in the fMRI study of chapter 7, where the ordinal position in the cartographic mapping of the Mandarin Left-Periphery of the dislocated constituents will be correlated with brain activity.

Chinese will allow to investigate the cerebral representation of a constituent in the CP sentence domain, by testing the representation of the CP syntactic-tree layer in absence of embedding, movement and gaps. This syntactic configuration allows, in fact, to disentangle movement complexity from the representation of a higher syntactic position

in the sentence's syntactic-tree function skeleton. It will be then possible to observe the neuronal correlates of the linguistic distinction between base-generated Topics vs. moved Topics (by wh-movement).

2.5.0.4 Establishing an Antecedent-Trace dependency-link

The different syntactic aspects characterizing the presence of empty syntactic positions in the sentence, that we introduced in section §2.4.3, will constitute different testing hypotheses for the fMRI experiment in chapter 6, especially in relation to the number of abstract silent syntactic traces present in movement-derived French Interrogative sentence.

As observed in section §2.1.1 co-referential relationships inside the sentence are constrained by rules linked to clausal boundaries. Moreover, overt/cover realization of dependency will be investigated in chapter7 asking if there are brain areas that selectively respond to *traces* (i.e. empty categories) compared to a Resumptive Pronouns to realize the dependency link between the Topic and the Comment clause ? (cf. §2.4.3.1)

We will investigate three fundamental characteristics of the sentential linguistic unit, through and across the *parametric* variations of French and Mandarin Chinese:

1. the representation of the sentence's architecture in the domains that determine basic sentence structure - namely VP, IP and CP and their interfaces;
2. the processes and operations that determine complex syntactic structures - what we have been calling syntactic transformations - ; and
3. the dependency relations within a sentence achieved by overt or non-overt linguistic means.

In sum, these are the research question that motivated the cognitive and experimental research approach to the sentence-unit that will be carried over in Part II. These issues rooted in linguistic theory allow the drawing of specific experimental hypotheses about what the mind or the brain (minimally) need in order to encode a sentence-unit and its internal structure. In this lies the indisputable "trademark" of our theory-oriented neuro-linguistic approach: have a linguistic approach and carrying along the question of how the above four main syntactic processes are instantiated in the human brain.

Next chapter will present a more linguistically oriented analysis of Topic-Comment sentences, a syntactic configuration that offers a optimal testing ground for all our five main research questions.

Chapter 3

Sentences with a Topic

All the sentences of our everyday lafnguage, just as they stand, are in ***perfect logical order****.*

Ludwig Wittgenstein, *Tractatus* (5.5563)

In our preliminary delineation of language faculty and man emerged that one of its most distinctive character was its fundamental capacity of conveying his thought into a precise linguistic encoding: the sentence. This fundamental aspect granted him the name of *Homo phraseologicus* – a man to which Sapir (1921) would attribute the basic linguistic activity to say something about something, in other words, to predicate.

This chapter has its roots in the reflection that Baroque thinkers developed about predication and propositional logic, that we briefly addressed in chapter (§1.3.2, p. 39). The initial contribution of philosophers and logicians to the study of the propositional content and predicational mechanisms of the sentence-unit, identifying two separate predicative functions (i.e. a referring one and qualifying one) will serve as notional starting point. Thereby, we will introduce how several generation of linguists defined the different syntactic roles around which the sentence-unit is built, and how the syntactic articulation of Topic-Comment sentences was characterized[1].

Ultimately, our goal is to give a linguistic and cognitive characterization to the *perfect logical order* Wittgenstein is identifying in *All the sentences of our everyday language* in order to generate clear hypothesis to further investigate its mental representations and neural underpinnings.

While chapter 2 presented the theoretical and broad linguistic grounding of our experimental hypotheses, chapter 3 outlines the detailed syntactic and psycho-linguistic motivations (§3.1) behind our experimental research on Topic-Comment articulation. It provides a typological (§3.2) and syntactic analyses (§3.4) of the Mandarin Chinese Topic structures and Left-peripheral phenomena. Along this chapater we will gradually build the foundation of the two experimental designs on Chinese, presented in chapters 4, 5 and 7.

1. As argued in previous chapter a sentence-unit can host both a the Subject of its verb, and a Topic to play an aboutenss role or frame the sentence (cf. §2.2.3 and 2.2.4).

Structure of the chapter

The function and characteristic of topical phenomena are rooted at the interface of various domains, penetrating sentence, discourse and semantics. As an all-purpose definition of Topic is hard to come by, section §3.1 explores the extensive literature produced by different frameworks on the broad notion of Topic to offer relevant insights into the linguistic and cognitive characterization of Topics[2]. The outcome of this notional exploration will be relevant for our experimental approach and will facilitate the development of discussion on the typological features of the sentence-unit in Mandarin Chinese in §3.2.

After having shown to what extent this notion pertains to the cognitive/psychological domain, ever since its early formulations, we will present the central role played by Topics in "Chinese sentence" and enter in the details of the typological claims about the articulation of sentence structure in Mandarin. Linguistic phenomena linking the Topic and the Comment (i.e. anaphoras, resumptives and empty categories) will be analyzed, and the derivation of Topic-Comment articulation (i.e. base-generation vs. movement) will be discussed for Mandarin (§3.4). The introduction of psycho-linguistic evidence – form language acquisition, language development and agrammatic linguistic production– will prevent this chapter form being a linguistics only chapter by offering some cognitive grounding to the issue of the Topic-Comment articulation.

All in all, this chapter will stand in the manuscript as bringing in both linguistic and psycho-lingusitic evidence for the centrality of the sentence-discourse interface in the articulation of the utterance.

Chapter 3, studying the locus of the discourse-interface to understand the sentence.

> Le dehors permet
> au dedans de se réfléchir.
>
> *[The outside allows*
> *the inside to be reflected.]*
>
> François Dagognet, Les Noms et les mots (2008:14).

By addressing the issue of the internal structure of the sentence across languages in chapter 2 (§2.2.4), different sentence articulations were presented. Specifically, Topic-Comment articulations were identified as carrying interesting discourse properties. This crucially brought our questioning about sentence structure to be enriched by taking into consideration the linguistic phenomena that witness of an interface between sentence-level and discourse-level.

Far from any hasty or simplistic pragmatic interpretation of these interfacial linguistic phenomena, what was actually identified is the *structural* possibility of establishing a link between the sentence-unit and discourse, which will be at the heart of our experimental research on the cerebral representation of syntactic complexity in chapters 5, 6 and 7.

2. In this section the description of Topic linguistic phenomena will be mainly using English examples.

As stated in the above epigraph by a French poet, in this chapter we will try to convince the reader that investigating sentence-discourse interfacial phenomena, like Topics, can reveal a lot about the structural organization of the sentence-unit.To illustrate this approach and strategy, we propose to use the architectonic metaphor illustrated in Figure 3.1.

One could metaphorically view the sentence as a cupola and syntactic-layer dedicated to sentence-discourse interface as the top-most architectonic element - the lantern. Interestingly, this last small cupola-like structure mounted on top of a dome, has different roles in architecture: (1) letting light enter to illuminate the cupola structure, (2) letting smoke escape, but most of all it has a third essential role, (3) that is a *static* and *structural* one, it statically holds the cupola standing. The lantern connects all the vertical ribs (i.e. stony structural elements maintaining the cupola, in blue) of the cupola by weighting on them at the point of their intersection.

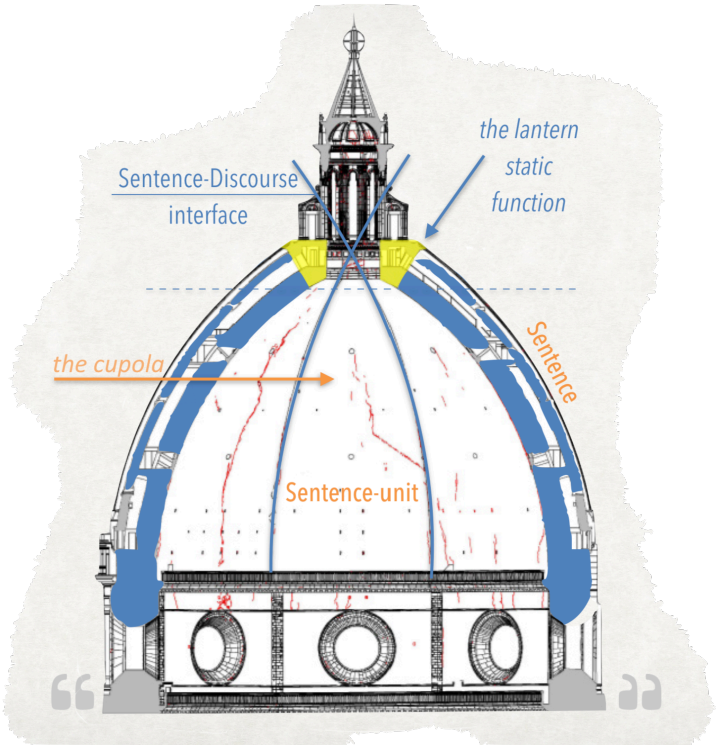


Figure 3.1 – Architectonic metaphorical representation of the sentence-discourse interface – *the lantern* – and its '*impact*' on sentence-internal articulation -the cupola. Lines and structural elements in blue represent the lines along which the static forces expand starting from the point of application of the lantern weight in yellow. Drawing of Santa Maria del Fiore (Florence) adapted from architectural drawing of Haines and Battista (2015).

As in this architectonical configuration[3], we will argue that the syntactic domain dedicated to the sentence-discourse interface (i.e. the CP) plays a comparably important role as the lantern for the structural architecture of the sentence.

Concretely, focusing on sentence-discourse interfacial phenomena (i.e. the 'lantern') will reveal the hidden balance of static and structural forces that make the sentence-cupola stand. Moreover, being a point where the different 'forces' inside the sentence can converge to, investigating this sentence domain will make emerge a number of hidden internal balance of forces, like the one linked to co-referential links in the sentence (ch. 7), syntactic movement (ch. 6 and 7), prosodic sentence-level patterns (ch. 4 and 5) and interrogative force (ch. 6).

This third chapter will go on to illustrate how the syntactic complexity parameter linked to the presence of a sentence layer dedicated to the sentence interface with discourse (cf. end of previous chapter, §2.4.4) are a rich testing ground to isolate and test the distinct complexity parameters that we presented in chapter 2: (a) syntactic hierarchy, (b) syntactic movement transformations, (c) presence of resumption versus

3. Note that the parallel with architecture will also have the advantage to stress the eminently syntactic dimension we choose to adopt to describe the interfacial linguistic phenomenon of Topic-Comment articulations.

syntactic-empty positions (i.e. gaps), (d) the cerebral representation of the sentence-discourse syntactic layer without other complexity parameters.

While chapter 2 briefly reviews the interest for investigating the cerebral bases of the sentential linguistic phenomena we selected as the object of our neuro-linguistic research (i.e syntactic movement transformations in French and Topic comment articulations in Chinese), this chapter will clarify why studying Topic-Comment articulation in contemporary Mandarin Chinese is important to investigate the cerebral representation of the sentence and to understand the determinants of activation of the sentence's cerebral network.

This chapter will provide a background discussion to set the foundation for the current neuro-linguistic research on Topic-Comment articulation in contemporary Mandarin Chinese, including and introduction to Topic-prominence typological claim.

An Initial intuition

Before entering the linguistic details of Chinese Mandarin Topic comment sentences, it might seem the right time to offer the reader as short explanation of why we decided to focus on Topic-comment sentence articulation in this pluri-disciplinary research project on the *sentence as a cognitive object*.

In a short temporal flash back, we might spend now a few words on the very first intuitions that brought us to start this research work several years a go. The original intuition that structuring the content of a sentence in Topic-Comment articulation - an extremely productive and frequent sentence structure in Mandarin Chinese- could coincide with a particular the mental construal (or representation) arrived actually quite early when I started to learn Mandarin Chinese.

This very embryonic idea came with the intuition that structuring an utterance according to this sentence pattern -the Topic and what is being said about it in the Comment clause- was not only giving a secondary role to the subject, but to the main verb too. Namely, the discovery of the possibility (and the necessity in Chinese) of this way of structuring the sentence-unit articulating the predication of a sentence by grammatically "*ex-centering*" the verb, was perceived as sentential Copernican Revolution in , where the verb and subject (earth and moon) are in a way ex-centered from the sentence system.

Namely, the interrogation that emerged at this very early stage of my reflection could be expressed in this metaphorical way: What does it means for the mind to structure the utterance *around* a Topic instead than a Subject or the Verb? In my mind, the sentence started "turning around" the Topic planet. Learning Chinese, implied also learning that the propositional meaning of the sentence was not uniquely constructed on the relationship between subject and verb, but could also be constructed by the interplay of a Topic and a comment in certain constructions, where the Topic cannot acquire a semantic role in the comment because it has no selectional relations with the verb of the sentence[4].

4. Note that this concept of a predication that is independent from the selectional relation of the verb and it theta-role assignment, is called *syntactic predication* in that it is opposed to theta-role assignment (i.e. the kind of predication distinguishing the relationship between the verb and its arguments). This term is inspired from Caroline Heycock's (1993) article on Japanese Topical phenomena and non-theta *ga*-phrase, where the author defines these phrase as a kind of *syntactic predication*, that she calls non-

Later, I realized that this was not only the case in Chinese, but sometimes in every day French. For this reason, and to help the non-Chinese-speaking reader I will systematically make reference to French and English in the following linguistic chapter. This will not only help the readers coming from different linguistic backgrounds to grasp what is at stake in our neuro-linguistic investigation, but it will also prove that the linguistic phenomena we investigate can bear on universal questions about the cerebral encoding of sentence structure.

Hence, this chapter 3 will retrace how digging in the linguistic literature transformed my original questioning about the particular sentence structure I later encountered in Chinese Linguistics into a wider interrogation about sentence's structure neural underpinnings. Namely, a close acquaintance with the literature helped formulating our initial interrogation in terms that could nourish the more general syntactic research question about what is a sentence for the brain and how its internal complexity is represented and processed in the brain. In this regard, a particularly noticeable fact is that since the very first formulations of the notion of Topic, its logic, psychological and cognitive dimensions have been immediately put forward (Section3.1, 235).

Importantly, this recasting of my initial questioning made me depart from the linearity of Topic-Comment structures to grasp the fundamental hierarchical relation that exist between the Topic and Comment. This step implied to embrace a layered representation of the sentence structure and of its interface with discourse to analyses Chinese Mandarin Topic-comment sentences. This led to precise experimental hypotheses that PartII will develop in three chapters 4, 5 and 7.

In conclusion, this manuscript will study the cerebral bases of Chinese sentence structure to investigate the internal hierarchical organization of Topic-Comment articulations, while the neural underpinnings of French sentence will be investigated to cast a light on the sentence complexity engendered by syntactic transformations linked to interrogation. Focusing on these two syntactic configurations though the lens of linguistic typology and of linguistic formal approaches will contribute to the understanding of the determinants of cerebral activation to syntactic complexity. The comparison however will only be done through linguistic theory and only in an indirect way testing localizers designs (in the Annexes H).

This first Part I will bring the reader to acknowledge that working on certain languages, like French and Chinese, that have the appropriate characteristics to isolate certain linguistics configurations and syntactic elements, is essential to understand how hierarchical syntactic structures and some of their features are represented and processed in the brain.

theta predication. Predication appears then to be independent of theta-role assignment, and "in that it was independent of the selectional relation of the verb and it theta-role assignment" (1993:188).

Overview of the contents of this chapter

Dans le passé, en effet, à travers plusieurs problèmes, nous avons valorisé l'extériorité [...] [or] l'enveloppe mérite d'être relevée et réfléchie, car elle enferme en elle de riches significations à découvrir, ce qu'elle doit d'ailleurs à son statut, celui de l'interfacialité ; elle se situe entre le dedans et le dehors, elle traduit l'un et occupe l'autre.

[In the past, in fact, across different issues, we valorized exteriority [...] but the envelop merits to be considered and thought about, because it contains rich significations to be discovered, which it acquires thanks to its status: inter-faciality. As it is situated between the inside and the outside, it translates the first one and occupies the second.]

François Dagognet, *Les noms et les mots*, ed. Les Belles lettres, coll. Encre marins (2008:13).

Topic is in the air

Far from being an isolated linguistic phenomenon, that would attract only the attention of '*sinotropic*' minds, Topic-Comment articulations can be easily found in common everyday linguistic practice, even in languages where this sentence articulation has not been formalized by grammarians.

As previously pinpointed in (§2.2.4, p. 118), Topic-comment structure is a universal phenomenon; this relationship is encoded with various formal linguistic devices in the grammar, namely, morphological markers, syntactic structures and intonation (Gundel, 1988:216).

Belonging to every day conversational use this articulation of the sentence-unit is stylistically considered as restricted to the oral register and sometimes considered as poor use of language especially in France. The oral and conversational character, makes it an interesting testing ground to deepen our understanding of the pervasiveness of the sentence-discourse interface in very natural linguistic contexts.

Consider for example a few French sentences making use Topic-Comment articulations in (59).

(59) Former French President's favorite syntactic construction

a. "La France, elle a des atouts..."
'France, it has assets...'

b. "Cette politique, elle coûte à la croissance"
'This political measures, they cost to the country growth.'

c. "Les résultats, ils tardent à venir"
'The results, they are late to come.'

Examples like (59) have been the 'fabric mark' of the former French President eloquence. The the last five years of French presidential governance have indeed, allowed an effortless enlarging of our sample of Topic-Comment sentences in French and provided us with simple examples that everybody knew to explain our research topic. When questioned about my research, one could often hear the following reply: "I investigate how the brain encodes the President's way of speaking, his favorite syntactic construction."

Topics is in the press

Syntax is in the air The French press dedicated several newspaper articles to this way of building sentences, and have raised an great amount of interrogation about the actual motivation of such a poor language use, to the point that several intellectuals and academic personalities have raised up against it. Famous academicians of the highest cultural institutions of the county have risen against this language use, saying that this kind of sentence articulations are good for children utterances[5].

Nonetheless, "the Topic-Comment perseverance" of the former French President, François Hollande, has to be acknowledged here for non-political linguistic reasons: it will have made the syntactic object (or topic) of this PhD a newspaper one[6]. For the enjoyment of all linguists in France, syntax and its sentence structures became for some weeks the focus of newspaper debates[7].

Considerations about to what extent anaphorical resumption could be viewed as a stylistic process, and yet how the redundancy linked to the doubling of the subject by a pronoun couldn't be considered as a proper way of forging sentences animated the usually linguistics-free newspaper's debates[8].

More linguistically 'enlightened' newspaper articles did carry out a detailed syntactic explanation of the sentence structure. We can read a linguistically driven analysis:

> **Un procédé nommé "dislocation":** *le redoublement d'un nom ou d'un pronom qui est détaché de la proposition s'appelle "dislocation". Si le nom*

5. Alain Finkielkraut, *Le Point* n° 2222: 51, 9 avril 2015 : « Difficile d'incarner la nation quand on pratique systématiquement le redoublement du sujet. "La France, elle a des atouts." : Cette syntaxe sied aux enfants, pas au chef de l'État ! » "It is hard to embody the French nation when you systematically utter such sentences as 'France, it has assets'. This kind of syntactic construction suits children, it does not suit the Head of State".

6. This was enough for hijack me of my intellectual *turris ebournea* and for making feel the research I was conducting on Topic-Comment articulations was highly relevant for how real people were using language, how real utterances were structured in every day life and the president's words of democratic decadence.

7. This fancy debates about syntax and language that the president's the linguistic habits raised -the so-called french "*engouement pour la langue*", "Enjoyment for language-oriented debates"-, reveal that syntax is not a kind of abstraction only intellectuals remark, it lies at the core of society. We underline the central position of language in society, can be stressed even without explicitly quoting the famous french social critic Pierre Bourdieu (1930 - 2002). See one of his cornerstone books *Language et pouvoir symbolique*, 2001. Syntax appears here as a real object, so prominent that people and cultures can project on it values, sentence structure in itself ironically can even bear a moral value as these newspaper articles and social debates demonstrate.

8. The same François Hollande used to take the upper hand over Nicolas Sarkozy during the TV debate of the 2012 presidential election campaign thanks to is anaphoric discourse beginning with "Moi, Président de la République", "I, President of the Republic".

ou le pronom est détaché à droite de la proposition, on parle de dislocation à droite. Exemple : « Elle est belle, la France ». Si le nom ou le pronom est détaché à gauche de la proposition, on parle de dislocation à gauche. Exemple : « La France, elle est belle ». Ironie du sort, notre président, issu du parti socialiste, pratique la dislocation à gauche ! [9]

Topics: an everyday mark of oral natural utterances

Although the journalist's attempt to syntactically explain these sentences should be prized Linguistic research went further in the definition of the former President's favorite declarative sentence structure. What has been defined as "le français ordinaire" – ordinary french by Gadet (1994), actually does not use this sentence syntactic configuration to obtain a pragmatic saliency effect that is typical of dislocation. On the contrary, early studies by Ashby (1982) and Lambrecht (1984) showed that these sentence structures are highly frequent: 70% of nominal subjects of oral French are doubled by a subject clitic that does not refer to a previously salient discourse-entity (Zibri-Hetz, 1994; Auger, 1995). For example, the famous french writer Marcel Proust (1871 - 1922) made use of this sentence articulation to picture the vividness of oral and popular discourse of his characters: *"Il faut que le bœuf, il devienne comme une éponge [...] les soufflés ils avaient bien de la crème"* (M. Proust, in *À la recherche du temps perdu.*, 1913)[10].

Interestingly, Culbertson and Legendre (2008) show through the prosodic analysis of double subject sentences what several linguistic studies had already put forward, namely that subject doubling shouldn't be analyzed as a dislocation (Blanche-Benveniste, 1997; Zibri-Hetz, 1994; Auger, 1995)The authors compared the prosodic patterns obtained for the three experimental condition in Figure 3.2 with the prosodic-acoustic signature for French Left-Dislocated Topics identified by Doetjes et al. (2002): (i) F0 increase on the dislocated constituent and (ii) lengthening of topic last-syllable[11]. Culbertson and Legendre's results show that doubled subjects are not separated by a pause from the clitic pronoun, and that the intonational pattern linked to last-syllable lengthening, attested for left-dislocation, is observed for object Topics as shown in Figure 3.2. We will resume to this kind of analyses to perform our phono-acoustic study of Chinese Topic-Comment construction in next chapter (ch. 4).

These findings and debates bring some evidence that the kind of sentence articulation under analysis is a natural and spontaneous linguistic stimulus although"good french use" name it as bad speaking, a denotation that is absent in Mandarin Chinese where these sentence articulations are paramount in both oral and written production.

All these examples of French are here to show that the sentence articulation that we selected for neuro-linguistic investigation carries the mark of language vividness in real

9. A process named 'dislocation': adding a noun or a pronoun outside the boundaries of the clause is called dislocation. If the noun or the pronoun is postponed, it is a right dislocation, for example: "It is a beautiful country, France". If the noun or the pronoun is advanced, it is a left dislocation, for example: "France, it is a beautiful country". Ironically, our president, coming from the "Left Wing" Socialist party, is keen on uttering left-dislocated sentences!.

10. It remains that for the commonly used reference grammar of French (e.g. Grevisse or Bescherelle) if this sentence articulation cannot be considered as ungrammatical, the advice is to moderate the use of it both in oral and in written style.

11. We will resume to the study by Doetjes et al. (2002) in chapter 4 when analyzing the prosodic pattern of a particular type of Chinese topic-comment structure (i.e. Scene-setting Topic).

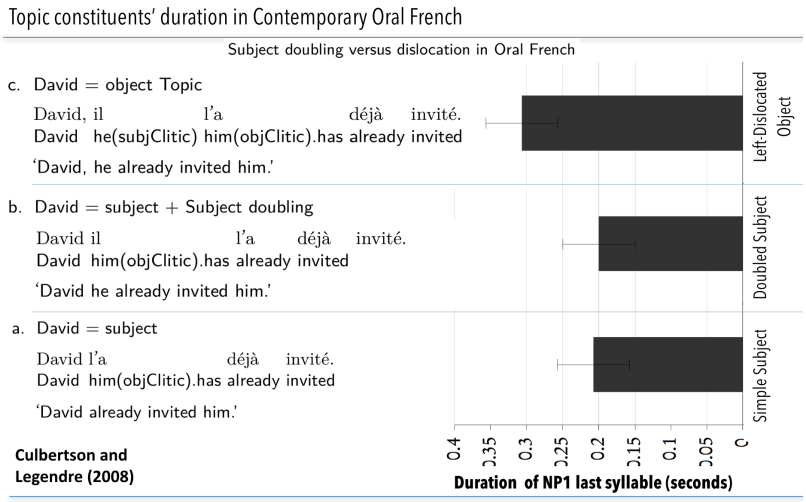


Figure 3.2 – A significant effect of the sentence is found for the the length of last syllable ($F(2,57)=17.4$, $p<0.001$). The average length of of dislocated objects was longer compared to simple subjects and double subject constructions ($F(1,57)=35.5$, $p<0.01$), but no significant difference was found between simple subjects and double subjects ($F(1,57)=0.14$, $p=0.71$). Hence, the expected last syllable lengthening for Left-dislocated elements was observed only for objects and not for Doubles subjects. This phono-acoustic property gives additional evidence to the linguistic analyses arguing that Double subject are base-generated. Adapted from Culbertson and Legendre (2008).

people, real situations and real interactions[12]. Hence, Chinese will serve as an excellent testing ground for studying these distinct linguistic processes, without the controversial aspect of register that is present in French.

A recurrent criticism towards linguistics, is that the field relies on intuitive grammaticality judgments, based on meta-cognitive knowledge of the informer which might not be reliable. Although, we haven't undertaken large-scale judgment surveys, the fact of creating several hundreds of stimuli (see Annexes C, p. 823.) corresponding to certain experimental conditions does not actually correspond to the typical '*armchair linguist*' activity: we went through systematical test for naturality with half a dozen native informants, whose judgments discriminated the most natural sentences through our large experimental corpora. This process has been rather informative, in that I happened to test the very same sentence structure in approximately a hundred different sentences which revealed some interesting linguistic aspects we will address in chapter 4.

Hence, braving the general criticisms that linguistics is undergoing lately[13], Topic-

12. It goes without saying that one of the most important preoccupation for a linguist doing cognitive neuro-imaging of language should be to select vivid and natural linguistic stimuli. Experimental neurolinguistics has often recourse to ungrammatical stimuli as baselines to the linguistic phenomenon under analysis. It wont be the case in this dissertation (Except for the ERP study on prosody for external reasons to my will). The main reason for this is that I find the relation between acceptability judgments and linguistic behavior very difficult establish clearly. Moreover, the graded character of grammaticality is difficult to explain psycho-linguistically.

13. For an overview on these debates see Phillips (2008) "Should we impeach Armchair linguistics?" in *Japanese/Korean linguistics* n°17.

Comment linguistic phenomena has the advantage of being one of the outstanding fruit of a linguistic tradition that grew in the transition zone from the so-called armchair-linguistics to field linguistics. To add a further step to this "growth", this doctoral dissertation submits the typological-claim linked to this type of sentence articulation under the observation lens of experimental neuro-syntax.

3.1 The notion of Topic and its sentence's articulation

> *Languages may not be loved platonically; one must live in and with them before one may venture to form opinions about them.*
>
> Georg von der Gabelentz (1840 - 1893)

3.1.1 The notion of topichood

Focusing on the notion of Topic, we will first show how the Topic-Comment articulation of the sentence-unit is an essential and acknowledged feature of human language, by characterizing its predicational mechanisms (§3.1.2), function (§3.1.4) and syntactic articulation (§3.1.5).

As the structuring of utterances into a Topic part and a Comment part is a pervasive phenomenon across languages, it is worthwhile to read the initial attempts to conceptually define this fundamental way of articulating the sentence by scholars and linguists over the last centuries. The reason for reporting here an notional outline is that it has proven essential to elaborate our experimental approach to these linguistic phenomena. Namely, this crucial step of notional clarification, the linguistic overview on Topichood (§3.2) and the syntactic articulation that characterizes it (§3.4) will offer some central elements to approach this concept from the point of view of Cognitive Psychology (§3.4.2.2), and to address its syntactic realization in Mandarin Chinese (§3.4.6).

By presenting the birth of the notion of topichood, we will emphasize the psycholinguistic and cognitive relevance of such utterance articulation (§3.2.2.3).

A panorama of definition at different levels of analysis

If the notion of topic dates back to Ancient Greece, when the concept of logical subject was first formulated at the beginning of classical logic, its very first modern formulation of the notion of Topic has been taken over with the term *theme* by Charles Bally in 1932.

Such an insight initiated in the heyday of functional linguistic tradition a series of linguistic studies from the School of Prague (see Sgall, Hajiová and Panevová 1986 for an overview). Since then part of the work done by European functionalists was continued by north-American functionalism, and Hockett (1958) introduced the term Topic and deepened the exploration of the structure of the utterance into a Topic and a Comment

part. This led the the Topic or Theme to received a number of successive definitions: as the "starting point of the message conveyed by the utterance" (Halliday, 1967), as the "reference frame where the predication is realized" (Chafe, 1976), or as "what the proposition is about", or simply as "what is talked about" (cf. Kuno, 1972; Dik, 1989; Reinhart, 1982; Gundel, 1988; Lambrecht, 1994).

Despite the considerable quantity of studies on the notion of Topic, there exist an great variety of diverging definitions of what a Topic is, moreover what the level of analysis that shall be chosen to understand topical phenomenon is still a matter of debate. In our view, this *definitory* issue is less a matter of confusion around the concept of Topic than a matter of stratification of the level of analysis.

There exist many ways of defining the nature of the Topic, but globally there is consensus in saying that, at the level of the sentence, it is a sentence segment about which the concern of the remaining part of the sentence is.

In the linguistic literature, the term Topic is accompanied by notion associated to the idea of "remaining part" of the sentence, or more precisely to the idea of sentence's predicative bi-partition, in a kind of internal dialogue crystallized in this sentence articulation between the Topic about which the comment is. The different notions illustrated by Theme-Rheme, link-tail, or psychological subject terminology, namely reflect that at least the dichotomy between Topic and Comment is uncontroversial.

Hence, one of the first question to be addressed in approaching this sentence articulation is that of the choice of level of analysis of these linguistic phenomena[14].

Although our focus is doubtlessly syntactic, we will address in this rich panorama of definitions of the notion Topic a series *satellite concepts* that often accompany it, to delineate our cognitive and experimental approach of the sentence-unit, of its internal structural organization and its specificity in Mandarin Chinese.

As a all-in-one definition of Topic-Comment is hard to come by, and instead of considering a single feature, we will direct our attention to several factors

In the main aim is to delineate the linguistic notions and concepts at stake, that will theoretically ground our neuro-linguistic experimental approach. The notional overview offered by this section will be organized thematically around four main axes pertaining to our cognitive approach:

1. Logic analysis of the sentence-unit predication (§3.1.2);
2. Introducing the speaker's and hearer's mind in analysis of the linguistic utterance (§3.1.4);
3. Information structure the sentence discourse interface the relation between the topic referent and the context (§3.1.4);
4. Syntactically Tailoring sentence according to informational packaging concentrating on the sytnacitc encoding of sentence-discourse intefacial phenomena (§3.1.5).

We will present here past contributions to the establishment of this notion in the light of a tripartite delineation distinguishing three main level of analysis: the semantic, the pragmatic and most of all the syntactic analysis.

While sections §3.2and §3.4 will be entirely dedicated to the two main syntactic approaches we pursue in this research, namely the typological and formal approach, our notional overview will integrate both semantic and pragmatic approaches when their analyses will shed light on our research issues.

14. As for terminological choices, we decided to avoid the term Theme because in the Anglo-Saxon literature this term is often associated to Discourse-oriented analyses.

The notion of Topic has long been suffering of a foggy *definitory panorama* and this section is a tentative to unravel these different levels of analysis surrounding Topic-Comment linguistic phenomena, and to situate it back into the realm of our fundamentally syntactic approach. Hence, exposing some cross-linguistic considerations on the propositional/predicative articulation of the sentence will continue fueling our central discussion on sentence as a cognitive object.

During this notional exploration of Topic-hood, we will focus on general examples to illustrate topical linguistic phenomena and the different levels of linguistic analysis they are linked to.

Being interested in the mental representations of sentence-unit, we will review in this section the most significant contributions that helped us to understand what comes into play when listening or producing an utterance according to Topic-Comment articulation, and more particularly, those that are highlighting the different dimensions of topical linguistic phenomenon that are likely to be relevant for our psycho-linguistics approach, in order to derive precise experimental hypothesis for both ERP and fMRI studies presented in Part II and chapters 5 and 7.

3.1.2 Logicians, Grammarians and Sinologists' Topic

The genesis of the term

The notion of Topic deepens its roots in the first steps of classical logic theory, where it was first defined as the logical subject of an utterance.

The term Topic was surely borrowed from the lexicon of Aristotelian rhetoric, in which Ancient Greek etymology clearly states a reference to "what concerns the place", in spatial terms "something established in one place". It is only starting from 1372, that this term has been employed in the sense that we know today.

However, one of the first putative description of Topic-comment sentence articulation could be identified in the meta-language used by medieval Arabic grammarians descriptive, who were distinguishing between '*mubtada*' (lit. *beginning*) and 'xabar' (lit. *news*) in the canonical sentence articulation. These two sentence elements were crucially differing from grammatical subject and grammatical predicate, in fact Arab grammarians refer to the initial and left-most nominal constituent also as Aas al-muHadda , literally '*the person or thing being talked about*'[15]. as shown in the following example[16]:

(60) al-waladu, Hindun darabathu
def.boy.NOM., Hind.NOM. hit.3f.SG.ACC.
'(As for) the boy, Hind hit him.' (Hussein Abdul-Raof, 1998)

3.1.2.1 Philosophical account of the sentence-unit articulation: the Logic of the proposition

It is a matter of facts that philosophers like Hobbes and Leibniz being involved in a logic theorizing of language predication (i.e. propositional logic) it allows, went beyond the

15. Goldenberg, G. (1988), "Subject and predicate in Arab grammatical tradition", Zeitschrift der deutschen morgenländischen Gesellschaft 138, 39-73.

16. Abdul-Roaf, Hussein (1998) Subject, Theme and Agent in Modern Standard Arabic. London: Routledge Curzon.

grammar categories of subject and predicate Hobbes used: *continens* for the predicate and *contentum* for the subject (and the reverse for Leibniz). Later on logico-philosophical investigation of the proposition introduced a more detailed account of the logical form a *statement* can have.

For instance, a philosopher like Clemens Brentano[17] contributed to logic by developing a theory of *Judgments* laying the basis of what he defined as "*the science of mental phenomena*" (in Brentano's *Psychology*, 19874:18 – English ed. 1924), what will be later known under the name of 'Brentano's Theory of Mind'[18]. Directly asking the question of what happens in our mind when we make *judgments*, he delineated the following claims:

1. Judgments require that some object is given in *Presentation*, but not necessarily that something is predicated of it.
2. Judgments are most essentially expressed in sentences of the form "A exists/does not exist", where the term 'A' denotes the presented object which is also the object of the judgment, and the rest of the sentence indicates its quality.

Given this judgments' typology, we can say that his theory of judgments concerns the relation between judgment and predication, determining a canonical form which all judgments can be expressed. It holds that statements are either simple [*Presentations*], i.e. the most basic kind of act happening each time man is directed towards an object, or alternatively being a [*Presentation + another judgment*] (this will be called a Double judgment, Doppel Urteil)[19].

Hence, the act of predication is then made of two steps: (1) fist accepting the the existence of something, and then (2) adding a second judgment saying either that the presented object of the predication has or lacks a property. As noted by Huemer (2015), this definition of predication can be seen as close to the fregean account distinguishing between '*grasping a proposition*' and then '*judging it to be true*'. "*In a judgment we accept or deny the existence of the presented object. A judgment, thus, is a presentation plus a qualitative mode of acceptance or denial.*" from Psychology of Brentano (1924).

Importantly, this bipartition of predication shows that some judgments can have an non-predicational content as shown by subjectless sentences, which instead of being expressed according to the subject-predicate relationship uttering "**The weather is rainy**" or "*The moon is lacking water*", directly express a judgment by specifying an object which is given in presentation, 'rain' or 'water on the moon') and by indicating whether this object is accepted or rejected as in (61).

(61) *subjectless* sentences in Brentano's and Marty's philosophy of language

17. Franz Clemens Brentano (1838–1917) is well-known for his contribution in philosophy of psychology. One could say that he is one of the latest intellectual delicatessen of the Austro-Hungaric empire and its "Mittel-Europa" Kultur. He made important contributions in many fields in philosophy, we will consider here his contribution to logic in his Theory of Judgments. It has to be noted here that Brentano is often considered one of the forerunners of the phenomenological movement and of analytic philosophy. His charismatic teaching exerted a strong influence on the later work of Edmund Husserl and Anton Marty, who will contribute to the logic investigation of Topic-Comment articulations.

18. One of his overtly declared goals was namely to lay the bases for a scientific psychology.

19. In his theory of mind Brentano characterized mental phenomena according to the three ways in which a man can be directed towards an object, namly Presentation, judgments and phenomena of love and hate. Starting from *Presentations*, the most basic kind of act happening each time man is directed towards an object, judgments can happen. Namely, "*In a judgment we accept or deny the existence of the presented object. A judgment, thus, is a presentation plus a qualitative mode of acceptance or denial*".

a. It is raining.

b. There is no water on the moon.

According to this refined view of predication, predication is made not by combining two ideas or presentations, but by combining two judgments, this point will be developed by the logician and philosopher Anton Marty (1884–1895).

Marty: Thetic and categorical judgment

Short after these two initial distinctions, Anton Marty defended the Brentanian theory of judgments in the linguistic domain. In 1884, he analyzed whether all sentences could be structured distinguishing the two types of Brentanian "judgments", and thus formulated the notions thetic judgment and categorical judgment that we still u.

Specifically, Marty distinguished *categorical* sentences, for which a "categorical judgment" can be established, from *thetic* sentences, that do not have a constituent that first identifies a 'psychological subject'. Hence, leaning back on the theory of psychology and of the mind of his teacher, Clemens Brentano, Marty[20] described the *thetic* proposition as implying a *'unique judgment'* - the state of things denoted by the proposition is presented in one and only piece, instead of being 'double' and identifying in a first place an object, a logical proposition or a state of things and then in a second time something is predicated about it[21]. Given the definition of the Topic-Comment articulation we already sketched, this double-judgment configuration is easy to put in correspondence with the bi-parted articulation of the statement found in Topic-Comment.

Marty's conception of Statements (*Aussagen*) took as a starting point the problem posed by statements which apparently have no subject, so-called "impersonals" (*Impersonalien*). This issue, already addressed by Brentano, was reinterpreted in the linguistic domain by his work.

In this way, Marty analyzed subject-less sentences investigating their relationship to logic and psychology. Thereby, he defined them as simply expressing the acceptance of something. For instance, an impersonal sentence like "It is raining" was understood as expressing the judgment that there exists 'an event of raining', as perceived by the speaker, without needing to theorize an entity to which the dummy subject 'It' would actually refer.

Hence, in Marty's view thetic judgment are better expressed in the *existential* form, rather than the *predicative* one, while double judgments are genuinely categorical, in that they cannot be reformulated without the bipartition given by the Double judgment constituting categorical sentence and featured by the subject-predicate or Topic-Comment relation. From this follows that the peculiarity of 'double judgments' resides in the fact that the subject or topic term in the sentence like "*This cat is purring.*", "this cat",

20. Anton Marty (1847- 1914) was a philosopher of language (ontology) and psychologist. His work is generally understood as an application of Brentano's psychology to the study of language clearly departing form the mainstream currents in the linguistics and philosophy of language of his time, that were mainly focusing on historical Linguistics and phonological or semantic diachrony. His philosophy of language is viewed as having a prominently psychological character, as a reflection on linguistic phenomena as arising from individual human minds and intentionally directed to objects. For further presentation of Brentano's philosophy see Huemer, Wolfgang, "Franz Brentano", The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.), Fall 2015 Edition.

21. We should however not here, that according to him even 'thetic' sentences may have a psychological subject, that is just not realized as part of the utterance, because it is given in the situation of utterance.

already expresses an instance of acceptance and "is purring" represents a predication built upon this acceptance.

It is noteworthy that the kind of examples that were given for *subjektslose* 'subjectless' sentences[22] are either *dummy subject sentences* or featuring unaccustive verbs. Exactly, these two types of linguistic phenomena will be further taken into consideration in the second section of this chapter (§3.2), when presenting the arguments for the typological analysis of sentence structure according to Topic-prominence parameter.

Linguistic loan to philosophy: thetic/categorical judgments

In linguistics, this pair and the propositional logic it conveys were made popular by Kuroda (1972) and Sasse (1987). Kuroda (1972), namely adapted the concepts from the philosophers Brentano and Mary to the two fundamental predicative relationships in the sentence, first to subject-predicate relation and later to Topic-Comment articulation.

> "This theory assumes, unlike either traditional or modern logic, that there are two different fundamental types of judgments, the *categorical* and the *thetic*. Of these, only the former conforms to the traditional paradigm of subject-predicate, while the latter represents simply the recognition or rejection of material of a judgment. Moreover, the categorical judgment is assumed to consist of two separate acts, one the act of recognition of that which is to be made the subject, and the other, the act of affirming or denying what is expressed by the predicate about the subject. With this analysis in mind, the *thetic* and the *categorical* judgments are also called the simple and the double judgments [in german *Einfaches Urteil* and *Doppelurteil*]." Kuroda (1972:154)

The following examples (62) illustrate in (a) categorical double-judgment, where the first act is to recognize the subject/topic neko '*a cat*', and the other act to affirm (or deny) what is expressed by the predicate/comment about the cat '*sleeping there*'; and in (b) thetic judgment involving as single judgment: an act expressing the recognition of the existence of a specific entity or situation '*a sleeping cat*'[23].

Note that in drawing this analysis in Japanese, Kuroda's most compelling argument is the presence in this language of two distinct morphemes to mark this logic distinction: the Topic marker -wa and the subject marker -ga.

(62) Categorical and Thetic morphosyntactic marking in Japanese following Kuroda

a. Thetic

neko *ga* asoko de nemutte iru
cat SUBJ$_{GA}$ there at sleeping is

'The/A cat is sleeping there.'

b. Categorical

neko wa asoko de nemutte iru
cat TOP$_{WA}$ there at sleeping is

'The cat is sleeping there.'

22. Marty, Anton (1884) "Über subjektslose Sätze und das Verhältnis der Grammatik zu Logik und Psychologie", Vierteljahresschrift für wissenschaftliche Philosophie.

23. Note that the bare noun marked by *-wa* cannot be indefinite nonspecific, which follows from the presuppositional nature of the subject.

Importantly, the linguistic encoding of thetic and categorical judgments is linked to Topic-prominence parameter in the framework developed by Kiss (1995). Namaley, according to here classification of languages along the notion of Discourse-configurationality, a language that would present categorical an thetic propositions in different linguistic forms would be a Topic-prominent language.

(63) Categorical and Thetic word-order marking in Chinese following Xu

a. Thetic: 屋里有一条狗。

wū-lǐ yǒu yī-tiáo gǒu
Room-inside YOU$_{there.is}$ one-CL. dog

'There is a dog in the room.'

b. Categorical SVO: 这条狗在啃骨头。

zhè-tiáo gǒu zài kěn gútou
This-CL. dog PROG. chew bone

'This dog is chewing a bone.'

As noted by Xu (2007), Chinese meets Kiss's definition of Topic-prominence, and uses SVO word-order or Topic-comment to for categorical propositions, but not for thetic propositions, as illustrated by examples in (63)[24].

In conclusion, it is interesting to observe that since the very first attempts to define the Topic notion, even the philosophical reflection about the logic of predication and statement in different epochs already formulated its psychological or even intuitively its cognitive dimension.

The fact that the roots of the very first sketch of the Topic-Comment notion are to be found in the realm of logic and philosophy is, in a way, one of the first pieces of evidence we want put forward to prove that the intimate relationship between predication itself (i.e. the possibility to say or predicate something about an entity) and syntactic sentential encoding is at stake in the sentence's articulation we selected for our neuro-linguistic investigation of the internal organization of the sentence-unit.

3.1.2.2 Topic-comment: a Predicative construction

The fist steps of a linguistic account

After these early investigation of how a proposition is logically articulated, we can retrace the first steps of its the linguistic account and better understand Hockett's statement on taking Topic-Comment articulation as the most general predicative construction (§2.2.4, p. 118), that we report here under:

> "The most general characterization of ***predicative constructions*** is suggested by the terms "Topic" and "Comment" [...]: The speaker announces a topic and then says something about it."

24. While this first feature make Chinese belong to Type A discourse-configurational languages. The fact Mandarin also has a specific position for Focus classifies it as a Type B discourse-configurational language too (as Hungarian).

Figure 3.3 – The first generation of Linguists exploring the linguistic realization of the notion of Topic.

This being said, the first linguist having suggested the difference between subject and topic was probably Henri Weil in 1844[25]. We can read in his chapter *Du pincipe de l'ordre des mots*:

> There is then a point of departure, an initial notion which is equally present the one who speaks and the one who hears, which forms, as it were, the ground upon which the two minds meet; and another part of discourse which forms the statement properly called *Enunciation* [French: *l'énonciation*]. This division is found in almost all we say. Henry Weil (1844:29)

He probably was the one who inspired the notion of ***Theme***, later introduced by Charles Bally[26] in 1932. Although the notion of *Theme* had a fundamentally intuitive character, Bally will also carry out a broader analysis of language and discourse, he will choose for this a perspective, that we will later call a pragmatic one. Bally's work is often presented as one of the forefathers a "psychological linguistics"[27].

After Henri Weil, **Georg von der Gabelentz**(1840-1893) will introduce the distinction between psychological subject (i.e. the Topic) and the psychological object (i.e. the focus). What is rarely said about this author, is that he was not only a well established linguist and philologist - a common conjunction at that time - but he was also a sinologist, teaching oriental languages and general linguistics in Berlin.[28]

At the end of XIXth century Vilém Mathesius will bring again this psychological aspect of the definition explicitly forward by opposing in the sentence a psychological subject to a psychological predicate, before the other exponents of the Prague School will start structure the notion into a more complete framework for sentence analysis.

25. Weil, Henri, chapter 1 : "Du pincipe de l'ordre des mots", in "De l'ordre des mots dans les Langues Anciennes Comparées aux Langues Modernes", Paris, 1844, 3e ed. 1879. p.11-30.

26. Charles Bally (1865-1947) was one of the students of Ferdinand de Saussure in Geneva. He and Albert Sechehaye, were co-editors of the famous "*Cours de linguistique générale*", published in 1916, three years after the death of F. de Saussure.

27. Charles Bally in his *Linguistique générale et linguistique française* (1932) transfers the distinction between theme and rheme to an other level of analysis, that of the 'enunciation' an no more that of the internal structuring of the utterance inside the discourse. He analyzed the enunciation into two parts a *modus*, where a modal subject asserts 'I think/I belive that...', and a *dictum* or propositional content. In this way, he states that "the *modus* is the theme and the *dictum* is the comment of the explicit enunciation". Quoting it in French: *"nous dirons donc que le modus est le thème, et le dictum le propos de l'énonciation explicite"*.

28. Author of a grammar of Chinese language (i.e. Anfangsgründe der chinesischen Grammatik mit Uebungsstücken, T.O Weigel: Leipzig, 1883, 152 p.) and second son of the linguist and politician Hans Conon von der Gabelentz (1807-1874) - it was the good old times where linguists used to rule the world. In general linguistics , he also formulated some typological and evolutionary principles going back to von Humboldt, and inaugurated the study of what is called grammaticalization.

Table 3.1 – Topic notion.

Author	Date	Definition	Definition and quote by the author
Weil	1844	*le point de depart de l'énonciation*	There is then a point of departure, an initial notion which is equally present the one who speaks and the one who hears, which forms, as it were, the ground upon which the two intelligences meet; and another part of discourse which forms the statement (in french : l'énonciation) properly called. This division is found in almost all we say. (Weil 1844, p.29)
Gabelentz	1869	*psychologisches Subjekt*	Evidently I first mention that which animates my thinking, that which I am thinking about, my psychological subject, and then that what I am thinking about it, my psychological predicate. (von der Gabelentz 1869, 370f , author's translation)
Paul	1880	*psychologisches Prädikat*	The psychological subject is [...] that which the speaker wants the hearer to think about, to which he wants to direct his attention, the psychological predicate that what he should think about it. (Paul 1880, author's translation).
Marty	1884	*thetic vs. categorical judgment*	The psychological subject is not expressed in the sentence *'Es brennt'*, 'there's fire'. But it would be wrong to believe that there is none. In this case we find a combination of two ideas as well. On the one hand there is the realization of a concrete phenomenon, and on the other the notion of burning and fire which already rests in the soul and under which the phenomenon can be subsumed. (Marty, 1884:§91, author's translation).
Mathesius	1915/1927	*theme of the enunciation*	"the grammatical subject, that of the doer of the action expressed by the predicative verb and that of the theme of the enunciation contained in the predicate."

3.1.2.3 Prague school: Topic as the psychological starting point

The fundamental contribution of the Prague School to this issue was the analysis of sentence-unit in terms of Theme and Rheme it initiated. This direction was taken to account for the mechanism attributing pragmatic functions to linguistic elements in the sentence.The main concern of this linguistic tradition was namely to analyze the function of linguistic elements inside the utterance, by this kind of analysis Linguists from the Prague circle observed that word order in language was also affected by factors like discourse context ans speaker's intention.

Linguists like Mathesius, František Daneš, Jan Firbas, Petr Sgall and Eva Hajičová stated that the *informational structure* of the sentence had to be considered as bipartite, where the Theme would be the starting point of the utterance. Mathesius (1927:61) used the term theme to distinguish between two different roles in the sentence:

> "the grammatical subject, that of the doer of the action expressed by the predicative verb and that of the Theme of the enunciation contained in the predicate."

After Daneš (1970) equations of topic=given and comment=new, later work from the Prague School (Mathesius, 1939), defined the Theme as has showing the following characteristics:

- Announcing what the sentence or the predication is about ;
- Carrying *mainly given* information, while in the rest of the sentence -the Rheme- new information is usually found;
- Occurring in sentence-initial position.

Not only this characterization recalls the logico-philosophical account of the sentence-unit articulation addressed in previous section, but these notions survived the Prague School period and have been pervasive in subsequent linguistic analysis. For example, we can trace back the first one in Dik's analysis (1997) of sentence-initial position (P1). He claims that P1 universally has the special purpose to host the placement of constituents with Topic or Focus function. Although all languages are claimed to use P1 position for special purposes, not all languages necessarily do it in the same way. Thus, in general terms, the pattern for the use of PI in SVO languages would be the following: [P1 + S + V + O] (Dik 1997: 408-409).

Another direction of linguistic analysis where the Prague school's Theme-Rheme distinction have been instrumental is that investigating the focal and contrastive interpretation of sentence elements, where not only new information is provided, but also salient or contrastive interpretation is required on these sentential elements.

In conclusion, the Prague School inaugurate a informationally- and psychologically-driven approach that was only latent in the previous definitions. Ever since this definition, the view that word-order in languages is not only determined by purely syntactic constraints, but also affected by other factors such as discourse-context and speaker's communicative intention will be prevailing and further developments in very different linguistic framework will prolong it life as witnessed by next section §3.1.4.

3.1.2.4 Agrammatic production and the structure of the utterance

Exploring the conversational behavior of speakers with agrammatic aphasia Topic-Comment constructions have been often advocated to explain their linguistic behavior. We will sketch a few examples both from the so-called Conversation Analysis (Wilkinson, 1995) used in aphasiology and from an experiment comparing cross-linguistic aphasic utterances.

Conversation analytic approach to Aphasic speech production

Several studies in aphasiology have applied the a qualitative methodology advocated by Conversation Analysis to the everyday conversation and utterances produced by speakers with agrammatism[29].

A study by Beeke, Wilkinson and Maxim (2007b) analyzed the conversational grammar in agrammatic patients and found that it essentially consists of utterances built out of elements like nouns, adjectives, conjunctions, temporal elements, reported speech or mime that the agrammatic speaker combines is a systematical way, such that the order of elements is related to the *interactional* function of the utterance.

Consider the following example from Connie, Roger and Roy, who both had Left hemisphere cerebro-vascular accident (CVA)[30]:

(64) Agrammatic utterances and their structures
   a. The structure of: [Bare temporal element + topic + comment]
      i. 'June, three tier wedding cake, I make it.' (Connie's utterance in Beeke et al. (2003b)
      ii. 'July, no June, three tier wedding cake, I make it.' (Connie's utterance)
   b. The structure of: [adjective + 'because' + reason/evaluation/opinion]
      i. ''amazing, because, two years or three years.' (Roy's utterance)
      ii. 'interesting actually, because me, I think no, special. honestly.'(Roy's utterance)
   c. The structure of: [talk + mime/enactment = event] (without verb)

29. Note that everyday sentences exchanged with a family member cannot but differ significantly from utterances produced in response to out-of-context assessments protocols or therapy tasks (Beeke et al., 2003a), and this aspect makes Conversational analysis a valuable tool to address the syntactic competences of aphasics and agammatic patients.

30. Connie had these Conversation with Sam: 2 years after her stroke and the Conversation with Jane 3 years and 11 months after. Roy had these Conversation with Di and Sally 7 years after his stroke. Roger had these Conversation 4 years 1 month after.

i. yes but me, ((points to self)), ((mimes hiding face in embarrassment while smiling)) Roger

ii. plate,((enacts drying a plate))and day dreaming. and man, ((enacts puzzlement)) Roy telling Cinderella's story

Utterances constructed as (64a.i) are Topic-Comment structure, and patients recurrently use the topic function to establish a new object, like 'June' and 'three tier wedding cake', about which predicates a message like 'being the maker of' in the above example: *'June, three tier wedding cake, I make it.'* (Beeke, Wilkinson and Maxim, 2003b).

Another frequent utterance pattern in speakers with agrammatism is illustrated in (b.ii) where an adjective expressed a result that is commented by [adjective + 'because' + reason] (Beeke et al., 2010).

Moreover, the grammatical practice of person reference plus action displaying by means of *kinesic enactment* as in (64c.i and c.ii) without reporting any verbal element, is common in the data-set explored by Beeke, Wilkinson and Maxim (2003b) and generally reported in agrammmatic speakers.

In this regard, Goodglass (1976:259) noted that "In some instances these patients will simply name the speaker and follow with a direct quote, as in the text of a play."[31] Hence, the sentence constructions in (c.i) and (c.ii) can be seen as typical topic-comment structures, with an object reference as the "Topic" and the enactment as the "comment." Note that this analysis of aphasics utterances is not only ours, but Topic-comment structure are described as a feature of both fluent aphasic talk (Wilkinson et al., 2003) and non-fluent aphasic talk in agrammatism (see Beeke et al., 2007; Saffran et al., 1989).

This analysis of agrammatic utterances echoes Hockett's statement on Topic-comment articulation as the most general predicative construction (§2.2.4, p. 118): "The most general characterization of ***predicative constructions*** is suggested by the terms "Topic" and "Comment" [...]: The speaker announces a topic and then says something about it."

The reason for this linguistic behavior has been found in the light of these speakers' difficulties to produce grammatical relations in side the sentence in an ordinary way. As an alternative the looser aboutness relation between topic and comment allows them to avoid for example verb agreement complexity. Hence, the ordering principles that are observed in their utterances seems to rely much more on the sentence-discourse interface and its pragmatic articulation than on subject-predicate grammatical relation. This allows the hearer to build a model of the sentence where all the pieces of information given after the first utterance element are to be linked to it as comments:

(65) 'June, three tier wedding cake, I make it.'

This evidence from aphasiology for the basic and general character of Topic-Comment predicative construction introduces us to the next section where we will further develop the notion of Topic by integrating the the speaker's and hearer's mind in the definition of this the sentence-unit's articulation.

31. Wilkinson, Beeke and Maxim. Formulating Actions and Events With Limited Linguistic Resources: Enactment and Iconicity in Agrammatic Aphasic Talk, 2010

3.1.3 The speaker/hearer's minds and the sentence

> *"The initial constituent of a clause serves as 'the point of departure' for the clause as message"*
>
> Halliday, M. A. K. (1967:212)

3.1.3.1 Psychological subject, Grammatical subject and Logical subject in the mind

In the 1960s, the work of Michael Halliday (cf. Halliday, 1985/1994 in *Introduction to Functional Grammar.*) will carry on the tradition of the functional paradigm, by defining the Theme-Rheme couple as **the way a message is conveyed by the clause articulation** : the Theme plays the role the starting point of the message from which the speaker proceed; and the Rheme, the remainder of the message, is conveying the major point[32].

This definition introducing the *speaker's mind* and his formulation of the utterance (having a starting point from which to proceed) accomplishes a major shift in considering the linguistic utterance as the locus of calculating the speaker's mental representation of the sentence's referent and to then structure the sentence-unit starting from it.

Halliday offers an analysis of the sentence elements at the traditional subject position, showing that they can have three different functions:

1. *Psychological subject* "the concern of the message" or the Theme according to functional analysis.
2. *Grammatical subject*, a syntactic concept having a number of grammatical features such as case and its concord in person and number with the verb
3. *Logical subject*, "the doer of the action", carrying the semantic role of the "Agent".

The identification of these three main roles (situated at different levels of analysis) allows the author to put forward an important observation on clause initial-position, by moderating its cardinality in the definition of Topic-hood. He notes, in fact, that the clause-initial element often overlaps with the subject of the clause and with the actor of the action described by the proposition (see examples in Table 3.2). Based on this observation he concludes that the positional argument does not constitute a unique and appropriate criteria for the definition of the Theme, only certain languages would uniquely use sentence-initial position to realize and host the Theme.

Considering the sentences in Table (3.2), we can note that the syntactic encoding (i.e. SVO, Passive and Topic-Comment) of the sentence attributes a different partition of the three fundamental roles identified by Halliday: In the SVO structure, "Betty" accumulates all the roles; in the passive sentence, 'The onions' play the role of starting point of the message and of grammatical subject agreeing with the verb, while 'Betty' is only the logical actor of the action described by the verb. And in the Topic-Comment construction 'The onions' remain the starting point of the message but as the agreement information with the verb is carried by 'Betty', Betty becomes both the logical and grammatical subject and a third partition of the three sentential roles is shown.

32. In a similar way Antoine Culioli will call the Topic a "*un repère constitutif*", a landmark reference.

Table 3.2 – Cognitive considerations about a noun in a sentence.

DECLARATIVE SVO SENTENCE	**Betty**	peeled	*the onions*
	Grammatical subject Logical subject Psychological subject		
PASSIVE SENTENCE	**The onions**	were peeled	*by Betty*
	Grammatical subject Psychological subject		Logical subject
TOPIC-COMMENT SENTENCE	**The onions,**	*Betty*	peeled
	Psychological subject	Logical subject Grammatical subject	

These different levels of analysis at which the sentence-constituents are take in to account in Table 3.2 are a good illustration of the different roles the mind of the interlocutor should assign to the noun 'Betty' during listening to these simple sentences. We can additionally comment the different roles attributed to the noun 'Betty' with the words of Wallace Chafe, saying that *There are two broad kinds of considerations that enter into the identification of noun's status or statuses. It would be fashionable to call them syntactic and cognitive considerations* (Chafe, 1973). We will now turn to the more cognitive ones.

3.1.3.2 Cognitive characterization of Topics: Aboutness, Frame and definiteness

> *There are two broad kinds of considerations that enter into the identification of noun's status or statuses. It would be fashionable to call them syntactic and cognitive considerations*
>
> WALLACE L. CHAFE, 1973

By continuing along this path, questioning the notion Subject-hood and Topic-hood, one should say that many attempts in linguistics and psychology were made to characterize in cognitive terms these notions of "sentential subject-hood" (i.e. psychological, grammatical and logical) between the 50s' and the years 2000.

After the *aboutness* definition, that we can summarize in one sentence following Reinhart (1981), the topic is defined as "the entity that the sentence adds new information about" [33], we can briefly recall two of the most representative psychological contributions to sentential subject-hood phenomena. Carroll (1958) paired this notion to the *gestalt* figure-ground distinction, while Johnson-Laird's (1968a/b) definition of subject-hood identified the subject as the most prominent/important element of the sentence. Later, Linguistic investigation formulated in the 60's some more fine-grained distinctions of these sentential roles, by introducing the notion of *status of a noun*.

Gradually taking more and more into consideration the speaker/hearer's minds in linguistic analysis, Chafe (1967:27) discusses the notion of the status of a nouns in a sentence in the following way:

33. The term aboutness topic refers specifically to topics in this sense.

> The **statues** to be discussed here have more to do with how the content is transmitted than with the content itself. Specifically, they all have to do with **speaker's assessment of how the addressee is able to process what he is saying against the background of a particular context**. Not only *people's minds* contain a large store of knowledge, they are also *at a moment in certain temporary states with relation to that knowledge*
>
> [...]
>
> Language functions effectively only if the speaker takes account of such states in the mind of the person he is talking to.
>
> Wallace Chafe (1967:27, emphasis mine)

While Halliday's definition as the starting point of the message was bringing in the speaker's mind, here, the hearer's mind enters the scene too: the effectiveness of the conveyed message resides in the speaker's assessment of the hearer state of mind with regard to background contextual knowledge of what the speaker utters.

Also more pragmatically-oriented approaches like that of Reinhart (1982), Gundel (1988), Lambrecht (1994) and Vallduví (1992) bring in the concept of *mental state of the hearer* in the definition of Topic. They generally situate the Topic at a communicative-level of analysis –independent from that where the truth-value of the proposition is determined– This level depends on the representation that the speaker has of the mental state of the hearer.

The insight brought by these cognitively-oriented definitions, and the above epigraph from Chafe, give us the occasion to indulge in a short digression about our experimental approach and hypotheses.

One could, namely, say that what was only intuitively cognitive and '*fashion*' at the time of Chafe can be today experimentally addressed. Neuro-imaging of sentence comprehension becomes the field where detailed linguistic analysis and Cognitive Science converge to generate a *total picture* of how language 'cognitively works' at the level of the sentence-unit.

By unifying linguistic, psychological and cognitive evidence, this section wants to put forward the tailoring of our disciplinary background during these scholarly years consisted into intersecting multiple points of view on language phenomena and address through quantitative experimental observations these issues about speakers' and hearers' mind, during on-line comprehension of Topic sentences embedded in context.

The notion of Topic Frame and of the contextual status of the Topic referent will be investigated in next chapter 4, collecting behavioral measures reflecting the status of the Topic referent in the hearer's mind according to different context in which the same Topics are presented. Chapter 5 will record cerebral (ERP) on-line activation during listening to Topic-comment sentences featuring a Frame-setting relationship to their comment clause to directly ask the question of the cerberal on-line processing of this predicational relationship.

This work resolutely wants to give empirical backing to the cognitive considerations Linguistics has been formulating on this sentence articulation when attributing it a [***syntactic and cognitive structural dimension***]. Therefore, the theoretical and

linguistic considerations found in this more notional section are gradually introduced to show how they became experimental hypotheses.

The notions of definiteness and aboutness

As we can read in the above quotation of Chafe, people's mind (and its "large knowledge") became a place to be taken into account when speaking, to then evaluate if singling out an elements from it will be efficient for communication. Following the definition of Topic as "what is being talked about" we can engage in two different kinds of perspective, that of pragmatics and that of syntax. The first will concentrate on the discourse continuity and old versus new information analysis, and the second will structurally define the place in the sentence of the element, asking a narrower question "what the sentence is about?".

Hence, in the efforts to drag the definition of Topic, from a vague and intuitive one, to a more circumscribed one, the notion of *definiteness* adds a central informational characterization to the Topic referent. Early analyses already pointed out that Topic position hosts lexical elements whose referent are definite and known, as we can read in the foundational article of Chafe "Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View":

> [...] in the communicative situation whether I think you already know and can identify the particular referent I have in mind. If I think you can, I will give this item the status of *definite*."

However, the intuitive equivalence between definite and already appeared in the context is mistaken, namely according to Kuno (1972), a generic Nominal Phrase is also a definite one in that its referents can presumably be know by the interlocutor[34].

In Chinese the definteness of an Noun Phrase can be a determining factor to chose an informational packaging variant over another. As shown by the following Mandarin examples (66) a definite NP precedes the verb (a), while an indefinite one follows it (b).

(66) Definiteness and linearization in Chinese

a. 死人了。

sǐ rén le
die person PERF.

'A man has died.'

b. 人死了。

rén sǐ le
person die PERF.

'The man has died'

In this regard, we can cite a psycho-linguistic study on Topic-Comment development in Japanese children (from age 4 to age 6) showing that adding the definiteness feature to the Topic element (*sono* in the below examples) facilitates the comprehension of object

34. Kuno (1972), Functional sentence perspective: A case study from English and Japanese. Linguistic Inquiry 3 (3): 269-320.

Topicalized sentences bringing them tot the same comprehension rate of basic SVO (with *Sono*) sentences (Sano, 2005 and 2012)[35]

(67) a. Canonical SOV with sono

Sono zou-ga buta-o kettobashi-mashi-ta.
the elephant-NOM. pig-ACC. kick-Polite-PAST

'The elephant kicked the pig.'

b. OSV topicalization with sono

Sono buta-wa zou-ga kettobashi-mashi-ta.
the pig-TOP elephant-NOM. kick-Polite-PAST

'As for the pig$_i$, the elephant kicked [it$_i$].'

3.1.3.3 The Topic as a Frame for predication

Other kinds of Topic feature an element in sentence-initial position that appears only loosely related or associated to the comment-clause. Considering the sentence-initial Noun Phrase, Preposional phrase and phrase in (68) it is easy to find some shared characteristics between these two sentence-initial constituents, they all represent a frame for the comment clause:

(68) Frame-setting in Chinese

a. NP: 六点，他报告已经写好了

Liùdiǎn, tā bàogào yǐjīng xiě-hǎo le.
6-o'clock he report already wrote-RES(finish) LE-PRT.

'At 6 o'clock, he had already finished writing the report' from Li (2007)

b. clause: 头一天上工，祥子就差一点发了昏。

Tóuyītiān shànggōng, Xiángzi jiù chàyīdiǎn fā-le-hūn.
first-day go-to-work Xiángzi then almost make-PERF.-faint

'Xiángzi almost fainted after his first day of work.' (from *Rickshaw Boy*, Lao She, 1937)

c. (在) 台湾你可以吃到很多种水果

(zài) Táiwān nǐ kěyǐ chīdào hěn duō zhǒng shuǐguǒ
(at) Taiwan you can eat very many kinds fruit

'In Taiwan, you can eat many kinds of fruit.'

35. Children at three years of age showed some difficulty in comprehending OSV topicalization sentences as compared to canonical SOV sentences (75% vs. 87.5%), but performed much better than in those without *sono*. In the condition with sono, the asymmetry disappeared from age four to age six (e.g., at age four, 82.1% vs. 89.3%). Given these patterns the author argues that the presence of topic marker *-wa* is not sufficient to grant children's successful comprehension of OSV topicalization, an that the definite marker *sono* appears crucially necessary. Anticipating on a notion we will present in next section, we could speculate that *sono* attributes a greater salience and identifiability to the Topic referent. Note that the example 67 would be more correct with kettobasu.

The common notional property making them alike is that all express the domain in which it is pertinent to say what follows in the Comment predication. Wallace Chafe will introduce the notion of ***Frame*** in order characterize this notional relation from the point of view of the semantic relationship that the Topic and the Comment entertain, we can namely read:

> "What the topics appear to do is limit the applicability of the main predication to certain restricted domain [...]another way is to say that the topic sets a spacial, temporal, or individual framework within which the main predication holds."
>
> W.L. Chafe in Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View, in '*Subject and Topic* Ch. Li (ed.), Academic Press, 1977.

The function of this sentential Frame is more specific than the general *aboutness* definition of the Topic, it plays a crutial role of establising the limits of the applicability of the predication.

One looked at such sentences form the point of view of incremental processing, a more psycho-linguistic consideration, could be that the Topic's role is constant throughout the sentence, specifically because it does not directly participate to the event structure of the clause. Hence its framing role limits the applicability of the predication of the comment as a whole, and not only at a particular moment in time as it would be the case if the object of the verb had been pre-posed in sentence-initial position.

This observation is one of the first reasons that made us choose to first experimentally investigate the on-line processing of Frame-setting Topic-Comment configuration in Mandarin Chinese (cf. chapter 5), our aim was to uncover some of the cognitive aspects and cerebral processing mechanisms underlying this Frame-like limitation of sentential predication.

Moreover, it should be noted here that this framing function also participates to connecting the sentence-unit to the discourse-level, in that the Topics playing this role can hardly be indefinite, as we just saw.

Anticipating over the next Section dedicated to 'Chinese style' Topics, we can note that Chafe's notion of Frame, together with the aboutness one presented above will constitute the two theoretical definitions of the Topic, on which Li and Thompson (1981) will base their typological definition of Topic-prominence. However, this does not implies that the frame-setting function of topical elements is not to be observed in subject-prominent languages, but that the frequency observe in Topic-prominent languages of building the sentence according to his semantic relation, is mainly only observed in oral register.

In conclusion, we can say that the notion of Topic has received a wide variety of different definitions, successively making reference to psychological subject-hood, aboutness, definiteness, starting point of the message, presupposed information, or contextual giveness. These different contributions to the definition of Topic maybe arguably be considered in the next coming sections as phenomena of a different nature or level of analysis. However, from this last definition of Chafe (1976), and the more recently developed classification by Jacobs (2001/2004), we will further argue in next section that one should differentiate between a notion of Topic that more vaguely identifies the entity

about which a Comment is made -the Aboutness Topic-, from another Topic notion more specifically settings the *Frame* within which a proposition holds -the Frame-setting topic.

3.1.3.4 More than aboutness: a Typology of Topics

Exploring the notion of topicality in previous section, we repeatedly came across the concept of *predication*, that lies at the frontier between philosophy, logic and linguistics, but most of all semantics.

The semantico-pragmatic relationship Topics have with their Comment can help identifying different types of Topics, mainly according to different interpretative properties that characterize them. Jacobs (2001) proposes four dimensions that are at stake in the notion of Topicality, whose interplay would actually define Topichood: (1) informational separation; (2) predication; (3) addressation and (4) frame-setting. Hence, based on this typology Jacobs proposes three main sentential configurations presenting different distributions of the four semantico-pragmatic dimensions, namely Left-Dislocation, I(ntonational)-Topicalization and free Topics.

Importantly, such a logico-semantic typology will also have a impact on their syntactic encoding across languages and for the analysis of "Chinese Topics". We will therefore indulge outside the field of syntax for some brief considerations on the semantic relationship the Topic entertains with the Comment and thereby specify the notion Topic and its different syntactic encodings.

It would be namely difficult to grasp the notion of Topic without defining the different kind of predicative relations it entertain with the Comment. However, as this overview in the notion of Topic has probably already demonstrated, sentence Topics have been proven difficult to define, and one of the reasons we haven't mention it yet is that the way Topics limit the applicability of the main predication is not homogeneous. Namely, the difficulty to of assigning a fixed *interpretational value* to Topics –specifically in Chinese (for further arguments on this point see Paul, 2010) and when no co-referential of syntactic dependency–link can be established between the Topic and Comment (cf. Gapless topics, §3.4.1, p. 361), implies that there can difficulty be a distinct semantic feature associated with Topic position, if not the already discussed quite general *aboutness* or frame-like relation.

Hence, we briefly consider here three semantically-defined different types of Topics that will nonetheless allow to characterize a bit more precisely the relation that Topic and Comment entertain. In this, we wont forget syntax for too long because these different Topic's types appear to manifest different structural marking strategies across languages, as put forward by formal linguistic approaches to Topic-Comment structures. In section §3.4.5, these types of Topics will be mapped to a syntactic position in the Left Periphery of the sentence, following the cartographic approach of Chinese Mandarin Topic Field in Badan (2015).

We can consistently identify across languages at least three types of Topics:

A Hanging topics[36]: Typically introduced across languages by "As for" expressions, they frequently are at the beginning of the sentence.They have been analyzed pragmatically as a speech act of its own, that is used in situations where the speaker wishes to address something that was not addressed in the immediately preceding context, and operates a *shift* in the Topic of discourse. The relation

36. These Topics are sometimes semantically called Relevance Topics

between Topic and Comment is that the Comment *must say something relevant in relation to the topical element*. An example of this type of Topics was given in (28) repeated here under in (69) for convenience:

(69) a. As for American self-confidence, Columbia gave people a lift.
b. * American self-confidence, Columbia gave people a lift.

B Frame-setting topics: The Frame-setting relation restricts the domain of reality with regard to which the truth of the Comment-clause is to be evaluated[37]. Interestingly, these kind of Topics have been often analyzed in Sign-languages (for example in Russian Sign-Language (RSL) see Kimmelman, 2012 and for Hong-Kong Sign-Language see Sze, 2008).
As we can see in the following examples in (a) the predication is restricted to the physical dimension of 'the driver'. More generally, the first constituent be it an adverb (a), a Noun Phrase (c) or a spatial NP (b), feature a relation with the comment-clause that is of restricting the domain where the Comment predication is true:

(70) Frame-setting
a. Health-wise, the driver is fine.
b. Körperlich geht es Peter gut. (Jacobs 2001:656)
c. 温布尔登网球场大雨 (早就) 停了。
Wēnbùěrdēng wǎngqiúcháng dàyǔ (zǎojiù) tíng-le.
Wimbledon tennis-court heavy-rain already stop-PERF.
'On Wimbleton tennis ground the rain already stopped falling.'
d. 温布尔登网球赛大雨停了。
Wēnbùěrdēng wǎngqiúsài dàyǔ tíng-le.
Wimbledon tennis-competition heavy-rain stop-PERF.
Interpretation: 'During the Wimbledon competition the heavy rain stopped.'
Interpretation 2: 'During the Wimbledon competition there was heavy rain and it stopped.'

C Contrastive Topic: What is particular about contrastive Topics is that they not only identify a referent about which a Comment is made, but they additionally convey that there are other alternatives (competing entities) among which the speaker chose the Topic to contrastively characterize his message, for which the predication of the comment does not hold by implicature.
Hence, this type of Topic can be accompanied by explicit alternatives or not. A typical example of overt contrastiveness marking of Topic is classically found in Japanese. Interestingly the literature on its prosodic marking is showing that Japanese topical marker -*wa*, its subject marker -ga and the emphatic prosodical marking yield different interpretation of the same sentence as illustrated by examples in (71). These different interpretation are prosodically conveyed by different intonational patterns as was demonstrated by a series of studies by Nakanishi

37. This definition recalls the one we found at the beginning of our research in an inspirational article by Sophie Prévost (2003/1998) defining the Topic as a domain of validation (in French: '*domaine de validation*', inside which we the 'Pertinence' of the comment's information is validated.

(2000/2003). The results of this study are illustrated in Figure 3.4. Note that the Contrastive function for *-wa* marked constituents is only attested in the case where a *-wa* phrase is the answer to a wh-question (Kuroda, 2005) example reported here-under by Hara (2006:101), or when it is followed by an alternative that is also marked by -wa as in Figure 3.4.

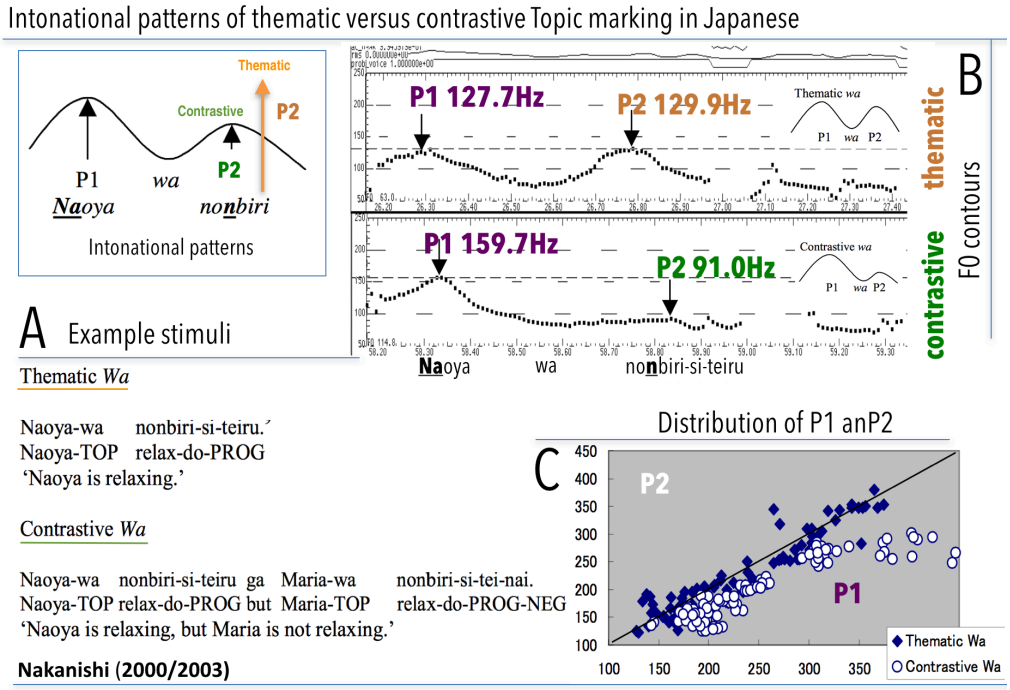


Figure 3.4 – (A) Example stimuli of thematic (a) and contrastive -wa (b); (B) F0 contours of thematic -wa and contrastive -wa in Participant KO (male); (c) The distribution of the thematic and the contrastive cases of the five participants, where the X-axis indicates the value of P1 (Hz), and the Y-axis indicates the value of P2 (Hz). When -wa marks theme role, the highest value of F0 contour after wa (P2) is as high as or even higher than the highest value before wa (P1), while when it marks contrastiveness, P2 is much lower than P1. Adapted from Nakanishi (2000/2003).

(71) Topic's Contrastive implicature and its overt marking

a. Japanese non-contrastive Topic

Mary-wa ukat-ta.
Mary-TOP. pass-PST.

'As for Mary, she passed.'

b. A : Who passed the exam?

i. B1: Japanese contrastive Topic [-wa + intonation]

Mary-wa ukat-ta.
Mary-CONTR. pass-PST.

'Mary passed.' > The implicature is 'Possibly, others didn't pass.'

ii. B2: Japanese non-contrastive

Mary-ga ukat-ta.
Mary-NOM. pass-PST.
'Mary passed.' > with no implicature

iii. B3: German contrastive Topic [+ intonation]

Maria /HAT bestanden.
Mary has passed
'Mary passed.' > The implicature is 'Possibly, others didn't pass.'

The same effect of contrastiveness[38] can be obtained in other languages like French, German (see example [71b-iii]) or Italian with an appropriate intonation contour. For instance, a contrastive accent on the answer term can be viewed as the neutralization of the exhaustiveness implicature that is generally attested for of the standard intonation. Interestingly, when syntax or morphology do not encode contrastiveness prosodic realization, intonational patterns can express information structure.

These different Topic's types appear to manifest different structural marking strategies across languages. For instance, the syntactic realization of aboutness Topics is debated in German (see Frey 2004b on middle field), and based on these considerations the following diagram in Figure3.5 shows how the different types of Topics that have been described here by analyzing the essential features they share.

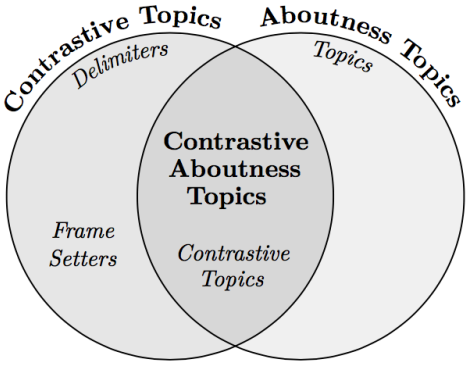


Figure 3.5 – Diagram representing the analysis by Krifka (2007), in italics, and Gyuris (2012), in bold, of the different features of Contrastive Topics compared to Aboutness topics, adapter from Constant's PhD (2014).

3.1.3.5 The cognitive role of Subjecthood and Topichood

To further clarify the notion of Topic, we will here pinpoint its characteristics in comparison with the one of the subject, by doing this, the fundamental difference between topics and subjects will emerge from a more structural point of view, to which we will parallelly add some psycho-linguistic evidence.

Although various authors have been defining topic and subject simply by analyzing them as being respectively a pragmatic concept at discourse-level for the first, and a semantic concept at the sentence-level for the second[39], we want to delineate a more

38. More precisely, the distinction between the thematic and the contrastive -wa is valid only when the topic marker is attached to the subject in canonical SOV word-order. When the topic marker is attached to the object in canonical position, the object is exclusively interpreted as a contrastive element, as shown in (1).

(1) John-ga ringo-wa tabe-ta.
John-NOM. apple-TOP. eat-PAST.
'John ate apples' (but there were some food that he didn't eat)

39. Different terminology does not always means different concepts , when topic and theme are *applied to the analysis of the clause structure*, certain authors treat them as equivalents or at least referring to the same type of element in different theoretical frameworks (see Reinhart, 1982 or Chu, 1998). It is anyway hardly the case that different theoretical approaches do not develop these concepts with different underlying assumptions and taking into account different parameters. For the Topic theme opposition

precise view of Topic-hood departing from the elementary definition of "what is being talked about" and present in the different visual angles offered by semantics, pragmatics and the functional perspective on sentence's structure.

Pragmatic approaches and logico-semantic analyses

Several definitions issued from logico-semantic analyses allow to give a clearer definition of the difference between the notion of topic and that of subject.

The Topic is part of the Sentence's presupposition As we already saw in our account of the inaugural philosophical approach to these concepts in section §3.1.2 (p. 237) the type of judgment posed by Topic-Comment sentences is a categorical one. Opposed to thetic judgment where an entity is posed like in 'It's rainin' or 'Train's arrived', categorical judgment is two-tailed. In the first step a phenomenon is selected, and in a second step a description or comment is given. To this should be added that, selecting a purely informational perspective, the informational status of the Topic differs from that of subject in that it is part of the utterance's presupposition, and it is for this reason that question and negation operator cannot scope over it.

The Topic is extending its truth-conditional value over the whole sentence Among the semantic approaches to the issue of topicality, the works of Kiss (1995) and Erteshik-Shir (1997/1999) illustrate a type of linguistic analysis where topicality is analyzed as giving the sentence-initial element a truth-conditional value that has an effect over the whole utterance. For instance, the truth-value of a Topic-Comment propositions is to be evaluated by determining the truth of its predicate in relation to the Topic (Erteshik-Shir, 1997/1999).

> "the topic has a semantic function: it acts as a restrictor as to when, where or with respect to who or what, the truth value of the predication is to be evaluated" (Erteschik-Shir, 1997:130).

In this perspective adverbs and even conditional phrases are viewed as Topics. We can add that it is mainly in this perspective that the Topic can be considered as corresponding to the proposition's *logical* or *semantic subject* (cf. Jacobs, 2001)[40] .

Topics and Context of the utterance Additionally, discourse-oriented approaches like the informational and communicational ones, take into account the broad context of enunciation, and are therefore able to introduce the notion of the "informational load" of the sentential elements against the discourse background information. To the role of Topics in the truth-conditional value of the Comment-clause, Reinhardt adds that they play a role in the storage of New Information given by the speaker hearing, so that one has to "Assess by what you already know about the topic; store under an entry corresponding to this topic" (1982:24).

---

this is namely the case. While the topic is generally indicating the element in the sentence, the term theme frequently is meant to refer to the discourse topic in a multi-clausal perspective.

40. Resuming to the criteria introduced in previous section (e.g. definiteness, ect.) we can say that what Jacobs explains is the reason why only referential topics are allowed. "Addressation is incompatible with nonspecific, quantified, and negative phrases".

In this regard, Jacobs further specifies the notion of Pragmatic aboutness initially raised by Reinhardt (1982) by defining the concept of *Address*, which indicates the storage mechanism implied by Topichood. Reinhardt had formulated it in the following terms:

> A useful metaphor for the procedure involved [...] is the organization of a library catalog [...]. The propositions admitted into the context set are classified into subsets of propositions, which are stored under defining entries. At least some such entries are determined by DP-interpretations. DP sentence-topics, then, will be referential entries under which we classify propositions in the context set and the propositions under such entries in the context set represent what we know about them in this set [...]. (Reinhart, 1982:24)

The individual contribution of these different approaches to Topic linguistic facts contributed to our experimental approach in the following terms:

1. While the logic-semantic approach (1) alimented our interest for the neural correlates of the "two-stepped predication" found in Topic-Comment articulations, considering it as another way of logically articulating the utterance,
2. the communicational approach inspired our idea to vary the *informational load* of the Topic referent by forging adapted narratives to embed our experimental sentences, as illustrating in Figure 3.6 here under),
3. the above semantic analysis, and specifically the definition of Topic as truth-conditionally overarching the whole comment-clause's interpretation, inspired us for designing a behavioral task in our ERP study on Mandarin Topic-Comment (chapters 4 and 5).

Hence, in order to leverage on this truth-conditional mechanism, we asked our participants to evaluate if our Topic-Comment sentences were true given the background they read in the narratives, assuming in this way the Topic would be used as a basis for the evaluation of the truth-value of the whole utterance in relation to the Comment predicate.

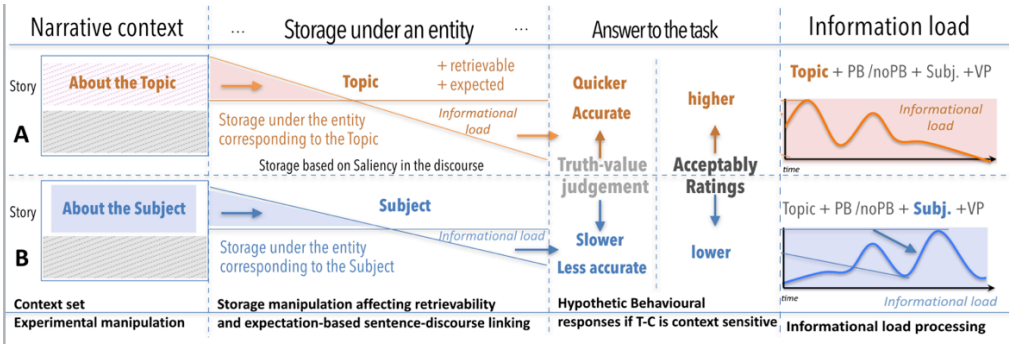


Figure 3.6 – Experimental manipulation of the Context Set informational load processing and behavioral task.

As shown by examples in (72), when a narrative was telling the story of an old farmer was sleepy in a field of watermelons while a grocery store boy was stealing a watermelon

(Story n°2), the participants had to judge the truth-value according to the context of mainly tree types of possible configuration, as illustrated by the following examples:

(72) Examples of the sentences in the task from our ERP study on Chinese Topic-comment sentence

a. Filler sentence 西瓜田老农对小明进行了批评教育。(PBTop): Wrong place

Xīguā tián, lǎonóng duì xiǎo míng jìnxíng le pīpíng jiàoyù.
watermelons field, old.farmer to Xiǎo Míng carry.on ASP. criticize education

'In the watermelon field the old farmer preached morality to Xiao Ming.'

b. Filler sentence 西瓜田，小明睡了一觉。(0201f PBTop): Wrong Subject actor

Xīguā tián, xiǎo míng shuì le yī jiào.
watermelons field, Xiǎo Míng spleep ASP. one CL.

'In the watermelon field Xiao Ming had a nap.'

c. Filler sentence 杂货铺，警察打了小明一顿。(0202f PBTop): wrong action

Záhuòpù, jǐngchá dǎ le xiǎo míng yī dùn.
Grocery.store, police hit ASP. Xiao Ming one CL.

'At the grocery store, the police hit Xiao Ming.'

Last but not least, restricting the perimeter of the linguistic analysis to the intra-sentential domain, the introduction of the notion of Topic gives also rise to a syntactic the bi-partition of the utterance, for instance the Topic and the rest that can already be observed in child utterances. We will dedicate next two sections to this syntactic approach.

This syntactic perspective appears not only to have been extensively theoretically and formally grounded in Linguistics[41], but the developmental studies proliferating in the 60's have been contributing a psycho-linguistic point of view that can reveal some of the semantico-syntactic criteria characterizing Topics in contrast with Subjects, as we will see in the following.

3.1.3.6 Topic-Comment in Child language

Lacking the complexity of syntactic configurations present in adult language, Child linguistic behavior offers interesting insights for an investigation into the cognitive basis of Topic-Comment sentence articulation.

First Language Acquisition studies converge in showing that English children produce Topic-comment structures as soon as they reach the two-word stage, like in "Mommy go" or "Water hot". Some authors also claim that the Topic function is already set in one word-stage prior to any use of word-order (Bates, 1976; Bates and MacWhinney, 1978), in communication motivated pointing gestures. In Bates (1976) we can namely read the following: "[...] in the absence of clear-cut semantic-syntactic constraints, the first ordering principle adopted by children will be Comment-Topic[42]" (Bates, 1976:210).

41. As we already saw in general terms in previous chapter in §2.4, and as we will concretely set for Mandarin in the next sections §3.4 and §3.2.

42. Not that in Bates (1967) the comment precedes the topic because she posited that this utterance configurations derives from the given/new dichotomy we already presented and that new information should occupy, in name of its psychological salience, the first position in the utterance.

Similarly, Meisel suggests that "[in multiword utterances up to age 2;0] the sequencing appears to be determined to a large extent by theme-rheme ordering" (Meisel, 1994:93).

Since the 60's, several linguistic and psycho-linguistic studies in language acquisition (see Bates, 1967 or Gruber, 1967) have analyzed two-words utterances across languages, and what emerges form these studies is that the first Topic-Comment two-words utterances display a number of interesting feature that allow to distinguish Topic and subject.

Two-words stage is still lacking the rich morpho-syntactic tools of adult speech, and typically do not produce tense or subject-verb agreement (Owens, 1988; Givon, 1979; Fuller and Grundel, 1987). Hence, early Topic-Comment utterances feature logico-semantic relationships between the two words that compose them (e.g. 'Mommy go') that are only loosely reflected in grammatical marking as in the below two-words Italian examples from Bambini and Torregrossa in (73).

(73) CHILDES corpus - Italian Topic-Comment two-words utterances

a. gatto/grosso (Martina 1;10)

gatto/grosso
cat/big

'the cat is big'

b. zia/trattore (Martina 1;9)

zia/trattore
aunt/tractor

'Aunt's tractor' meaning the tractor [she is playing with] belongs to her aunt'

c. dentro/chicco (Rosa 1;11)

Dentro/chicco
Inside/grain

'a grain is inside'

d. Fello/gol (Raffaello 1;10)

Fello/gol
Fello/goal

'Raffaello did a goal' to report his action of kicking

Considering the contribution of each constituent and the nature of the relation that binds them in an uttered-unit, topic/comment structures can indeed provide a general predicational pattern for these early utterances.

Crucially, this developmental configuration results in a tendency to systematically avoid of subject-verb agreement in children utterances favoring Topics which have the prototypical characteristic of not necessarily be arguments of the main verb in the Comment. It is namely because of this characteristic that Topic-Comment construction have been argued to be learned earlier, because the kind of morpho-syntactic rules they involve are simpler than those required by subject-predicate relations. Some authors namely argue that at this age the abstract grammatical notion of subject may not have developed yet (Bowermann, 1973; de Villiers and de Villier, 1978).

Such a pattern is still observable when children reach the stage of uttering more complex utterances, like the following English and French examples, respectively in (74) and

(76). These linguistic behavior, avoiding subject-predicate relationship, could actually be seen as paralleling what we observed in the utterance strategies of fluent and non-fluent agrammatic patients (cf. §3.1.2.4).

As a matter of facts, the following English and French examples from Gruber (1976) and De Cat (2002), witness a subject-predicate avoidance even if morpho-syntactic mechanism like plural and past tense are already established:

(74) Children Multi-words Topic-Comment constructions (Gruber, 1976)

a. It broke, wheels

b. Car, he take the wheels

c. Where went, the wheels

(75) French Topic Comment utternances (De Cat corpus)

a. (Max 2;0.14)

Lui, ça va là.
He, that goes here

'That one goes there.'

b. (Anne 2;2.00)

Maman, [ ] fait ça %# moi, [ ] fais [ ] drapeau
Maman, [elle] fait ça %# moi, [je] fais [le] drapeau

'Mum, [she] makes [the] xxx and me, [I] make [the] flag.'

Moreover, investigating the relationship between the semantico-discursive functions (i.e. Information Structure) of Topic and Focus notions, and their corresponding syntactic devices during French first language acquisition, Cécile De Cat and Karen Lahousse (2002 and much related work) have set the emergence of these linguistic encoding of the utterance early in French children's second year. In this regard, the authors identified in their corpus (i.e. 'De Cat corpus') mainly two strategies, illustrated by child utterances in (76):

1. the first *dislocation*,
   (a) or without (76a)
   (b) with an explicit resumptive element. (76a)
2. the second *clefting*,
   (a) introduced by '(c')est' meaning 'it is' (76b-e) and
   (b) by '*il y a*' 'there is' (76a), whether the relative clause is expressed (76a-d) or left implicit (76e), and whether the relative pronoun (or complementizer) is expressed (76a-b) or not (76c-e).

(76) a. Il y a maman qui arrive!
'There's mum who is arriving !'

b. C'est lui qui fait dodo sur lui.(Anne 2;4.20)
'It's him who sleeps on him.'

c. C'est Bruno, a mise comme ça. (Tom 2;1.13)

C'est Bruno [qui] a mise comme ça.
It.is Bruno [who] has put like that

'It's Bruno [who] put [it] like that.'

d. C'est Mamy, a mangé. (Tom 2;3.22)

C'est Mamy, [qui] a mangé.
It.is Mamy, [who] has eaten

'It's nanny [who] ate.'

e. est Maman, ça. (Anne 1;11.13, from De Cat corpus)

[C'] est Maman, [qui le fait] ça
[it] is mum [who it does] that

'[It]'s mum [who does] that.'

Importantly, examples from Mandarin Chinese L1 development in (77) show converging evidence for the subject-predicate avoidance even in a language that does not have an overt verbal marking of tense, agreement and number.

Chen (1983) observes that, between age 2,5 and 5, children have in fact a preference for the kind of Topic structures where the notion of Topic and subject overlap in one word *xiǎo háizi*, as in (77a).

(77) a. early acquisition: 小孩子，也上学，也游戏。

Xiǎo háizi, yě shàngxué, yě yóuxì
small kid, also go.to.school, also play.games

'The little kid goes to school and also plays games.' (Chen, 1983)

b. later acquisition: 小狗，尾巴喜欢动来动去。

Xiǎo gǒu wěiba xǐhuan dòng-lái-dòng-qù
small dog tail like move-around

'The small dog, [its] tail likes to move around'

The kind of Topic-Comment structures involving the grammatical presence of a subject, like the 'double subject constructions' in (77b), appear later in Chinese L1 acquisition, although they could seem simpler from the semantic point of view, as only one referent is to be identified. While semantically the Topic and the subject are in part-whole relation, syntactically the Subject ('dog') receives a distinct linguistic coding from the Topic ('tail') and this is the reason that Chen (1983) advocates for their late appearance in Chinese linguistic development.

Early utterances from French/English/Chinese first language acquisition clearly demonstrate how Topic-comment syntactic construction are accessible since early age. Taken together these examples further illustrate that the distinction between Subject and Topic is based on a certain amount of semantico-syntactic criteria, crucially it appears in an evident way in the different relations the two functional elements establish with the Comment's predicate: Subjects undergo the verbal selectional restrictions while Topics do not.

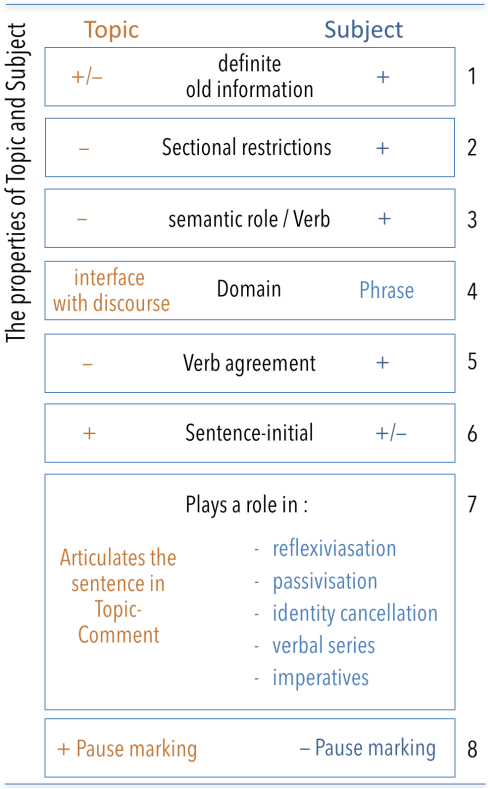


Figure 3.7 – The properties of Topic versus Subject as developed by Li and Thompson (1981), table adapter form M.C. Paris (1998).

To conclude, we can anticipate here what Li and Thompson contributed in the 70's to delineate the linguistic distinction between Topic and Subject. These authors put forward cross-linguistic evidence from 30 languages that Topic and Subject are distinguishable along a certain amount of features reported in figure 3.7. They are, namely, the first linguists having established a list of criteria and distinctive features[43] allowing to linguistically distinguish at the sentence-level between Topic and Subject (see Table 3.7).

Topic-Comment Articulation: simple or complex?

Before turning to next section, we would like to highlight a major discrepancy emerging from the relative simplicity that the above developmental studies attribute to Topic-comment articulation (e.g. its presence since two-words utterances, its pervasiveness in children multi-word utterances and in agrammatic talk), and the the rich set of arguments presented earlier in chapter 2 (§2.4.4, p.197) attributing to the sentence–discourse domain – the CP-layer housing the Topic – a complex cerebral representation.

The evident contrast between the developmental account of the pervasiveness of Topic-Comment and the fact that elements hosted in the CP-layer of the sentence cause more problems to agrammatic patients, is further attested in second-language learning by beginner's sentence patterns (cf. §3.1.2.4 and §3.2.1.3) and by several observations on the language development of pidgins and Creoles (Givon, 1979 and 1984).

A possible way to resolve this discrepancy is to hypothesize that while child grammar is still rudimentary in these early stages of language development, toddlers initially perform acts of predication (i.e. predicating something about something) based on some broad cognitive principle underlying the topic-comment articulation, leaving for later informational and syntactic patterns that are characteristic of the later phases of development we illustrated in the French examples (76).

Although a more thorough review of the literature on L1 acquisition focusing on the development of the CP-layer in children linguistic behavior is needed to answer this issue, by lack of space, we briefly advance the above speculations about the observed subject-predicate relation avoidance strategy in Children. And we attribute it to the gradual establishment of the complex morpho-syntactic rules characterizing agreement relations and not directly to the complex representation of the CP-layer. Thus Topic-Comment articulations are a good predicative strategy in absence of detailed grammatical knowledge as we saw in the case of aphasic conversational analysis too (§3.1.2.4, p.244).

43. The prosodic one -number 8- is our own addition based on our observation and on subsequent literature.

3.1.4 The context and the Sentence

> *"the speaker tries [...] to make the structure of his utterances congruent with his knowledge of the listener's* ***mental world****"*
>
> Clark and Haviland, 1977

From the early developments of the notion of Topic and Theme[44], the subsequent linguistic analyses will diverge in two distinct directions, one focusing on the sentence-level and the other on the infra-sentential one, mainly formulating on discourse-pragmatic considerations.

The notion of Topic we are referring to in this section is subject to a typical definitional problem in linguistics that is the stratification of levels of analysis[45]. So that it is useful to clarify that a major distinction is to be done between the syntactic place occupied inside the sentence by the element called Topic, from the informational-role played by the same element inside discourse [46].

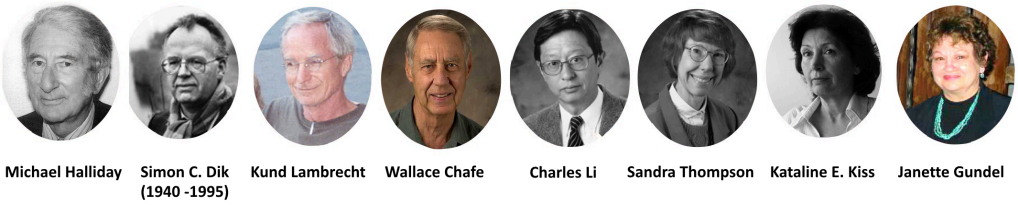


Figure 3.8 – The second generation of Linguists exploring the linguistics realization of the notion of Topic.

In the linguistic literature of the last 30 years, the term Topic can refer to linguistic phenomena taking place at different levels of the linguistic analysis. Across different authors, we can find terms like "sentence topic", "discourse topic", "communication theme", "logic theme", "informational theme", accordingly to the level at which we analyses the language : discourse level, sentence structure, communicative interaction, pragmatic context, textual, logic or informational dimension of the utterance.

The previously addressed definitions of Topichood not only put froward an analysis of the speakers/hearer mental representations but also introduced the context as an im-

44. As a terminological side note, we can easily observe that the XIXth and the beginning of the 20th century's linguists have frequently used them as similar concepts in the analysis of the utterance's structure. However, a more careful comparison of the notion of topic with that of theme (and the sets of concepts Topic-comment versus Theme-Rheme) can highlight major differences that are not our focus here.

45. "terme fourre-tout" in Fuchs, C., Marchello-Nizia, Ch., Présentation, Cahiers de praxématique, n°30, 1998:3.

46. The notions of thematization, theme or topic, have been frequently and generally used for the last 30 years of linguistics by both functionalists and formal analyses of the clause structure by a variety of highly diverse theoretical frameworks: Functional perspective, Typology, Textual Corpus linguistics, Discourse Analysis, Pragmatics, Information Theory, *Théorie des Opérations Énonciatives* (Theory of Enunciative Operations) (Culioli, 1990) and Generative Grammar (specifically in Cartographic project), just to name a few. Here is a sample of the different linguistic approaches making use of this notion, developing it on very divers conceptual basis and with different aims and perspectives.

portant element to be take in consideration to understand Topical linguistic phenomena. Namely, one of the functions of Topics inside the sentence-unit is to delineate an interface between the sentence and the discourse, defining a *relevance* relationship between the clause and the context. The investigation about the contextual status (e.g. accessibility, activation, identifiability, ect.) of the referent in the extra-sentential discourse environment has been studied under different points of view.

This information-oriented framework defines the Topic as referring to known, given, retrievable and accessible elements in the discourse. We will now consider these properties characterizing Topichood.

3.1.4.1 Information structure: a tale of given/new Referents ?

Referents and the informational taxonomy

The late developments of the Prague School (Firbas, 1964, Daneš, 1970, Sgall, 1986) and the influential article by Chafe in 1976 where already prefiguring the information structure analysis of the 80's and 90's. The functional grammar perspective had already demonstrated that utterances across-languages showed a varying tendency to group information according to Topic-Comment articulation, and had already identified the terms *theme* and *rheme* with 'old' and 'new' information, thus postulating that sentences showed a given-new information pattern.

Note that several analyses derived for information structure, like those subsumed under the name of cooperative communication approach, go beyond the sentence-level analysis, to investigate the way information is packaged within discourse to optimize communicative speech acts. Clark and Haviland (1977) had early formulated the idea of efficient communication as a given-new contract, where: "the speaker tries [...] to make the structure of his utterances congruent with his knowledge of the listener's *mental world*" (Clark and Haviland, 1977).

It is a matter of fact that Topic's referents appear frequently as known information and the majority of the definitions of Topic in the subsequent literature (see Lambrecht, 1994) call into play the notion of given/new information, and applied it to the different referents of sentential elements. This dichoto,y introduces another notion contrasting with that of Topic, the Focus, providing new or salient information in a sentence.

Focus

Contrary to Topic-Comment structure, focalization typically involves a focus-background articulation, where the first is the 'informative' part of the utterance and the second is the 'non-informative' one. The background is defined as the knowledge that the speaker presupposes to be shared by the hearer (Halliday, 1967; Jackendoff, 1972; Vallduvì and Engdhal, 1996 among others).

There exist different uses of focus, among which the one carrying new information is its simplest. A classical way to evaluate the so-called 'information focus' of a sentence is to verify which sentential element can correspond to a wh-question, as for 'John' in example (78a).

(78) Focus types

a. Q: Who did you meet yesterday? A: I met [John]$_{\text{F}}$.

b. Q: Did you meet Paul yesterday? A: I met [John]$_{\text{F}}$.

c. It is [John]$_{\text{F}}$ who broke the vase.

d. I could ate *even* [a wedding cake]$_{\text{F}}$.

A more contextually-orient focus is illustrated in (c) by a cleft sentence (Krifka, 2007; Kiss, 1998), which is singling out that 'John' is the only one that leads to a true proposition. *Exhaustive focus* denotation implies an identity statement, as in 'the one who broke the vase is John, and not someone else'.

However, the link with context can also be more complex and imply a more sophisticated pragmatic calculation on contextual information. Focus can be realized also by means of scalar particles (see König, 1991). In '*scalar focus*', the alternatives are ordered on a scale and the focal denotation is identified by either the least likely or biggest element on the scale. Consider for example (d), where the focused element preceded by 'even' is indeed the least probable element one could eat on an hypothesized scale of eatable entities. We will resume to this kind of Focus in section §3.4.3 to present its syntactic specificity in Mandarin Chinese, and we will test for it in our fMRI experiment in chapter 6.

Another type of focus is the one typically used to correct or confirm a piece of information as in (b). This phenoma, known under the term of 'contrastive focus' (see Kiss, 1998 among others), contributed to attract linguists attention on the prosodic factors yielding contrastive interpretation on *in-situ* sentential elements, that are marked through a prominent accent.

This introduces us to an other contrastive phenomenon that we eluded so far that will also be part of our fMRI investigation of the sentence-unit underpinnings.

Contrastiveness

Importantly, for our understanding of topical phenomena, several authors pointed out that Topic elements can bear a contrastive interpretation across languages (e.g. Jacobs, 1984 and 1996; Lambrecht, 1994; Büring, 1998), and nowadays these phenomena are generally subsumed under the **name and notion** of *contrastive topics*.

Carrying out an analysis of the informational roles, contrastive Topics are usually defined as new elements linked or contrastively bound to a given context information, or event. Hence, what is special about contrastive Topics is that they do not only identify an entity about which a Comment is made, but signal that, there are other 'competing' entities in the context about which a Comment could have been made, and was not made. Hence, contrastive Topics indicate that the speaker chose among a number of alternative Topic candidates to utter the sentence according to this pattern.

The pragmatic function of the topic in its contrastive dimension can be exemplified in the following example by Daniel Büring. Given a context like [This university has 20,000 students.], the utterance under analysis would be "*The freshmen, live in the dorms.*", where "The freshmen", plays here a role that obligatory realizes 'a dialogue' with the context. To understand the propositional information given by this sentence the dialogue with the context information is unavoidable to single out "The freshmen" from "the 20,000 students from that university". Our fMRI experiment in chapter 7 will attempt to tackle the cerebral underpinnings of this 'dialogue'.

Topic: not so 'old' information

Even though the correlation between the Topic as referring entities mentioned before or previously used expressions, and the Comment as referring entities being newly introduced can hold in many cases, it is not always the case. This equation has been largely contradicted by a substantial number of studies. For example Halliday (1967) showed that the Comment part of the sentence can contain given expressions, and Reinhart (1982) showed that topicality cannot be reduced to old information, typically when the topic is used to shift the main object that is being discussed (i.e. Topic-shift)[47].

Lambrecht: what is new and old?

Importantly it should be clarified that the notion of informational value/status is a relational concept. Thus, evaluating the more or less known/given character of a referent does not help defining its informational value. In fact, if we take a cleft sentence like 'It's Leo who won the contest", 'Leo' in this sentence is a known referent but it is its relation with the predicate that constitute a new information (cf. Prévost, 1998).

Hence we can say that to carry out an analysis concerned with the processing of information does not mean just categorizing the discourse referents in terms of given/new. This erroneous vision associating a simplified version of given/new taxonomy to information structure analysis is mainly imputable to a confusion about the object that is being qualified as given/known or new. The majority of the times it is the discourse referent that is qualified in these terms, but not the information itself. Hence, when it is the referent conveyed by certain elements of the utterance, that is analyzed as known or new, the analysis takes into account the speaker/listen couple and the amount of knowledge they share. This is not explicitly or necessarily and analysis of the information structure of the utterance.

Lambrecht is the author that more explicitly pointed out the clarification we just gave. In his theory, he distinguished the mental representation that one has of the referent from the relations in which it can be established, one can namely have an old referent and a new relation as in the above cleft example.

Old information in Lambrecht is a 'presupposition', while new information correspond to informatively adding some element, and it takes the abstract form of an 'assertion'. Hence, considering the same example as above 'It's Leo who won the contest", the presupposition corresponds to the fact that somebody won and the assertion (i.e. new) would be that it is Leo. This analysis further attests that information structure defines the status of the referents according to the role they play in the proposition.

Communicative Dynamism of the elements in the sentence: Firbas

The position proposed by **Firbas** in 1974 (and then in 1992), diverges from previous analyses by the refusal of the clear opposition between given-known. Advocating an informational continuum to solve this dichotomy, he adopts a functional perspective on the sentence, and defines it as taking mainly two forms, either the Presentation of a phenomenon or the attribution of a quality to the phenomenon in question (cf. the Philosophical accounts of Topic-Comment in §3.1.2). On this basis, Firbas develops a

47. See Paul (2006) for examples in Mandarin Chinese.

notion of communicative dynamism where the three main criteria are identified. Contextual dependence, word order and the content of the semantic relations altogether allow to establish several degrees of communicative dynamism. Firbas' analysis assigns it a degree of communicative dynamism, according to how much an element of the sentence contribute to the communicative development and accomplishment of the utterance.

Interestingly, only the Topic, as "presentating a phenomenon", according to Firbas definition, does not contribute to the development of the communicative dynamism, and it is therefore assigned a zero degree of communicative dynamism, while the Comment is attributed the highest level.

This last point on Topic communicative "stillness" can be put in parallel with Reinhart's (1982) formal model of information and communication, where information is modeled as a set of file cards identifying an entity and listing the properties of that entity. In her framework the Topic element is a file card which only contributes to the naming of the entity it collects information about, while the Comment is interpreted as an expression adding information to the Topic file card [48].

3.1.4.2 Contextual Accessibility of the referent: Chafe (1987)

On this bases, Chafe (1987) introduced the possibility of having an intermediary level between given and new, that he termed as "accessible", having two possible activation statuses, i.e. active and inactive.

Active referents are indeed active in the discourse, they are belonging to the interlocutors' "focus of consciousness", in this they are readily identifiable as Topics in that no additional mental effort is required to process the sentences containing them. Accessible referents on the contrary are those referents that are less readily 'sizable', but still frequently uses as Topics. They can be viewed as situated in the interlocutors' peripheral consciousness belonging to some background awareness, hence their retrieval is demanding a minimal mental effort. Finally, inactive referents are absent from the preceding discourse and they are only available in the interlocutors' long-term memory, and therefore require to be inferred with a greater effort.

Hence, the centrality of considerations about textual, situational or intra-/extra-discursive utterance's context, and the consequent informational status they assign, determine the accessibility and retrievablility of the topic referent. Namely, referents can be textually, situationally or inferentially accessible and it is the sentence's context that becomes the criterion the determines the given, old, identifiable, retrievable, active, new or inactive of a referent that would be a candidate to play the Topic role in an utterance.

The importance given to contextual pre-activation of the referent refers back to the concept that was introduced by Firbas under the name of "immediately relevant context", a concept the author defines as follows:

> "[...] immediate relevance being assessed in relation to the point in the flow of communication at which a new sentence is produced and/or taken in, and which separates the mass of information accumulated so far from the mass of information to be further accumulated."

48. This formal information setting has been developed more precisely in the framework of *file change semantics* (Heim, 1983) by Portner and Yabushita (1998), but we do not have the palace here to deepen this very interesting point on formal modeling of Topic-Comment predication. This could be, by the way a perspective to address experimentally in further research.

As made explicit in these last considerations, the concept of contextual accessibility brings us closer and closer to cognitive and incremental understanding of sentence comprehension in context, an issue that will be experimentally addressed in chapter 5.

3.1.4.3 Identifiability and its cognitive flavor

In the attempt to extract some of the cognitive and psycho-linguistics features from the different definitions the literature gave of Topic-Comment articulations, our attention was first retained by another Topic feature theorized by the informational approach: *Identifiability*.

While the first analysis of Topic sentence in terms of 'activation states of the discourse referents is to be attributed to Chafe (1974 and 1987), Lambrecht and Prince[49] parallel this term with that of *Identifiability*.

For Chafe, activation status concerns only identifiable referents and from a cognitive point of view, this author underlines that activation state presupposes a mental representation in the head of the interlocutor (i.e. the receiver):

> "[...] our minds contain very large amounts of knowledge or information, and only a very small amount of this information can be focused on, or be 'active' at any one time."

Thus, these two criteria define now the status of mental representations in the hearer. As we can read here under, they appear in closer analogy with some cognitive processes linked to the choice of sentence structures, which could be in turn mapped to the cerebral processing of the sentence.

> "The first (information structure category) is identifiability, which has to do with a speaker's assessment of whether a discourse representation of a particular referent is already stored in the hearer's mind or not. The second is activation ,which has to do with the speaker's assessment of the status of the representation of an identifiable referent [...] in the mind of the hearer at the time of speech act." Knud Lambrecht *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*, (3.1), 1994

These two criteria not only contribute to giving an important place to the speaker's mind and its mental processes to asses and infer the possible referent the interlocutor has in his mind at the moment of the enunciation, but they also delineate the definition of a common frame inside which every referent independent of its degree of indentifiability (i.e. be they unique or not) can be identified. In other words, they all share the cognitive property of belonging to a frame of reference, a relevant context (e.g. inter-textual or situational enunciation context).

In sum, we can delineate the following system of indentifability scale: an 'Unidentifiable' referent will be unanchored or anchored and an Identifiable one will vary from inactive, to accessible, and then active (e.g. textually > situationally > inferentially).

Clearly, the two parameters are correlated in certain ways, for instance an unidentifiable referent is necessarily outside the possible domain of the activation, since an activation state requires the existence of a mental representation in the addressee's mind. However, Identifiability and Activation can be further characterized as two independent cognitive categories:

49. Note that in the terminology of Prince the category of identifiability is defined as "anchored", while the concept of new is called with the term of "brand-new".

1. the first linked to the knowledge (i.e. the set of representations) which the speaker evaluates to have in common with an addressee, at the very moment of uttering a sentence in context, and
2. the other has to do with consciousness and the psychological factors determining the activation states of the discourse referents, responsible for discourse processing and understanding (cf. Chafe, 1974 and 1987).

In this framework, the topicalization –i.e. detachment construction following Lambrecht's terminology– is informationally defined as a grammatical device that promotes in these identifiability scale a referent from the status of accessible to the status of active, to further become a Topic at discourse level, to be talked about in the following.

Topic accessibility and continuity in Givon

In parallel to the reflection on the activation states of individual referents in discourse (cf. 'active', 'inactive' and 'accessible' in Lambrecht, 1994 and Chafe, 1987), a reflection on linguistic encoding of Topics referents was developed.

Givon (1979, 1983) contributed to this research direction by putting forward various degrees of Topic continuity across multiple sentences. Interestingly, this author correlated what he defined with the term of 'Topic discourse continuity' with the different types of syntactic elements and structures that can encode it.

As shown in figure 3.9, according to the syntactic encoding of the degrees of Topic continuity analyzed by Givon (1983), a pronoun will be usually coding for a more continuous Topic than a fully referential NP like 'John', and a Zero Anaphora will code for the most accessible referent to the interlocutors when a sentence is uttered (for an illustration see examples in Chinese (88) and in English (87), p.293).

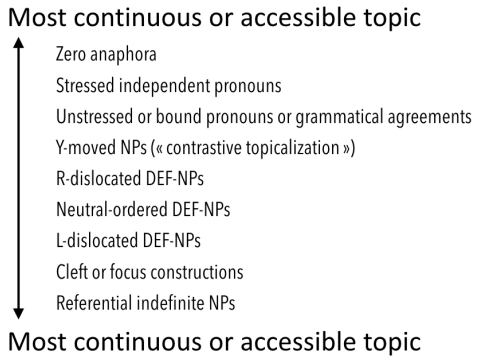


Figure 3.9 – Topic continuity Scale in Givon (1983) coding from the most continuous (top) to the least continuous Topic (down).

These considerations on the gradient nature of Topic continuity and accessibility, that is encoded by various syntactic means will be interesting for our experimental approach to Chinese topic sentences in context. In chapters 4 and 5, we will namely manipulate the accessibility of the Topic referent in the narrative context preceding the experimental sentences in order to observe the behavioral and cerebral Electro-Physiological Response (ERP) to the different accessibility of Topic referents in Chinese Scene-Setting Topics.

We can note that among the different morphological or syntactic features coding for Topic accessibility, prosodic features have been identified as carrying important information about the accessibility of the Topic referent. Hence, it is precisely this prosodic information that we will manipulate as the second parameter of our experimental approach to 'in-context' Chinese Topic-Comment sentences. In chapter 4 and 5 our aim will be in fact to uncover how prosodic information about sentence articulation and the hierarchy it conveys, can influence the relation to discourse-level information that is yielded by the sentence-discourse properties of Topics. In in broader terms this experimental attempt

will contribute to account for the sentence-discourse interface mechanisms of 'Chinese's style' Topic-Comment articulations during their on-line processing.

3.1.5 Sentence's Informational packaging and its syntactic encoding

Intermediate Interim Summary

As we pointed in previous section, the fact of considering the hearer speaker's minds and the context in which a sentence is embedded (§3.1.4), introduced in the linguistic analysis of the sentence-unit the internal point of view on the mind of the speaker and his representation of the hearer's representation of the context.

Such an approach crucially brings us to consider some cognitive processes and decisions that are prior to the formulation of the utterance: (1) after evaluating of the beliefs/representations of the hearer, the speaker (2) selects the adequate linguistic form and structure to convey his message. To this should be added, that the (3) suppositions of the speaker about the hearer depend also from the presence and the relative salience of the message's referents in the *context*, an aspect that is indeed likely to make the speaker presuppose that a given referent is more or less active in the hearer's mind: "Specifically, they [the referent statuses] all have to do with speaker's assessment of how the addressee is able to process what he is saying *against the background of a particular context* [...]" (Chafe, 1976)[50].

Given the clarifications on the given/new character that were expressed in the above section, we can now approach the theory whose aim is to formalize the way information is conveyed within the utterance, and how sentential elements are linguistically marked with respect to their assumed prominence in a given discourse-context.

3.1.5.1 Tailoring Sentences: the Context and hearer mental state into play

Information structure

Lambrecht's theory accomplishes the additional step of introducing in the utterance's linguistic analysis the idea that it linguistic encoding is influenced by the suppositions about mental state of the hearer. Borrowing Lambrecht's words, the hypothesis is formulated that the speaker's beliefs about the hearer's mind are linguistically translated in a given "*linguistic encoding*" of the utterance. This led to the following definition of Information Structure:

> 3. INFORMATION STRUCTURE: That component of sentence grammar in which propositions as *conceptual representations of states of affairs* are PAIRED with lexico-grammatical structures in accordance with the mental states of interlocutors who use and interpret these structures as units of information in given discourse contexts. Lambrecht (1994:5)

Hence, a particular presentation mode of the referent would indicate that the speaker presupposes the presence or the absence of a certain mental representation in the mind of the hearer. If we may afford using a pun the speaker may wonder if the state of mind of the hearer allows him *to-pic* or not *to-pic* a certain referent to start his sentence.

50. As a side note, exactly as Chafe, Lambrecht does consider the pragmatic circumstances only if they reach the level of construing linguistic constrains for the speaker's utterance.

Unquestionably, a cognitive approach to information structure would be concerned with those psychological mechanisms mirroring speaker's hypotheses making about the hearer's **mental states** and the hearer on-line information structure assignment to the linguistic message. In the psychology literature of the 80s, it was namely proposed that the listener structurally represents the important aspects of information, delivered by the speaker, within a mental model in which subsequently incoming information is constantly being integrated (Johnson-Laird, 1980)[51]. This idea was further developed through electro-physiological investigations by Cowles (2007).

We will focus here on the major changes in word-order and syntactic structure that unequivocally reflect the utterance's information structure, and, as argued before, we view the 'syntactic encoding' of the sentence's interface with discourse as even more relevant for both linguistic theory and a cognitive experimental investigation of the sentence structure, than simple pragmatic considerations.

Specifically, what will be the center of our experimental and cognitive interrogation are those situations in which the "*conceptual representations of the state of affairs*" are reflected in grammatical structures. Specifically, our experimental manipulation of the context of our Topic-Comment sentences in chapter 4 and 5 aims exactly at creating two different contexts and states of affairs that vary the accessibility of the Topic's referent, to be able to observe how the different conceptual representations of the sentence's context influence the acceptability of Frame-setting Topic-Comment sentence in Chinese.

information packaging

This idea, that the sentence-discourse interface has an influence on how the sentence is articulated, has received multiple and successive reformulations. Since the Prague school and Halliday (1967), the linear ordering of information at the sentence-level is analyzed as being influenced by aspects linked to information structure analysis like topic-comment and focus-background (see Büring, 1999 and 2007, for an argumentative discussion in this point). But its first clear development can be traced back to Chafe (1977:25-55), who had namely pointed to Topic-Comment structure as a particular *information packaging*:

> I have been using the term packaging to refer to the kind of phenomena at issue here [topic-comment articulations, etc.], with the idea that they have to do primarily with how the message is sent and only secondarily with the message itself, just as the packaging of toothpaste can affect sales in partial independence of the quality of toothpaste inside. (Chafe 1976:28)

Gundel (1985) will continue in this direction by defining the Topic of a sentence as the information that the speaker want the increase the listeners knowledge about. And Reinhart (1981), followed by Gundel (1988) will formalize the definition in the following terms, defining a clear cut informational division of labor between Topic and comment: "Topic element is a file card which only contributes to the naming of the entity it collects information about, while the Comment is interpreted as an expression adding information to the Topic".

The notion of linguistic Information Packaging in Lambrecht (1994/1996) is substantially developed in a syntactic direction, to the point that the author explicitly treats

51. Note how Johnson-Laird's proposal can be seen as parallel to Reinhart (1982).

information packaging of the utterance as a *component of grammar*, that enables speakers to mark an element with a specific information status with respect to the utterance context[52].

Successively, Prince (1981:227, 1986) leveraged on Chafe's packaging concept and developed its definition by introducing the image of the speaker *tailoring* the sentence accordingly to his assumptions about his interlocutor's mental states (1998:1), a tailoring invoked both in the explication of many sentence-level phenomena (Gapping, Dative, Pronominalization, Left and Right Dislocation, sentential subjects, it-clefts, wh- clefts, Topicalization, 1998: 255) as we will see in the next section and in the explication of how discourses are structured and understood (1981: 225).

The following examples from Reinhardt can be illustrative of the mechanisms we have been describing so far about the contextual motivation for choosing a particular tailoring one's sentence. Reinhart (1982:26) offers a concrete example motivating the the choice of topic-comment informational packaging and the retrieval of contextual information to understand Topic-Comment sentence.

(79) Topic-Comment informational packaging.

a. As for Carter, it's no wonder that he is considering to withdraw the American athletes from the Olympic games.

b. As for the Olympic games, it's no wonder that Carter is considering to withdraw the American athletes from them?

c. He is such a naïve politician.

d. The Olympic games were a mass and the security was not ensured.

(79a) and (79b) can seem truth-conditionally identical, but the situation described in (79a) that is that Carter thinks to boycott the Olympic Games is not astonishing of this politician, given for example that it was a Cold War American president. While in (b) the same proposition actually relates that the president choice is not astonishing of the Olympic Games, in this case, one could imagine that something particularly politically incorrect happened during the Olympic games to justify this withdrawal.

We can therefore state that the fact of being hosted in Topic syntactic position is giving to a sentence's element a *truth-conditional value effect over the whole utterance.* Erteshik-Shir (1997, 1999) addresses this issue saying that the truth-value of the propositions is to be evaluated by determining the truth of the predicate in relation to the Topic. Thsi mechanism is namely what make as understand in (79b) that something during the Olympic Games might have happened to justify the withdrawal of the American athletes.

We can further note that the discourse interface featured by Topic-Comment articulations is so fundamental/'natural' that, when sentences are out of context like in (79), it automatically yields a presuppositional calculation making us imagine, for example, that something during the Olympic Games. This shows how contextual information retrieval or world knowledge is anyway needed at the sentential level, in order to understand the relation between Topic and Comment.

52. Note that the focus here on information packaging wont be about pragmatic competence but on how pragmatic functions are to be reflected in the structuring of sentence syntax, in its interface with discourse.

Hence, the somehow major interpretative difference between (a) and (b) is syntactically yielded by the Topic-Comment articulation, and the speaker's choice of the sentence constituent to assign the Topic role and syntactic position. While in (a) the Comment-clause should be relevant information to characterize 'Carter', in (b) the Comment-clause information should be relevant for 'Olympic Games'. Further evidence for this analysis is given by the relative awkwardness of continuing these two sentence in the same manner by adding (c), which is more appropriate or natural discursive continuation of the sentence where the topic in 'Carter'. We namely see here how "local entries corresponding to sentence topics can be further organized under more global entries, thus constructing the discourse topics" (Reinhart, 1981:24).

In sum, although the given-new distinctions with various acceptations can be found on different levels of linguistic analysis -the sentence, the discourse, the participants of the discourse-model,it should be noted that at all levels -and this is probably not only universal, but also distinctive of human language- the crucial factor is that the speaker is tailoring the utterance to meet the particular assumed informational needs and of the intended interlocutor. In other words, information-packaging in natural language reflects the sender's hypotheses about the receiver's assumptions and beliefs or strategies, and shows how the sentence-unit is internally modified by its interface with discourse.

3.1.5.2 Allosentences: packaging the internal sentence's structure

Given the informational relation between the sentence-unit structural encoding and the discursive considerations about the hearer/speaker's minds, Lambrecht (1996) concretely defines the *pragmatic markedness status of grammatical structures* giving the following examples. A canonical SVO sentence, like *"She likes Germans."* which is unmarked for the feature '*argument focus*', has a clefted counterpart such as *"It is Germans that she likes."* that is marked for this feature and therefore introduces discursive considerations of the hearer/speaker mind in its syntactic encoding.

In this way, if we had to reconstruct the questions these sentences answer to, the *canonical sentence*, so-called 'neutral allosentence' can be the answer either to the question *"What kind of person is she?"* or to a question like, *"Does she like Italian or Germans?"*. On the contrary, the 'clefted allosentence' only allows narrow (or contrastive) focus interpretation, and can uniquely be the answer to the question: *"Does she like Italian or Germans?"*. It is noteworhty that these two sentence carry different propositional information also under normal prosody.

Hence, Lambrecht (1996) proposes the following definition of the *markedness* concept we are describing:

> Concerning the pragmatic markedness status of grammatical structures, we can state the following general rule : given **a pair of *allosentences***, one member is pragmatically unmarked if it serves to discourse functions while the other member serves only one of them. While the marked member is positively specified for some pragmatic feature, the unmarked one is neutral with respect of this feature. [...] This approach entails that the marked member of a given pair of *allosentences* may in turn be the unmarked member of another pair.
>
> Lambrecht (1996:17-18).

Introducing the concept of *allosentences* allows to show how variants of a the same propositional content –equal in both truth-conditions and illocutionary meaning– can

have different syntactic realizations each one conveying differences in the informational content. Crucially, differences in sentence's Information Packaging is not a difference in what is said about the world in two *allosentences* but in how the sentence's predication articulates what it conveys about the word.

Hence, comparing pairs of *allosentences* does not bring to light truth-conditional difference between to sentences, but the markedness/prominence mechanisms linked to the syntactic expression of the sentence-discourse interface. This kind of comparisons will be one of the focus of our fMRI study on Chinese Left-Periphery (chapter 7), where we will compare marked and unmarked sentences having non-truth conditional differences like (c6) and (c7) in the below Figure 7.14:

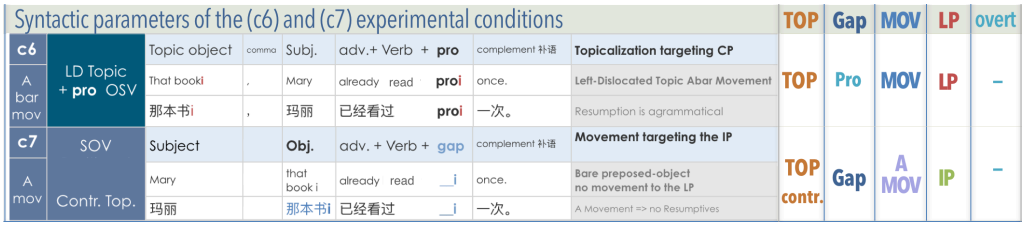


Figure 3.10 – Marked and unmarked sentences having non-truth conditional differences like our experimental conditions (c6) and (c7) in chapter 7.

1. PRE-POSING
   (a) Antonio accepted [this one].
   (b) [This one] Antonio accepted.
2. POST-POSING
   (a) I made [all the changes the supervisor wanted] without delay
   (b) I made without delay [all the changes the supervisor wanted]
3. INVERSION
   (a) [Two doctors] were [on board]
   (b) [On board] were [two doctors]
4. EXISTENTIAL
   (a) [A Jaguar] is in the courtyard.
   (b) [There] is [a Jaguar] in the courtyard.
5. EXTRA-POSITION
   (a) [That Adam is guilty] is clear.
   (b) [It] is clear [that Adam is guilty].
6. LEFT-DISLOCATION
   (a) [That money I gave her] must have disappeared.
   (b) [That money I gave her], [it] must have disappeared.
7. RIGHT-DISLOCATION
   (a) [The people from next door] are still here.
   (b) [They]'re still here, [the people from next door].
8. CLEFT
   (a) [The cat] broke it.

(b) [It] was [the cat who] broke it.

9. PASSIVE

(a) [Isabel] took [the car].
(b) [The car was] taken by Isabel].

Importantly, what these pairs of *allosentences* witness is that the marked ones in (b) imply different kinds of *syntactic operations* that are actually not triggered in order to satisfy 'purely' structural requirements –like the Case filter, agreement, or thematic structure, but by the interface of the sentence with discourse context , as we can read in Vallduví's (1992:12) attempt to clarify he role of Topic notion in a model of linguistic competence.

> "It has long been noted that there are syntactic operations which are not triggered by the need to satisfy any known 'purely structural requirement' – like the Case filter, agreement, or thematic structure –and which are logico-semantically vacuous as well. The operations include Topicalization, VP-preposing, left-dislocation, right-dislocation, adverb-preposing, gapping, it-clefting, pseudo-clefting, heavy NP-shift, and probably many others. A large number of studies, within the 'functions of syntax' approach, have pursued the task of establishing a *raison d'être* for all these non-structurally-motivated syntactic operations. This *raison d'être* is generally called the 'functional load' of a sentence." Vallduvì (1990:12)

In other words, the motivation underlying these "non-structurally" motivated syntactic operations is to be found in the fact that some constituents carry contextually relevant information. There exist, in fact, a the level of the sentence a certain set of formal properties that can make reference to the linguistic and extra-linguistic context in which the sentence is embedded. This is what we are going to call all along this manuscript *the Sentence-Discourse Interface*.

These observations about allosentences clearly bring to light in an intuitive manner the *effects* that the interface between syntax and discourse-level has on the internal structural organization of the sentence-unit. It is precisely to this kind of sentential phenomena we where pointing when introducing the architectonic metaphor of the cupola at the beginning of this chapter (p.226). Concretely, what we see in (3.1.5.2) transformations are the effects the sentence-discourse interface, producing a number of linguistic phenomena, like the fronting of a particular constituent into the Left-Periphery of the sentence, question formation or Focalization.

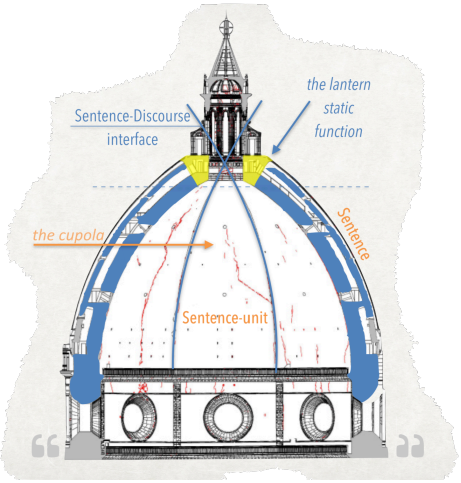


Figure 3.11 – The sentence-unit iconic metaphor of the cupola and its sentence-discourse interface (i.e. the lantern).

To summarize the crucial points made since the beginning of the chapter, we can say that the properties of interface with context carried by certain the structural articulation of sentential elements actually reflect the *informational packaging* that the speaker gives to his utterances according to his conscious representation of the referential status of the linguistic

expressions in the mind of some specific interlocutors. In fact, one of our interests in Topic-comment articulation namely lies in the possibility to study the neuro-cognitive substrates of the syntactic encoding of these interfacial mechanisms, giving to the sentence unit a room with a view onto the receivers mind and the context of the utterance.

Hence, while linguistic theories concerned with information structure investigate the way this functional properties are formally expressed into the sentential structure, the Cartographic approach maps them onto the sentence tree-structure skeleton of functional projections (cf. §2.4.4.1, fig. 2.44, p.201)[53], offering in this way a syntactic account of these interfacial phenomena.

We can conclude by resuming to our metaphor of the sentence as a cupola, and say that the above sentences in (3.1.5.2) offer a representative set of examples of how informational packaging yields different syntactic configurations of the sentence-unit. This is showing how the sentence is transformed by the interface that the sentence-unit can establish with discourse-level. Iconically, we could illustrate this by saying that allosentences namely show how the internal structure of the sentence-cupola lies under the influence of the 'equilibrium of forces' established by the sentence-discourse *lantern* (see Figure 3.11). It is namely adopting this perspective that we will now engage in the typological understanding of the Chinese sentences' articulation in the next section.

3.1.6 Summary Topic: a cognitive friendly notion

After having harvested the notional richness and complexity surrounding Topic-hood, we can concentrate on the various cognitive aspects that were brought to light, that make the Topic-Comment sentence's articulation a valuable candidate for a neuro-cognitive and experimental investigation of the sentence-unit.

The reader with a background in Cognitive Science has probably already identified some of them along this notional overview, we will address them directly in the following by summarizing the reasons that brought us to select this sentence articulation as a research object. This section contributed with a collection of circumscribed pieces of evidence for Topic cognitive relevance, it was namely argued that:

It was argued that:

- Topic-Comment articulation has been identified already by Grundel (1988) as a ***Universal*** structure across languages. Frequent in everyday communication, going from commercial advertisements to presidential talks (lately on fashion in France), its universal presence across languages makes it perfect candidate for a neuro-cognitive investigation about language sentence structures.
- Topic-Comment is proto-typically an oral construction across languages. As such it is a paramount of linguistic ***natural*** stimulus. Its functional role is grammaticalized through morpho-syntactic marking in several languages (e.g. Japanese, Korean, Russian and Hungarian in particular cases), while other languages feature an optional or international marking like Chinese.
- The psychologically pertinent concepts of accessibility and identifiability of the Topic referent, that were brought to light by linguistic analyses, show the fact that

53. As we already had the occasion to note, this approach has the advantage of assuming that specific discourse properties are directly encoded in the initial syntactic computation and force the particular elements having the same properties to move to their specifier (Spec) to realize their interpretative contribution to the sentence.

when uttering such sentences the speaker is taking into account the hearer and his ***mental representation***.

- At the level of logico-semantics, and in the philosophical account we reviewed, the notion of Topic is identified with a ***structural bi-partition*** of the propositional *predicative pattern*, that makes the Topic-Comment structure "the most general characterization of ***predicative constructions***", where "the speaker announces a topic and then says something about it." (Hockett, 1960), or where the Topic plays the role of a Frame limiting the applicability of the predication of the Comment-clause (Chafe, 1967).
- Last but not least, specific Topic-comment sentences feature syntactic encoding of the sentence-discourse interface (cf. informational packaging). This additionally makes them an cognitively interesting testing ground, to uncover the cerebral representation of constituents carrying contextually relevant information and the syntactic structure that encodes this discursive functional properties. In this way, Topic-comment sentences can be considered as the locus where the sentence functional interface with context motivates syntactic operations yielding syntactic complexity in the sentence-unit.

This listing reveals in one sentence that Topic-comment articulations are Natural, Universal, Structured and Complex Syntactic objects as the sentence-unit. To these elements, it should be added, that Topic-comment structuring or articulation is a good candidate for a typically human enunciation structure. Namely, in recent years, the kind of analyses on Topic-Comment articulation we have been reviewing have been brought to Cognitive Sciences, as illustrated by a very original and inspirational article by Manfred Krifka (2007b).

Noting that human interactions exhibit a pervasive structuring of utterances into Topic and Comment, the author points out striking similarities between the bimanual coordination and the structuring of utterances into Topic-Comment. The dominant hand can be seen as physically predicating something on the object held by the non-dominant one. To support analogy, Krifka gives arguments showing that bimanual coordination is used in the gestural expression of Topic-Comment structure in Sign Languages. This leads him to further claim that asymmetric bimanual coordination played a role in the rise of the Topic-Comment structures in communication.

From an evolutionary point of view, the author engages in the correlation between visual processing and the Topic-Comment articulation, by noting that the task sharing between the ventral stream –*the what pathway*– involved with object identification and the dorsal stream –*the where pathway*– processing spatial locations could be reinterpreted as a Topic-Comment system (cf. Mishkin et al., 1982). In this way, the *where stream* would be viewed as predicating the spatial articulation of the object identified by the *what pathway*. Although we do not subscribe to these fairly speculative remarks, they point to an informational understanding of cerebral pathways, which is a view sometimes adopted in Cognitive Science. It still holds true that animal communicative behavior sharply contrasts with he pervasiveness of Topic-Comment articulations in human linguistic behavior. Animal communication, seems to be lacking the communicative behavior to first identify an object and then comment on it, what could be interpreted as a proto-formulation of the topic-comment predication mechanism. Tomasello and Zuberbühler (2002) clearly state:

> "Virtually no ape gestures are referential in the sense that they indicate an

external entity (i.e., there is no pointing in the human fashion)."

We can retain a second argument made by the author going in this same direction. Krifka proposes an interesting reflection on a topic that has recently attracted much attention in Cognitive Science with the publication of an article on the warning calls of Vervet monkeys (Barceló-Coblijn and Gomila, 2012). The author notes that their signaling, for example, "danger from above / an eagle", or "danger from the ground / a snake" (cf. Struhsaker, 1967) manifest some form of possible compositionality, but crucially does not illustrate the sentence predicational configuration shown in the Topic-Comment articulation. For instance, these alarm calls do not *first* identify a particular region, or a certain type of animal, *and then* say something about it[54]. Similarly, Tomasello (2003) notices that chimpanzees produce attention-getting gestures, but do not appear to have any strategy for combining these gestures with other ones communicating more specific content which could be viewed as a precursor of Topic-Comment structure in human language or child utterances, like the pointing gestures of children or the non-fluent aphasic talks illustrate.

These facts on animal behavior and their interpretations, overtly contrast with what we put forward in human linguistic behavior, where the facts send still that Topic-Comment articulation is an essential feature even in the early process of language acquisition, as illustrated by French children's early use Topic-Comment sentence (De Cat and Lahousse, 2002). All this brought together makes Topic-comment articulation a good candidate for typically human enunciative structure.

To conclude this overview of the richness of satellite notions that surround Topichood (i.e. informational structure, syntactic encoding, predicative Frame, definiteness, aboutness, referent's identifiability/accessibility), we can say that even though this definitional challenge could have been easily solved theoretically just by choosing a certain level of analysis, a given field of research or framework, this would have been problematic in the perspective of experimentally addressing its neural implementation. This forced us to adopt a wider perspective and urged a notional clarification to circumscribe the cognitive aspect at play and orient our experimental methodology.

In conclusion, the different arguments and examples that were brought to light in this section showed that Topic-Comment articulations have a major cognitive relevance for the investigation of human syntactic competence.The next two sections of this chapter will address the arguments that make Mandarin Chinese an ideal testing ground for a neuro-syntactic investigation of the sentence-unit.

54. One instance of animal communication that could be remotely comparable to Topic-Comment articulation is bee communication, in that bringing pollen to the hive, which could represent the Topic, bees indicate through their dance the direction and distance where more of that pollen can be found, which in turn is the Comment.

3.2 Topic-prominence and what is basic across languages

> "[Une] Grammaire générale et raisonnée contenant [...] les raisons de ce qui est commun à toutes les langues et des principales différences qui s'y rencontrent"
>
> *[[A] General and reasoned grammar, including the principles of all what is common to all languages and the principal differences that one can observe.]*
>
> Antoine Arnault and Claude Lancelot, the so-called *Grammaire de Port Royal*, 1660.

Intermediate Interim summary

In previous section, we brought to light the notional elements contributing to the definition of the concept of Topic-hood. This overview revealed how this notion is deeply rooted in a cognitively-oriented understanding of linguistic behavior and of the syntactic articulation of the sentence-unit.

The extensive amount of satellite notion and definitions gravitating around Topic-Comment articulation witness how linguistically crucial is this sentential structure. A long list of cognitive-friendly notions have been established to characterize it: going from Topic-Comment universality, its being the most general predicative construction, to considering the speaker/hearer's mind, the concept of cognitive accessibility of a Topic referent, the calculation about shared knowledge that is necessary to chose the Topic referent. To this the idea of psychological subject should be added, together with the calculation linked to considering contextual information to understand Topic-Comment utterances, the pragmatic competence needed to infer foregrounding or contrastiveness, and the psychologically pertinent concept of identifiability. We can finally state that all these aspect linked to Topic-comment sentence-structure contribute to qualify topical phenomena as being eminently relevant for a cognitive investigation of the syntactic faculties of the *homo phraseologicus*.

Importantly, this notional trial brought us to the notion of informational packaging of the sentence-unit and to the syntactic encoding of discourse dimensions in Topic comment articualtions, so that we can define them as being both the manifestation of the essential faculty of saying something about something, and of the interface between the syntactic capacity of building sentences and all the cognitive calculations that should be implied in dealing with the sentence-discourse interface.

We will now concentrate on the contributions made to these issues by the wide range of previous analyses of Topic-comment articulation in the Chinese Linguistics and particularly on the cross-linguistic comparisons that have been made to understand the meaning and the distribution of this sentence articulation.

A Syntactic Typological approach

In the spirit of Arnault and Lancelot's epigraph, we will now adopt a typological approach parallely outlining what is common to all languages and the typological differences linguistic analyses have revealed about this way of articulating the utterance[55].

In its approach to sentence structure across languages, Typology offered a central contribution to the delineation of the notion of Topic. By pinpointing some cross-linguistic patterns linked to the prominence of this syntactic role in the articulation of the basic sentence structure of a language, Li and Thompson (1976) inaugurated a review process of a number of cross-linguistic patterns that could be better explained with the notion of Topic and advocating that certain patterns in a language were linked to the fact the basic sentence articulation was a Topic-Comment one. This notion introduced in linguistics in the 60s, interacts with syntax and has a place in the cognitive interrogations about language faculty.

Starting from the seventies, the detailed overview on Topic phenomena across languages offered by Li and Thompson, made linguists start to talk about Topic-prominent languages and Subject-prominent languages. At the same time the major role of Topic in characterizing Chinese's sentence structure in a typological perspective was revealed. Hence, the Topic-prominence typological claim applied to Mandarin Chinese: the Topic rather than the Subject is the central notion in the syntactic analysis of the sentential structure will be one of our fundamental focus in understanding topical phenomena. Sections §3.2 and §3.4 are here to show how this notion is particularly important for the analysis of Chinese and its basic articulation of the sentence's structure. We hope we will convince the reader that carrying on neuro-imaging studies on Mandarin Chinese without ceasing the opportunity of observing one of the key typological issues naturally offered by this language would be a pity.

Importantly, this cross-linguistic approach adds to the universal notion of Topic a special place inside the organization of a linguistic system, and thus inside the syntactic articulation of the sentence, which both make it an ideal candidate for a cognitive investigation of the sentence's structural articulation.

Adopting this sentence-based approach, we will concentrate on showing how sentence as-a-syntactic-unit is not only an*extraordinary* way to syntactically package information out of smaller syntactic units (i.e. constituents, cf. §2.3), but it is also characterized by the possibility of having a discourse interface inside it, that is a powerful mean to anchor the sentence in an articulate discourse or communicative space[56] and reveals many aspects of its internal structural organization (as theoretically introduced in §2.4).

Given the very rich notional background linked to Topic-hood that was over-viewed in previous section, the issue of "appropriate level of representation" will be important in the linguistic characterization of topical phenomena we address in the following. Our syntactic focus will therefore mainly restrict our attention on '*sentence Topics*', addressing only marginally issues liked to Discourse-Topics or semantically-oriented analyses. This will allow us to give an overview of the typology of topical phenomena present in Mandarin Chinese, considering their syntactic characteristics, and to conclude by show-

55. As for Antoine Arnault and Claude Lancelot one has to remember that the twin book of the *Port-Royal Grammar* is the *Port-Royal logic*: here again surfaces the preoccupation for a approach of the grammar of sentence-unit considering logic.

56. Krifka (2008) argues that this particular aspect of human linguistic ability (i.e. sentence-discourse interface) is what animals can not achieve.

ing which syntactic and linguistic aspects motivated our choice of making this syntactic articulation the object of our experimental approach.

Paralleling the structure of chapter 2, this linguistic chapter will show how Mandarin Chinese offers a series of suitable linguistic configurations in different types of Topic-comment sentences for our theoretical neuro-linguistic interrogations about:

1. the building of sentence's articulation and hierarchy under minimal marking conditions, in §3.2.2 and 3.2.3;
2. the representation and processes linked to the Sentence's Interface with Discourse, in §3.4.5;
3. the representation of syntactic-tree's complexity metrics, in §3.4.2;
4. and the establishment of Antecedent-Gap dependency-link, in §3.4.4.

Hence, the main theoretical neuro-linguistic issues exposed in chapter 2 will resumed to in order to show how Chinese constitute an optimal testing-ground our experimental hypotheses.

3.2.1 Topic-Comment structures across languages

In every day language use, the communication between speakers of a given language stands on a certain amount of tacit understandings about a certain number of linguistic aspects, such as rhetorical structures, social registers, and some commonly used syntactic patterns that determine the linguistic form propositions and clauses are given.

Among this type of structures used to convey propositional content that constitute the core of the linguistic knowledge of the speakers of a given language. we namely Topic-Comment articulations.

These syntactic patterns and structural rules are central to language system and are convergently expressed in a basic linguistics unit -the sentence. This unit and its articulation are central to the study of language faculty and of the *Homo Phraseologicus* we depicted in chapter 1.

3.2.1.1 Topic-comment articulation of the sentence is universal

In this respect Topic-prominence plays an important role in the organization of the sentence's articulation, and while Topic-Comment articulation is considered a universal linguistic phenomenon and an important feature of all human languages, the syntactic forms it takes in particular languages are quite diverse.

In many languages there are specialized syntactic constructions to express the syntactic relationship between the Topic and its Comment. Namely, the Topic-Comment sentence articulation is encoded with various formal linguistic devices in the grammar of a language, through (1) morphological markers (cf. §3.2.3.2, Topic markers), (2) syntactic structures and (3) intonational patterns (Gundel, 1988:216).

However, the use of syntactic structures are the most frequently used device to encode a Topic-Comment across languages (Gundel, 1988:223). And the tendency to place Topics in sentence-initial position also represents a common denominator to many languages particularly in the oral register.

Before showing how the Topic-Comment articulation plays a central in the organization of Chinese sentence, a few English examples can introduce us to the universality of the syntactic configurations we will meet in this section dedicated to Mandarin Chinese.

mlkjhg Consider the series of sentences presented in chapter 2 (§2.2.4) in (28) reproduced here for convenience in (80a, b and c).

(80) a. As for American self-confidence, Columbia gave people a lift.
b. * American self-confidence, Columbia gave people a lift.
c. * It was American self-confidence that Columbia gave people a lift.

In (69a and b) the agrammaticality is yielded by the failure to assign a semntic/fucntional role to the sentence-intial NP. Thus, the agrammaticality patterns shown (69) suggests that English syntactically encodes the aboutness Topic role by morphological means -'*As for*'- whenever the Topic cannot possibly play the role of an argument of the verb in the comment as in (81)[57], and not positionally by just by having a bare NP in sentence-initial position as in (69b), where the topic NPs cannot acquire a semantic role given the fact that they have no selectional restriction with the verb.

(81) a. Topicalization
Spaghetti on toast I find ___to be an absolutely revolting way to begin the morning.
b. Left-Dislocation
Word-processor, I sometimes think *they* should be recycled into Space Invaders machines.
c. Des voyages, Olivia Bransbourg en a réalisés de nombreux.

These examples can be taken to indicate that the grammar of English does not specify any association between topic sentence-initial position and a semantic mapping be it an aboutness or a frame-setting relationship. Topics that are not introduced by '*as for*' preposition are grammatical only when the topic is semantically or grammatically bound to an element in the comment as in It-cleft, Topicalization (81a) and Left-Dislocation (81b and c).

In fact, the topics in (81) acquire a semantic role in the sentence by virtue of being co-referential with a covert or overt pronoun in the comment-clause, which is selected as an argument of the verb 'think' in (b) and 'find' in (a).

Contrastively, the Mandarin Topic-Comment sentences in (82) are representative of the same 'aboutness' relationship the Topic entertains with the Comment as in the above English examples (69) (they can in fact only be adequatly glosed in English by using 'as for' or 'speaking of', but do not need any morpho-syntactic marking making the aboutness relation between topic and comment explicit.

(82) Topic has no selectional relation with the verb
a. 那场火，幸亏消防队来得快。

Nà-cháng huǒ, xìngkuī xiāofángduì lái de kuài.
that-CL. fire, fortunately fire-brigade come DE quick

'That fire fortunately the fire brigade came quickly'
(Li and Thompson 1978)

b. 那场火，消防队来得晚了。

Nà-cháng huǒ, xiāofángduì láide wǎn le.
that-CL. fire, fire-brigade come late ASP.

57. Note that "*As for*" construction is often analyzed as a switch-Topic device (Chafe 1976:49).

'That fire the fire-brigade came late.' (Shi 2000, ex. 41)

c. 那场车祸我们太幸运了。

Nà-cháng chēhuò wǒmen tài xìngyùn le.
that-CL. Car-accident 2PL. too luncy Prt.ASP.

'In that car accident we had such luck.' (Her 1991)

While the grammaticality of "*As for*" marking is sometimes promoted as a test for the aboutness relation established by "topic-hood" (Gundel, 1974, 1985; Reinhart, 1981), the grammaticality of the Chinese sentences in (82) shows that Chinese does not need to introduce the kind of aboutness topics in (82a) with the preposition '*as for*'[58].

Hence, a second fundamental difference between these English and Chinese Topic-Comment sentence is that in Chinese bare topics are licit even when no selectional relation with the verb can be established. Namely, according to the abundant research on these type of topical phenomena in Chinese linguistics (Li and Thompson, 1981; Shi, 2000; Tsao, 1979 just to cite the early works), the Topic must be related to the comment semantically, but not necessarily grammatically. This last type of Topic-Comment structure will retain our attention in next chapter and will actually be the object of our ERP study on Chinese sentence processing in context.

Thus, the major distinction between a Topic and Subject here is that, the subject has a grammatical relation with the predicate, while this is not required for a Topic. The Topic is related to the comment semantically, and may or may not be grammatically related. In English this And as showed by sentence (69c) forcing the selection of the topic by the verb results in agrammaticality.

Although these kind of sentences in English or French are fairly frequent, they are nonetheless generally regarded as restrictively representative of the oral and expressive register and (69a) have anyway a minor place in the linguistic system of English, while the above Chinese Topic-Comment (T-C) articulations are non-marked and show how this sentence articulation plays a central in the organization of Chinese sentence, to the point it has been considered as the predominant form of sentence in ordinary use:

(83) a. Every day use if T-C sentences in Mandarin: 今天很冷。

Jīntiān hěn lěng
Today very cold

'Today it's very cold.'

b. Multiple Topics: 今天城里有事。

Jīntiān chénglǐ yǒushì
Today in-town have-business

'I have buisness in town today.'

c. Verb-less T-C sentence: 这辆车，两门。

58. For this reason (among others), we opt for frame-setting definition of Chafe (1976) to better describe the relations between topic and comment: The function of topic is that 'the topic sets a spatial, temporal or individual framework within which the main predication holds' (Chafe, 1976). Moreover, we won't rely on this diagnostic also because in Chinese this kind of preposition serve to mark a particular type of Topic implying a Topic-shift and introducing a new discourse topic, and as shown in §3.1.3.4 (p. 252) for '*as for*' Topics (Jacobs, 2001).

Zhè-liàng chē liǎng mén
this-CL. car two doors
'This car has two doors.'

These verb-less sentences mainly bring to light a major difference between English and Chinese in the organization of the sentence. As we already argued for example (69), English the verb is characterized by being the semnatic core of the clause, setting up the propositional frame, while Chinese can actually build propositional content without having a verb in the comment, as shown in the above common Mandarin sentences (83). we could therefore say that in Chinese topics and comments can be filled by phrases of any major lexical category and verbs are not necessarily required in the comment.

Having given these introductory overview on the extreme difference between Topic-Comment sentences and their somehow extreme development in Chinese, we can conclude by saying that all in all these comparisons indicate the fundamental difference between topic and Subject, which we could summarize with the words of Li and Thompson (1981) before addressing their typological claim in the next section.

"The topic at the clause-level is 'typically a noun phrase that names what the sentence is about, is definite generic, occurs in sentence-initial position and may be followed by a pause or a particle"
The subject is "the noun phrase that has a 'doing' or 'being' relationship with the verb in the sentence" L

3.2.1.2 The notion of Topic-Prominence: a typological claim

As we outlined the Topic was first proposed as a linguistic concept that could exhibit at first sight some similarity to the notion of subject, but that in the topic-comment pair "suggest the most general characterization of predicative construction" (Hockett, 1958:201), where the Comment is described as the predication about the Topic.

Successively, Chao Yuanren's Grammar (1968:67-104)[59] applies it to the analysis of Chinese sentence-structure:

"The grammatical meaning of subject and predicate in Chinese sentence is topic and comment, rather than actor and action."[60]
Chao Yuanren, 1968:69

Based on the fundamental difference between Subject and Topic, Typologists from the 70's identified major differences between languages in their sentential-level articulation according to their strategies in building sentence structure relying either on the subject-predicate relation or on the Topic-Comment relation,

Li and Thompson (1976) show that the basic sentence type in languages like Chinese Lahu and Lisu presents a Topic-Comment articulation instead that subject-predicate one, and argue that :

"This evidence shows not only that the notion of topic may be as basic as that of subject in grammatical descriptions, but also that languages may differ in their strategies in construction sentence according to the prominence of the notion of topic and subject." Li and Thompson, 1976.

59. Chao Yuanren's Grammar, whose title by the way already announces a focus on orality *A Grammar of Spoken Chinese* Zhongguohua de wenfa 中國話的文法 (1968).

60. Chao Yuanren [Zhao Yuanren 趙元任]. A Grammar of Spoken Chinese. Zhongguohua de wenfa 中國話的文法. Berkeley: University of California Press, 1968.

Hence, the typological claim was first made on descriptive grounds by simply proposing that certain linguistics facts in those languages (we will review them in section §3.2.3) could be better explained in the light of Topic-Comment relation. We can namely read in the following to statements this first descriptive claim:

> "Our typological claim will simply be that some languages can be more insightfully described by taking the concept of topic to be basic, while others can be more insightfully described by taking the notion of subject as basic." Li and Thompson, 1976.

> "One of the most striking features of Mandarin Chinese structure, and one that sets Mandarin apart from many other languages, is that in addition to the grammatical relations of "subject" and "direct object", the description of Mandarin must also include the element "topic". Because of the importance of "topic" in the grammar of Mandarin, it can be termed a Topic–prominent language."[61] Li and Thompson (1981:15)

The next two sections are dedicated to present a handful of linguistic evidence that led to this typological perspective, and to demonstrate how considering the Topic notion allows to explain a number of structural aspects of the linguistic and syntactic system of Topic-prominent languages, and particularly of Mandarin Chinese.

Given this framework, two types of languages were isolated: Topic-prominent languages like Chinese, and Subject-prominent languages, like English. Li and Thompson (1976:459) introduce a certain graduality between the two poles based on the opposition between the subject-prominent and topic-prominent parameter. This led the authors apply four basic typologies to 30 world's languages with regard to their preferred strategies in constructing the sentence-unit:

1. Languages that are subject prominent (Li and Thompson, 1976: 459): English, Indo-European, Niger-Congo, Finno-Ugric...
2. Languages that are topic prominent: Chinese, Lahu (Lob-Burmese), Lisu (Lob-Burmese)...
3. Languages that are both subject-prominent and topic prominent: Japanese, Korean...
4. Languages that are neither subject prominent nor topic prominent: Tagalog, Illocano...

Ever since this founding paper on Topic-prominence by Li and Thompson[62] , Topic has been systematically identified in the literature as a key notion to analyze Chinese and Topic structures have been widely discussed (Chao, 1968; Huang, 1982; Huang et al., 2009; Li, 2000; Li and Thompson, 1976, 1981; Shi, 2000; Xu, 2000; Xu and Langendoen, 1985). Only later the issue about its function at different linguistic levels came the trend

61. Our experimental approach on topic-comment articulation in Chinese could also be seen as an experimental verification of the initial typological claim stating that in Chinese Topic rather than Subject is the central notion in the analysis of the sentential structure. Seeking to verify, understand and tentatively justify it experimentally.

62. We should note here that the founding paper by Li and Thompson (1967) has not been the first to suggest that the notion of Topic was central in Mandarin's basic organization of the sentence-unit. Several authors have been paving the way to Li and Thompson's analysis by mentioning the notion of Topic as being explanatory of a number of linguistic phenomena (see James Tai, 1973; Vivianne Alleton (1973) and Huang, 1976). Specifically, James H-Y. Tai in his Paper presented at 1973 Annual Meeting of Linguistics Society of America in 1973, titled "On the center of predication in Chinese verb-complement construction" and Viviane Alleton in her 1973's "Grammaire du chinois".

to consider and analyze Topics at the interface with discourse[63], under the influence of Lambrecht (1995) informational analysis, and particularly of Kiss (1995) for Hungarian, followed by Liu Danqing and Xu Liejiong (1998) in Chinese.

While research along the contrasting views of pragmatics and syntax on topicality continues to generate debates among the tenants of the pragmatic or informational analysis and those of the syntactic and typological view, the analysis of this sentence articulation in the formal framework of generative linguistics, clearly posits the Topic as a syntactic position in the sentence-domain dedicated to discourse interface (i.e. the CP), which has the advantage of harmonizing these points of views and analyses, actually revealing the rich linguistic functional characterization of the Topic-Comment articulation of the sentence.

Our study will promote a reconciliation among these views to generate precise neurolinguistic experimental hypotheses about the internal organization of the sentence-unit. Importantly, Chinese is not only considered as typologically prone to structure sentence in relation to the interface with discourse. That is to say that not only the notion of Topic plays a role at the discourse-level, as it is the case in any language, but also Topic has major manifestation at the level of the articulation of the sentence-unit.

Importantly, next session addresses typological and syntactic analyses showing how certain typological characteristics of Chinese topical phenomena do not allow to derive Topics from other basic sentence types and stress their non-marked (non-emphatic) use in other words, arguing for the "syntactic basicness" of Topic-Comment sentential articulation in Chinese.

Moreover, many researcher underscore that this typological feature is not only to be observed in the sentence structure, but also in *inter-clausal* organization (Chu, 2006). The so-called *Topic-chain* has been defined as a multi-clausal unit of text, where clauses are not linked by conjunctions but by co-referential relations between overly expressed NPs and implied zero NPs. These functional units in Chinese will not be the object on any further analysis here, but they surely are one of the most difficult feature to be taught to intermediate L2 learners and to be translated into foreign languages, and will be therefore considered as sign post for Topic-prominence in the next section dedicated to some psycho-linguistic aspects of Topic-prominence emerging for the study of second-language acquisition. Inter-language linguistic behavior will be now explored to gain some cognitive insight on what it means to structure a sentence into a Topic-Comment articulation.

3.2.1.3 Topic-prominence in Second Language acquisition

A universal topic-comment stage in Language Acquisition

Leaving on a side the debate on a universal Topic-Comment stage in First Language Acquisition, which could also be linked to pointing behavior of young children[64], we can briefly refer back to the French examples that Information Structure[65] (cf.§3.1.3.6, p. 76)

63. There does not seem to be any contradiction in affirming that the fundamental role of Topic notion in Chinese, for example, works at various linguistic levels of analysis.

64. See Clarke (2003) for an interesting interpretation of pointing where the author argues it is linked tot the predicative necessity of first setting what is salient to them to then say something about it: "Children focus initially on conveying what is most salient to them.". For a predicative analysis of French early Topic-Comment utterances see also Danon-Boileau and Morgenstern (2009).

65. Information Structure: information about the flow between speaker and hearer.

to say that Topic-comment constructions are indeed mastered early and use pervasively during the first stages of First language acquisition.

As for the Topic-comment patterns observed in second language (L2) learning (i.e. inter-language), a general trend has been observed in L2 adult learning process over time: learners show a progressive acquisition of the target language sentence structure, starting from utterances that are mainly characterized by loose grammatical knowledge and the structuring of the sentence-unit according to a Topic-Comment articulation.

To verify this locally attested observation and hypothesis, a large scale study was conducted for 30 months by the European Science Foundation on a hundred immigrants (Perdue, 1991). Shortly after arrival in different European countries (France, England, the Netherlands, west Germany and Sweden), the every day conversation with coworkers and researches interviews of the immigrants showed that their general utterance structure at early stages of acquisition was characterized some typical learners strategies, such semantically evident and transparent form to meaning mappings and simplification of complex sentential relationships.

Most strikingly, despite the huge inter-individual differences of the learning profiles, and independently from the different learner's mother tough (i.e. the source and target language variation in morpho-syntactic marking, word inflection, basic word-order, etc.), their linguistic production showed an over-all Topic-Comment sentence pattern (Klein and Perdue, 1989; Perdue, 1991).

The universal L2 Topic-Comment stage and Topic prominence.

Since then, early stages of L1 and L2 acquisition have been typically and cross-linguistically characterized by the learners' production of Topic-Comment utterances. Yet, the question may rise about what happens when a learner of a non-Topic-prominent start acquiring a Topic-prominent one, does the general tendency to utter Topic-Comment sentence in L2 first learning stages facilitate their acquisition of the Topic-prominent language?

Learner L1 typological Transfer on L2 A restricted number of studies directly addressed the above question about the existence of a 'Typological transfer' in L2 learning. Among the most significant contributions to this issue, we can cite for instance Fuller and Gundel (1987), who conducted an experiment to measure the oral and written Topic-prominence or Subject-prominence of the linguistic production of different populations of L2 learners of English.

In the written part of the experiment, participants from the Topic-prominent mother-tong group (i.e. Chinese, Japanese and Korean), and Subject-prominence mother-tong (i.e. Arabic, Farsi and Spanish) group were asked to write narratives in English that were then compared with those from a native English control group.

The results were analyzed according to the seven criteria (i.e. Topic-Comment features) by Li and Thompson we presented in Table 3.7 (§3.1.3.5, p. 262), and clearly show that regardless of the L1 linguistic background of the learners there were more Topic-Comment features in their English written production, compared to that of English native speakers.

Moreover, Subject-Prominent Arabic, Farsi and Spanish learners of English used Topic-Comment articulation not only in their written expression, but Topic-Comment

encoding was also frequent in their spoken utterances, which were also lacking regular subject-verb agreement and presented Left-dislocations like: '*The car he wants to eat the bird*".

In conclusion this results suggest that there exist a universal second language acquisition Topic-Comment stage that does not depend on the topic-prominence parameter of the First language of the learner[66].

The question may now raise of what are the inter-language patterns featuring the shift between Topic-prominent language and Subject-prominent language and viceversa.

L1 Mandarin L2 English

A progressive shifting towards the Subject-prominence parameter can be observed in the following examples in (84) from Rutherford (1983). These learner's utterances show an initial Topic-prominence stage in (a), transforming into a mixed stage where topic and subject roles are differentiated at the intermediate stage in (b), to finally blend into a target-like (i.e. English) utterance structure where an overlap strategy between Topic-Comment and Subject-predicate sentential organization is observable in (c).

(84) Progressive setting of subject-prominence in Mandarin L2 learners of English from Rutherford (1983)

a. Level 3 - Advanced beginner:
[In my country man and women chooses husband or wife] [no subject] [comment is very simple].

b. Level 4 - Intermediate:
[hoose a good husband or wife], [this subject] is very important problem for everybody.

c. Level 5 - Advanced:
[choosing husband and wife] [is one of the essentials of life]

By examining the English written production of topic-prominent speakers of Chinese, Japanese and Korean, Rutherford detected an overproduction of topic sentences by topic prominent speakers, especially in Chinese speakers, and the gradual increase in then number of syntactic and morphological rules of English.

Interestingly, at different L2 proficiency levels are delineated distinct stages:

1. Level 3 (a) features the presence of complex sentential topic 'In my country man and women chooses husband or wife', and no syntactic category of subject.
2. Level 4 (b) displays the realization of a grammatical subject 'this' being co-referential with the Topic 'choose a good husband or wife'.
3. Level 5 (c) is characterized utterances in which the topic functional role is reanalyzed as a subject 'choosing husband and wife'. A structure that might be preferred as a middle may between SVO English features and Topic-Comment Chinese one.

The results from a translation study study by Chen Caicai (1989) replicate the progressive departure from Topic-prominence, and the once learners get to an advanced-level. Table 3.3 shows how the present in Topic position of different Phrases gradually decrease their written production.

66. See also Bates, 1976; Huebner, 1983 and Givon, 1979.

	Low level of English proficiency group		High level of English proficiency group	
Type	Number*	Percentage	Number*	Percentage
NP-Top	160	26.67%	36	6%
Clause-Top	114	19%	64	10.67%
VP-Top	40	6.67%	30	5%
PP-Top	64	10.67%	122	20.33%
Total	**378**	**63.00%**	**252**	**42.00%**

Table 3.3 – The number and percentage of the sentences with TopPC (topic prominence construction). *Total number of each group is 600. Adapted from Chen Caicai (1989).
NP-Top = noun phrases as topics; Clause-Top = clauses as topics; PP-Top = prepositional phrases as topics; VP-Top = verb phrases as topics.

Note that Preposition Phrases that are semantically typical Frame-setters across languages do not show the same decrease. This last point critically indicates that what is signifieds in these evolving patterns is the shift toward target language parameter, and not just an homogeneous reduction of Topic-prominence parameter.

L1 English L2 Mandarin In the nighties, three studies focusing on Mandarin Chinese learning (Xie 1992a; Jin, 1994 and Polio, 1995), addressed the issue of subject/topic-prominence transfer by quantifying more advanced criteria characterizing the Topic-prominence of the Chinese language system, like the presence of dummy subjects or the use of *zero NPs* across L1 Mandarin speakers and L2 English learners.

The fine-grained linguistic approach of these studies was much more detailed than previously performed investigations that only considered Topic-Comment articulation in broader terms.For instance, the set of criteria chosen by Xie (1992) to measure and evaluate the various degrees of subject-prominence (i.e. the presence of dummy subjects) and of Topic-prominence are linked to a number of grammatical constructions that although particularly representative of the Mandarin grammatical system already require an advanced-level of mastery of Chinese language in that these Topic-comment configurations feature a wide range of syntactic and semantic characteristics[67]:

1. Topicalization:
2. zero anaphoras in successive sentences:
3. double nominative constructions:
4. left-dislocations:
5. cleft constructions:

It should be noted that these syntactic criteria are highly advanced in that the above structure respectively require:

67. Similarly, the study by Jin (1994) also focused a set of criteria to evaluate the shift from one prominence parameter to the other selecting grammatically advanced knowledge in both directions, like specificity/definiteness marking, article use for subject-prominence and zero anaphora and double-nominative constructions for topic prominence.

Table 3.4 – Percentage of NP forms used for non-initial-mention referents by Chinese native speakers and L2 learners of Chinese.

	L2 learners of Chinese (Polio, 1995)			
Proficiency levels	low	mid	high	**Chinese native speakers (Chen, 1986)**
Zero NPs	11.6%	17.4%	20.6%	37.7%
Pronouns	44.5%	31.5%	37.7%	35.8%
Full NPs	43.9%	50.1%	41.5%	26.5%

1. for topicalization and clefts, some naturally complex sentence-level transformations and word-order changes for L2 learners ,
2. for double nominatives[68] featuring part-whole relationships, the mastery of some complex semantic features in learners whose lexical vocabulary size is still minimal, and
3. for zero anaphora or co-referential assignment in Left-Dislocations,the deletion of overt infra-sentential dependency-links in topic chains as required by Chinese.

Xie (1992) shows that at the beginning of their learning English speakers use the same percentage of zero NPs in their L2 Chinese production in English (respectively 4,49% and 4,09%, and as their proficiency increases the percentage of use of zero NPs progresses too, reaching a ceiling at 17,94% while Chinese native speakers are at around 34,07%.

The hardship in mastering Topic-controlled deletion strategies in Chinese multi-utterances units is shown in the results from the three story-telling studies by Xie, Jin and Polio that unanimously characterize L2 difficulty in tracking reference at the level of multi-utterance discourse units, as show in the Table 3.4.

All in all, these studies generally show a tendency for English learners of Mandarin to transfer subject-prominence, by adding dummy subjects (e.g. wo 'I' pronouns are added to every sentence)[69] to their utterances and by over-generalizing the SVO word-order in utterances where Chinese native speakers would adopt a Topic-Comment articulation as illustrated by the following examples from Li (1996/1999): (85).

68. To these considerations should be added some evidence for Chinese children linguistic developmental patterns, showing for example that double nominatives, is one of the last type of topic comment structure to emerge in spontaneous Chinese children production at about the age of 4, while topicalized sentence appear early at the age of 2;2, (cf. §3.2.4.2, p. 325).

69. An alternative account for the overuse of pronouns in L2 learners of Chinese comes from Polio (1995), who interpret them as place-holders to pause on while thinking about the rest of the sentence. I would also argue in this direction for French learners of Chinese in my Chinese classes. Students and learners typically utter very long first person pronoun, emitting a sound like "woooooo". This linguistic inter-language feature echoes even remember a Chinese joke was done in a Chinese new year's eve TV show precisely on the foreigners tendency to pause on first person pronoun. A foreigner interacting with a Beijinger is starting his interaction by a long "woooo" that is interpreted by the Beijinger as a sign that he wants to shake hands -like he was taught foreigners like to. Namely, the Chinese for 'shake hands' sounds like the first person pronoun "wo + shou". The misunderstanding is set, in that the foreigner was just reflecting about how to formulate his sentence: "Wo shi laowai", meaning "I"'m a foreigner' using a pejorative expression Chinese people use to define foreigner. The joke namely continues by the foreigner not wanting to shake hands, the beijingers taking umbrage that he didn't want to shake hands with him and the laughter for the use of a pejorative expression to start an interaction with. Anyway, this scene was very funny, and represents a argument found in Chinese popular culture about the characteristic long pronominal pausing of the so-called "*foreigner talk*" in Mandarin.

(85) Utterances of English learners of Chinese vs Native Chinese

a. Chinese native utterance: 那辆车我买了。

Nà-liàng chē wǒ mǎi-le.
that-CL. car I bought-PERF.

'I bought that car.'

b. Chinese learner utterance: 我买了那辆车。

Wǒ mǎi-le nà-liàng chē.
I bought-PERF. that-CL. car

'I bought that car.'

Although this patent discrepancy with the hypothesized early Topic-prominent stage in L2, by advancing that that the relative lack of morphology in Mandarin transfers learners attention on lexical, phonological and word-order features[70], it should be noted that the detailed criteria these studies chose, might be the reason for the fact that their results are in overt contradiction with those revealing a universal Topic-Comment stage across a wide range of languages and learners profiles.

The above studies clearly do not consider Topic-Prominence features focusing on highly grammaticalized ones that mainly reside on advance grammatical knowledge of the target language at a multi-utterance level.

Several studies by Li Wendan (PhD 1996 and subsequent work see Li, 1999), have considered these advanced characteristics of Topic-Comment constructions as distinct from the basic ones found in studies calming for a universal Topic-prominent stage in early L2 utterances, like: (1) lake of morphological marking on subject-verb agreement and (2) Topic-comment utterance strategy. Her studies reported a greater number of topic-prominence criteria linked to these basic topic prominence-features in beginners and distinguished between those that where present in early L2 acquisition of Mandarin -that she referred to as Topic-comment features- from those present only later in the learning process -that she names as Topic-prominent features. These arguments among others bought us to add several experimental condition in our fMRI study on Chinese Left-Periphery (chapter 7) enabling us to distinguish between these two main set of features.

In conclusion, the various pieces of evidence collected in this section on second language acquisition of topic-prominent languages like Chinese seem to confirm the Interface Hypothesis by Sorace and Filiaci (2006), which is proposing that grammatical properties at the interface between syntax and other cognitive domains cannot hardly be fully acquired and are often prone to variability even in the final state of L2 grammars. In our case, Antonella Sorace's proposition would namely suggest that the acquisition of the syntax-discourse interface involved in Topic-comment sentences is indeed difficult to obtain even at the final stage of Chinese learning, a pattern that seems confirmed in the experiments we reviewed, and progressive setting of Topic-Prominence illustrated by examples below (84) and by personal experience. The difficult task is namely to acquire the more complex topic-prominent features corresponding to typically Chinese syntactic structures that are typically Chinese (i.e. double nominatives, etc.), not even mentioning the management of topicality at the level of multi-utterance.

70. These can be resolved by considering the above reviewed acquisition data from native Mandarin speakers in English (cf. example 84).

3.2.2 The basicness of Topic-Comment constructions

As argued through the comparison between English and Chinese, Topic-Comment construction can be considered as a universal phenomenon even if its role inside the linguistic system across languages is subject to important variations. One of the most important variation linked to topic-prominence is to be observed in the different degree of basicness this sentence articulation has in the linguistic system of a language.

Hence, this constructions can either be the result of a stylistic/rhetoric/pragmatic operation emphasizing one element of the sentence, and altering the basic canonical word-order on emphatic grounds, or be "a sentence favorite type", as Hockett (1958:202) had defined it in Mandarin Chinese.

It has been widely observed that Mandarin generally tends to organize sentences into Topic-Comment structures, and Chinese "sentence topics" are not necessarily interpreted contrastively nor emphatically. The use of Topic-comment articulation is so pervasive that even lexicalized element like four-characters idiomatic expressions -*chengyu* - have been characterized as featuring Topic-Comment patterns. Consider the following two in (86):

(86) Topic-Comment *Chengyu*

a. 等级森严 [T-C] pattern

děngjí-sēnyán
rank-strict

'rigidly stratified'

b. 他口蜜腹剑 [T-C + T-C] pattern

tā kǒu-mì-fù-jiàn
he mouth-sweet-breast-sword

'He is hypocritical.'

This emphasis made on the basicness of this sentence articulation in Chinese is central to our neuro-linguistic approach. We will argue that this particular linguistic configuration permits to test in a non-pragmatically marked utterance the cognitive process linked to the cerebral representation of the hierarchy that it featured by the Topic-Comment articulation.

3.2.2.1 Basicness and Frequency of T-C articulation in Chinese

One of the most remarkable properties of Topic-prominent languages compared to Subject-prominent ones is the extent to which Topic-Comment sentences are to be considered as belonging to the repertoire of most basic types of sentences[71].

Since Topic-Prominence claim was first made in the 70's, typologists identified major differences in the sentential-level articulation between languages. However, the inaugural characterization of cross-linguistic Topic-Comment constructions made by Li and Thompson did not consider the fact that Topics in Chinese often extend their semantic

71. Anticipating on next section, developmental studies on language acquisition in Chinese Children show that children at early age of 3 perform equally well on OSV topicalized sentence than SVO ones in picture-sentence matching task, and reach ceiling performance at the age of 5 (Hu Shenai's PhD, Exp.5, 2015). This findings suggest that children at age 3 have already set and master the topic-prominence parameter of their language.

domain over several clauses to form a so-called Topic-chain, by which they control what has been called Topic-NP deletion in the chain (for a complete overview of topic chains see Wendan Li, 2005).

In fact, in Chinese the notion of Topic plays such an important role in the organization of the sentence-unit that even textual organization overwhelmingly bears the mark of Topic-prominence. As we saw for Chinese acquisition as a second language the use and management of empty syntactic element (i.e. zero anaphoras) at the multi-utterance level was unanimously reported as the most difficult aspect to master for learners of Chinese (cf. §3.2.1.3, Table 3.4, p. 290).

Consider the following sentences and their English translations in a contrastive manner[72]:

(87) "That car is too expensive. (b) **The** color is not good **either**. (c) I don't like **it**. (d) Yesterday I went to take a look. (e) I also drove **it** for a while, (f) but still didn't like **it**. (g) **And** I didn't buy **it**."

(88) (a) 那辆车，价钱太贵，(b) 颜色不好，(c) 我不喜欢。(d) 昨天去看了，(e) 还开了一会儿，(f) 还是不喜欢，(g) 没买。

(a) nà liàng chē$_i$, ___ jiàqian tài guì, (b) ___ yánsè bù hǎo, (c) wǒ bù xǐhuan
(a) That CL. car$_i$, Ø$_i$ price too high, (b) Ø$_i$ color not good, (c) I$_j$ don't like
___. (d) ___ zuótiān qù kàn le ___, (e) ___ hái kāi le yī huǐr ___, (f) ___
Ø$_i$. (d) Ø$_j$ yesterday go see PRT Ø$_i$, (e) Ø$_j$ also drive PRT one moment Ø$_i$, (f) Ø$_j$
háishi bù xǐhuan ___, (g) ___ méi mǎi ___.
still not like Ø$_i$, (g) Ø$_j$ did.not buy Ø$_i$.

'That car [its] price was too high, (b) [its] color not good, (c) I don't like [it]. (d) [I] yesterday go see [it], (e) [I] also drive [it], (f) [I] still not like [it]. (g) [I] did not buy [it].'

The Chinese version in (88) shows a certain systematic trend in featuring empty syntactic elements (i.e. zero-anaphoras) in both subject and object position compared to the English. Namely, the difference between English and Chinese utterances mainly resides in a generalized Chinese strategy to drop clausal elements across different clauses, replacing those by a covert encoding of the syntactic dependency-links between the empty elements marked by Øand their antecedents (cf.co-indexation)[73], that English overtly expresses adopting a pronominal resumption strategy.

The Mandarin strategy shown here has been analyzed as linked to Topic-prominence, in that the multiples gaps following the Topic '*that car*' are implicit zero-pronouns co-indexed with the topic in sentence (a)(Li and Thompson, 1981 and later developed by Chu's Discourse Grammar, 1998). Hockett (1958) had already noted that repeated Topics (subjects) can generally be omitted without yielding agrammaticality[74].

While the English counterpart always requires the transitive verb to have their object, the Noun Phrases to have their determiner (see bold marking) and connectors like the

72. Parenthesized letters are added for the convenience of the discussion.

73. Note that the generalized occurrence of Zero NPs in Chinese has been analyzed in terms of accessibility, by characterizing the syntactic positions in which they occur and establishing a hierarchical scale of accessibility of these function position for Zero Anaphors. Chien (1983) proposes the following scale: Topic/Subject > direct object > indirect object > pivotal object > oblique object.

74. His example was a very basic conversational one based on the salutation exchange of Ni hao ma? *How are you?*, to which on should answer 'hen hao' *very well* and not 'Wo hen hao' *I am very well*, a pattern that is corresponding to the oral 'am fine'.

sentence-initial '*And*' in (g), Chinese adopts a covert syntactic strategy to mark the same grammatical information.

Although the pervasive use of zero-anaphora is not attested in English, its linguistic system does show the possibility of having discourse Topic crossing multiple clauses. Consider the case in sentence (b) for "color" which is implicitly understood as 'the color [of the car]', and the case of 'take a look' in sentence (d), where the object 'car' is taken for granted.

While the kind of interconnected sentences – called Topic chains[75] – are rare in the English version, they are very frequent in Chinese, statistical data show that approximately one third of clauses in narrative text involve the use of Topic chains (W. D. Li, 2005)[76].

We can note that in all the sentences in (88) there exist an aboutness relationship with the initial Topic NP, so that each sentence carries an additional piece of information -a comment- about the NP referent 'That car'.

Interestingly, contrasting the first two sentences (a) in Chinese and English, we can observe that Chinese does not uses copular structure as the English version does in 'That car is too expensive'.Interestingly, Chinese articulates the same information by using an additional Noun Phrase, *jiaqian* (litt. *price*), which gives rise to a [NP + NP + VP] configuration that we will extensively analyzed as being a typical example of Topic-Comment articulation. Hence, we can note that the predication using the auxiliary 'to be' in English is here replaced by a Topic-Comment articulation.

3.2.2.2 Frequency and Ambiguity of T-C articulation in Chinese

Topic-Comment sentence articulation is so frequent that ambiguity surfaces in very simple types of sentences. We will review a few cases that are representative in Mandarin Chinese of the pervasiveness of the sentence-discourse interface in the sentence's articulation of more or less basic sentence constructions.

As already observed in several examples, a constituent is fronted without the need of any morphological change in Chinese. When the nominal fronted is inanimate (feature), selectional restrictions between noun and verb do not allow the sentence to be ambiguous. However in case the fronted element carries an animate feature, selectional restrictions allow an ambiguous reading of the Topic-Comment structure, like in the following examples:

(89) 鸡吃了

jī chī le
chicken eat PRT.

'[I] ate the chicken.' or 'The chicken ate.' in Sun Chaofen (2006:184).

Note that, naturally, semantic differentiation is called upon when syntax is ambiguous or blurred.

75. Topic-chain analyses take an important step away from the sentence grammar level of analysis towards multi-clausal discourse-level. Shi Dingxu (2001) developed a syntactic analysis, arguing that Topic-Comment structures are a syntactic means to achieve particular kinds of discourse functions like the one shown in the Topic-Chain in example (88).

76. Tsao first observed (1977:92) that a Topic can extend its domain to several sentences, this can be viewed as the first formulation of Topic-chain in Chinese, where each sentence can be taken as an independent comment on the same Topic. In this way the Topic and the following Comment-clauses, form a discourse unit where certain co-reference properties are observable.

(90) 我鸡不吃

wǒ jī bù chī
I chicken NEG eat

a. Chicken as dish: 'I don't eat chicken.' or
b. Chicken as animal: 'My chicken does not eat.'

The phono-acoustic analysis of these kind of ambiguities offers an interesting insight in the syntactic encoding offered by intonational patterns of the sentence-unit and its word-order in Chinese. Chang Y-C (2001) behaviorally investigated the perception and production of several types of syntactically ambiguous sentences among which a category was the one reported in (91).

(91) 鸡吃了

a. 鸡，不吃了 (Conditon a08)

jī, bù chī le
chicken NEG eat PRT.

'I don't eat chicken anymore.'

b. 鸡不吃了 (Conditon b08)

jī bù chī le
chicken NEG eat PRT.

'The chicken does not eat anymore.'(Chang 2001)

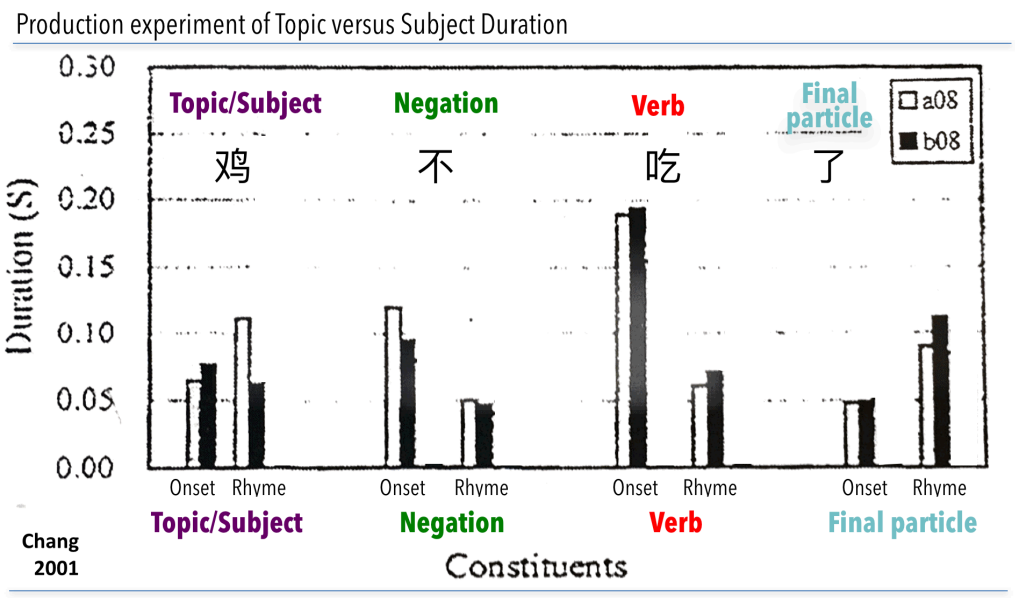


Figure 3.12 – Onset's and rhymes' duration of the constituents in the sentences in (91) *ji bu chi le*. Conditons a08 and b08 are respectively the topic-comment structure signifying 'I don't eat chicken any more' and the subject-predicate one, meaning 'The chicken does not eat anymore'.

The experimenters obtained from the participants to pronounce the two different readings (a) The chicken has not eaten.' and (b) 'The chicken was not eaten.' by a question

and answer experimental setting and recorded the intonational patters for the syntactic structures and consequent interpretations. The phono-acoustic analysis of the onset and rhyme duration of every syllable (i.e. syllable=word in this case) in (89) revealed that (1) the word 'ji' chicken showed a longer duration when pronounced as a Topic than when it plays the role of the subject of the verb chi 'eat', and (2) a lenghthening was observed at the onset of the post Topic syllable, as illustrated by Figure 3.12

These results show how the syntactic encoding of the topic function is tightly linked to oral intenational/prosodic patterns, we will deepen our understanding of the intonational pattern of by performing a phono-acoustic analysis on a corpus of 580 topic comment sentence in chapter 4.

Ambiguity between sentence-initial roles

Other examples of ambiguity given by the frequency of Topic- Comment structures can be found in more complex sentences like :

(92) 他谁都怕

tā shéi dōu pà
he who all fear

a. 'Everybody is afraid of him' or
b. 'He is afraid of everybody'

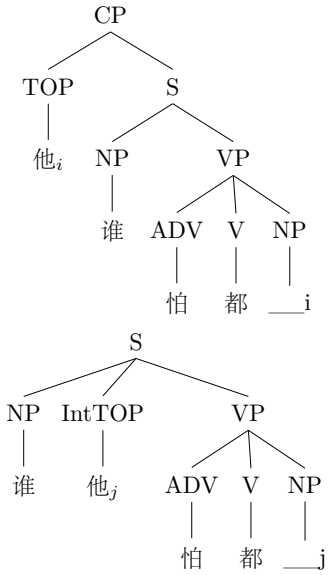


Figure 3.13 – Syntactic tree structure for reading of the above example (92a.) "tā shéi dōu pà" '*Everybody is afraid of him*'.

This sentence is structurally ambiguous. Depending on the specific structural analysis that is carried over the two NPs *ta* and *shei* at the beginning of the sentence can have different functions according to Chinese sentence structure rules. For reading (92a) the

pronoun ta is in Topic position and bears the role of the undergoer having moved from post-verbal object, while for reading (92b) the pronoun *ta* is in pre-verbal subject position as the doer of the action and the object of the verb pa is fronted in pre-verbal position in secondary topic position, following Sun Chaofen analysis (in 'Chinese: A Linguistic Introduction', 2002).

Ambiguity between object versus subject Topics

Another case of ambiguity in the structural analysis of out-of-context Topic-Comment articulations is observable in cases like in (93). While (93) can be analyzed as the extraction of an object from a complex NP (i.e. the Topic NP is interpreted as the object of the verb in the Comment clause), according to native speakers another analysis of the same sentence can be drawn as shown in (94), where the Topic is understood as the subject of the verb *criticize*.

(93) a. gap: 张三，批评的人很多

Zhāngsān$_i$, [[e$_j$ pīpíng e$_i$ de] rén$_j$] hěnduō
Zhangsan$_i$, [[e$_j$ criticize e$_i$ REL.] people$_j$] a.lot

'Zhangsan$_i$, people who criticize [him$_i$] are a lot.' from Huang , Li and Li (2009).

b. resumptive: 张三，批评他的人很多

Zhāngsān$_i$, [[e$_j$ pīpíng tā$_i$ de] rén$_j$] hěnduō
Zhangsan$_i$, [[e$_j$ criticize ta$_i$ REL] people$_j$] a.lot

'Zhangsan$_i$, people who criticize him$_i$ are a lot.'

(94) a. gap: 张三，批评的人很多

Zhāngsān$_i$, [[e$_i$ pīpíng e$_j$ de] rén$_j$] hěnduō
Zhangsan$_i$, [[e$_i$ criticize e$_j$ REL] people$_j$] a.lot

'Zhangsan$_i$, people who [he$_i$] criticizes are a lot.'

b. resumptive: 张三，他批评的人很多

Zhāngsān$_i$, [[ta$_i$ pīpíng e$_j$ de] rén$_j$] hěnduō
Zhangsan$_i$, [[ta$_i$ criticize e$_j$ REL] people$_j$] a.lot

'Zhangsan$_i$, people who he$_i$ criticizes are a lot.'

In (93) the relative clause *who criticize him*$_i$ ([[e$_j$ pīpíng e$_i$ de] rén$_j$]), contains two empty categories, one referring to the head of the relative, *rén* 'people', and the other to the Topic *Zhāngsān* (see co-indexation in the above examples).

As shown in both (93b) and (94b) the presence of a resumptive pronoun in either preverbal or post-verbal position in the Relative clause offers a syntactic disambiguation to this ambiguous configuration.

The fact that the choice between a subject/object analysis of the empty syntactic element reveals that contextual factors can at least partly determine the syntactic analyses in (93) compared to (94). This property could be linked to the kind of linguistic analyses that have been developed around the notion of discourse-configurationality for Mandarin Chinese to which we will resort in next section (see Discourse-Configurationality in Kiss,

1995 and for Chinese see Xu and Liu, 1998), where we will argue with Li and Thompson (1976) that this ambiguity between the main functional roles of subject and object in Chinese follows from its typological characterization as a Topic-Prominent language in section §3.2.3.

3.2.2.3 Topic-prominence versus contrastiveness

The frequency of Topic-comment constructions and the consequent ambiguity this can yield in the several circumstances we reviewed continue together to the grounding the widely observed fact that Mandarin tends to organize sentences into Topic-Comment structures.

As noted by Constant (PhD, 2014) "one way of understanding why Mandarin Topic-Comment structures are so prevalent is in terms of the relatively wide range of interpretations available to syntactic topics" (cf. 3.4.5, p. 404), and in particular *Mandarin sentence topics are not necessarily interpreted contrastively*, unless marked by a dedicated contrastive particle *-ne* as revealed by several authors (Li B. 2006, Contant's corpus study, 2014; Feng Shengli, in press).

(95) 热力学，大部分的人可能都没听说过。Out-of the-blue Context

Rèlìxué, dàbùfen de rén kěnéng dōu méi tīng-shuō-guo.
thermodynamics most de person possible even have.not hear-say-exp

Lit:'Most people have probably never even heard of thermodynamics.' (Constant, 2014)

Hence, considering the Chinese sentence in (95), it has the characteristinc of being possible even in what we would call an out-of-the-blue context, where the sentence topic rèlìxué 'thermodynamics' wouldn't be contrasted with other referents or topics. For example, (95) could play the role of the opening sentence of a thermodynamics class. Notably, this non-contrastive claim – arguing for the 'basicness' of Mandarin sentence Topics – holds also for the type of syntactic configuration where topics are arguments of the main verb in the Comment-clause and are associated with a gap, and for sentence-initial scene-setting Topics too.

(96) Contrastive interpretation of Topics in English

a. Most people have probably never even heard of thermodynamics.

b. Thermodynamics, most people have probably never even heard of.

By contrast, the English counterpart of (95) illustrated in (96) does not allow a non-contrastive interpretation. For instance, (96b) would sound unnatural in the mount of a thermodynamics professor opening his course, but becomes acceptable as soon as a contrastive context is given to support a contrastive understnaidn of the Topic '*thermodynamics*' as shown by (97).

(97) Contrastive context for sentence (96b)

a. Some areas of physics, like relativity, are quite famous. But thermodynamics, most people have probably never even heard of.

As Prince (1999) observes, specific discourse conditions can also allow the construal of English fronted material as focus, as in her example "She was here two years. <checking transcript> Five semesters she was here." Note however that the thermodynamics example (27b) does not support such a construal.

Marking topicality and contrastiveness configurations in Mandarin Chinese

We just saw that Mandarin syntactic topics are not always interpreted contrastively. However, when *-ne* appears following a Topic, only a contrastive interpretation is available[77]. Thus, (29) is possible in a contrastive context like (28), but impossible as the opening sentence of a thermodynamics class. In this respect, the sentence with -ne patterns with the English counterpart where the topic has been fronted.

(98) 热力学呢，大部分的人可能都没听说过。

Rèlìxué **ne**, dàbùfen de rén kěnéng dōu méi tīng-shuō-guo.
thermodynamics CONTRAST. most de person possible even have.not hear-say-EXP.

Lit:'Thermodynamics NE$_{contrast.}$, most people have probably never even heard of.' (Constant, 2014)

The approach to contextual factors determining topic interpretation in Chinese has been extensively investigate by Chauncery C. Chu, who even formalized a "Discourse grammar of Mandarin Chinese" capturing many interesting sentential facts linked to the interface between sentence and discourse (see Chu, 1998). Chu (2006:21) observed that topic-marking by the particle *-ne* permit a direct contrastive reading as shown in the below example (99a).

(99) Contrastive use of *-ne* particle

a. 妈妈每天晚上很晚才回家。爸爸呢，干脆就不回来。

Māma měi-tiān wǎnshang hěn wǎn cái huí-jiā. Bàba ne, gāncuì jiù bù huí-lái.
mom every-day night very late only.then return-home dad CT. simply just not return-come

'Every day mom doesn't come home until late. Dad NE, doesn't even come back at all.'(Shao 1989:174)

b. 学校五点才放学。爸爸 (# 呢)，干脆就不回家。

Xuéxiào wǔ-diǎn cái fàng-xué. Bàba (# ne), gāncuì jiù bù huí-jiā.
school five-o'clock only.then let.out-school dad CT. simply just not return-home

'School doesn't let out until 5pm. Dad (# NE), doesn't even come home at all.'Chu (2006:21)

The Topic 爸爸 bàba 'dad' is in opposition with contrasts with the sentence-initial Topic 妈妈 māma 'mom'. The specificity of the contrastive use of -ne is further illustrate by the impossibility to establish a contrast between 学校 xuéxiào 'school' and 爸爸 bàba 'dad' in (b).

77. As we saw in section §2.1.2.5, question-final *-ne* is often treated by the formal linguistic literature as a clause-typing particle that marks a clause as a wh-question. (See Cheng, 1991:21; Li, 1992:139; Aoun and Li, 1993:210; and Cheng et al. , 1996:80). However, there are a wealth of arguments showing that it not only plays this role (e.g. Lin, 1984:220-221; Shi, 1997:133–134; Gasde, 2004:315–318, and Li, 2006:13–15). Refer to Constant's (2014) and Li B.'s (2006) PhDs for further discussion.

In the last two examples and the corpus study by Constant (2014) in Figure 3.14 the use of -ne as a contrastive-topic marker is made evident and confirms previous analyses found in the literature (Lee-Wong, 2001; Lee, 2003; Li B. PhD, 2006)[78].

Importantly, we see here that Mandarin, unlike Japanese *-wa* (cf. example 71) and Korean *-(n)un*, marking both contrastive and thematic (non- contrastive) topics, has a dedicated contrastive marker in the particle *-ne*.

Given these elements and the the strong variability of native informers on topic morpho-sytnactic markers (see Notice to the reader, p.xxiii reproduce in the footnote)[79], we decided to avoid testing in our experimental material Topic sentences with overt morpho-syntactic marking, be it *-a*, *-ya* or *-ba*. Although the literature on Mandarin Chinese identified these topic markers (Shi, 2000; Paul, 2005; Li 2006), we decided to use the comma marking in our reading paradigm.

Not only, we chose the Topic marking that native speakers mostly agree on across the different dialectal areas is the *pause* marking, which can be easily transcribed in the written style by a simple comma. But, as our research focus is on the mental representation of the *syntactic* encoding of Topic-Comment sentence articulation, we resolutely tried to control for any possible emphatic or contrastive pragmatic reading in our 'Chinese style' Topics (when tested out of context), this was mainly meant to to avoid the possible confound linked to a particular informational-load processing that would have been additionally implied by contrastive interpretation of Topics. For this very same reason, we also chose Scene-setting Topics which are generally not contrastive and performed a preliminary phonological study on the prosodic contour of Frame-setting Topics in chapter 4.

After having tentatively unraveled the notional richness and complexity surrounding Topic-hood, we can concentrate on the various cognitive aspects that were brought to light, that make the Topic-Comment sentence's articulation a valuable candidate for a neuro-cognitive and experimental investigation of the sentence-unit.

The reader with a background in Cognitive Science has probably already identified some of them along this notional overview, we will address them directly in the following.

Summarizing the reasons that brought us to select this sentence articulation as a research object, a collection of well circumscribed pieces of evidence for the claim of topic cognitive relevance was contributed.

78. For instance, Lee-Wong (2001:139) states that topic-marking -ne "signals topic shift and focuses on new, contrastive information". By claiming that Topic-marking by the *-ne* particle "shows an explicitly expressed or listed contrast between two elements of the same type", Lee (2003:357) first establishes a link between the topic-marking function of *-ne* with the class of Contrastive markers found cross-linguistically, followed by Li B.'s PhD (2006).

79. Since Mandarin is spoken and written by a wide range of speakers with different linguistic backgrounds, having being raised by speakers of different dialects (e.g. grand-parents), there is an important issue as to the extent of variation in how topic markers like *-ne*, *-a*, *-ya* are used across different dialects, and, social groups. The previous literature on topic marking in Mandarin Chinese has tended to abstract away from these different kinds of variations. One remarkable exception is the deep investigation of Topic markers in Chinese varieties in a book by Xu Liejiong and Liu Danqing on Shanghai dialect's Topics. While we won't be able to offer any characterization of the nature of this variation here, it is almost certain that some variation does exist, and we leave this for future research. At the beginning of my research, when eliciting judgments about a particular use of *-ne*, *-a*, *-ya* as Topic makers, speakers would accept the sentence resisting a bit, saying that it sounded like something other might have been said before, or required a certain particular context, or under some very precise circumstances, or even that it seemed not standard Chinese to their hears to use such and such particle. Hence, given that my initial interest for Topic-Comment sentence was related to the "basicness" these sentences recovered once the Topic marker was dropped, I haven't attempted to categorize these intuitions here.

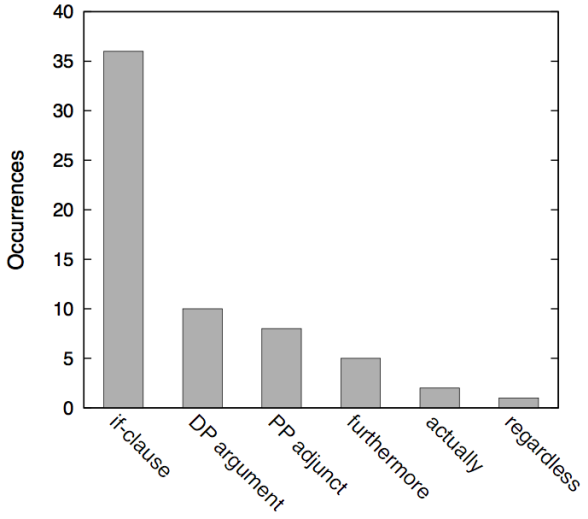


Figure 3.14 – Types of constituents marked by *-ne* particle in the 'Richshaw boy' by Lao She (1936). Intrestingly, the large majority is not represented by Bare-NPs.

It was argued that:

- Topic-Comment articulation has been identified already by Grundel (1988) as a ***Universal*** structure across languages. Frequent in everyday communication, going from commercial advertisements to presidential talks (lately on fashion in France), its universal presence across languages makes it perfect candidate for a neuro-cognitive investigation about language sentence structures.
- Topic-Comment is proto-typically an oral construction across languages. As such it is a paramount of linguistic ***natural*** stimulus. Its functional role is grammaticalized through morpho-syntactic marking in several languages (e.g. Japanese, Korean, Russian and Hungarian in particular cases), while other feature an optional or international marking like Chinese.
- The psychologically pertinent concepts of accessibility and identifiability of the Topic referent, that were brought to light by linguistic analyses, show the fact of taking into account the hearer and his ***mental representation***.
- At the level of logico-semantics, and in the philosophical account we reviewed, the notion of Topic is identified with a ***structural bi-partition*** of the propositional *predicative pattern*, that makes the Topic-Comment structure "the most general characterization of ***predicative constructions***", where "the speaker announces a topic and then says something about it." (Hockett, 1960), or where the Topic plays the role of a Frame limiting the applicability of the predication of the Comment-clause (Chafe, 1967).
- Last but not least, the syntactic encoding of the sentence-discourse interface (cf. informational packaging) that Topic-comment sentences feature, additionally makes them an cognitively interesting testing ground, to uncover the cerebral representation of constituents carrying contextually relevant information and the syntactic

> structure that encodes this discursive functional properties. In this way, Topic-comment sentences can be considered as the locus where the sentence functional interface with context motivates syntactic operations yielding syntactic complexity in the sentence-unit.

This listing reveals in one sentence that Topic-comment articulations are Natural, Universal, Structured and Complex Syntactic objects as the sentence-unit.

To these elements it should be added, that Topic-comment structuring or articulation is a good candidate for typically human enunciation structure. Namely, in recent years, the kind of analyses on Topic-Comment articulation we have been reviewing have been brought to Cognitive Sciences, as illustrated by a very original and inspirational article by Manfred Krifka (2007b).

Noting that human interactions exhibit a pervasive structuring of utterances into Topic and Comment, the author points out striking similarities between the bimanual coordination and the structuring of utterances into Topic-Comment. The dominant hand can be seen as physically predicating something on the object held by the non-dominant one. To support this Krifka gives arguments showing that bimanual coordination influences the gestural expression of Topic-Comment structure in Sign Languages, further claiming that asymmetric bimanual coordination played a role in the rise of the Topic-Comment structures in communication.

From an evolutionary point of view, the author engages in the correlation between visual processing and the Topic-Comment articulation, by noting that the task sharing between the ventral stream –*the what pathway*–involved with object identification and the dorsal stream –*the where pathway*–processing spatial locations could be reinterpreted as a Topic-Comment system (cf. Mishkin et al., 1982). In this way, the *where stream* would be viewed as predicating the spatial articulation of the object identified by the *what pathway*.

Although fairly speculative this remarks points to an informational understanding of cerebral pathways -a largely adopted view in Cognitive Science. Furthermore it still holds true that animal communicative behavior sharply contrasts with he pervasiveness of Topic-Comment articulations in human linguistic behavior. Animal communication, seems to be lacking the communicative behavior to first identify an object and then comment on it, what could be interpreted as a proto-formulation of the topic-comment predication mechanism. Tomasello and Zuberbühler (2002) clearly state:

> "Virtually no ape gestures are referential in the sense that they indicate an external entity (i.e., there is no pointing in the human fashion)."

We can retain a second argument made by the author and going in the same sens. Krifka proposes interesting reflection is given on a topic that has recently attracted much attention in Cognitive Science with the publication of an article on the warning calls of Vervet monkeys (Barceló-Coblijn and Gomila, 2012). The author note that their signaling, for example, "danger from above / an eagle", or "danger from the ground / a snake" (cf. Struhsaker, 1967) manifest some form of possible compositionality, but does not illustrates the sentence predicational configuration shown in the Topic-Comment articulation. For instance they do not *first* identify a particular region, or a certain type of animal, *and then* say something about it[80]. Similarly, Tomasello (2003) notices that chimpanzees produce attention-getting gestures, but does not appear any strategy

80. One instance of animal communication that could be remotely comparable to Topic-Comment

for combining these gestures with others communicating more specific content which could be viewed as a precursor of Topic-Comment structure in humans, like the pointing gestures of children or the non-fluent aphasic talks illustrate.

These facts and their interpretations overtly contrasts with what we put forward in human linguistic behavior, where the facts send still that Topic-Comment articulation is an essential feature even in the early process of language acquisition, as illustrated by French children's early use Topic-Comment sentence (De Cat and Lahousse, 2002).

All this put together makes Topic-comment articulation a good candidate for typically human enunciative structure.

Although ground-breaking and capturing interesting and essential aspects of the sentence-unit's structure and of human syntactic competence (e.g. sentence-discourse interface and syntactic operations liked to displacement), the hardship to experimentally approach the Topic notion lies in the richness of satellites notions that surround it (i.e. informational structure, syntactic encoding, predicative Frame, definiteness, aboutness, referent's identifiability/accessibility). Even though this notional/definitional challenge can easily be solved theoretically just by choosing a certain level of analysis, field of research or theoretical framework, experimentally addressing its neural implementation forced us to adopt a wider perspective and urged a notional clarification to circumscribe the cognitive aspect at play and orient our experimental methodology.

In conclusion, the different arguments and examples that were brought to light in this section showed that Topic-Comment articulations do have a cognitive relevance for the investigation of human syntactic competence, to build propositions predicating meaning in the structured utterance called the sentence.The next two sections of this chapter will address the arguments that make Mandarin Chinese an ideal testing ground for a neuro-syntactic investigation of the sentence-unit.

Word-orders and contrastiveness

Given the way Chinese systematically exploits word-order for syntactic and discourse-oriented purposes it has been described as positional syntax (or grammatical word-order) language with an SVO basic word-order, in that core grammatical roles like subject and object are defined in terms of the syntactic position in which occur[81].

The Topic-prominent parameter attributes to Chinese Topics a significant role in the sentence articulation (Li and Thompson, 1981)[82], but Mandarin also features at the sentence-level some word-order variation that convey contrastiveness as shown by the following series of examples:

(100) Word-orders in Chinese

a. 我很喜欢音乐。(SVO)

articulation is bee communication, in that bringing pollen to the hive, which could represent the Topic, bees indicate through their dance the direction and distance where more of that pollen can be found, which in turn is the Comment.

81. This is also true on the phrase-level where word-category distinctions are determined on the basis of word-order too, like in noun and classifier ambiguities. Moreover, Word-order plays also a role at the level of the utterance informational structure too. If we consider for example the linearization criteria linked to 'definitness', NPs with definite referent tend to be pre-verbal while those with indefinite referents tend to surface post-verbally.

82. The typological emphasis on Topic centrality in the organization of The 'Chinese Sentence' -if we may afford the term- can in fact be at first sight in contrast with the unanimously acknowledge fact that Mandarin Chinese has a canonical SVO word-order.

Wǒ hěn xǐhuan yīnyuè.
I very like music

'I like music.'

b. 我音乐很喜欢 (SOV)

Wǒ yīnyuè hěn xǐhuan.
I music very like

Lit: 'I, music, like.'

c. 音乐，我很喜欢 (OSV)

Yīnyuè, wǒ hěn xǐhuan.
music I very like

'Music, I like.'

If we consider the different word-orders in (100), the SVO order behaves differently when a negative polarity is inserted compared with the other two word-orders. This extensively attested phenomenon (Huang, 1982; Huang, Li and Li, 2009; Li and Thompson, 1981; Qu, 1994; among others) is illustrated by (101), where the sentential negation *mei* is added to the three sentences in (100)

(101) Word-orders in Chinese

a. 他没写什么/任何书 [S neg VO]

Tā méi xiě shénme/rènhé shū.
he not write what/any book

'He did not write any book.'

b. * 他什么/任何书没写 [SO neg V]

Tā shénme/rènhé shū méi xiě.
he what/any book not write

Lit: 'He any book did not write.'

c. * 什么/任何书，他没写 [OS neg V]

Shénme/rènhé shū, tā méi xiě.
what/any book he not write

'He did not write any book.' (from Huang, Li and Li, 2009:200)

Although the SOV and the OSV orders share the same features in sentential negation, the two are differ greatly in interpretation. SOV has the features of focalization, requiring a contrastive or a focus interpretation (Xu and Li, 1993).

The focal interpretation of pre-verbal object further attests how Chinese language structurally attributes to given word-orders and linarizations some sentence-discourse interface interpretative properties, a property that has been proposed by Liu and Xu (1998) as meeting the conditions to qualify as Discourse Configuration language.

The issue of discourse-semantics and the linearization principles of the structure of the sentence across languages has been dressed by E. Kiss introducing the notion of Discourse-Configurationality (1995). In this framework, a language categorizes as discourse configurational when the discourse-semantic function of Topic is expressed through a particular structural relation associated to a particular structural position. And parallely when

the discourse-semantic function of Focus is realized through a particular structural relation implying for instance the movement into a particular structural position in the sentence. These two criteria being met in Chinese, Liu and Xu (1998) proposed to qualify Chinese as discourse configuration language. Hence, from formal point of view, we will consider the SOV structure is derived by A-movement following Shyu (1995) and Badan (2008), while Left-dislocation of the object yielding the OSV structure is derived by Abar-movement (Huang, Li and Li, 2009).

In conclusion, this extraordinary simple mapping between linearization and discourse-semantic interpretation is one of the core reasons for our choice to carry on a neuro-linguistic investigation of how the brain represents the Chinese sentence-unit. The way Chinese systematically exploits word-order for syntactic and discourse-oriented purposes allows to uncover in absence of explicit morpho-syntactic cues how the brain manages sentence structure building and the complex calculation linked to the sentences-discourse interface only by linear and positional syntactic encoding.

These remarks on the default non-contrastive interpretation of Topics in Chinese feed our on-going reflection on the sentence-discourse interface as a property of the sentence-unit. The clear contrast between Topic-comment basicness in Chinese and its essential contrastiveness in English illustrated above is essential in delineating the fundamental basicness of Topic-comment articulation and its interface with discourse. Topic-comment utterances appear namely a "favorite sentence type", to say it à la Hockett, without necessarily requiring a contextual contrastive reading. We could therefore conclude that in Chinese the link to discourse is part of the basic predicative articulation these constructions embody more that a informational one implying necessarily a contrastive reading of Topics.

These observations introduce us to next section where a more thorough overview of the the typological characteristics of Mandarin Chinese will help us to better identify some of the factors determining this neutral or basic interpretation of Topics.

3.2.3 Typological characteristics of Topic-prominent languages

> "Pour ce qui est de langue du pays, je puis vous assurer qu'il n'y a que pour Dieu qu'on puisse se donner la peine de l'apprendre."
>
> *["As far as the language of this country is concerned, I can tell you that only for god would one bother to learn it"]*
>
> Joseph Henry de Prémare
> (Jesuit,1666-1736)

Since its early formulation in Chinese Linguististic (Chao, 1968) and the more cross-linguistic approach of Li and Thompson in the 70's, the notion of Topic has been widely used to analyze Chinese sentence structure, and Topic-Comment articulation has been generally recognized as a distinctive typological feature of Chinese language.

While the initial claim on topic prominence is based on the fact that introducing the notion of topic-comment is descriptively convenient, claiming that "some languages can be more insightfully described by taking the concept of topic to be basic, while others can be more insightfully described by taking the notion of subject as basic."(Li and Thompson, 1976)[83].

The cross-linguistic survey on 30 languages conducted by Li and Thompson revealed a four category classification of languages according to several syntactic characteristics that concretely delineated a number of linguistic trends overriding the simple descriptive convenience. These syntactic trends and characteristics have been argued to follow from the prominence of the notions of Topic and Subject in the articulation of sentence-units.

Among the syntactic characteristics that first attached our attention is the fact Topics are in control of co-reference. We already addressed introducing the notion of Topic Chain in example (88) (§3.2.2.1, p. 292) and in Second Language acquisition (§3.2.1.3 , p. 286)[84].

(102) 那棵树叶子大，所以我不**喜欢** (那颗树)。

Nà-kē shù yèzi dà, suǒyǐ wǒ bù **xǐhuān** ___[nà-kē shù]
this-CL. tree leaves big, so I NEG like ___(this tree)

"This tree its leaves are big, and I don't like it."

While we already noted that Topics can be distinguished from Subject in that they have no selectional restrictions with verbs in the comment, and that they can take as Comment a whole sentence[85], and extend their 'semantic domain' to more than one clause (see example [102]), Topic-Prominent languages manifest a series of syntactic trends and patterns that represent the distinctive character of their typological classification (Li and Thompson, 1967:466-471)[86]:

1. Surface coding of topics by sentence-initial position or with dedicated Topic morpho-syntactic markers (but not necessarily for subject);
2. A trend to marginalize the importance of Passive constructions;
3. The absence of "dummy pronouns", lack of expletive subjects;
4. The co-occurence of subject ant topic in Double nominative constructions, and their pervasive use (i.e. sentences with a subject and a base-generated topic);
5. The control of co-reference and control of the deletion of co-referential elements with the Topic;
6. The possibility to have multiple topics and the lack of constraints on what can serve as a Topic;
7. Ambiguity in marking of the distinction between subject and object;
8. Do not govern clause-internal phenomena such as reflexivisation;
9. The basicness of Topic-Comment sentences.

83. Other examples of Topic-Prominent languages are found in East-Oriental Languages like Japanese, Korean, Vietnamese, Malaysian, Indonesian, Singlish (Singaporean English), Malaysian English, Hungarian, Somalian, and the Siouan (Amerindian language) are all considered Topic-Prominent languages together with a certain number of American Sign Language.

84. We illustrate only briefly this feature in the following example (102), and differ the discussion of the different patterns of resumption to the section dedicated to Topic-Comment syntactic analysis (§3.4, §3.4.4)

85. To the point that the structure of a topic comment construction have been described as "a loosely formed structure of NP + S" (Chappell, 1996:465).

86. To the following list should be added also: A tendency to be verb-final; Absence of articles denoting the given/new informational status;

We will briefly review and illustrate some of the distinct set of syntactic characteristics of topic-prominent languages in the following and dedicate two subsections to topic marking (§3.2.3.2), to the different sentence configurations yielded by the concurrence of Subject and Topic (§3.2.3.3), and a subsection to develop some sentential phenomena linked to the presence scene-setting topics, we analyze as subject-inversion (§3.3).

3.2.3.1 Main syntactic trends in Topic-prominent languages

To introduce our overview of the main syntactic trends in Topic-prominent languages, it should be acknowledged that among the most striking examples of the basicness of Topic-Comment constructions are doubtlessly, verb-less sentences illustrating the the possibility to have multiple topics as the sentences, below:

(103) 这辆车，两门。

Zhè-liàng chē liǎng mén
this-CL. car two doors

'This car has two doors.'

(104) Multiple topic: 这个大学，两个学院，一个理工，一个文史。

Zhè-ge dàxué liǎng-ge xuéyuàn, yīge lǐgōng, yīge wénshǐ.
this-CL. university, two-CL. colleges, one-CL. sciences, one-CL. humanities

'This university [has], two colleges, one Science and Technology, the other Arts and Humanities' Li (2007)

The possibility illustrated by 169 to have a layering of Topic comment one inside the other is what Hockett was defining by the term "Chinese box style". These sentences not only show that the notion or concretely the presence of a subject is not as 'vital' in Chinese as in English for example but that sentence can be built verb-less just by relying on the aboutness relation conveyed by Topic comment predicative articulation. What other argument should we need to say that the subject-predicate relationship is not at the core of the sentence linguistic system?

No subject sentences and the passive meaning in Topic-Comment

Another kind of subject-less Topic sentence type brings some evidence to relative weakness of the subject definition inside of Topic-prominent linguistic systems.

Examples in (105a), illustrate how a passive content like 'having been published/broadcast', is expressed by having the direct object of the Comment's verb in Topic position playing nonetheless the semantic role of a patient, which brings some evidence to the marginality of passive constructions in Topic-prominent languages.

(105) Pseudo-passive sentences in Chinese

a. 这件新闻广播了。

Zhè jiàn xīnwén guǎngbō le.
this-CL. news broadcast ASP

'This piece of news has been broadcast.'

b. 那本书已经出版了。

Nà-běn shū yǐjīng chūbǎn le.
this-CL. book already edit ASP

'This book has bee already edited.'

These sentences have been analyzed under a multitude of point of views, eight as having no subject or as having no object (or as having no topic), and are generally called Pseudo passives, in that they can easily be analyzed as conveying a passive meaning as shown by their passive form translation but no passive form is grammatically adopted in Mandarin Chinese. Although the subject of the verb is absent, it can be easily understood from the context as shown by (106) (Li and Thompson, 1981). These type of configuration are presumably tight to the ambiguity phenomena addressed in section §3.2.2.2.

(106) Sentence with no subject: 衣服烫完了

yīfu tàng wán le
Cloth iron RES.finish ASP

Lit.: The clothes [someone] has finished ironing [it].
'The clothes are ironed.'

Similarly, the contrast between (107a) and (b), offer an additional example of how passive meaning can be expressed by a so-called impersonal Topic-Comment constructions like (c), where the actor NP is not of interest and can be omitted. These constructions echo non-adversative *bei* constructions in (c) which feature the inversion of the agent undergoer NP in its Word-order (OSV) and topicalizes the undergoer in the following syntactic configuration: Object-NP(undergoer) + bèi + Subject-NP(actor) + VP transitive+ Result/Adv.

(107) a. 茶被喝完了

chá bèi hē-wán le
tea PASS. drink-RES*finish* PART.

'The tea was finished.' or 'The tea has been finished.'

b.（我）把茶喝完了

(wǒ) bǎ chá hē-wán le
(I) BA tea drink-RES*finish* PART.

'(I have) finished the tea.' or 'I drunk out the tea.'

c. 茶喝完了

chá hē-wán le
tea drink-RES*finish* PART.

'The tea was finished.'

d.（我）喝完茶了

(wǒ) hē-wán chá le
(I) drink-RES*finish* tea PART.

'The tea was finished.'

Tentatively, the translation of these syntactic options differ along the lines of their interpretation:

1. through *bei* a change of state of the undergoer is stressed;
2. though *ba* the result 'there's no more tea left' is put forward;
3. SVO expresses a simple sequence of event [drink + finish], and
4. Topi-comment construction expresses the pseudo-passive meaning about 'tea' and its resulting state of being finished.

No topic sentences and Null-topics

Parallel to this trend, another linguistic phenomenon is observed in sentence with no topic but featuring an unusual post-verbal subject in Chinese:

(108) a. Sentence with no topic: 进来了一个人

jìn-lái le yī ge rén
enter-DIR ASP one CL. person

'A person came in.'

b. Sentence with topic: 王冕家进来了一个人

Wángmiǎn-jiā jìn-lái le yī ge rén
Wangmian-maison, enter-DIR ASP one CL. person

'A person entered in Wangmian's house.'

This type of no topic sentences in Chinese are usually analyzed as having undergone *topic-drop* (cf. Hunag, 1982, §2.2.4.3). In (108a), 'a person' is analyzed as the subject rather than a Topic since it has a selectional relationship with the verb 'come in', it isn't in sentence initial position, and it is neither definite nor generic (Li and Thompson, 1981).

Interestingly, Japanese explicit morpho-syntactic marking for Subject and Topic roles, shows that a direct object of the verb *suki* 'like/appreciate' in (109) can be attributed either the subject role. This not only participate to the discussion developed in section 2.2.3 about the fact that subject is in general terms a difficult notion, but also contribute an argument for to the general trend topic-prominent language shows in the indifferentiation of their subject and objects.

(109) 苺が好きです

ichigo-ga suki desu
strawberry-SUBJ. appreciated COP.

Lit: "the strawberry is appreciated"

(110) 私は/友達は苺が好きです

Watashi-wa/tomodachi-wa ichigo-ga suki desu
me-TOP./my.friend-TOP. strawberry-SUBJ. appreciated COP.

Lit. "As for me/As for my friend, the strawberry is appreciated"
"I/My friend like/s strawberry"

Note that in (109) the subject appreciating strawberries is not present, the whole pattern resembling to the pseudo-passive sentences we just illustrated in Chinese. However, example (110) brings in some evidence that a Topic can be added to this sentence articulation when wanting to express that the taste for strawberries is one's own of that of some friend.

Lisu's subject/object ambiguity

According to Li and Thompson's classification Lakhota and Lisu belong the group of languages where the extreme role of Topics classifies them into the entirely Topic-prominent group. Form the examples below we can better understand what kind of phenomena are subsumed under the 'Ambiguity between subject and object role' -the 7th distinctive characteristic of Topic prominent languages

This feature was aready adressed for Chinese in §3.2.2.2 (see examples (94) and (93)), but finds a somehow extreme expression in Lisu.

Lisu (Lolo-Burmese language) has been namely reported as a language where it is frequently impossible to distinguish the subject from the direct object or even the agent from the patient (Rosen, 2007)[87].

(111) Lisu

làthyu *nya* ánà khù-a.
people TOP. dog bite-DECL.

a. 'People, they bite dogs.' or
b. 'People, dogs bit them.'

(112) Lisu

ánà *nya* làthyu khù-a.
dog TOP. people-DECL. bite-DECL.

a. 'Dogs , they bite people.' or
b. 'Dogs, people bit them.' from Li and Thompson (1967:475)

As illustrated by the example in (111), subjects in Lisu are difficulty diagnosed, which brought Rosen to define this language as a *discourse-oriented languages*, which according to her definition "licenses arguments on the basis of the role each plays in the discourse rather than in the event denoted by the predicate" (2007:197-198).

In (111) the topic 'people' is interpreted as either the agent or patient of biting, with the non-topicalized argument 'dogs' as the other argument of the verb. Lisu is an SOV language lacking verb agreement, making it impossible to tell whether the non-topicalized constituent is the internal or external argument of the verb.

Lakhota (Sioux Language) offers another example of the cardinality of Topic in sentence structure building and of the fact the subject plays a secondary role in the articulation of the sentence, Lakhota makes in fact a morpho-syntactic distinction between

87. Sara Rosen has a thoughtful discussion in a chapter titled "Structured event, structured discourse" where we can read: "Not all languages use the A-positions TP and nP/AspP to license and interpret the clausal arguments; instead they license arguments in the A-bar positions in the CP layer. The languages that license arguments in the A-bar system fail to interpret the arguments via event structure, but rather use the discourse structure as encoded syntactically in the CP layer. I call these 'Discourse languages'." 2007:197-198). Ritter and Rosen (2005a and b) argue that some languages are more discourse-oriented and organize the arguments of the clause around discourse principles related to topic or point of view.

semantic subjects that "perform, enact, or instigate" the action and those that do not (Mithun, 1991).

(113) Lakhota

Miyé ṡuŋkáwaƙaŋ eya owićabluspe yeló.
be-the.one-that.1.sg horse.DET.PL. capture-UND-1.sg.ACT. ASSERT.(decl).masc.

Lit. 'As for me, some horses, I have captured' or 'It's me that has captured some horses' or 'I'm the one that captured some horses.' from Li and Thompson (1976)

3.2.3.2 Marked and unmarked topics: Japanese vs. Chinese

Topic-prominent languages dispose of a range of different means for the surface encoding of Topic function. Certain languages present overt marking of the Topic role, while they have none for the Subject role, others present the opposite pattern, For example, Lisu (傈僳族 Lìsùzú) and Lahu (拉祜族 Lāhùzú) from the Lolo-Burmese linguistic family, make use of morphological means to mark Topics, while they do not have any for subjects.

Japanese and Korean present an intermediary marking strategy between the two typological patterns exposed, in that they both mark topics with the particle *-wa* and subjects, with the particle *-ga* (cf. Kuroda (2005) and discussion on categorical and judgment in §3.1.2.1, p.239). Yet, inside the category of Topic-prominent languages there exist different marking strategies for the topic constituents by overt morphological marking or using a positional strategy like Chinese, that only has an optional Topic marking.

Contrasting the following sentences in Chinese (114) and Japanese and Korean (115a and 118), we can observe that while they differ in their overt morpho-syntactic marking, their structure and word-order are fundamentally the same.

(114) 张三我已经见过了

Zhāngsān wǒ yǐjīng jiàn-guò-le.
Zhāngsān I already see-EXP.-ASP.

'Zhang San, I already saw him.'

(115) a. 魚は鯛が美味しいです。

Sakana-wa tai-ga oishi-i desu
fish-TOP. snapper-NOM. tasty-NPST. desu

'Among fishes, the snappers are [the most] tasty.'
Or : 'The snappers are the most delicious fishes.' (1976:468)

Chinese adopts a positional syntactic means to encode the Topic role in the sentence, it is marked by sentence-initial position (for a argumented discussion on the possibility to have Topics in other positions in the sentence see Paul, 2002)[88]. While Japanese and Korean have an uncontroversial marker for topicality, Chinese topical marking is optional and the particles present in the dialects are often having also other functions like sentence-final-markers .

88. Paul (2002) namely shows that in the area between subject and verb can host a more restricted number of topicalized positions than the area to the left of subject. Only one IP-internal Topic position, only one bare pre-posed object is available.

As for Topic positional marking, Fuller and Gundel (1987) propose an interesting parallel with English strategy of coding Topics either by sentnece-intial positon or by non-inital positons as illustrated by the repetiton of 'The goldfish' in the sentence reported in (116).

(116) Non-sentence-intial Topic
  a. The goldfish lay on the table next-under the aquarium.
  b. And it looked like the gold fish was dead.

In the obligatoriness of the morpho-syntactic marking of the Topic function resides the most important linguistic evidence for advocating the essential role played by Topics in the building of the sentence-unit in topic-prominent linguistic systems. Namely, Japanese obligatorily marks the topic role by the particle *-wa*, even in Double subject configuration where topic and subject are in a part-whole semantic relation. Typological investigations, has indeed argued that this overt morpho-syntactic or positional marking attests that in those linguistic systems the Topic is indeed treated as a grammaticalized function having its own morpho-syntactic marking in the same way as subjects and objects.

Consider the following Japanese sentences morpho-syntactically differentiating the role of Totoro, through は *-wa* (topic) or が *-ga* (subject) marking, in two situations where the referent of marked NP 'Totoro' can conceptually both be the subject and the topic:

(117) a. トトロは学生である

Totoro-wa gakusei de aru
Totoro-TOP. student de aru

'As for Totoro, [he] is a student' (as opposed to other persons)

b. トトロが学生です。

Totoro-ga gakusei desu
Totoro-SUBJ. student COP.

'Its Totoro that is a student.' or 'Totoro is a student.'

(118) Korean Topic marking (from Gundel, 1988:217)
  a. John-nin i salam-lil manna-ass-ta.
  John-TOP. this man-OBJ. meet-PAST-DEC.
  'John (topic), met this man.'

  b. I salam-nin John-ka manna-ass-ta.
  this man-TOP. John-SUBJ. meet-PAST-DEC.
  'This man, John met.'

  c. Suyeng-nin Waikiki-ka coh-ta.
  swimming-TOP. Waikiki-SUBJ. good-DEC.
  '(As for) swimming, Waikiki is good.'

d. Ecey-nin John-ka naksicil-lil ka-ass-ta.
yesterday-TOP. John-SUBJ. fishing-OBJ. go-PAST-DEC.
'(As for) swimming, Waikiki is good.'

e. San-ey-nin namu-ka manh-ta.
mountain-at-TOP. tree-SUBJ. many-DEC.
'At the mountain, there are many trees.'

f. San-ey ka-1 ttay-ey-nin moca-lil ssi-la.
mountain-to go time-at-TOP. hat-OBJ. wear-IMP.
'When you go to the mountain, wear a hat.'

Let us turn to Topic structures in Chinese. If we contrastively consider the languages that are found in the founding article of 1976 by Li and Thompson we can already observe that Lisu, Japanese and Korean obligatorily encode the Topic function of sentence-initial Topics with dedicated markers, while Chinese displays a word-order strategy, even having multiple Topic markers, like *-ya*, *-a*, *-me*, *-ne* and *-ba* (Li and Thompson, 1981), they are all considered as optional the literature on Mandarin (Shi, 2000; Paul, 2005 ; Li B., 2006), while they are not in dialects like Shanghainese (Xu and Liu, 1998).

Interestingly, resuming to the previous discussion on -ne topic marking and its contrastiveness (cf. §3.2.2.3, p. 299) Mandarin Chinese, unlike Japanese -wa and Korean -(n)un, marking both contrastive and thematic (non- contrastive) topics, has a dedicated contrastive marker in the particle *-ne*.

The comparison between (119a) and (b) shows that the topic marker does not need to be phonetically realized. However, when the topic marker is null a comma can occur after the topic, a typographical fact that is sometimes also observed after the topic marker, it is interpreted as a pause or an intonational break when speaking (contra Xu, 2000).

(119) Topic marking with -a 啊 in Mandarin

a. John 我喜欢他

John wǒ xǐhuān tā
John 1SG. like him

'John, I like him'

b. John 他喜欢他自己

John tā xǐhuān tā zìjǐ.
John 3SG. like 3SG. self

'John, he likes himself'

c. John 啊他喜欢他自己

John a tā xǐhuān tā zìjǐ.
John PART.(TOP) 3SG. like 3SG. self

'John, he likes himself'

The particles identified by the literature as Topic markers, however, as demonstrated for -ne (cf. §3.2.2.3) have a semantic meaning, and do contribute a semantic interpretation to the topic phrase.

The examples in (119b) and (c) tentatively show how topic markers like *-a* and *-ya* demonstrate in the difference between (119b) and (c) to carry nonetheless some additional pragmatic information.

Further research needs to be carried out to solve this issue (cf. previous discussion in section §3.2.2.3), especially because there exist consistent dialectal evidence for the link between markers like -a and -ya and topic function. Many Chinese dialects that posses Topic marking and in Sinitic languages of Northwest China and Chinese varieties, for instance the marker *ha* in (120) derived from the Mandarin Topic markers *a* and *ya*, as proposed by Xu (2011) is systematically used as topic marker[89]. Consider the following examples in Xu (2015) in linguistic data from Linxia, Qinghai, Tangwang and Wutun dialects:

(120) Chinese varieties with a derived topic/accusative markers

a. Linxia: 這個慫哈壞得很 from Li Wei (1993:437)

zhe ge song **ha** huai de hen
DEM CL. coward TOP. bad degr-part very

'This guy is terrible.'

b. Qinghai: from Wang and Wu (1981:51)

ni **ha** bao qu shuo
2SG. TOP. NEG. go say

'Don't go there, he said.'

c. Tangwang: Xu Dan (2011)

jɑ̃ xa tʂhɿ tʂɛ
sheep TOP. eat DUR.

'Sheep are grazing (on the grass).'

d. Wutun: from Janhunen et al. '2008:63)

gu ha e di li
3SG. TOP. hungry PROGR. OBJ.

'S/he is hungry.'

89. Interestingly, [xa] (or *ha* in pinyin) has been reported to have two functions: topic marker, called a "focus marker" by Janhunen et al. (2008), and accusative case marker, which is among the elements that brought Xu (2015) to propose that the accusative marker [xa] originated from a Chinese topic marker.

3.2.3.3 Topics versus Subjects and their co-occurrence

> *In a subject-prominent language "the structure of sentences favors a description in which the grammatical relation subject-predicate plays a major role.", by contrast, "topic-prominent languages are those in which "the basic structure of sentences favors a description in which the grammatical relation Topic-Comment plays a major role".*
>
> Li and Thompson (1976:459)

After having pointed to the several sentence-level phenomena that couldn't fit into the subject-predicate sentence model, we turn now to the analysis of the fundamental linguistic difference between Topic and Subject and to a particular configuration that can be found when Subject and Topic co-occur, that of Double nominative constructions.

The characterization of Topic against Subject has already allowed to uncover a number of properties defining these two sentential functions. Thus, to begin our analysis we briefly recapitulate in Table the different characteristics reviewed so far on the difference between Topic and Subject, among which the most distinctive one remains the one of exhibiting or not an agreement with the verb. However, Mandarin being an o inflectional language does not exhibit any agreement with the verb this property is transferred on the level of selection relation with the main verb. It should be noted that this criteria is sometimes also defeated by a language like English in sentences like "There's Vanna, Laurence and Stanislas coming out of the lab", where the subjects of the action do not agree with the verb because of the presentational sentence in which they are embedded in.

Another important aspect distinguishing Topic and Subject at the sentence-level (n° 7 on our list in Table 3.7) is linked to the exclusion of topics in strictly clause bounded processes like reflexivization (cf. chapter 2, §2.1.1 see ex [4]). The following example show that the subject controls relativization (a) and topic has no control over this clause-bound linguistic phenomena :

(121) a. John，我喜欢他

John, wǒ xǐhuān tā
John 1SG. like him

'John, I like him'

b. John，他喜欢他自己

John, tā xǐhuān tā zìjǐ.
John 3SG. like 3SG. self

'John, he likes himself'

c. * John，我喜欢他自己

John, wǒ xǐhuān tā zìjǐ.
John 1SG. like 3SG. self

'*John, I like himself' Li and Thompson (1976:478)

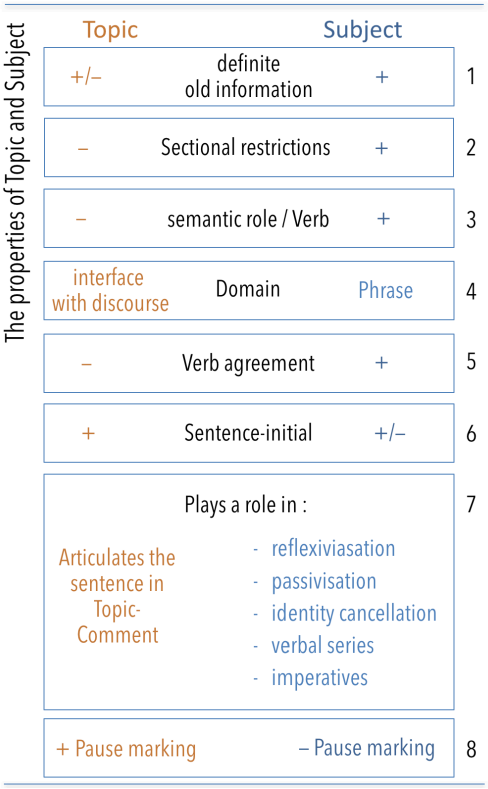


Figure 3.15 – The properties of Topic versus Subject as developed by Li and Thompson (1981), table adapter form M.C. Paris (1998).

Topics versus dummy subjects

Li and Thompson's typological claim puts forward an analysis of French sentence like (122 a) as being subject-oriented constructions in that they grammatically need expletive subjects for unaccusative verbs like raining that do not require any semantic subject. These kind of non-referential subject caled dummy subject are found in many subject-prominent languages like the German 'es', the English 'it' in "It's raining".

Hence, the contrast between (122a) and (b) is explained arguing that in a subject-prominent language "the structure of sentences favors a description in which the grammatical relation subject-predicate plays a major role."

(122) a. French dummy subject

Il fait beau.
It make-3SG. nice

'The weather is nice.'

b. 下雨

xià yǔ
falls rain

'It is rainy.'

c. 今天/这里好冷

jīntiān/zhèlǐ hǎo lěng
Today/here very cold

'It's very cold today/here.'

Although the Chinese version of the sentence "It is raining." in (b) features an Null-Topic, these kind of subject-less expressions about the whether are also typically uttered with a Topic setting a spatial or temporal frame like in (c), but still no (referentially empty) expletive subject has to fill the logical subject role in pre-verbal position.

The below examples further attest that subject-less sentence are totally commonplace in Mandarin:

(123) a. 今天城里有事

Jīntiān chénglǐ yǒushì
Today town-in have-business

'I have buisness in town today.'

b. 一个包子两口就吃完了

Yīge bāozi liǎng kǒu jiù chī-wán le
one-CL.. bun two bites then eat-RES.*finish* PRT.

'One bun, in two bits, it's gobbed.' or 'One bun is gobbed in two bites.'

Contrastively, if we consider the above Chinese basic sentence in (123a), where the subject is genuinely omitted and the following Japanese utterance in (124), we can clearly see Topic-prominent utterances are easily subject-less and specifically in (124) that Topics do not necessarily implies an anaphora in the Comment clause.

The sentence in (124) has no subject, and the word unagi 'eel' does not plays the role of a subject it is the predicate with the simple copula da 'to be'. Therefore, (124) shouldn't be translated in a 'subject-oriented manner' "As for Totoro, *he* is an eel" by adding a dummy subject, but " *Its for Totoro the eel*", and with the right syntactic analysis this utterance acquires its full coherence in the context of a dialogue at the restaurant.

The Topic has a proper function role in the Japanese sentence as a Subject of an object has, and this full-fledged functional role does not necessarily conveys emphatic information, overtly obtained though grammatical morpho-syntactic marking は *-wa* as opposed to が *-ga*, and for instance in (124) no particular emphasis is put on the eel, just that "the eel plate is for Totoro" when the waiter arrive at the table, which in fact in English is expressed by an anaphora in the Comment[90].

(124) トトロは鰻だ

Totoro-wa unagi da
Totoro-TOP. eel be

Interpretation: 'The eel, it's [for] Totoro'

90. It could either be a List Topic.

All the phrases hosted in Topic position

Among the salient characteristics Li and Thompson pointed out for Topic-prominent Languages is the fact that virtually any phrase type can be found in Topic syntactic position. Xu and Langendonen (1981:5; Jin, 1991) classified six categories that can be used as topics, to which we add wh-phrases topicalization. 'wh-topicalisation' (Xu and Langendoen, 1985; Wu, 1999) has been a recently studied phenomenon by Pan (2006 and 2007), who brought to light the existence of some contextual conditions for wh-fronting in Chinese.

Interestingly for our on going discussion on the sentence-discourse interface embodied by topic syntactic position, the value assigned to the wh-variable must be either taken from a pre-established discourse (see [125i] and [j]), or complex wh-phrases such as the construction *shenme NP* 'what NP' in (125h) are inherently linked with the discourse and can thus be fronted without being already present in discourse context (Pan, 2007).

We present here a list of examples showing the variety of Phrases types that can be hosted in Topic syntactic position in Mandarin Chinese, going from simple NP (noun phrase), to whole sentences S (sentence), VP (verb phrase), PP (prepositional phrase) or Post.P (post-positioned phrase), and wh-phrases:

(125) types of Phrase in Topic position (Mandarin)

a. NP: 这些话我不相信

Zhè-xiē huà wǒ bù xiāngxìn.
ThisPLUR.CL. words I not believe.

'I do not believe these words.'

b. Prep. P Locative Phrase: 在桌上他放了几本书

Zài zhuōshàng tā fàng le jǐ běn shū.
LOC. desk-on he put ASP someCL. books.

'He put some books on his desk.'

c. Post. P Localive Phrase: 桌子上有书，床上不会有书

Zhuōzi shàng yǒu shū, chuáng shàng búhuì yǒu shū.
table on have books, bed on not will have book.

'There are books on the table while there can not be any on the bed.'

d. ADV: [CP[TopP 其实 [Top'[Top° 呢][TP 他不来也好]]]]

[CP[TopP Qíshí [Top'[Top° ne] [TP tā bù lái yě hǎo]]]]
actually TOP. 3SG. NEG. come also good

'In fact, it's as well that he doesn't come.' Lü Shuxiang (1980:413).'

e. VP: 说这些话我不赞成

Shuō zhè-xiē huà wǒ bú zànchéng.
Say ThisPLUR.CL. words I not agree.

'I do not agree on saying such words.'

f. VP: [vP 吃饭]，张三很会，[vP 做事]，他一点都不会

[$_{vP}$ Chī fàn], Zhāngsān hěn huì, [$_{vP}$ zuò shì], tā yīdiǎn dōu bú huì
eat food Zhangsan very know, do matter 3SG. a.bit all NEG. know.

'Zhangsan, he for sure knows how to eat, but he doesn't know at all how to work.' Huang (1982:164,[93])

g. S: 他会说这些话，我不相信

tā huì shuō zhè-xiē huà, wǒ bù xiāngxìn.
He MODAL say these words, I not believe.

'I do not believe that he said these words.'

h. wh-words: [$_{TopP}$ 什么书张三买了]?

[$_{TopP}$ Shénme shū$_i$ Zhāngsān mǎi-le t$_i$]?
what book Zhangsan buy-PERF. t$_i$

'What book$_i$ did Zhangsan buy t$_i$?' (Pan, 2007)

i. context for (j): 小王、小刘、小李，这几个女孩当中

Xiǎo Wáng, Xiǎo Liú, Xiǎo Lǐ, zhè jǐ-ge nǚhái dāngzhōng
Xiao Wang Xiao Liu Xiao Li this few-CL. girl among

'Xiao Wang, Xiao Liu, Xiao Li, among these several girls'

j. context for (j):[$_{TopP}$ 谁张三最喜欢]?

[$_{TopP}$ Shéi$_i$ Zhāngsān zuì xǐhuān t$_i$]?
who Zhangsan most like

'Who does Zhangsan like the most?' (Pan, 2006)

All in all, examples in (125), show among others that Chafe's frame-like conception of the role of topic limiting the applicability of the predication (§3.1.3.3) is compatible with the large range of different phrases occurring in Topic position, including non-referential constituents such as clauses, verbal phrases, and adverbs and wh-phrases.

We differ the analysis of the syntactic manifestation of the different topic-comment relationships and configuration above, and of the dependency-links to following section §3.4.

In conclusion, as note by Paul (2010 and subsequent work), although multiple phrases and multiple sub-types of Topic constructions can be classified, the aboutness[91] or frame-like relation of topic and comment remain the two functions Topic plays in Chinese without being associated with a particular informational content (new versus old information).

Topic+Subject or Double Nominative constructions

Although topic and subject are different notions, they are not mutually exclusive inside the sentence-unit, and often co-occur.

Double nominatives constructions feature for instance the topic and the subject as two distinct NPs juxtaposed in clause-initial position, where both NPs could be viewed as carrying a subject-like relationship with the verb -this is why they are called double subjects by certain authors-, and the Topic is not co-referential with any syntactic element in the Comment-clause.

91. Including the shift towards a new Topic, called Aboutness-shift.

(126) その椰子は葉っぱが大きい。

Sono yashi-wa happa-ga ookii.
This palm-tree-TOP. leaves-SUBJ. big

'This palm-tree, its leaves are big."

(127) 那棵树叶子大，所以我不**喜欢** (那颗树)。

Nà-kē shù yèzi dà suǒyǐ wǒ bù **xǐhuān** ___[nà-kē shù]
this-CL. tree leaves big, so 1SG. NEG. like ___(this-CL. tree)

'Cet arbre ses feuilles sont grandes, donc je ne l'aime pas.'

Considering the sentence pattern [NP1+NP2+VP][92], it can be divide into to main types depending on whether a certain semantic relationship holds between the two sentence-initial NPs. One would therefore distinguish on the one hand, the so-called "Double nominatives constructions" and on the other hand the Frame-setting or Left-Dislocated Topics.

Double nominatives constructions's most frequent semantic relationships between the two sentence-initial NPs are part-whole, possessor-possesee or subset-item relations and the verb in the comment clause is often a stative or an intrasitive predicate. Their high frequency of Double nominatives constructions is a proto-typical and salient feature of Topic-Prominent languages, As shown by the examples in (126) and (128a), these constructions semantically characterize the whole by means of marking a predication about a property (Chappell 1996:467). While (126) is the Japanese version of the Chinese sentence in (102), reported in (127) for convenience, (128a) features a typical inalienable possession relationship[93].

(128) Part-whole Double Subject

a. 象鼻子长

xiàng bí-zi cháng
Elephant[topic] trunk[subject] big

'The elephant, its trunk is long'.

b. 象的鼻子长

xiàng de bí-zi cháng
trunk of.GEN elephant[subject] big

'The elephant's trunk is long.'

The interpretation subtlety of the contrast between (128a) and (128b) could be expressed as follows: while (a) leads the interlocutor consider the elephant as a whole and than focus on an aspect of it its nose, (b) only attracts the attention on the nose of the elephant.

92. A pattern that in Chinese Linguistics is called "主主谓" *zhu zhu wei*. For an overview on unconventional [NP1+NP2+VP] in Mandarin see Lin (1994), and 杨启光 (2012).

93. Two main corpus studies were carried over to determin the use or non use of the genitive particle DE in possessor possessee relations between NP1 and NP2 in Double subject constructions. See the Chappell and Thompson (1992) for a written corpus analysis and Liu H-Y (PhD 2002) for a oral corpus analysis.

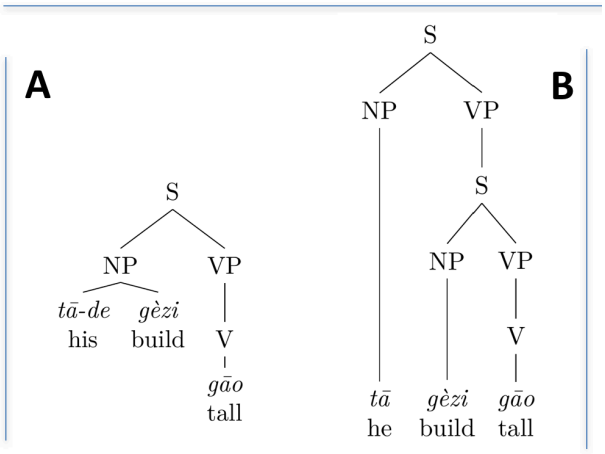


Figure 3.16 – Double subject constructions and simple SVO with a genitive subject have two distinct underlying forms and are non-derivable one for another.

The possessive and metonymical link between the two sentence-initial NPs is so central in the definition of these construction [94] and the interpretative similarity too that one could be tempted to consider them as derive one form another, so that the Double nominative construction would be the result of the ellipsis of the genitive marker DE, yielding the pattern [NP + NP + VP] (Hasimoto, 1971; Tang, 1972). However, the two different underlying structures in Figure 3.16 can be analyzed following Teng (1974).

(129) a. 这个人个子很高。

Zhè-ge rén gèzi hěn gāo.
This-CL. man height adv.very tall

'This person his build is tall.'

b. 这个人的个子很高。

Zhè-ge rén DE gèzi hěn gāo.
This-CL. man DE height adv.very tall

'This person's build is tall.'

c. 这个人啊/呢个子很高。

Zhè-ge rén a/ne gèzi hěn gāo.
This-CL. man TOP(a/ne) height adv.very tall

'This person his build is tall.'

d. * 这个人的啊/呢个子很高。

Zhè-ge rén DE a/ne gèzi hěn gāo.
This-CL. man DE TOP(a/ne) height adv.very tall

94. Chappell (1996:466) advanced that double subject constructions serve as a major syntactic means to express in Mandarin the notion of inalienability.

'This person's build is tall.'

As repeatedly observed in the literature (Mullie, 1933; Teng, 1974 and Chappell, 1996), the main linguistic tests proving the non-derivability are namely the separability of the two sentence-initial NP by adverbs (e.g. *hai* 还) or negation (e.g. *bu* 不), and the possibility to add contrastive or emphatic topic particle (-ne and -a) in between like in (129d)[95].

Thus, the impossibility of inserting a particle in the genitive NP of the SVO pattern grounds the analysis that double subject constructions and simple SVO with a genitive subject have two distinct underlying forms like in Figure 3.16 and are non-derivable one for another.

In our fMRI study on Chinese sentence's Left-Periphery (chpt. 7), we will leverage on this contrast to distinguish two case of external merge in which the Fist-NP role is in case (a) the determiner of the genitive subject and in (b) the Topic of the whole sentence with the advantage of having both a minimal surface difference between the two structures and secondly a minor interpretative difference as previously noted.

Applying this test to sentences where no possession relationship exist between Topic and subject like in (130), but a more complex part-whole one- further attests the non-derived character of these Topic-Comment type.

(130) Impossible derivation Double Subjects

a. 他们谁都不来

Tā-men shéi dōu bù lá
3SG.-PLUR. who adv.all NEG. come

Lit. 'Of them, none came. None of them came.' (Teng, 1974 ; Chappell, 1996).

b. 他们啊谁都不来

Tā-men a shéi dōu bù lái
3SG.-PLUR. TOP who adv.all NEG. come

NONE of them came.

c. * 他们的谁都不来

Tā-men de shéi dōu bù lái
3SG.-PLUR. DE who adv.all NEG. come
*

We can effortlessly admit that (a) is difficult to paraphrase into a genitive construction and that there would be no advantage in considering the above structure as being derived from a subject-predicate sentence like (130c.), having a genitive phrase as subject.

In sum, the fact that these sentences are irreducible to subject-predicate ones adds linguistic evidence to a central argument for Topic-prominence, namely that Topic-Comment articulation is the proto-typical and basic way to build the sentence-unit in Chinese. Hence, next section will be dedicated to bring arguments from Chinese First-language acquisition studies that further attest how the different Topic-comment constructions in Chinese are indeed *basic* sentence's articulations.

95. As noted already by Tsao (1990), double nominative constructions are not an homogeneous linguistic phenomenon. The frequently cited example for this in Chinese Linguistics is the case where the the second subject is part of la lexicalized compound like 'head-hake' *tou-teng* 头疼. For a discussion see Shi (2000).

3.2.4 The notion of Topic and its Acquisition

> *"In psychology one may or may not be a behaviorist, but in linguistics one has no choice. Each of us learns his own language by observing other people's verbal behavior and having his own faltering verbal behavior observed and reinforced or corrected by others. We depend strictly on overt behavior in observable situations. As long as our command of our language fits all external checkpoints, where our utterance or our reaction to someone's utterance can be appraised in the light of some shared situation, so long all is well. Our mental life between checkpoints is indifferent to our rating as a master of the language."*
>
> W. V. Quine, 1987:5.

A universal stage of Topic-Comment utterances has been cross-linguistically observed in First Language Acquisition (cf. §3.1.3.6) starting from the early two-word combinations children produce to convey their communicative intentions. Child language represents a privileged observation ground to investigate the emergence of linguistic patterns typical of a given language. As previously mentioned children use the specific word-order of their language a early stage even when it show relative complex syntactic configurations. The case of Sesotho children who acquire and produce passives at early of 2;8 (Demuth, 1992) was instrumental in showing that if exposed to a complex syntactic construction that is central to the grammar of their mother-tough (i.e. produced frequently in child directed speech) children have no difficulties in acquiring the schema of the canonical sentence in their mother-tongue even at early stages of the acquisition of syntax (Slobin and Bever, 1982) (cf. §2.2.2, p. 105)

Initial evidence has been gathered that Chinese children begin to produce strictly canonical SVO order sentences at an early stage presenting deviations from basic word-order until they access to the control of the basic sentential relationships (Erbaugh, 1992:416). However, parallely to the SVO basic word-order, the prevalence of topic-comment structures is one of the most distinctive features of Mandarin Chinese, and given the lack of morphological markers for agreement, number, gender or case, word-order has been put forward in the literature as the most important syntactic device for sentence interpretation in Chinese (cf. Chang, 1992). From this two central characteristics of Chinese language the question rises of how the SVO and topic-prominence features of Chinese can be observed and interact in children linguistic development.

We will concentrate on two experimental studies that have shown that Chinese-speaking children in the early stages of syntax acquisition are capable of distinguishing the notions of Topic and Subject (Chien, 1983; Chien and Lust, 1985). Then, we will address the question of which Topic-comment's syntactic configuration is first spontaneously produced and mastered (Chen 2009) and then contrast topic and Subject

topicalization in Children comprehension (Hu, 2015).

3.2.4.1 Acquisition of the *notion* of Topic in Chinese children

By first considering the issue of the difference between Topic and subject in child grammar, the study by Chien (1983) represent the one that most directly addressed this point. Examining the comprehension of subject and topic by children acquiring Chinese as their first language in a quite unusual way using meta-linguistic grammatical judgments, she was thus able to investigate the notions of topic and subject. Children were asked to make judgments on the grammaticality of sentences based on their knowledge of subject and topic, and the results indicated that children are sensitive to the distinction between the notions of subject and topic at their early stages of syntax acquisition.

To further examine the sensitivity of young Chinese children the distinction between Topic and Subject, Chien and Lust (1985) used an an elicited imitation task to test 95 children (aged from 2;6 to 5;0). They asked children were ask, this time, to imitate coordinate sentences and control-sentences, respectively reported in (131a) and (b).

(131) Chien and Lust (1985) test and control sentence

a. Test: 宝宝，脚很小；宝宝也很可爱。

Bǎobao jiǎo hěn xiǎo; *bǎobǎo* yě hěn kěài.
baby feet very small baby also very cute

'As for the baby, the feet are small; as for the baby, (he) is also very cute.'

b. 小花，姐姐喜欢小花戴帽子

Xiǎohuā jiějie xǐhuān *Xiǎohuā* dài màozi .
Xiaohua older sister like Xiaohua wear hat

'As for Xiaohua, (his) older sister likes Xiaohua to wear a hat.'

While in (131a), the first occurrence of the NP baobao *'baby'* is the Topic having as Comment jiǎo hěn xiǎo '*feet very small*', the second occurrence of baby is redundant in Chinese, because yě hěn kěài '*also very cute'* can directly be a comment of the first occurrence of the NP baobao '*baby*'. Importantly the second occurrence of this NP was added intentionally for the needs experimental setting.

The logic behind this experimental task was to observe what sentential element between topic and subject would have been deleted in the reformulation required by the experimenter. So that if children interpret the first NP as the Topic, they should repeat the sentence by deleting its second occurrence (in italics, above), as shown in (132), which gives rise in Chinese to a perfectly grammatical sentence having the same interpretation as (131a).

Similarly, in the control sentence (131b) the first occurrence of the NP 小花 *Xiǎohuā* is the sentence topic and its comment is a control sentence, its second occurrence controls the PRO subject of dai maozi 'wear a hat'. If the children were to omit it, such reduction would result in a change of meaning in that the phrase 姐姐 jiějie '*old sister*' would actually control the PRO as shown in (132b), which is well-formed, but but does not mean the same as (131b).

(132) Chien and Lust (1985) patterns of omission of the repeated NP

a. 宝宝，脚很小；也很可爱。

Bǎobao$_i$, jiǎo hěn xiǎo; e$_i$ yě hěn kěài.
baby feet very small also very cute

'As for the baby, the feet are small; [he] is also very cute.'

b. 小花，姐姐喜欢戴帽子。

Xiǎohuā jiějie xǐhuan PRO$_i$ dài màozi.
Xiaohua older-sister like wear hat

'As for Xiaohua, (his) older sister likes to wear a hat.'

The results showed that children dropped less the first NP in a coordinate construction like (131a) than in a control-construction like (131b), respectively 2.02% against 43.16% of the times. Secondly, children were reported to omit the second NP more frequently in the coordinate (21.71%) than in the control (2.49%).

From these patters, the authors proposed that Chinese children were indeed able to distinguish the notion of topic and subject at least from the age of three.

3.2.4.2 Types of Topic-Comment in children: Chen 2009

A detailed account of the different types of syntactic configurations fond in Chinese Topic-comment sentences is offered by an large scale study by Chen (2009), who analyzed the utterances of 44 children from the CHILDES database (MacWhinney, 2000) as shown in Table 3.2.4.2.

As shown by the above Table 3.2.4.2 Children were divided into four age groups of a dozen of children each : the 2;2 age group, the 2;8 age group; the 4;0 age group and the 6;0 group. Topic-Comment constructions are present already since children early utterances in the first group, and age 4 appear to assist to an increase in frequency of production, but the frequency does not differ statistically between ages 4;0 and 6;0.

Table 3.6 – Total number of clauses and topic-comment clauses produced by children of four age groups. Adapted from Chen (2009).

	Age 2;2 (10)	**Age 2;8** (10)	**Age 4;0** (12)	**Age 6;0** (12)
Total number of sentences	387	227	875	1009
Mean number of sentences	38.7 (SD = 31.0)	22.7 (SD = 14.7)	72.9 (SD = 59.3)	84.1 (SD = 48.0)
Total number of topic-comments	5	8	34	35
Mean number of topic-comments	0.5 (SD = 0.71)	0.8 (SD = 0.92)	2.83 (SD = 3.76)	2.92 (SD = 3.00)
Average % of topic-comments produced per child	1.20% (SD = 1.84)	2.89% (SD = 3.69)	3.55% (SD = 2.85)	3.38% (SD = 3.41)

The above results were broken down by age and types of topic-comment structures, and the observed different types of topic structures that where found are summarized in table in Table 3.7.

Table 3.7 – Breakdown of types and of age groups and Topic-Comment structures.

		Types of Topic-Comment structures			
Age	**Occurences**	Resumptive	Topicalized	Double nominatives	Adverbial
2;2	5	0	100% (5)	0	0
2;8	8	0	87.5% (7)	0	12.5% (1)
4;0	35	32.4% (11)	14.7% (5)	38.2% (13)	14.3% (5)
6;0	34	45.7% (16)	11.4% (5)	14.3% (4)	28.6% (10)

The result from Chen (2009) suggest that there is little variation in the type of Topic-Comment structures produced by the children in two earlier age groups, respectively the 2;2 and 2;8 age.

Topicalization Children begin to produce topicalized structures shown in (133) as early as age 2;2, and out of the 8 topicalized sentences (3.5%, out of 227 utterances) produced by the 2;8 age group 7 object topicalization sentences. We can take this finding to confirm form a developmental point of view that Null-Topic parameter correlates with the drop of the objects - the so-called object drop Parameter in Chinese as it was proposed by (Huang, 1984 and 1989) (cf. §2.2.4.4 p. 124).

(133) Gapped type: 李先生我认识

Lǐ xiānsheng$_i$ wǒ rènshi e$_i$.
Li Mister I know
'As for Mr. Li, I know [him]'

Namely, for the two earlier age groups (2;2 and 2;8 age), the majority of the topic-comment structures produced are topicalized clauses, while topic-comment structures with resumptive pronouns or double-nominatives were not produced.

The production frequency patterns in Table 3.7 show a decline of topicalized clauses between ages 2;2 and 4;0. This could be due to the appearance of other topic-comment types. Namely, the 4;0 age group, features the emergence of other types of topic-comment clauses, which could have caused the decline in percentage of topicalized utterances we can observe at age 4;0 and 6;0 .

Resumptives From the age of 4;0 to 6;0, the resumptive topic-comment structures and double nominatives are the most represented Topical utterances (topicalized structure being the most infrequent at this stage). Within the resumptive pronoun category, the majority clauses have a topic NP and a resumptive pronoun in the comment clause (25/27, 92.59%), whereas only a few utterances feature a unusual configuration where the Topic is a pronoun and the subject in the comment is a resumptive NP.

(134) Resumptive type: 李先生我认识他

Lǐ xiānsheng$_i$ wǒ rènshi tā$_i$.
Li Mister I know he

'As for Mr. Li, I know him'

Double-nominatives Double nominatives, appear only after age 4 they are the latest type to appear. As for the semantic relationship between the Topic and the Comment, out of the 17 items identified, 12 had a possessor-possessed relationship like in (136a), one was a domain-subset type (136b), and in the last 4 Topic and Subject were semantically related NPs.

(135) Double nominaitve - Aboutness

a. possessor-possessed relationship: 长颈鹿脖子长。

Chángjǐnglù bózi cháng.
giraffe neck long

'As for giraffes, their necks are long.'

b. domain-subset: 水果，我最喜欢樱桃。

Shuǐguǒ wǒ zuì xǐhuan yīngtáo.
fruit I most like cherry

'As for fruits, I like cherries best.'

Adverbial Phrases Within the adverbial phrases type, temporal adverbials account for the large majority of this category (80% - 12/15), while spatial adverbials account for 20% (3/15).

(136) Double nominaitve /Aboutness: 昨天晚上我没睡觉。

Zuótiān wǎnshang wǒ méi shuìjiào.
yesterday evening I not sleep

'As for last night, I did not sleep.'

As for the evolution in the types of topic-comment structures produced there is some increase between ages 2;2 and 2;8, but not all four types are produced yet. Importantly, by age 4;0, all four types of topic-comments are present in child speech, and in the age group 4;0 and 6;0, more topic sentences were found, 34 sentences (3.9%, out of 875 utterances) and 35 sentences (3.5%, out of 1009 utterances) respectively.

Although Table 3.7 seems to indicate a certain number of differences in the use of different topic constructions between age 4 and age 6, the statistics hardly show any significant difference. For instance, the fact that four year old children seems to produce significantly more double nominatives than six year old children is not statistically supported[96].

Another central result from this study is the observed similar frequency distribution between early children utterances and adult speech in terms of frequency. The author put

96. A total of 13 tokens were found among four year olds, and among these eight were found in the same child.

Table 3.8 – Comparing different movements types.

Type of TV show	Percentage of Topic-Comment clauses
1. Political	5.17%
2. Leisure (food)	4.69%
1. Fashion	4.27%
4. Entertainment	2.02%

children acquisition fequency of production in perspective with a small sample of adult data from four TV talks and observed that the percentage of topic sentences produced by adults was not significantly different from those reported across the different children age groups, it ranged from 2.02% to 5.17% across four talks as show in Table 3.8. All in all this comparison shows that, although topic structures are produced from a very early age, they are not more abundant in the spontaneous speech of Chinese children compared to adults.

In conclusion, these results can be taken as suggesting of the representational or processing difficulty of the different types of Topic-Comment syntactic configurations that Chinese offers.

The late appearance of Double Nominatives and resumptives echoes the observed production patterns of Topic-Comment structures in L2 learners (cf. §3.2.1.3). While the astonishing early appearance of object topicalized sentences could be interpreted as a mark of Topic-prominence and it echoes the analysis proposed by Huang (1984) about Mandarin Chinese object drop (cf. §2.2.4.4, p. 124), a phenomenon that is also widely attested trough the Chinese linguistics literature. He noted that Null-Topic parameter was correlated with another one featuring the drop of the object - the so-called Pro-object drop Parameter[97].

3.2.4.3 Comprehension of Object and Subject topicalization

As we just saw, the analysis of spontaneous production of Chinese children utterances showed that topic structures are produced by children starting from their first multi-word combinations and although this study couldn't put forward an enhanced frequency of usage of T-C constructions in this population, it revealed nonetheless an interesting tendency to produce more object topicalizations than Subject ones.

In a recent study, Hu Shen 'ai (PhD, 2014) addressed the question of the basicness Topic-Comment constructions in Mandarin Chinese by examining the comprehension of topicalization, where crucially, the difference between object an subject topics was tested as illustrated by the two different kinds of Topic-comment structures in (137).

(137) Experimental sentence in Hu (2015)

a. Topic-Object: 这个孩子 (呀)，外婆在画

Zhège háizi$_i$ (ya), wàipó zài huà PRO$_i$.
this-CL. child (TOP.) grandma PROG. draw

'As for this child, the grandma is drawing [him].'

97. Note that this type of object -drop trend in children linguistic behavior has been observed alos in other languese see Pérez-Leroux et al. (2011)

b. Topic-Subject: 这个孩子 (呀)，在画外婆

Zhège háizi$_i$ (ya), PRO$_i$ zài huà wàipó.
this-CL. child (TOP.) PROG. draw grandma

'As for this child, [he] is drawing the grandma.'

In (137a) the Topic and the subject are two different referents and differ in terms of argument roles in relation to the verb, while (b) displays an overlap between the Topic and the subject (actor), and the subject is instantiated by a silent pronominal element, *pro* (in bold; cf. §3.4.4.2 where we will analyze in details this silent pronominal element, p.390).

Notably, we already mentioned this configuration, where topic and actor role overlap, when reviewing the utterance construction strategy of Chinese advanced learners of English and English learners of Chinese, who adopt a similar strategy of preferring word-orders where the subject coincides with the Topic (cf. §3.2.1.3, see example [84]).

As illustrated by Table 3.17, experimental results of this picture elicitation show that, although the accuracy rates of OSV topicalization sentences were numerically lower than SVO topicalization sentences (e.g., at age three, 88% vs. 76%; at age four 89% vs. 84%), no difference between the two structures reached significance.

Groups	OSV topicalization sentences				SVO topicalization sentences			
	%	N	M	SD	%	N	M	SD
3 y.o.	76	121/160	6.05	1.50	88	141/160	7.05	0.89
4 y.o.	84	134/160	6. 70	1.26	89	142/160	7.10	1.02
5 y.o.	93	149/160	7.45	0.10	96	154/160	7.70	0.57
6 y.o.	98	156/160	7.80	0.52	100	160/160	8.00	0.00
Adult	100	80/80	8.00	0.00	100	80/80	8.00	0.00

Table 3.9 – Percentages (%), raw scores (N), means (M) and standard deviations (SD) of correct responses in each age group for the picture matching task evaluating children's comprehension of who is doing what to whom in Topic-Comment sentence like in examples above.

Interestingly no statistically significant difference is found between the two Topic-Comment constructions in a quite large number of participants (80)[98] aged from 3;0 to 6;11 and divided into 4 age groups of 20 participants each (see Table 3.17). The only statistically robust result is that a significant improvement is observed at age 5.

This being said the most important result from this study is to have shown that Chinese children do not seem to have trouble establishing the relation between the topic and the empty category in the comment clause. Critically, it demonstrated that there was nothing intrinsically hard about Topic-comment structures, either when the object is found in anon canonical word-order position or when the canonical SVO linearization is

98. Note that the participant s lived in Zhejiang province (PRC). To this population applies our remark on southerns dialects influence on Topic marking acceptability in Mandarin (see §3.2.2.3, p.300 or p.xxiii).

respected. this acquisition pattern suggest that topic-comment is not only a optimal way to analyze Chinese sentences but that is early established in first language acquisition.

To sum up, acquisition studies have shown on the one hand that Chinese children sensitivity to topic subject distinction emerges as age 3, demonstrating in this way an early ability to distinguish between the grammatical function topic and the subject (Chien); on the other hand, that Chinese children are at ceiling in the comprehension of Topicalized sentences at age 5, and that no difference in comprehension performance is observed between subject and object topicalized sentence structures.

In addition, database analysis of Chinese children early spontaneous utterances showed that :

1. (1) they start spontaneously producing as early as age 2;2 topic structures having the following pattern [NP+NP+verb] ,
2. (2) that among the topic structures produced by children up to three years of age, the majority were object topicalizations, and that
3. (3) it is only after age 2;8 that they begin to produce adverbial topics and objects topicalization.

All in all these findings corroborate the argument that Topic-Comment sentences belong to the basic utterance structure of the linguistic system of Chinese, and suggest that the Topic-prominence feature of Chinese makes it easier for children to acquire both subject/topic notions and syntactically complex topicalized gapped-sentence structures, and that no particular comprehension difficulty is observed between Subject and object topicalization.

3.3 Topics and Post-verbal subjects in Mandarin, French and Neapolitan

After having reviewed the linguistics studies surrounding the phenomenon of Topic-Comment constructions in Typology, and more particularly the way it appears in Mandarin Chinese, we undertake here a linguistic study, to put forward some linguistic aspects that can be good candidates for having an effect during the online processing of Topic-Comment articulations in Mandarin Chinese. We will put together linguistic evidence for a possible interplay between Topic-Comment syntactic structure with verbal argumental structure of the main verb in the Comment-clause. Our analysis is achieved by delineating an account for the observed tendency of the sole arguments of unaccusative and unergatives verbs to be constrained by Topic construction in their surfacing to the left or to the right of the verb.

A close look at the linguistic data will be instrumental to interpret the data obtained during brain processing of Chinese Scene-setting Topics in chapter 3 and 4. A more linguistic account may also reveal some of syntactic aspects of Topic-Comment sentence articulation. The basic point we will argue here is that the presence of a Frame-setting Topic in the sentence has an influence on the comment-clause and specifically on the linearization of the argumental structure of the verb it hosts. The details of what this observation could reveal about the Topic-comment syntactic articulation will remain open to discussion.

3.3.1 Double-subject sentences, cross-linguistically

Cross-linguistically, double-subject construction (or double nominatives sentences) are attested in a number of languages that haven't been described as topic-prominent and semi-topic-prominent languages (e.g. Mandarin Chinese, Korean and Japanese), like some dialects in Europe. Neapolitan dialect will introduce us to a peculiar linguistic phenomenon illustrating the impact of topics in the internal structure of the Comment clause, which we will be further analysis in Chinese.

This frequent sentence type inside the group of dialects spoken in the Campania region (Southern Italy), features two subjects, one a third person pronominal *chillo* as first Subject, meaning "that one," and a lexical DP as second subject, that are co-referential and hence displaying agreement for person (third) and gender. However, compared to Japanese, Korean and Chinese, in Neapolitan, these Double-subject constructions represent a marked sentence type, mainly playing the role of Topic-shift introducing a new topic[99].

Consider the following example from Ledgeway (2011):

(138) Neapolitan Double-subject construction:

a. DSubj1 . . . (DPSubj2) . . . Vfin . . . (DPSubj2)

Chella$_i$ [$_{DP}$ 'a fibbia]$_i$ s'è rotta.
that-one.FEM the.FEM buckle.FEM self=is broken

Lit. *That one$_i$ the buckle$_i$ has broken.* 'The buckle has broken.'

b. Neapolitan Double-subject and post-verbal Subject-2 (DPSubj2)

99. In information-structure terms Subject-2 is introducing a 'new topic'.

Chella$_i$ s'è rotta [$_{DP}$ 'a fibbia]$_i$.
that-one.FEM self=is broken the.FEM buckle.FEM

Lit. *That one$_i$ has broken the buckle$_i$.* 'The buckle has broken.' (Ledgeway, 2010)

The two sentences in (139) exemplify a frequent pattern in Neapolitan where the second Subject (Subject-2) can alternatively surface left of its finite verb, as in (139b), or may occur in post-verbal position as in (139c). These sentence structure have been analyzed as Topic-Comment articulations, in that intervening material between the two subjects confirms they do not for a constituent. Moreover, prosodically, the intonational contour of such constructions is that of canonical topic-comment structures and not that of rhematic sentences (i.e. no single prosodic contour arching over both nominals is found), and syntactically the Topic status of the Subject-1 is confirmed by Ledgway's analysis stating that Subject-1 targets the Topic field within the Left-Periphery.

Other examples of Neapolitan double-subject construction show that the surfacing of the Subject-2 in post-verbal position is found out across different argumental structures:
.

(139) Neapolitan subject inversion in the Comment

a. Neapolitan Double-subject construction

Chillo$_i$ [Arturo]$_i$ m'ha mannata na lettera.
that-one Arturo me.has sent a letter

'Arturo sent me a letter.'

b. Neapolitan Double-subject construction + subj. inversion - Transitive

Chella$_i$ me vede [$_{DP}$ muglierema]$_i$.
that-one.FEM me sees wife.FEM=my

'My wife will see me.'

c. Neapolitan Double-subject construction + subj. inversion - Unaccusative

Aspetta Rafè, chillo$_i$ mò esce [[$_{DP}$ lo patrone]$_i$.
wait Raffaele, that-one.m now exits the.m.sg boss.m

'Wait Raffaele, the boss is just leaving.'

d. Neapolitan Double-subject construction + subj. inversion - Unergative

Chillo$_i$ sta aspettanno [$_{DP}$ lo forastiere]$_i$ abbascio.
that-one.m is waiting the.m.sg stranger.m down

'The stranger is waiting downstairs.' Ledgeway (2010)

Double subject construction is not limited to unaccusative clauses like in (139) -it can equally occur with transitives (139a) and Unergatives (d), and crucially, informants report preverbal and post-verbal Subejct-2 as interpretively identical.

Ledgeway (2011) carried out an in-depth examination of the distributional and co-occurrence restrictions operating on both subjects in the double-subject construction and showed that although double subject construction is not limited to unaccusative, the vast majority of post-verbal Subject-2 has been found in Unaccusatives: there were only four transitive and three unergative examples in the corpus.

In this respect, the behavior of unaccusative predicates in Italo-Romance languages should be clarified, in that it generally contrasts with that of transitives and unergatives. Unaccusative in Italo-Romance languages generally license rhematic subjects, a distinction that is overtly marked in the syntax by their typical post-verbal position (see Hulk and Pollock, 2001 for an overview)[100]. Yet, this property was observed in the corpus study by Ledgway (2010) too, unaccusative subjects in Double-subject constructions appear mainly (71% of a corpus of 111 sentences with unaccusatives) in pre-verbal position when topical and keep their post verbal positon in 44% of the cases. A pattern (cf. ex [139c and d]) that is interpreted by Ledgway:

> "to mean that raising of Subj2 to the preverbal subject position is not a necessary condition, insofar as the intended topical interpretation of Subj2 is in any case already ensured by the presence of Subj1 which licenses the preverbal subject of predication feature" (Ledgeway, 2010:262).

3.3.2 Subject-inversion and Topic in French

The observation of Neapolitan offered an insight on some linguistic phenomena tight to the presence of a Topic in the sentence-unit, showing that the presence of a Topic in the sentence has an impact on the Comment-clause's internal organization and particularity on the surfacing of the comment subject in post-verbal position.

Following the thread build though Ledgway's arguments, if we can consider some cases of French subject inversion in particular declarative sentences as illustrated in (140):

(140) French subject inversion in simple declaratives

a. Subject inversion with spatial prepositional Phrase

Dans la cour, régnait l'animation habituelle.
In the courtyard reigned the.animation habitual

Lit.: 'In the courtyard reigned the habitual animation.' (Brincourt)

b. Subject inversion with a modal adverbial

Soudain sont entrés les enfants.
All-of-a-sudden are entered thePL kids

Lit.: 'Suddenly entered the kids.'

c. Modal adverb

Soudain les enfants sont entrés.
All-of-a-sudden thePL kids are entered

Lit.: 'Suddenly the kinds entered.' Lahousse (2003)

d. Subject inversion with an Unaccusative

La plage était calme, incroyablement calme. Arrivait l'heure où le soleil
The beech was calm, incredibly calm. Arrived the.time where the sun
exténué glissait vers l'ouest.
exhausted slid toward the.west

100. For detailed accounts see Benincà, 1988:168–170; Salvi, 1988:54; Lepschy and Lepschy, 1994:146.

'The beech was calm, incredibly calm. Lit.:'[had arrived] the time when the sun was sliding down toward west had arrived.' (Déon)

e. Subject inversion with a spatial deictic adverbal

Je m'inscrivis au club de tennis de V. **Là** se nouaient des idylles avec
I myself.subscribed to club of tennis of V. There were tided thePL idylls with
[...]
[...]

'I subscribed to the tennis club of V. There were tided the idylls with [...].' (Hébrard et Velle)

In this regard, a number of authors have established a link between Subject Inversion information conveyed Scene-setting Topics, establishing a locative frame or an adverbial frame to the predication of the Comment. Tasmowski and Willems (1987:182) have analyzed subject inversion in main clauses establishing a tripartite structure like [Loc + V + S] where the first constituent a pre-positional phrase is a Frame Topic. Similarly, Fournier (1997:119) interprets this pattern in French as a configuration where the Spatial preposition is topicalized, an analysis that can be found also in De Bakker (1997:206), for whom it is the presence of a topicalized constituent that explains the subject inversion in main clauses starting with a Prepositional Phrase.

The topicality of adverbials However, as noted by a number of studies (e.g. Le Bidois, 1952; Marandin, 2001 and Lahousse, 2003) prepositional phrases are not the only elements whose presence Topic position correlates with an inversion in the comment[101], as shown in the examples above. As noted by Lahousse (2003), all together these French examples in (140) converge in showing that the notion Topic allows to have a unified approach of all types of phrases yielding subject inversion. Going from spatial prepositions 'Dans la cours' *in the courtyard*, modal adverbs 'soudain' *suddenly*, deictic adverbs 'là'[102], all these types of adverb can constitute the frame of the comment proposition[103].

Emphasizing the topical status of sentence-initial adverbials, Le Querler (1993:177) analyzes this French topics as "circumstantial complement have a framing function when

101. We have to distinguish here situations where the inversion is frequent (the so-called free inversion) from those where the Inversion is obligatory. Frequently attested cases of Inversion are found after locative and temporal and modal adverbs (e.g. ici, là, derrière; alors, après, enfin, puis, parfois, bientôt; brusquement, lentement, soudain, du coup, peu à peu). While contexts where inversion is obligatory are principally the following:

1. (a) in direct interrogatives starting with wh-word, like in the following examples *Partira ton ami? Quand partira ton ami? or 'Où est Jean ?' versus '*Où Jean est ?'. Kayne and Pollock (1978:596-597)
2. (b) in indirect speech like in: Ele dit que partira ton ami.
3. (c) the so-called complex inversion in French is also blocked in these contexts as shown by the agrammaticality of the following sentences: *Quand ton ami partira-t-il? Similar restrictions are observed in other Roman languages (cf. Pollock et al., 1999).

102. 'Là' *there* is called by Fournier (1997) with he term of anaphoric adverb.

103. However, see Rizzi 2002: "The assumption that preposed adverbials may be full-fledged topics is not very plausible on interpretive grounds. [...] A preposed adverb seems to have something in common with a topic, the fact of being made prominent by movement to the left-periphery, but it does not share with the topic the necessary connection to the background".

it fixates a spatial localization or a temporal one inside which the remaining of the utterance takes place[104], a definition that echoes the one by Nikolaeva (2001) of "scene-setting clause-external topics".The grammaticality contrast between (141a) and (b) against (141) shows that inversion is allowed by the presence of a scene-setting adverbial in (a), an either its absence in (c) or its non-canonical topic position in (b) are infelicitous.

(141) Spatial-Frame Topic in French

a. Topic position of the spatial adverbial

Ici coexistent les données anciennes.
here coexist the ancient data

'The ancient data coexist here.'

b. ? Clause-internal position of the spatial adverbial

Coexistent ici les données anciennes.
coexist here the datas ancient

'The ancient data here coexist.'

c. * Agrammaticality in absence of adverbial framing

Coexistent les données anciennes.
coexist the data ancient

'*Coexist the ancient data.'(Lahousse, 2003)

Crucially, the notion of scene-setting topic by Erteshik-Shir (1997:26-27) brings an additional definition to these Topical framing phenomena:

> "the stage topic (sTOPt) defines the spatio-temporal parameters of the utterance. Stage topics may be overt ('this afternoon', 'on Park Avenue'), or discoursally implied. [...] The term "stage" here [...] refer[s] to the Time/Place at which the event expressed by the sentence takes place. [...] the event can be viewed as taking place on the stage defined by this topic."[105]. Erteshik-Shir (1997:26-27)

(142) Obligatory subject-inversion with Topic spatial framing

a. Au coin de la rue Beaune et du quai Voltaire se tient un restaurant où l'on mange médiocrement.
Lit.: 'At the corner between Beaune street and quai Voltaire stands a restaurant where one can eat badly.'

b. * Au coin de la rue Beaune et du quai Voltaire se tient un restaurant. Lit.:'At the corner between Beaune street and quai Voltaire stands a restaurant.'

c. ?? Au coin de la rue Beaune et du quai Voltaire un restaurant où l'on mange médiocrement se tient.
Lit.: 'At the corner between Beaune street and quai Voltaire a restaurant where one can eat badly stands.'

104. Le Querler (1993:177) in French "pour un circonstant thématisé, on dira qu'il y a cadrage quand le circonstant en première position fixe une localisation spatiale ou temporelle dans le cadre de laquelle le reste de l'énoncé se situe".
105. This definition evidently recalls that of Chafe (1976): "The topic sets a spatial, temporal [...] framework within which the main predication holds (the frame within which the sentence holds)"

From (142) we can see that under certain condition Sentence requiring a spatial framing are un-grammatical when inversion does not take place, inversion is obligatory int hes situation unless some additional information is attached to the pre-verbal subject like in (142c) (cf. Kayne et Pollock [2001:155] for a discussion, and Fournier [1997:101] for similar examples).

Movement and Subject-inversion

Importantly, Subject inversion is allowed in presence of Scene-setting topics, while its not in the case of other pre-positonal phrases that count as Left-dislocated Topics as show by the sentences below:

(143) Topicalized arguments do not allow subject-inversion

a. À Jean, Marie pensait souvent le soir, quand il faisait noir .
To John, Mary thought often in.the evening, when it was dark
To John, Mary often thought about in the evening when it was getting dark.

b. À Jean, Marie plaisait souvent le soir, quand il faisait noir.
To John, Mary was.appealing often in.the evening, when it was dark
'To John, Mary was often appealing in the evening, when it was dark.'

c. * À Jean, pensait Marie surtout quand elle portait sa robe bleue.
To John, Mary thought especially when she was.wearing her dress blue
'John, Mary often thought about when she was wearing her dress blue.'

d. * À Jean, plaisait Marie surtout quand elle portait sa robe
To John, Mary was appealing especially when she was.wearing her
bleue.
dress blue
'To John, Mary was often appealing when she was wearing her dress blue.'

(144) 'As for' Topics do not allow subject-inversion

a. * À propos de Jean, est arrivé son père.

'*As for John, is arrived his father.'

b. */? En ce qui conserne Josef, c'est terminé son mandat.

'*For what concerns Joseph, it's ended his charge.'(Adapted from Marandin,2001)

The comparisons between examples (141) and (142) allowing subject inversion and, the un-grammaticality of subject inversion in (143) and (144) shows how it is a linguistic phenomenon specifically linked to scene-setting topics rather than to the broad notion of topic. Namely, subject inversion is allowed in presence of Scene-setting topics, while its not in the case of Left-dislocated Topics (143) or Aboutness topics introduced by 'as for' (144). As confirmed by the opposition between (145) and (146) below featuring some French cases of the so-called '*locative inversion*':

(145) NO locative inversion when the topic is an argument

a. À Paris, j'y vais .
To Paris, I.there$_{clitic}$ go
Lit.: 'To Paris, I go there'

b. * À Paris, va Jean.
To Paris, go John
Lit.: 'To Paris, goes John'

c. ?? À Paris, je vais .
To Paris, I go
'To Paris, I go'

d. Paris, j'y vais/*je vais.
Paris, I.there$_{clitic}$ go
Lit.: 'Paris, I go there'

e. ? Paris, Jean y va, Marseille, y va Camille.
Paris, John there$_{clitic}$ goes Marseille there$_{clitic}$ goes Camilla
Lit.: 'Paris, John goes, Marseille goes Camilla.'

(146) subject-inversion in Scene-Setting Topic with resumptive clitics

a. Sur le pont d'Avignon, on y danse.
on the bridge of.Avignon, 3SG. there$_{clitic}$ dance
Lit.: 'On Avignon's bridge, people there dance'

b. Sur le pont d'Avignon, dansent les jeunes mariées.
on the bridge of.Avignon, dance the young brides
Lit.: 'On Avignon's bridge, dance the young brides.'

c. Sur le pont d'Avignon, on danse.
on the bridge of.Avignon, 3SG. dance
Lit.: 'On Avignon's bridge, people dance'

d. Le pont d'Avignon, on y danse.
the bridge of.Avignon, 3SG. there$_{clitic}$ dance
Lit.: 'Avignon's bridge, people there dance.'

e. * Le pont d'Avignon, dansent les jeunes mariées.
the bridge of.Avignon, dance the young brides
Lit.: 'Avignon's bridge, dance the young brides.'

f. Le pont d'Avignon, y dansent les jeunes mariées.
the bridge of.Avignon, there$_{clitic}$ dance the young brides
Lit.: 'On Avignon's bridge, there dance the young brides.'

(145) and (146) further attest that non-movement derived Topic fame-setter are the only allowing subject-inversion.

Another crucial aspect revealing the unacceptability of subject inversion in syntactic configurations where movement has occurred is shown in (147). Even-Focus presupposes the presence of a reference subset, in this case all the hostels in which Napoleon has slept, and the element to which 'meme' even is attached in implicitly in contrast with this reference sub-set.

(147) Context: [Un guide accompagne des touristes en Belgique, il est fasciné par Napoléon et leur montre tous les hôtels dans lesquels Napoléon a dormi. Le guide montre une auberge et dit:]

a. Dans cette auberge dormait Napoléon quand il était de passage à Waterloo.
'In this hostel slept Napoleon when he was passing by in Waterloo.'

b. ?* Même dans cette auberge dormait Napoléon quand il était de passage à Waterloo.
'Even in this hostel slept Napoleon when he was passing by in Waterloo.'

Hence, not only Left-dislocated Topics block subject inversion as in (143), but subject inversion is not allowed after constituents having a contrastive or Focal interpretation, which in turns confirms the impossibility to have subject inversion after moved constituents.

In this regard, we could draw a parallel here with the work of Belletti (2001, 2002 and much related work) on the issue of Italian Subject inversion in case unaccusatives verbs. The author demonstrated that unaccusative subjects move in a low focus position, the so-called Low periphery, instead of moving up in its 'classical' IP position. In other words, in cases like *Arriva Piero* literally '*arrives Piero*', the subject does not move to its position in the IP-layer, staying low in the vP, therefore getting a focal interpretation. Following this proposal done for Italian, the un-grammaticality of (147b), could actually be accounted for: two focus cannot be found in the same sentence (Badan, p.c.). Under this analysis both post-verbal subject 'Napoléon' and the constituent introduced by *even* 'même' are focuses.

Topic + Unaccusatives = subject-inversion

In fact, an aspect we left on the side so far in this overview of French Scene-setting topical phenomena is the prevalence in the above sentences of Unaccusative verbs, as observed in Neapolitan Double Subject construction. This known fact in French grammatical tradition (Le Bidois, 1950) features a 'postverbal realization of the external argument of the verb' in indicative independent clauses, where the post-verbal NPs have object-like properties:

(148) Locative inversion

a. Spatial Scene-setting: Dans une forêt lointaine *vivait* un vieil ermite.
'In the remote forest was linving an old hermite.'

b. Prepositional Phrase: À chaque élément *correspond* un dossier différent.
'To each element correspond a different file.'

c. Prepositional Phrase: Avec le patriotisme instinctif, decline l'amour de la terre.
Lit.:'With instinctual patriotism, declines the love of earth.' Jonare (1976:72)

d. Conditional Phrase: Même s'il s'agit d'une rumeur, deumeure le doute de la probité de Fillon.
'Even if it's all rumors, remains the doubt about Fillon's probity.'

Similar examples of presentative and Locative Inversion are reported for French and other romance languages (e.g. Portuguese, Italian, Spanish and Catalan) in Marandin (2001, 1997, 2005), who defines the Subject inversion in these sentences as an instance of "Unaccusative construction"[106].

(149) Locative inversion in French unaccusatives

106. As for the association between arguments and their function, the unaccusative construction is non-canonical, in that the 'first argument' (i.e. external argument) is realized as an object which is an internal argument.

a. * Dans la salle, commencèrent immédiatement à travailler les candidats.
'In the room, started immediately to work the candidates.'

b. Dans la salle, se turent immédiatement les candidats.
'In the room, became silent immediately the candidates.'

Marandin (2001) namely observes a the correlation of two features "object properties of the NP" and "non-agentive semantics of the verb" with subject inversion (not only locative one). As shown in examples (149a) and (149b) agentive verbs do not yield inversion (Marandin, 2001).

Among the generally acknowledge syntactic properties of locative inversion we can find, in fact, that in English it is not permitted if the verb takes a direct object and that it is generally restricted to unaccusative verbs (e.g. Coopmans, 1989)[107]. In his analysis Marandin (2005)[108], the link between information structure and unaccusative inversion is straight, and he distinguishes between the argumental status and the grammatical function in a direct and simple way: "the status of first argument accounts for the subject-like properties and the functional status accounts for the object properties, most notably the postverbal placement." (Marandin 2001:196). Thus, in the "unaccusative realization" of the verbs in locative inversion constructions, the verb is subjectless and its external argument is realized as a complement. The author mainly argues that this kind of locative inversion are thetic utterances and that they are pervasively authorized in presence of intransitive verbs, and particularly unaccusatives, in that they yield an unaccusative realization of the verb.

Yet, pursuing the above parallel between our French examples of Subject inversion and Belletti's analysis of subject occupying the Low Focus position in Italian unaccusatives (2001/2002), we can tentatively argue the following: the fact that the above French Topics yield the unaccusative subject to be stay in post-verbal potions (i.e. in the Low Periphery) seems indeed to be confirmed by a possible focal interpretation of the above post-verbal subjects, which are all new information in (148)[109] and possibly contrastive like in (149b). However, to confirm that the analysis of Belletti holds for these French syntactic configurations one would need to further test for the possible contrastive interpretation of these post-verbal subjects to clearly set such an analysis, we leave this point as a open issue for future research[110].

In conclusion, we find in French base-generated Scene-setting Topics once again - as it was the case in Neapolitan Double subject constructions- a correlation between base-generated topics[111] and unaccusative post-verbal subjects, we will now continue to explore this direction in Mandarin Chinese to add some cross-linguistic evidence for the correlation between the presence of a sentence-topic and the Comment's linearization, specifically on the fact the unaccusative argument is kept in-situ. In one sentence,

107. Typical English examples like "Down the hill the baby carriage rolled." are given with Unergative verbs, like run, race, fly, dive, and, unaccusative ones, like 'drop', 'slide', 'fall' and 'bounce'.
108. Note that his analysis is cast in the framework of HPSG, and it it showed that inversion follows from the non-canonical mapping of the argumental status of first argument onto the syntactic function of object.
109. Except may be (148d) which can anyway be understood as contrastive.
110. Note that Badan (p.c.) adds to this initial sketch of analysis, the remark that contrastive Focus is not activated in the French Left-Periphery. Importantly, this fact would additionally push in the direction of analyzing these post-verbal subjects as being moved to the Low Periphery, which would indeed give them a focal (and possibly contrastive) interpretation.
111. Be they scene-setting ones or hanging topics as in Double nominatives.

Scene-setting Topics in French impact the surface linearization of subjects in post-verbal position.

The above linguistic arguments show that under certain conditions in French the unaccusative sole argument can be realized in post-verbal position. Hence, this overview also introduces us to one of the main focuses of our neuro-linguistic investigation of NP-movement in unaccusatives to be addressed in chapter 6. In our fMRI study on French syntactic transformations, we look for the neural underpinnings of the representational complexity yielded by the syntactic transformation moving the sole argument of Unaccusative verbs in preverbal position (i.e. NP-movement, cf. section §2.4.2.1) and not in-situ as in the above locative inversion constructions.

All in all, this overview on how French Frame-Setting topics impact on the linearization of grammatical subject in the Comment-clause, revealed three main aspects at stake in 'Subject inversion' the linguistic phenomena:

1. the fact that Frame-setting Topics allow to account for a wide range of free Subject inversion phenomena in French.
2. that Base-Generated and non-movement derived Topics allow Subject inversion, and
3. that in case of unaccusative verbs (and unaccusative constructions) subject inversion is observed

3.3.3 Subject-inversion and Scene-setting Topic in Chinese

3.3.3.1 A particular type of Base-Generated *in-situ* Topic

Drawing on the parallel with Neapolitan Double subject structures and French Scene-setting Topics yielding both subject-inversion in the Comment-clause, we want to sketch a possible analysis of a debated syntactic phenomenon in Mandarin featuring an unusual post-verbal subject:

(150) Topics and unaccusative verbs in the Comment

a. 王冕死了父亲。

Wáng Miǎn sǐ-le fùqin.
Wang Mian die-ASP. father

Lit.: 'Wang Mian died a father.' or 'Wang Mian's father has died.'

b. 他死了个儿子。

Tā sǐ-le ge érzi.
he died-ASP. one-CL. son

Lit.: 'He died a son' or 'One of his sons died.' La Polla (1995)

c. 昨天死了一条狗。

Zuótiān sǐ-le yītiáo gǒu.
yesterday died-ASP. one-CL. dog

Lit.: 'Yesterday died a dog.'

d. 动物园跑了一只熊猫。

Dòngwùyuán pǎo-le yīzhī xióngmāo.
yesterday escaped-ASP. one-CL. panda

Lit.: 'At the zoo a panda escaped.' La Polla (1995)

The sentences in (150) together with pseudo-passive construction (cf. section §3.2.3.1), existential constructions and non-causative ones, participate to a general trend that is usually expressed in a generalization stating that Chinese displays 'surface unaccusativity'. In these constructions the verbs lacks an external argument, and the sole theme (patient) argument not only shows object like properties (e.g. indefiniteness), but is also linearized like the object of a transitive verb in post-verbal position, but given the the unaccusative nature of the verbs in (150) the post-verbal argument is not assigned the role of the object. These linguistic phenomena have been successively described with different terms. Among the most common we find "transitive use of unaccusative verbs" or "surface unaccusativity" and some authors suggest that a better translation for (150a) should be "he was affected by the death of [his] father" to convey the non-topicality of the post-verbal subject (cf. Kuno, 1978:206).

The derivation of these non-canonical post-verbal subjects in Chinese such as in (150) has been a puzzle in Chinese linguistics since its early description in Li and Thompson (1981), which call these structures as existential presentative structures where the locus can either be spatial (150c and d) or a person a sentence-initial possessor NP (150a and b).

In the 90's, La Polla brings back on the stage the distinction between categorical and thetic propositions and terms these linguistic structures as 'event-central' sentences[112]. He advances that an incorporation of the subject noun into the intransitive verb helps converting the categorical judgment of a Topic-Comment constructions into a thetic statement (see also Sasse, 1984:260)[113].

Debates on the parallel between possessor NPs and Scene-setting Topic have been recently revived (see Pan and Han, 2005; Pan and Hu, 2000/2001; Shen Jiaxuan, 2006; Yang Berry, 2016; Huang and Her, 1998), and several alternative explanations have been developed, mainly analyzing the surfacing of the subject in sentence-final position as the mark of a presentational construction or as mirroring the informational status of the post-verbal subject as indefinite.

Among the different analyses that have been adduced for these sentence constructions we can cite, the possessor raising hypothesis (Xu, 1999; Han, 2000), the Middle applicative analysis by Tsai and Yang (2008), the light verb analysis (Zhu, 2005; Lin, 2008), the base-generated subject by Huang (2007), and finally the one we subscribe to and will assume by Pan and Han (2005). These authors analyzed the first NP as a Base-Generated *in-situ* Topic, as represented in Figure 3.17:

As illustrated in (3.17) by the syntactic-tree diagram reproducing the analysis of Pan and Han (2005), the post-verbal argument is incapable of receiving accusative case because of the unaccusative nature of the verb in the Comment. Thus, it inherits nominative case from the vacant subject position. The authors argue in their analysis that Chinese unaccusative sentence allow a possessor Topic NP while keeping the NP argument *in-situ*.

112. La Polla (1995) adduces evidence for the event-centrality of the type of constructions observing the kind aspectual marking the verbs allow (i.e. perfective marker *-le* or experiential aspectual marker *-guo*, but never iteratively achieved actions); see page 319: ex. (27-28).

113. The notion of theticity allows to account for the presentative effect of these sentences. Sasse (1987) opposes two discourse values in thetic utterances, namely the fact that they convey an eventuality, the so-called 'eventuality-centered' utterances or one of the participants to that eventuality, the so-called 'entity-centered' ones.

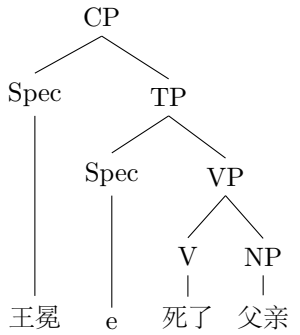


Figure 3.17 – As illustrated by the syntactic-tree reproducing the analysis of Pan and Han (2005), the post-verbal argument is incapable of receiving accusative case because of the unaccusative nature of the verb in the Comment. It inherits nominative case from the vacant subject position. The authors argue that Chinese unaccusative sentence allow a possessor Topic NP while keeping the NP argument in-situ.

Furthermore, thanks to the above observation by Li and Thompson – equating spatial locus and possessor locus in Topics – we can draw an initial parallel between the above Chinese Topic-Comment sentences and the frame-setting topics yielding subject in version in French[114]. Similarly to what was argued for French (cf. Lahousse, 2003) the Chinese examples converge in showing that the notion of Frame-setting Topic allows to have a unified approach of the two types of phrases yielding subject inversion: the first featuring a Possessor NP in (150a and b) and the second featuring typical Scene-setting topic spatial and temporal information in (150c and d). For this reason, the reformation of (150a) in (151) is not only grammatical but shows that subject inversion is kept:

(151) Topics and unaccusative verbs in the Comment
Spatial adverbial: 王冕家死了父亲。

Wáng Miǎn jiā sǐ-le fùqin.
Wang Mian family die-ASP. father

'In Wang Mian4s family the father has died.'

3.3.3.2 Chinese Scene-setting Topics and Spatial Preposition Drop

Before moving forward to draw a parallel between the patterns we observed in Neapolitan and French, we should note a fundamental difference between Chinese Scene-Setting Topics and the ones observed in French so far.

(152) scene-setting Topics and Bare Nominals

a. Spatial Scene-setting Topic: 床边一张小几，放着个小坐灯。

chuángbiān yīzhāng xiǎo jī, fàng-zhe ge xiǎo zuòdēng.
bed-side one-CL. little table, put-DUR CL. little table-lamp

'A little lamp was on the little table beside the bed.' Li W.D Corpus (2005)

b. Temporal scene-setting Topic: 六点，他报告已经写好了

114. Remember French examples also involved non-spatial/temporal adverbials like modal adverbs or even conditionals as Frame-setting Topics.

Liùdiǎn tā bàogào yǐjīng xiě-hǎo le.
6-o'clock he report already wrote-RES(finish) LE-PRT.

'At 6 o'clock, he had already finished writing the report' from Li (2007)

The locative sentence in (152a) presents a Place-word in Topic sentence initial-position that is not introduced by any spatial preposition, similarly sentence (b) is framed by a temporal expression that does not present a preposition. comparatively French examples can be found in which no spatial preposition introduces the scene-setting Topic (see [153]) but they are considerably less frequent.

(153) Spatial framing with no subject-inversion

a. Le métro, avec la carte orange tu vas partout.
'The underground, with the orange card you can go everywhere.' (Prévost, 2003)

b. Le quartier de l'église, ils ont refait les trottoirs.
'The church neighborhood, they have done the pavement.'

The drop of the preposition in Topic position in French and Chinese examples however yields a shift in the referential denotation of the Topic NP: from a real spatial location to entity denotation[115]. The variable interpretation of Nominals at Topic position has been comprehensively addressed in Mandarin Chinese for Quantificational Phrases in Topic position by Audrey Li (1997 and much related work). Interestingly, Li (1997) formulated similar considerations on the variable interpretation of Nominals at Topic position. If one considers quantified expressions found in the Chinese topic field: they can not be understood as referential expressions singling out a quantity in an individual denoting one, but have a more abstract quantity reading only as illustrated by the following.

Drawing a parallel between the referentiality shift occurring for quantified expressions at Topic position for bare space expressions (i.e. space expressions having no locative prepositions or post-positions), it seems that there is a similar tendency to shift from a more referential space denoting [+ ref] reading to a more abstract notional-domain denoting [- ref] reading[116], that we will just briefly and intuitively address it in the

115. See French ex. (145): *Paris, j'y vais.* 'Paris I [there$_{clitic}$] go.'

116. Starting from a very interesting example given by Audrey Li in her thesis (1997), we can notice that being at topic position for a constituent like the quantified expression in (1a) and (1b) is not without consequences: this position changes the referential denotation of the hosted quantified expression from individual, more referential in (1b) and (1c) (hence forth [+ref]) to quantity denoting [- ref] interpretation (see (1a)).

(1) a. 三个学生，我想是不够的。Quantity denoting [- ref]

Sān-ge xuésheng wǒ xiǎng shì bùgòu de.
3-CL. student 1SG. think be NEG. enough DE

'Three students, I think are not enough.' (Li, 1997)

b. * 三个学生，我以为 e 吃完蛋糕。*Individual denoting [+ ref]

Sān-ge xuésheng, wǒ yǐwéi chī-wán dàngāo.
3-CL. student 1SG. believe eat-RES$_{finish}$ cake

'Three students, I thought ate up the cake.' (Li, 1997)

c. * 三个人，他们吃不完你给他们的饭。*Individual denoting [+ ref]

Sān-ge rén tāmen chī-bù-wán nǐ gěi tāmen de fàn.
3-CL. person 3PL. eat-not-finish 2SG. to 3PL. DE food

'Three persons, they can't finish eating what you gave them to eat.'

following examples[117].

Namely, the possibility of having Frame-setting Topic *bare place words* (i.e. with no spatial preposition) whose referentiality can shift from spatial- to entity-denoting in Chinese is what can be advocated to yield the grammaticality pattern observed in the pairs of sentences in (154).

---

d. 有三个学生，我以为 e 吃完蛋糕。Individual denoting

Yǒu sān-ge xuésheng, wǒ yǐwéi chī-wán dàngāo.
YOU$_{(There-is)}$ 3-CL. Student 1SG. thought eat-RES.$_{finish}$ cake

'There are three students, who I thought ate up the cake'

e. 三个人，我知道都吃不完你给他们的饭。Individual denoting

Sān-ge rén wǒ zhīdao dōu chī-bù-wán nǐ gěi tāmen de fàn.
3-CL. person 3PL. I know eat-not-finish 2SG. to 3PL. DE food

'There are three persons who can't finish eating what you gave them to eat.'

f. 有三个人 i 吃不完你给他们 i 的饭。Individual denoting

Yǒu sān-ge rén$_i$ chī-bù-wán nǐ gěi tāmen$_i$ de fàn.
YOU$_{(There-is)}$ 3-CL. person eat-not-finish 2SG. to 3PL. DE food

'There are three persons who can't finish eating what you gave them to eat.'

As we can observe in (1a) the Quantificational Phrase [Num + Cl + NP] at Topic position looses its referential individual interpretation gaining a quantificational denotation. It displays what is here defined as referentiality shift: in (1a) it's no more a matter of the three persons in the sentence, but it's their being three that is taken into account. When the predicate within the comment is forcing an individual reading of the quantified expression in Topic position like in (1b), the resulting topic-comment sentence is ungrammatical. So that individual denotation is an infelicitous reading for quantified expressions at Topic position. This is true even in presence of a co-referential pronoun in the comment-clause as shown in (1c). The ungrammaticality of the binding of a pronoun with the QP at Topic position reveals that it's impossible to influence the quantity denoting of the bare Quantificational Phrase (QP) at Topic position. The ungrammaticality of (1c) could be mainly due to the impossibility for an indefinite quantified expression to have a non-referential quantity denotation. In (1c), *sān-ge rén* 三个人 'three persons' at Topic position cannot bear a referential index and cannot be the antecedent of a pronoun in the comment. Interestingly, in presence of the existential verb you 'exist, have' preceding the sequence [Q Cl+ N], the Quantificational expressions in (1d) and (1f) recover their individual denoting reading (see also Li, 1998). In fact, (1d) must be interpreted in terms of the existence of three individuals (students) who finished the cake, and (1f) shows that the quantified expression introduced by you is co-referential with the pronoun in the comment clause.

Crucially, the quantity denoting reading of Quantificational Phrases appears to be extremely resistant to all sorts of manipulations in the comment. This means that any individual denotation of the sequence [Num Cl N] is excluded from Topic position. The only possibility to recover a referential individual reading seems to be: (1) The presence of the verb you, this introduces an existential or presentational construction before the QP (as in [1d] and [1f]); or (2) the presence of the universal quantifier *dōu* 都 in the comment clause, that ranges over the entire set of individuals (as shown in [1e]). Therefore, (1e) must be interpreted in the sense that each of the three individuals couldn't finish the food that was given. For lack of space we cannot thoroughly investigate if the same interpretation can be driven for Scene-Setting Topics. Pursuing in this direction could help understanding if when spatial information is hosted in Topic position it acquires the role Scene-setting role of localizing the content of the Comment-clause, or if it becomes like an Aboutness Topic having a pure referential interpretation.

117. cf. Fabre (2014) Conference on Space and Quantification, Barcelona. A more comprehensive analysis of the distribution and interpretational variability of bare space-nouns and stative locative Phrases at Topic position in Mandarin Chinese should be undertaken. Further research is needed to possibly parallel this with the interpretational variability of bare Quantificational Phrases in the same position as described by Li (1997), confer to previous footnote. This could tentatively show that the interpretational referentiality shifts of Bare/Definite spatial NPs and Quantificational Phrase that is observed in Topic-Comment constructions can be actually correlated to some of the properties of the Topic-Comment syntactic articulation.

(154) Topic Spatial Frame and Entity denoting

a. locative nominal + come or go

i. locative nominal + come: 前面来了一个人

Qiánmian lái-le yī-ge rén
ahead come-ASP. one-CL. person
'Someone came from ahead.'

ii. * locative nominal + go: * 前面去了一个人

Qiánmian qù-le yī-ge rén
ahead go-ASP. one-CL. person
'*' Xu Dan (2012)

b. locative place word + come or go

i. locative place word + come: 家里来了一个人

Jiā-lǐ lái-le yī-ge rén
home-inside come-ASP. one-CL. person
'Someone came into the house.'

ii. locative place word + go: 家里去了一个人 *Topic is the source*

Jiā-lǐ qù-le yī-ge rén
home-inside go-ASP. one-CL. person
'Someone from the family went (as a representative of the family).' Xu Dan (2012)

While sentences locative words in (a) cannot have a referentiality shift, Bare space words in (b) are grammatical in that the different verbs in the comment activate a different Denotation of the topic NP: in (i) a spatial one and in (ii) a entity-denoting one.

Hence, (bi) and (bii) can be analyzed as both representing the *locus* frame of the comment, and the verb in the Comment interacts with the semantics of the topic to change its interpretation from a spatial frame to an entity-denoting Topic as illustrated by the translations.

Specifically, in (154) the topic changes role from (a) and (b), according to the verb in the comment. In (ai) it features the source of the verb *lai* 'come', while in (aii) the verb 'go' forces the the interpretation of the locative word 'ahead' in topic position as a goal which is semantic infelicitous[118]. Interestingly, replacing the locative noun by a Bare Place word yields the grammaticality of (bii) because the Bare Place word is interpretable as a source, yielding the interpretation of the Topic DP as "some from the family".

3.3.3.3 Topic-Comment, Argument structure and Subject-inversion

After having considered the effects of Topics on the predicate thematic structure in French and Neapolitan, this section is dedicated to the analysis of *native speakers' preferences* for certain surface word-orders of the Comment's arguments and to what these can tentatively reveal of the structural effects topic-comment constructions can have on the comment predicate and its arguments, in unaccusative, unergative and Transitive Comment-clauses.

118. Xu Dan (2012) notes that in (a) the action is accomplished in the speakers' view, while this is not the case in (b). In this resides the semantic infelicitous interpretation of (aii).

Unaccusatives

Considering unaccusative verbs in (155), our seven informants report such examples to be interpretatively identical, but when asked to establish a preference between pre-verbal and post-verbal Unaccusative subjects they preferred post-verbal subjects' sentences to those with pre-verbal subjects.

(155) a. 咱们工厂, 病了一个工人。Post-verbal subject preferred to (b) pre-verbal one

Zánmen gōngchǎng, bìng-le yī-ge gōngrén.
our firm ill-ASP. one.CL. worker

Lit.: 'In our firm caught the illness a worker.' or 'In our firm a worker caught the illness.'

b. ? 咱们工厂, 一个工人病了。[- pref.]

Zánmen gōngchǎng, yī-ge gōngrén bìng-le.
our firm one.CL. worker ill-ASP.

Lit.:'In our firm a worker caught the illness.'

b'. 咱们工厂有一个工人病了。有 Rescues (b).

Zánmen gōngchǎng $\text{YOU}_{there.is}$ yī-ge gōngrén bìng-le;
our firm YOU one.CL. worker ill-ASP.

'In our firm there is a worker that caught the illness.'

The patterns reported in example (155) lead us to an initial observation: when a spatial expressions is occupying Topic position, the subject of intransitives (i.e. the only argument of these verbs) can occupy both pre-verbal (155b) and post-verbal position (155a). Interestingly, in (b'), adding *you* to explicitly transform the sentence into a presentational structure, actually rescues the sentence (b).

As revealed in the comparisons presented in (156), pre-verbal subject NP may stand adjacent to Topic bare NP as in example (155) or can be separated by intervening materials (156c), including circumstantial (156a-b), like temporal adverbs.

(156) a. 咱们工厂昨天病了一个工人。[+ pref.]

Zánmen gōngchǎng zuótiān bìng-le yī-ge gōngrén.
our firm yesterday ill-ASP. one.CL. worker

Lit.: 'In our firm caught the illness a worker yesterday.' or 'In our firm a worker caught the illness yesterday.'

b. 咱们工厂昨天一个工人病了。

Zánmen gōngchǎng zuótiān yī-ge gōngrén bìng-le.
our firm yesterday one.CL. worker ill-ASP.

Lit.: 'In our firm a worker caught the illness yesterday.'

c. ?? 咱们工厂听说一个工人病了。

Zánmen gōngchǎng tīngshuō yī-ge gōngrén bìng-le.
our firm hear-saying one.CL. worker ill-ASP.

Lit.: 'In our firm I heard saying that a worker caught the illness yesterday.'

In this case speakers preference is still for post-verbal subject (156a) compared to (156b), and as shown above, the presence of intervening material does not change the preference for post-verbal subject position (see 156c).

Testing a classically unaccusative verb Particularly revealing in this respect, is the behavior of a classical unaccusative predicate like 'die'. In contrast to unergatives and transitives, these verbs thematically license an object as their sole argument. This distinction is generally non-overtly marked in the syntax because the surface structure is that of a pre-verbal subject because of NP-movement, like in *Lao Wang si-le* 'Laowang died', and not *si-le Lao Wang* 'died Laowang'. However, when the verb 'die' is in the Comment-clause of a Topic-comment construction like (157), its sole argument can surface in post-verbal object position, and most importantly a clear preference for (a) post verbal position over (b), the classical pre-verbal subject position, is reported by our native informants.

(157) a. 王冕家昨天死了一个人。Post-verbal: preferred

Wáng Miǎn jiā zuótiān sǐ-le yīgerén
Wang Mian house yesterday die-PERF. one.CL.person

'A person died yesterday at Wang Mian's place.' modified from Shen Jiaxuan (2006:295)

b. ??* 王冕家昨天一个人死了。Pre-verbal: native speakers' ratings very very low

Wáng Miǎn jiā zuótiān yīgerén sǐ-le
Wang Mian house yesterday one.CL.person die-PERF.

'A person died yesterday at Wang Mian's place.'

b'. 王冕家昨天有一个人病/死了。reparing b.

Wáng Miǎn jiā zuótiān yǒu yīgerén bìng/sǐ-le
Wang Mian house yesterday YOU$_{there.is}$ one.CL.person die-PERF.

'At Wang Mian's place there is a person that died yesterday.'

Comparing sentences (157a) and (b) the only grammatical option according to speakers ratings is (a), even if the possibility of inversion of the verb and its sole argument is theoretically predicted. Notably, native speaker's ratings are very low for sentence (157b). And again adding *you* to explicitly transform the sentence into a presentational structure rescues sentence (b)[119].

Thus, we may now ask how to account for this acceptability asymmetry. Why is Subject-Verb canonical order in (157b) less acceptable than the inverted one (Verb-Subject) when a Topic is found in sentence-initial position?

Relying on the comparison with French and Neapolitan data presented so far, we can tentatively understand this clear-cut preference for post-verbal subject linearization as resulting from the interplay of unaccusative predicates and Topic-comment construction. In fact, when the word-order remains the same as it is outside a Topic-comment construction, the acceptability of the sentence appears to be reduced (157b). We could therefore argue that this configuration may show that some of the properties of Topic-comment constructions interact at some stage with the predicate argument structure of the Comment's Verb Phrase. The next coming sections

A way to account for this phenomena could be to refer – as we did for French Subject Inversion phenomena – to Belletti's analysis (2001/2002) of Subject inversion in Italian

119. We should say that the fact the subject *yi-ge ren* is indefinite may interfere in this test. We will see in the following examples other configurations.

unaccusatives, like *Arriva Piero*"arrives Piero" (V-S). Drawing this parallel would again imply to analyze the post-verbal subject linearization in the above Chinese unaccusatives as the mark that the subject moves to a Low-Peripheral Focus position where it can get a focal interpretation, or alternatively remains low in it internal argument position (i.e. in little v).

This appears to be confirmed by the above examples (157), where the subject *yīgerén* 'one person' is definitely both new and contrastive information. Moreover, Pan and Han (2005) noted that for the examples in (157), raising the subject NP into the IP vacant subject position would actually produce a patient-subject sentence, an option that is semantically dis-preferred in their analysis.

Topicalization and relativization in Mandarin This point can be further illustrated by the fact that the preferred word-order is also the only one which is relativizable (cf. Huang, Li and Li 2009). This not only shows that some gapless topic constructions, namely Scene-Setting Topics, can be relativized, but also probably suggesting the reason why the post-verbal subject word-order is preferred and maybe also the a direct link between topic-comment constructions and word-order constraints for the comment's arguments.

Relativization test is proposed by Huang, Li and Li (2009). The well-formedness of the relativized counterpart of Topic sentences is taken by Huang Li and Li (2009) from Kuno (1976) as a test for derivability of relative clauses from Topic ones.

This link between Topic and relative constructions has been advanced by early work of Kuno (1976) who proposed the-so-called '*Thematic Constraint*'. Not only these two constructions are both subsumed under wh-constructions, but the *Thematic-Constraint* states that a relative clause is well-formed only if there is a corresponding grammatical Topic structure, where (i) the head that is modified by the relative clause can be the Topic, and (ii) the relative clause can be the comment of its Topic-Comment counterpart. We will use this test in the following to investigate the acceptability of Subject inversion we vary the types of verbs in the Comment-clause.

The main idea underlying this test can also be found later in Tang (1979), who proposed that to form a relative sentence in Chinese an argument has first to be topicalized within the relative clause. A similar claim was also made by Jiang (1990) who observes that Topic structures are closely related to a relative ones, he argues indeed that when an element is relativized, it is always relativized from a Topic position[120].

While gapless Aboutness Topics, where no element in the Comment is co-indexed with the Topic constituent, generally turn to be ungrammatical when relativized[121] –

120. For a comprehensive discussion of the derivation of relative clauses from Topic structures refer to Huang, Li and Li (2009:212-235).

121. The fact that gapless Aboutness Topics generally show up to be ungrammatical when relativized can be show through the following examples by Li et al. (2009):

(1) NON-relativizable Aboutness Topics

i. * [[幸亏消防队来得快的] 那场火]
[[Xìngkuī xiāofángduì lái de kuài de] nèi chǎng huǒ]
fortunately fire-brigade come DE fast DE(rel) that-CL. fire
'*The fire such that fortunately the fire brigade came fast.'

ii. * [[我最喜欢香蕉的] 水果]
[[Wǒ zuì xǐhuan xiāngjiāo de] shuǐguǒ]
I most like banana DE fruit

proving in this way that the aboutness relation that holds between Topic and Comment is not sufficient to license a relative construction (Li et al. 2009) –, we clearly see in the following examples (158) that (1) Frame-setting Topic are relativizable, and that (2) dis-preferred word-orders in the Comment-clause yield un-grammaticality when relativized.

(158) Relativizable Scene-setting Topics

a. [[昨天死了一个人的] 那家]

Zuótiān sǐ-le yīgerén de nèijiā
Yesterday die-PERF. one.CL.person DE that-house

'The house where a person died yesterday.'

b. * [[昨天一个人死了的] 那家]

Zuótiān yīgerén sǐ-le de nèijiā
Yesterday one.CL.person die-PERF. DE that-house

*'The house where a person yesterday died.'

(159) Gapless Topic and Subjet linearization with unaccusatives

a. 王冕家，昨天死了一个人。Post-verbal subject + unaccusative verb

Wáng Miǎn jiā, zuótiān sǐ-le yīgerén
Wang Mian house yesterday die-PERF. one.CL.person

'A person died yesterday at Wang Mian's place.' modified from Shen Jiaxuan (2006:295)

b. ?* 王冕家昨天一个人死了。Pre-verbal subject > native speakers' ratings very very low

Wáng Miǎn jiā zuótiān yīgerén sǐ-le
Wang Mian house yesterday one.CL.person die-PERF.

'A person died yesterday at Wang Mian's place.'

Using the Thematic constraint test in (158), we can observe that in (158a) the head modified by the relative clauses can be used as a Topic and the relative clause as a Comment about the Topic for the sentence with the preferred word-order (157a) replicated for convenience here under in 159, but not for the sentence with the dis-preferred one as shown by the ungrammaticality of (159b).

This asymmetry has also been noted by Huang, Li and Li (2009:213). The authors showed that when Experiencer Topics are found with unaccusative verb like *fasheng* 'happen', their well-formed relativized counterpart appears to be only the one where the Comment word-order is having the sole unaccusative argument in post-verbal object position [122], like we see in our example (158a).

'The fruit such that I like bananas.'

122. Example (i) illustrates the unaccusative use with one argument (Theme), and (ii) the transitive use with an additional argument (Experiencer) occurring in a topic-comment construction, and (iii) and (iv) show that only the post-verbal subject word-order has a well-formed relativized counterpart:

(1) Experiencer Topics and unaccusative verb

i. 他们，意外发生了

Tāmen, yìwài fāshēng-le.
they accident happen-LE

The above evidence leads us to continue in this direction by using the Thematic constraint test to contrast the relativization behavior and subject linearization preference patterns we just observed in unaccusative predicates, that generally license post-verbal subjects as in example (157), with that of unergatives and transitive verbs.

Unergatives

The main difference between Unaccsuatives and Unergative verbs is that, while unaccusative verbs thematically license their sole argument as a internal argument, that has to move through NP-movement from this position in the underlying structure to be fond in pre-verbal subject surface position, Unergative verbs only have an external Agent argument. Given the fact that the argumental configuration between Unaccusative and Unregative verbs is one the mirror image of the other, it is interesting to observe that when comparing the two subject linearization we are investigating in the Comment-clause, the opposite pattern of grammaticality is shown between the unaccusative case in (157) and the uneregative case in (160).

(160) a. * 王冕家昨天晚上哭了一个人。Unergative verb: *post-verbal subject

Wáng Miǎn jiā zuótiān kū-le yīgerén
Wang Mian house yesterday cry-PERF. one.CL.person

'A person cried yesterday at Wang Mian's place.'

b. 王冕家昨天一个人哭了。Unergative verb: preverbal subject position

Wáng Miǎn jiā zuótiān yīgerén kū-le
Wang Mian house yesterday one.CL.person cry-PERF.

'A person cried yesterday at Wang Mian's place.'

While the linearization preference in unaccusatives (157a) was featuring the verb preceding it's sole argument in a non-canonical word-order, with an unergative verb in the Comment the argument distribution we observe in (160b) and (160a) is the opposite: the sole argument (which is thematically a subject) of the unergative verb 'to cry' 哭 *kū* is felicitous only if occurring in preverbal subject position. In sum, post-verbal unergative subject is ungrammatical, while pre-verbal subject is felicitous[123].

'(As for) them, an accident happened.'

ii. 他们$_i$, e$_i$ 发生了意外了

Tāmen$_i$, e$_i$ fāshēng-le yìwài le.
they happen-LE accident LE

'(As for) them, they had an accident.'

iii. * [[意外发生了的] 那些人]

[[Yìwài fāshēng-le DE] nèi-xiē rén]
accident happen-LE DE those person

'The people such that an accident happened.'

iv. [[e$_i$ 发生了意外的] 那些人]

[[e$_i$ Fāshēng-le yìwài de] nèi-xiē rén$_i$]
happen-le accident DE those person

'The people who had an accident.'

123. This holds true also with the verb 'laugh' in example (1b) compared to 1a).

Back to our demonstration on the the contrast of linearization of subject in unergatives and unaccusative, an additional confirmation for the fact these particular argument word-orders are yielded by the Topic-Comment articulation of this sentences is again to be found in the Thematic Constraint relativization test also proposed by Huang, Li and Li (2009): the preferred and grammatical word-order is the only grammatical when relativized (161).

(161) a. [[昨天哭了一个人的] 那家]

Zuótiān kū-le yīgerén de nèijiā
Yesterday cry-PERF. one.CL.person DE that-house

'The house where a person cried yesterday.'

b. * [[昨天一个人哭了的] 那家]

Zuótiān yīgerén kū-le de nèijiā
Yesterday one.CL.person cry-PERF. DE that-house

*'The house where cried a person yesterday.'

Interestingly, unaccusative-unergative distinction is not only shedding a light on the linear distribution of subjects of intransitive verbs, but also offering some evidence for what we could call some Topic construction-specific constraints on the surfacing of the sole argument of these intransitive verbs. The picture we can draw on the basis of the above tests is that in Topic-Comment articulations containing intransitive verbs

(1) a. * 王冕家，昨天晚上笑那个人。

Wáng Miǎn jiā, zuótiān wǎnshang xiào-le nàgerén
Wang Mian house yesterday evening laugh-PERF. that.CL.person

'A person laughed yesterday evening at Wang Mian's place.'

b. ?? 王冕家昨天晚上一个人笑了。

Wáng Miǎn jiā zuótiān wǎnshang yīgerén xiào-le
Wang Mian house yesterday evening one.CL.person laugh-PERF.

'A person laughed yesterday evening at Wang Mian's place.' or 'Mian's family made fun of someone yesterday evening'.

c. 王冕家昨天晚上有一个人笑了。

Wáng Miǎn jiā zuótiān wǎnshang yǒu yīgerén xiào-le
Wang Mian house yesterday evening YOU one.CL.person laugh-PERF.

'A person laughed yesterday evening at Wang Mian's place.'

However, it has to be noted that the sentence in (1b) is given another meaning in absence of the comma separating Topic and Comment[124], the verb 笑 xiào 'to laugh' is interpreted as having a transitive meaning, i.e. 'make fun of s.one', and 王冕家 *Wang Mian-jia* 'Mian's family' is analyzed as the agent-subject of the transitive verb 'make fun of s.one'.

In such a configuration where the verbs is altering a transitive meaning, we can note in (1c) that the only way to rescue the locative function of the sentence initial DP is the presence of the existential verb *yǒu* 'to have' or 'there is'. Following Huang (1987, 1990), the function of 有 *yǒu* would be that of asserting the existence of an event/state denoted by the VP complement and to select the clause 一个人笑了 *yīgerén xiào-le* 'a person laughed' (i.e. the Comment-clause) as its complement[125]. Commenting the effect of inserting 有 *yǒu*, Hu and Pan (2008a) proposed that the properties of this configuration having an intransitive verb preceding it's sole argument in non-canonical word-order (what they call argument inversion) could be analyzed in such a way that an abstract verb *yǒu* 有 'to have' would be postulated.

some syntactic properties should be interacting at some stage with predicate argument structure of the Comment's Verb-Phrase, yielding a preferential post-verbal surfacing of the sole argument of these two types of intransitives: the argument is to be found in the original underlying structure of the verb, respectively post-verbal for unaccusatives and pre-verbal for unergatives.

The reasons for this preference are still to be explained, the parallel with Belletti's proposal for Italian unaccusatives Subject inversion is a promising direction we leave here for future research.

In the meantime, we can just describe that what seems to be at stake here, is the existence of a construction-specific interaction with the argument structure of the verb. The linguistic analysis of this relation between Scene-setting Topics and the argumental structure of the verb in the Comment will reveal crucial for the interpretation of our ERP data in section §5.4, to allow us to pinpoint the structure-driven constraints yielded by the presence of a Topic in the sentence-unit. Namely, time-locking Event-Related Potentials at Comment-Subject position and at Comment-VP position will offer us a window to further describe at a brain processing-level these apparent interplay between Topic-Comment articulation and argumental structure of a verb, even in cases like in Scene-setting Topics, where the Topic does not undergo the verbal selectional restrictions.

Transitives

We will now move to consider what happens to the two arguments of transitive verbs when they are found in Frame-setting Topic-Comment articulation. As for frequency of use, it would be very interesting to be able to test for the correlation between argumental structure types and the frequency of use of the Topic-comment constructions in spontaneous language as undertaken by Ledgway (2010) for Neapolitan Double Subject constructions.

Although the tests for the linearization of the subject in we performed in Topic comment articulations containing intransitive verbs, are less probing in the case of Transitive ones, we nonetheless continue reproduce the tests we performed because they bring to light other aspects of Mandarin Chinese like its pro-drop property. For instance, comparing (162a and b) to (160c and d) we can observe that subject omission is preferred irrespective of the position of the object.

(162) a. 郊区政府已经盖了很多新房子。[Topic+Agent+V+Object]

Jiāoqū zhèngfǔ yǐjīng gài-le hěnduō xīn fángzi.
outskirt government already construct-LE many new apartments

'In the outskirts the government already built many new apartments.'

b. 郊区很多新房子已经盖了。pre-verbal object + null subject

Jiāoqū hěnduō xīn fángzǐ yǐjīng gài-le.
outskirt many new apartments already construct-PERF.

'In the outskirts [the government] already built many new apartments.'

c. 郊区已经盖了很多新房子。subject omission VO [+ pref.]

Jiāoqū yǐjīng gài-le hěnduō xīn fángzǐ.
outskirt already construct-PERF. many new apartments

'In the outskirts many new apartments have already been constructed.'

d. 郊区新房子已经盖了很多 subject omission OV [+ pref.]

Jiāoqū xīn fángzǐ$_i$ yǐjīng gài-le hěnduō ___$_i$
outskirt new apartments already construct-PERF. many

'In the outskirts many new apartments have already been constructed.'

Yet, if we consider the examples in (163), where the subject is being kept, native speaker's ratings show a clear preference for verb-final constructions in the Comment, be it in a Multiple Topic construction like in (b) or in a pre-posed object construction with a contrastic reading like in (c).

(163) a. 办公室，我们收不到网上买的货。TopSVO [- pref.]

Bàngōngshì, wǒmen shōu-bù-dào wǎng-shàng mǎi de huò.
office we receive-not-POT. internet.on bought DE goods

'In our office, we cannot receive incoming goods bought on internet.'

b. 办公室，网上买的货我们收不到 ___。Top+Top$_{obj.}$SV the preferred word-order [++ pref.]

Bàngōngshì, wǎng-shàng mǎi de huò wǒmen shōu-bù-dào ___.
office internet.on bought DE goods we receive-not-POT.

'In our office, incoming goods bought on internet we cannot receive [them].'

c. 办公室，我们网上买的货收不到 ___。TopSOV [+ pref.] contrastive object pre-posing

Bàngōngshì, wǒmen wǎng-shàng mǎi de huò shōu-bù-dào ___.
office we internet.on bought DE goods receive-not-POT.

'In our office, the incoming goods we bought on internet cannot be received.' (while other goods/letters can)

The above preference is verified also when the Topic NP can be a complement of the comment's Verb, as in (164). The topic NP *bàngōngshì* 办公室 is the complement (i.e. goal) of the verb *sòngdào* 送到, and again the preferred word-order appears to be the verb final one in (164a and b):

(164) a. 办公室，网上买的货快递送不到。Top+Top$_{obj.}$ the preferred word-order [++ pref.]

Bàngōngshì, wǎng-shàng mǎi de huò kuàidì sòng-bù-dào
office internet.on bought DE goods express.courrier deliver-not-POT.

'In our office, goods bought on internet the express courrier cannot deliver.'

b. 办公室，网上买的货送不到。TopOV [+ pref.]

Bàngōngshì, wǎng-shàng mǎi de huò sòng-bù-dào
office internet.on bought DE goods deliver-not-POT.

'In our office, goods bought on internet cannot be delivered.'

c. 办公室，快递送不到网上买的货。TopSVO

Bàngōngshì, kuàidì sòng-bù-dào wǎng-shàng mǎi de huò
office express.courrier deliver-not-POT. internet.on bought DE goods

'In our office, the express courrier cannot deliver goods bought on internet.'

d. 办公室，快递网上买的货送不到。TopSOV [+ contrastive Word-Oder]

Bàngōngshì, kuàidì wǎng-shàng mǎi de huò sòng-bù-dào
office express.courier internet.on bought DE goods deliver-not-POT.

'In our office, the express courier cannot deliver goods bought on internet (while other goods or letters he can).'

One could explain these phenomena by the well known fact that Chinese languages allow bare nouns to appear as arguments and that various interpretations of Bare-NPs in Mandarin have been claimed to be highly correlated to pre-verbal or post-verbal position[126] (Xu Liejiong, 1995; Sybesma, 1999) especially as far as definiteness feature is concerned[127]. In general pre-verbal position is yields a definite interpretation of pre-verbal objects while post-verbal is associated with indefinite. Hence, this preference can be explained by a whole array of arguments described by previous literature on indefinite non-specific subject opposed to definite Topics (see Qu Yangfeng, 1994; Shyu, 1995; Yang Barry, 2012).

126. Generally speaking Bare-NP following the verb are interpreted as indefinite, while when they precede it they get a definite interpretation, and topicalized elements are always characterized by a certain degree of definiteness.

127. An interesting point for our reflection on argumental structure and topicalization has been brought to light by Yuan Yulin (1996). There exist a restrained class of verbs that obligatory require a post-verbal object and resist to topicalization, as illustrated by one of her examples here under:

(1) Non-frontable objects

a. 新政策会繁荣金融市场

Xīn zhèngcè huì fánróng jīnróng shìchǎng
new policy will boom financial market

' The new policy will cause the financial market to boom.'

b. * 金融市场，新政策会繁荣

*Jīnróng shìchǎng, xīn zhèngcè huì fánróng
financial market new policy will boom

' *.'

c. 他的话温暖了我们的心

Tāde huà wēnnuǎn-le wǒmende xīn
his words warmed-ASP. our heart

'His words warmed up our hearts.'

d. * 我们的心，他的话温暖了

*Wǒmende xīn, tāde huà wēnnuǎn-le
our heart his words warmed-ASP

' *.'

We could speculate that this is due to their causative interpretation.

However, this difficulty explain why in presence of a Scene-setting Topic objects would preferentially be considered as definite and therefore not be found in post-verbal position but in topic position or in contrastive pre-verbal position and agents be systematically dropped. Interestingly, discussing the kind sentences where agent are systematically dropped Pan and Han (2005) note that Cheng (1989) resumes to an argumentation on unaccusativisization to account for Transitivity alternations in Chinese[128]. Hence, for the moment we can only speculate along the arguments given by Marandin (2001) about the correlation between Locative inversion with the "unaccusative realization" of the verbs in these locative constructions.

Given these additional observations on transitives Comment-clauses, the preference of native speakers for the subject linearization in Comment-clause across verb types seems to indicate some Topic-Comment construction-driven effects on the linearization of subjects and objects in the Comment surface word-order.

In sum, the Neapolitan, French and Chinese examples presented in this section have been showing a certain pattern in the linearization of the subject in the Comment-clause where intransitive verbs are found. Some properties Topic-comment construction can be tentatively pointed out to play a role in the ordering of arguments in the Comment.

To conclude, the different linguistic facts we brought to light in Chinese and our cross-linguistics analysis of Scene-setting Topics aimed at:

1. settling down clearly the Topical character of place words found in sentence-initial position in Mandarin Chinese. These scene-setting Topic constructions will be selected as experimental sentence for our neuro-linguistic investigation of the on-line processing of Topic-comment sentence in chapter
2. delineating their derivation as Base-Generated Topics, and lastly
3. we wanted to contribute some linguistic arguments showing how though the presence of a Topic in a sentence-unit, the 'activation' of the sentence-discourse interface can actually have an effect on the Comment-clause internal organization, and more specifically on the surfacing of subject in post-verbal position.

We leave the discussion of the pattern presented here and its relevance for the debates about unaccusatives Subject inversion and Locative inversion to future investigation. All in all, a detailed account of these open issues remains largely to be done and Belletti's proposal for Italian unaccusatives Subject inversion can be indicated as a promising direction we leave for future research.

Reviving our metaphor of the sentence-unit as a cupola as illustrated in Figure 3.18, we can just say that these linguistics facts that emerged from this section indicate how the sentence-discourse properties of Topics (here in the case Scene-setting topics) appear to have the structural and static functional role of a lantern that deepens its 'architectonic' influence into the sentence's structure, for example influencing the different linearization of subjects across different verb types.

128. An argument going in the same direction was given by Huang and Her (1998), who developed some arguments to account for these word-order preferences in transitive Comment-clauses inside the Lexical-Functional Grammar framework. The authors argue that locative inversion verbs require a locative argument but also constrain or forbid the presence of other roles in the argumental structure.

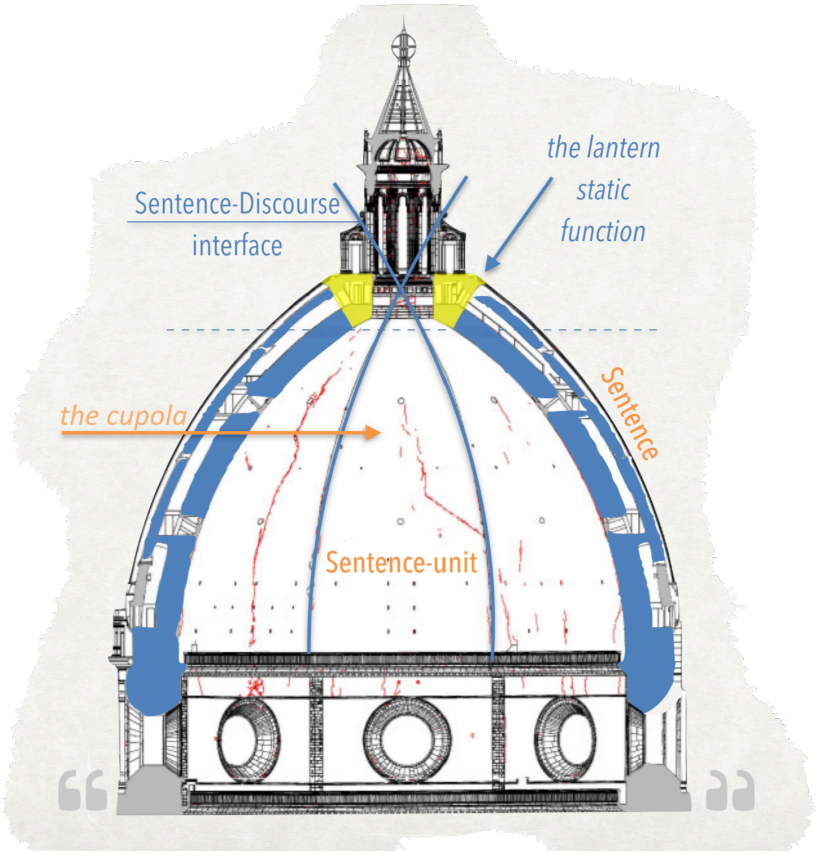


Figure 3.18 – Architectonic metaphorical representation of the sentence-discourse interface -*the lantern*- impacts on sentence-internal articulation -the cupola-. Lines and structural elements in blue represent the lines along which the static forces expand starting from the point of application of the lantern weight in yellow. Drawing of Santa Maria del Fiore (Florence) adapted from architectural drawing of Haines and Battista (2015).

3.4 Syntactic properties of Topical linguistic phenomena: Topic and Topicalization

> "Quot homines tot sententiae: suo' quoique mos."
>
> *["As many men, so many sentences: to every one his own way."]*
>
> TERENCE (161 BC) *Phormio*, L.45

After having delineated the characteristics of the Mandarin Chinese linguistic system that have been correlated with the possibility of having 'Chinese-style' Topics, we now move to the syntactic analysis of the Topic-Comment configurations that were brought to light in our typological overview (§3.2). For this we will adopt a more formal syntactic point of view on the notion of Topic, namely the one developed in the framework of Generative Grammar.

Following the basic lines of the Cartographic Project (cf. Cinque 1999, 2002; Rizzi, 2004; Belletti, 2004) and the functional mapping of the Mandarin sentence structure proposed by Badan (2007 and much subsequent work), we will address its cross-linguistic universality, by discovering how the Topic-Prominence claim translates in a theory of the fine-structure organization of CP-layer in Mandarin Chinese, the so-called Left-Periphery.

So far the discussion has concentrated on establishing an overall description of the different syntactic encoding of topicality in Chinese and across languages. We now turn to a more formal approach addressing the issue of the syntactic derivation of Topic in the different configurations we went through.

Syntactic formal approach to Topical phenomena

Previous chapter sketched the formal analysis of Topic-Comment structure issued by generative approach (§2.4.4). To briefly summarize the syntactic account argued for this sentence articulation in §2.4.4.3 (p. 206) we can first resume to in the tree-structure formal representation of Topic-Comment sentence's articulations (Cinque, 1990 and Rizzi, 1997/2002), where the Topic corresponds to a functional projection (TopP) in which the the specifier position is occupied by the topical element and the complement of the projection (IP) is the Comment – the rest of the sentence –, as represented here under:

To this formal representation expressing the fundamental bipartition of the sentence into Topic and Comment-clause, we can add the following general definition of the Topic proposed by a foundational article by Rizzi's (1997:285):

> "The topic is a *pre-posed* element characteristically set off from the rest of the clause by '*comma intonation*' and normally expressing old information, somehow available and salient in previous discourse"

A- Movement versus in-situ analysis for Topics In the above general definition we can note that Topical phenomena are characterized as dislocated constructions, which is allegedly accepted in Generative Grammar as well as in general linguistic theory.

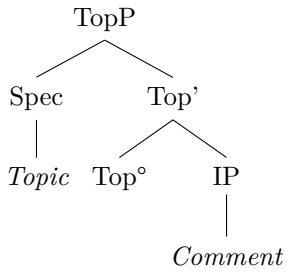


Figure 3.19 – Syntactic-tree structure of the Topic functional projection. In Xu (2000): Topic Phrase (TopP) is the maximal projection of the Top head (Top°). The Topic element occurs in its specifier position (Spec), followed by a functional category Top and the complement of Top is the IP, the Comment-clause.

While the Movement-base account for Topic-Comment articulations is often taken for granted in English, this point has been hotly debated in Chinese linguistics as well as in other languages. The literature on Chinese Topic structures is divided into two contrasting approaches to the issue of the syntactic derivation of Chinese topics. Namely, one arguing for a movement only analysis to sentence-initial position (Huang, 1982; Shi, 2000) and the other for base-generation stating that all Topics are generated in their surface position. (Xu and Langendoen, 1985; Xu, 2000). We will address this in section §3.4.1 by arguing that Topic-prominence parameter yields a different syntactic derivation for 'Chinese-style' Topics following Paul and Whitman (2015).

Section §3.4.2 will be namely dedicated to the issue of whether topical constituents move through wh-movement to their Left-periphery syntactic position or are actually found in-situ (i.e. Base-generated) in that syntactic position. This will be done the perspective of experimentally addressing the issue of the syntactic complexity linked to movement derivation versus base-generation from the cerebral point of view, in chapter 7 for Chinese and in chapter 6 for French.

B- Topic-Field, a locus for the Sentence-Discourse Interface The sentence-discourse interface property of Topics that was put forward by our itinerary in the notion of Topic (§3.1), constitute a further argument to adopt the syntactic analysis of the Cartographic approach presented in chapter 2.

As already pointed out this approach has the important advantage of postulating a syntactic encoding of the Scope-Discourse properties that are specific to different kinds of Topics and syntactic phenomena like questions for example. Namely, in the Cartographic project these sentence-discourse interface properties are directly represented in the structural syntactic-tree skeleton of the sentence-unit, in what is called the sentence's Left-Periphery (i.e. the CP layer).

Hence, when a Topic head (Top°) is occurring in the Left-Periphery of the sentence, the following sentence-discourse interface mechanism is activated and an interpretative instruction at the interface with semantics-pragmatics is given for building the sentence into a Topic-Comment articulation, yelding a kind of order that could be expressed like: "please *interpret my specifier as the Topic, and my complement as the Comment*" (Rizzi, Paris 2014). This forces particular elements with the same features to move into their specifiers (Spec), in order to share the same interpretive import. The syntactic theoretical

formalization of the Sentence-discourse interface domain allows to address the cerebral representational issue linked to syntactic syntactic complexity of the sentence-unit in a unitary way, which is essential in our neuro-linguistic approach of sentence as a cognitive object to allow to investigate the neural correlates of Topic-Comment articulation in Chinese and its syntactic complexity.

C- Ordering constraints in the Topic-field An additional advantage of the Cartographic framework is offered by its approach to the different types of Topics we met in previous section. Following the rule "*one morpho-syntactic property = one feature = one head*" the cartographic study of the sentence structural articulation establishes a typology of syntactic positions dedicated to Scope-Discourse features that we will illustrate in section 3.4.5 following the work of Badan (2007 and much subsequent work) and Badan and Del Gobbo (2010).

While the Topic projection [Top] is broadly speaking preceding the 'grammatical subject' in the Left-Periphery, the Cartographic approach defines a fine characterization of the so-called Topic Field where the ordinal position of each functional projection is precisely determined compared to other Topic types or elements that occupy the Left-Periphery. A type of ordering configuration we already observed for relative or interrogative pronouns, in the previous chapter 2 (§2.4.4.1).

In the perspective of investigating the neural substrates of these ordering constraints, section §3.4.5 will carefully demonstrate the exact position in the Left-Periphery of various kinds of Topics in Chinese to be then able to correlate the relative ordering of these positions with brain activity, a complexity measure corresponding to the representation of the height in the syntactic-tree (cf. §2.4.4.5, p.217). Reviewing the different kind of topics will also be the occasion to offer a description of the Mandarin Left-Periphery (§3.4.5) and to introduce the characteristics of all the Left-Periphery phenomena that will be the experimental conditions in chapter 7. A particular focus will be dedicated to Scene-setting Topics, which will further be at the center of our concern in chapters 4 and 5.

D- Overt versus Covert dependency-links between Topic and Comment Based on the overview presented in previous section, we can draw an initial typology of the Topic constructions that are going to be analyzed in the following sub-sections. We will mainly distinguish four types syntactic configurations used to express the sentence's Topic-Comment articulation in Mandarin Chinese: (3) Gapless Topic-Comment articulations, (1) Left-Dislocated elements, (2) Hanging Topics, (4) Scene-setting Topics, and (5) double subjects constructions.

(165) Topic-comment dependency-links

a. Aboutness Topic-Comment: shows a generic aboutness relation with the Comment, where no gaps, no grammatical link or dependency-link are found between Topic and Comment

b. Double-subject constructions: are often featuring a part-whole relation between the two sentence-initial NPs (cf. domain-subset in Kroeger, 2004). As no dependency-link is found between the Topic and the Comment, we will therefore call Aboutness Topics

c. Hanging Topics: co-referential with a overt pronominal

d. Left-Dislocations (Topicalization): the Topic is related to an empty element in the Comment clause

e. Scene-Setting Topics: the Topic position hosts Bare NPs or Adverbials setting the spatial or temporal frame of the Comment, and no dependency-link is found between Topic and Comment

These five different syntactic types correspond to different overt and covert linguistic means to achieve the link between Topic and Comment: going from resumption, null pronominals, gap strategy, to *dangling* gap-less option and semantic link (e.g. part-whole, set-subset, etc.). Section §3.4.4 is dedicated to the analysis of these dependencies, in order to present the syntactic details of our experimental hypothesis for chapter 7.

We will now start by analyzing the syntactic properties of some of these types of Topic-Comment sentences, to give evidence for the basicness of this sentence articulation in Chinese thereby show the non-derivational nature of some kind of Topics. For this, we will put forward a formal argumentation based on a series of syntactic tests differentiating Topic-Comment articulations for which a movement analysis (i.e. A-bar Topicalization) is available from those where no-movement analysis is possible (§3.4.2, p. 366).

3.4.1 Syntactic "basicness" of Topic-Comment in Chinese

One of the most remarkable properties of Topic-prominent languages compared to Subject-prominent ones is the extent to which Topic-Comment articulation is to be considered as belonging to the repertoire of most basic types of sentences. Since the very first moves of this typological description the structural 'basicness' of Topic-Comment sentence's articulation has been put forward in following terms:

> "The pervasiveness of sentences like this type provides very clear evidence against a process of topicalization." Li and Thompson (1976:479)

.

In this way, Li and Thompson were clearly stating (1) that in Chinese, and other Topic-prominent languages, the Topic could be analyzed as a non-dislocated element, and (2) that the Topic-Comment articulation is not founded on the subject-predicate relation.

Since the pioneer work of Li and Thompson (1976), Topic-prominence definition has be refined in many aspects and even recast in the framework of Principle and Parameters. The properties of Topic-prominent languages described in the typological approach have been reinterpreted, for example by stating that a topic-prominent language is characterized as one language that tends to activate the Topic Phrase (TP) (Xu Liejiong, 2001). Most recently, Paul and Whitmann (2015), have been reinterpreting the topic-prominence parameter stating that the salient feature of the languages traditionally labeled topic-prominent is that they do not fill the functional head of the Topic projection by movement, but this head is filled via External Merge (Base-Generation) of the Topic (and its particles when present).

3.4.1.1 Aboutness Gapless Topics and their irreversibility

A first and central piece of evidence for the a non-movement account of some of the above Chinese Topic-Comment configurations comes doubtlessly from the existence of "gapless" Topic sentences.

We clearly see that examples (166) to (169) present some Topics which feature the kind of "aboutness" relation between Topic and Comment that is not established by any grammatical encoding of the link between Topic and Comment, if not the purely structural one of being in a Topic-Comment construction.

(166) Double Subject constructions

a. part-whole: 长颈鹿 (呀)，脖子长。

Chángjǐnglù (ya), bózi cháng.
giraffe (TOP.) neck long

'As for giraffes, their necks are long.'

b. set-subset: 水果 (呀)，我最喜欢樱桃。

Shuǐguǒ (ya), wǒ zuì xǐhuan yīngtao.
fruit (TOP.) I most like cherry

As for fruits, I like cherries best.'

(167) Aboutness Topic

a. 那场火 (呀)，幸亏消防队来得快。

Nà-chǎng huǒ (ya), xìngkuī xiāofángduì lái de kuài.
that-CL. fire (TOP.), fortunate fire-brigade come DE quick

'As for that fire, fortunately the fire brigade came quickly.'

(168) Scene-Setting topics

a. spatial frame: 绍兴县宋和村泥石流已经停止泛滥。

Shàoxīngxiàn Sònghécūn níshíliú yǐjīng tíngzhǐ fànlàn.
Shaoxing-district Songhe-village landslide already stop spread

'In Shaoxing Songhe village the landslide already stopped spreading.'

b. Temporal frame: 六点，他报告已经写好了

Liùdiǎn tā bàogào yǐjīng xiě-hǎo le.
6-o'clock he report already wrote-RES(finish) LE-PART.

'At 6 o'clock, he had already finished writing the report' from Li(2007)

c. Adverbial: 1968 年八月二十二日 (呀)，我正好二十一岁。

1968 nián bā yuè èrshí'èr rì (ya), wǒ zhènghǎo èrshíyī suì.
1968 year 8 month 22 day (TOP.) I exactly 21 age

'As for August 22, 1968, I was exactly 21 years old.'

(169) Multiple topics: 这个大学，两个学院，一个理工，一个文史。

Zhè-ge dàxué liǎng-ge xuéyuàn, yīge lǐgōng, yīge wénshǐ.
this-CL. university, two-CL. colleges, one-CL. sciences, one-CL. humanities

'This university [has], two colleges, one Science and Technology, the other Arts and Humanities' Li (2007)

This rich set of examples present a panoramic view of the main kinds of *in-situ* Topics, that since Chafe (1976:50) have been called "Chinese style" Topics. Having considered the functional description of these types of Topic-Comment sentences in previous section, we can now analyze the details of their syntactic derivation.

The Chinese's linguistic system offers some linguistic evidence for considering some Topics to be non-dislocated elements, and the above sentence configurations are probably the syntactic articulations where this non-derivational aspect appears in a quite evident way, as they feature no gap and no extractions site.

All these different Topic-Comment configurations are namely called 'gapless'[129], in that they can difficulty be reverted or converted into a Subject-Predicate articulation where, for instance, the Topic NP could enter in an argumental (theta-)relation with the verb in the Comment. As shown in examples (170) the Topic cannot be an argument of the Comment's verb, and cannot involve movement, as no extraction site is to be found.

The following examples taken from Paul (2015), clearly illustrate this point, though the ungrammaticality of the corresponding sentences in (b) where the Topic phrase is tentatively incorporated into the Comment-clause (TP)[130]:

(170) Gapless-Topics no possible reconstruction[131]

a. Gapless-Topics are irreversible (1)

i. 这件事，你就放心吧

[DP Zhè-jiàn shì], nǐ jiù fàng xīn ba
this-CL. matter 2SG. then put heart SFP

'[Concerning] this matter, you can put your mind at ease.' Lü Shuxiang (1986:334)

ii. * 你就放心这件事吧

Nǐ jiù fàng xīn [DP zhè-jiànshì] ba
2SG. then put heart this-CL. matter SFP

'You can put your mind at ease this matter.'

b. Gapless-Topics are irreversible (2)

i. 明天的会议，大家都通知到了

[DP Míngtiān de huìyì], dàjiā dōu tōngzhīdào-le.
tomorrow SUB. meeting everybody all notified-PERF.

'Tomorrow's meeting, everybody has been notified.'

ii. * 大家都通知到了明天的会议

*Dàjiā dōu tōngzhīdào-le [DP míngtiān de huìyì]
everybody all notified-PERF. tomorrow SUB. meeting

'Everybody has been notified tomorrow's meeting.'

c. Gapless-Topics are irreversible (3)

129. By choosing the term Gapless we want to stress, for the non linguist reader, the fact that this kind of Topics are not the result of movement derivation. For the linguist reader we should clarify that we subsume under this term both Scene-scetting Topics, Topics that are in an Aboutness relation with the Comment (like in examples (170)and Hanging Topics.

130. These "gapless" sentences have been studied by a number of linguists (e.g., Chao, 1968; Li and Thompson, 1976, 1981; Xu and Langendoen, 1985; Xu, 2000).

131. Reconstruction test, is a syntactic test in which the movement operation is undone, thus "reconstructing" the pre-movement representation, and allowing the binding principles to apply as if the movement had not occurred (from Andrew Barss, 2003).

i. 我们的中文班，十个学生已经毕业了
Wǒmen de zhōngwén bān, shí ge xuéshēng yǐjīng bì yè le
1PL. SUB. Chinese class, ten CL. student already finish study SFP
'Our Chinese class, ten students have already graduated.'

ii. * 十个学生已经毕业
Shí ge xuéshēng yǐjīng bì yè [DP wǒmen de zhōngwén bān le]
ten CL. student already finish study 1PL. SUB. Chinese class SFP
'Ten students have already graduated our Chinese class.'

This first syntactic test is demonstrating the 'basicness' of this sentence's articulation: these topics are generated in their surface position (i.e. based-generated). This adds further evidence for considering Topic-Comment structures as syntactically basic inside the linguistic system of Chinese (cf. §3.2.2, p.292).

This basic character is one of the fundamental (if not the one) aspects that determined our choice of focusing on this sentence articulation in Chinese for our neuro-linguistic investigation of the sentence-unit. In more trivial terms, by discovering it existed another basic way of articulating the structure of the sentence unit, we decided to try and see how the brain could manage that.

The absence of gap is not only interesting from the point of view of establishing a syntactic derivation, but it also shows the Topic and the Comment can establish an 'aboutness' predication relation even when no antecedent-gap relation exists which is, from the point of view of cognitive processing of the sentence, even more interesting. This means namely that the aboutness relation encompass the whole Comment-clause and not just one element in it. In this regard, Aboutness Topics differ remarkably from the other Topic types, in that the sentence-initial Noun Phrase is semantically related to the Comment but not co-referential with any empty syntactic position, or pronoun or full Noun Phrase, the comment as a whole 'is about' the Topic.

c1	SVO Baseline	Subj. determiner	DE	Subj.	adv. + VP	complement 补语	**Baseline** no Left-Periphery
SVO		That tree	DE modif	leaves	already became yellow	in great quantity.	no TOP, no Left Periphery
		那颗桃树	的	叶子	已经黄了	一大片。	SVO sentence with a 'complex' subject
c2	Base-Generated gapless Topic	Topic	comma	Subj.	adv. + VP	complement 补语	**Aboutness Topic (AT)** position in the Left Periphery > higher in CP
about-ness B-G		That Tree	,	leaves	already became yellow	in great quantity.	Base-Generated Topic in LeftPeriphery
		那颗桃树	,	叶子	已经黄了	一大片。	Pause marking

Table 3.10 – To compare the neural correlates of SVO and Topic-Comment articulation.

Considering the first two experimental sentences of our design, (c1) and (c2), in 3.10, we can observe how this minimal pair will be used to try and to give experimental evidence to the difference between Topic-comment and Subject-predicate relation. In the contrast (c2) > (c1), firstly we do not expect not to find a very broad effect given the basic character of this sentence construction in the linguistics system of Chinese, and secondly we expect to observe an activation linked to the representation in the Left-Periphery where the Topic element *nà-kē táoshù* 那棵桃树 this tree' is hosted, while (c1) does not have any element in the sentence-discourse inter-facial layer.

3.4.1.2 Scene-setting Topic as a Spatial Frame

We turn now to an issue that has been greatly debated in the literature relative to the function of 'Chinese-style' Topics and their informational content and role as frame-setters.

We saw that Topichood is a much debated issue in Chinese linguistics and beyond, and even if the very first attempt to define Topic-Comment constructions dates back to Father Mullie (1932) and his first formulation of 'double subject' constructions, one of the most widely agreed definition of sentence Topic is "what the sentence is about" (Hockett, 1958:21).

The characterization of Topics in the linguistic literature has revealed that topics can be both given and new information (Reinhart, 1982; Paul, 2014 for Mandarin Chinese examples) and that aboutness and frame-like definition are still the two main functions a of Topichood. Although Chafe's foundational definition seems giving precedence to Frame-setting function, we can namely read: "real topics (in Topic-prominent languages) are not so much what the sentence is about as the frame within which the sentence holds" (1976:51).

(171) Scene-setting Topics in English

a. On the way home, he ran into Murgatroyd.

b. In Brazil, such situations are not common.

c. Under other conditions, I would agree to your request.

For instance, the above English clause-initial non-subjects constituents in (171) are not always Topical discourse entities, they are very often Frame-Setting expressions, as in (171a) and (b), or they may serve a contrastive function, as in (171c).

Thus, this difficulty to establish a one to one correspondence between this syntactic position and a particular unction and informational content is a interesting testing ground to investigate the interaction between the syntactic and semantic properties of Topic-Comment constructions. The case of Scene-setting topics is particularly relevant in this regard as extensively documented in previous section on Scene-setting Topics.

We can briefly sum up that in Chinese the only cases where adverbials have topical status are when they appear as temporal or spatial phrases, and adverbial fronting has been regarded as 'a special case of topicalization' by Xu and Langendoen (1985)[132]. Contrary to what was showed for French (in §3.3.2), connectives such as *ranhou* 'then', *keshi* 'but' and *yinwei* 'because' although appearing in the pre-subject position, they are not identified as topics and are just referred to as 'sentence-linking adverbs' (Li and Thompson, 1981) .

Under Chao's (1968:73) analysis, a sentence-initial locative phrase represents as topic [133], since it 'needs not represent the actor, it can, among other things, represent the place at, place to, object for.' Li and Thompson (1981) also suggest that temporal and spatial adverbial phrases are Frame Topics in that their interpretaitonal properties are to set the frame for the sentence. They may be followed by a pause and optionally by a particle.

(172) 昨天晚上我没睡觉。

Zuótiān wǎnshang wǒ méi shuìjiào.
yesterday evening I no sleep

132. For a comprehensive account of adverbials in Chinese see Lu Peng's PhD Dissertation (2003)
133. Note that under his terminology it is the subject of the predication.

'Last night, I did not sleep.'

(173) a. 在台湾你可以吃到很多种水果

Zài Táiwān nǐ kěyǐ chīdào hěn duō zhǒng shuǐguǒ
at Taiwan you can eat very many kinds fruit

'In Taiwan, you can eat many kinds of fruit.'

b. 台湾你可以吃到很多种水果

Táiwān nǐ kěyǐ chīdào hěn duō zhǒng shuǐguǒ
Taiwan you can eat very many kinds fruit

'In Taiwan, you can eat many kinds of fruit.'

Interestingly, these kind of Topics can co-occur with other kinds of topics like Aboutness topics, as illustrated by sentence in (174) and when this happens they preceded Aboutness Topics.

(174) Scene-Setting Topic > Aboutness Topic

a. 我的花园，那棵树，叶子大

Wǒ de huāyuán, nà-kē shù, yèzi dà.
I DE garden that-CL. tree leave big

'In my garden, that tree, leaves are big.'

b. * 那棵树，我的花园，叶子大

Nà-kē shù, wǒ de huāyuán, yèzi dà.
that-CL. tree I DE garden leave big
'That tree, in my garden, leaves are big.'

Badan (2007) notes that the difference between the left-periphery structure proposed by Benincà and Poletto (2004) from Italian (and serveral of its dialects), i.e. [Topic-field[HT][Scene Setting] [LD] [LI]]]]] [FocusField [FocusP....]...], and the one attested in Chinese differs in the ordering constraints of scene-setting topics probably partly because prepositions are obligatory in Scene-setting topics in Italian. We can read: "As for the reasons that cause this parametric difference, at this point it is only possible to speculate that such a variation may have to do with the existence of Aboutness Topic in Chinese and the different use of prepositions for this kind of "scene-setting" Topics in the two languages (Chinese and Italian)." (2007:44).

c2	Base-Generated gapless Topic	Topic	comma	Subj.	adv. + VP	complement 补语	**Aboutness Topic (AT)** position in the Left Periphery > higher in CP
about-ness B-G		That Tree	,	leaves	already became yellow	in great quantity.	Base-Generated Topic in LeftPeriphery
		那颗桃树	,	叶子	已经黄了	一大片。	Pause marking
c3	scene setting Gapless Topic	Topic *Space*	comma	Subj.	adv + VP	complement 补语	**Hanging space Topic** position in the Left Periphery > lower in CP
space B-G		That Garden	,	leaves	already became yellow	in great quantity.	Base-Generated scene-setting Topic
		那座果园	,	叶子	已经黄了	一大片。	Topic localizing the Comment

Table 3.11 – Comparison Scene-setting Topic vs Gapless Abooutness Topic.

Addressing this issue linked to the informational content of Topic, our experimental design in chapter 7 contrasts the Scene-Setting Topic in condition (c3) in Figure 3.11,

where the first DP plays the role of a spatial frame localizing the following Comment clause, with (c2) Gapless Aboutness Topic condition. The two minimally differ in terms of lexical material, so that, comparing (c3) >(c2), will tentatively show the difference between a Topic setting an individual framework and a Topic setting a spatial one following Chafe's definition: "what topics appear to do is to limit the applicability of the main predication to a certain restricted domain. [...] Typically, it would seem the topic sets a spatial, temporal, or individual framework within which the main predication holds (1976:50).

3.4.2 Topicalization versus in-situ Topics

3.4.2.1 Movement-Derivation of Topics in Mandarin

Topicalization and wh-movement: a brief brush up in English

By asking the question whether the Topic occupies the sentence-initial position as the result of syntactic movement from a position within the Comment or whether it is base-generated in that position (i.e. a Base-Generated in-situ Topic), we introduce one of the crucial aspects that led us to investigate the cerebral representation of this derivational distinction[134].

For the reader that would have forgotten or skipped some essential part of chapter 2, a brief reminder is offered in the next pages to be able the grasp the main syntactic aspects linked to topicaliztion and wh-movement we already presented there. We will take a few English examples to recall the arguments that brought to analyze topicalization as a wh-movement.

Presenting movement related word-order transformations in §2.4.2, we underlined that considering topicalization firstly implies to acknowledge the existence of a default or canonical word-order in sentence (§2.2.2). Yet, the default for Chinese, in terms of frequency, is the SVO word-order (cf. §3.2), as in English.

If we consider the following examples in (175) comparing them to their corresponding dislocated versions in (176), we clearly see that the sentences in (175) are in canonical order the subject (in boldface) occupies clause-initial position, while their dislocated version are not, non-subjects (in italics) precedes the subject.

(175) a. **She** totally detests Charlie's parents.
b. **A single lion** can devour a little wildebeest like that in less than a minute.
c. **I** recognize this kind of situation from my previous job.

(176) Topicalization in English
a. Charlie's parents, **we** heartily detest.
b. A little wildebeest like that, **a single lion** can devour in under an hour.
c. This kind of situation, **I** recognize from my previous job.

134. The title of this section is here to remind the reader about this cardinal derivational distinction. Topicalization will refer to the syntactic transformation/operation moving a constituent in the sentence to the topic position through movement. While the term Topic, will only refer to the element occupying Topic syntactic position, and Base-Generated *in-situ* Topics will indicate the kind of Topics where no movement-base analysis can be driven to derive the Topic-Comment articulation of the sentence.

Clause-initial non-subjects are often analyzed as discourse Topics in English, and the movement illustrated in (176) has therefore come to be known under the standard term of as Topicalization.

Topicalization as an instance of wh-movement As we saw in chapter 2 (§sections 2.4.2 and 2.4.3, p.185), since Chomsky 1977's influential article, Topics have been analyzed as an instance of A-bar (non-argumental) movement called wh-movement. The term wh-movement is generally used to refer to any instance of movement to Spec CP, regardless of whether the moved constituent is a wh-phrase. Namely, topicalization obeys to a series of constrains on the syntactic environment out of which a topicalized element can be extracted from, the so-called *Island Constraints* (§2.4.3, p. 186) as illustrated in example (177).

Moreover Gapped Topic structures can be derived by Abar-movement that is defined as long-distance movement, and it establishes an antecedent-gap relation that can cross multiple clause boundaries, like in "That girl$_i$, I think John believes that Bill likes ___$_i$". This dependency relation is sensitive to subjacency Condition (Chomsky, 1973) as witnessed by impossibility of extraction from Subject Island (177e) or the violation of the complex NP constraint (Ross, 1967) in example (177c). The following English examples in (177) briefly recapitulate what we have been presenting in 2.4.2 and 2.4.2.1.

(177) wh-movement and Island constraints

a. Wh- complement: *[For next class]$_i$, they have forgotten [which problem they should solve t$_i$].

b. Noun complement: *[Yves Bonnefoy]$_i$, they don't believe the claim [that he has met t$_i$].

c. Complex NP: *[That scholar]$_i$, you brought [the book that criticize t$_i$].

d. Relative clause: *[Yves Bonnefoy]$_i$, I dislike the journalist [who interviewed t$_i$ for Causeur].

e. Sentential subject: *[His addiction to alcohol]$_i$, [admitting t$_i$] nearly destroyed his career.

f. Left branch of noun phrase: *[Julia's]$_i$, we heartily detest [t$_i$ parents].

g. Coordinate structure: *[Tiramisu]$_i$, Murielle ordered [coffee and t$_i$].

3.4.2.2 Debates about *in-situ* Topic and Topicalization in Chinese

While the Movement-base account for Topic-Comment articulations is generally taken for granted in many languages like English, the literature on Chinese Topic structures has been divided into two contrasting approaches for the last 30 years in the issue of the syntactic derivation of Chinese topics: one side arguing for movement to sentence-initial position and the other arguing for base-generation stating that Topics are generated in their surface position.

The question whether the Topic is hosted in sentence-initial position, in the so-called Left-periphery of the sentence, as the result of a fronting movement from a position within the Comment or whether it is Base-Generated in the sentence-initial position, has

been long debated[135] resolved into a third direction that is now largely agreed on: some topics are derived by movement and others are not (see Audrey Li, 2000; Huang, Li and Li, 2009, p.198).

Abar-Movement criteria for Gapped-Topics

According to the consensus analysis, not all topic in Chinese Topic-comment structures are derived in the same manner, we already went through the Gapless in-situ ones and will concentrate now on the topic structures containing a gap that can be interpreted as co-referential with the Topic.

We will call "*Gapped-Topic sentences*" the Topic-Comment constructions that show the distinctive characteristics of wh-movement that were previously recapitulated for English Topicalization (cf. ex. 177).

In general terms, Topics will be regarded as movement-derived, if (i) leaving an empty syntactic element behind, (ii) showing to be sensitive to Island Conditions and (iii) showing reconstruction effects.

According to Xu (2007) the crucial test is to see whether a Topic can be related to a gap in a syntactic island, such as complex Noun Phrases and sentential subjects. These two main diagnostics for movement-analysis, can be reinforced by some tests put froward by Badan (2008), who showed among others that in Chinese Prepositional Phrases can be topicalized and parts of idiomatic expressions can be topicalized too.

Concretely, to attribute A-bar properties to Gapped Topic , the following syntactic tests, proposed by Rizzi (1997), can be performed in Chinese as shown by Badan and Del Gobbo (2010) and Shyu (2001)

(178) Abar-Movement properties for Gapped Topics

a. 1- Reconstruction is possible as illustrated by the binding of anaphors, Topics can have a resumptive element in the sentence:
张三，李四说玛丽很喜欢他。

Zhāngsān$_i$, Lǐsì shuō Mǎlì hěn xǐhuan tā$_i$
John Lisa said Mary very like him

'John, Lisa said Mary likes him very much.'

b. 2- Topics don't show WCO effects:
张三，他的妈妈批评了。

Zhāngsān$_i$, tā$_i$ de māma pīpíng-le e$_i$.
John, his mother criticized e$_i$.

' John, his mother criticized [him].'

c. 3- No bare Quantificational element can be fond in topic position:
* 每个人，李四都没看见。

* měi-ge rén, Lǐsì dōu méi kànjiàn.
Every-CL. person, Lisa all NEG. see

'*All the people, Lisa haven't seen [them].'

135. Some linguists have argued that all topics in Chinese are Base-Generated (Xu and Langendoen, 1985; Xu, 1986; Xu Liejiong 2006), while others have reported movement analyses (Huang, 1987; Shi Dingxu, 2000).

d. 4- Topic it does not interact with another + Q operator such as a wh-element[136]:
这本书，你还给了谁?

Zhè-běn shū, nǐ huán gěi-le shéi?
This-CL. book, 2SG. restitute to-ASP. who

'This book, who did you restitute [it].'

Gapped Topic-Comment sentences and long distance movement Moreover the Gapless Topics in example (179) illustrates a kind of Topic associated to an empty syntactic element that has been extracted from the Comment-clause

(179) 这个孩子，张三知道李四看见外婆在画。

Zhège háizi$_i$, [$_{CP}$ Zhāngsān zhīdào [$_{CP}$ Lǐsì kànjiàn [$_{IP}$ wàipó zài huà t$_i$]]].
this-CL. child Zhangsan know Lisi see grandma PROG. draw

'As for this child$_i$, Zhangsan knows that Lisi saw the grandma drawing [him]$_i$.' (Hu, 2014)

As we can see in (179) the displaced element – the object of the sentence *zhège háizi* 'this child' – has reached the sentence-initial position from the object position across several clause boundaries establishing a long-distance dependency, and left a gap in its original post-verbal object position[137].

This test further shows that in Mandarin the Topics that are associated to gaps in the comment-clause are in an A-bar *antecedent-gap* relation that can cross multiple clause boundaries, where the topic element original position is deeply embedded, and has been moving through a number of intermediate CPs to reach the Topic Field in the Left-Periphery.

We can leverage on the fact Topics cannot be extracted from Complex NPs (cf. English example (177c)) to further attest the movement-derivation of Gapped Chinese Topics. Examples like (180) have been provided in the literature to show that Gapped Topics observe Subjacency. For instance the agrammaticality of extracting the nominal *Lisi* from a Complex NP Island (in bold) is indeed confirming this:

(180) Complex NP Constraint in Mandarin:
* 李四，我很喜欢唱歌的声音。

* Lǐsì$_i$, wǒ hěn xǐhuan [[[**t$_i$ chàng gē**] **de**] **shēngyīn**].
Lisi I very like sing song DE voice

*'As for Lisi$_i$, I like the voice with which ___t$_i$ sings.' from Huang, Li and Li (2009:210).

This test indicates that Gapped-Topics are sensitive to some of the Island conditions recapitulated in (177), hence, it further permits to establish that these Topic-Comment structures are derived by movement.

To this example (181) adds further evidence showing that the presence of a resumptive in the same position of the gap yields a felicitous and grammatical sentence. While the presence of the resumptive is optional in the pair of simple Topic-Comment sentences in (182), it is here necessary to save the structure where a complexe NP is present, as witnessed by (181) – from Badan (2008) and Shyu (1995):

136. Note that this last test allow us to say that Chinese Topic is not a [+Quantificational] Operator, which would be essential in the comparison with Focus conditions in chapter 7.
137. Sentences with such long distance dependency are allegedly acceptable even though the naturallity can be questioned out of context

(181) Complex NP Constraint in Mandarin: 李四，我很喜欢她唱歌的声音。

Lǐsì$_i$, wǒ hěn xǐhuan [[[**tā**$_i$ chàng gē] de] shēngyīn].
Lisi I very like she sing song DE voice

'As for Lisi$_i$, I like the voice with which she$_i$ sings.'

(182) Optional resumptive pronoun in simple Topic-Comment sentences

a. 张三啊今天下午我看見了。

Zhāngsān$_i$ a, jīntiān xiàwǔ wǒ kànjiàn e$_i$ le.
Zhangsan TOP. today afternoon 1SG. saw e$_i$ PART.

'As for Zhangsan, I saw [him] this afternoon.'

b. 张三啊今天下午我看见他了。

Zhāngsān$_i$ a, jīntiān xiàwǔ wǒ kànjiàn tā$_i$ le.
Zhangsan TOP. today afternoon 1SG. saw 3SG. PART.

'As for Zhangsan, I saw him this afternoon.'

Importantly, whenever a resumptive pronoun is inserted in the gap position of an otherwise ill-formed Topic-Comment sentence it becomes acceptable. We will resume to the importance of resumptives and to their distribution patterns in next section 3.4.4, in order to illustrate the importance of overt versus covert encoding of the dependency-link between Topic and Comment in our neuro-linguistic investigation of the cerebral representation of the sentence-unit.

On the basis of the different syntactic tests presented here, we can conclude that in Mandarin Chinese Topic structures are derived by movement when a gap occurs and only then. Anticipating over what we will extensively argue for in §3.4.4, two cases where no movement occurs can be distinguished where the gap is a null pronoun or when there is no gap at all.

Psychological reality of Gapped Topics

As extensively documented in chapter 2 (§2.4.3.2), gaps left by movement have a psychological reality. Previous psycho-linguistic studies have found that the moved wh-word (i.e. the filler) is reactivated immediately after the verb (the gap, i.e., the hypothesized trace site), lending support to the psychological reality of wh-movement (Zurif et al. 1993; Nicol et al. 1994). Priming effects described in this literature (i.e. a facilitation effect) indicate a reactivation of the meaning of the object in its original post-verbal position, meaning that the object meaning is re-accessed and lending support to the theory of syntactic movement.

As for this kind of evidence in Mandarin Chinese, we can report only two of the rare psycho-linguistic and neuro-linguistic studies that were carried out on Topic-Comment structures.

Firstly, a moving window self-paced reading experiment by Huang and Kaiser (2008) investigated the on-line processing of constructing filler-gap dependencies in Chinese topicalized sentences demonstrated that the processing course of Chinese topicalized constructions involves the construction of a dependency relation between the moved constituent and its gap (trace) (at the verb site) giving some psycholinguistic evidence for movement analysis for topicalized Topics in Chinese, in that the parser slow dow at the

pregap site in conditions where the verb was semantically implausible can indicate that the parser is trying to construct a filler-gap dependency at this point (the ParasiticGap point in the sentences[138].

Moreover, this study is showing not only that the parser actively searches for a gap but also that it is also sensitive to syntactic restrictions linked to grammatical knowledge, namely grammatical restrictions on parasitic gap constructions (Huang and Kaiser, 2008)[139].

Secondly, a recent study by Yang and Liu (2014) explored the on-line electro-physiological response to the Topic gaps in Topicalized sentences in Chinese like:

(183) 桌子$_i$ 经理踢了 ___$_i$ 两脚。

Zhuōzi$_i$ jīnglǐ tī-le ___$_i$ liǎng-jiǎo.
table manager kick-PAST. ___$_i$ two-CL.foot

'The table$_i$, the manager kicked twice.' Yang and Liu, (2014, ex.4)

138. Huang and Kaiser (2008) investigated a particular type of gap. Parasitic gaps in English and in Chinese are found when in a topic structure, a gap in an adverbial clause can be saved if there is a gap in the main clause. While extracting the topic from within a sentence is subject to locality constraints and therefore gaps inside islands such as adverbial clauses are not allowed,the fact a gap is also found in the main clause license the presence of a gap. Concretely, the examples below illustrate this opposition (a) gap in the adverbail caluse is ill-formed, (b) gap in the main clause is grammatical and the combination of the two yields the grammaticality f (c):

(1) Huang and Kaiser (2008): Parasitic Gap (PG)

a. * 那个员工在老板见过之后大家继续开会。

Nà-ge yuángōng [zài lǎobǎn jiàn-guò zhīhòu] dàjiā jìxù kāihuì.
That-CL. employee at boss meet-ASP. after everyone continue meeting

'That employee, after the boss met (him) , everyone continued the meeting.'

b. 那个员工，在老板见过经理之后立刻就被开除了。

Nà-ge yuángōng, [zài lǎobǎn jiàn-guò jīnglǐ zhīhòu], lìkè jiù bèi-kāichú
That-CL. employee at boss meet-ASP. manager after immediately JIU was-fired
le.
PART.

'That employee, after the boss met the manager, [he] was immediately fired.'

c. 那个员工在老板见过之后立刻就被开除了。

Nà-ge yuángōng [zài lǎobǎn jiàn-guò ___ PG zhīhòu] ___ lìkè jiù bèi-kāichú
That-CL. employee at boss meet-ASP. ___ PG after ___ immediately JIU was-fired
le.
PART.

'That employee, after the boss met [him], [he] was fired immediately.' from Ting and Huang (2008)

Specifically, the reported participants' reading times prove to be sensitive to a plausibility manipulation made on the verb before the gap. The transitive-implausible conditions had significantly slower reading times only at the (V) region than the transitive-plausible conditions. The authors interpret this slowdown at the verb in the transitive-implausible condition as indicating that the parser is trying to construct a filler-gap dependency at this point but has faced a semantic mismatch between the fronted object and the sub-categorizing verb.

139. Given this finding, we expect that a comparable sensitivity to grammatical restrictions would be observable in our experiment comparing condition (c5), where the gap in the comment clause is in free alternation with a Resumptive pronoun, and, condition (c6) where the presence of the inanimate Resumptive Pronoun ta 它 is agrammatical at the gap position.

The results indicate a syntactic dependency between the Topic and the *trace* position comparing the three condition reported in (184):

(184) Yang and Liu (2014) experimental sentence

a. grammatical topicalized sentence condition: 桌子经理踢了两脚

Zhuōzi jīnglǐ tī-le liǎng-jiǎo.
table manager kick-PAST. two-CL.foot.

'As for the table, the manager kicked (it) twice.'

b. Ungrammatical topicalized sentence condition: * 桌子经理踢了王五

* Zhuōzi jīnglǐ tī-le *Wáng Wǔ.
table manager kick-PAST. Wang Wu

* 'As for the table, the manager kicked Wang Wu.'

c. Adverb + SVO s
entence as controls: 昨天经理打了王五

Zuótiān jīnglǐ dǎ-le Wáng Wǔ.
yesterday manager hit-PAST. Wang Wu

'As for yesterday, the manager hit Wang Wu.'

As shown in Figure 3.20B the authors observed is an enhanced N400 component at the sentence-initial NP in (a) and (b), but not in (c). As underlined by the authors, the first NPs across conditions differ mainly in their semantic feature: (a) and (b) are inanimate, but the first NP in (c) is a temporal NP, an adverbial. Hence, this semantic component (the N400) has been interpreted as being associate to the semantic feature of the experimental condition (a) and (b), and not to the Topic function.

Importantly, a *negativity component* is elicited at the verb (see 3.20C), reflecting the retrieval of the topic from working memory in the first two conditions and that a sustained anterior negativity is elicited in the topicalized sentence starting at the NP2 and continuing to be present till sentence-final position, reflecting the processing of maintaining the sentence-initial Topic constituent in syntactic working memory. Importantly, a Sustained Anterior Negativity (SAN) was found wen averaging the whole Comment clause ERPs starting from the Subject position to the end of the sentence (3.20A), it was interpreted as reflecting the processing of storing and holding the moved constituent in syntactic working memory. Last but not least, a P600 component is elicited by the sentence-final word, which may reflect the establishment of syntactic dependency between the moved topic and its post-verbal trace (see (3.20D).

These two studies provide further experimental evidence for movement analysis of Left-Dislocated Topics, like (c5) and (c6), and for the existence of a neural encoding of filler-gap dependency between Topic element and the gap in the Comment clause.

3.4.3 Other moved elements: Contrast and Focus expression in Chinese

In contrast to Topic, the Focal articulation of the utterance involves the focus-background partition of the sentence-unit. From the informational point of view, Focus is generally

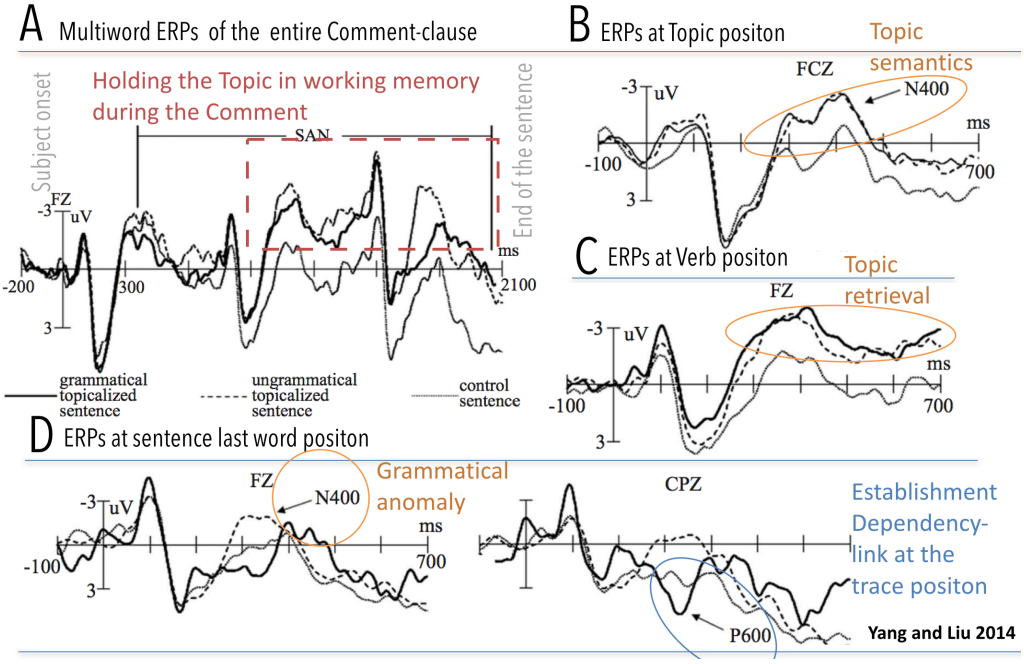


Figure 3.20 – (A) Multiword ERPs Averaged from the Onset of the NP2 to the Position of the Sentence-final Word. (B) Grand Average ERPs at the Position of the Sentence-initial NP1. (C) Grand Average ERPs at the Position of the Verb. (D) Grand Average ERPs at the Position of the Sentence-final Word. Adapted from Yang and Liu (2014).

described as the salient or new informative part of the utterance and 'background' is the non-informative part of it (cf. Prague School; Halliday, 1967 among others).

There exist various types of Foci and the most basic one is the so-called 'informative focus', it can be evaluated by a simple wh-question and answer pair, where the focused element is the object of the question, like in 'Who did you call yesterday?', to which the answer reveals the focus 'I called Antonio from the Lab.'

The overt focalization of an element to express new information, can be manifested in different ways in Chinese: either by a bare in-situ focus, marked by intonational contrastive stress and with no morpho-syntactic markers, or by the structurally pre-verbal even-Focus, that will call *lian-Focus*[140]. Following Gao (1994), we can say that virtually any element in the sentence can be stressed and contrasted.

Bare-focused elements Differently from Topics, the *bare focused element* bearing prominent stress cannot be moved to the Left-Periphery in Chinese as shown in (185c), they remains in-situ, thus preserving the canonical-word order. Consider the question and answer pair in (185) – from Shyu (1995) and Paris (1998):

(185) Q: 张三喝橘汁了马?

140. As for the in-situ focus strategy, Badan and Del Gobbo note that "the fact that Focus and question formation in Chinese adopt the same in situ strategy tells us that they belong to the same class: operator-like elements." (2010:88).

Zhāngsān hē jú-zhī le ma?
Zhangsan drink orange-juice PART. Q.

'Did Zhangsan drink an orange juice?'

a. 不是，(他) 喝葡萄酒了! [SVO]

Bù shì, (tā) hē pútáojiǔ le!
Not to.be 3SG. drink wine PART.

'No, he drank wine!'

b. * 不是，葡萄酒他喝了!

Bù shì, pútáojiǔ tā hē le!
Not to.be wine 3SG. drink PART.

*'No, wine he drank!'

c. * 苹果，张三吃了一个

PÍNGUǑ, Zhāngsān chī-le yīge
APPLE Zhangsan eat-ASP. one.CL.

*'APPLE, Zhangsan ate one.' From Shyu (1999)

In (185), the element bearing contrastive focus can never be fronted to the Left-Periphery, as shown by the ungrammaticality of answer b (in 185b).

As we already briefly introduced while discussing the contrastiveness of SOV word order, pre-posed object, moved to the position between subject an verb, are implicitly and often even prosodically put in contrast with an other item in the context. As shown in the examples below:

(186) SOV versus pre-verbal *lian*-construction

a. Subj + lian-Obj + dou/ye + V: 我连这一本书都/也没有看

Wǒ lián zhè yì běn shū dōu/yě méi yǒu kàn.
1SG. even this one CL. book all/also not have read

'I didn't read even this book.'

b. SOV: 我这一本书没有看

Wǒ zhè yì běn shū méi yǒu kàn.
I this one CL. book not have read

'I didn't read **this book**.' Badan (2007:ex.118)

While (186b) is not a focus but just a contrastive topic – as we will soon demonstrate – the even-construction in (a) features a *lián* 连 is generally associated with the meaning of even. A second element, *dōu* 都, of the focal even-construction is found in pre-verbal position. It is a quantificational element that literally means 'all', and is quasi-fully interchangeable with yě 'also' (see Hole, 2004).

Given the contrastive (186b) and focal interpretation (186a), the above sentences show that the Chinese Left-Periphery properties present also the opportunity to obtain a contrastive interpretation without moving constituents in the Left-Periphery. In cases of sentence-internal even-Focus (*lian-Focus*) and of pre-posed object to pre-verbal position, pragmatic interpretation is actually encoded without implying the movement to the Left-Periphery (Badan and Del Gobbo, 2015 and much related work)[141].

141. This Focus is indeed said to be in the Low Periphery (*à la* Belletti)

Importantly, from the experimental point of view, this property allows the decorrelation of pragmatic and contrastive interpretation from the position in the Left-Periphery. Although featuring a strong contrastive and pragmatic interpretation Pre-posed objects and clause-internal even-constructions do not move to the LP and therefore do not occupying a syntactic position in the sentence CP-layer. For this reason, we selected them as critical conditions for our fMRI experiment on the cerebral representation of syntactic position in the Left-Periphery. Secondly, the fact that, for these two syntactic configuration, A-movement analysis has been given will further allow us to compare cerebral activation to A-movement and A-bar Movement in Mandarin.

Semantic account of Chinese 'even' *lián* 连 scalar focus

As we saw in the above examples in Mandarin Chinese, two lexical items are responsible for conveying the meaning of the English 'even' scalar particle (Koenig, 1991). *Lián... yě* focus construction namely implies scalar focus which is principally defined as based on a series of alternatives to the focused elements that are order on a scale where the focused item is either the least likely or the biggest.

In Paris (1979) we can read:

> "The role of *lián* consists in picking out of a class of elements that are scanned by *yě* and *dōu* one (or more) element(s) which has/ve the least probability of possessing the property predicated of the whole class and in asserting that this/these elements also possess(es) this property. To the class of elements which serves as a reference class, one more element —which was not expected to possess the same property as predicated of the reference class- is added." (Paris 1979:66).

If we consider the semantics of the two elements constituting the Chinese even-construction *lián* 连... 也 *yě*/都 *dōu*, the first *lián* is the focalizer[142], and the other one is the adverb *dōu* or *yě*, respectively 'all' and 'also'.

Specifically, the even-interpretation arises from the generally called 'scalar particles' that accompany *lián* rather than *lián* itself. If the element that needs to be semantically focused receives prosodic stress and moves to a preverbal position, to the left of the adverb *dōu*, the sentence still receives anyway an *even* interpretation :

(187) a. 我连这一本书都没有看

Wǒ lián zhè yì běn shū dōu méi yǒu kàn.
1SG. even this one CL. book all not have read

'I didn't read even this book.'

b. 我这一本书都没有看

Wǒ ZHÈ YÌ BĚN SHŪ dōu méi yǒu kàn.
I this one CL. book all not have read

'I didn't **even** read this book.'

142. The nature of lián, and most of all its syntactic behavior are controversial topics in Chinese linguistics. Traditional Chinese grammars label *lián* either as a preposition (Chao, 1968), or as a 'focusing adverb' (Tsai 1994, 2004), we keep neural calling it a focus marker.

Although the above examples seems to indicate the optionality of *lián*-Focus marker, Paris (1979a) and Tsai (2004) (with reiterated confirmation in the literature) state that the *even*-interpretation of the *lián... dōu* construction is the result of the interplay of *lián* in quality of as a focus particle, with *dōu*.

The function of *dōu/yě* has successively been defined as twofold: (i) a maximality operator (Cheng and Giannakidou, 2006) and (ii) the trigger of movement of the element in focus, the *lián*-XP (see Badan 2008; Xiang, 2010; and Constant and Gu, 2008).

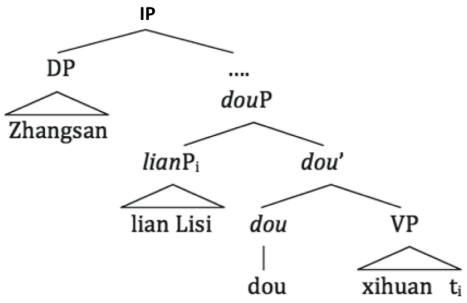


Figure 3.21 – According to Shyu (2001), the sentence-internal lian-XP moves to the left of *dōu* via A-movement, leaving a A-bind trace in object position. According to Badan (2008), *dōu* is not the head of the FocusP but is a maximality operator and *lián* is a semi-lexical projection. Adapted from Badan and Del Gobbo 2010.

Following Badan (2008), and in line with Xiang (2008) and Constant and Gu (2010), we assume that *dōu/yě* adverbs are ultimately responsible for triggering movement. As illustrated by the treee diagram in figure 3.21 the element in focus – the *lián*-XP – needs to move to the Spec of the adverb *dōu* in order to check semantic features. We instead propose that *dōu* operates on the set of alternatives determining the maximal set[143]. Let us concretely see the following example to grasp the essential lines of the semantics of this construction:

(188) 李四连张三都很喜欢

Lǐsì lián Zhāngsān dōu hěn xǐhuan.
Lisi even Zhangsan all very like

'Lisi even likes Zhangsan'

From the point of view of the semantics and of the scalarity it expresses, one can say about sentence (188) that: for every possible alternative to Zhangsan, it holds that Lisi likes the given alternative (more than Zhangsan) and that Zhangsan is the least likely among the alternatives to be liked by Lisi[144].

143. Xiang (2008) proposes that the input to the maximality operator is the alternative degrees ordered on a scale about unexpectedness.

144. Hence, the Focus is Zhangsan, the Scope is '*hěn xǐhuan x*' 'very much like x', and the Scalar implicature can be summarized in the following: For all x under consideration beside Zhangsan, the likelihood that Lisi likes x is higher than the likelihood that Lisi likes Zhangsan. For a more comprehensive and formal description of the semantic meaning of 'even'-*lián* and scalar particles in Chinese, see Constant and Gu (2010)

Hence, in the *lián* 连... *dōu* 都 construction, *dōu* 都, plays the role of a maximality operator, providing existence and exhaustivity of the possible alternatives under consideration. While the focus particle *lián* 连 and alternatively the constituent marked by prosodic focus have the function of indicating the range of what is focalized, which is in general the entire phrase that *lián* takes as its complement[145] (Badan, 2007 & 2008).

Given the complexity of calculating the scalar implicatures that are linked to this focus construction, we expect it to elicit relatively wide spread cerebral activation patterns in more temporally distributed areas, that have been reported as been involved to semantic calculations or to focal intonational patterns both in oral or written presentation (e.g. Lœvenbruck et al., 2005), as we will describe in chapter 7.

3.4.3.1 Sentence-internal *lián*-XP and A-movement

As for the movement analysis of *lián*-Focus, two cases should be distinguished, in that *lián* marked constituents can either be pre-posed to the left of *dōu* 都 and the main verb like in (189a), or can be found in pre-subject position in the Left-Periphery of the sentence like in (189b).

(189) IP-Internal and IP-External Lian-Focus

a. 我连这些书都/也看完了。我连这些书

Wǒ lián zhè xiē shū dōu/yě kàn-wán le.
1SG. even this CL.PL book all/also read-RES.finish PART.

'I read even these books.'

b. 连这些书我都看完了

Lián zhè xiē shū [wǒ dōu kàn-wán le].
Even this CL.PL book 1SG. all read-RES.finish PART.

'I read even these books.'

The syntactic and interpretive differences between sentence internal *lián... dōu* and sentence initial have been the object of some debates (see Badan, 2008 and Gu and Constant, 2009). Although these can marginally be considered to still be open issues, we will adopt here the approach by Badan and Del Gobbo (2010 and 2016), which appears to be the most comprehensive one (see figure 3.22).

Firstly, based on (Shyu, 1995 and 2001) Badan and Del Gobbo note that IP-internal *lian*-XP is clause bound. It can only move within the clause it is generated in, as illustrated by the agrmmaticality of the following sentence where the even-constituent marked by lian crossed a clausal boundary in (190):

(190) a. * 张三连玛丽认为 [李四都不喜欢]

Zhāngsān lián Mǎlì$_i$ rènwéi [$_{CP}$ Lǐsì *dōu* bù xǐhuan e$_i$]
Zhangsan even Mali think [$_{CP}$ Lisi *all* not like e$_i$]

*'John thinks that Lisa doesn't like even Mary.' (Shyu 2001:3–5)

b. * 张三连玛丽都认为 [李四很喜欢]

145. Badan proposes in fact that in the *lián* 连... *dōu* 都 construction the maximality operator, *dōu*, is the quantificational element, and not the focus particle *lián* 连.

Zhāngsān lián Mǎlì$_i$ *dōu* rènwéi [Lǐsì hěn xǐhuan e$_i$]
Zhangsan even Mali *all* think Lisi very like

*'John thinks that Lisa doesn't like even Mary.' (Shyu 2001:80)

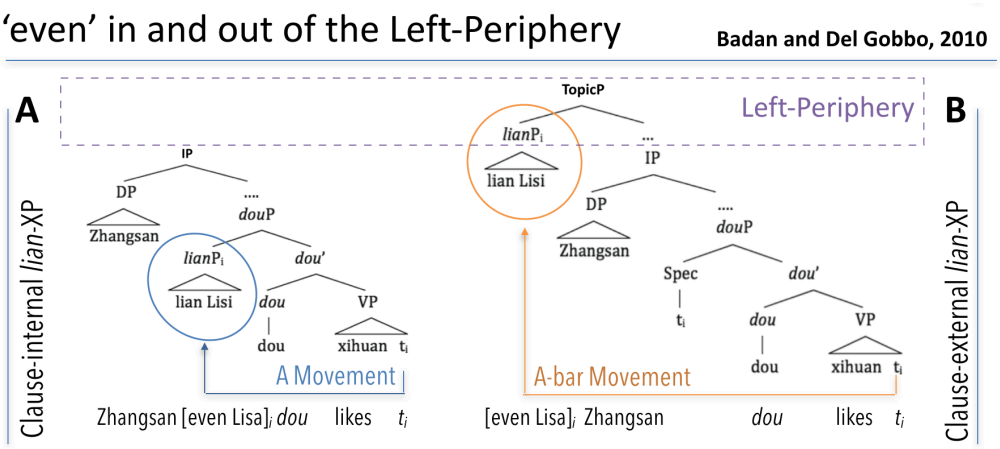


Figure 3.22 – (A) is an instance of A-movement. (B) is an instance of Abar-movement, because the lian-XP can move long distance. Note that, according to Badan (2008), *dōu* is not the head of the Focus Projection but is a maximality operator and that Focus is assigned by *lián* as a semi-lexical projection.

Example (190) features an embedded object that cannot be pre-posed with *lián* across a tensed boundary. This shows that the kind of movement involved in *lián... dōu* construction cannot go long-distance to the position between the subject and the verb in the matrix clause, regardless of the position of the adverb *dōu*, be it in the matrix clause as in (b) or in the embedded as in (a)).

If we consider reconstruction effect as a diagnostic test for A-movement, we observe in (191) that for sentence-internal lián-XP no reconstruction effects for principle C of the Binding Theory are found, in that the co-reference between the pronoun *ta* and its antecedent Zhangsan is not possible in (191).

(191) * 我连 [张三$_i$ 的书]$_j$ 都被他$_i$ 抢走了

Wǒ lián [Zhāngsān$_i$ de shū]$_j$ dōu bèi tā$_i$ qiǎngzǒu le
I even Zhangsan$_i$ DE book$_j$ all by him rob.away PART.

'I was robbed of even [Zhangsan$_i$'s book]$_j$ by him$_i$' (Shyu, 1995:83 ex.105)

Secondly[146], no resumptive pronoun is allowed with the internal *lian*-XP construction (Shyu, 1995:90; Ting, 1995:295) as illustrated by (192).

(192) * 李四 [连玛丽]$_i$ 都很喜欢她

Lǐsì [lián Mǎlì]$_i$ dōu hěn xǐhuan tā$_i$
Lisi even Mary DOU.all very like her

146. Additionally Shyu (1995) shows evidence for the fact *lian*-XP focalization remedies Weak-Crossover effects is sensitive to locality conditions like Complex NP extraction, and that the object of Idioms (V-O compounds) can be moved.

Lit: 'Lisi even Mary likes very much her$_i$.'

To sum up, as already pointed out by Shyu (1995), all the properties emerged from the above tests are evidence that in sentence-internal position, lián-XP has undergone A-movement.

Sentence-external lián-XP is A-bar movement A closer look to example in (189) reveals that *even*-constructions can be found in pre-subject position, suggesting that it can be hosted in the Left-Periphery of the sentence too[147].

Interestingly, Sentence-initial *lian*+XP compared to its IP-internal counterpart displays a different set of properties. As illustrated in (193), this construction allows long-distance dependencies that are typical of A-bar-chains as we saw in chapter 3:

(193) a. 连玛丽，张三认为 [李四都不喜欢]

Lián Mǎlì, Zhāngsān rènwéi [$_{CP}$ Lǐsì dōu bù xǐhuan t$_i$]
even Mary Zhangsan think [$_{CP}$ Lisi all not like t$_i$]

'Zhangsan thinks that Lisi doesn't like even Mary.' (Shyu, 2001: 3–5)

b. 连玛丽，张三都认为李四不喜欢

Lián Mǎlì$_i$, Zhāngsān dōu rènwéi Lǐsì bù xǐhuan t$_i$
even Mary$_i$ Zhangsan all think Lisi not like t$_i$

'It is even Mary$_i$ that Zhangsan thinks that Lisi doesn't like [her]$_i$.'

Moreover, in sentence-initial position, lian+XP can co-refer with a resumptive pronoun in the 'original' object position (Paris, 1999) as illustrated by (194):

(194) 连玛丽$_i$，李四都很喜欢她$_i$

Lián Mǎlì$_i$, Lǐsì dōu hěn xǐhuan tā$_i$
Even Mary Lisa all very like her$_i$

'Lisi likes even Mary$_i$.' (Shyu, 1995:139)

Importantly, sentence initial lian+XPs can also be followed by Topic markers, something that is not allowed for sentence-internal lian-XP (Paris, 1999):

(195) a. 连这本书 (啊)，[张三都已经买了]

Lián zhè-běn shū (a), [Zhāngsān dōu yǐjīng mǎi-le]
Even this-CL. book TOP Zhangsan all already buy-PART.

'Zhangsan has already bought even this book.'

b. * 张三连这本书啊，都已经买了

Zhāngsān lián zhè-běn shū a, dōu yǐjīng mǎi-le
Zhangsan even this-CL. book TOP all already buy-PART.

'*'

147. Paul (2005) proposed that no Topic can be found under IP-external lián-Focus. Paul (2006 and successively 2012) argues that topic positon in Mandarin is not associated to a fixed informational vale (i.e. old/new), and that Phases carrying both new information or old one can be hosted in Topic position.

The facts just outlined allow us to claim that the movement of the sentence-initial *lián*-XP is an Abar-movement. The different syntactic behavior of the *lián*-XP construction (i.e., sentence-initial vs. sentence-internal) can be nicely accounted for by proposing that in the case of sentence-initial *lián... dōu* we have an instance of topicalization[148]; while in the case of sentence-internal *lián... dōu*, focalization has occurred (Badan, 2008).

3.4.3.2 Bare preposed Object

Considering sentences where the word-order is the same as in the IP internal *lián*-Focus (i.e. SOV word order) shows that Chinese achieves by simple word-order a kind of interpretation that at first sight can be considered as similar to *lián*-Focus in many aspects. Several properties of pre-posed object construction are worth notice. Firstly, we can note that a similar, even if pragmatically and semantically more simple, contrastive interpretation is achieved in Chinese by simple word-order linear configuration by moving the object in pre-verbal position. However, when [+animate] NPs are in pre-verbal position basic word-order marking raises some ambiguity. Consider the following sentence in (196):

(196) a. 他 [张小姐]$_i$ 不喜欢 t$_i$

Tā [Zhāng xiǎojie]$_i$ bù xǐhuan t$_i$.
he Zhang miss not like

Interpretation 1: 'Miss Zhang does not like him.'
Interpretation 2: 'He does not like Miss Zhang.'

b. 他 [张小姐]$_i$ 不喜欢 t$_i$

Tā [Zhāng xiǎojie]$_i$ bù xǐhuan t$_i$.
he Zhang miss not like

Contrastive interpretation: 'He does not like Miss Zhang.'

The possibility of having a switch in Theta-roles in the [NP1+NP2 V] configuration witnessed in (196a) has been repeatedly attested in the literature (e.g. Tsao, 1977; Qu, 1994; Shyu, 1995). The natural interpretation is to take the second NP as the Subject and the first as a Topic, but whenever a contrastive stress is found on NP2, functions are reverted, and an SOV interpretation is attested, as illustrated by (196b).

A closer look at bare pre-posed object constructions leads us to briefly sketch the main tread of the extensive and documented debates the literature offers on this syntactic configuration (e.g. Ernst and Wang; 1995, Shyu, 1995; Paul, 2006 among others)[149].

In fact, different syntactic derivations have been proposed for bare objects found in the linear position between the subject and the verb, and the different positions mainly diverge in their interpretation of pre-posed objects as Topics or as Foci:

148. Paul (2006 and 2012) argues that topic position in Mandarin is not associated to a fixed informational vale (i.e. old/new), and that Phases carrying both new information or old one can be hosted in Topic position.

149. Syntactic studies are not the only to be concerned with Object pre-posing patterns, evidence has been gathered in Old Chinese and some dialects, specifically Bai dialects, that consideration linked to old information and focalization (i.e. new information are implied in the object pre-posing in Old Chinese, while Bai dialect only pre-pose old-information. Interestingly, Bai dialects' object-preposing interacts with negation and yes-no questions syntactic configuration as noted by Ernst and Wang (1995) for Contemporary Mandarin. For a detailed account of dialectal licensing to pre-pose objects and object pronouns, see Xu (2009).

- Focus analysis by Shyu (2001)
- Positional Focus by Huang, A. Li and Y. Li (2009)
- Inner Focus in Tsai (2015): future adverbial associated with the IP layer test (zuotian, mingtian)
- Paul (2002, 2005) analyzes attribute to it the role of sentence-internal Topic
- Badan (2008) analysis identifies it as sentence-internal Contrastive Topic

For instance the analysis of pre-posed Objects as Contrastive Topic analysis, from Badan and Del Gobbo (2015), emphasizes that preposed-objects must be in comparison with items in a set, as illustrated by (197a)[150].

While the contrastive aspect of this word-order is difficult to equivocate as illustrated by the contrast between (197a) and (b), it holds true that this type of sentence construction displays from a syntactic point of view two Focus properties, while most of its characteristics are typical of Topic-like elements. Therefore it naturally qualifies as the possible object of controversies.

It is mainly for these reasons that pre-posed object construction has been successively analyzed as expressing either focus (Ernst and Wang 1995; Shyu, 1995, 2001; Tsai, 1994; Zhang, 1996) or Topic (Paul, 2002, 2005 an much related work).

(197) Contrastive alternatives

a. SOV: 我酒喝 (可乐不喝)

Wǒ jiǔ hē ('Kělè' bù hē).
1SG. liquor drink Coke not drink

'Liquor I drink ([but] Coke I don't drink).' Ernst and Wang (1995:22)

b. OSV: 酒，我喝

Jiǔ, wǒ hē.
liquor 1SG. drink

'As for Liquor, I drink.'

It is true that the object needs a contrastive stress as a focus, but it can be followed by topic particles (Paul, 2002; Badan, 2008) as shown in (198).

(198) 张三，他 [这本书] 啊已经看完了

Zhāngsān, tā [zhè běn shū] a yǐjīng kàn-wán le.
Zhangsan 3SG. this CL. book TOP. already read-RES.finish PART.

'As for Zhangsan, he already read THIS BOOK.'

Movement analysis As for the movement type deriving this SOV sentence construction, we rely on Shyu (1995), who put forward that Bare-preposed objects are instances of A-movement, and proposed the following tests in her analysis.

Firstly, the bare pre-posed object cannot co-refer with a resumptive pronoun, while we saw that OSV topic construction can:

(199) 这只狗，[自己的主人]$_i$ 咬了（* 他$_i$）别人却不咬

150. Bare pre-posed Object can actually appear without any strong stress, but they must be accompanied by an emphatic element like the negation *bù* 不 'not' or the adverb *yě* 'also' (Ernst and Wang, 1995).

zhè-zhī gǒu [zìjǐ de zhǔrén]$_i$ yǎo-le (*tā$_i$) biéren què bù yǎo
this-CL. dog self DE master bite-ASP (*he) others but not bite

'This dog bit its own master, but not others.' (Shyu, 2001:50)

According to Li et al. (2009), differences in resumption between OSV topic structure and SOV structure, seem to indicate a focus analysis in that Focus does not allows for a co-indexed pronoun in post-verbal object position :

(200) a. * 我张小姐$_i$ 不想追她$_i$

Wǒ Zhāng xiǎojiě$_i$ bù xiǎng zhuī tā$_i$
1SG. Zhang Miss NEG. want court 3SG

Lit.: 'I, Miss Zhang don't want to court her.'

b. 张小姐，我不想追她

Zhāng xiǎojiě$_i$, wǒ bù xiǎng zhuī tā$_i$
Zhang Miss, 1SG. NEG. want court 3SG.

'Miss Zhang, I don't want to court her.'

From the above contrast we see that the resumptive pronouns are not allowed. This fact further indicates that the bare pre-posed object is derived by A-movement[151] and because there is a violation of the principle B[152], this indicates that the Topic is not outside of IP.

Furthermore, A-movement is subject to Locality restrictions, and unlike topicalization, Object pre-posing never crosses sentence boundaries. As show by Qu (1994) and in the following by Xu (2007) the object of the embedded clause can actually be fronted at the beginning of the main clause, but it cannot be located between the matrix subject and the verb:

(201) a. 这本书$_i$，老师知道 [学生读过 ___$_i$]

Zhè-běn shū, lǎoshī zhīdào [xuésheng dú guò e].
this-CL. book teacher know student read ASP.

'This book, the teacher knows the students have read.'

b. * 老师这本书$_i$ 知道 [学生读过 ___$_i$]

Lǎoshī zhè-běn shū zhīdào [xuésheng dú guo e].
teacher this-CL. book know student read ASP.

*'The teacher, this book, knows the students have read'

In sum, it appears that the empty syntactic element on the right of the verb is A-bound, since the movement displays several A-properties (see Fu, 1994; Qu, 1994[153]; Ting, 1995; Shyu, 1995, 2001; Zhang, 1996), the same that was diagnosed for sentence-internal even construction above:

151. Moreover, there can be only one bare pre-posed object and multiple ones are banned.
152. see Glossary
153. Argues that in in pre-posed object construction the object NP lands in a A-position base on diagnostic test like No reconstruction and Locality restrictions. Specifically she proposes the following analysis; subject NP and object NP overtly rises to [Spec ArgSP] and [Spec ArgOP], respectively. This entails that there should be an other A-position after the subject.

1. clause-boundness, i.e. Locality restrictions (ex.),
2. absence of Reconstruction for Binding Principle C,
3. absence of resumption.

Nonetheless, as previously noted, this construction is controversial because it also presents clear topic-like properties:

1. presence of topic markers (see ex. 198);
2. impossibility to be cleft by means of shi ...de "be ... DE" (Paul and Whitman 2001);
3. co-occurrence with a focus in situ,
4. definiteness requirement.

All in all, these characteristics brought to assign to preposed objects the the Specifier position of a contrastive topic projection in the Low Periphery (Badan, 2009). Already, Paul (2005) showed that the preposed object in Chinese is located above VP and below IP, in a low periphery position and proposed the following hierarchy for the low-periphery of Mandarin:

(202) Paul's hierarchy for the low-periphery in Chinese:
IP > inner TopicP > even-Focus > vP

These two trends put together led Badan (2009)[154] to propose that preposed objects are moved within IP being dislocated through an A-movement (following Shyu, 1995, 2001; Ting, 1995; Zhang, 1996) and considers them as Contrastive Topics: syntactic Topics that can get contrastive stress, on the basis of the above listed syntactic behaviour and their unequivocable pragmatic/semantic interpretation. This is the position we adopt.

In conclusion, the contrastiveness obtained by simple word order in bare pre-posed object will be contrasted to the Focus interpretation obtained by *lian*-Focus in the same pre-verbal potion in our fMRI experimental design in chapter 7 (see Figure 3.12).

c7	SOV Positional Focus or Contr. Top.	Subject	comma	**Obj.**	adv. + Verb + gap	complement 补语	**Movement targeting the IP**
A mov		Mary	,	that book i	already read __i	once.	Bare preposed-object no movement to the LP
		玛丽	,	那本书i	已经看过 __i	一次。	A Movement => no Resumptives
c8	Sentence-internal *lian* Focus	Subject	***lian***	**Obj.**	***ye*** adv.+Verb+ gap	complement 补语	**sentence-internal *lian* Focus**
A mov		Mary	even	this booki	***ye*** already read __i	many times.	**A-Movement**
		玛丽	连	这本书i	也 已经看过 __i	好几次。	Movement targeting sentence-internal position

Table 3.12 – *lián*-focus vs preposed object i.e. Contrastive Topic.

This will allow to observe the different discursive/pragmatic interpretations that are assigned to morpho-syntactically marked and unmarked pre-verbal objects that crucially are not hosted in the Left-Periphery.

154. Badan 2009 also tests Weak Cross-Over, whose results are not stable.

3.4.4 Topic Anaphoras

> "Le nez de Cléopâtre: s'il eût été plus court, toute la face de la terre aurait changé"
>
> *["The nose of Cleopatra, had it been shorter, the whole face of the world would have been different."]*
>
> Blaise Pascal (1623 – 1662)

Resumption strategy: what is Overt and Covert in sentence structure

Addressing the issue of resumptions strategies of Chinese leads us to put forth a syntactic analysis of the overt and covert dependency-links that are tiding the Topic and the Comment together. We will thus contrast Gapless Topics, Gapped Topics and Topics that are actually adopting different resumption strategies, i.e. null pronominals and overt pronouns.

Going back to some French examples, and avoiding this time the oral performance of our former French president, we can choose these of the thinker Blaise Pascal[155].

(203) French *proleptic* subject

Le nez de Cléopâtre : s'*il* eût été plus court, toute la face de la terre aurait
the nose of Cleopatra : if.it had been more short, all the surface of the earth had
changé.
changed

'The nose of Cleopatra, had ***it*** been shorter, the whole face of the world would have been different.'

In (203), the topic '*the nose of Cleopatra*' is the subject of the following sentence[156] is accompanied by a redundant information in the pronominal '*il*' referring back to it and associating it to the right functional role (i.e the subject) in the Comment clause, which is also the case in the following german example in (205) between 'Peter' and the pronoun 'er'.

As we already noted, in the majority of European languages this sentence articulation is restrained to the oral register and is quite systematically associated to a pause, manifested in the written by a comma or by semi-colon like in the above French example, and sometimes accompanied by changes in intonation or prolonged sounds like 'euh' as illustrated by (204):

(204) French oral

155. Blaise Pascal born and lived in his earthly existence in Clermont-Ferrand the hometown my father family side. He was a great thinker, philosopher and litterateur, and is remembered by History because of his scientific work in Physics on atmospheric pressure, whose unit of measure is indeed the 'Pascal', in his honor.

156. traditionally called proleptic subject in French

ta chemise euh... j'ai oublié de la repasser, j'la repasserai
your shirt er... I.have forgotten to Clitic-3SG. iron, I.Clitic-3SG. iron-FUT.
après, hein?
after, ok?

'your shirt ... I've forgotten to iron it, I'll iron it later, is it fine for you?'

In the following we will define and analyse the properties of these Topic-Comment constructions that are accompanied by some information redundancy, linked to the presence of resumptives and epithets or by Clitic doubling.

3.4.4.1 Hanging Topics: a resumptive pronoun or an epithet as anaphoric device

As illustrated by the above French example, the Topic type we haven't analyzed yet, witness a particular type of relation with the Comment, requiring the Comment-clause to really be about the topic, and not just relevant in relation to the Topic element, as we saw it was the case in '*as for*' topics (cf. §3.1.3.4). This is syntactically manifested by the fact they adopt resumptive or epithet as anaphoric device. Hanging Topics can in fact come at the end or at the beginning of a discourse that is about the referent of the Topic, as in the below examples (205) and (203):

(205) Peter ist in einem Modelleisenbahnverein. Er hat eine CB-Funkanalage, sammelt alte Faxgeräte und liest jeden Elektronikblog, den er finden kann. *Peter, er ist ein ganz schöner Geek*.
'Peter is in a model train club. He has CB radio, collects old fax machines and reads every electronics blog he can get hold of. *Peter, he is quite a geek*.'

The choice between overt and covert The following English examples in (206), the above French and German ones are showing how a Topic NP followed by a clause that relates to it in some way, can feature two distinct patterns, which are formally distinguishable by the presence or the absence of a pronominal element referring back to the Topic.

(206) Pronominal referring back to the Topic Hanging Topics
  a. That man, I saw *him* run away after overhearing that conversation for an hour.
  b. My Fiat Panda, *it* couldn't make it to Brittany without overheating.
  c. Your crazy Japanese knife, I cut myself with *it*.

(207) Silent syntactic element referring back to the Topic in Left-Dislocated Topics
  a. That dump, Bill wouldn't live in ___for anything.
  b. Fred I can't stend ___, but Ursula I can get on with ___quite well.
  c. This French cheese we haven't touched ___yet.

While Hanging-Topic sentences in (206) display an anaphoric element referring back to the Topic, those in (207) and are termed as Topicalized and feature a Gap that is the result of Left-Dislocation.

Hence, what defines Hanging Topics is that there must be a direct predication relation between Topic and Comment and the topical referent must be taken up in the comment, and this role can be played by an epithet too. To illustrate this last point, we report in (208) another German example by Altmann (1981:49).

(208) Die schlanke Blondine da drüben, ich glaube, ich habe ***dieses Gesicht*** schon
the.NOM. slim blonde over there, I believe, I have ***this face*** already
einmal gesehen.
once seen.
'The slim blond-girl over there: I think I've seen *that face* somewhere before'

(208) is particularly representative of a configuration where at first sight there does not seem to be a predication relation between the Topic and the Comment. However, the linguistic strategy of marking the Topic-Comment relation through an epithet *this face* is a frequent anaphoric device for linking the topic *slim blond-girl* to the comment[157], we can find this in Chinese too as illustrated by (209a):

(209) Epithets: a new link between topic and comment

a. HT + epithet 傻子: 张三$_i$，我给 [那个傻子]$_i$ 寄了一封信!

Zhāngsān$_i$, wǒ gěi [nà-ge shǎzi]$_i$ jì-le yī-fēng xìn!
Zhangsan I to that-CL. fool send-PERF. one-CL. letter

'John, I sent a letter to that fool!'

b. * [LD + epithet]: 给张三，我给那个傻子寄了一封信!

?* Gěi Zhāngsān$_i$, wǒ gěi [nà-ge shǎzi]$_i$ jì-le yī-fēng xìn!
to Zhangsan I to that-CL. imbecile send-PERF. one-CL. letter

* 'To John, I sent a letter that fool!' Badan, 2007

The possibility to be resumed by an epithet has been interpreted by Badan (2007) as a distinctive feature of Hanging Topics in Chinese. By following the set of diagnostic features established by Benincà and Poletto (2004) for Italian, Badan (2007) was able to distinguish between Hanging Topics and Left dislocated ones in Mandarin Chinese, and put forward the following contrastive set of features illustrated in table 3.13.

Table 3.13 – Syntactic Tests for Hanging Topics vs. Left-Dislocated Topics.

Hanging Topic (c4)	Left-Dislocated- Topics (c5)
1- only be a DP always requires a resumptive clitic expressing the type of argument no case agreement only number and gender	1- Either a DP or a PP
2- Resumed by epithet and tonic pronoun	2a- No obligatorieness of a resumptive, when PPs resumptive agree in number gender and case 2b- Cannot be Resumed by epithet and tonic pronoun
3- Multiple Hanging topics are not possible	3- Multiple Left-Dislocations are possible
4- HT co-occur with Left Dislocation	

While the test based on the presence/absence of a resumptive clitic is not possible for Chinese, Left-Dislocated Topics were expected to leave a Gap, and the author tested for the Left-Dislocation of Prepositional Phrases (210a), and the possibility to have multiple LDs illustrated respectively in (210) and in (212).

157. For this reason this kind of topics are called Hanging topics and are often assumed to resemble a pointing gesture.

(210) LD topics

a. Left-Dislocation of Prepositional Phrases:
给张三，我寄了一封信。

Gěi Zhāngsān, wǒ jì-le yī-fēng xìn.
To Zhangsan I send-PERF. one-CL. letter

'To Zhangsan, I sent a letter'

b. Left-Dislocation of Prepositional Phrases no resumtive:
给张三，我寄给他一封信。

* Gěi Zhāngsān, wǒ jì gěi tā yī-fēng xìn.
To Zhangsan I send to him one-Cl. letter

* 'To Zhangsan, I sent to him a letter.'

On the contrary HT were not expected to agree in case with the resumptive in the Comment-clause (211a), and were shown to be resumed by an epithet as we saw above in example (209). Moreover, it was shown that HT can co-occur with Left-Dislocations in a given relative order (HT>LD), see example (211c).

(211) HT

a. Case-agreement with the resumptive:
张三，我给他去买东西

Zhāngsān$_i$, wǒ gěi tā$_i$ qù mǎi dōngxi.
Zhangsan I to him go buy thing

'John, I go buy things for him.'

b. no Case-agreement with LDs:
张三，我去买东西。

* Zhāngsān, wǒ qù mǎi dōngxi.
Zhangsan I go buy thing

* 'John, I go buy things.'

c. HT co-occur and precede Left-Dislocations:
张三，从这家银行，我知道我们可以为他借到很多钱

Zhāngsān$_i$, cóng zhè-jiā yínháng$_j$, wǒ zhīdao wǒmen kěyǐ wèi tā$_i$ jièdào hěnduō
Zhangsan, from this-CL. bank, I know we can for him b lot
qián.
money

'John$_i$, from that bank, I know we can for him$_i$ borrow a lot of money.'

substraction

This categorization is even more evident when we try to apply to these to sentences the multiple topic test as illustrate in (212). HT with a resumptive yield agrammaticality while multiple Left-Dislocated Topics give a grammatical sentence.

(212) multiple Left-Dislocated Topics

a. multiple HTs: * 张三$_i$，[李四和小雨]$_j$，他$_i$ 在阅览室看见他们$_j$。

* Zhāngsān$_i$, [Lǐsì hé Xiǎoyǔ]$_j$, tā$_i$ zài yuèlǎnshì kànjiàn tāmen$_j$.
Zhangsan Lisi and Xiaoyu he in reading-room saw them

*

b. multiple LDs: 张三$_i$，[李四和小雨]$_j$，t$_{i/j}$ 在阅览室看见 t$_{j/i}$。

Zhāngsān$_i$, [Lǐsì hé Xiǎoyǔ]$_j$, t$_{i/j}$ zài yuèlǎnshì kànjiàn t$_{j/i}$.
Zhangsan Lisi and Xiaoyu in reading-room saw

'Zhangsan, Lisi and Xiaoyu, he saw them in the reading room.' or
'Zhangsan, Lisi and Xiaoyu, they saw him in the reading room.' Badan (2007)

Notice that the sentences in figure 3.14 are differing only by the presence or absence of a resumptive pronoun in object position, and can therefore be respectively classified into Hanging-Topic type for (a) and Left-Dislocated type for (b). Given the above diagnostics, we will be able to compare the neural underpinnings of different kind of Topics in our fMRI study on Chinese syntactic complexity in chapter 7.

c4	Topic + resumptive	Topic object	comma	Subj.	adv. + VP + **Resumpt.**	complement 补语	**Overt Resumptive Pronoun**
Res pron B-G		(As for) Mister Wu	,	Mary	already knows **him**	for a longtime.	**Base-Generated Topic + RP**
		吴先生i	,	玛丽	已经认识 他i	很久。	Hanging Topic
c5	LD Topic + optional **gap** OSV	Topic object	comma	Subj.	adv. + VP + ***trace***	complement 补语	**Optional Resumptive Pronoun**
A-bar mov		As for Mister Wu	,	Mary	already knows **__ei**	for a longtime.	**Left-Dislocated Topic Abar Movement**
		吴先生i	,	玛丽	已经认识了 **__ei**	很久。	Island test

Table 3.14 – Gap vs. Resumptive in our Experimental design in chapter 7. The contrast between c4 and c5 will feature the syntactic opposition between a HT and a LD Topic.

Hanging Topic vs Aboutness Topics As observed at the beginning of this section the particular type of relation that Hanging topic have with the comment requires more than a vague relevance (cf. '*as for*' Topics). On the basis of the different predicative relations they feature, and also on syntactic grounds -as we will see in the following- Hanging Topics do not qualify as Aboutness Topics. We already put forward that Aboutness topics do no have a grammatical link with the comment-clause as illustrated by (213).

(213) Part-whole and set-subset relations between the Topic and the Subject in the Comment

a. Part-whole relation: 十个梨，五个烂了。

Shí ge lí, wǔ ge làn le.
ten CL. pear five CL. spoil PERF.

'Of the ten pears, five have spoiled.' Xu and Langendoen (1985, ex.75a)

b. Condition c2 in our fMRI experiment: 那颗树，叶子大。

Nèi-kē shù, yèzi dà.
that-CL. tree leave big

'That tree, the leaves are big.' Li and Thompson (1976, 23)

c. set-subset relation: 家人，我不能跟姐姐开玩笑。

Jiārén, wǒ bù néng gēn jiějie kāi wánxiào.
family-people I not can with big-sister make joke

'In the family, I can't joke with my sister.' adapted from Badan (2008)

d. Possessor-possesee: 这个人，我觉得记性特别好。

Zhè-ge rén, wǒ juéde jìxing tèbié hǎo.
this CL. man I feel memory exceptionally good

'As for this man, I feel his memory is exceptionally good.' Xu and Langendoen (1985, ex.75h)

Example (214) namely shows that Aboutness topics can co-occur with both Left-Dislocated Topic and Hanging ones.

(214) AT co-occurs with LD and HT

a. Aboutness Topic > Hanging Topic:
家人，妈妈，他昨天看见她了，爸爸还没看见了。

Jiārén, māma, tā zuótiān kànjiàn tā le, bàba hái méi kànjiàn le.
family-people mother he yesterday saw her MOD., father yet not saw MOD.

'As for family, (his) mother, he saw her yesterday, (his) father, (he) hasn't seen (him) yet.

b. Aboutness Topic > Left-Dislocation:
家人，为妈妈，颖艺已经借到很多钱了。

Jiārén, wèi māma, Yǐngyì yǐjīng jiēdào hěnduō qián le.
family-people for mother Isabel already borrow very-much money PART.

'As for family, for her mother, Isabel already borrowed a lot of money.

To conclude, given the characteristics that are illustrated in (213) and (214), Badan (2007) established the contrastive set of syntactic diagnostics tests that are reported in Table 3.15.

Table 3.15 – Syntactic Tests for Hanging vs Aboutness Topics.

Hanging Topic (HT) c4	**Aboutness Topic (AT) c2 and c3**
1- only be a DP always requires a resupmtive clitic expressing the type of argument no case agreement only number and gender	1- a DP
2- Resumed by epithet and tonic pronoun	2a- No *trace* or resumptive pronoun. No agreement in Case with the element in a part-whole relation 2b- They are not even subcategorized for by the verb they have a part-whole, set-subset or possesor-possessee relation with an element in the comment
3- Multiple Hanging topics are not possible	3- Multiple aboutness topics are possible
4- Can co-occur with Left Dislocation	4- Can occur both with HT and LD.

3.4.4.2 Resumptives in Chinese: why and when are they needed

In Mandarin topicalized structures, the presence of the resumptive pronoun within the Comment-clause is optional, differently from what happens in Romance languages, in French where we saw clitics to be overtly expressed, compare (203) French example and (215).

(215) gap versus resumptive pronoun: two syntatic derivations

a. 张三啊今天下午我看 (他) 见。

Zhāngsān$_i$ a, jīntiān xiàwǔ wǒ kànjiàn e$_i$ le.
Zhangsan TOP. today afternoon 1SG. saw e$_i$ PART.

'As for Zhangsan, I saw [him] this afternoon.'

b. 张三啊今天下午我看他了。

Zhāngsān$_i$ a, jīntiān xiàwǔ wǒ kànjiàn tā$_i$ le.
Zhangsan TOP. today afternoon 1SG. saw 3SG. PART.

'As for Zhangsan, I saw him this afternoon.'

A third type of syntactic element for the dependency-link between Topic and Comment: Null pronominals *pro*

While analyzing the derivation aspect of different Topic-Comment configuration structures (§3.4.2) we observed the unexpected agrammaticality of sentences violating the Complex NP Constraint. Specifically, the ill-formedness of sentence (180), reproduced for convenience in (216), lies in sharp contrast with the acceptability of (217), where the apparent extraction of the Topic element from a Complex NP should structurally actually show a violation of the Complex NP Constraint.

(216) Complex NP Constraint in Mandarin:
李四，我很喜欢唱歌的声音

* Lǐsì$_i$, wǒ hěn xǐhuan [[[**t$_i$ chàng gē] de] shēngyīn].**
Lisi I very like sing song DE voice

* 'As for Lisi$_i$, I like the voice with which ___t$_i$ sings.' from Huang, Li and Li (2009:210).

(217) pro: 李四，唱歌的声音很好听

Lǐsì$_i$, [[[t$_i$ chàng gē] de] shēngyīn] hěn hǎotīng.
Lisi sing song DE voice very nice-to-hear

'As for Lisi$_i$, the voice with which t$_i$ sings is very nice'. from Huang, Li and Li (2009:210).

This apparently surprising fact, is explained by analysis drawn about the syntactic movement derivation of different Topics in section §3.4.2 -Topic structures are derived by movement when a gap occurs and only then- are explained by the fact the empty/silent syntactic element found in the comment is a *null pronominal* (pro), which allows the empty category to be co-referential with its closest antecedent, the Topic *Li Si*.

On these grounds Li (2000) states that in Mandarin Topic structures where movement occurs and the gap is a *trace*, where movement does not take place and the gap is a pro or there is no gap as we saw in Gapless topics Hanging topics.

A further confirmation, using the Subjaceny Condition as a diagnostic in (217), the movement analysis is discarded, and, sentences like (218) are not movement-derived, but feature a particurlar type of covert *null pronominal* – a (*pro* [-anphoric] [+pronominal]) – as we saw in chapter 2 (cf. §2.2.4.4, Table 2.37 §2.4.3).

(218) pro: 李四，唱歌的声音很好听

Lǐsì$_i$, [[[**pro**$_i$ chàng gē] de] shēngyīn] hěn hǎotīng.
Lisi sing song DE voice very nice

'As for Lisi$_i$, the voice with which [**pro**$_i$] sings is very nice.' from Huang, Li and Li (2009).

In short, the properties of a *pro* are the following, it is co-indexed with a nominal element like any overt pronoun, and the binding relation between the Topic and the *pro* is an anaphoric one not an antecedent-gap one like in a movement analysis[158]. Given the interesting syntactic characteristics this type of null syntactic element features, it will be retained to be inserted in our fMRI design on syntactic complexity cerebral representation in Chinese (chapt. 7).

Syntactic derivation of Resumptive Pronouns, gaps and pros in the Comment-clause

To summarize, as attested in the following examples (219 - 221), Left-Dislocated structures with a gap give rise to Island and Crossover effects[159], while interestingly LD structures with RP do not give rise to any of these effects. This has been analyzed as evidence supporting the fact that these two sentences structures are derivationally distinct from one another, ad lately documented in Pan (2015). As illustrated in (219), and like in our experimental conditions (c5) (Table 3.16), when a trace is present we observe island effects:

(219) * 那位法国影星$_j$, 我碰到了 [小明认识 [拥抱过 t$_j$]] 的那位女同学。

158. A further property of a *pro* is illustrated in the following. Consider (1) the following examples from Huang, (1989:187). It illustrate a different pattern of referential assignment for the two different empty syntactic element under analysis: According to the comment of Xu and Langendoen (1985), the null pronoun *pro* can be interpreted pragmatically by referring to a salient entity in the context, a discourse referent or topic.

(1) Referentiality in pro versus Gap

a. Null pronominal *pro*:
张三说很喜欢李四

Zhāngsān$_i$ shuō [e$_{i/k}$ hěn xǐhuan Lǐsì$_j$].
Zhangsan say very like Lisi

Interp. 1: 'Zhangsan$_i$ said that [he$_i$] liked Lisi.
Interpretation 2: 'Zhangsan said that [Mary] liked Lisi.'

b. Object gap:
张三说李四很喜欢

Zhāngsān$_i$ shuō [Lǐsì$_j$ hěn xǐhuan e$_k$].
Zhangsan say Lisi very like

'Zhang San said that Lisi liked [Mary$_k$].'

159. See Glossary.

Nà-wèi fǎguó yīngxīng$_j$, wǒ pèngdào-le [Xiǎomíng rènshi [yōngbào-guò t$_j$]]
that-CL. French star I meet-PERF. Xiaoming know embrace-EXP 3MSG.
de nà-wèi nǚtóngxué.
DE that-CL. female.student

'(As for) that French star$_j$, I met the school-girl that [Xiaoming knows ___[who embraced t$_j$]].' (adapted from Victor Pan, p.c.)

While in the case of Resumptive Pronoun, like in our condition (c4), no island effects is observed, as in (220).

(220) 那位法国影星$_j$，我碰到了小明认识 [拥抱过他$_j$] 的那位女同学。

Nà-wèi fǎguó yīngxīng$_j$, wǒ pèngdào-le [Xiǎomíng rènshi [yōngbào-guò tā$_j$]] de
that-CL. French star I meet-PERF. Xiaoming know embrace-EXP MSG. DE
nà-wèi nǚtóngxué.
that-CL. female.student

'As for that French star$_j$, I met the girl that [Xiaoming knows ___[who embraced (him$_j$)]].' (adapted from Victor Pan, p.c.)

Importantly, in the following case (221), where an Island effect is expected but the sentence remains grammatical. We can therefore conclude that the empty element inside the island is null pronoun, a *pro*.

(221) 那位法国影星$_j$，[[唱歌的] 声音] 很好听。

Nà-wèi fǎguó yīngxīng$_j$, [[e$_i$ chàng gē de] shēngyīn] hěn hǎotīng.
that-CL. French star, sing song DE voice very good

'That French star$_i$, the voice with which (he$_i$) sings is very good.' (Adapted from Huang, Li and Li, 2009: 210)

c4	Topic + resumptive	Topic object	comma	Subj.	adv. + VP + **Resumpt.**	complement 补语	**Overt Resumptive Pronoun**
Res pron B-G		(As for) Mister Wu	,	Mary	already knows **him**	for a longtime.	**Base-Generated Topic + RP**
		吴先生i	,	玛丽	已经认识 **他i**	很久。	Hanging Topic
c5	LD Topic + optional **gap** OSV	Topic object	comma	Subj.	adv. + VP + ***trace***	complement 补语	**Optional Resumptive Pronoun**
A-bar mov		As for Mister Wu	,	Mary	already knows **__ei**	for a longtime.	**Left-Dislocated Topic Abar Movement**
		吴先生i	,	玛丽	已经认识了 **__ei**	很久。	Island test
c6	LD Topic + **pro** OSV	Topic object	comma	Subj.	adv. + VP + **pro**	complement 补语	**Topicalization targeting CP**
A bar mov		That booki	,	Mary	already read **proi**	once.	**Left-Dislocated Topic Abar Movement**
		那本书i	,	玛丽	已经看过 **proi**	一次。	Resumption is agrammatical

Table 3.16 – The tree types of dependecy links in our fMRI exepriment in chapter 7. When there is no island, a gap and an RP can be in free alternation in Left-Dislocated structures.

Hence, we can conclude that in Mandarin Chinese, when there is no island, a gap and an RP can be in free alternation in LD structures, as we saw in (220) and (221) and we can see in our experimental conditions (c5) and (c4) in Table 3.16. As argued through

the above diagnostic test and since Huang (1987), the gap in sentences like condition (c6) is a resumption that remains phonetically covert.

This configuration where the dependency-link can be achieved by overt or non-overt linguistic means preserving the same surface word-order is particularly suitable for our experimental approach to the cerebral encoding of the sentence-unit. It allows to tackle the broader question of dependency-links inside the sentence that has undergone syntactic movement transformations.

From what is attested in the above formal analysis, the very minimal contrast between (c4) condition and (c5) will allow to observe the difference between a co-reference link achieved by means of an overt Resumptive Pronoun (c4) or by a non-overt linguistic mean (c5), a trace : Base-Generated Topic with resumptive [OiSV + Resumptivei] vs. Moved Topic (A-Bar) corresponding to the structure [TopicObjectiSV + tracei].

Finally, in these examples, we see once again that movement is not the only possible derivation of Topic-Comment sentence in Chinese. We can say that the contrast in grammaticality between (216) and (217) further attest that movement-analysis holds only when a gap in found in the Comment-clause.

In conclusion, to offer to this overview a diachronic perspective, we would like to share with the reader a relatively off-topic glimpse on some linguistics facts, that can nonetheless show some interesting comparisons with contemporary Mandarin Chinese resumption patterns along the overt/covert opposition. In (222), we can note that the syntactic properties and trends we are analyzing synchronically do show some variation if considered in a diachronic perspective. For instance, diachronic studies on French resumption strategy in Topic-Comment articulations (Troberg, 2004) show that Old French (OF) and Middle French allowed null resumptive pronouns in the matrix clause of a Topic-Comment structures, while Modern French (ModF) does not permit them.

Compare respectively (222a) and (b), and (c) with (d):

(222) Old and Middle French resumption patterns

a. Vostre terre$_i$ qui defandra ___$_i$? (OF)
your land$_i$ who will.defend ___$_i$
'Who will defend your land?'
from Arteaga (1997), 2: Chanson de Lyon 1617 AD)

b. Votre terre$_i$, qui *(la)$_i$ défendra ? (ModF)
your land$_i$ who (it)$_i$ will.defend
'Who will defend your land?'
from Härmä (1993), 772: La Prison 3710)

c. Car les letres$_i$ que li messages aportoit, c'estoit mes usages de
because the letters$_i$ that the messenger used-to-bring it-was my custom for
regarder ___$_i$ avant toute oevre. (MidF)
to-look ___$_i$ before all work

d. Car les lettres$_i$ que le messager apportait, c'était mon habitude de
because the lettersi that the messenger used.to.bring it-was my custom for
*(les)$_i$ regarder avant tout autre travail.(ModF)
(them)$_i$ to.look before all other work
'Because it was my custom to look at the letters that the messenger would bring before all other work'

3.4.4.3 Resumption and Topic NPs Semantic features

> *"Animacy, or the distinction between animate and inanimate entities, is so pervasive in the grammars of human languages that it tends to be taken for granted and become invisible"*
>
> Dahl and Fraurud (1996:47)

The above epigraph directly introduces us to a semantic feature determining the overt versus covert realization of the dependency-link between Topic and Comment in Mandarin Chinese. It goes without saying that Chinese is not an isolated linguistic case where animacy plays an important role in Grammar. Animacy feature is namely well known to affect the encoding of arguments cross-linguistically[160], and it is actually often observed that some semantic features of the Topic element determine the obligatoriness of resumption.

Resumption, object-drop and Animacy If we take, for example, the case of French object drop patterns, when non-referential direct objects and quantified expressions (lexically restricted) occur as Topics, they generally do not require resumptive pronouns in the original extraction syntactic position. When these lexico-semantic requirements are not met, we can find a limited class of verbs like *aimer* 'to love', *adorer* 'to adore', *connaître* 'to know', *acheter* 'to buy' that permit Topic-Comment structures without a resumptive pronoun, as illustrated by the examples below:

(223) a. Ce film, j'adore ___.
'This film, I adore ___.'

b. (*?) Ce film, qui adore ___?
'This film, who adores ___?'

c. Le caviar, j'aime ___!
'Caviar, I like ___!'

d. Le caviar, j'aime ça.
'Caviar, I like that.'

Fonagy (1985)[161] identifies a similar configuration with the verb *prendre* 'to take', as illustrated in example (224) and we can add that a growing trend in actual French is to allow preposition stranding in topicalized sentence like (224)

(224) a. La bleue, je prends ___.
the blue I take.3SG.
'The blue one, I'm taking'

160. For an overview of how animacy affect argument structure or the encoding of arguments across languages, see de Swart et al. (2008) in *Lingua*.
161. I have the pleasure to quote here the first linguistics book I ever read under the auspicious advice of the providential figure that initiated me to linguistics when I was 14 years old, Eddo to which this PhD is dedicated.

b. *? La bleue, qui prend _ ?
the blue who take.3SG.
'The blue one, who is taking?'(Fonagy 1985,9)

However, if a wh-element is introduced in (224b), the structure is considerably less acceptable, which is also observe in the above example (223b) with the verb 'to adore'.

As for animacy, we noted that Oral French is more permissive with inanimate object drop than with animate objects and along this animate-inanimate opposition it also allows preposition stranding as show by the following examples:

(225) a. [Ma trousse de toilette]$_i$, je voyage toujours avec ___$_i$.
My case of cleaning, 1SG. travel always with ___$_i$
My beauty-case i, I always travel with ___i.'

b. * [Ma trousse de toilette]$_i$, je voyage toujours avec elle$_i$.
My case of cleaning, 1SG. travel always with her
My beauty-case, I always travel with her.'

c. * Mon assistant, je voyage toujours avec ___.
My assistant, 1SG. travel always with ___
My assistant, I always travel with ___.'

d. Mon assistant, je voyage toujours avec lui.
My assistant, 1SG. travel always with ObjectM(clitic)
My assistant, I always travel with him.'

While (225a) is fully well-formed, the inanimate resumtion option in (225b) is illicit. If we change the Topic with an animate DP the opposite resumption pattern is observed. To this we can add that Oral french allows to have preposition stranding also in cases with null sentence Topics, where the Topic is not explicitly present in the sentence and the referential assignment of the gap after the preposition is done with a Discourse Topic as shown in Figure 3.23.

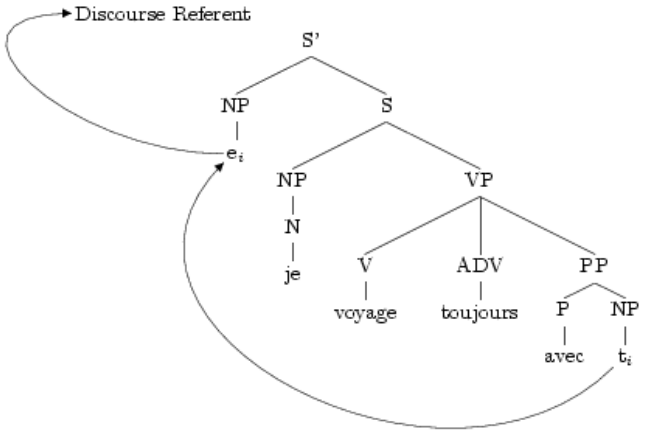


Figure 3.23 – An example of preposition stranding in Oral French having a null operator after the Preposition.

What about Mandarin? Mandarin offers similar resumption configurations where animacy gives rise to a [+/-overt] realization of the object constituent in the Comment-clause, preserving the same surface word-order[162].

Examples in (226) show an animate-inanimate contrast in the case of Topic extraction from an object relative. Comparing (a) and (b), the only way to render the inanimate topic sentence licit is to give an overt expression to the null pronoun inside the object relative as shown in (c). Namely, once a resumptive pronoun is present, Island conditions become irrelevant.

(226) Animate/inanimate Topics and resumption

a. 那个家伙$_i$，[我认识很多 [e$_j$ 不喜欢他$_i$ 的人$_j$]] 。[+animate] Topic

Nèi-ge jiāhuo$_i$, [wǒ rènshi hěnduō [e$_j$ bù xǐhuan tā$_i$ de rén$_j$]].
that-CL. fellow I know very-many not like him DE person

'That fellow, I know many people who don't like him.'

b. * 那本书$_i$，[我认识很多 [e$_j$ 看不懂它$_i$ 的人$_j$]] [-animate] Topic

Nèi-běn shū$_i$, [wǒ rènshi hěnduō [e$_j$ kàn-bù-dǒng tā$_i$ de rén$_j$]].
that-CL. book I know very-many cannot-understand it DE person

'That book, I know many people who can't read it.'

c. 那本书$_i$，[我认识很多 [e$_j$ 看不懂 e$_i$ 的人$_j$]] [-animate] Topic

Nèi-běn shū$_i$, [wǒ rènshi hěnduō [e$_j$ kàn-bù-dǒng e$_i$ de rén$_j$]].
that-CL. book I know very-many cannot-understand DE person

'That book, I know many people who can't read [it].'

On the contrary, examples by Xu (2007), in (227) and (228) show that animate NPs can be antecedent of overt pronouns in Chinese. The author also notes that there is no compelling rule for resumptive pronouns of [+animate] Topics to take a zero form.

(227) Animate resumptive and null pronoun

a. resumptive: 这个顽皮的孩子$_i$[我找不到 [e$_j$ 愿意收养他$_i$ 的人$_j$]]。

Zhège wánpí de háizi$_i$ [wǒ zhǎo-bu-dào [e$_j$ yuànyì shōuyǎng tā$_i$ de rén$_j$]].
this naughty DE child I can't.find willing adopt him DE person

'This naughty child, I can't find a person who is willing to adopt him.'

b. null pronoun: 这个顽皮的孩子$_i$[我找不到 [e$_j$ 愿意收养 e$_i$ 的人$_j$]]。

Zhège wánpí de háizi$_i$ [wǒ zhǎo-bu-dào [e$_j$ yuànyì shōuyǎng e$_i$ de rén$_j$]].
this naughty DE child I can't.find willing adopt DE person

'This naughty child, I can't find a person who is willing to adopt him.' (Xu, 2007)

(228) Animate resumptive and null pronoun

a. resumptive: 这位候选人$_i$[我们在各地见到了无数 [e$_j$ 拥护他$_i$ 的人$_j$]]。

Zhèwèi hòuxuǎnrén$_i$ [wǒmen zài gè dì jiàn-dào-le wúshù [e$_j$ yōnghù tā$_i$ de rén$_j$]].
this candidate we PREP. every place see-ASP. numerous support him DE person

162. For a discussion on the choice between overt and covert, see Xu (2003).

'This candidate, we saw numerous people who supported him everywhere.' (Xu, 2007)

b. null pronoun: 这位候选人$_i$[我们在各地见到了无数 [e$_j$ 拥护 e$_i$ 的人$_j$]]。

Zhèwèi hòuxuǎnrén$_i$ [wǒmen zài gè dì jiàn-dào-le wúshù [e$_j$ yōnghù e$_i$
this candidate we PREP. every place see-ASP. numerous support
de rén$_j$]].
DE person

'This candidate, we saw numerous people who supported him everywhere.' (Xu, 2007)

Given these linguistic considerations, we decided to compare the internal neural of the optionality and the obligatory null pronominal forms that the animate/inanimate lexical contrast imposes on the overt and covert realization of anaphoras in Mandarin Topics. Table 3.16 p.392 shows the experimental conditions we selected for our fMRI neuro-linguistic investigation in chapter 7

Animacy issues in Resumption: Brazilian Portuguese As for animacy issues in resumption, an interesting parallel can be drown with Brazilian Portuguese, that exhibits null objects in varied syntactic contexts, as opposed to European Portuguese.

Bianchi and Figueiredo (1994)[163] found that Brazilian Portuguese has separate animate and inanimate paradigms for object drop as shown in the examples below where the interpretation of the sentence changes according to presence of the null pronoun in (a) and of the use of the strong pronoun in (b).

(229) null objects and the [+/- animate] semantic feature

a. Gap for inanimate interpretation

Eu nunca vejo o meu pai. Não me lembro d[a cara
I never see-pres-1SG. the my father. Not CL-reflexive remember of[the face
dele]$_i$. Vou esquecer ___$_i$
of-he]$_i$. Go-pres-1SG. forget-INF. ___$_i$

'I never see my father. I don't remember [his face]. I'm going to forget (it).'

b. strong pronoun for animate interpretation

Eu nunca vejo [o meu pai]$_i$. Não me lembro da cara
I never see-pres-1SG. [the my father]$_i$ not CL-reflexive remember of-the face
dele. Vou esquecer ele$_i$
of-he. Go-pres-1SG. forget-INF. him$_i$.

'I never see my father. I don't remember his face. I'm going to forget him.' from Cyrino (2003)

Example (229) illustrates that the presence or absence of an overt object leads to different co-indexation patterns: in (a) the object of 'forget' is interpreted as the [-animate] 'his face', while in (b) the use of a strong pronoun leads to interpret the object as the [+ animate] 'my father'.

163. Bianchi, V. and M. C. Figueiredo Silva. 1994. On some properies of agreement-object in Italian and in Brazilian Portuguese. In M. L. Mazzola (ed), Issues and theory in Romance Linguistics: selected papers from the Linguistic Symposium on Romance languages XXIII. Georgetown University Press.

Several linguistic analyses have advanced that the null objects in the Brazilian Portuguese example above (229) are empty pronoun, *pro*, because they show the particular requirements of identification and licensing of this empty category[164]. Specifically, as for the identification requirement on *pro* (Rizzi, 1986), the *pro* in (229) receives its grammatical specification from the element (i.e. the licensing head) it is co-indexed with: in (a) the [- animate] 'his face' and in (b) the [+ animate] 'the my father'..

To capture some of the syntactic aspects of the animacy constraints illustrated above, we can compare these cases with European Portuguese, where the availability of discourse-bound null objects has been attested, and Raposo (1986) indeed showed that null objects are ruled out in a strong island context like in (230).

(230) European Portuguese

a. A: E a Maria?
What about Maria?

b. Answer

B: O Pedro está triste porque o Zé *(a) beijou.
the Pedro is sad because the Zé herCL. kissed

'Pedro is sad because the Zé kissed her.'

If we compare the two kinds of Portuguese, we see that, in Brazilian Portuguese, the animate paradigm obeys islands like Object Drop in European Portuguese (see [230]), while inanimate paradigm is insensitive to Islands as illustrated in (231), which according to Bianchi and Figaredo (1994) would make a null clitics analysis appropriate for the inanimate paradigm only.

(231) ungrammatical in European Portugese, but grammatical in Brazilian Portugese

a. Compre i o casaco depois que experimente i ___.
Buy-past-1SG. the coat after that try.on-past-1SG. ___
'I bought the coat, after I tried [it] on.'

b. O rapaz que trouxe ___agora mesmo pastelaria era o teu afilhado.
The boy that bring-past-3SG. ___now just pastry-shop was the your godson
'The boy that brought [it] just now from the pastry shop was your godson.'

Given these patterns, the generalization for Brazilian Portuguese is the following: it is possible to have a null object both if it is animate or inanimate, but in the animate interpretation the null pronoun is binded by a null operator -and therefore sensitive to island constraints- while in the inanimate interpretation it is not the case (Bianchi, p.c.)[165].

All in all these cross-linguistic patters linked to an animacy constraint on the overt vs covert syntactic realization of co-referential pronouns show that the Chinese pattern of resumption is not an isolated case. Thus, given the syntactic theoretical considerations exposed in our analysis of 'Topic anphoras it will be interesting in chapter 7 to

164. Following Rizzi (1986), Bianchi and Figueiredo (1994) assume that pro is subject to a formal licensing requirement and identification requirement. While the first is given a complex explanation liked to the agreement system, the second (i.e. the identification one) can be easily understood if we consider the above examples in (229).

165. More recent work on Brazilian Portuguese confirms with greater detail these patterns and a comprehensive review on the null objects, see Cyrino (1997), Cyrino and Reich (2001) both written in Portuguese, and Cyrino (2003) in English.

contrast the cerebral activations elicited by the presence of a resumptive or a gap in the two syntactic configuration offered by Mandarin Chinese, where [+/-animate] feature determines resumption patterns.

3.4.4.4 Complexity effects linked to movement and *features*: Relativized Minimality

Psycho-linguistics of Chinese topics

The derivational difference that was proposed for Mandarin Topic-Comment constructions has been the object of a recent psycho-linguistic study on Chinese children acquisition of Topic-comment articulations (Hu Shen'ai, PhD, 2015). In order to gain some insight on the psycho-linguistic reality of the difference between Base-generated Topics and movement-derived ones, we will present its main results and introduce the approach to syntactic-movement this study was based on: Relativivized Minimality.

Notably, considering the featural specification of Topicalized elements is not only relevant to characterize the different patterns of resumption it yields (cf. previous section), but also to understand an additional complexity parameter linked to syntactic-movement linguistic phenomena[166]. We will namely consider Relativivized Minimality framework (RM) (Rizzi 1990 and 2001), because it can offer a principled account for some movement-related complexity effects observed in the linguistic behavior of certain populations.

Relativivized Minimality

According to a probe-goal approach to syntactic-movement, the movement of an element is triggered by attractors, that are endowed with certain features to attract a syntactic element sharing those features[167].

In this framework, movement is allowed only when some specific configurations of the features that characterize (a) the moved item and (b) the other possible items found in between the extraction-site and the landing-site (the so-called *interveners*) are gathered. Furthermore, according to Relativivized Minimality principle of economy of syntactic representation, syntactic relations are restricted to the closest possible element capable of bearing that relation.

More formally, this theory posits that syntactic relations have to be satisfied in the smallest possible environment in witch they can be satisfied. Thus, the following formal definition specifies (Rizzi, 2004a) that given the minimal configuration: ... X ... Z ... Y ... :

(232) Y is in minimal configuration with X if there is no Z such that :
  i Z is of the same structural type as X, and
  ii Z intervenes between X and Y.

Notably, the intervening elements Z is defined as having the same structural type as X in terms of identity of morpho-syntactic features. Hence, the fundamental idea of applying these locality principles to understand linguistic behavior is that the *representation* of the full array of morpho-syntactic features is needed in order to distinguish, for

166. This further analysis of movement-based of syntactic phenomena, and of their complexity parameters, leads us to deepen our understanding of the constraints on movement presented above in previous chapter 2 (§2.4.2).
167. In chapter 2 we already addressed similar issues concerning French question formation, (§2.4.4.2).

example, a moved-object from the intervening-subject in case of Topicalization as will be illustrated in Chinese.

This linguistic theory (Rizzi, 1990/2001/2004; Starke, 2001), formalized within the generative approach, was successfully reinterpreted in the realm of agammatic studies by Friedmann and Shapiro (2003:295), and was then further specified by Grillo's (2005) study on agrammatic behavior. It was namely proposed that the source of interpretative difficulty of object dependencies, like object relatives, or more generally non-local dependencies, is related to the intervention of a noun phrase between the displaced constituent and its original position (Grillo, 2009). Thereby, it already provided insightful explanation to agrammatic sentence comprehension patterns in Broca's aphasics.

Specifically, Grillo (2003, 2005) proposed that a loss of syntactic processing abilities can compromise the representation of the full array of morpho-syntactic features that are normally associated with syntactic elements in a sentence. This impaired representation would give rise to *Minimality Effects* in precisely definable syntactic configurations, where the feature structure of the NPs involved in movement are not distinguishable[168].

Relativized Minimality configurations and complexity effects in different populations

In the perspective of interpreting the locality principles expressed by the RM approach in terms of features (Rizzi, 2004 and Starke, 2001), a more recent refinement of this approach to linguistic behavior showed that a particular set of syntactic complexity effects can actually be captured.

Figure 3.24 illustrates that the configuration introduced above is a particular case of the possible configurations of abstract morpho-syntactic features triggering the syntactic movement of an item Z : in (i) the local relation between X and Y cannot be established if Z, has the same feature as Y, and acts therefore as a potential candidate for the same relation, but in (ii) when the Z element, intervening in-between, only shares a subset of the features of Y, the relation can be established and this configuration yields a behavioral complexity effects only in certain populations.

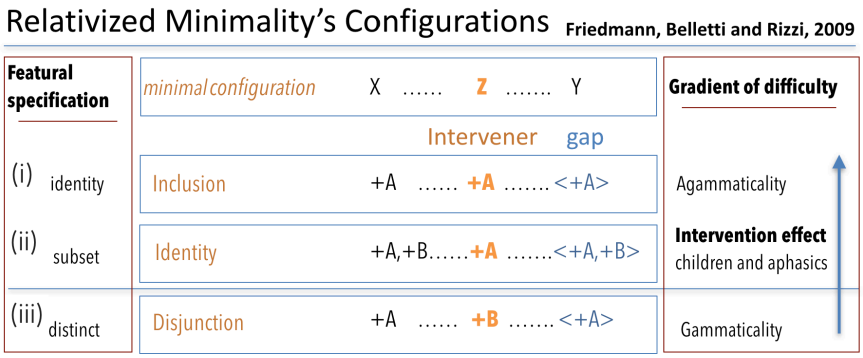


Figure 3.24 – A and B stand for abstract abstract morpho-syntactic features triggering the syntactic movement of an item Z. Adapted from Fiedmann et al. (2009).

168. We can note that the approach developed by Grillo PhD and following publications (2005, 2008, 2009) has a interesting theoretical advantage of trying to unify processing and representational accounts.

Friedmann et al. (2009) showed that children (22 Hebrew speaking children) have troubles computing the subset relation present in (ii) configuration (i.e. the so-called *intervention effect*), as they have greater trouble in understanding and producing the kind of syntactic constructions, such as object-relative clauses and object wh-questions, that involve such featureal configurations[169].

According to the authors, this difficulty is related to the limitations in the operative syntactic memory, as we can read in the following: "disjointness is easier to determine, as it can be calculated feature by feature, whereas calculating a *subset-superset relation* requires holding in operative memory and comparing the whole featural specifications associated to different positions, an operation which may exceed the capacity of the early systems [...]." (Friedmann, Belletti and Rizzi, 2009:84, emphasis mine).

Given this background, Hu (2015) formulated different predictions with regard to the acquisition of Topic constructions by Chinese children: Movement derived Topic constructions should show Intervention effects, while the one where a non-movement analysis is established should imply the absence of Intervention effects in child grammar.

Concretely, applying Relativized Minimality analysis to the Topic structures in (233) gives the configuration in (b), where the Determiner and Noun constituting the silent object (i.e. <D+NP>) are attracted by a composite attracter that present both Topic [+TOP] feature and [+NP] lexical specification features.

(233) Hu (2015)

a. 这个孩子$_i$，外婆在画 t$_i$

Zhè-ge háizi$_i$, wàipó zài huà t$_i$.
this-CL. child grandma PROG. draw

'As for this child, the grandma is drawing [him].'

b. feature configuration of (a)

D+NP ..... D+NP ..... <D+NP>
[+TOP,+NP] ..... [+NP] ..... < >

Topic ..... Subject ..... gap

c. Topic-Subject: 这个孩子 (呀)，在画外婆

169. Specifically, the accuracy rates of object RCs with gaps and those with resumptive pronouns were respectively 55% and 56%, while those of subject RCs reached 90%.

(1) intervention effects in Hebrew children

a. gapped object RC

Tare li et ha-pil she-ha-arie martiv.
show to-me ACC the-elephant that-the-lion wets

'Show me the elephant that the lion is wetting.'

b. resumptive - object RC

Tare li et ha-kof she-ha-yeled mexabek oto.
show to-me ACC the-monkey that-the-boy hugs him

'Show me the monkey that the boy is hugging.'

Zhège háizi$_i$ (ya), PRO$_i$ zài huà wàipó.
this-CL. child (TOP.) PROG. draw grandma

'As for this child, [he] is drawing the grandma.'

The Topic and the Subject of the sentence in (233) share some features defining them, they are in a configuration where a subset of the features of the moved item are shared by the intervening one. The object of the Comment-clause *crossed over the similarly lexically restricted position* of the subject of the Comment. This is precisely the configuration that was reported to be either more difficult or inaccessible to certain populations, like children and aphasics.

As presented above, this configuration implies the following step in order to interpret the subject as being distinct from the Topic: children and adults have to compute *a subset relation* that is more demanding in terms of computation resources. And, when these resources are limited (aphasics and children) the processing cost of objects' extraction can be too high for children or agrammatic aphasics, which in turn leads to a comprehension difficulty or failure of the grammatical sentence.

The experimental data on Mandarin reported by Hu (2015) appear to be more complex than predicted: although the accuracy rates of OSV topicalization sentences were numerically lower than SVO topicalization sentences (e.g., at age three, 88% vs. 76%; at age four 89% vs. 84%), no difference between the two structures reached significance. All in all, these findings confirm intervention effect in an A-bar movement construction such as Relative clauses, but not in subject and object Topicalized constructions.

The intervention configuration in Relative clauses affects reaction times in comprehension tasks for children from 3 to 5 years old. But, when examining the comprehension of Topic-comment structures in young children in Table 3.17, cases of topicalization of objects or subjects are reported to be at ceiling in whatever configuration: [Topic$_{object}$ S V] order and [Topic$_{subject}$ V O] order are equally good in the performance of 5 years old Chinese children[170].

Groups	OSV topicalization sentences				SVO topicalization sentences			
	%	N	M	SD	%	N	M	SD
3 y.o.	76	121/160	6.05	1.50	88	141/160	7.05	0.89
4 y.o.	84	134/160	6. 70	1.26	89	142/160	7.10	1.02
5 y.o.	93	149/160	7.45	0.10	96	154/160	7.70	0.57
6 y.o.	98	156/160	7.80	0.52	100	160/160	8.00	0.00
Adult	100	80/80	8.00	0.00	100	80/80	8.00	0.00

Table 3.17 – Adapted from Hu Shen'ai et al. (2015).

The author interpret this result as showing that children have troubles only with structures where an Abar-moved constituent is in a position where an intervener with relevant feature is closer to the target landing site than the low element of the dependency,

170. As a side note we can follow the author by observing that the superficial word-order of subject-topicalized structure could actually be understood as assigning the same position to the Topic and the subject, while in object-topicalized structures the Topic is identified regardless of the presence of the topic marker because of the [NP + NP + V] word order.

which is the case in Relative Clauses. We understand the equal difficulty of these two as being related to Topic-Prominence in Chinese as argued in section §3.2.4.3 (p.323).

We can put in parallel the trend observed in Chinese acquisition of object and subject Topics with previous studies on the acquisition of topicalization across languages (e.g. French, German and Japanese). The picture that emerged from these studies presents an interesting pattern. Object-topicalization was reported as having lower scores than SVO structures (both in Hebrew and in German). While this asymmetry between OSV and SVO constructions can be explained in the realm of Relativized Minimality, children fail to establish that topicalized object and subject are distinct, it was further observed that the comprehension scores for object-topicalization were nonetheless higher than for object RCs (respectively 90% and 70% in Friedmann et al. 2009). .

Although it is difficult to conclude that a possible gradual pattern of complexity among these constructions is observed, as only a trend was shown for Mandarin Chinese, we can thus only say that the kind of configuration that object-topicalization features, where a subject is a potentiel intervener in the A'-chain that is connecting the moved Topic to its *copy* is easier to calculate compared to the same intervention configuration in Relative Clause.

Given these theoretical considerations and their initial psycho-linguistic grounding, it will be interesting in chapter 7 to compare the cerebral response patterns related to the presence of a resumptive or a gap in Mandarin moved and *in-situ* Topic objects, and in the particular syntactic configuration offered by Chinese, where [+/-animate] feature determines the different resumption patterns, as illustrated in previous section.

3.4.4.5 Aphasics' Topic-Comment articulation cross-linguistically

While we presented in chapter 2 (§2.4.2), neuro-imaging findings revealed substantial experimental evidence for considering topicalization as a determinant of complexity effects at the level of brain activation (e.g. Shetreet et al. 2014; Dongli et al., 2002, Ben-shachar et al., 2004; Liu and Yang, 2014; Ohta et al., 2017).

The theoretical hypothesis that base-generated Topic compared to moved ones would yield different cerebral activation patterns has found a first neuro-psychological evidence in a study by Beretta et al. (2001), on Agrammatic comprehension of scrambled sentences in Spanish and Korean (Beretta et al., 2001).

As previously shown in chapter 2, Agrammatic patients' linguistic behavior is an interesting testing ground to investigate the kind of cognitive and cerebral processes that are generally acknowledge to be on going in parsing of resumptive pronouns, clitics or gaps left by movement to the Left-Periphery (cf. §2.4.3.2 and §2.4.3.1, p.193). Crucially, the syntactic configurations respectively presenting resumptive pronouns, clitics or gaps, allow to test different [+/- overt] syntactic means for referential dependency inside the sentence-unit.

Hence, in this study, the comprehension of Spanish sentences with syntactic dislocations but also displaying clitics, was compared to that of control sentences featuring clitic doubling and no-movement analysis, as the following examples respectively show in (234) and (235):

(234) Clitic doubling with no-movement analysis

La	mujer	*la*	esta	empujando	a	la	girafa
the.3SG.f	women	*clitic.3SG.f*	is	pushing	to	the.3SG.f	giraffe

'The women is pushing the giraffe.'

Importantly, as Broca aphasic's sentence comprehension is highly sensitive to the linearization of verbal arguments for thematic-role assignment (cf. §1.4.4, p.57)[171] the the movement derived sentence selected (i.e. Clitic left-dislocation) in the study (235) preserved the relative order of arguments. In Spanish, namely, the left-dislocation of a direct object requires a resumptive pronoun in the *trace* position illustrated in bold in (235), hence the following syntactic movement analysis is acknowledged for the Clitic left-dislocation in (235): [A la girafa]$_i$ [la mujer]$_j$ t_j t_i la esta' empujando.

(235) Clitic left-dislocation: Movement

a. A la girafa la mujer **la** esta empujando.
to.DAT. the.3SGf giraffe the.3SG.f women cltic.3SG.f is pushing
'The women is pushing the giraffe.'

b. co-indexation and structure:[A la girafa]$_i$ [la mujer]$_j$ t_j t_i la esta' empujando.

The results of this experiment in Spanish show no impairment for non-movement derived clitic-doubling sentences having 'la mujer' as Topic and a resumptive clitic in Comment subject position as in (235)[172].

This sentence type and simple active control sentences were largely above chance in the tested Broca aphasics, while the other sentence types like Clitic Left-Dislocation, scrambled actives, Passives and scrambled passives were below chance. As we saw in the previous chapter 2 this patter can be understood in light of the other experiments we presented to show how the presence of a resumptive in a syntactic configuration can yield different impairments patterns and in both comprehension and production (cf. §2.4.3.2 and §2.1.1).

In conclusion, this first experimental evidence supporting our experimental hypothesis on the different cerebral substrate (or representation) of sentence with in-situ topics (i.e. Base-Generated) compared to moved ones (i.e. Left-Dislocated) encouraged us to pursue our initially theoretical and linguistic reflection. We can now move on to the experimental part of this manuscript.

3.4.5 Topic Types and their position in the Topic field

So far the discussion was oriented towards syntactic formal description of different topical phenomena in Mandarin, we now turn to consider the mapping of these Topic types based on their specific ordinal constraints and interpretative aspects (cf. Jacobs, 2001, §3.1.3.4).

As we saw in previous section, similarly to what has been found in Romance languages (see Frascarelli and Hinterhölzl, 2007), also in Chinese it is possible to distinguish between different types of Topics not only on the basis of the interpretation, but also on syntactic grounds (Badan 2007; Badan and Del Gobbo, 2010).

171. See also §2.2.2.3 about the psycho-linguistics of canonical word-order of sentences, and §2.2.2.1 cognitive linearization principles across languages.

172. The author concludes that this impairment pattern indicate that competing referential dependencies are good or preferable candidates for the locus of impairment in Broca's' aphasics. We can add, *en passant*, that this results could also be understood in the light of Relativized Minimality.

Having reviewed the behavior of Chinese Topics in the preceding sections, we were able on the basis of characteristic syntactic and interpretational properties to differentiate several Topic and Focus types and the corresponding functional projections (in unspecified order):

(236) Topic-Filed different types of functional projections in unspecified order

a. (i) a functional projection for Hanging Topics
b. (ii) a functional projection for Scene-Setting Topics
c. (iii) a functional projection for Left-Dislocations
d. (iv) a functional projection for Aboutness Topics typical of Chinese Topic Prominence
e. (v) a functional projection for Contrastive Topics
f. (vi) a functional projection for *even*-Focus

Following Badan and Del Gobbo (2010/2015), we assume that contrastive Bare focus in Chinese Mandarin does not move to the Left-periphery of the sentence, and that more generally Chinese does not allow focalization strategies in Left-Periphery. Hence we assume Badan's Low periphery mapping:

(237) Badan's Low periphery: Inf Focus > IP > inner TopicP > even-Focus > vP

The Topic Field in the Left-Periphery

Aiming at discovering how Topic-prominence claim translates into the fine structure of the LP, Badan and Del Gobbo (2010) propose that:

> "The fact that Chinese is a Topic-Prominent language seems to play a role in the sense that it seems to allow more types of Topics than Romance languages do, in particular, it allows base-generated Topics without either gaps or strictly resumptive elements ("Aboutness Relationship" ones). [...] In a sense, then, the LP of Chinese takes a heavier burden as far as Topics are concerned, but a lighter one regarding Foci." (2010:88).

Given the array of characteristic properties we reviewed, the mapping of Chinese topical facts onto the Functional structure of the sentence-unit will be experimentally tested in chapter 7. For this we chose to assume partly the work of Paul (2006), and to mainly rely on the comprehensive Functional structure of the Left-Periphery by Badan (2008) and Badan and Del Gobbo (2015), who compared Mandarin Chinese Left-periphery to that of a typologically unrelated language like Italian based on the work Benincà and Poletto's (2004) following proposal of the Italian Left Periphery:

(238) [[Hanging-Topic][[Scene-Setting Topic] [[Left-Dislocated Topic] [[List Topic] [[Contr. Focus] [[Inform. Focus]]]]]]].

It should be noted that, from a more theoretical point of view, comparing the Left-periphery and Topic Field across languages can be informative about the existence of

a kind of universal sentence structure template from which languages could select or activate some, if not all, functional projections[173].

Put together, the strong agramamticality yielded by the violation of certain of these ordering constraints and also their learnable character, brought us to hypothesize that these ordering would be somehow represented by the brain. Hence, we will adopt and test for the following Left-Periphery ordinal positions cerebral representation in the brain:

(239) Mandarin Chinese Left-Periphery mapping

a. Paul's Low periphery IP > inner TopicP > *lian even*-Focus > vP

b. Badan's Left Periphery and Low periphery : Aboutness Topic > HT > scene setting > LD > *lian* even-Focus > Inf Focus > **IP** > inner Contrastive TopicP > *lian* even-Focus > vP

c. Badan's Left-Periphery and Low periphery (revised by us) : **Scene-setting** > **Aboutness Topic** > **HT** > **LD** > *lian* even-Focus > Inf Focus > **IP** > **inner Contrastive TopicP** > ***lian*** **even-Focus** > vP

The hierarchical pointer hypothesis

Furthermore, the a one-to-one relation between position, function and the specific semantic/pragmatic interpretation of each Functional Head represent together important aspects the sentence-structure building. Thus, we hypothesized that the sentence comprehension system should in some way represent these aspects to produce and understand utterances.

In other words we hypothesized that the hierarchical organization of the linguistic encoding of these different semantic properties should be somehow cerebrally represented and be part of the neural implementation of syntax. We call this hypothesis about the cerebral implementation of the syntactic linguistic system ***the hierarchical pointer hypothesis***, which can specifically be defined in this terms:

(240) ***The hierarchical pointer hypothesis***: the hypothesis according to which there exist in the cerebral organization of syntax, a brain area playing the role of a pointer in charge for encoding the hierarchical organization of the different Functional Projections (each one encoding different semantic properties) put forward by the cartographic analysis.

The different positions in the Left-Periphery realized by the different conditions (in bold in ex. [239c]) in our experimental design, will enable us to test for the presence of a brain area that shows a linear activation according to the height in the syntactic-tree and specifically by the height in the Left-Periphery that a particular Topic type occupies int he sentence structural skeleton.

173. Yet, notice that Chinese does not seem to activate the List Interpretation Projection, we can speculate that this is due to the existence of a Contrastive Topic projection. Following this thread we can also emit the hypothesis that the relative ordering of these functional projections could not only form certain patterns across-languages, but could also show certain invariable order between Functional Heads that could cognitively be specified. Cinque's (1999) general assumptions was actually that Universal Grammar would not allow any variation among languages neither in the number and the type of the Functional Heads nor in their relative order across languages and clause types. However, the '*mother hierarchy*' being determined by cross-linguistic empirical investigation, it is thus based on order constraints between the different specialized positions of very different languages, which could lead to admit to some extent a parametric change in the ordering constraints that could be learned and an some invariable ordering guidelines that could constitute a core universal nucleus.

In conclusion given our general research direction, investigating the cerebral implementation of the hierarchical and ordinal organization of the different Topic functional projections we observe in the Chinese Left-Periphery, will tentatively bring further light on how the cerebral representation of the sentence-unit's hierarchical organization.

3.4.6 Conclusion: What is special in "Chinese Topic"

In order to recapitulate the reasons that brought us to consider Mandarin Chinese Topic-Comment articulations as an interesting testing ground to investigate the cerebral representation of sentence structure and hierarchy, we briefly summarize the different aspects that have emerged form this broad overview of topical linguistic phenomena.

1- Building sentence's articulation and hierarchy by Minimal word-order marking

Chinese Topic-Comment sentences offer the possibility to test for the *syntactic encoding the Sentence-discourse interface* in a configuration where minimal surface marking (optional morpho-syntactic marking and non-contrastive interpretation) and linear positional encoding based on word-order, together *permit to study how the brain manages this incredible equilibrium the sentence-unit can achieve between linearity and hierarchy.*

2- The representation and processes linked to the Sentence's Interface with Discourse

By bringing evidence to the syntactic encoding of the sentence-discourse interface property of Topic-comment sentences in Chinese, we observed the possibility to have two different types of 'Chinese style' Gapless Topics, featuring either an Aboutness or a Frame-setting role in the articulation of the propositional content of sentences. Contrasting these two distinct kind of Gapless Topics, we may be able to observe the neural correlates of these two different Topic functions.

As for the Topic position in the functional sentence skeleton, although the sentence-initial position has been identified as one of the fundamental properties of Topic (Li and Thompson, 1981), following Paul (2002) we assume a pre-verbal Topic Position in the SOV contrastive pattern. Given this property of Mandarin Chinese, we will be able to investigate the cerebral representation of Topic when it occupies the Left-Periphery and when it is in clause-internal position. Moreover, as stated by Badan (2008) *even*-Focus in Chinese can be found in clause-internal position, we will leverage on this low Focus position and on sentence-internal Topic position to be able to disentangle cerebral response patterns linked to complexity of representing of elements high in the sentence syntactic-tree (in the CP layer) from the rich pragmatic interpretation that an element hosted in the CP-layer can be attributed.

3- Representing Syntactic-tree complexity metrics: movement and height in the Sentence-hierarchy

Chinese also offers the opportunity to test for the cerebral representation of some complexity metrics liked to syntactic-tree representation format. The phenomena observed in Mandarin sentence-initial position are either directly generated by External-Merge (i.e. Base-generation), in the case of Aboutness Topics and Hanging Topic, or they are derived through A-bar movement (i.e. External-Merge + Internal-Merge) in the case Left-dislocated Topics. Crucially, this means that in Mandarin the Topic field in the

syntactic-tree can be occupied as the result of movement or as the result of *in-situ* base-generation.

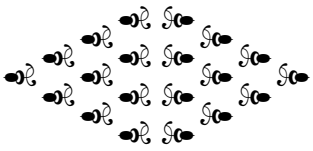
Hence, the fact that two distinct syntactic derivation for Topic-Comment articulations are possible for different types of Topics is a characteristic will allow to disentangle, at the cerebral processing level, syntactic movement-related complexity effects from those possibly linked to height of the targeted position in the syntactic-tree hierarchy (i.e. the landing-site).

4- Achieving the grammatical link between Topic and Comment though both overt and covert syntactic elements

Most interestingly, in Topic-prominent languages the so-called 'Chinese-style' Topic (Chafe, 1976) constitute a sentence construction where in the Comment-clause no constituent can be syntactically and sometimes even semantically bound by the Topic. Yet, this crucial aspect of Mandarin Aboutness and Hanging Topics allows the first element not have an explicit selectional relationship with the main verb, which permits to have Topic-comment articulations with no gaps or resumptives. Hence, in our quest for the neural underpinnings of empty and phonologically covert syntactic elements, Chinese offers a perfect configuration where -thanks to an animacy-related grammatical rule- the same Topic-comment surface structure configuration alternatively features a resumptive, a null pronoun or a gap.

Moreover, these characteristic will theoretically allow to observe in minimally differing sentences the effects of establishing a filler-gap dependency-link or pronominal reference linking a resumptive pronoun to its antecedent.

In sum, we can conclude from this brief recapitulation that focusing on Chinese and on its Left-Periphery will allow to isolate and test the distinct syntactic complexity parameters that we presented in chapter 2: (a) syntactic hierarchy, (b) syntactic movement transformations, (c) presence of resumption versus syntactic-empty positions (i.e. gaps), and finally (d) the cerebral representation of the sentence-discourse CP syntactic layer without other complexity parameters like movement or rich pragmatic interpretation.



PART

Experimental Development

Experimental approach: a change of point of view on linguistic facts

> "[L'homme] au lieu d'observer les phénomènes naturels tels qu'il lui sont naturellement donnés, place la nature dans les conditions de son entendement."
>
> *["Man, instead of observing the natural phenomena as they are naturally given, places nature in the conditions of his understanding."]*
>
> HANNAH ARENDT, *Condition de l'homme moderne, 1983, p.299.*

We are now going to face the sentence-unit, not as a finish product of enunciation and of a given linguistic system, but from the side of the locutor and his brain. This will mainly imply to change point of view on linguistic facts and embrace a new setting where the locutor's cognitive system is ***incrementally***[1] set in motion to process the sentence-unit and comprehend it.

The change in our view angle will make us depart from the description of the sentence-unit as a linguistic *fact*, and we will address how the sentence-unit is processed (chapt. 4 and 5) and how it is represented (chapter 6 and 7) through experimentation.

This experimental approach will tentatively offer a few insights into the characterization of the functional architecture of syntax in the sentence-unit cerebral network. This is namely the moment where the reader who hasn't detached yet the Big Brain Picture that is at the beginning of this manuscript, can do it and play the game of the functional mapping of the sentence brain network.

The continuous back and forth between (i) linguistic explanations of linguistic facts and (ii) psycho-linguistic, (iii) neuro-psychological and (iv) cognitive neuro-imaging explanations of linguistic behaviors and facts we have been putting into practice so far, will now take the form of concrete experimental designs. Every design will be extensively motivated at the beginning of each chapter referring back to the theoretical and linguistic chapters of Part I, in order to guarantee the cohesion of our approach in theoretical framework, preventing chorality from transforming into cacophony. Thus, according to the organization of the theatrical action of the manuscript in Figure 3.25, *Kommos* moments will thus be found at the beginning and at the end (the discussion) of each experiment.

1. One of the most remarkable aspects of language understanding is how quickly and accurately utterances are understood and produced in one's mother tongue. Comprehension process is often put in place without being aware of the exact linguistic formulations and of their complexity. To account for this, Levelt (1989:24) suggests namely that sentence production and understanding is incremental, in that "*the next processor can start working on the still-incomplete output of the current processor.*".

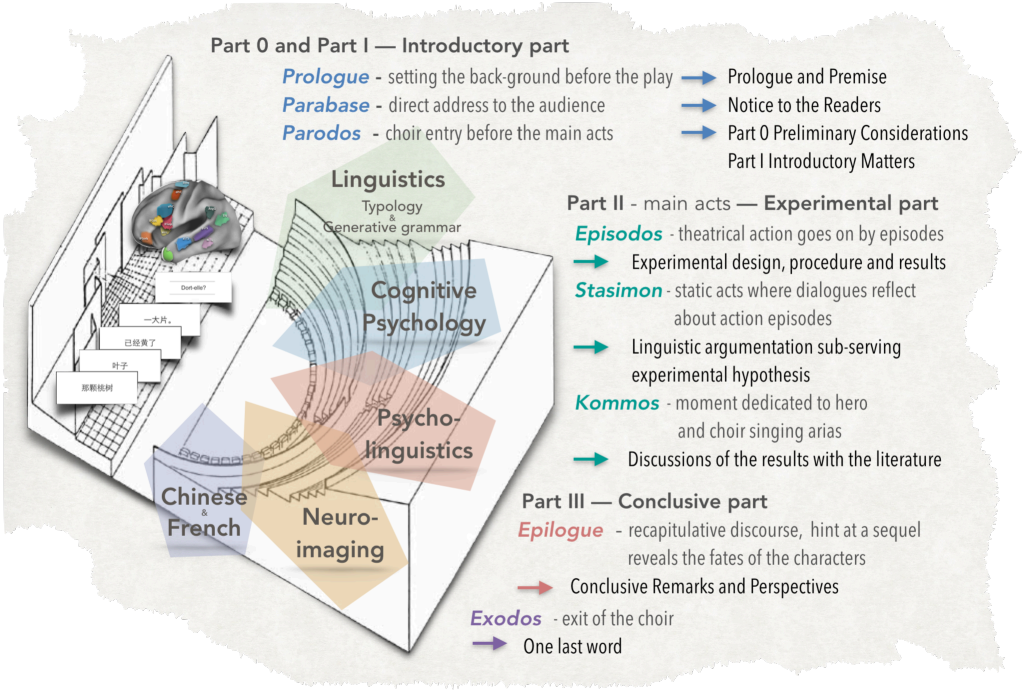


Figure 3.25 – Greek theater organization of Part II.

In conclusion, we can note that this has a great deal of epistemological consequences and delineates the role of linguistics as hypothesis-formulating-discipline. This epistemological shift has been addressed in a more comprehensive way in the Epilogue of this manuscript (§III, p. 702), in the meantime we can just rise the issue that this shift toward experimentation actually profoundly questions what it means to *"explain" a linguistic fact*.

Chapter 4

The prosodic signature of Chinese Topic

"[...] Die Sprache liegt nur in der verbundenen Rede, Grammatik und Wörterbuch sind kaum ihrem todten Gerippe vergleichbar."
"Language actually only exists in spoken discourse, its grammar and dictionary are hardly even comparable to its dead skeleton."

Wilhelm von Humboldt, GS Vol 6:147.

In order to complete the linguistic analysis of Topic-Comment structures in Mandarin Chinese, this chapter is dedicated to the study of the prosodic inflection characterizing a particular type of Gapless Topic. Mandarin Scene-Setting Topics will be given a psycho-linguistic behavioral description when they are embedded in the discourse-context.

The question of the prosodic marking of Topic structures in Mandarin Chinese has scarcely been addressed by the literature. Our study aims at providing a start to fill this gap by giving also a neuro-linguistic description of the Prosodic Boundary marking (henceforth PB) yielded by the pause separing Topic and Comment.

The first part of this chapter presents a detailed phono-acoustic study of Topic-comment construction-specific Prosodic Boundary marking (henceforth PB). This investigation is mainly carried out through acoustic-phonetic speech analysis, and will be instrumental for the neuro-linguistic approach we are going to adopt in the next chapter when we will focus on Topic-comment sentences using ERP recordings.

The second part of the chapter is dedicated to a psycho-linguistic study to answer the open question of how and if a certain discourse context, where the Topic referent is given explicit saliency, crucially facilitates the overall comprehension of Chinese basic Topic-comment articulations. These behavioral results will also be instrumental in the next chapter to understand context-induced effects elicited during online processing of these sentence structures.

Why study Topic-Comment articulation in context

As the aim of this chapter is to give psycho-linguistic account of the sentence-discourse interface characterizing Topic-Comment (T-C) articulation, we will try to determine the weight of prosodic information to cue for syntactic structure, and the effects of discourse-context information on the comprehension of Chinese T-C sentences with Gapless Scene-setting Topics. The core of our investigation will be the issue of the interplay between Topic-Comment of intonation patterns, of structures and their triggering of the **interface** with context during in-context sentence comprehension.

Hence, our two main research questions will be:

1. What is the impact on in-context sentence understanding of a clear intonational marking of the sentence syntactic structure, when a Topic construction that can be anyhow parsed on the basis of word-order information;
2. What is the impact of the saliency of the Topic referent in discourse context on the sentence comprehension and its behavioral responses. In other words, is the Topic-related context licensing the Topic-Comment construction?

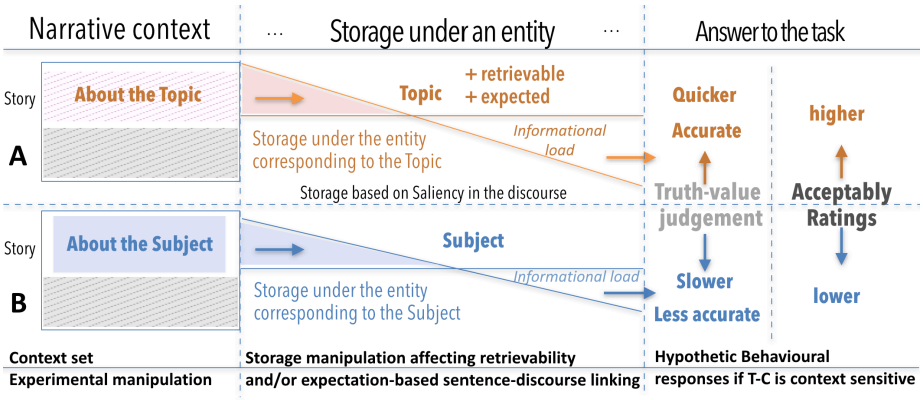


Figure 4.1 – Experimental setting testing how context, giving different saliency to Topic (context A) or to Subject (context B) referent affects the comprehension of Scene-Setting Topic-comment sentences differently marked by the typical "comma intonation" of Topic (i.e. PB) or flat reading (i.e noPB). Truth-value judgments and their reaction times will indicate to what extent Topic-comment articulation are actually constrained by certain licensing context, or are only or facilitated by the activation of the sentence-discourse interface characterizing Topic comment-articulation..

Based on the overview offered by the previous chapter, we can start by saying that three dimensions come into play in the online comprehension of Topic-Comment structures :

1. The syntactic dimension linked to the [NP+NP+VP] word-order,
2. The prosodic dimension linked to the typical 'comma intonation' of Topic-Comment articulation, and
3. The informational status of the Topic's referent in the discursive context.

While the first dimension, the syntactic one, will be held constant in our experimental design, the second -the prosodic one- will be varied by the presence or absence of

Prosodical Boundary. The experimental sentences will have the same syntactic structure and the same propositional content. The third and last dimension, linked to sentence-discourse interface and the informational status of the Topic referent, will be the object of a manipulation in informational-load, meant to uncover if there is in Mandarin Chinese a licensing or favorable context to 'Chinese style' Topic-Comment articulations. In other words, this experiment will allow to say if Topic-Comment in Chinese needs to be contextually licensed or if it is the "preferred sentence structure" as Hockett used to say.

Overview of the contents of this chapter

4.1 Hosting space in Topic position in Chinese

> "*Cross-linguistically, there are very substantial differences between languages not only in the semantic parameters utilized in spatial description, but also in the syntactic means to introduce spatial information in the sentence concatenation [i.e. Topic vs PPs], and that makes it natural to ask how these parameters correlate with the building of mental model during the understanding of an utterance.*"
>
> Levinson, *Space in language and cognition* (2005), *parenthesis mine.*

Why did we choose Scene-Setting Topics

The reasons we chose Scene-setting Topic are semantic, syntactic, psycho-linguistic and methodological.

Scene-setting Topics as Prototypical frame-setters The first and fundamental reason is that this Topic type embodies the prototypical 'frame-setter'. A close look at its in-context behavior and online processes will tentatively allow to tackle the logico-semantic process that is characteristic of Topic-Comment predication, the truth-conditionality of the Topic spans on the whole Comment-clause: "*the Topic has a semantic function: it acts as a restrictor as to when, where or with respect to who or what, the truth value of the predication is to be evaluated*" (Erteschik-Shir, 1997:130).

This means that, to calculate the truth value of a sentence, the hearer has to evaluate its propositional content within the frame set by the Topic which functions as the "pivot for truth value assessment" (Erteschik-Shir, 2007:15).

A valuable example of the restriction the truth value of the whole predication can be found in Sign-Language, where a particular modality-specific means of the non-dominant hand, called *preservation* illustrates this frame-setting Topic information. Not only Topics are found in Sign-Language too[1], but they even show similar types of syntactic derivations. Aarons (1994) has analyzed both Base-generated and moved Topics, and put forward that the opposition between Scene-setting and Aboutness Topics can be found in Sign Language too.

Interestingly, when Topic function is coupled with background spatial information, several Sign Languages use *preservation* handness device to hold the Topic. This consists in the preservation of one sign on one hand, while the second hand articulates a series of other signs.

1. Topics are claimed to be present in the vast majority of Sign Languages, and they are often marked prosodically and non-manually. Some Sign Languages have been claimed to be Topic-prominent (Rosenstein, 2001), but the question is actually fairly debatable (see Sze's PhD, 2008).

Corpus studies on Russian Sign Language (RSL) compared with Netherlands Sign Language (NGT) by Vladimir Kimmelman (Kimmelman and Sáfár, 2012)[2] show that preservation is used for Scene-setting function in locative construction like the one in Figure 4.2.

The sentence illustrated in Figure 4.2, shows a Scene-setting Topic with a the preservation by the left hand of the Topic location "the house", while the right hand enacts the comment "there is a plate on it".

Figure 4.2 – Scene-setting Topics in Russian sign language. Courtesy of Vladimir Kimmelman.

The preservation of the Topic by the left hand is kept until the end of the utterance as to signify that frame-setting restriction applies to the whole sentence, its function is indeed that of the "pivot for truth value assessment", which is further confirmed by the glossing and the 'pointing gesture' in the comment (see Figure 4.2) that co-refers to the Topic: "there is a plate on ***it***".

Specifically, Russian Sign language uses this syntactic means (i.e. preservation) to indicate the background in locative constructions, but also simultaneity of actions, Prominence marking, pointing for discourse Topic across several clauses and clauses as Topics[3].

The frame-setting function of Topics is what inspired Erteshik-Shir's definition of Stage Topic: "*the stage Topic (sTOPt) defines the spatio-temporal parameters of the utterance. Stage Topics may be overt ("this afternoon", "on Park Avenue"), or discoursally implied. [...] The term "stage" here [...] refer[s] to the Time/Place at which the event expressed by the sentence takes place. [...] the event can be viewed as taking place on the stage defined by this Topic*"(1997:26-27). Importantly, framing function is found cross-linguistically, as we will attest in the next sub-section.

Scene-setting Topics and experimental task Secondly, given the above semantic definition of Topic as restrictor of the truth value of the predication is to be evaluated, it appeared to us that choosing an experimental task as a locative restriction (Erteschik-Shir, 1997:130) -e.g. action taking place in one location instead of another, persons and details linked to that place-, would have made clearer for the participant how to perform the truth-value judgment evaluating the coherence of the critical sentences according to the story

2. I should thank here Vladimir Kimmelman who shared with me the video of an example of Spatial Scene-setting Topic.

3. Similar examples can be found in Sze (2008, PhD) for Hong-Kong Sign Language.

narrative context. To this we can add that subsuming a great amount of information under a location entity, rather than to anchor it to an individual, would have made it easier to build stories. Methodologically, we chose Gapless Scene-setting Topics to maximize the possibility to observe informational load effects.

Scene-setting Topics are non-movement derived Syntactically, this Topic type has been analyzed as Base-Generated, which makes it a suitable candidate to investigate Gapless 'Chinese style' Topics and to understand how the relation between Topic and Comment is established in absence of the syntactic complexity linked to movement and the consequent filler-gap dependency-link movement derivation yields.

Thirdly, from the point of view of incremental sentence parsing the clear-cut semantic difference between the two sentence-initial NPs should facilitate the discovery of the [NP + NP + VP] sentence pattern and the identification of their different role. Namely, when no early prosodic cues are given, the sentence comprehension process would be mainly carried on word-order cues and the semantic difference between the Topic referent (spatial information) and the subject referent (an actor) will make the sentence structure more evident. By this we also wanted to exclude any confounding effect linked to linearlization preferences in terms of animacy (§2.2.2.1, p. 109), and avoid actorhood ambiguity in absence of Prosodic cues of the sentence structure, which in turn prevent from having processing cost linked to the reanalysis of the sentence syntactic structure.

Yet, to linguistically ground our choice we would like to briefly add some cross-linguistic evidence showing that information about spatial localization is frequently found in Topic position. Scene-setting Topics are widely attested across languages and the use of Topic syntactic device to express stative space location is not an isolated case found in Mandarin Chinese.

4.1.1 Cross-linguistic evidence for hosting Space in Topic position

An initial question may rise about why there exist different syntactic devices, in our case Topic vs. Prepositional Phrases to introduce spatial information in the sentence. The issue of the link between Spatial localization and Topic syntactic position could also be expressed in the terms used by Levinson (2005), reported in the above epigraph, where he state that languages differ in "syntactic means to introduce spatial information in the sentence concatenation, and that makes it natural to ask how these parameters correlate with the building of mental model during the understanding of an utterance."Although we will not attempt to go any deeper in the understanding of the mental models linked to syntactic encoding of space, it is true that a linguistic system like Mandarin Chinese use the Topic device for locative expressions, when it does have another syntactic device to express stative location and place, in a whole array of locative prepositional phrases.

The use of Topics as a spatial frame-setter, often characterized by the presence in Topic position of bare spatial nominals (NPspace) that can appear without locative preposition and still have a stative locative role in the sentence, and therefore be interpreted as Locative Phrases. Several aspects of Scene-Setting Topics have already been addressed in §3.3 for French and Chinese, where we highlighted its effect on the linearization of subjects in the Comment-clause, the so-called Locative inversion (cf.§3.3, p.331).

We already addressed the impossibility of assigning a fixed interpretational or semantic value to Topics in Chinese 3 (§3.3), this crucially implies that there shouldn't be a distinct semantic feature associated with Topic position, apart from the very well known,

and vague aboutness or frame-like relation with the Comment. We namely came to the conclusion that Gapless Topics should have an exquisitely syntactic function "casting out predication over the whole Comment-clause". We can nonetheless say that spatial framing function observed in Scene-setting function shares a lot with the frame-setting relation in general, and therefore point to Scene-setting Topics as Prototypical frame-setters.

There is cross-linguistic evidence that Topic syntactic device is indeed not unique to French or Chinese to express location. A first example of spatial Topic marking is given in (241) showing that Topic-marker are frequently found in spatial location:

(241) Tagalog

Malapit sa Maynila ang Pasay City
near OBL Manila TOP Pasay City

'Pasay City is near Manila.'(Shopen, 1985)

We can also find topical syntactic configurations for spatial expression in west African Gbe Languages, and more generally Kwa languages (Aboh, 2010).

(242) Place resumptives in west African Gbe

a. Tavo lo **yà** Asiba xe **é** ji.
table DET TOP Asiba climb 3SG Part2.
'As for the table, Asiba climbed on top of it.'
Lit.: 'on top/surface of the table.'

b. Kojo xe tavo lo ji.
Kojo climb table DET Part2.
'Kojo climbed on top/surface of the table.'

c. Tavo lo ji **yà** Asiba xe **é** flen.
table DET Part2 TOP Asiba climb 3SG.
'As for the top of the table, Asiba climb there.' Aboh (2010)

As shown in (242a and c) a definite place NP occupies Topic position. Spatial Scene-Setting Topic is marked by a Topic particle and is also the antecedent of a resumptive pronoun or an adposition generally encoding location in the Comment-clause. *Place resumptive* (Rplace) are also found in German main clauses, when a Place Locative Phrase is in sentence-initial position, as illustrated by the following German example (243). Interestingly, Noonan (2010) analyses '*drin*' as being a remanent constituent allowing scambling or topicalization:

(243) In dieser Kiste sitzt mein Kater am liebsten drin.
In this.DAT box sits my tomcat always with preference DR-in.
In this box sits my tomcat always with preference [in].(Noonan, 2010)

According to the author the '*in*' in '*drin*' is interpreted as an un-deleted copy of Place[4].

In conclusion, we can say that all these examples contribute to ground our assertion about the strong link between Spatial Framing value found in Scene-setting Topic and

4. It would be interesting to make a parallel between the German place resumptive *drin* and the use in Chinese of 那儿 *na'r*. As noted by Xu Dan (1988)

Topic-Comment articulation cross-linguistically. This not only motivates our choice but shows how close Scene-Setting Topics are to the definition of Topic given by Chafe (1976), stating that 'The Topic sets a spatial, temporal [...] framework within which the main predication holds (the frame within which the sentence holds)".

4.2 Prosodic marking of T-C : intonational contour and pause

> *The physical nature and ontological status of utterance accent have been subject to much debate. The protagonists in this debate can be very roughly divided between those who argue that accent is essentially a linguistic, rule-governed phenomenon, derived from syntax, and those who argue that it is entirely non-linguistic, and determined by speaker choice, with no reference to grammar. What is not in dispute is that the placement of accent, whatever its status, can have striking effects on what is communicated by a speaker.*
>
> Chapman (1998:9)

Natural auditory language comprehension naturally happens in context. Under this natural conditions listeners' interpretation of sentence meaning is also guided by factors like prosodical and pragmatic-induced expectations. In the auditory modality, listeners can in fact derive syntactic information, like constituency, chunking, attachment and sentence structure, respectively from local and sentence-level prosodic information, which is generally comprising intonational contours, prosodic boundaries or breaks and accent cues (Cutler et al., 1997).

The prosodic characteristics of the utterance, such as pauses and changes in the fundamental frequency (F0), contribute in the chunking of the utterance stream into prosodic phrases (Fraizer et al., 2006). Spoken language appears then to be hierarchically organized into prosodic units which enable listeners can reliably on intonational contour and prosodic boundaries as cues for the sentence syntactic structure. The prosodic features of Intonational Phrase Boundaries (IPBs) are generally defined as acoustic cues comprising (1) pre-boundary lengthening, (2) pith change, and (3) pausing. These Boundaries mainly coincide with syntactic clause boundaries (Selkirk, 1984/86; Nespor and Vogel, 1986), and have proven to be instrumental in resolving syntactic ambiguities and syntactic preferences (Warren, 1995; and others like Marsel-Wilson 1994), which can temporally arise as sentence unfold.

Prosody, markedness and syntax processing

It is generally acknowledged that the prosodic realization of the utterance manifests both syntactic and pragmatic information. Prosodic cues (e.g. pitch accent, pauses,

intonational patterns, stress) signal both syntactic structure and informational status of sentence's constituents.

However, although pitch reliably gives clues to the syntactic structure of a sentence, a one-to-one mapping between intonational patterns and the syntactic and pragmatic dimensions is difficultly found. For this reason, one should be extremely cautious to distinguish discourse Topic vs. sentence Topic intonational patterns, and to distinguish Topic contrastive pattern, and focal pitch accent as being distinct in terms of markedness from sentence Topic.

4.2.1 Markedness, Topicality and sentence comprehension

For what concerns topicality in mandarin sentence, our discussion on the basicness of Topic-comment articulations brought forward their fundamental un-markedness. Hence, focusing now on the intonational patterns of these basic sentence articulation, we will adhere to the strict distinction between sentence-Topic and discourse-Topic, and concentrate on sentence-topicality prosodic patterns. The sentence-initial element carrying a frame-like function will be considered within the boundaries of the sentence-unit, which of course does not preclude from drawing consideration on how sentential topicality can be further utilized in discourse in different ways.

In this regard, an interesting experiment by Netz and colleagues (2011) evaluates Topic sentence comprehension based on how participants continue intonationally marked and unmarked sentences taken from a recorded natural dialogues corpus (Santa Barbara Corpus of Spoken American English, SBCSAE).

The authors selected Topic types marked at different degrees, that fulfill different functions and therefore implied a different development of the discourse: Left-Dislocated, Object fronted and 'Subject marking', as illustrated in (244).

(244) Effects of Markedness and Topicality on comprehension

a. Left-Dislocation (LD):

a. Cathy –

a'. She's not a good friend herself.

(245) Object Fronting (OF):

b. These shoes we never did put on a horse.

(246) Subject Marking (SM): the subject is stressed and occupies a separate intonation unit

b. ...Dad,

b'. ..you know,

b''. has done some of it.

Note that this corpus transcribes and understands prosodical markedness in terms of intonation units, which in example (244) is rendered having at each line one intonation unit. Fundamentally, this experiment is based on the assumption that listening to a marked Topic construction, listeners have different expectations regarding the development of discourse, which would lead subjects to successfully identify the continuations.

As for the prosodical manipulation of the experiment, the authors considered, following Chafe (1994:58), that the features that characterize delimitation of an intonation unit include changes in acoustic such as pitch, duration, intensity and voice quality, as well as pauses, and thus built an experimental design opposing the above three types of

Topics (see ex.244) featuring prosodic encoding through intonation boundaries while non-marked conditions were artificially transformed from the one fond in the corpus. Hence, to test listeners ability to correctly continue the discourse in marked and non-marked cases, the participants were given context, input sentence and two possible continuations, as illustrated in a marked case in (247):

(247) LD Topic context: (Daughter (Lynne) telling her mother about her training in blacksmithing):
a. Lynne: Horses at college,
they just haven't been disciplined enough.

(248) Continuations:
a. Lynne: And they're just, it's like a kid, they're just ornery,
you know?
b. * Lynne: And that's another thing we had to learn in class you know,
just had to learn our safety of where to stand... Netz et al (2011:ex.8)

The results presented in Table 4.3 show that the accuracy in the discourse continuation task is higher in case of prosodically marked Topic-constructions and show a clear effect of Prosodic markedness on the way listeners perceive and understand these Topic-Comment constructions. This findings lead the authors to conclude that, when a marked sentence structure is in its appropriate context it facilitates listeners' comprehension, while as un-marked structures' function is less clear, people find it significantly harder to reconstruct a coherent development of discourse. Moreover, a second task evaluated participants general understanding of dialogues and reported that for the three marked Topic structure had a facilitation effect on subjects understanding of the whole the discourse meaning[5].

	Experimental (+) marked mean score	Std. deviation	Control (–) marked mean score	Std. deviation	*T*
(±)LD	85.9%	Across subjects: 7.8 Across items: 12.5	77.3%	Across subjects: 13.9 Across items: 10.3	*2.95 *2.49
(±)SM	66.4%	Across subjects: 8.2 Across items: 19.3	59.9%	Across subjects: 11.7 Across items: 22.6	*2.50 ns

* $p < .05$.

Figure 4.3 – Mean scores and standard deviations of the accuracy in the discourse continuation task. The table presents the standard deviations, both across subjects and across items for each of the sentence types, and SM sentences show a larger variance. The percent of correct answers was computed separately across participants, summing over test items, and across items, summing over participants.

We understand these results in the following terms: Intonational units appear here to have a strong impact on the sentence understanding and therefore on the ability to continue a discourse. Interestingly, the authors note that the most stable results are found for Left-Dislocated Topics which are marked not only as sentence-Topics but also as discourse Topics.

These findings are instrumental in showing that erasing natural intonational units in different kinds of Topic-comment constructions has an impact on sentence comprehension

5. Across subjects statistics: the means of all three experimental groups were significantly greater than were the means of the equivalent control groups (for LD, $t(58) = 2.36$, $p < .05$; for SM, $t(58) = 6.49$, $p < .05$; for OF, $t(58) = 4.44$, $p < .05$).

and that this is particularly evident in sentences where the Topic not only sub-serves a sentential function but also plays a role at the discourse level for the continuation of discourse, a function that Netz et al. identify with the fact that in English LD Topics have been described as 'reference-highlighting devices' (Geluykens, 1992), 'Topic establishing devices' (Gregory and Michaelis, 2001) or 'opening moves' according to their own definition.

Similar conclusions can be drawn from a Danish experiment that we presented in chapter 2 that found that sentential word-order featuring either subject initial or object Topic did not have a significant effect on Response Time (see 2.43 (B)), while *context* had, thus revealing faster responses to sentences with a supportive context. As shown in 2.43 (B), accuracy measures revealed that context had a important facilitating effect on the comprehension of object-initial clauses.

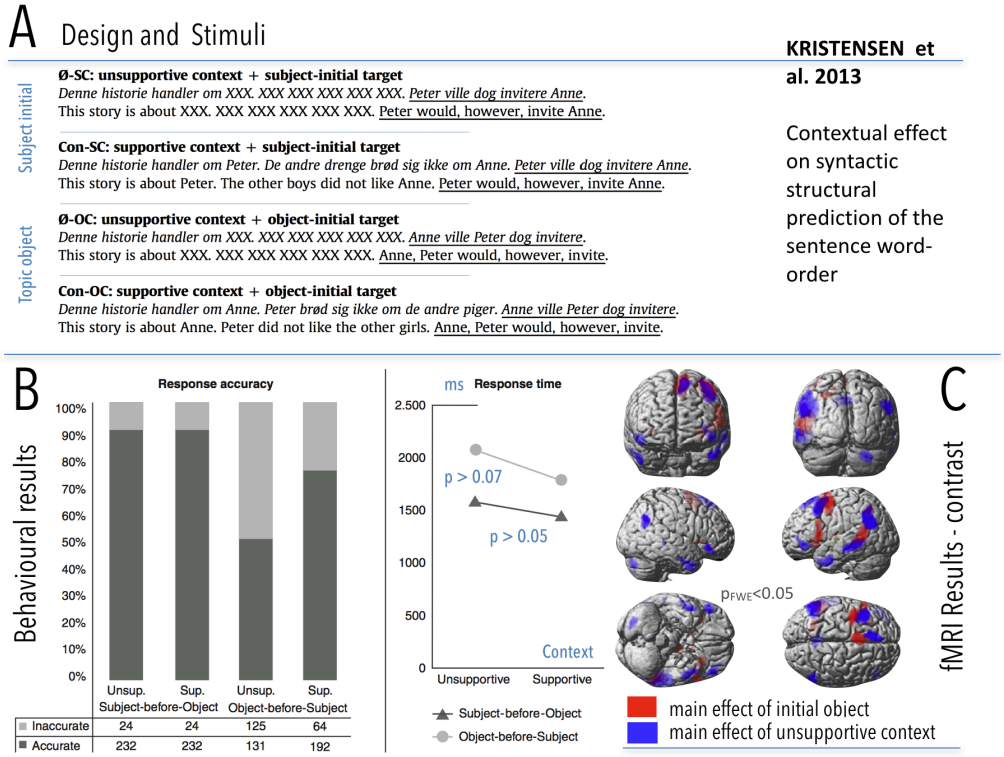


Figure 4.4 – (A) Experimental design and sentence stimuli examples. Participant were invited to read a short context of three sentences and after each target sentence, a comprehension question of the type "Ville Peter invitere Anne?" "*Would Peter invite Anne?*" was presented and responses were collected by button press. (B) Behavioral results for the comprehension task, showing Accuracy (left) and Response Time (right) by sentence types and context. (C) Group-average brain map for the main effect to object-initial sentence in red, and for the main effect of un-supportive context in blue. Adapted from Kristensen et al. (2014a).

From these findings we concluded that considering the sentence-discourse dimension for the investigation of word-order changes in the sentence-unit can actually reveal that sentence processing is highly responsive to discourse information both at the level of

behavioral responses and at the cerebral-level, as we can see from the broadly distributed cerebral activation in 2.43 (C) encompassing the quasi totality of the sentence network.

We should however not hastily generalize these results to our study case in Chinese, because, in a non-Topic-prominent language like Danish, topicalized objects have to occur in a highly constrained context, which often implies contrastivness interpretation. This can also be found in English, with for instance the object-initial clause "*beer I like*" should appear in a contrastive context like "I don't drink wine and shots, but beer I like". Nonetheless, it holds true that these behavioral findings clearly indicate that context information can facilitate syntactic processing and this is namely what we are going to test in Chinese Topic-Comment base-generated constructions to test their context sensitivity.

From this we can draw a series of predictions for our experimental design: the absence of Prosodic Boundary will possibly have an impact on the understanding of Scene-setting Topics and on the truth-value judgments we asked the participants to perform. However, this effect could be relatively small in that the tested Topic constructions does not count as carrying strong discursive functions. Importantly, to ground this last claim, it will be instrumental to compare our findings on the prosodic signature of Scene-setting Topics by first considering the fundamental difference between contrastive in-situ Focal reading and Intonational Topic (I-Topic) profiles. We will consider in the following the results from a phono-acoustic study by Wang and Xu (2011) and other studies investigating Topic-Comment prosodic patterns in other languages.

4.2.2 Chinese Topic prosodic signature in the literature

The intonational profiles of Topic-Comment constructions in Mandarin Chinese and especially that of 'Chinese style' or '*Dangling*' Topics have hardly been addressed in the literature. Probably because this sentence construction is understood as a basic sentence structure, the few phono-acoustic studies offered by the literature we are aware of, have mainly investigated the distinct prosodic patterns of discourse Topic and Focus in Chinese, leaving on the side the sentence Topics in Chinese. By this we namely resume to the the important distinction we previously highlighted between discourse Topics and sentence Topics and will consider these patterns to contrast them with our own results.

As a starting point, we can here summarize previously discussed elements on the marking of sentence-Topics in Chinese as described by non-quantitative linguistic methods (cf. §3.2.3.2). Topics can be separated from the rest of the sentence by means of a special intonational contour or by a clearly audible pause, which can be reinforced by a particle acting as a Topic marker *(y-)a, ne, me, ba* (Shi, 2000; Gasde, 1999; Paul, 2005; Li, 2006 and Badan Del Gobbo, 2010). However, recent investigations (cf. §3.2.2.3, Fig. 3.14) show that Topic markers like *ne* carry a contrastive interpretation and Constant (2014) put forward, a series of tests to show the contrastive-like behavior of Topic-marking -*ne* (2014:306)[6] As we already had the occasion to highlight (cf. §3.2.2.3) we based ourselves on these analyses and on the informers variability of acceptance of the different

6. Namely, contrastive-like Behavior of Topic-marking -*ne* is shown by the fact it:

1. resists exhaustive foci,
2. resists non-contrasting Topics,
3. resists maximal elements,
4. marks if-clauses, but resists because-clauses.

Topic markers, for instance *-a* and *(y-)a*, to decide to avoid marking the Topic-comment articulations with this explicit morpho-syntactic marking in our experiments.

Distinct prosodic patterns for Topic and Focus in Chinese

In a phono-acoustic study, Wang and Xu (2011) investigated how Intonational Topic and *in-situ* Focus are distinctly encoded in Mandarin Chinese. Their study concentrated on these two types of intonational marking of SVO sentences. Avoiding any morpho-syntactic encoding of these two functions the authors aimed at investigating the phonetic means by which one can encode these two different sentence-discourse interpretations, contrastiveness and topicality which can otherwise be expressed syntactically through overt morpho-syntactic means with particles or through dedicated sentence constructions.

Embedding the same sentence into different contexts, Wang and Wu aimed at modifying the degree of newness in the discourse-context of the First Word (w1) to then be able to elicit the corresponding acoustic encoding by 6 native speakers reading aloud. Their experimental corpus comprised a total of 60 sentences with three lengths and five tone combinations, that were recorded by six speakers after four different contexts yielding four different Topic-focus reading of the same sentences. These four discourse contexts were built in order to prime respectively for four different readings: (1) initial focus, (2) new Topic, (3) implicit Topic and (4) given Topic.

As shown in Figure 4.5, the results of this acoustic analysis show that Intonational Topic and in-situ Focus differ both in the prosodic patterns they yield on sentence's First and Second Word. While the literature on these issues in Mandarin Chinese is scarce, the literature on other languages offers several studies on how intonation is used to realize information structure. As for intonational Focus, it is widely accepted that focused element show a higher F0 and longer duration. As for Topic construction, the topicalized constituents are know to carry a prominent rising accent on their accented syllable and to be separated from the rest of the sentence by a boundary tone, or by duration (Féry, 2005).

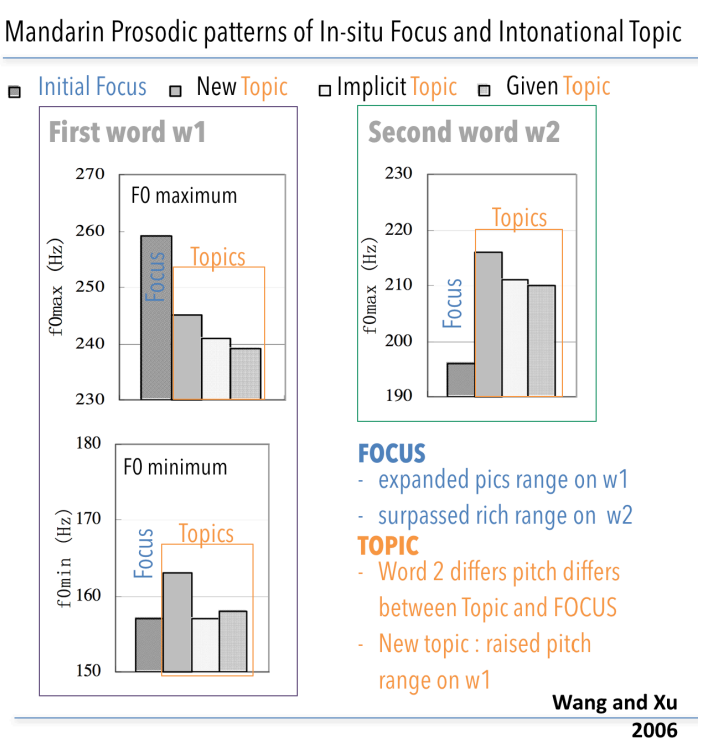


Figure 4.5 – Analyses the speakers' phonetic realization of Topic, initial focus (IF), new Topic (NT), implicit Topic (IT) and given Topic (GT). (A) First word F0 min max; (B) Second word F0 max; Adapted from Wand and Xu (2011).

Specifically, Topic raises the F0 register at the beginning of the sentences while allow-

ing F0 to drop gradually afterwards[7], while *in-situ* Focus is encoded with an expanded pitch range on the first word and a suppressed pitch range on the subsequent word, that is to say it both raises on focus F0 and lowers post focus F0[8]. Moreover, Topic has higher pitch register in isolated and discourse-initial sentences than in no-initial context, namely New Topics are encoded with a raised pitch range on the sentence-initial word.

From these results we can conclude that even in absence of explicit syntactic encoding of the Topic and Focus functions, their intonational profiles significantly differ when conveyed in basic SVO Mandarin sentences.

This is instrumental for our experimental design, because it shows that the syntactic encoding we selected will not risk to have a contrastive/focal intonation, if the prosodic signature we bring to light differs from the the one observed by Wang and Xu (2011).

In this regard, a study on the interrogative and affirmative intonational patterns in Mandarin Chinese by Lin (林茂灿, 2006) addressed the issue of the prosodic patterns testing a sample of sentences sentence structures, among which we can find a similar kind of Topic construction, an Aboutness Topic, whose F0 profile of two affirmative items is reported in Figure 4.6. Lin reports the following properties for these sentence structures: the last syllable of the Topic is lengthen 厅 *ting*, a pause is present between Topic and Comment, 厅 *ting* and 吃 *chi*, and that an F0 variation intervenes between Topic and comment, which form two separate intonational units. The author also notes that although the two items are affirmative sentences, the fact that the first word of the comment can show a higher pitch in Item 1 (Figure 4.6) does not imply any change in meaning.

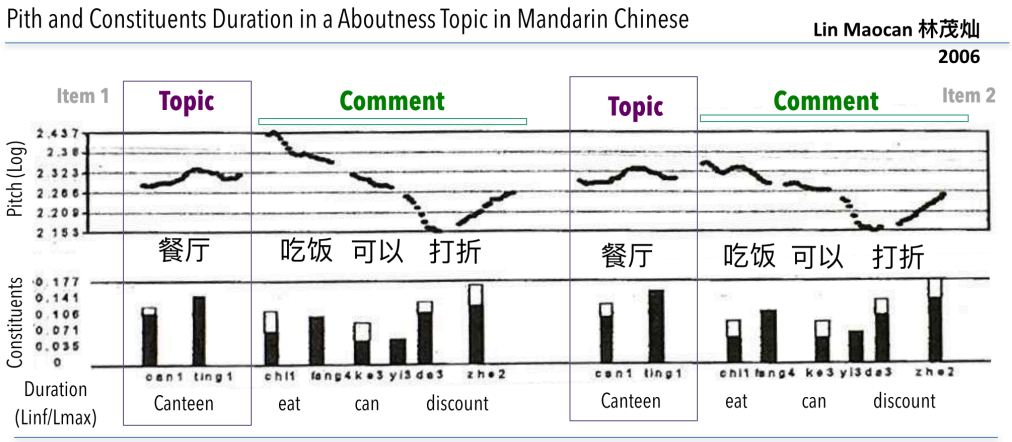


Figure 4.6 – Prosodic profile of Mandarin Aboutness Topic. Lit.Trans: 'At the canteen, meals can be discounted'. Upper part F0 profile (Pitch, Log); bottom part duration of each syllable. Adapted from Lin Maocan (2006).

From the comparison between the findings by Wang and Xu in Figure 4.5 and the above profiles (4.6), we can tell that the profile of the Aboutness Topic in the study of

7. The authors in another article describe Chinese Topic prosodic contour in contrast with focus at the syllable level in interaction with the five tones, and obtain the same results. Wang and Xu, Journal of Phonetics 39 (2011).

8. This is coherent with previous findings showing a sharp post-focus F0 drop, and an expansion of pitch range of focused element, consequently suppressing pitch range of post-focus components (Xu et al. 2005 and 1999).

Lin (2006) highlights a clear intonational boundary between Topic and Comment marked by a pause, and no drastic drop is observed on the second word (i.e. the subject) which seems to be nearer to that reported for Topics that for Foci both by Wang and Xu (2011) and by a previously reported Japanese study on the thematic and contrastive intonational patterns of *-wa* Topic marking in Japanese reproduced here under in 4.7. Nakanishi (2002/2003), namely found that when *-wa* marks theme-role, the highest value of F0 contour after *-wa* particle (P2) does not show a drastic drop compared to the value before *-wa* marking (P1), while when it marks contrastiveness, P2 is not only much lower than P1, but never rises until the end of the sentence.

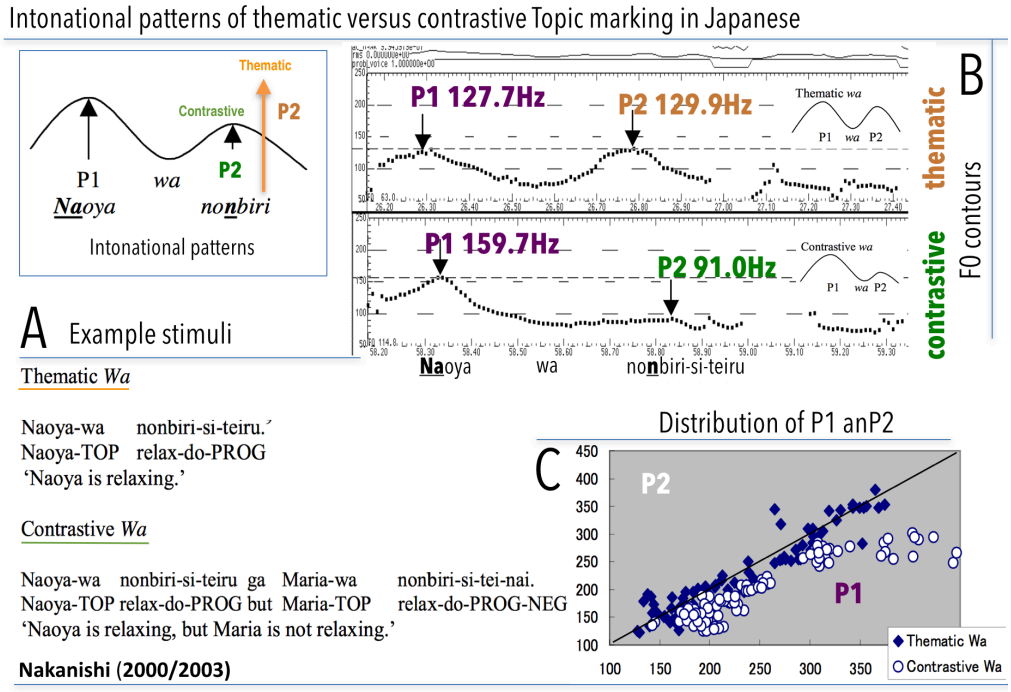


Figure 4.7 – (A) Example stimuli contrasting Thematic *-wa* marking and Contrastive *-wa* marking in Japanese. (B) F0 contours for the two *-wa* interpretation and schematic representation of the to intonational patterns on the left. (C) Distribution of the height of the F0 Topic Peak (P1) and F0 Subject Peak (P2). Blue points represent that there is hardly any difference between the height of P1 and P2 in Thematic *-wa* while white points show that P1 is higher that P2 in Contrastive *-wa*. Adapted from Nakanishi (2000/2003).

It should be noted that the prosodic signature of the syntactic configurations presented in this sub-section are contrastively different, although very different from the one we are going to analyze: the Japanese case illustrates the differential intonational patterns associated to Topic explicit morpho-syntactic marking (which is not our case), and, the study by Wang and Xu points to intonational patterns of *in-situ* Focus discourse function and discourse topicality (except for new Topic profiles).

All in all, from these findings the only patterns that can be retained as informative for our study are those reported by Lin (2006), we expect the the presence of a marked intonational unit with a pause between Topic and Comment. A second indirect hypoth-

esis can be drawn from the study of Wang and Xu: we expect the drastic drop in F0 after Scene-Setting Topic.

Yet, a brief comparison of the different intonational patterns of French Topic-Comment sentences whose syntactic structure is similar to our Scene-setting Topics will be instrumental in guiding our analysis.

4.2.3 The Prosody of Left-Detached Constituents in Modern French

The intonational patterns of French non-focal Left-Detached constituents and Hanging Topics (HTLD) offer some relevant insights to consider some syntactic aspects of the Topic-Comment prosodic encoding. In order to determine the respective weight of syntax and discourse information in shaping the intonation of constituents, the author first distinguished the three constructions given in 4.2.3 on syntactic grounds:

(249) Topic-comment constructions in French with +/- active and +/- accessible Discourse referents (DR)

a. Topicalization: *A Bernard j'ai donné un livre, à Marie, j'ai donné une cassette.* 'To bernard I gave a book, to Mary I gave a tape.'

b. Hanging Topic Left-Dislocation: *(Tiens, j'ai un truc à te dire.) Marie, je connais le flic qui lui a retiré son permis.* '(Hey, I have something to tell you) Mary, I know the cop that took her driving license.'

c. Hanging Topic Left-Dislocation: *Marie c'est clair qu'elle sera fâchée contre son frère.* ' Mary it is sure that she will be angry at her brother.'

d. Hanging Topic assertion: *La Bretagne on y va souvent.* ' Brittany, we often go [there].'

e. Clitic-Left dislocation: *A Jean, Marie lui a offert un compact-disque.*' To john Mary offered a CD.'

The prosodic realization associated with each construction is given in Figure 4.8 and at first sight we can observe that only Hanging Topics and Clitic Left-Dislocated Topics mark to the Topic element with a clear independent intonational unit, which contributes to the clear separation between Topic and Comment. For example compare (A) against the other prosodic profils in Figure 4.8.

Secondly, Deotjes (2002) shows how the syntactic dimension and status of the discourse referent interact in delineating different intonational patterns. For instance, the syntactic dimension allows to distinguish topicalization (A) from left-dislocations (B). This distinction being made, we can observe the continuation rise (H%) boundary tone can be assigned only to the left-detached element in case of topicalization.

The second dimension highlighted by the author is linked to the informational status of the discourse referent. Following Lambrecht (1994), the author distinguishes three main statuses for the Discourse referent of the Topic element: [+active] when the referent is the current center of attention, [+accessible] when the referent is the is contextually evoked or inferable from preceding discourse, and [-active] when it is a new Topic in discourse and it can only belong to the speaker's knowledge.

Given the typology of Discursive Referent described above, when the syntactic construction at play is a Left-Dislocation, be it a HTLD as in (4.8C) or a ClLD in as in (4.8D), the continuation rise (H%) is chosen when the Discourse Referent (DR) is [+ active], this is the case only in ClLD, and H(L)% when it is [- active, + accessible], which is the case in both ClLD and HTLD as shown in 4.8 D.

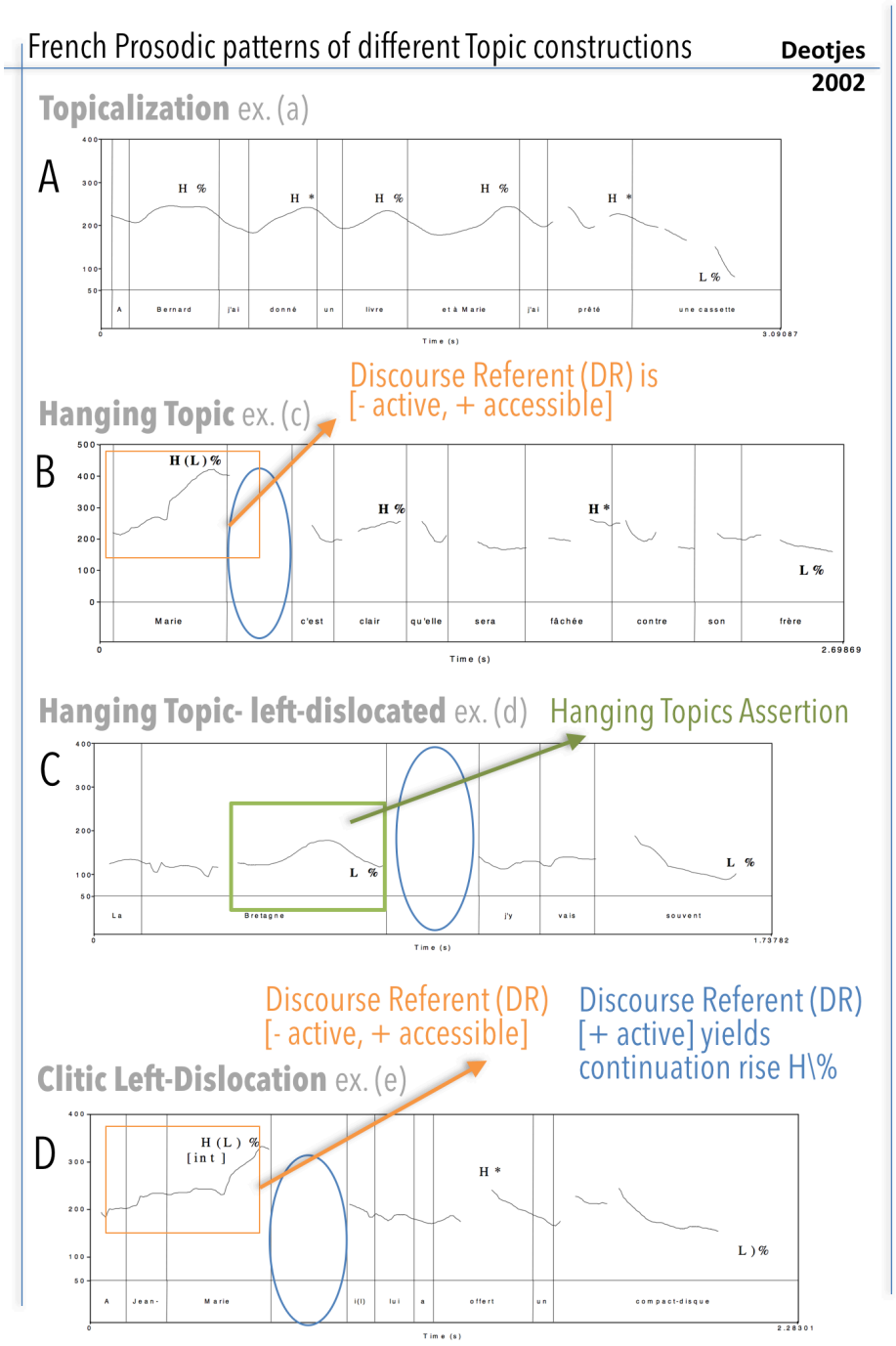


Figure 4.8 – Non-focal left-detached constituent in French an their Intonational patterns in relation to the informational status of the discourse Referent and the syntactic configuration of the Topic-comment constructions. Adapted from Deotjes (2002).

Following Deotjes et al. (2002) analysis of French Left-Dislocation prosodic patterns, two dimensions come into play in the realization of non-focus left-detached constituents in French: (1) the syntactic dimension, and (2) the status of the discourse referent. We will adopt this analysis to perform our phono-acoustic analysis in the following section. We will keep the accessibility status of the Topic-referent constant to manipulate its [+/- activite] status in our narratives.

In sum, from different prosodic analysis of Mandarin and French Left-Periphery phenomena present in the literature, we see that careful experimental work is needed to ground many of the claims made about the link between discourse and topicality[9].

What we could nonetheless harvest from this review are some basic considerations or predictions:

1. we might expect to find a right edge boundary at the end of the Topic intonational phrase
2. an intonational break should separate a left-peripheral syntactic Topic from the rest of the clause
3. there shouldn't be a significant break between the subject and the predicate.
4. we expect to observe a difference in F0 pattern between Topic and Comment

The central questions now are (1) how prosodic pattern of Topic-Comment articulations differs in presence or absence of the pause between Topic and Comment sentence parts, and (2) how this interacts with the four lexical tones present in mandarin to determine the prosodic signature of Topic-Comment articulation in Chinese. We will thus present our experimental corpus and prosodic manipulation in the next session.

4.3 Phono-Acoustic analysis of the experimental corpus

4.3.1 Experimental corpus

Experimental data included behavioral ratings of 37 Mandarin-speaking native subjects listening to a corpus of 270 x 3 Topic sentences, which were declined into three different auditory presented conditions, preceded by (21 x 2) short narratives, constituting two informatively different contexts (A or B).

The two narrative contexts differed according to the different saliency status given to the discourse referents of the Topic NP and Subject NP of our critical sentences. The + active status of one discourse referent over the other was obtained by manipulating the amount of information given about the Topic or the subject as schematically illustrated in Figure 4.9. Contextual information load was manipulated by creating two length-matched versions of short narrative stories that each subject read, with no time constraint, before listening to the critical sentences. The two versions were obtained varying the information load they conveyed about the Topic (A,Rich context) and the subject (B, Plain context). The 3 x 2 conditions meant to manipulate the contextual

9. For an overlook on Mandarin prosodic structure see Peng et al. (2006) in Peng, Shu-hui, Marjorie K. M. Chan, Chiu-yu Tseng, Tsan Huang, Ok Joo Lee, and Mary E. Beckman (2006). Towards a Pan-Mandarin System for Prosodic Transcription. In Prosodic Typology: The Phonology of Intonation and Phrasing, ed. Sun-Ah Jun, 230–270. Oxford University Press.

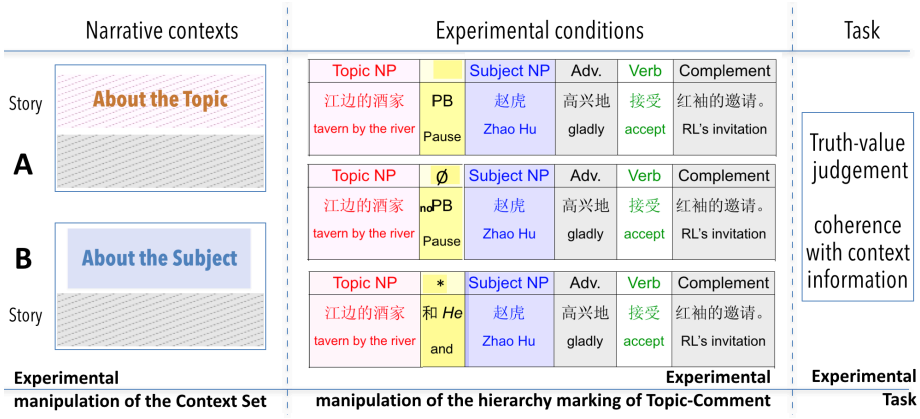


Figure 4.9 – Critical sentence Examples.

saliency (i.e. contexts A or B), while the Prosodic Boundary (PB) marking was meant to manipulate prosodic cues for the syntactic hierarchy between the two sentence initial NP respectively the Scene-setting Topic carrying the spatial information and the subject of the Comment clause (condition c. *He) giving the following conditions:

a. [+ PB-marked] Topic condition, with a pause after Topic NP, the grammatical and natural condition;
b. [- no PB] Topic condition sustained reading of the sentence, but with no pause between Topic and Comment, the grammatical and prosodically under-informative condition, where we nonetheless for the intelligibility of the Comment-clause;
c. Coordination Structure condition, where the relation of the two initial NPs is syntactically cued by the coordination conjunction 和 hé, ungrammatical condition.

Narrative contexts

Experimental corpus comprised 21 sets (Rich/Plain context) of medium length narratives, that were constructed in 3 paragraphs each, counting 320 characters in average (See Annexes §823, p.823).

The one containing more information about the Topic referent of critical sentences (A-Rich context) was meant to give contextual saliency to the referent of Topic NP making it more active in the discourse context, while the second version of the story had the same length and had by half the same content as the first. Critically, it also contained less information about the Topic NP and more about the referents of the subject NP in the other half. While the whole corpus and its critical sentences are reported in the Annexes (§C.1, p.823), a translated example is given in the following page from Story n°2 (Figure 4.10).

From these examples we see that the rich narrative context (A) contains an equal number of sentences with a detailed description of two different spatial settings where different actions or facts are happening to the characters presented in the story. Specifically, this manipulation of the spatial setting of actions was also meant to constrain the participants to rely on the information given about the Topic referent. This in order

A Rich context — **Story n° 2** — Plain context **B**

The watermelon field was plenty of round and big watermelons, given the soft sound cicada and the gentle breeze it would have been a pity to stay indoor, and Xiao Ming went out to do his re-stock for the grocery store. On the way back, he passed by the watermelon field and saw that the old farmer was asleep. Suddenly, some wicked thoughts appeared in his heart, and he stole a watermelon from the field. Leaving the field Xiao Ming inadvertently made some noise, and the old farmer woke up shouting "Stop, stop!", but he couldn't see clearly who was the watermelon robber.
The old farmer promptly went to the police to report these facts. The police came on the watermelon field and found a receipt from the grocery store of Xiao Ming. Immediately after the watermelon grower and the police went to the grocery store to investigate.
In Xiao Ming's grocery store there was always a big mess and the prices were relatively high, but there was a rich variety of goods which actually made his business work well and the clients were numerous.
Once the policemen arrived they showed the receipt to Xiao Ming, who felt guilty and inadvertently hit into the garbage can. And, this is how the police noticed a half eaten watermelon rolling out of the garbage. The old farmer, full of anger, started to criticize the young Xiao Ming telling him it was bad to steal things.

The family of Xiao Ming has always been very poor, his elder sister and him have been depending on each other for survival since early childhood, she has been like a real mother for him. Now that she married and founded her own family, and now lives on her own. Recently she had a child, and Xiao Ming went to visit her to see if everything was fine during her post-birth month. Relieved to see that she was fine he returned home.

Passing by a watermelon field he saw that the old farmer was asleep. Suddenly, some wicked thoughts appeared in his heart, and he stole a watermelon from the field. Leaving the field Xiao Ming inadvertently made some noise, and the old farmer woke up shouting "Stop, stop!", but he couldn't see clearly who was the watermelon robber.

The old farmer promptly went to the police, and the police found in the watermelon field a receipt from the grocery store of Xiao Ming. Immediately after the watermelon grower and the police went to the grocery store to investigate.
Once the policemen arrived they showed the receipt to Xiao Ming, who felt guilty and inadvertently hit into the garbage can. And, this is how the police noticed a half eaten watermelon rolling out of the garbage. The old farmer, full of anger, started to criticize the young Xiao Ming telling him it was bad to steal things.

A Rich context — **Story n° 2** — Plain context **B**

西瓜田里有好多又大又圆的西瓜，虫鸣微风，呆在里面好不惬意。小明去为杂货店进货，回来时路过瓜田，看到老农睡着了，顿起贼心，在田里顺了一个西瓜。离开时小明不小心弄出了动静。惊醒了的老农冲着他大叫"站住站住！"，但没有看清是谁偷了瓜。老农连忙报了警。警察在西瓜田里发现了一张小明杂货铺的进货收据，就迅速和瓜农赶往杂货铺调查。

小明的杂货店里面向来很杂乱，东西的价格也不便宜，但是种类很丰富，因此生意也还不错，总有很多的客人。警察掏出了收据给小明看，小明十分心虚，不小心碰倒了垃圾箱，警察看到里面掉出来吃剩一半的西瓜。瓜农生气地指责了小明不该偷东西。

小明家里一直穷苦，姐姐和他从小相依为命，对他如母亲一样。现在嫁做人妇，有了自己的生活。最近姐姐生了孩子，小明去看望正在坐月子的姐姐，知道姐姐一切安好，十分放心地回来了。

路过西瓜田时小明看到了看瓜的老农睡着了，顿起贼心，在田里顺了一个西瓜。离开时小明不小心弄出了动静。惊醒了的老农冲着他大叫"站住站住！"，但没有看清是谁偷了瓜。老农连忙报了警。警察在西瓜田里发现了一张小明杂货铺的进货收据，就迅速和瓜农赶往杂货铺调查。警察掏出了收据给小明看，小明十分心虚，不小心碰倒了垃圾箱，警察看到里面掉出来吃剩一半的西瓜。瓜农生气地指责了小明不该偷东西。

Figure 4.10 – Examples of narratives from story n°2. (A) Rich Topic-related context. (B) Plain context.

to be able to perform the truth-value judgment task, in that the actions or situations described in the Comment-clause are critically happening in one or in the other location.

For generating Context B narratives, the detailed spatial descriptions giving a central role to the location in the story, e.g. the watermelon field in the above example, were replaced by other information not related to Topic referent, but for example related to the subject's character or by anecdotes making them more salient, e.g. the robber has been growing up in poverty.

Maximal rhetorical coherence of narration and story plot was guaranteed in the two versions by submitting them to three independent readers for verification. Importantly, the actions or situations described in the narratives never used the same verbal lexical items that were present in the critical sentences, but synonyms. Thereby we wanted to prevent the participants to adopt root memorization strategy, which would have possibly yielded an un-natural sentence processing of the critical sentence. Moreover, this last lexical manipulation was designed to ensure that the participants had to fully comprehend the story to be able to perform the truth-value judgment task, and were not just evaluating if they had already seen a given verb associated to some characters in the prior context.

Sentence stimuli

Sentence structure The selected syntactic structure is that of Base-Generated and Gapless spatial Scene-setting Topic construction, featuring a Topic DP semantically conveying the name of a place in the story. In sentence-initial position the Topic DP carries a locative value indicating where the Comment action is taking place and it is not subcategorized by the verb in the Comment-clause, as illustrated by the following examples taken from the experimental stimuli for Story n°2.

(250) Experimental critical conditions

a. PB condition: 西瓜田，老农正在打盹休息。(0201y PBTop)

Xīguā tián , lǎonóng zhèngzài dǎdǔn xiūxi.
watermelon field [PB] old.farmer PROG. take.a.nap rest

'In the watermelon field the old farmer is taking a nap.'

b. noPB condition: 西瓜田老农正在打盹休息。(0201m PBTop)

Xīguā tián lǎonóng zhèngzài dǎdǔn xiūxi.
watermelon field old.farmer PROG. take.a.nap rest

' In the watermelon field the old farmer is taking a nap.'

c. * He condition: 西瓜田和老农正在打盹休息。(0201h PBTop)

Xīguā tián hé lǎonóng zhèngzài dǎdǔn xiūxi.
watermelon field and old.farmer PROG. take.a.nap rest

'The watermelon field and the old farmer taking a nap.'

(251) PB condition: 西瓜田，小明不小心掉落了收据。(0204y PBTop)

Xīguā tián , Xiǎo Míng bù xiǎoxīn diàoluò le shōujù.
watermelon field [PB] Xiao Ming NEG. carfull fall.down ASP. receipt

'In the watermelon Xiao Ming inadvertently made a receipt fall.'

(252) PB condition: 杂货铺，小明把西瓜吃掉了。(0206y PBTop)

Záhuòpù , Xiǎo Míng bǎ xīguā chīdiào le.
grocery.store [PB] Xiao Ming BA watermelon eat.RES.(finish) ASP

'At the grocery store, Xiao Ming finished the watermelon.'

Fillers

In addition to the experimental conditions, filler conditions were included, exactly 15 out of 25/27 sentences per short narrative. Fillers comprised correct sentences, in which the conjunction was followed by a correct noun phrase, that could be confirmed by the first NP, to ensure that participants could not automatically determine the sentence ungrammaticality hearing the conjunction. The average duration of sentences was 2954.1 ms, respectively, critical sentences' length: 3233.5 ms; and filler sentences' length: 2488.5 ms.

Other fillers conditions consisted of syntactically different sentences that were all grammatical in order not to explicitly draw subjects' attention to grammaticality by adding ungrammatical filler sentence and thereby increase the probability of sentences including a morpho-syntactic violation. Fillers sentences were in turn constructed to fit or not with the content of the two different narrative context.

(253) Filler conditions

a. Filler sentence 西瓜田老农对小明进行了批评教育。(0216f PBTop): Wrong place

Xīguā tián, lǎonóng duì xiǎo míng jìnxíng le pīpíng jiàoyù.
watermelons field, old.farmer to Xiǎo Míng carry.on ASP. criticize education

'In the watermelon field the old farmer preached morality to Xiao Ming.'

b. Filler sentence 西瓜田，小明睡了一觉。(0201f PBTop): Wrong Subject actor

Xīguā tián, xiǎo míng shuì le yī jiào.
watermelons field, Xiǎo Míng spleep ASP. one CL.

'In the watermelon field Xiao Ming had a nap.'

c. Filler sentence 杂货铺，警察打了小明一顿。(0202f PBTop): wrong action

Záhuòpù, jǐngchá dǎ le xiǎo míng yī dùn.
Grocery.store, police hit ASP. Xiao Ming one CL.

'At the grocery store, the police hit Xiao Ming.'

d. Filler sentence with another syntactic structure
小明的杂货铺货物价格特别便宜。(0212f PBTop): VP wrong

Xiǎo Míng de záhuòpù huòwù jiàgé tèbié piányi.
Xiao Ming DE grocery.store products price particularly cheap

'The prices at Xiao Ming's grocery store are particularly low.'

e. Filler sentence with another syntactic structure
小明吃完西瓜之后回到了杂货铺。(0208f PBTop): totally wrong sentence

Xiǎo míng chīwán xīguā zhīhòu huídào le záhuòpù.
Xiao Ming eat.RES.(finish) watermelon after go.back ASP. grocery.store

'Xiao Ming after having finished eating the watermelon went back to the grocery store.'

4.3.2 Phono-acoustic Analyses

Recording procedure

The experimental sentences were read aloud in a sound-proof room in Beijing University. After having read the context narratives they were spoken at a normal speed and in a natural way by a male native speaker of Standard Mandarin Chinese. Recordings were digitized at a sampling rate of 44100 Hz (16 bit, mono, normalized to/the peak intensity to -30 dBs using CoolEdit Pro 2.0).

Phono-acoustic measurement

In order to calculate the pitch difference between Topic and Comment, we proceeded to phono-acoustic analysis using the package tuneR of R software.

To calculate the Fundamental Frequency (F0) contour of each sentence in the corpus, we used a function that estimates periodo-grams (spectral densities) of the time series contained in the sound file of each sentence, we first estimated intervals in the signal to be able, then, to calculate the F0 in each of those periods. A running window of non-overlapping 100 ms (width 4096) selected through the time series the samples from which the periodo-grams were then calculated.

Examples are reported in the following Figures, presenting the waveform Spectrogram and F0 contour of some stimuli examples. These were obtained using PRAAT software. Each of the following Spectrogram is accompanied by its respective periodogram calculated with tuneR, so that the reader can visualize the kind of data on which the statistical analyses of the phono-acoustic intonational pattern was performed to determine the F0 average difference between Topic and Comment presented in next section §4.3.5.

Spectrograms of the four different tones in Topic position

The first analysis we performed aimed at verifying if our Topic-Comment construction was actually encoding the Topic as a separated intonational unit.

We therefore proceeded to at first-pass comparison of the spectrograms of our two experimental conditions. Figures from 4.11 to 4.14 distinctly show that natural Topic condition (PB) displays an independent intonational unit for the Topic constituent. The pause time interval in highlighted in yellow, and the cursor is located at the same time point in the two conditions. As we expected from what Lin (2011) illustrated in Aboutness Topic profiles, presence of a pause between Topic and Comment indicates that the two parts of the Topic-Comment articulation are indeed two distinct intonational units.

This first-pass observation of the periodograms and in the estimated F0 superimposed in blue already confirms that, the absence of the drastic drop in F0 after scene-setting Topic that Wang and Xu (2011) reported in-situ Focus intonation.

Moreover, although Wang and Xu (2011) reported that different lexical tones didn't affect the encoding of *in-situ* Focus and Topics, we decided to verify this point by simply inspecting the intonational patterns of sentences with Topics whose last syllable (the place where right edge boundary is situated) have an differed in lexical tone. Examples are given in Figure 4.11 to 4.14.

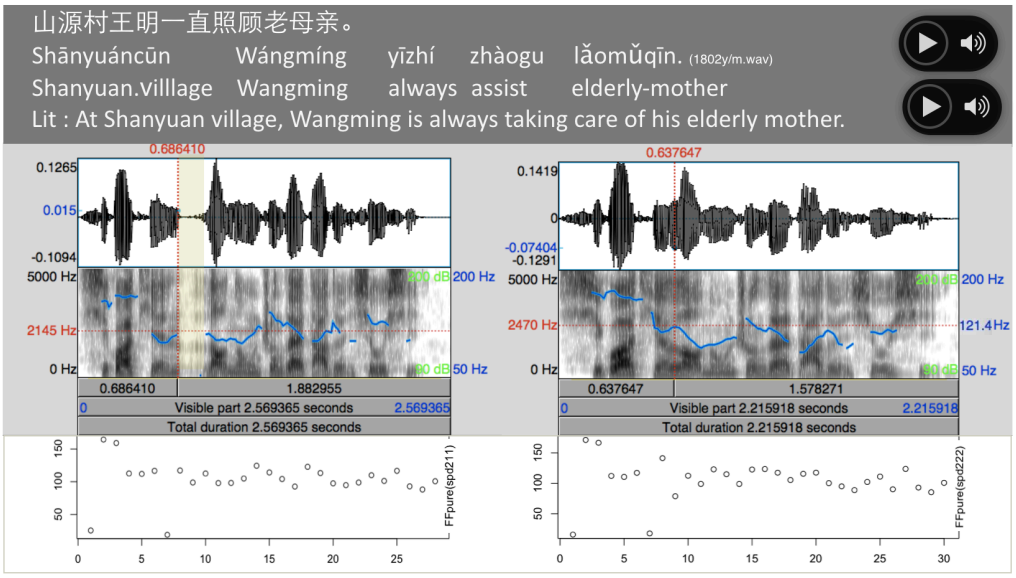


Figure 4.11 – Spectro-gramme and F0 analysis for T-C sentences 0201y and 0201m in our experimental Corpus PBTop, with last Topic Syllable having a 1st tone. Upper section: Sentence and glosses. Mid section: sound-wave and spectrogram with estimated F0 superimposed in blue. Inferior section: each dot indicates the average pitch in the 100 ms intervals.

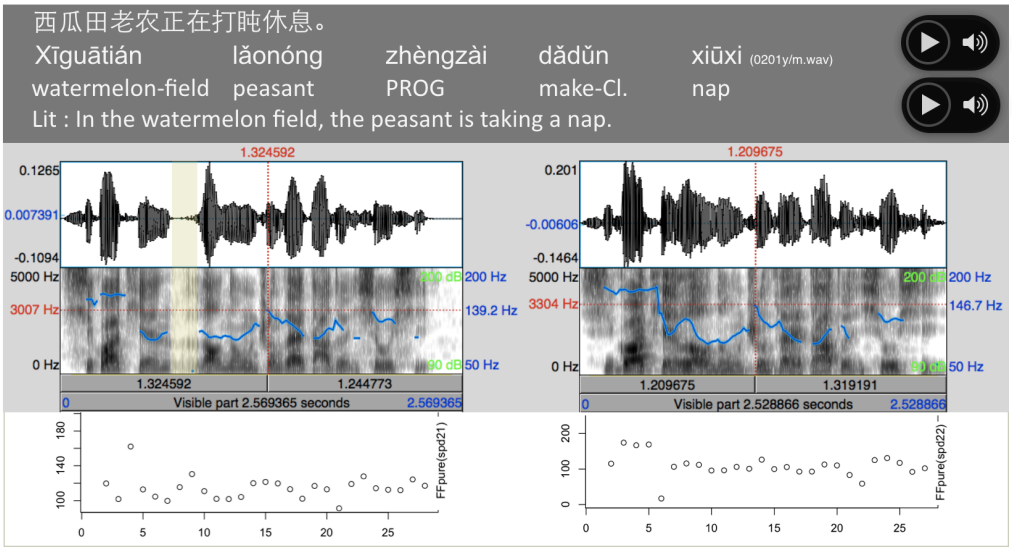


Figure 4.12 – Example of sentence in our experimental Corpus PBTop. T-C sentences 0205y and 0205m with last Topic Syllable having a 2nd tone. **Upper section**: Sentence and glosses. Mid section: sound-wave and spectrogram with estimated F0 superimposed in blue. **Inferior section**: each dot indicates the average pitch in the 100 ms intervals.

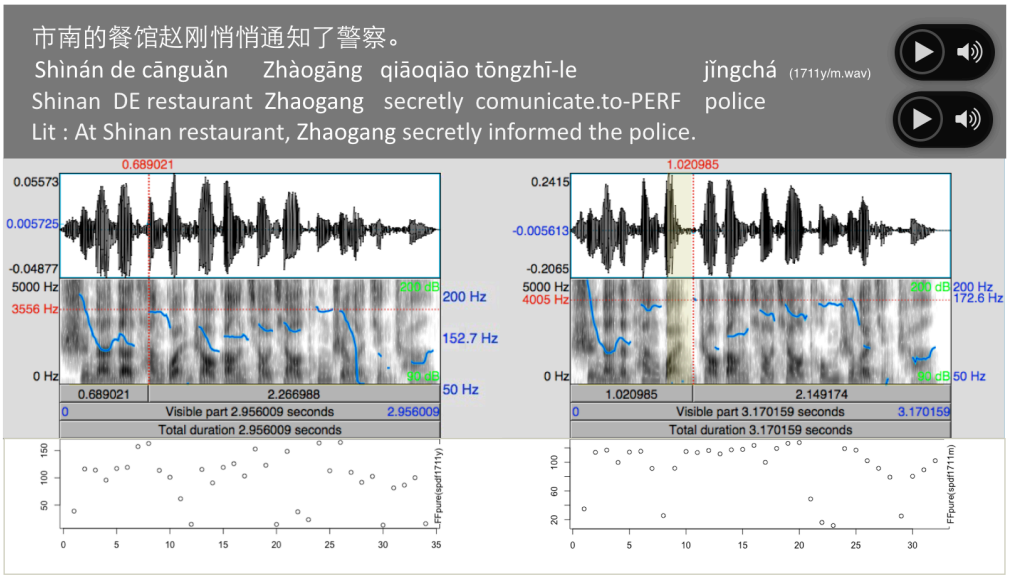


Figure 4.13 – Spectrogram, F0 in blue (PRAAT) and intonational patter (R - tuneR). for T-C sentence 0205y and 0205m with last Topic Syllable having a 3rd tone. Example of sentence used as a stimulus in PBTop ERP experiment (T-C sentence 0205y and 0205m with last Topic Syllable having a 3rd tone). **Upper section**: Sentence and glosses, Mid section sound-wave and spectrogram with estimated F0 superimposed in blue; **Inferior section**: each dot indicates the average pitch in the 100 ms intervals.

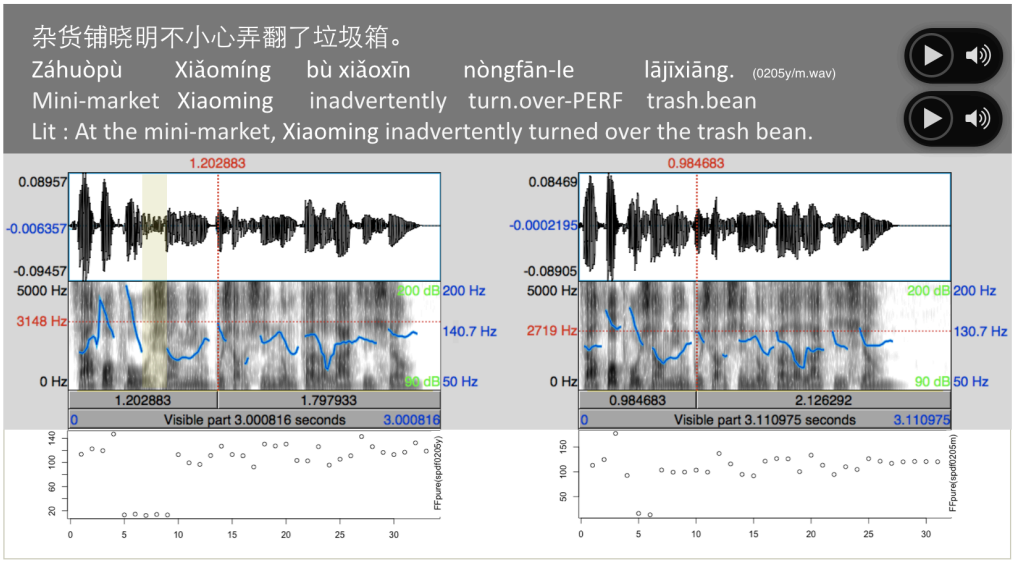


Figure 4.14 – Spectrogram, F0 in blue (PRAAT) and intonational patter (R - tuneR). for T-C sentence 0205y and 0205m with last Topic Syllable having a 4th tone. **Upper section**: Sentence and glosses, Mid section spectrogram; **Inferior section**: Average pitch, quantile 4096.

4.3.3 Average pause length

Our analysis strategy focused on investigating the classical prosodic properties related to prosodic boundaries like pauses, final lengthening and F0 peaks. In the literature, pauses are often used to signal intonation phrase boundaries, while final lengthening and F0 are often used to signal different levels of the prosodic hierarchy. (Cutler et al., 1997)

Hence, to further attest the above observation that Topic is prosodically encoded as a separate intonational unit, we performed a series of statistical analyses on the average pause duration separating the Topic from Comment-clause. Notice that to perform these statistics the whole corpus was manually annotated for the beginning and the end of each sentence constituent and for the beginning and the end of the pause and of Topic last syllable (see Annexes §C.1, p.823).

Table 4.1 – Corpus statistics - Pause duration between Topic and Comment.

Topic Last Syllable		**Descriptive Statistics**		**Anova results**
	PB pattern	Mean Length (ms)	SD	nb sentence items
	He	194,838	42,8430	240
	noPB	,000	,000	240
	PB	**68,940**	61,7787	240
	Total	87,926	91,6244	720

As shown in the Table 4.1, an average pause length of 68 ms is observed and between Topic and Comment.

4.3.4 Last syllable of the Topic

Having a closer look at the duration of the different constituents, we observed a significant lengthening of the last syllable Topic NP in prosodically natural conditions. With an average length of 300 ms, the last syllable of the Topic NP marked by a prosodic boundary, shows a significant average lengthening of 79 ms compared to the non-prosodically marked (NoPB) conditions ($F(2,717) = 246{,}014$, $p<0.001$). This lengthening of the last syllable resulted also in a significant lengthening of the whole Topic constituent ($F(2,717) = 23{,}716$, $p<0.001$), as shown in the following Tables.

Table 4.2 – Corpus statistics - Constituents duration.

Topic Last Syllable	**Descriptive Statistics**			**Anova results**
PB pattern	Mean Length (ms)	SD	nb sentitems	F(2,717) =246,014 p<0.001
He	233,596	36,3133	240	
noPB	228,138	47,7931	240	
PB	307,679	46,6342	240	
Total	256,471	56,9071	720	

Topic NP	**Descriptive Statistics**			**Anova results**
PB pattern	Mean Length (ms)	SD	nb sent. items	F(2,717) =23,716 p<0.001
He	526,773	136,9488	240	
noPB	518,742	144,3833	240	
PB	599,232	141,6784	240	
Total	548,249	145,4245	720	

Subject NP	**Descriptive Statistics**			**Anova results**
PB patterns	Mean Length (ms)	SD	nb sent. items	F(2,717) = 4,780 p<0.01
He	509,470	96,4945	240	
noPB	482,494	101,3324	240	
PB	487,776	105,8480	240	
Total	493,247	101,8282	720	

Table 4.3 – Table corpus statistics on sentence constituents duration: Topic Last Syllable, Topic NP, Subject NP. Note that no significant differences were observed for the Verb in the three experimental conditions.

Topic Last Syllable	**Pair-wise Analysis**			
Condition I	Condition J	Diff. Mean (I-J)	Standard Error	Significance
He	PB	5,457	4,0061	.521
	noPB	-74,083*	4,0061	.000
noPB	He	-5,457	4,0061	.521
	PB	-70,540*	4,0061	.000

Topic NP	**Pair-wise Analysis**			
Condition I	Condition J	Diff. Mean (I-J)	Standard Error	Significance
He	PB	8,031	12,8749	1.000
	noPB	-72,459*	12,8749	.000
noPB	He	-8,031	12,8749	1.000
	PB	-80,490*	12,8749	.000

Subject NP	**Pair-wise Analysis**			
Condition I	Condition J	Diff. Mean (I-J)	Standard Error	Significance
He	PB	26,976*	9,2471	.011
	noPB	21,694	9,2471	.058
noPB	He	-26,976*	9,2471	.011
	PB	-5,282	9,2471	1.000

These findings converge with what found in a previously presented phono-acoustic study by Chang Y-C (2001), where the author was able to differentiate SVO from Topic-comment structure's intonational patterns in Mandarin pointing to a significant

lenghthening of the last Topic syllable, even if the type of Topic construction was different from ours. Namely, the phono-acoustic analysis of the onset and rhyme duration of every syllable in sentences like (254a), than when it was playing the role of the subject of the verb *chi* 'eat', thus, yielding the 'The chicken has not eaten.' and (254b) 'The chicken was not eaten.' revealed

(254) 鸡不吃了

a. 鸡不吃了 Topic

jī, bù chī le
chicken NEG eat PRT.

'I don't eat chicken anymore.'

b. 鸡不吃了 subject

jī bù chī le
chicken NEG eat PRT.

'The chicken does not eat anymore.'(Chang 2001)

that the word 鸡 *jī* 'chicken' showed a significant longer duration when pronounced as a Topic (yielding the interpretation in (254a), than when it was playing the role of the subject of the verb *chī* 'eat', thus, yielding the interpretation in (254b). The convergence of our last finding with those reported for other types of Topic-Comment constructions are indicating that our analysis is nailing down a prosodic pattern that could be attested for various Topic types.

In sum, we can briefly list the characteristics of the prosodic contour of Scene-setting Gapless Topic observed so far. Our experimental condition marked by a Prosodic Boundary shows (1) a significant average Pause length of 68 ms between the Topic and the Comment; and (2) a significant Topic NP last syllable lengthening of 80 ms. These two measures qualify the intonational pattern of out Topic-Comment construction as a classic-Prosodic Boundary. We will now move to the analysis of the the F0 Average difference between Topic and Comment in order to test for the marking of a prosodic hierarchy between Topic and Comment.

4.3.5 F0 Average difference between Topic and Comment.

Having established the prosodic profile of Topic-Comment as presenting two distinct intonational units, now to investigate for the hierarchy encoding of the Topic-Comment articulation, we further tested if a significant difference in pitch height between the Topic and Comment-clause is observable compared to noPB conditions.

Based on the periodograms calculated for each sentence (sampling every 100 ms) we obtained two average measures of the pitch (F0), one for the Topic constituent (3 to 5 syllables length) and one for the Comment-clause.

Statistical analyses reported in Figure 4.15 show a significant 18 Hz average pitch difference between Topic constituent (irrespective of its lexical tone) and comment, an effect that we would like to call a *double-decker* prosodic signature signaling the syntactic hierarchy between Topic and Comment in Base-generate Gapless Scene-Setting Topic.

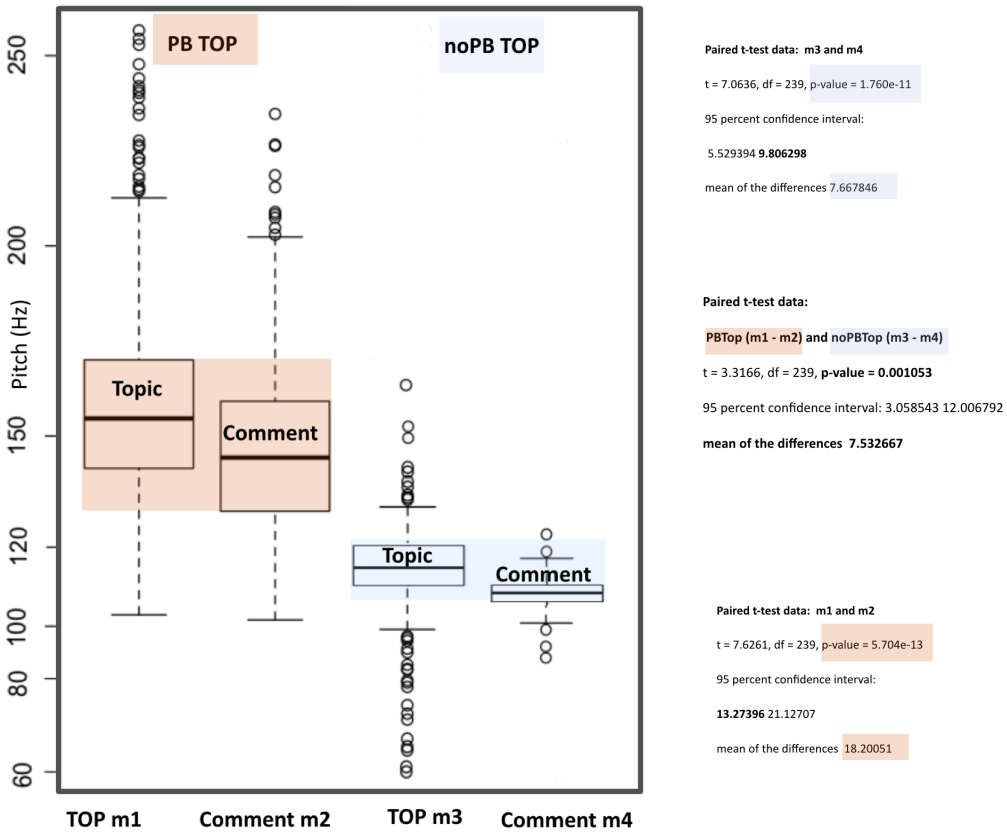


Figure 4.15 – Density distribution of Average FO in Topic versus Comment.
Paired t-test data: PBTop [Top (m1) - Comm (m2)] and noPBTop [Top (m3) - Comm (m4)]
t = 3.3166, df = 239, p-value = 0.001053 ; 95% confidence interval: 3.058543 12.006792 ; mean of the differences 7.532667.

In sum, the Topic prosodic signature in Mandarin Chinese of T-C articulation revealed by our phono-acoustic analysis leads us to the following main conclusions. The characteristics of this prosodic contour are:

1. Significant lengthening of of 80 ms of the last Topic syllable
2. 20 Hz of significant average pitch difference (F0)
3. Average pause length of 70 ms between Topic and Comment

For future research, it will be interesting to verify if these prosodic parameters can be reproduced in the different Mandarin Topic types we reviewed in previous chapter. Now that our phono-acoustic analysis extracted a prosodic signature for our Scene-setting Topic-comment construction, we can move to the different behavioral responses it yields in context.

4.4 Behavioral study of T-C sentences in context

Aiming to give a psycho-linguistic account of the sentence-discourse interface property of Topic-Comment articulation in Mandarin Chinese, we decided to embed them in context. We first conducted a behavioral study investigating the sensitivity of native speakers to the Prosodic Boundary marking characterizing them (i.e. Pause-marking duration and prosodic contour), to be able in a second time to observe how Topic-Comment sentences are understood under different prior contexts in presence or absence of PB.

The close link between information structure, Topicality and sentential intonational patterns is generally acknowledged in the literature. Studies on Chinese parsing ambiguity, like Chang (2001) (or Li, 2011) already showed that especially in absence of explicit Topic making, the syntactic encoding of the Topic function is tightly linked to oral intonational patterns, remember the different interpretation between "I don't eat chicken" and "my chicken does not eat".

Hence, having a clear picture of the effect of the prosody-syntax interface characterizing our Topic sentence articulation, is an essential step to then be able to account for the sentence-discourse interface yielded by Topic-Comment articulation, both behaviourally and during online sentence processing.

4.4.1 Experimental Procedure

Participants were seated comfortably in a sound-proof, electric-shielded chamber at Peking University, in Laohuaxue building. They were about 80 cm away from the CRT screen, where the narratives were presented in the center of the screen, and were invited to move their heads or bodies as little as possible.

Participants were first asked to read a short narratives (i.e. context) carefully and attentively for comprehension, with no time limits for the reading. Then a fixation cross was presented (for 400 to 900 ms) and they listen to 27 critical sentences per story in average (11.43 critical sentences and 15.57 fillers in average), whose presentation was jittered in time so that the subjects could not predict the exact onset of the auditory stimuli.

All together 1440 sentences were tested per subject, 240 per experiential condition, and 720 fillers sentences comprising a great variety of linguistic structures (SVO, *bei* passives, Disposal *ba*, etc.).

During auditory stimulation, EEG signal was recorded, and 300 ms after each sentence a question mark appeared on the screen and participants were asked to judge whether each sentence was true and consistent with the story content they just read.

The experimental session was divided into 5 sessions with 4 breaks, the distribution of the stories in each block was randomized selecting for one half of the stories context A (rich context) version and counterbalancing them with the other half with context B version. Eight lists (4 pairs) of randomized stories were created and for each context, the three conditions were counterbalanced using Latin square design: one third of the sentences were assigned as PB, one third as No-PB, one third as He, making sentence list A, as well it's counterbalanced counterparts list B and list C. The two Layers were then mixed and matched, in total 8 (story lists) * 3 (Sentence Lists), 24 presentation lists. For presentation to the subjects, the sentences were randomly ordered within each story, and it was made sure that no three sentences in a row were all critical sentences.

Post-experiment grammaticality ratings

After the EEG experiment, participants rated the naturalness of all the critical sentences they heard during the experiment. In this post-experiment grammaticality rating tasks subjects were asked to explicitly judge the naturalness of the critical sentences auditory stimuli on a scale from 1 to 5 (1: Very unnatural; 5: Very Natural).

Every block had four stories only the second block had five. The entire experiment including the completion of post-experimental grammaticality questionnaire and electrode preparation lasted 1.5 hour.

4.4.2 Experimental task, manipulations and their hypotheses

Before entering this section dedicated to the two behavioral measures we collected, a few well spent words will clarify the different hypothesis that are underlying the experimental design.

4.4.2.1 Prosody manipulation

Our first experimental manipulation aimed at making the syntactic structure of Topic-comment more or less evident in auditory presentation. Its aim was studying the sensitivity of native speakers to the duration of pause marking and prosodic boundary contours in Topic sentence comprehension. The behavioral response collected will put forward if its absence is crucial for comprehension and to what extent listeners can simply rely on word-order during auditory sentence comprehension.

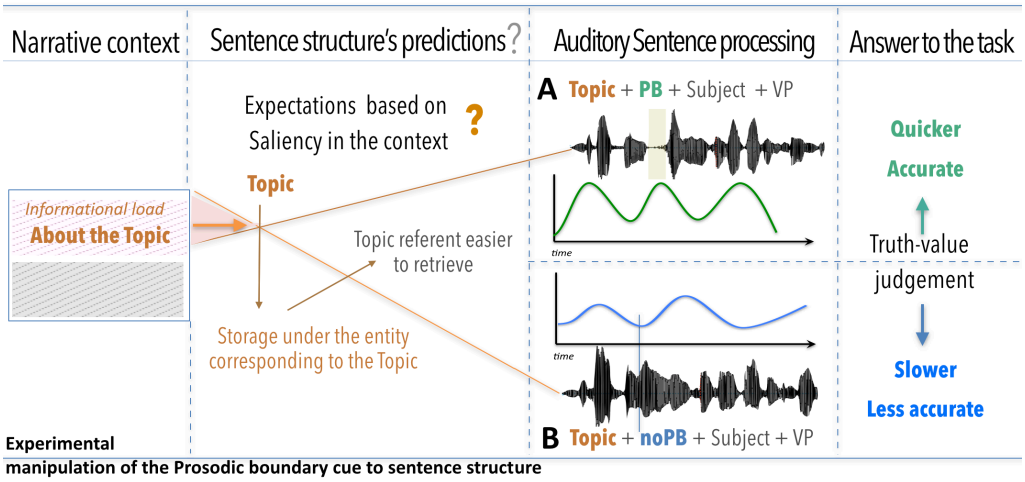


Figure 4.16 – Experimental manipulation of the Prosodic boundary in Topic sentence and behavioural response.

4.4.2.2 Discourse manipulation and truth-value Task

By embedding Topic-comment sentences in context, our aim was mainly to investigate behaviorally the sentence-discourse interface of this sentence articulation, and test how

previously presented context, giving different saliency respectively to Topic (context A) and to Subject (context B) referents in a fictitious narrative scene, would affect the comprehension of subsequent basic Topic-comment sentences differently marked by the typical "comma intonation" of Topic or flat reading (i.e noPB).

The experimental setting described in Figure 4.17 will allow to uncover if the acceptability on Topic-Comment sentences by native speakers of, otherwise out-of-context felicitous Topic-Comment articulations, are context-sensitive or actually licenced by a particular context. Specifically, we will ask if this sentence structure articulation is sensitive to the saliency given to the Topic referent in previous context, despite its basic status, inside the Chinese Linguistic system. Concretely we will test for how the link with discourse context influences the accuracy of a Truth-value Judgment performed during online sentence processing.

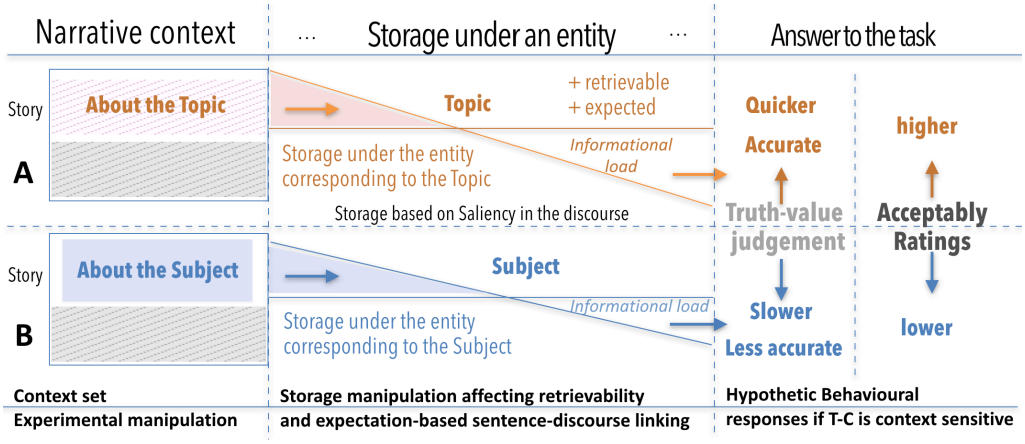


Figure 4.17 – Experimental hypothesis and predictions on behavioral ratings and reaction times of Gapless Topic-Comment sentences embedded in discourse.

The error rates of truth-value judgments will indicate to what extent Topic-comment articulations are actually constrained or facilitated by certain licensing context where the Topic discourse referent is +/- active and salient, as illustrated in Figure 4.17. While on-line electro-physiological measures will inform us about (1) the processing mechanism of the Frame-setting relation between Topic and Comment in Gapless Topics, (2) the underlying effects of context-load during incremental processing, and (3) the processing costs linked to the absence of PB usually marking the Topic-comment articulation.

It should be noted that based on linguistic literature reviewed in previous chapter (cf. §3.1.3.5), two main mechanisms can be hypothesized to take place during in-context Topic-Comment linguistic phenomena. The first linked to the information storage of the narratives content, following Reinhardt (1982), and the second hypothesizing context-driven expectations in the hearer regarding Information Structure's syntactic encoding of the critical sentences.

Discourse information storage Reinhardt (1982), pointed to a discourse storage mechanism where the Topic is considered as the entity/unit under which discourse information is stored, which would make the salient information easier to address and retrieve to

then comprehend subsequent sentence. We can namely read:

> A useful metaphor for the procedure involved [...] is the organization of a library catalog [...]. The propositions admitted into the context set are classified into subsets of propositions, which are stored under defining entries. At least some such entries are determined by DP-interpretations. DP sentence-Topics, then, will be referential entries under which we classify propositions in the context set and the propositions under such entries in the context set represent what we know about them in this set [...]. (Reinhart, 1982:24)

As show in Figure 4.17 in context A, we assumed following Reinhart (1982) that half of the narrative's information would be stored under the subsequent sentence's Topic referent, and under the subsequent sentence's Subject referent in context B.

Context-induced expectation on sentence structure

The second mechanism highlighted by the linguistic analyses we reviewed concerns the possible discourse-driven exceptions a certain context can yield on upcoming sentences and their structure. In this regard, Buering (1999) claims that in German a contextually introduced Aboutness Topic is expected to be found in sentence-initial position, whereas a 'neutral' context would not generate such syntactic expectation.

As we will see in next chapter, several psycho-linguistic and ERP studies have also revealed that prior context causes language listeners to generate expectations on the particular informational status of discourse entities, and, such expectations can in turn generate expectations on the pitch accent of certain constituents in the up-coming sentences (see Heim and Alter, 2006; or Toepel and Alter, 2004). For example, in a reading experiment Cowles et al. (2007) observed that a contextually unexpected word placed as the focused constituent in an it-cleft sentence like "It was the rabbit that ate the salad." would yield an N400 component, which was interpreted as the evidence for a strong expectancy in terms of informational status about the focal informational status of the clefted referent.

As for the tested base-generated Gapless Topic-comment sentences, three scenarios can be hypothesized. The first, the simplest, would show that Scene-Setting Topics are behaviorally not context sensitive and represent a basic structure type.

The second will hold if the tested Topic-comment sentences are observed to be strongly context-sensitive. In this scenario the contextual saliency would generate expectations on the syntactic Topicality of the Topic NP-discourse referent and increase the acceptability of sentences embedded in a narrative context where the Topic is salient. This would lead us to conclude that the contextual informational status of the Topic referent plays a licensing role for 'Chinese style' Topics. However, another alternative linked to our Discourse manipulation would also show some context sensitivity but not linked to a licensing role of the context. This third option could actually show that the saliency of the Topic referent simply facilitate the retrieval of information about the Topic and widely increase the comprehension of truth-value judgments of the Topic sentences embedded in a narrative context.

4.4.3 Truth-value Judgments of Contextual coherence

Truth value-judgment - Discourse coherence judgment

A recent study by Dimitrova and colleagues (2012) gave empirical evidence to the claim that a certain number of the response variation reported in the ERP literature on prosodic processing were actually task-dependent, highlighting in this way the critical role of the choice of task to investigate prosody related components.

In order to avoid to have task-dependent ERP prosodic components and the risks highlighted by Dimitrova and collegues, we decided to give our participant a novel task that would avoid the participant to focus on the prosodic or syntactic parameters manipulated in the experiment. We had them rate the truth-value of the critical sentences according to the content of the sentence with the previously read short narrative.

4.4.3.1 Topic-comment and truth-value judgment

Leaning on the analysis given by Kiss (1995 and Erteshik-Shir (1997/1999) who pointed that Topics are used as a basis for the evaluation of the truth-value of the utterance, we chose a task that could maximize or at least preserve the mechanism of truth value judgment that appears to be one of the core definition of the Topic-Comment sentential articulation (cf. section on Topic notion, §3.1.1)[10]

This truth-value judgment task, evaluating the propositional coherence of target sentences with the narrative context (i.e. fit between context content and target sentence's content), was meant to evaluate the comprehension of the stimulus sentences. Moreover, by rating the coherence of the sentence with the prior short narrative through a binary choice "Correct" (Coherent) and "False" (Incoherent) we would observe to what extent the absence of Prosodic Boundary would impact the accuracy of this judgment process and hinder the sentence-discourse relation.

Importantly, we tried to force the explicit contextual-discourse judgment to happen toward the end of the comment-clause by constructing our sentence materials in order to present new information or different lexical items (i.e. synonyms of the verb presented in the story) only at the verb. This would force the participants to make the judgment only after the verb is encountered. In informational terms, the first two noun phrases were background information and the verb was either a contextually incongruent verb or a synonym of the verb appearing in the context narrative.

Hence, a closer look at these two behavioral responses will be namely informative about the impact on in-context sentence understanding and processing of a clear prosodic and intonational marking of the sentences syntactic structure.

4.4.3.2 Truth-value ratings

Behavioral responses was recorded during ERP experiment and than analyzed. Behavioral data were analyzed using a linear-mixed effect model, including subject as random effect and using Maximal Random effect structure as justified by the data.

As show by accuracy measure of the Truth-value judgments results in Figure 4.18, no main effect of context was found. Planned contrasts showed that Sentence With

10. Moreover, we expected to find some particular ERP effects at verb position as this truth-evaluation of propositions is essentially determining the truth of the predicate in relation to the Topic (Erteshik-Shir, 1997/1999).

No-PB were less likely judged to be coherence compared to Sentence With PB (t=4.937, p<0.001), although the comparaison of the means actually shows a very minimal difference.

Both sentences with PB and without PB were significantly more likely to be judged as congruent than sentences where we inserted a 'conjunction violation' between Topic and Comment: percentage of critical sentences judged as consistent in PB condition was slightly lower for sentence without PB (t=4.92), while unequivocally both prosodic contours were higher than the conjunction violation condition.

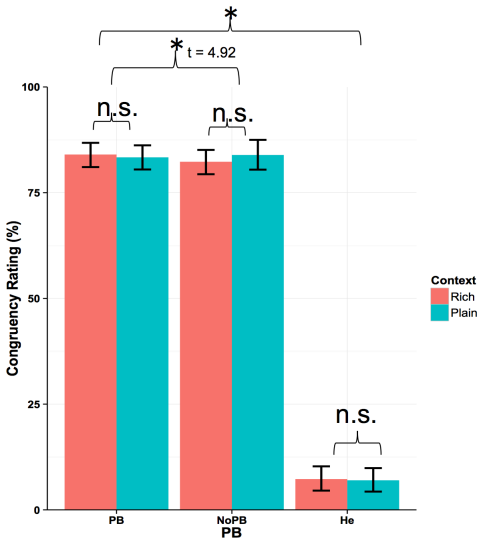


Figure 4.18 – Behavioural Ratings: Percentage of critical sentences judged as consistent.

As no main effect of context was found, we can conclude that our base-generated Gapless Scene-setting Topics are perfect for every discursive context and that this Topic-comment articulation is not constrained by certain licensing context. However, the descriptively small but statistically significant facilitative effect that led subject to correctly judge the critical sentences to be consistent with the story under the PB condition, suggests that the presence of PB facilitated comprehension.

This effect of Prosodic boundary on our behavioral measures of context coherence judgments can also be interpreted by briefly resuming to the issue of sentence syntactic hierarchy tree-like representation addressed in chapter 2 (§2.3.1). Following Wagner (2010) we can sketch the following explanation. According to the framework put forward by the author the syntax-prosody alignment is viewed as being achieved by what he calls 'syntactic readjustment'[11].Thus, sentences with no prosodic boundary would be generate flatter syntactic-tree structures, while a prosodically cued Topic NP would immediately cue for the right syntactic parsing of our T-C articulation.

Given these results showing a facilitatory effect of Prosodic Boundary, the next question that we will address in the following chapter is: What does the neural underpinning of this intonational pattern can reveal about the online syntactic parsing of Chinese Topic structures? But before, let us consider the acceptability ratings for noPB sentences.

11. On this point, see also Chomsky and Halle (1968), contra Selkirk (2000).

4.4.4 Acceptability Judgments

The third and last behavioral measure was collected off-line, in order to assess the acceptability and naturalness of sentences in presence or absence of pause (Prosodic contour vs. flat reading without pause).

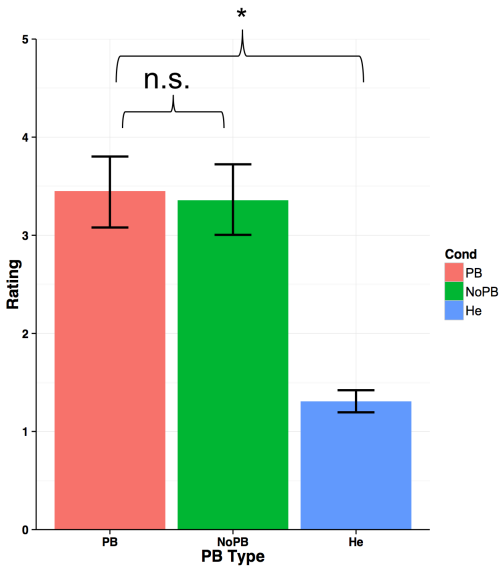


Figure 4.19 – Ratings results from naturalness judgment for critical sentences. Ranging from 1: Very unnatural to 5: Very Natural. Sentences w/o PB were both judged to sound more natural than sentences containing a conjunction violation (He condition). Participants (n = 23).

These results show subjects judged sentences with or without PB to be comparably natural, and that the flat reading conditions were significantly judged as being more natural than ungrammatical condition with a conjunction between NP1 and NP2. This finding could also explain the alleged optionality of comma-marking for written Topic-comment constructions.

From the results in Figure 4.19 we conclude that the impact of absence of Prosodic and intonational cues to syntactic hierarchy was not critical at the level of behavioral ratings, thus showing that participant could rely on word-order cues to parse and understand the sentences even without prosodic boundary information between Topic and comment.

To conclude, it will be important to asses if these findings are confirmed analyzing the ERP brain potentials evoked by the two prosodic conditions (PB and noPB) during auditory sentence comprehension. Cerebral online sentence processing might reveal processing costs that are not observable at the level of behavioral response. Importantly, while off-line measures like acceptability ratings or truth value judgments can be used to attest whether a certain context licenses certain syntactic structures and their prosodical contour, by providing hints on the sentence processing on a global level, an online cerebral recording methods, like ERP, will be instrumental in characterizing the underlying mechanisms of context linking during incremental processing of Topic-Comment

syntactic articulation.

In sum, by embedding Topic-comment sentences in context, we aimed at investigating behaviorally the sentence-discourse interface of this Scene-setting Topic articulation. The offline behavioral measures converged in indicating a facilitative effect on sentence comprehension in presence of Prosodic information cueing for the syntactic boundary between Topic and Comment, we interpret this as a mark of the early activation of the sentence-discourse interface.

Moreover, given the effect observed in truth-value judgments, we can say affirm that the hypothesis of contextual licensing of 'Chinese style' Scene-setting Topic is discarded, there is no specific licensing context for Scene-Setting Topic-comment sentence.

Next chapter will show if and how the preceding discourse-context and the informational load it carries about the different sentence discourse referents (i.e.Topic and Subject) has an impact on the online Topic-comment processing.

We might ask now if the results would have been the same if we had chosen an Aboutness Topic instead of a Frame-setting one. We leave this interesting comparison for further research.

4.5 Summary and Discussion

In this chapter, we have sought, on the one hand, to complete the linguistic analysis of Topic-Comment structures, presented in chapters 2 and 3, by studying the prosodic contour characterizing them in Mandarin Chinese, and, on the other hand, we gave a psycho-linguistic and behavioral description of Topic-Comment constructions embedded in discourse-context.

In conclusion, the aim of this study was to behaviorally determine the weight of prosodic syntactic cues and discourse-context information on the comprehension of Chinese T-C sentences with Gapless base-generated Scene-setting Topics in context. We can affirm that by investigating the intonation patterns of Topic-Comment structures and their interface with discourse context behaviorally, we answered the following questions:

1. We identified the phono-acoustic signature of a particular type of 'Chinese style' Topic, Scene-Setting Topic.
2. We observed a positive impact on in-context sentence comprehension of a clear prosodic and intonational marking of the syntactic structure of Topic-Comment articulation.
3. We demonstrated that the impact of Topic referent saliency in discourse context has no impact on sentence comprehension by native speakers.
4. We put forward that in absence of the natural PB marking of Scene-setting Topic-Comment sentences acceptability judgments are made on word-order grounds and are not affected. In other words, the Frame-setting predicational relation of Gapless Topic-Comment constructions and the contextual integration processes it requires, are not by disrupted by the absence of PB in Mandarin, they only show to be mildly impacted.

4.5.1 This chapter in one page

The prosodic pattern of T-C articulations

The main findings of this chapter are two-folds.

Firstly, the question of the prosodic marking of Topic structures in Mandarin Chinese has hardly been addressed by the linguistic literature. This chapter of my dissertation aimed at providing a start to fill this gap by giving an acoustic-phonetic speech analysis of the prosodic pause intervening between Topic and Comment.

The phono-acoustic analysis of the oral corpus provided a starting point to describe the prosodic signature of all Topic-Comment articulations in Mandarin Chinese.

Specifically, our study put forward the following prosodic signature for Scene-Setting Topics:

a. a significant lengthening (79 m) of the last syllable Topic NP;
b. an average pause length of 68 ms between Topic and Comment separating the sentence unit into two distinct intonational units, and
c. a significant 18 Hz of average pitch (F0) difference between Topic and Comment-clause

A psycho-linguistic description of Topic-Comment articulations in context

Secondly, the main result obtained by the behavioral ratings can be briefly summarized by saying that Topic-Comment constructions in Mandarin Chinese didn't show discourse-context licensing effects, giving us an experimental argument to pursue in the direction of defining Topic-Comment constructions as a particular way to structure the predication in basic non contrastive sentence-units.

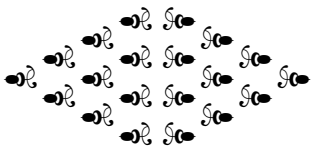
Namely, in contrast with part of the literature claiming that Topics have to be salient in previous discourse context, contextual adequacy judgments do not show any advantage in favor of the association between Topic-Comment articulation and a particular contextual/informational-status of Topic NP in context, thus confirming the basicness of Topic-Comment constructions in Mandarin Chinese.

As for the sentence-discourse interface property of T-C articulation, we can retain that Contextual judgments in different informational-load contexts show that Topic construction resists to context-load experimental manipulation. We interpret this response pattern as confirming the basicness and non-emphatic or contrastive nature of this sentence articulation in Mandarin Chinese.

To put this study back in the broad frame of this PhD, it must be recalled that it was essentially meant to disentangle the contributions of informational discourse-context load from the purely syntactic aspect of sentence structure, in order to provide experimental evidence for the definition of Chinese Topic as a basic sentence articulation.

In conclusion, the following results indicate the fundamental syntactic nature of Topic-Comment construction in Chinese, discarding pragmatics-centered analysis, even under the particular circumstance of Topic-Comment sentence parsing in narrative contexts that should informationally impact on this pragmatic phenomena. These findings are thus allowing us to go a step further into the psycho-linguistic description of Topic-Comment constructions and to continue by studying the cerebral processing of sentence-initial field, the so called Left-Periphery (LP). The following chapter is, namely, devoted

to an ERP experiment (PBTop) where we examined the temporal unfolding of the comprehension and building of Scene-setting Topic-Comment constructions in the auditory mode.



Chapter 5

Context meets the sentence at the Prosodic Boundary: ERP Experiment on Chinese Topic in context (PB Topic)

I do not define time, space, place and motion, as being well known to all.

Isaac Newton, Scholium to the Principia, 1687

Deeply rooted in the thorough linguistic overview presented in chapter 2, the experimental concern of this chapter is namely to address at the level of cerebral sentence processing the question of what are the cognitive mechanisms underlying Topic-Comment articulation and how the Topic-Comment relation in case of Gapless 'Chinese style' Topics is incrementally achieved.

The main idea being that of tracking back the typological claim of Li and Thompson on a brain processing-level, and shed a light on how the brain builds up the kind of Topic-Comment predication implying the interaction of different syntactic hierarchy levels (i.e. CP and IP), instead of subject-predicate relation. We will thus investigate the parsing of this typologically relevant sentence articulation with a particular focus on the cognitive mechanisms of its on-line incremental building in time.

As the sentence is not only the linguistic level at which we can observe the real-time syntactic processes of assembling linguistic building blocks, but it is also the locus of the interface with the discourse context level, an additional aspect of Topic-Comment articulation will retain our attention, namely that of its interface with discourse. By examining the brain evoked responses linked to the temporal unfolding of Topic-Comment constructions embedded in discourse, we will be able to carry out a deeper neuro-linguistic investigation of the Sentence-Discourse interface of Topic-Comment articulations in this chapter.

Hence, the questions we asked through the lens of psycho-linguistic methods in previous chapters are here re-proposed under a new light, that of the online incremental mechanisms for Topic-Comment sentence structure building, using a high-temporal res-

olution imaging method EEG (Electro–Encephalography)[1]:

1. What is the impact on in-context sentence processing of a clear prosodical and intonational boundary, marking sentence syntactic structure? Can we observe at the cerebral processing level a sensitivity to the intonational (i.e. duration of pause) and prosodic boundary marking (Topic syllable lengthening and Prosodic contour) during online comprehension?
2. What is the impact of the saliency of Topic referent in discourse context on the processing of Topic-Comment articulations? when and how is the sentence discourse interface observable during online sentence processing?
3. In more general terms what are the processing mechanisms establishing the frame-setting predication tiding together Topic and Comment?

Why choosing ERPs brain-imaging technique

Electro-physiological data from event-related brain potential (ERP, see Annexes **??**) studies not only already provided relevant insights about the time course of the incremental updating of the sentence comprehension, and permitted to take a further step towards understanding of sentence comprehension, by stretching beyond the simple decoding of the propositional content of isolated sentences and bringing to light contextual effects on sentence processing, and online interactions whit discourse context.

Several electro-physiological studies namely showed that the slightest mismatch between contextual prediction and sentence parsing could result in an immediate brain response recruiting additional processing resources to salvage the on-line interpretation process. This put forward that discourse contextual expectations can guide syntactic parsing.

Moreover, ERP literature has been giving experimental grounding to the fact that prosodic information is essential in spoken language comprehension to guide listeners' syntactic parsing, showing immediate constraint of prosody on online syntactic parsing.

Not only sentence structure building a can be facilitated by prosody, but a increasing number of ERP studies highlight that also the context information integration processes needed to understand a sentence in a broader discourse are affected by prosodic encoding.

Although the time-course of the cognitive mechanisms related to syntax-prosody-context interface are not yet thoroughly identified in the ERP literature, this high-temporal resolution imaging method EEG (Electro – Encephalography) appeared the more suitable one to tackle:

1. Our broad interest: understand the cognitive mechanisms of Topic-Comment articulation
2. Our chapter question on Sentence processing: Investigate the effect of prosodic marking on on-line processing of the sentence-discourse interface, to study to what extent prosodical marking or its under-informative counterpart (with no syntax-prosody mismatch) influence immediate context retrieval

1. We present here the experimental results of Event-Related Potentials (ERP) in the auditory modality designed in Beijing with two colleagues to test Gapless T-C sentences in context : Shaorong Yan (University of Iowa, Iowa City, United States), Yingyi Luo (Hiroshima University, Higashi-Hiroshima, Japan), Xiaolin Zhou (Peking University, Beijing, China). Acknowledgments: We thank Marco Buiatti and Baptiste Gautier for helpful advice on the ERP data preprocessing. Duan Yuan for the help in presentation randomization and for running the experiment for some participants.

3. Our neuro-syntax question: How word-order syntactic information can guide sentence parsing in absence of sentence-level prosodic boundary marking and of morpho-syntactic marking
4. Our contribution to the ERP state-of-the-art: disentangle the prosody-syntax interplay in the early and in late integrative stages of spoken language processing when sentences are embedded in context, and more broadly to contribute to shed light on the interplay of Syntax, Prosody and Context during on-line sentence processing.

Overview of the contents of this chapter

5.1 Syntax, Prosody and Context and their related ERP Effects

We will now briefly consider how the two core experimental questions of this chapter have been addressed by the ERP literature. The first is about the (1) neural basis of processing PB in Chinese Topic-Comment construction, and (2) the second is a more general one about the way sentence-level prosodic boundaries modulate the use of contextual information during speech comprehension, addressing in this way a broader question on the online sentence-discourse interface. Therefore the broader aim of the present investigation is to track the triple interaction of syntax, prosody and context information during online auditory sentence comprehension.

5.1.1 Syntax and the effects of its prosodic cues

Electro-physiological studies of language comprehension have been largely dominated by research issues on semantic and syntactic integration processes. Only in the last one and a half decade, language comprehension models that can be derived from online sentence processing ERP studies have concentrated on auditory language processing and its relying more on sources of information like prosody and context (Eckstein and Friederichi, 2006).

Up to now, the study of the functional poly-valence of prosody information (e.g. lexical-level, structural local-level in sentence, global-sentence-level modality and pragmatic functions) has followed two main research directions in the ERP literature: (1) the study of the processing of intonational boundaries, comprising syntactic Prosodic Boundary (PB) processing in phrasing and attachment, and (2) that of pitch accents/stress (see studies on accentuation), accent linked to the information structure of the utterance and pragmatics (i.e. focal prominence).

Facilitation through prosody

From a cognitive and psycholinguistic point of view, these two directions mirror two acknowledged (structural) linguistic functions/properties attributed to supra-segmental marking, namely (a) prosodic grouping, which is constraining syntactic grouping and affecting hierarchical structure decisions by means of prosodic boundaries and (b) prosodic prominence conveyed by pitch accent intensity, having the role of determining accessibility of a constituent and influencing among others the processing of information structure (see Calson et al., 2009 in *Lingua* for a discussion and for multiple empirical evidence).

In natural speech, linguistic structure building can be facilitated by prosodic (Frazier et al., 2006) and statistical cues, and some underlying neural signatures have been demonstrated (Buiatti et al., 2009 ; Steinhauer et al., 1990). The same has been proven in children, phrasal prosody and function words represent crucial cues for sentence processing in adults (i.e. syntactic analysis) (Christophe et al. 2008) and in children acquisition and to constrains syntactic analysis even in toddlers (Carvalho et al., 2017 ; for a review see Millotte et al., 2013).

The ERP literature offers nowadays a substantial body of electrophysiological evidence in favor of interaction between syntactic parsing, prosody boundary and pitch accent processing, so that it is possible to declare that the processing of prosodic information does influence syntactic processing (see Fraizer et al., 2006 for a review), and that both aspects interact at an early stage in sentence processing (Eckstein and Friederici,

2006). Previous work investigating prosodic violations reported in fact early effects, giving evidence for interaction of prosody and syntax in early stages of sentence structure building.

Syntactic Parsing and Prosodic Boundaries in ERP

Prosody can surely facilitate sentence parsing, but it also conveys a great variety of different functions and for this reason it has been proved difficult to study. The growing interest devoted to prosody focus namely on the rich information it provides for sentence comprehension: prosodic boundary (PB) information is considered to benefit lexical access and syntactic parsing as properly signaling segmentation. And the fact extra-syntactic factors affect or have a tight interaction with syntactic processing, most notably prosodic information and discourse-context informational load is the object of a glowingly amount of ERP studies.

Naturally, prosodic and syntactic information are highly intertwined, and the fact that prosodic information carried by the above mentioned supra-segmental parameters extend longer in time is a challenge to experimental independent manipulations of syntactic and prosodic parameters. Disentangle these superimposed processes during parsing has proven difficult, and the solution to this in the literature has been to use temporarily parsing ambiguous sentences, with all the experimental drawbacks this can yield.

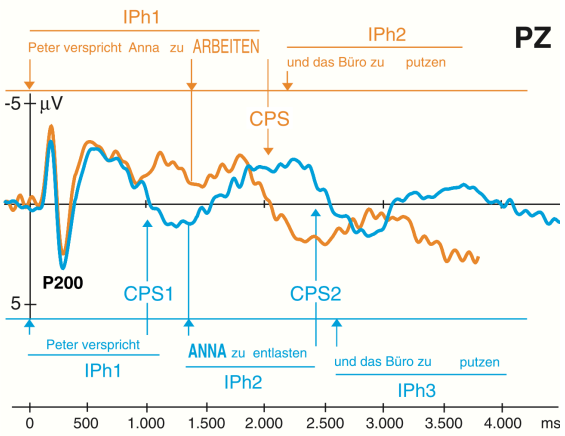


Figure 5.1 – Closure positive shift. Grand-average ERPs of both experiments (n = 40) at the PZ electrode. The wave-forms of conditions A (orange) and B (blue) are superimposed. The word onsets of the sentence examples are aligned to the time axis. Both conditions evoke closure positive shifts at their respective Intonation and prosodic boundaries. Only one shift is observable in condition A, following the second verb "arbeiten", whereas two such shifts occur in condition B, before "Anna"and after the second verb "entlasten". Adapted from Steinhauer et al. (1999).

5.1.1.1 Syntactic ambiguity and prosody disambiguation in parsing

Psycho-linguistic research on prosodic cues to parsing has repeatedly shown that temporal and local syntactic ambiguities can be resolved by appropriate boundary tones. ERPs have proven to be a sensitive tool for the investigation of the time-course of the incremental prosodic influence on sentence comprehension, the actual state of the art offers a multiform general picture of the ERP signature of the Prosodic Boundary detection and of the temporal dimensions of the processing difficulties engendered by prosody-syntax mismatches (see Annexes §E for an overview p.925).

Since Steinhauer and Friederici's (1999) seminal work on PB processing, number of Event-Related brain Potentials (ERP) studies have undertaken to described the neural underpinnings of prosody-syntax interface in adults and children (in adults Steinhauer et al., 1999, 2001, 2003 and much related work; in children: Männel and Friederichi, 2009; Männel, 2013). Thus, establishing the reality of immediate use of prosodic cues in spoken sentence processing to modify syntactic parsing preferences. However, the large majority of these studies large majority of the studies have used *syntactically ambiguous sentences to investigate PB role.* For a comprehensive review

on ERP component elicited Sentential prosodical boundaries ERP see the Table of the different latencies, polarities and scalp distributions in Annexes §E.1 (p.925)

The presence of prosodic boundaries is often accompanied an ERP component the Closure Positive Shift (CPS) in a variety of EEG studies.

The nature of this presumably PB-related ERP effect and its relation to the detection of intonational and prosodic boundaries is actually a fairly disputed Topic in the auditory ERP sentence processing literature. Point of views mainly diverge on the cognitive process underlying to this ERP component. Evidence in support that the acoustic variations of PB rather than the co-occurring morpho-syntactic computation engender the CPS has been reported on one hand, and on the other hand, recent studies report the absence of the CPS for some type of prosodic boundaries (e.g. Holztgrefe et al., 2013). Several accounts for CPS have been advanced, and the CPS is generally interpreted as related to prosodic parsing and alternatively as a effect of information segmentation because it appears at phrase boundaries that give rise to a new (prosodic) parsing.

In this regard, a relatively consistent amount of studies (e.g. Roll and Horne, 2011) have successfully discriminated the differential ERP signature of Right-edge boundaries -i.e. indicating the closure of the clause- and Left-edge boundaries -marking the beginning of a new constituent. They indicated that only Right-edge PB only elicit Closure positive shift (henceforth CPS), while Left-edge PB elicits a N100 in absence of PB and that both edges show an early P200 component. Authors argue that this difference is functionally interpretable in terms of the different roles these boundary tones play in the sentence, namely closure of the clause and the marking of the beginning of a new constituent.

However, such prosodic cues, are not always available, and even when available, they can generally not be sufficient, in that sentence structure building relies also on a listeners' (tacit) syntactic knowledge as repeatedly shown in the literature on the revision process happening when listeners parsing and linearization preferences are contradicted in German (e.g., Hemforth, 1993; Bornkessel and Schlesewsky, 2006/2009; Hung and Schumacher, 2012/2014 in Chinese).

Parsing on syntax only

However, such prosodic cues, are not always available, and even when available, they can generally not be sufficient, in that sentence structure building relies also on a listeners' (tacit) syntactic knowledge. Hence, given our research interest to isolate PB contribution to the processing of the sentence interface with contextual information and to avoid the above mentioned debates on CPS interpretation, we decided not to include ambiguous sentences but to study the contribution to sentence-level syntactic boundary given by prosodical information in a situation where the listener can parse the sentence base on word-order information. Namely, finding at the beginning of the sentence two NPs the listener can deduce quite early on what is the syntactic structure of the sentence.

Based on the behavioral results presented in previous chapter, we built our experimental paradigm on the hypothesis that in absence of a clear-cut prosodic boundary between Topic and comment the listeners would rely on the their tacit syntactic knowledge of Mandarin word-order syntax alone. To further ground this hypothesis, we can consider the results of a very recent article investigating mental representation in the brain of hierarchical linguistic structures in Chinese. Ding et al. (2016) namely provide initial insights into the neural representation of abstract linguistic structures that are,

according to the authors' words, "internally constructed on the basis of syntax alone"[2].

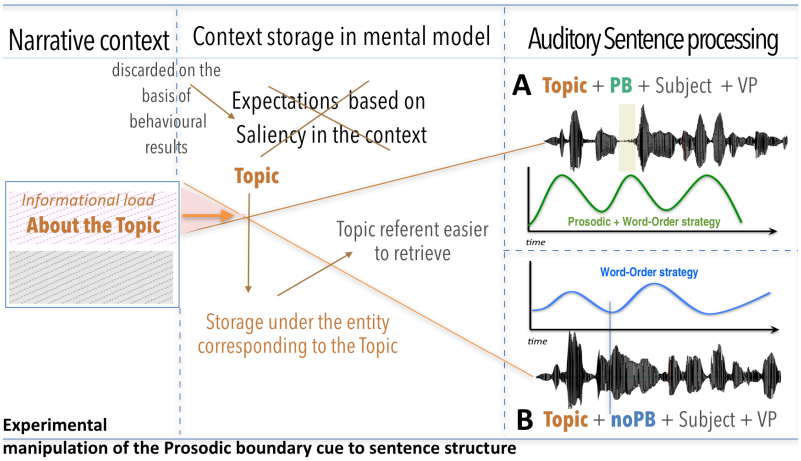


Figure 5.2 – Experimental manipulation of the Prosodic Boundary cue to sentence structure.

We lean partly on this initial evidence to hypothesize that the experimental manipulation of the presence/absence of prosodic boundary and intonational profile of Topic-Comment articulations in Mandarin Chinese will bring listeners'to tacitly rely on their syntactic knowledge of word-order patterns in Mandarin Chinese for sentence structure building and opt for a word-order strategy as illustrated in Figure 5.2 (B) compared to (A).

5.1.1.2 '*Comma-intonation*' and ERP effects from Punctuation

Among the different prosodic boundaries that can cue for syntactic parsing Comma-related effects[3] have been recurrently reported in ERP studies, showing that commas are indeed essential to sentence parsing and syntactic integration[4].

2. Specifically, by showing the entertainment of different neural populations (recorded though intracranial electrodes ECoG) to the different time-scales (i.e. rates) corresponding to different linguistic levels (syllable, word and syntax), the authors give an initial proof of the existence of multiple and probably specified linguistic processing cerebral networks dedicated to the different levels of linguistic representation. In other words, the fact the cortical activity of at different time-scales tracks the time course of different linguistic levels, namely word, phrases and sentences, can be interpreted, maybe in a way too speculative way as the biological possibility to rely only on certain information to perform language understanding.

3. Notice that as comma can signal a multitude of syntactic structures, like intonational units in sentences, isolated clauses or even parenthetic components, if syntactic components in sentences are separated by commas the meaning of the sentence is changed mainly because of a change in syntactic structure.

4. As presented in chapter 2, Medieval copyists introduced the three punctuation marks, dating back to Aristophane of Byzantium (2nd B.C.), to primarily meet the demands of oral reading marking breath marks, that gradually evolved into the comma, colon, and period. The need to make pauses in oral speech will long be the prevailing and most immediate function of punctuation, showing how syntax and sentence-level intonational patterns are tightly linked. Ever since Renaissance, the punctuation is assumed to be basically an orthographic device which signals syntactic patterns to the reader, an in the early 18th century we can read in writing of English or French grammarians like Nicolas Beauzée or John Brightland, statements like : "The Use of these points, Pauses, or Stops, is not only to give a proper

Comma Punctuation marking and its effects have been studied by Steinhauer and Friederici (2001) studied in German sentences. This permitted to point out that during silent reading, commas elicited the above mentioned positive shift, the "closure positive shift"(CPS). This brought Steinhauer (2003) to consider that commas (in visual modality) shared the same mechanisms in the brain as prosodic boundaries during auditory sentence parsing. Importantly the multiple function of commas was attested by the fact that inserted commas could elicit different components from the CPS, namely a P600 effect was elicited when the comma marking could rise uncertainty on the sentence structure (Steinhauer, 2003; Steinhauer and Friederici, 2001).

Further evidence on the role of commas during sentence processing comes from eye-movement studies. Specifically, Hirotani et al. (2006) investigated the function of commas in sentences, and pointed out that commas influenced sentence parsing and prosodic boundary and played an important role in the so-called wrap-up effect.

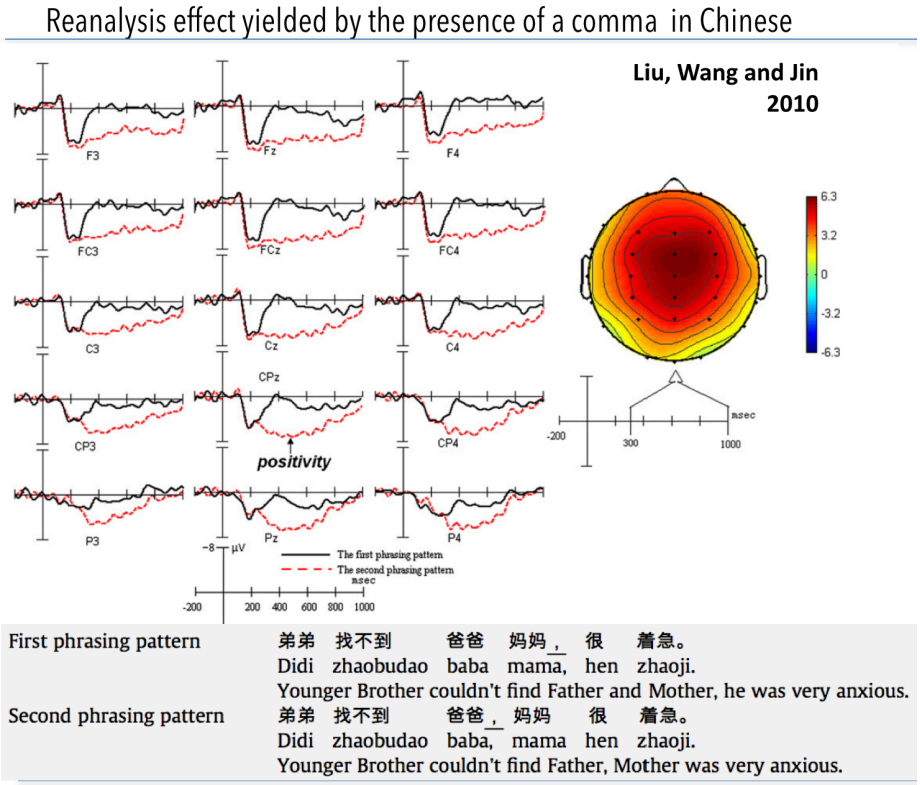


Figure 5.3 – Grand average ERPs in response to the words after the critical commas in the sentences with the first phrasing pattern (e.g. the word '很 hen') and in the sentences with the second phrasing pattern (e.g. the word 妈妈 *mama*). In the time window of 300-1000 ms post-onset, the second pattern gave rise to a positive shift with two peaks at about 400 ms and 650 ms. This late positive shift was similar with the one in task 2. The topographic maps of the ERP differences were plotted in the right part of the figure. The two peaks were involved in the two time windows of 350-550 ms and 600-800 ms respectively. Adapted from Liu et al. (2005).

Time for Breathing: but to avoid Obscurity, and Confusion of the Sense in joining Words together in a sentence" (Brightland, 1711:149).

As for comma marking in Chinese, Li and Yang (2009) studied the perception of prosodic hierarchical boundaries and found that acoustic prosodic boundaries could evoke the CPS effect, and this effect was modified by the variations of acoustic stimuli. But what remains still unclear about the relationship between the ERP effects evoked by prosodic boundaries and punctuations in Chinese, is if the effect of PB takes place in early sentence structural decision eliciting P2 component or LAN effect (100- 500 ms)[5], or if it is taking place later as a syntactic analysis effect (e.g. P600), as shown in Figure 5.3, where the ERP results (A) by Liu et al. (2009) reported a reanalysis effect of the two sentence structures in (B).

In this regard a study by our colleague Luo Yingyi, reported a P2 ERP component evoked to the right of PB by the presence of PB relative to absence of this PB (Li et al., 2010).

5.1.1.3 Prosody and sentence-level syntactic hierarchy

The studies presented until now have investigated sentences either with locally ambiguous syntactic parsing or with a single and grammatically required prosodic boundary, which were both to be interpreted strictly locally as the signal of the end of the current syntactic unit. However, prosodic information can also carry non-local sentence-level hierarchy marking, and other studies provide evidence for the interaction between syntactic parsing strategies and sentence-level prosodic contours -comprising the whole sentence-like question intonation.

These studies suggest that the global pattern of prosodic phrasing is critical for sentence comprehension and observed that a number of different ERP components are actually linked to these processes, namely RAN, P800, etc.

For instance, Astésano et al. (2004) showed that a prosodic mismatch (in a French statement ending with an F0 pattern typical for questions) can elicit a positive deflection with a peak at around 800 ms after the onset of the prosodic pattern (P800). This P800 component differs greatly from the CPS. First, the CPS is generally functionally linked to the processing of more local prosodic boundaries, whereas the P800 is elicited by more sentence-level prosodic contours. Second, contrary to the CPS, the P800 is a narrow peak, and, third, its topographical distribution shows it is a left-lateralized ERP response, which further differentiates it from the typical centro-parietal distribution of the CPS.

Focusing on the sentence-level prosody, Eckstein and Friederici (2005) investigated prosody marking of sentence-final intonation. When the prosodic manipulation occurred at the final word, a right anterior negativity followed by a late positivity (P600) was observed. Interestingly, the right anterior negativity was present independent of sentences' grammaticality, and the P600 was not, which was only present for straightforward prosodic and syntactic violations, but increased for the combined violations (as is to be expected for a a late positivity component). This suggests that the right anterior negativity, and not the P600, should be considered as a pure prosodic effect.

Moreover, in a follow-up study Eckstein and Friederici (2006), the late integrative time-window for prosody-syntax interaction (i.e.late-positivities) during sentence comprehension was replicated from a first study Eckstein and Friederici (2005).Interestingly,

5. Männel et al. (2009) reported namely a Fronto-central negativity component, evoked in presence of inconsistency between PB and syntactic structure. Specifically, it was a fronto-central distributed negativity around 430- 630 ms.

this study also reported a second prosody-related component showing an early interaction of prosody and syntax. A more fine-grained distinction was done between early effects present in the syntax-only condition showing left-lateralized negative polarity component, and bilateral early negativities present only in case of in the case of both prosody and syntax mismatch. From these studies we see that syntax-prosody interaction can be observed both in early and late time-windows.

Taken together these findings converge in showing important variability among studies but mainly that different prosodic boundaries do elicit different ERP components, especially in case of sentence-level syntactic information.

5.1.1.4 Revision processes elicited by the manipulation of PB

In case of mismatch between prosody and the syntactic structure in which the parser was engaged when meeting upon the PB, the ERP literature has described a wide range of repair processes. For example, Eckstein and Friederici (2005) found that a mismatch between prosody and syntactic structure at disambiguating word can give rise to an N400 component, while a combined prosodic/syntactic violation elicits a P600 effect. Similar findings were reported by Männel et al. (2009), who observed a Fronto-central negativity component evoked around 430- 630 ms in presence of inconsistency between PB and syntactic structure. In syntactically correct sentences, a prosodic manipulation of the penultimate word gave rise to a late centro-parietal negativity resembling the classical N400 component, which was interpreted as a correlate of lexical integration costs for the prosodically unexpected sentence-final word arriving too early. Further comparisons with syntactically incorrect sentences revealed that this effect was dependent on the sentences' grammatical correctness.

A finer-grained distinction was tested by Pauker et al. (2011), who compared the ERP responses to missing versus incongruent prosodic boundaries. This study reported a stronger biphasic repair effect (i.e. N400 and P600) for incongruent PB. Based on the observed ERP response patterns the author formulated the Boundary Deletion Hypothesis (BDH), stating that it is more costly to delete an existing PB than mobilize the repair processes for a missing one. This processing cost-related evidence represents an important cue to understand the cognitive processing hiding behind the CPS and could arguably be interpreted as evidence for the fact that CPS mirrors the use and integration of the syntactic significance of PB, rather than the pure presence of a PB. In conclusion, this last piece of evidence has been guiding our predictions: the absence of PB in our experiment would not yield the complex repair processes reported in the ERP literature, and the typical word-order distinguishing the Topic-comment, plus the semantic characterization of the Topic NP in Scene-setting constructions, would trigger sentence-level structure building in absence of PB.

5.1.2 The sentence-unit embedded in Context and Contextual effects

While the vast majority of previous ERP studies have highlighted the importance to understand the precise online timing of the prosody-syntax interaction, contributing to identify the timing at which both prosodic and syntactic cues are processed online, and demonstrating how prosody can influence syntactic preferences using context-free sentences, some studies explored the influence of context and prosody on sentence processing by embedding sentences in different pragmatic contexts (e.g. Kerkhofs et al.,

2007; Toepel et al., 2007).

Improving our actual understanding of online auditory sentence processing essentially needs to tackle the complex issue of when and how during sentence understanding, prosodic information and context information do interface. This question is of course tightly linked to our now familiar research question about the cerebral representation of the sentence-discourse interface.

It is now clear for the reader that to build a coherent representation of the sentence-unit, the listener not only has to retrieve and integrate in the sentence lexical and syntactic information or world knowledge stored in long-term memory, but crucially needs discourse contextual information too, which may also involve a further complex mechanism of building a representation of the ongoing discourse itself.

More and more context-driven and top-down processes are now explored by the actual literature, which gradually attributes a role to these kind of predictions in the architecture of sentence processing (see Steinhaurer and Drury, 2013).

5.1.2.1 Context-driven prediction on the sentence-unit

ERP studies have provided initial evidence that discourse context preceding the sentence can result in predictions about syntax-related features. Early studies on these issues, like Van Berkum et al. (2005) showed that context provides grammatical gender expectations, while Lau et al. (2006) showed that the ELAN component elicited by a word category violation, is actually modulated by context-driven expectation for a particular word category in the relevant place in the sentence. As for word-level expectations, DeLong et al. (2005) showed context-driven prediction for a the word form of the next word, and these authors interpret this as a mark of a pre-activation during incremental sentence comprehension.

Pursuing this research direction, some studies asked to what extent context driven expectations/predictions cue for a given syntactic structure or at least facilitate the processing because of this possible pre-activation, namely one of the main issues of our experimental paradigm.

For instance, several psycho-linguistic and ERP studies have also revealed that prior context causes language listeners to generate expectations on the particular informational status of discourse entities which would in turn generate expectations on pitch accent of certain constituents in the the up-coming sentences (see Heim and Alter, 2006; Toepel and Alter, 2004) and sometimes even the syntactic structure where they could be embedded.

For example, in a reading experiment Cowles et al. (2007) observed that a contextually unexpected word placed as the focused constituent of an it-cleft sentence like “It was the rabbit that ate the salad.” would yield an N400 component, which was interpreted as evidence for the strong expectancy in terms of informational status that an it-cleft construction generates about the the focal informational status of the clefted referent.

fMRI findings going in the same direction been reviewed in chapter 2 (§2.4.4), showing the influence of context on syntactic priming (e.g. Schoot et al., 2014), it can bee briefly said that several studies indicate a right hemisphere lateralization for certain context-based process during sentence comprehension. The Danish experiment we already presented in §4.2.1 by Kristensen et al. (2014a) showed (see Fig. 2.43, p.199) that context gives rise to expectations, and that context has an important facilitating effect on the comprehension of object-initial clauses given the accuracy patterns. Concretely,

they demonstrated that topicalized object-initial sentences are more context-sensitive than subject-initials in Danish.

Moreover, there has been increasing experimental evidence that non-linguistic information produced during the processing of the sentence like gestures or speakers voice information elicit particular ERP responses (see Dimitrova et al. (2016) for online integration of gesture and focus accent; Willems et al. 2007 for co-speech gesture; van Berkum et al. (2008) for the speaker's voice information on sentence parsing; Hagoort et al. (2004) or Nieuwland and Van Berkum (2006) for word-knowledge).

5.1.2.2 Prosody-Context interplay and the Sentence-Discourse interface

The processing of prosody in context requires sophisticated experimental designs to be studied in a setting that would yield as natural as possible sentence processing effects.

Prosodic information influencing Sentence-Discourse interface

A series of studies on Focus accent have underlined the importance of prosodic information to normal processing and integration information, revealing reanalysis processes linked to prosodic incongruity (Dimitrova et al., 2012). Plausibility and expectedness have been show to interact with focus accent (Wang et al., 2011) by generally affecting the amplitude of N400 component.

Context-induced expectations on prosodic realization of the utterance

Of particular relevance for the present investigation are the studies by Toepel, Pannekamp, and Alter (2007) and by Kerkhofs, Vonk, Schriefers, and Chwilla (2007), which demonstrated the influence of context-induced expectations on the syntactic structure and pragmatics of a sentence over the amplitude of the CPS.

Instead of testing isolated sentences or embedding them in short context (1 or 2 sentences), these studies decided to have context-embedded sentences, using question and answer pairs. Through listening to questions and answers prosodic mismatches, Toepel et al. (2007) investigated whether the use of prosodic cues to structure in contextually embedded utterances was the same as in single isolated ones. The informative status of question and answer pairs was manipulated by creating both prosodic mismatches and contextual mismatches.

Similarly, Kerkhofs et al. (2007) found in two experiments that the closure positive shift (CPS) was reduced in size when a prosodic break was aligned with a syntactic break and when context information was biasing for a certain sentence constructions (Sentential coordination versus coordination of two NPs). Thus, showing evidence for the immediate integration of prosodic information with syntactic information in the interaction with the activated contextual information.

Interestingly, in our data we will be able to test whether the modulation of the CPS by context can generalize to situations in which the presence or absence of a prosodic break is in line with the syntactic disambiguation.

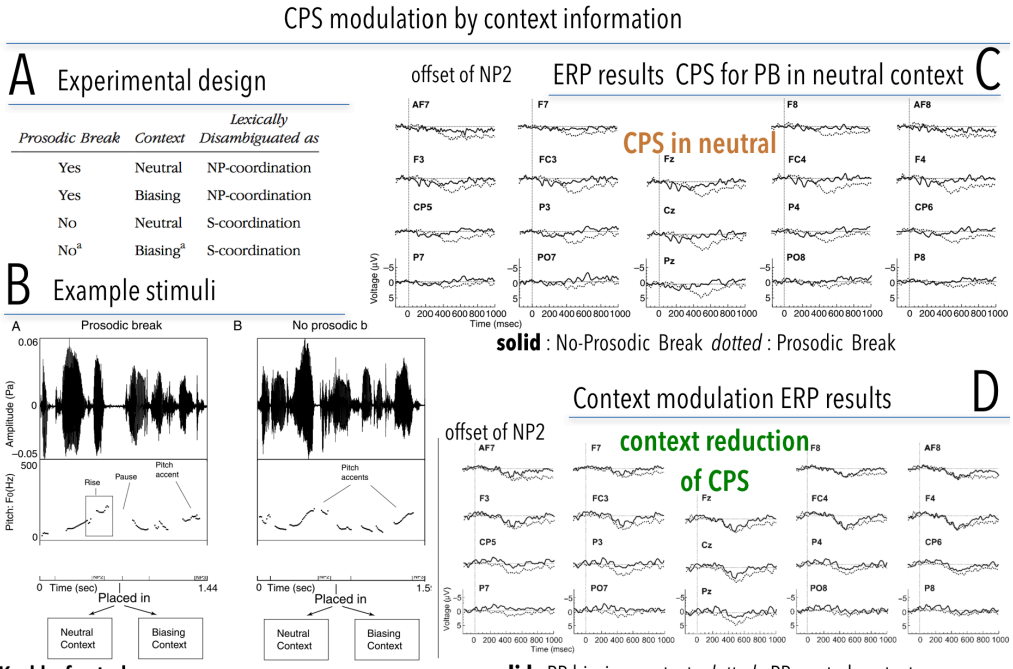


Figure 5.4 – *Lead-in sentence*: Hardly a day after 'Mad Dog' Johnson's release from prison, trouble started again in Gulch. *Neutral context*: After Johnson summoned his men, the largest ranch in the neighborhood was targeted for their raid. *Biasing context*: When they heard the farmer cry for help, the sheriff and the farm adjutant rushed to the ranch. *Target sentence*: *The sheriff protected the farmer and the farm adjutant bravely defended the ranch against Johnson's gang.* Adapted from Kerkhofs et al. '(2007).

5.1.2.3 NPs referential processing in Context

The impact of discourse-level information is not only observable in the expectations it can yield on the incremental processing of sentence-unit, but also in the processes underlying referential assignment to the noun-phrase found in the sentence. Namely, connection to previously introduced information also ensures coherent sentence comprehension. From the literature emerges that Referential processing evokes two temporally distinguishable neural responses, N400 and Late Positivity, showing it relies on two core stages broadly named as referential processing and context updating.

Several ERP studies were able to assess contextual information effects both on syntactic structural parsing and on the referential interpretation of NPs int he sentence. Early studies identified a sustained negative shift (Nref) for referentially ambiguous items at around 300 ms after NP onset (Van Berkum et al., 1999, 2007). Furthermore, this research direction demonstrated that listeners make use of contextual information very early in time, after having heard only the first two or three phonemes, and in many cases before the end of the word. (Van Berkum et al., 2003). Overall, we can say that these findings show that referential processing is highly incremental.

Another series of electro-physiological studies investigated the type of referential de-

pendency (i.e. identity versus inference) and the discourse saliency -intended as the degree of accessibility of an entity in a mental model- of the contextual anchor for referential interpretation of NPs (Burkhardt, 2006; Burkhardt and Roehem, 2007) and found that a broadly distributed negativity peaking around 400 ms was modulated by these factors.

As shown in the below examples demonstrated the electo-physiological evidence for the processing cost as a function of the accessibility of a noun-phrase. The computational demanding integration of the NP (i.e. the bride) yielded an N400 enhanced in amplitude that was interpreted as a cost of 'inferential bridging' of the referential dependency of the 'the bride' Noun Phrase:

(255) Modulation of N400 component by the degree of accessibility from Burkhardt (2006)
   a. Identity relation, reduction of N400: *Tim watched a bride.* THE BRIDE ...
   b. Inference based link to context, + negative: *Tim went to a wedding.* THE BRIDE ...
   c. Lack of connection to context, N400: *Tim met Paul.* THE BRIDE ...

However, several ERP studies have found that N400 effects are not uniquely triggered by the informational opposition between given-new status (Burkhardt, 2006; Schumacher, 2009), but by factors like as referential identifiability or ambiguity, and crucially also by sentence structure (Ledoux et al., 2007; Schumacher and Baumann, 2010; Yang et al., 2007).

Broadly speaking ERP studies showed that the sentence processing system is generally sensitive to contextual saliency of an entity that serves as anchor to establish referential dependency, and that it is constrained by the information encoded in the mental model of discourse level information and that discourse context retrieval is inhibited if the saliency of the a referential anchor is inappropriate[6]. All in all these studies show that the N400 in referential processing is modulated by the effort during referent identification, and more generally it is seen as reflecting expectation-based processes, which are anyway mainly formed on the basis of cognitive accessibility of discourse referents.

Based on these preliminary results one can argue that incremental processes of discourse information matter in the selection of competing referential candidates for pronouns or NPs interpretation -eliciting a sustained negative shift for referentially ambiguous items, the so-called Nref (starting at 300 ms after NP onset). Hence, we hypothesized that discourse context informational load would also play a role in online sentence comprehension.

On the contrary, when the functional contribution of an entity to discourse structure is assessed, enhanced Late Positivity is observed with the increasing updating effort elicited either by the introduction of novel informational units (Burkhardt and Roehm, 2007; Schumacher, 2009) or in coincidence with the need of reorganizing the information in the mental representation of discourse (Burkhardt, 2007).

In sum, the relatively flourishing literature that has been lately emerging on these context-related issues on anchoring the discursive referent in the sentence and inferential

6. Note that ERP investigation on the saliency of an entity in the mental model of discourse information of the listener provided evidence that N400 component is sensitive to saliency manipulations only when the noun-phrase integration happens by inference and not by identity: inferential relations showed a more pronounced N400 with less salient anchors (Burkhardt and Roehem, 2007). Note in our case the referential bridging will happen in an identity condition and in longer context narratives, we then expect to observe this type of negativity even if the relation with the context is not an inferential one but a simple establishment of dependency as in other studies like (Schumacher, 2011 or van Berkum, 2007).

bridging relations between context and sentence-unit, allegedly identified factors that have been proved to influence referential processing: (1) the degree of activation of an entity in the discourse domain, (2) syntactic and sentence structural factors like serial position. To these we would add a supplementary and still unquestioned factor, namely *context informational load*.

In conclusion, during natural auditory sentence comprehension and parsing, context information retrieval and integration are subject to constant dynamic, and what our study wants to explore are the online sentence-discourse interfacial processes triggered by Topic-Comment articulation in the particular case of 'Chinese style Gapless' Topic-Comment articulations. In other words the question sake is when and how listeners relate relate the Topic and comment to wider discourse in the kind of Gapless Topics where no grammatical or semantic direct link between Topic NP and the comment-clause is observable.

Transcending given/new information status opposition, we addressed the issue of the triple online interplay between syntax, prosody and context informational load during online sentence processing. In order to uncover the influence of early triggering of the Topic-Comment sentence-level syntactic boundary by prosodic cue on context bringing and integration processes, we tested two strategies to build the syntactic hierarchy characterizing Topic-Comment structure that activates the sentence-discourse interface property of this sentence articulation, either by Prosodic Boundary (with no accentuation device) or by simple word-order cues. An to highlight the context bridging mechanisms activated by the Topic sentence discourse interface property we decided to test for the possible modulation yielded by prosodic information of the information load access during online sentence processing.

The second critical question of the present study was indeed whether prosodic under-informativity (absence of PB) would modulate the contextual information retrieval and integration, and therefore indicate a clear and direct link between prosody and contextual information load processing during online auditory sentence parsing.In next section we will present our model of the online parsing mechanisms of Topic-Comment articulation.

5.2 The present study: Design, Materials and Predictions

As the linguistics and intonational-prosodic contours of the stimuli were thoroughly presented in previous chapter (§4.3 , p.430) we will present the relevant feature of the experimental conditions from the point of view of incremental cerebral processing, and introduce the three manipulated parameters in the ERP experimental design, thus showing how our experimental material is suited to answer the research questions raised above and is actually grounded on a detail set of hypotheses on how the online incremental processing of the retained Topic-Comment construction unfolds.

The discussion will focus on three aspects of information structure: (i) the neural processing of syntactic boundaries conveyed by prosodic ones or by word-order (ii) the cognitive processes underlying the processing of informational load and their modulation by the presence of a Prosodic Boundary, and (iii) the role of discourse context salience of Topic referents in the interpretation and possible anticipation of topicality.

Design and Materials

Materials All critical sentences (see Examples in 5.5) began with a NP/DP that depicted the location of the event described by the comment sentence (i.e. the syntactic Topic of the sentence). The Topic NP was followed by another NP (i.e. the Subject of the comment clause), and a predicate (Verb), followed by a complement. Critically, the verb in critical sentences did not appear in the story, it was a synonym verb that could be easily deduced from the story content.

Three versions of each critical sentences were recorded by a native speaker:

- 1) one with a PB between the Topic NP and the Subject NP of the Comment clause, the PB Condition,
- 2) one without a PB, the NoPB Condition, see Fig. 5.5, and
- 3) one with a conjunction 和 *hé* 'and' to elicit a (morpho-syntactic) violation, the He Condition.

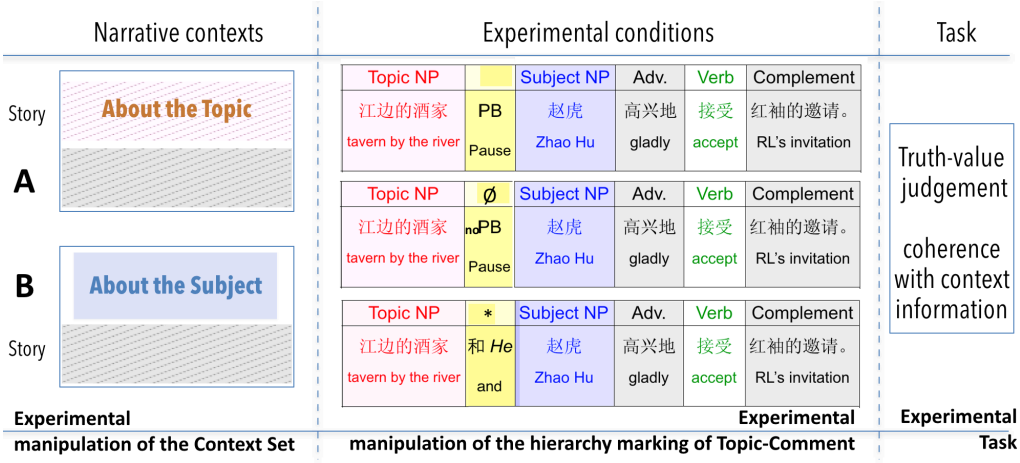


Figure 5.5 – Experimental design and Examples of Critical sentences.

As shown in the above Figure 5.5, we designed our experimental materials mainly manipulating three parameters in order to answer to the three questions raised in previous section about :

- (1) the sole contribution of prosodic boundary information to the parsing of non-structurally ambiguous sentences,
- (2) syntactic grammaticality vs. prosodic under-infomativity, and
- (3) the contribution of context informational load about Topic ans subject NPs discourse referents during online sentence comprehension.

Fillers Conditions

In addition to the experimental conditions, filler conditions were included, exactly 15 out of 25/27 sentences.

Fillers comprised correct sentences, in which the conjunction was followed by a morpho-syntactically correct noun phrase, to ensure that participants could not automatically determine the sentence grammaticality hearing the conjunction. Other fillers conditions consisted of syntactically different sentences that were all grammatical in order not to explicitly draw subjects' attention to grammaticality by adding ungrammatical filler sentences and thereby increase the probability of sentences including a morpho-syntactic violation. These were in turn constructed to fit or not with the content of the two different narrative contexts.

5.2.1 Contextual information - Predictions

As shown in the brief review of the literature, a stable view on Prosodic Boundary related ERP components and their functional significance is challenged by the heterogeneity of the results. Additionally, considering the variable linked to contextual embedding further increments the complexity of the picture. However, we can still take as starting point at least some general conclusions to which previous ERP studies arrived to build on them and structure of our experimental design and paradigm on them and to further sketch some general predictions.

As a detailed description of the two different experimental narrative contexts was given in previous chapter in §4.4.2, we will concentrate here on clarifying the experimental arguments that brought us to manipulate discourse context in this ERP design.

Our model of in-context Topic-comment parsing mechanisms

By embedding Topic-comment sentences in context, our aim was mainly to investigate the online processes yielded by the sentence-discourse interface of this sentence articulation, and to test how previously presented contextual information in a fictitious narrative scene – giving saliency and higher accessibility to the Topic referent (in context A) or to the Subject's one (in context B) – would actually affect the comprehension of basic scene-setting Topic-Comments acoustically marked by the typical "comma intonation" of Topic (i.e. PB) or flat reading (i.e noPB).

The experimental setting and processing model described in Figure 5.7 will allow to uncover if ERP effects can confirm behavioral measures that Chinese Topic-Comment articulations are not licensed by a particular discourse context, but are syntactic articulations that structurally activate the sentence-discourse interface as thoroughly demonstrated in our linguistic overview in chapter 3.

Specifically, we will ask if Gapless 'Chinese style' Topic-Comment articulation, given their basicness inside the Chinese Linguistic system, are sensitive to the saliency given to the Topic referent in context .

Truth-value judgments and acceptability judgments indicated that Topic-comment articulation aren't really constrained by certain licensing context but probably only facilitated by the presence of a PB, now on-line electro-physiological measures will inform us about:

1. The processing mechanism of the frame-setting relation between Topic and comment in Gapless Topics,
2. How the sentence-Discourse interface property of this syntactic articulation unfolds during incremental processing, and

3. The processing costs linked to the absence of PB usually marking the Topic-Comment articulation.

It should be noted that based on linguistic literature reviewed in previous chapter (cf. §3.1.3.5, two main mechanisms can be hypothesized to take place during in-context Topic-Comment linguistic phenomena. The first linked to the information storage of the narratives content, following Reinhardt (1982), and the second considering some of the interfacial links that the sentence-discourse properties of Topic-comment articulation structurally carry.

This is mainly why we built a two fold manipulation of context: (1) the first is linked to the saliency that the discourse storage encoding in the mental model gives to the Topic or the Subject discourse-referent, which is meant to influence the cost of establishing the referential dependency of Topic or subject NPs (i.e. referential integration cost). And (2) the second is linked to the context informational-load, and is meant to highlight the interfacial linking mechanisms between sentence and discourse in Topic-comment articulations. Crucially, these two views lead to different predictions during online sentence processing that we graphically represent in the two right hand boxes of Figure 5.7.

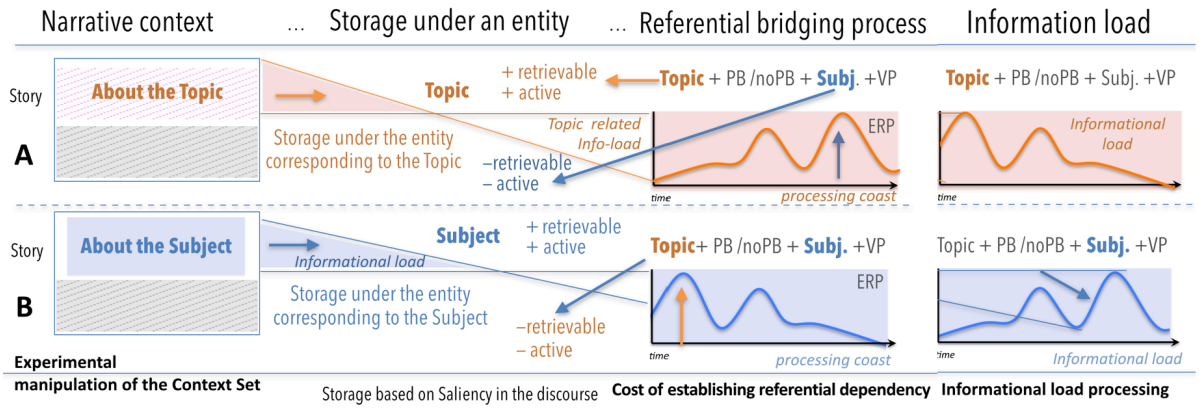


Figure 5.6 – In-context processing model of 'Chinese-style' Topic-Comment sentence. Manipulation of context can be observed under two different points of view, one linked to the saliency attributed by the discourse storage encoding in the mental model of the narratives, which is meant to influence the cost of establishing the referential dependency of Topic or Subject NPs (i.e. referential integration cost). And the second linked to the informational-load, which is meant to highlight the interfacial linking mechanisms between sentence and discourse in Topic-comment articulations. Crucially, these two views lead to different predictions during online sentence processing that we graphically represent in the two right hand boxes of the diagram.

5.2.1.1 Discourse manipulation and Topic-comment sentence-discourse interface mechanisms

Informational load

Delineating a model of the on-line parsing mechanisms of 'Chinese style' Topic-Comment sentences, our first interest was to highlight the interfacial linking mechanisms between sentence and discourse-context that characterizes Topic-Comment articulations.

For this reason we decided to manipulate the informational-load discourse-context was conveying about the Topic discourse referents to be able to observe in ERPs the use of contextual information this syntactic articulation yields. Informational load manipulation was namely meant to offer a quantitative observable measure for online contextual information processing (/retrieval) during sentence parsing.

Specifically, given the frame-setting and Aboutness role Topics play in the sentence-unit where they are found we hypothesized that to parse a Topic-Comment sentence one would have to live up and retrieve all the information available about the Topic to be able to understand the Comment and its relation with Topic referent, for this reason we predicted to find an effect of information-load on Topic NP critical words' the late time windows.

As show in Figure 5.7, our predictions for context-load manipulation is that in order to process sentence interpretation, evaluate its relevance with the Comment-clause and its fit with the discourse (i.e. coherence), the parser would need to live up all the related information to the different NPs in the critical sentence.

Hence, we will understand the notion of *context effect*, as corresponding to ERP effect measured at critical words, which are actually due to the integration of the informational load of preceding discourse context (plain vs. rich) carries about a given critical word. In short, *context-load effect observed in a given time window will show that at this step of incremental processing of the Topic-Comment sentence, the sentence discourse interface 'gateway' is opened and that the contextual information is accessed to parse the sentence.*

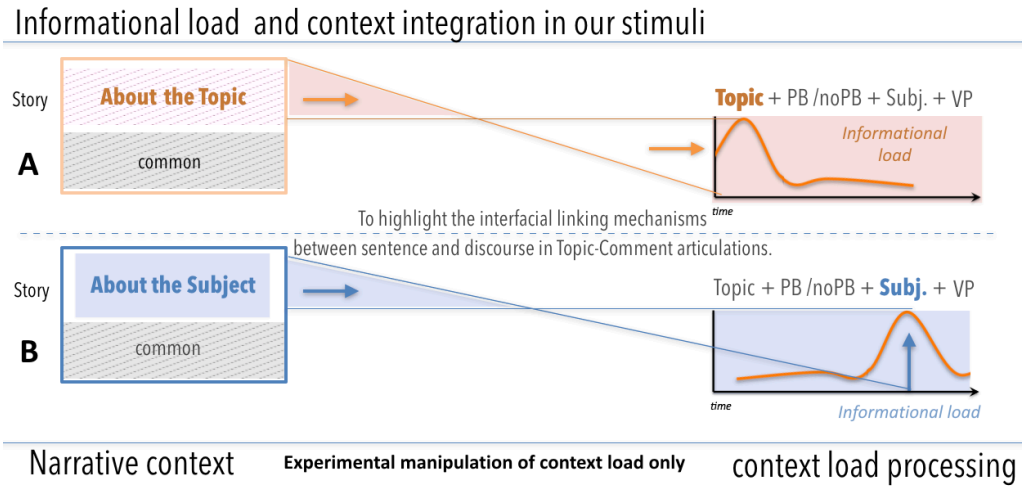


Figure 5.7 – Informational load about Topic referent and subject referents is differently distributed in the two narrative contexts (A more about Topic versus B more about Subject) and we predict this will be observable in ERP context-load effects as shown in the left hand part of the figure.

No Given - New information distinction Importantly, aiming at having direct evidence of the sentence-discourse interface mechanisms and for the use of contextual information load it yields during on-line sentence processing, we avoided Given - New information distinction to take place in our critical sentences.

Crucially, earlier findings in the ERP literature show a different processing of background information and new information (Dimitrova et al., 2012; Wang et al., 2011; Li and Ren, 2012), we avoided to have any new information in our critical sentences to control for any informationally driven focus effect. The selected sentence structure, in fact, has a default non-contrastive reading in Chinese.

Hence, transcending the traditional given-new information opposition (Lambrecht, 1994) , we did not manipulate the informational status (like in Toepel et al., 2007) of the different NPs in our critical sentences, but only their context information load (given by the narratives about NP1, NP2). In our experimental materials, we not only fixed as given the dependency relation between incoming information in the critical sentences and prior information in the narratives (i.e. all the NPs in the critical sentences were could be linked with accessible referent in prior context), but we also selected Topic-Comment syntactic structure, because it structurally requires the referred entity for the Topic NP to be given information (Chafe, 1976 ; Li and Thompson, 2976). Experimental evidence for this well-established Topic informational status linguistic feature was already given by Hirotani and Schumacher (2011), who argue that Topic-marking in Japanese represents a low-ranking salience feature during discourse linking stage (i.e. following the authors' definition the processing step that is hypnotized to sub-serve linking attempts with discourse representation).

For all these reasons, we avoided manipulating the particular referential relation with context in order to minimize information status interference with sentence processing, and to tentatively isolate the effect of sentence PB on discourse domain linking and the amount of information revived by online sentence-discourse interface mechanisms.

Moreover, this model predicts that the syntactic PB marking would have an effect, or would interact with the discourse integration process based on the available amount on information in the context at the verb or earlier at the subject position. Crucially, we see the PB as an early trigger of the sentence-discourse interface and its contextual information integration processes.

5.2.1.2 Discourse context manipulation and in-context referential interpretation

Discourse information storage However, one could argue that the manipulation of information-load would also impact simultaneously the saliency of the referential anchor of Topic referent, agreeing with this position we actually predicted that informational load would attribute a particular role to the Topic and the subject discourse-referents in the two different context narratives.

Delineating a model of the on-line parsing mechanisms of Chinese style Topic comment sentences, The discourse information storage framework proposed by Reinhardt (1982) for Topicality offered central contribution (cf. §3.1.3.5): Topic Phrase referent is the entity under which discourse information is stored:

> A useful metaphor for the procedure involved [...] is the organization of a library catalog [...]. The propositions admitted into the context set are classified into subsets of propositions, which are stored under defining entries. At least some such entries are determined by DP-interpretations. DP sentence-Topics, then, will be referential entries under which we classify propositions in the context set and the propositions under such entries in the context set represent what we know about them in this set [...]. (Reinhardt, 1982:24)

Following her framework where the Topic Phrase referent is the entity under which discourse information is stored, we assume in our in-context parsing model that, in the rich context narrative, the Scene-Setting Topic NP will assumes the role of being the entity under which discourse information is stored, which would make it more salient an [+active] referent and thus easier to address and retrieve to then comprehend our critical T-C sentence if embedded in the rich context.

As shown in Figure 5.8, following Reinhart (1982), we assume that half of the narrative's information would be stored under the subsequent sentence's Topic referent in context A, and under the subsequent sentence's Subject referent in context B, the gray part of the narrative being common to the two discourse context).

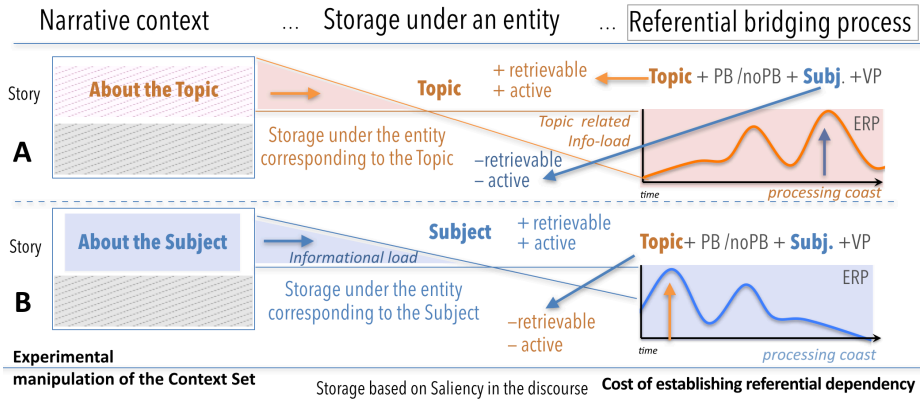


Figure 5.8 – In-context processing model of 'Chinese-style' Topic-Comment sentence. Manipulation linked to the saliency attributed by the discourse storage encoding in the mental model of the narratives, which is meant to influence the cost of establishing the referential dependency of Topic or Subject NPs (i.e. referential integration cost), which leads to the predictions during online sentence processing that we graphically represent in the right hand of the diagram.

As already put forward in the literature review §5.1.2.3, the ERP studies that investigated discourse saliency of the contextual anchor for referential interpretation of NPs, understood it intended as the degree of accessibility of an entity in a mental model (Burkhardt, 2006; Burkhardt and Roehem, 2007) and found that a broadly distributed negativity peaking around 400 ms was modulated by these factors.

ERP studies showed that the sentence processing system is generally sensitive to contextual saliency of an entity that serves as anchor to establish referential dependency, and that it is constrained by the information encoded in the mental model of discourse-level information and that discourse-context retrieval is more difficult if the saliency of the a referential anchor is inappropriate. In fact, Hirotani and Shumacher (2011) showed in Japanese Topics that the strength of the N400 elicited at critical NP varied as a function of the degree of accessibility of the corresponding referent in the discourse representation.

Moreover, an ERP study by Hung and Schumacher (2012) investigated how Chinese Topic affects referential processing incrementally. The authors explored the nature of context-induced topicality in Chinese discourse processing in question-answer pairs consisting of Topic and non-Topic questions followed by different continuations, yielding either Topic-continuity versus Topic-shift or novel-Topic.

Asking whether referential expressions in sentence-initial position – hosting Topic information in Chinese – are subject to different processing decisions compared referential expressions in other sentential positions, this study was able to show that the electro-physiological correlates of Topic position in Chinese the sentence-initial position is subjected to distinct neural demands from non-Topic positions. Namely, their result revealed that referential assignment at the two positions is indeed guided by different informational features: by topicality for sentence-initial Topic and by the given-new distinction for sentence-final object.

Hence, based on this study we expected that referential processing in our paradigm too would be showing position-specific constraints, and that at the Topic and at the subject position would be yielding different processes. As these authors found that the referential processing of sentence-initial position was guided by contextual topicality in Chinese we expected that the topicality of the narratives would either shape expectations or influence internal discourse structuring in the mental model as argued through the above discussion on the encoding of the narratives' informational-load under an entity.

Therefore, as illustrated by Figure 5.8, we predicted that the referential processing of the two sentence NPs (Topic and subject) would be influenced by contextual saliency and encoding in the mental model constructed for the narratives' content, yielding more activation for referential assignment of subject NP in sentences embedded in Topic-related narratives (story A), and vice versa in the case of embedding in subject-related story (B).

In sum, two central considerations can be drawn from the emerging picture of the ERP literature on discourse linking issues. The first is linked to two structural factors that influence the referential interpretation (e.g. Burkhardt and Roem, 2007): syntactic role (condition PB and noPB), degrees of activation in the discourse domain (Context A or B); and the second is that contextual Saliency helps individualizing/anchoring the discursive referent in the sentence (Schumacher, 2011).

5.2.1.3 Discourse manipulation and contextual licensing of Topic-comment sentences

The second mechanisms we wanted to test for, at the cerebral processing-level, is that of possible discourse-driven exceptions that the rich context (A) could yield on upcoming Topic-Comment sentences and their intonation contour.

Although this kind of effects were observable in the behavioral acceptability ratings and only leniently present in the truth vale judgments reported in previous chapter, the ERP literature reports several context-induced expectations on sentence structure and its information structure.

We hypothesized that no clear **context-expectation effects** would be implied by the different contextual information load given by the two contexts, but a modification of the saliency intended as the sum of factors that influences the degree of accessibility of of an entity in the mental model of the narratives.

In other words, in contrast with previous studies manipulating salience, using pitch accent and informational status in question and answer pairs (Magne et al., 2005; Dimitrova et al., 2012), the fact that the Topic NP linked entity in the discourse would be more or less described in details in one context (A) relative to the other narrative (B), would not result in a different expectancy in either syntactic or prosodic terms but in a difference of saliency and storage under a the Topic entity or under the Subject entity, as showed in the following diagram.

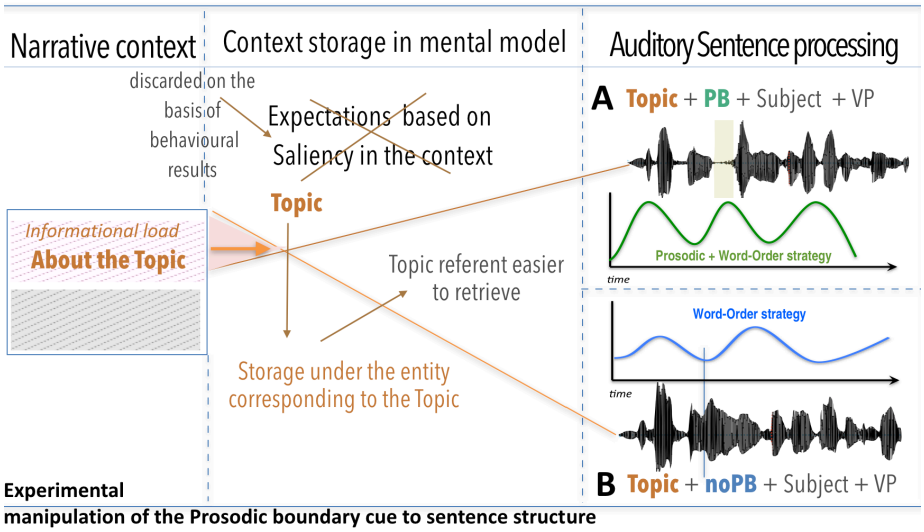


Figure 5.9 – Although discarded by the descriptively small (but statistically significant) difference between PB and noPB truth-value coherence ratings presented in previous chapter, the option of context licensing of Gapless Scene-setting Topic-comment sentences in Chinese will be overlooked in ERPs more in terms of context storage encoding in the mental model of the listener, than in terms of discourse expectations.

Given the linguistic facts and analysis presented in chapter 3 about Topic-comment sentence articulation's basicness inside the Chinese linguistic system. The different amount of information on NP1 would not induce, form a linguistic point of view, any specific expectancy of licensing neither regarding syntactic structure nor the presence of Prosodic Boundary. In contrast, with other languages like German, where it is claimed that a contextually introduced Aboutness Topic is expected to be found in sentence-initial position, whereas a 'neutral' context would not generate such syntactic expectation (Buering, 1999).

Three alternatives and their predictions

No context-sensitivity As for the tested base-generated Gapless Topic-comment sentences, three scenarios can be hypothesized. The first, the simplest, would show that Scene-Setting Topics are behaviorally not context sensitive and represent a basic structure type.

Repair processes in plain context due to Context-sensitivity The second will hold if basic base-generated Gapless Topic-comment sentences are observed to be context-sensitive (i.e. the contextual saliency generates expectations for Topicality) PB conditions inserted in a context where the Topic in less salient should show some additional processing effects

No repair processes in plain context only sentence-discourse interface mechanisms A third alternative is also possible, but not linked to a licensing role of the context, we could actually observe in on-line processing that the saliency of the Topic referent simply facilitate

the retrieval of information about the Topic referent and increase the comprehension, but not the rapidity of truth-value judgments of the Topic sentences embedded in a narrative context. A slowing down of the truth-value judgments would be expected in our informational load manipulation, because the amount of information to be revived to perform the judgment in the context where the informational load on the Topic referent is heavy would have a processing cost, and therefore slow down the decision.

5.2.2 Experimental Task and its predicted effects

Importance of the task in auditory ERP studies

A recent study by Dimitrova and colleagues (2012) gave empirical evidence to the claim that a certain number of the response variation reported in the literature on prosodic processing were actually task-dependent, highlighting in this way the critical role of the choice of task to investigate prosody related components.

In order to avoid to have task-dependent ERP prosodic components and the risks highlighted by Dimitrova et al. (2012), we decided to give our participant an novel task that would avoid the participant to focus on prosody or syntax parameters manipulated in the experiment and would make them rate the congruity of the sentence with the previously read shot narrative.

In the truth value judgment based on discourse congruity, the participants were asked to judge the degree of relatedness between the sentence and the proceeding narrative (i.e. the fit of context and target sentence) on a 5-point Likert scale with 1 representing the most related and 5 representing the least related.

5.2.2.1 Topic-comment mechanism and truth-value judgment

As exposed in §4.4.3, our choice for the experimental task was based on the analysis given by (Kiss, 1995 and Erteshik-Shir, 1997/1999) stating the Topics are used as a basis for the evaluation of the truth-value of the utterance. Importantly, as illustrated in Figure 5.10, from the point of view of online processing of Topic-comment articulation, a truth-value task had the advantage of possibly maximizing or just preserving the mechanism of truth-value judgment that appears to be one of the core definition of the Topic-Comment sentential articulation (cf. section on Topic notion, §3.1.1).

Specifically, in our paradigm the truth-value judgment implied evaluating the propositional coherence of target sentences with the narrative context (i.e. fit between context content and target sentence's content) and it was also obviously meant to evaluate the comprehension of the stimulus sentences in context. Furthermore, we predicted that by engaging in the decision about the rating the coherence of the sentence with the prior short narrative through a binary choice, "Correct" (Coherent) and "False" (Incoherent), the ERP wave-forms would also reveal to what extent the absence of Prosodic Boundary would impact online the context information bridging processes and hinder the sentence-discourse relation.

Importantly, we tried to constrain the explicit contextual-discourse judgment to happen during the comment-clause by constructing our sentence materials in order to present new lexical material only at the verb time-window (i.e. synonyms of the verb presented in the story). We hypothesized this would force the participants to make the judgment only after the verb is encountered. Namely, in informational terms, the first two noun

phrases were background information and the verb was either a contextually incongruent verb or a synonym of the verb appearing in the context narrative. Given this setting, we expected to find some task-related ERP effects at verb position as this truth-evaluation of propositions is essentially determining the truth of the predicate in relation to the Topic (Erteshik-Shir, 1997/1999).

The comparison between PB and noPB conditions ERP responses will be informative about the impact on in-context sentence processing of a clear prosodic and intonational marking of Topic-Comment syntactic structure.

5.2.2.2 Experimental strategy to tackle the time of the Sentence-discourse interface

Given these considerations, our experimental task could probably be considered as having the draw back of leading the participants to be more focused on contextual information, thus yielding stronger context-related ERP effects, as they had to to prepare an answer to the truth-value judgment following each experimental item, asking them to evaluate the sentence's propositional coherence with narrative discourse information.

However, what could be seen as a drawback was actually an advantage for us. We purposely manipulated this task related contextual influence and the information load dimension of narratives, precisely to enhance, highlight and be able to observe the moment in time where the integration processes linked to the sentence-discourse interface property of Topic-Comment articulation take place as shown in Figure 5.10.

This will hopefully bring to light the parsing strategy of Gapless Chinese style Topic-comment sentence and the moment at which the Comment is actually related to the Topic in Gapless 'Chinese style' Topic-comment sentence.

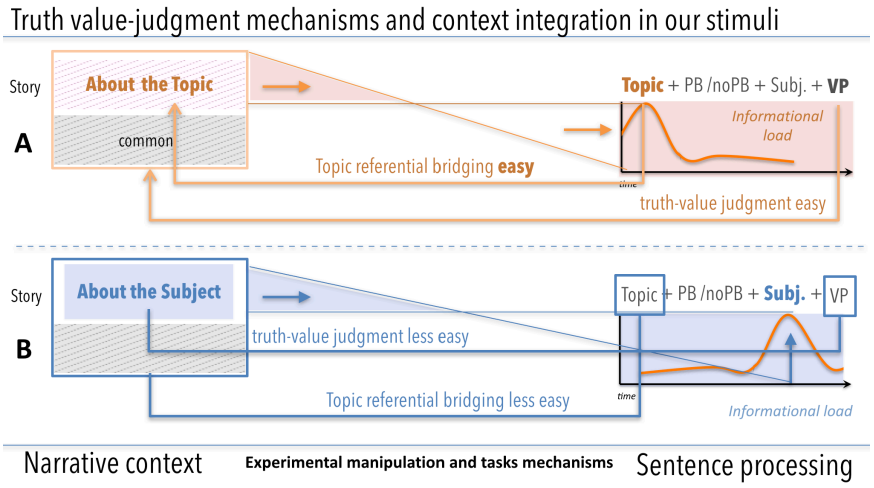


Figure 5.10 – Experimental manipulation and truth-value judgment tasks mechanisms.

5.2.3 Prosodic Boundary - Predictions

5.2.3.1 Prosody manipulation

Our first experimental manipulation aimed at making the syntactic structure of Topic-comment more or less evident in auditory presentation, by testing the sensitivity to the duration of pause marking and prosodic boundary contours during the comprehension of the Scene-setting Topic construction we selected Topic sentence comprehension.

The kind of Prosodic boundary (PB) we described in previous chapter has been identified in the psycho-linguistic and ERP literature as a type of prosodic information benefiting syntactic parsing and lexical access. Critically, the Prosodic Boundary in the Topic-Comment construction is just aligned with the syntactic boundary marking this sentence-level syntactic structure, and is not related to parsing disambiguation. Such sentences can only have one possible interpretation an no contrastive interpretation[7].

From a syntactic point of view Chinese Topic-Comment constructions encode the Topic function by simple linear word-order position, this minimal syntactic encoding assigns to the sentence-initial constituent features sentence/discourse interface properties.

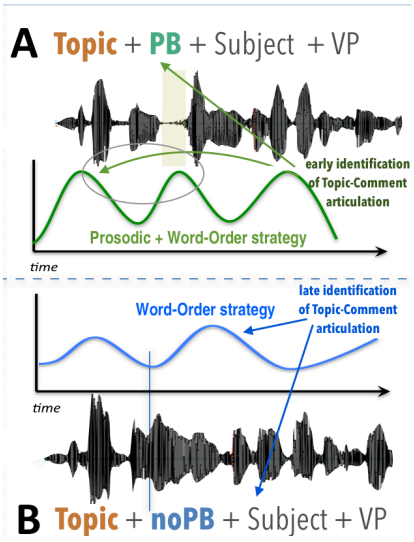


Figure 5.11 – Parsing prediction for presence of Prosodic boundary in (A) or its absence in (B).

Our psycho-linguistic investigation of the the behavioral effects of this boundary showed that it not only overtly marks sentence constituents' segmentation by a pause separating the utterance into two distinct intonational units, but it also marks the establishment of sentence-level hierarchy between the Topic and Comment, namely the CP sentence-discourse interface layer and the IP sentence domain, by a particular prosodic contour.

The main point of manipulating the prosodic cue was to be able to observe to possible parsing strategies that can be adopted in Mandarin Topic-comment comprehension, on guided by two syntactic cues, a prosodic one aligned with the positional one given by word-order, and the second where the parser is only guided by word-order cue.

Crucially, from the point of view of incremental processing this manipulation had the advantage of making happen these two different hierarchy building strategies at different time-windows. The prosodic cue should anticipate the sentence structure building at the boundary locus, starting already from the last syllable of the Topic NP, while the building on word-order encoding should start when enough evidence has been gathered to build a Topic-Comment articulation, namely reaching the second NP, the subject, as illustrated by the following Figure 5.11.

By inserting these sentences in context, we wanted to ask whether the PB located between the Topic and the comment would actually modulate the use of contextual information in comprehending the comment.

In conclusion, this experimental manipulation of the Prosodic Boundary will put

7. Note that as reported by recent work on Mandarin (Constant PhD., 2014) Chinese possesses a dedicated Contrastive marker making the abountness Topic shift to Contrastive Topic, namely 呢 *ne*.

forward: (1) if its absence is crucial for the comprehension mechanisms yielded by the sentence-discourse interface property of this syntactic structure and the contextual links it establishes during on-line sentence processing, and (2) to what extant listeners can simply rely on word-order during auditory sentence comprehension.

5.2.3.2 Predictions for the Prosodic Boundary manipulation

In this ERP study our main focus was three folds. We decided to investigate the effects of prosodic cues to syntactic structure and the effects of context informational-load during sentence processing, in order to elucidate the cerebral mechanisms underlying the interfacial links between sentence and discourse context in Topic-Comment articulation. For the present study this would lead to the following predictions for PB effect and Contextual load effects.

5.2.3.3 Condition PB and noPB

The Phono-acoustic difference having been detailed in previous chapter, we can just summarize that in PB Topic-Comment sentences compared to noPB conditions not only show two intonational unit, where Topic and comment are separated by in average 68 ms pause (significantly different from noPB) but a distinct prosodic contour accompanied this prosodic brake with a pre-final average lengthening of the last Topic syllable (79 ms), that was significantly longer that from noPB condition, $F(2,717) = 246{,}014$, $p<0.001$). This lengthening of the last syllable resulted also in a significant lengthening of the whole Topic constituent ($F(2,717) = 23{,}716$, $p<0.001$)

Furthermore, a third prosodic feature distinguished PB and noPB conditions, the average difference in pitch between Topic and Comment intonational units was on average of 18 Hz in PB. This pattern was significantly different from noPB ($t = 3.3166$, $df = 239$, p-value = 0.001053).

The examination of these prosodic features indicate that in PB condition the first acoustic information about the Prosodic Boundary becomes available as early as the Last Topic syllable before the pause between Topic and Comment.

It should also be noted that as shown in the behavioral acceptability test no outright prosodic violation was present in the noPB condition, which was naturally read by a male native speaker of Standard Mandarin Chinese[8] with an un-paused sustained flat reading.

Given this intonational under-informativity we just expect to observe some more subtle prosody-related ERP effects compared to the one reported in the literature in response to real Prosodic boundary violations or prosody-syntax mismatches.

Importantly, this manipulation and its different timing allowed us to manipulate the variable prosody as early as the onset of the last syllable of Topic constituent, whereas the factor word-order syntactic cue could be detected and came into play in the next time window at the subject position meeting the second NP, and thus detecting the [NP+NP+VP] pattern.

8. The sentences were spoken by a male native speaker of Standard Mandarin Chinese and digitized at a sampling rate of 44100 Hz (16 bit, mono, normalized to/the peak intensity to -30 dBs using CoolEdit Pro 2.0).For Figures of the Waveform of the intonation contour of the stimulus sentences in condition (a) and (b), see chapter 4, section 4.3.5, page 440).

Effect of prosodic boundary vs. Sustained flat reading at Topic last syllable If the under-informative flat intonational contour of noPB critical sentences is interpreted by the parser as the absence of a global sentence-level prosodic contour, we could expect to observe a late-positivity, following Astésano et al. (2004), who claim that some late semantic-prosodic processing interaction are observable in a P800 component with its left temporo-parietal distribution as being further be modulated by context information load manipulation.

Alternatively if the absence of PB is processed as a local prosodic mismatch we either expect to find some repair processes right after PB position at the point in the sentence when the syntactic structure can be detected upon syntactic word-order cues. In the most simple case, we would simply observe ERP components reflecting syntactic building on word-order grounds.

Interestingly, Eckstein and Friederici's (2005) experimental manipulation presented two different patterns for plain omission of an expected prosodic Boundary and the incongruent prosodic substitution by cross-slicing method. While the first elicited a RAN component, the incongruent prosodic substitution in Eckstein and Friederici (2006) elicited the so-called *Prosodic negativity* with a bilateral and broadly distributed from 300 to 500 ms after onset of the critical word.

Effect of prosodic boundary vs. Sustained flat reading at Subject

First of all, Subject time-window is in our model the point in the sentence where, in absence of earlier Prosodic cues, enough word-order information is gathered to detect the syntactic structure the parser engaged in.

If the default analysis that is pursued by the listener in absence of prosodic cues for syntax needs at this point to be reviewed based on the syntactic information now available on word-order grounds, we could predict that the prosodic under-informativity of noPB condition would be observable only at subject position, when it actually becomes structurally obvious. This would yield repair processes only when NP2 is encountered and not in the earlier time-window.

As the noPB condition presents a mild prosodic incongruity, it is indeed possible that it would become detectable by the parser only when it has enough evidence for word-order syntactic cues on how the sentence structure should be constructed. Consequently the Topic syntactic role to the sentence-initial NP, would be later attributed in the noPB condition than in the PB condition.

The specific ERP signatures of these processing difficulties at NP2 (the putative syntactic desambiguating point) are hard to predict, given the new syntactic structure tested in our study and the heterogeneity in the results reported in the literature. As argued above any of the ERP components linked to prosody syntax mismatch could actually be observed, the likely candidates would be mainly a LAN effect (Kerkhofs et al., 2008) and/or P600 effect (Kerkhofs et al., 2008; Pauker et al., 2011; Bögels et al., 2013).

Still, we wouldn't push this argument any further by identifying in this absence of Prosodic Boundary, a prosodic garden-path effect. We difficultly figure how such an early position in the sentence (i.e NP2/3) like the subject position could actually be a real *disambiguating point*. We namely believe this should not be the case: the under-informativeness of the prosodic pattern substituting the regular prosodic contour found in PB condition should not cue for any particular syntactic structure the parser

could be engaged in. In fact, even the generally considered default agent-first parsing strategy cannot anyway be not supported in our case, because of the inanimate feature of our Scene-setting Topic (i.e. a place word). Furthermore, if we assume the linguistic typological framework, the lack of a prosodic mismatch effect would be actually explained by a typologically default syntactic preference assigning the Topic role to sentence's first NP on the basis of Topic-prominence parameter observed for Chinese.

Anyhow, if any garden-path effect was to be observed, we would expect it to be a smaller P600 component than the one likely to be elicited after lexico-semantic violation in Coord-He condition. Namely, in case of a strong impact on sentence building processes of the prosodic under-informativity, the comparison between noPB sentences and lexico-grammatically incorrect sentences (Coord-He) would shed light on the extent of violation and reconstruction processes due to absence of Prosodic cueing for Topic-Comment construction. We predicted that a stronger neural activity was expected for syntactically incorrect sentences if the prosodic incongruence had only a small or no impact on the cerebral sentence processing. Probably the distinction made in the literature between early and late time-window effects for prosody-syntax interface processes will be relevant in our case.

In conclusion, our model predicts that the absence of early prosodic cues in the noPB condition leads to simple syntactic-role assignment (respectively Topic and Subject) on word-order grounds as soon as the second NP is encountered, thus we expect no misanalysis and no attachment ambiguity processes to occur, in turn this would not generate a garden-path effect at subject position.

No CPS predicted As for the highly debated ERP component following a Prosodic Boundary, the Closure Positive Shift (CPS) (Steinhauer et al. 1999), we predict an absence of CPS, mainly for two reasons.

First, because the information brought by the Prosodic Boundary is not making the parser engage in a new or different syntactic parsing strategy, as there is no syntax-prosody mismatch, and the Prosodic Boundary represents an additional cue to the syntactic one already given by word-order encoding. As revealed by the behavioral responses, the absence of PB does not make it impossible to parse the sentence: the off-line acceptability ratings were comparable and the discourse truth-value judgments (i.e. a proxy for comprehension ratings) showed that native speakers rated noPB slightly lower than PB sentences (descriptively small but statistically significant).

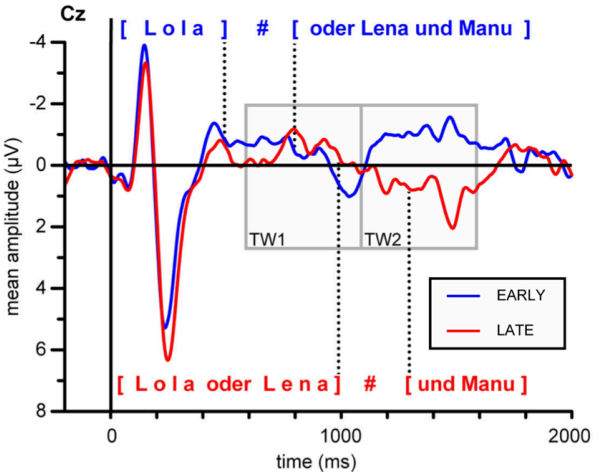


Figure 5.12 – Grand average ERPs for both conditions at electrode Cz. Gray boxes indicate the time windows used in the statistical analysis relative to stimulus onset. Dotted lines indicate the mean onset and offset of the pause at the Intonational PB in the respective condition, the silent pause is indicated by a hash mark. Adapted from Holzgrefe et al. (2013).

Taken together linguistic an psycholinguistic evidence converge in affirming that in noPB the parser should successfully rely on word-order only to process the sentence. It will be interesting to see the how the parser manages these two different ways (i.e. word-order versus prosody

strategy) of building the sentence-level hierarchy characterizing Topic comment sentence structure.

Secondly, as reported by Holztgrefe et al. (2013) the occurrence of CPS depends on the position of the intonational Prosodic boundary (pitch rise, final lengthening and a pause) in the stimulus sentence, namely no CPS was observed when the PB appeared early within the utterance, which in their case was the second name (see Figure 5.12). Given these results, the position of our PB being at the end of the first constituent in the sentence (varying from the 3rd to the 5th syllable) was then expected not to elicit CPS effect.

Thirdly, Kerkhofs et al. (2007) reported a contextual reduction of the CPS (see above §5.1.2.2, Figure 5.4). This result could actually constitutea prediction for what could be observed in the narrative (A), where the informational load about the Topic referent make it salient.

Not only the relation of CPS to Prosodic Boundary perception and processing per se has been repeatedly shown to be disputable, but some more recent studies reported the absence of the CPS for some type of PB (Holzgrefe et al., 2013) and not for others an different scalp distributions for different structures (Pannekamp et al., 2005). For a comprehensive review on CPS and the functional processes it underpins, see Annexes §E.1, p.E.1) [9].

Table 5.1 – Prosodic boundaries - overview of the cited ERP studies
NEG = negative shift in ERPs; POS = postive shift in ERPs; LAN = left anterior negativity; 1a = time-locking to sentence onset; 1b = time-locking to focus accent onset; 1c = time-locking to verb onset; 1d = time-locking to target onset; 1e = time-locking to offset of word before boundary; 1f = time-locking to offset of stressed syllable before break; 2a = targets not matched for frequency; 2b = targets not matched for lexical stress position; 3 = targets at phrase boundary.

Study	Task	Paradigm	Conditions	Effect	Interpretation	Possible problems
Kerhofs, Vonk, Schriefers, & Chwilla, 2007 (dutch)	None	Auditory; dialogues, with prosodic/ syntactic mismatch	Mismatch: prosodic/ syntactic break	POS	CPS	Time-locking (1e)
			Mismatch: prosodic/ syntactic break	POS	(right) CPS	Matching (2a, b)
Bögels, Schrifers, Vonk, Chwilla & Kerkhofs, 2009 (Dutch)	Comprehension	Auditory; prosodic breaks in single sentences	Prosodic break	POS	CPS (larger with object verbs)	Time-locking (1f) Boundary (3)
Li at al., 2010 (Chinese)	Comprehension	Auditory; dialogues, with prosodic/ syntactic mismatch	Missing prosodic boundary	NEG	LAN	Time-locking (1d)
			Superfluous prosodic boundary	NEG	LAN + N400	Matching (2a)
			Prosodic boundary	POS	P2 (fronto-central)	Boundary (3)

9. In this overview on Prosody related ERP studies, we mainly argue that CPS component is not a signature of prosodic information, but a signature of sentence syntactic restructuring based on the cue given by prosodic information when comparing prosody-syntax mismatch cases (e.g. prosodic garden-path sentences), which is not our case, in that word-order is cueing exactly for the same syntactic interpretation as PB marking.

5.2.3.4 Condition Coord-He

The ungrammatical condition of our experimental design merits some description, as it was uniquely inserted in the experimental design to constitute a term of comparison for the noPB condition. Namely, to be able to observe the difference between the kind of neural responses elicited by the detection of prosodic under-informativity and those elicited by the kind of morpho-syntactic violation corresponding to the insertion of a coordination conjunction between the Topic and the subject.

Crucially, from the point of view of incremental processing, this third condition (condition C) was detectable as ungrammatical starting from NP2 position.

Importantly, by this manipulation we were able to modify (before ungrammaticality was discovered) the hierarchy marking of of the two sentence-initials NPs, opposing two types of hierarchy cues, namely morpho-syntactic cue with natural prosody in the coord-condition and word-order cues with natural prosody in the noPB condition.

Not only coordination structures have proven to elicit CPS even in non-sentential sequences (Hotzgrefe et al., 2013), but they also have been the object under study of a number of behavioral studies investigating prosodic phrasing perception (see Wagner 2005; Bogel et al., 2013).

Prediction on Topic position noPB vs. He Contrast A direct comparison between condition PB and Coord at the PB could be informative about what kind of ERP structure building effects are elicited by to supra-segmental prosodic cues to sentence structure and what kind of ERP component underlay to structure building by means of explicit morpho-syntactic marking like the conjunction 'and'.

Moreover, another syntactic aspect differentiating greatly PB condition (a) and Coord condition (c) at the boundary position Top last Syllable (before encountering the grammatical error at NP2). While PB in condition (a) is marking a sentence-level hierarchy between the Topic NP hosted in the Left-Periphery of the sentence (i.e. in the CP-spec position) and NP2 the subject of the Comment sentence, the conjunction in condition (c) is marking a local hierarchy cueing the attachment of NP1 and NP2 in a coordinate structure.

Hence we expect the comparison between these two conditions (c versus a) to be informative of how the parser builds under natural prosody conditions a more local hierarchy in the coord condition and sentence-level hierarchy in the Topic noPB condition.

Prediction on subject position noPB vs. He Contrast If our model of the parsing strategy of these conditions is true, in this time window we expect he parser to be engaged in the syntactic hierarchy building of Topic-comment articulation based on word-order cues of noPB conditions, and in the discovery and repair process linked to the lexical violation of the Coord condition, e.g. 'the tavern by the river and Zhao Hu'.

Alternatively, if in noPB condition the parser engages in some repair processes of the Prosodic boundary cueing for sentence-level hierarchy (i.e. intonational pause, lengthening and pitch), the comparisons between noPB vs. Coord conditions at subject position may help to disentangle processes repairing grammatical violations from the kind of possibly repairing the prosodic contour under-informativity.

Under this improbable scenario- given the behavioral responses showing equal acceptability of PB and noPB conditions, a difference in the kind of repair processes yielded

by these two conditions, would contribute to give a further confirmation to what was argued before about the difference between word-order based (PB-marked or not) syntactic boundaries and morpho-syntactically marked ones.

Condition B and prosodic under-informativity Our hypothesis was that the prosodic under-informativity in condition (b) (compared to the baseline sentence with regular PB) would elicit a qualitatively different repair ERP pattern than the comparison between the ungrammaticality induced by the coordinative conjunction in condition C.

If contrary to what was found by Toeple et Alter (2004)[10], the prosodic under-informativity is identified and dealt by the parser in a similar way as grammatical illformedness, then the same type of repair ERP component should be elicited by no-PB and *Conj conditions, with the only expected difference being a quantitative larger response for the ungrammatical condition Conj.

Alternatively, if the under-informativity of the absent Prosodic boundary in condition noPB leads the parser to fail to recognize and build the sentence hierarchy between NP1 and NP2 (or alternatively if it changes the interpretation of sentence structure), then this should be seen in a repair response that could only have been caused by the use word-order cues to sentence structure building at NP2 position before the verb is encountered. P600 would be indicating the necessity of a subsequent structural revision concerning the place of NP1 in the sentence hierarchy.

Coord condition violation effect As for the nature of the violation discovered upon reaching Subject time-window (NP2), we can say that it should mainly be lexical, as the animate NP2 is semantically inadequate to be coordinated with the inanimate NP1 on animacy grounds. We could predict that the repair of the semantic feature of NP2 or alternatively the deletion of the conjunction would lead to substantial processing and reanalysis difficulties already at NP2 position (subject) and even more at Verb position. However if the violation was to be identified by the parser as as for the morpho-syntax violation effect in Conj. (c) condition (i.e. violation of the coordination structure), we clearly expected to find the classic ERP components reflecting difficulties in local phrase structure building compared to sentence-level structure building that should be seen in comparison with PB condition.

We thus hypothesized that N400 component should be elicited at the point of morpho-syntactic violation, both compared to PB and noPB conditions, where the second NP of condition c did not respect the semantic constraint of the preceding conjunction and (“he”) And in a later time window a further expected a P600 reflecting late processes of syntactic reanalysis and repair.

10. This study found that the under-specification of accent prosodic information in a given context requiring it did not affect the processing relative to the adequate accent, when participants performed a comprehension task, focusing on the content.

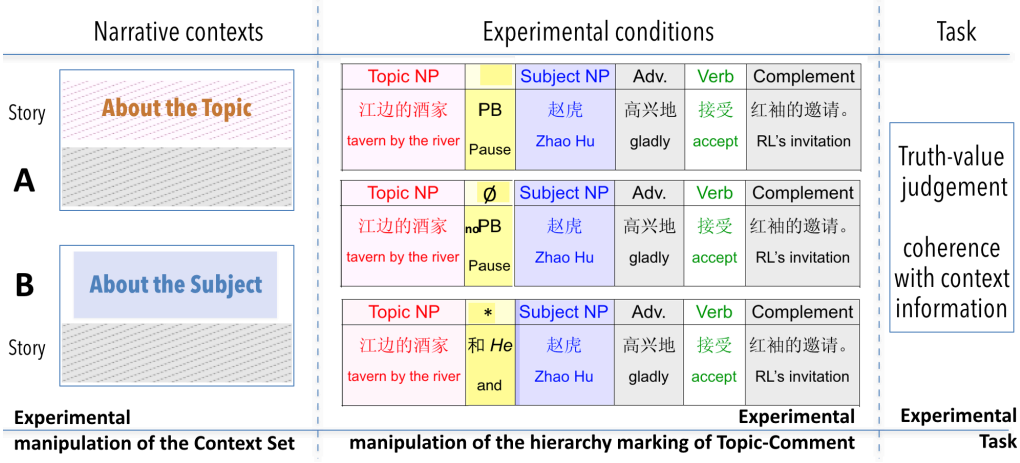


Figure 5.13 – Experimental manipulation of the Context set in Story A and B, experimental manipulation of the hierarchy marking of Topic-Comment and experimental task.

To sum up, by crossing the presentation of two types of contexts (neutral vs. biasing/rich) with the three types of critical sentences structures (PB, noPB, *Conj conditions) we obtained the following six experimental conditions (see Table hereunder), in a 3 (PB/No PB/He violation) x 2 (Rich Context/Plain Context) design.

As illustrated in Figure 5.14, this design will allow to investigate the time-course of contextual information integration during Gapless Topic-Comment sentence processing, we expect that the profile of ERP responses found in the three critical time-windows will inform us about (1) the impact of the saliency of the Topic discourse-referent, and (2) the impact of prosodic marking of the syntactic hierarchy between Topic and Comment on contextual intergation during online sentence comprehension.

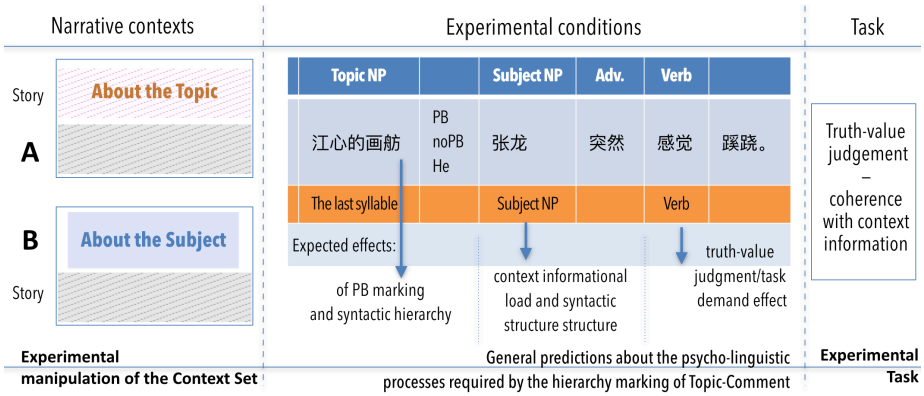


Figure 5.14 – General Experimental hypotheses on the localization of the effects of Contextual-load and PB information across the three critical time-windows, Topic last syllable, Subject onset and Verb onset.

5.3 Method: Protocol and Data analysis

5.3.1 Experimental Protocol

5.3.1.1 Participants

Twenty-four participants (including 11 females, Age ranging from 18 to 25 years old, Mean age = 22 years old) were paid for their participation to the ERP experiment. All of them had normal or corrected-to-normal eye sight and had no neurological or psychiatric disorders. All of them were native Mandarin speakers and were monolingual. This experiment was approved by the Committee of Academics in Peking University.

5.3.1.2 Procedure

Participants were seated comfortably in a sound-proof, electric-shielded chamber at Peking University, in Laohuaxue building. They were about 80 cm away from the CRT screen where the narratives were presented in the center of the screen, and had been asked to move their heads or bodies as little as possible.

The participants were first asked to read a short narratives (i.e. context) carefully and attentively for comprehension, with no time limits for the reading.

Then a fixation cross was presented (for 400 to 900 ms) and subjects listened to 27 critical sentences per story in average (in average 11.43 critical sentences and 15.57 fillers), whose presentation was jittered in time, so that the subjects could not predict the exact onset of the auditory stimuli. Average duration of sentences was 2954.1 ms, respectively, critical sentences' length: 3233.5 ms ; and filler sentences' length: 2488.5 ms.

The experimental session was divided into 5 sessions with 4 breaks, the distribution of the stories in each block was randomized selecting for one half of the stories context A (rich context) version and counterbalancing them with the other half with context B version. No participant saw the two versions of the same story. Every block had four stories, only the second block had five.

Eight lists (4 pairs) of randomized stories were created and for each context, the three conditions were counterbalanced using Latin square design. That is, one third of the sentences were assigned as PB, one third as noPB, one third as Coord-He, making sentence list A, as well it's counterbalanced counterparts list B and list C. The two Layers were then mixed and matched. Therefore, there were in total 8 (story lists) * 3 (Sentence Lists), 24 presentation lists. The sentences were randomly ordered within each story, and it was made sure that no three sentences in a row were all critical sentences.

During auditory stimulation EEG was recorded and 300 ms after each sentence a question mark appeared on the screen and participants were asked to judge whether each sentence was consistent with the story they just read. This discourse truth-condition judgment task evaluated the participants comprehension of the fit with the context of each sentence. The rating of the coherence of the sentence with the prior short narrative was a binary choice 'Correct' (*duì* 对, coherent) and 'False' (*búduì*, 不对 incoherent) by keyboard press.

After the EEG experiment, participants rated the naturalness of all the critical sentences they had heard during the experiment.

The entire experiment including the completion of post-experimental grammaticality questionnaire and electrode preparation lasted 1.5 hour.

5.3.1.3 Experimental Task

Discourse coherence judgment In the discourse congruity rating, the participants were asked to judge the degree of relatedness between the sentence and the proceeding narrative (i.e. the fit of context and target sentence) on a 5-point Likert scale with 1 representing the most related and 5 representing the least related.

Post-experiment grammaticality ratings A post-experiment grammaticality rating tasks was completed by participants after the EEG session, they were asked to explicitly judge the naturalness or well-formedness of the critical sentences auditory stimuli on a scale from 1 to 5 (1: Very unnatural; 5:Very Natural).

5.3.1.4 EEG recording

EEG was recorded using Brain Vision Analyser Software (Analyzer version 2.0.; Brain Products, Gilching, Germany), via a 64 cap-mounted Ag/AgCl-electrodes Brain Products Asian Cut EasyCap (64 channels fast EasyCap for BrainAmp, Brain Products, Germany), and a BrainAmp DC amplifier. The vertical electro-oculograms (EOGs) were recorded by placing an additional external electrode above the left eye. The horizontal EOG was recorded by placing an additional external electrode at the outer cantus of the right eye. All EEGs and EOGs were online-referenced to an external electrode which was placed on the tip of the nose, were band-pass filtered from 0.016 to 100 Hz. and sampling rate was at the frequency of 500 Hz. The ground reference was a medial frontal site electrode. The electrodes were referenced online, while recording, to an electrode placed at the tip of the nose. Impedance was kept below 10 kΩ for EOG channels and below 5 kΩ for all other electrodes. Bad channel signals were replaced offline using spherical spline interpolation with the surrounding electrodes.

5.3.2 Data Analysis

5.3.2.1 EEG Pre-processing

Offline EEG data analysis was performed in EEGLAB (Delorme and Makeig, 2004). The EEG was first filtered with a low-pass filter with a cutoff frequency of 30 Hz, and a high-pass filter of 0.5 Hz. Epochs of 900 ms around Critical words on the Last syllable of the Topic NP, Subject NP and Verb (from -200 ms to +700 ms) were extracted from the continuous EEG signal. Eye blinks and eye movement artifacts were compensated for/corrected using Independent Component Analysis (ICA) (Jung et al., 2000) and artifact detector algorithm based on joint use of spatial and temporal features, a ADJUST plug-in to EEGLAB toolbox (Mognon et al., 2010). Trials were removed whenever the signal amplitude exceeded +140 mV in the epochs.

The segmented EEGs were baseline corrected according to (1) the mean amplitude of the activity in 200 ms pre-onset of the critical words.

The mean amplitude of the ERP data was calculated per condition, participant and electrode, before grand averages were computed over all participants.

The electrodes were grouped together in nine Regions of Interest (RoIs), corresponding to three anterior/posterior regions in both hemispheres (left/right), and three midline regions from anterior to posterior (see Figure 5.15).

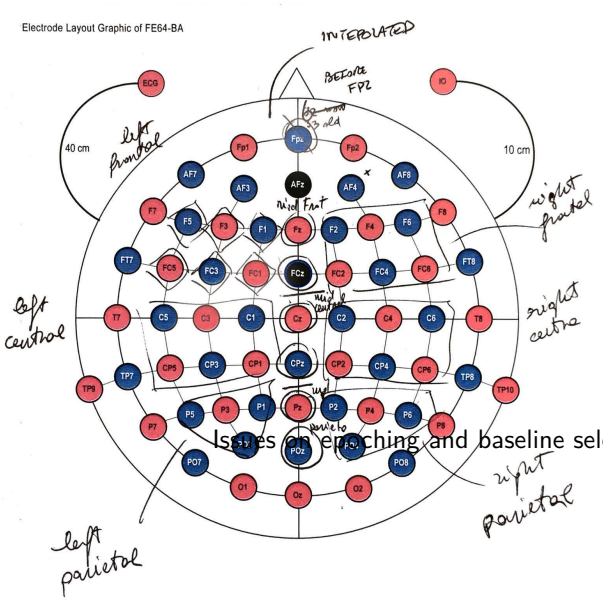


Figure 5.15 – Topographic factors in the Linear Mix Model.

Midline regions comprised: mid-frontal (Fz, FCz), mid-central (Cz, CPz), and mid-parietal (Pz, POz). The other electrodes were assigned to one of 6 lateral locations, which were classified into left/right hemisphere and frontal/central/parietal location factors:

1. left frontal (F1, F3, F5, FC1, FC3, FC5),
2. left central (C1, C3, C5, CP1, CP3, CP5),
3. left parietal (P1, P3, P5, PO3),
4. right frontal (F2, F4, F6, FC2, FC4, FC6),
5. right central (C2, C4, C6, CP2, CP4, CP6), and
6. right parietal (P2, P4, P6, PO4).

The investigation of prosodic boundary processing virtually always comes with critical latency differences in the stimulus material, for instance the difference between PB and noPB conditions is in our case on average 68 ms. Moreover, the onset latency of the CPS represent a substantial issue in the literature because it yields differential results across studies (see Bögels et al. 2010 for comparative approach). In a variety of studies this component was reported to be somewhat later than those present in previous studies (see Steinhauer, 2003; Steinhauer et al., 1999).

This difference in onsets is mainly caused by important differences between the averaging and time-locking procedures that are used by different authors. Studies on CPS (e.g., Steinhauer et al., 1999, Toepel et al., 2007 among others), indeed, time-locked and normalized the wave-forms to the onset of the sentences, computing an average ERP waveform over the entire sentences.

Consequently, the location of a prosodic break in the auditory signal is determined by computing the average location of the pause of the prosodic break in the auditory signal, and the latency of the CPS is then estimated by comparing the point in time at which the Prosodic Break condition and the No-Prosodic Break condition begin to differ at the average position of the prosodic break. Clearly, as noted by Steinhauer (2003), this latter procedure has the disadvantage of 'considerable latency variability across trials' (p.151), with respect to the onset of the pause of the prosodic break, a problem that actually does not occur when time-locking the ERPs to the offset of the word preceding the pause. It should also be noted, however, that in the visual modality (written presentation), the onset of the CPS in response to a comma (Steinhauer and Friederici, 2001) showed a similar onset latency as the one observed in previous studies that time-locked events at the the onset of the pause.

More recently a number of studies, described the CPS positive shift to start almost immediately at the end of the pre-boundary time window, right after the onset of the pause peaking at around 500 ms (see, Bögels et al., 2010; Itzhak et al., 2010; Pauker et al., 2011) and to last around 500-700 ms (e.g., Pauker et al., 2011, see also above).

In the present study, we used the default ERP methodology of time-locking each individual trial to a specific critical event and normalizing the wave-forms in a 200 ms interval directly preceding that event. However, as in our case prosodic boundary information also involved pre-PB pitch variations and last syllable lengthening, we chose to time-lock the first event of the sentence at the onset of NP1 (Topic) last syllable, i.e. the onset of the last stressed syllable right before the pause in the condition with a prosodic break (see Bögels et al., 2010 and 2013). The second event of the sentence was time-locked at point where syntactic word-order cue for Topic comment structure were available, i.e. the subject time window (NP2). This was meant in order to capture the cerebral responses to any prosodical information or word-order syntactic cue likely guide the parser in the building of the sentence-level hierarchy as early as possible. A third epoch was time-locked at the onset of the comment clause verb to be able to observe

Considering the results in the literature and the issue of pre-boundary prosodic cues to parsing, it would be interesting to run a more sophisticated analysis time-locked to the boundary position (offset of critical NP) in addition to the time-locking to stimulus onset that allows to (a) compensate for these inherent latency differences between conditions and to (b) disentangle the CPS from post-boundary onset components (P200) and Early negativities (EN or PMN prosodic mismatch negativity) or pre-CPS negativity (Paukeret al., 2011; Bogel et al., 2011).

5.3.2.2 Statistical analysis

Linear regressions were run to analyze ERP data in three epochs with t-value > 1.98 and associated p value of 5% considered as significant.

Analysis Method

Statistical analysis was performed with R software [11]. The experimental hypotheses formulated in the previous section were tested in a Linear Mix Model (LMM), since we were expecting to have interactions between effects in the different time windows at critical words between -200 ms to 800 ms, that is Last syllable of the Topic NP (NP1), Subject NP (NP2) and Verb.

Experimental factors were:

1. context (Rich vs. Plain) and
2. PB (two planned contrasts: a. PB vs. noPB; b. PB vs. he).
3. Topography. There were two topographic factors:
   (a) the first was for midline analysis: anterior-to-posterior (3 levels: Fz and FCz / Cz and CPz / Pz and POz)
   (b) The second was for lateral analysis: Hemisphere (left/right) * anterior-to-posterior (3 levels, similar to the midline analysis).

The topographic factors corresponded to the ROIs (see previous section, Figure 5.15). Midline (a) and Lateral (b) analyses were performed separately. Hence, Linear mixed-effects models were adopted to analyze average amplitude in each time window (Topic NP, Subject NP and Verb), had the following contrasts in the model:

1. - PB Effect (PB vs. NoPB condition)

11. R

2. - He conjunction Effect (PB vs. He condition)
3. - Context Effect (Rich vs. Plain context condition)
4. - Electrode position (Left vs. Right; Frontal, Central vs. Posterior)
5. - Interactions among these conditions

5.3.3 Results: Short epoch analysis

For the short epoch analysis, the epoch length around Critical words was fixed from -200 ms to 700 (/800) ms on Topic NP Subject NP and Verb (Random effect for subjects).

Last syllable of Topic NP

A larger negativity was found for the presence than for the absence of PB (0.42 V, t = 3.2), while a positivity was observed for Rich than for Plain context (0.32 V, t = 2.4) 100-500 ms after the onset of the last syllable of Topic NP.

Subject NP

On Subject NP, the interaction between the presence of PB and contextual load in 0-150 and 150-300 ms time window (t = 3.1) indicated that there was a negativity effect for increasing load when a PB was provided (0.44 V, t = 3.6), while no significant effect was obtained when the PB was unavailable.

Contextual load also marginally interacted with the presence of violation in our Coord-He Condition in 300-650 ms Time-window (t = 1.91), with the load-associated negativity again found only after a PB but not after the conjunction violation.

Comment-clause Verb

On the Comment-clause Verb, however, the PB by context load interaction (t > 5.0) unfolded differently: increasing load made no difference after a PB, but elicited larger negative responses in the no PB condition (0.76 V, t = 6.0).

Importantly, from the statistical analysis no lateralized response appears to resist statistical significance; Hemispheric interaction with the others manipulated parameters PB and Context.

We will proceed in the following to a one-by-one discussion of the three time-windows and the ERP components that can be observed for Topic prosodic marking and contextual effects.

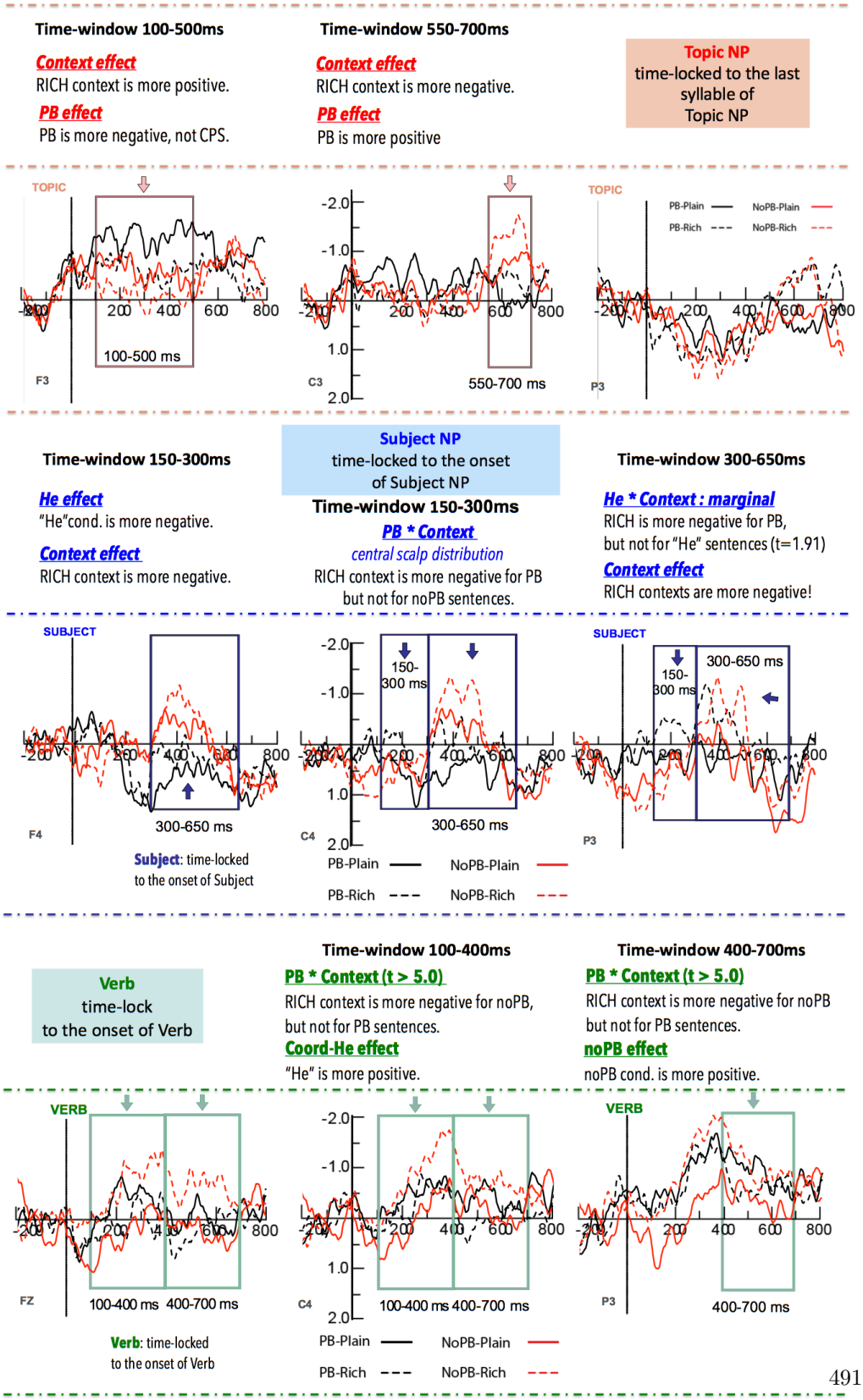


Figure 5.16 – Summary of Grand-average wave-forms over participants (n = 23), time-locked to the onset of Topic, Subject and Verb position for Neutral Context conditions (solid lines), the Biasing/Rich Context condition (dotted lines) in the three conditions, respectively PB (black), no-PB (red) and Coord-He (blue) condition.

5.4 Discussion: ERP Results and interpretation

5.4.1 Topic Last Syllable

As shown in Figure 5.17, different time-windows have been investigated upon ocular inspection:

- 50-200 ms: revealed no statistical significant effects
- 200-550 ms: PB effect in plain context
- 550-750 ms: Interaction between PB and Context

The ERPs at Topic Last Syllable (NP1) are presented in Fig. 5.17, adjusted to a 200 ms baseline preceding the critical epochs. The boxes indicate the time windows used in the statistical analysis relative to stimulus onset

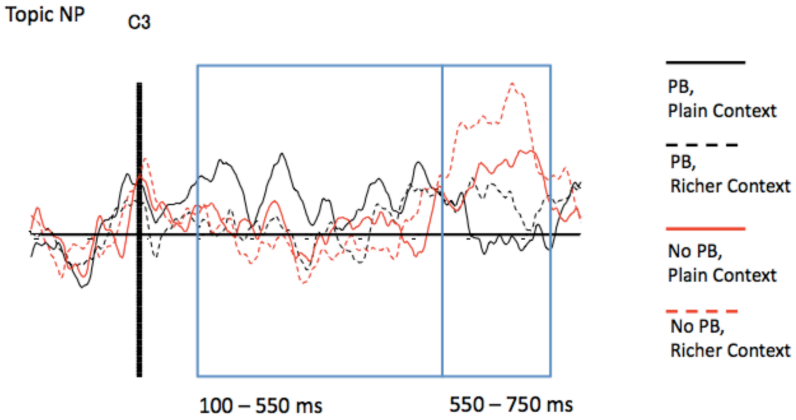


Figure 5.17 – Modulation by context and PB at Topic last syllable in C3 eletrode. Grand-average waveforms over participants (n = 23), time-locked to the offset of NP2 Topic's last syllable, for Prosodic Break in Plain Context condition (solid lines) and Prosodic Break in Rich Context condition (dotted lines) in two conditions, respectively PB (black), no-PB (red).

5.4.1.1 Topic's Earlier time-window 100 - 550 ms

In the earlier time window 100-500 ms no CPS is found, but we can observe a larger negativity for the presence PB relative to its absence (0.42 V, t = 3.2). While a positivity was observed for Rich than for Plain context (0.32 V, t = 2.4) 100-500 ms after the onset of the last syllable of Topic NP.

Clearly, while the Late time window shows an interaction between prosody marking and context informational load (interaction:1.8), during the earlier time window (50 - 550 ms) only the effect of prosody is observable in a frontally distributed negativity. The negativity we found might correspond to a pre-CPS negativity we had hypothesized and that is frequently reported (Kerkhofs, et al., 2008; Bogels et al., 2010; Pauker et al., 2011), however the amplitude and latency of the negativity in our study are not comparable to this component.

PB and the building structure through early prosodic cues

We have a very early effect for prosodic boundary starting as soon as the first pre-boundary cues are processed[12]. As previously reported by several studies (e.g. Boegels et al., 2010; Pauker et al., 2011) the prosody syntax interface seems to be implied in early stage of phrase structure building, as soon as Prosodic information is available.

Importantly, the fact we followed Boegels et al. (2010) and Kerkhofs et al. (2007) in selecting to time-lock the beginning of Topic Last syllable was mainly due to our interest in understanding how early the prosody syntax interface was happening. This has the disadvantage of not being able to compensate for the acoustic changes and latency difference between PB and noPB conditions (see Steinhauwer and Dury, 2012 for a thorough discussion of the importance of baseline selection in auditory ERPs).

However, as we can see from Figure 5.18 showing three different electrodes and from the zoom on the central region electrode C3 in Figure 5.18, this latency difference is not observable upon visual inspection during the prosodic sustained and long lasting negative effect taking place between 50 to 550 ms[13]. Il should be added that this short latency difference is anyhow reabsorbed in the second time-window as we clearly see that synchronously all conditions inverse their patterns starting at 550 ms for central electrodes and at 600 ms for frontal ones.

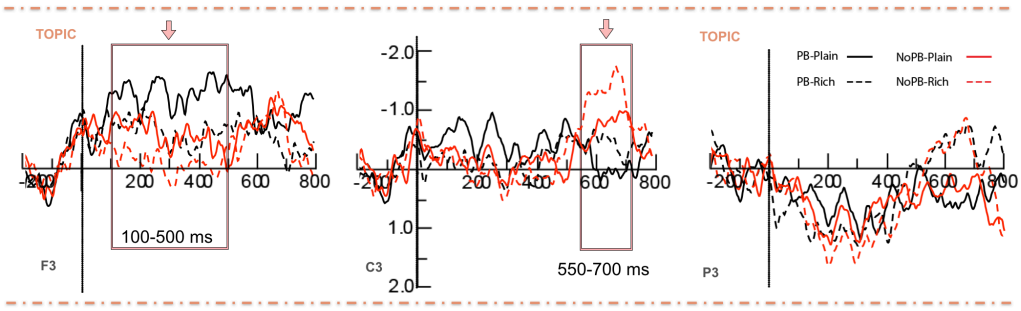


Figure 5.18 – Modulation by context and PB at Topic last syllable in three electrodes F3 for frontal scalp region, C3 for central scalp region and P3 for posterior scalp region. Grand-average waveforms over participants (n = 23), time-locked to the offset of the last syllable of the Topic NP, in Subject-related Plain Context (solid lines) and in Topic-related Rich Context (dotted lines) in two conditions, respectively PB (black), no-PB (red).

5.4.1.2 Topic's Late time-window 550 - 750 ms

The ERP effects in the last time-window show an inversion in pattern compared to the Topic's first time-window. The processes taking place from 550 to 700 ms seconds show a nice gradual effect of context-load and prosodic information, with a smaller significant interaction (t= 1.8).

As for the negative modulation of PB and noPB conditions by richer context-load, we can say that this response pattern suggests an ongoing activity of reviving/integrating

12. Average difference of pre-boundary last syllable length compared to noPB last syllable is 79.54 ms ($p > .0001$), and average significant pause length 68 ms

13. At 200 ms after on set the lengthening of Topic last syllable and the pause of PB condition are concluded.

informational-load related to the Topic referent in this late time-window[14].

This effect of context-load integration already at Topic position is in line with the predictions we had formulated in our model of the informational load processing of in-context Topic-Comment structures, where contextual information needs to be revived to be able to understand and truth-conditionally evaluate the Comment's predication. In the precise case of our Scene-setting Topics the truth-conditional evaluation of the Comment clause, should be based on a first step of integration of all that is known about the two spatial settings in the stories, and on a second step of verification to check if the action of the predicate in the Comment clause is actually taking place in the right setting in the context-narrative.

However, while our model predicted an early effect of informational-load integration, we had not emitted the hypothesis of such an immediate effect of the prosodic boundary marking on context integration. Namely, the absence of prosodic boundary makes it more difficult to integrate Topic-related information, and we can therefore interpret this response pattern as showing that since the very first constituent of Topic-Comment articulation the incremental processing strategy of Topic-comment integrates context information-load effortlessly when the PB early cues for the Topic-Comment syntactic articulation.

The gradual effect of absence of PB and contextual-load present namely the PB marked conditions in plain context at the lowest, then PB marked in rich context, followed successively by noPB plain context and noPB richer context. The context integration process is gradually more costly in absence of clear prosodic marking of the sentence-level hierarchy and when the amount of information linked to the Topic discourse referent is increased.

Hence, we can speculate that in presence of PB marking the sentence-level hierarchy between the CP layer and the Comment clause is built earlier, and that this activates the sentence-discourse interface mechanisms Topic-comment articulation. This has the consequence of making it easier to integrate the informational-load present in the context. It is based on this first finding that we say that ***context meets the sentence at the prosodic boundary***!

As a side remark in terms of cognitive resources, the interaction between contextual information load and Topic marking Prosodic Boundary, brought us, my colleague Luo Yingyi and I, to speculate about the biological need of a pause to live up and integrate all the information available from context about the Topic referent, to then be able to understand the relevance of Gapless Comment, whose link with the Topic can be somethings fairly inferential as we saw in many examples of Aboutness Topics in chapter 3.

Contextual-load effect

From 550 to 750 ms the negative component is affected by increasingly richer context and absence of PB.

The absence of CPS is, as predicted, due to the fact that the information brought by the Prosodic Boundary is not making the parser engage in a new/different syntactic

14. Note that if these effects were linked to referential dependency they would be linked to the salient encoding of the Topic referent in the mental model constructed to store the context information of Rich narrative A-, and they would therefore be inverted, as we had predicted in our Topic-comment processing model. Hence, what we see here is a context information-load effect.

parsing strategy, there is no syntax prosody mismatch the PB represents an additional cue to the syntactic cue already given by word-order. As word-order is cueing exactly for the same syntactic interpretation as PB marking, the absence of PB does not make it impossible to parse the sentence[15].

Furthermore, the behavioral responses of subjects already showed equal acceptability of the two prosodic versions which were comparably understood in context (noPB sentences slightly lower than PB sentences). These two measures and the ERP effects observed here converge in showing that listeners can successfully rely on word-order only to understand the noPB sentences.

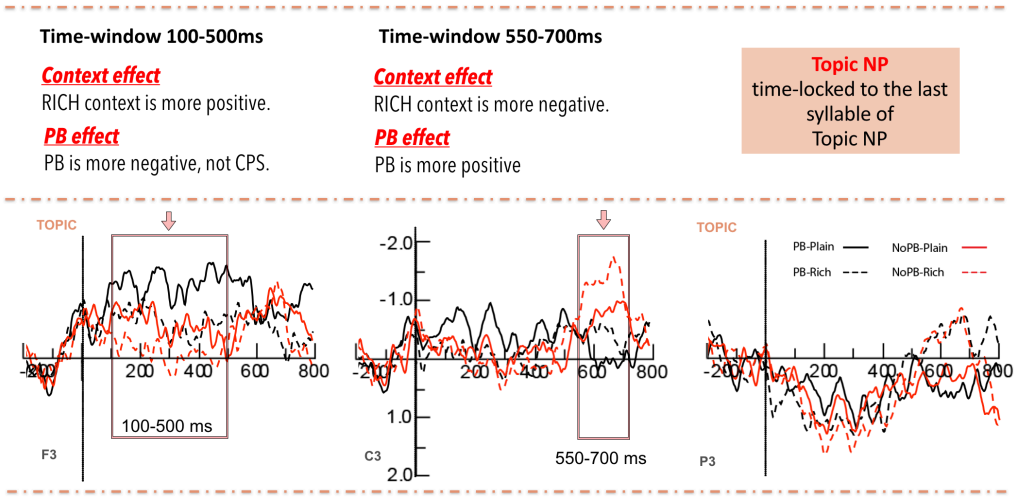


Figure 5.19 – Recapitulation of the modulation by context and PB at Topic position in C3. Grand-average waveforms over participants (n = 23), time-locked to the onset of NP2, the Subject of the Comment-clause, for the Plain Subject-related Context (solid lines), and the Rich Topic-related Context (dotted lines), in the two experimental conditions, respectively PB (black) and no-PB (red).

5.4.2 Subject of the Comment

On Subject NP, the interaction between the presence of PB and contextual load in 150-300 ms time window (t = 3.1) indicated that there was a significant negativity effect for increasing load when a PB was provided (0.44 V, t = 3.6). An effect of un-grammaticality for the Coord-He condition effect with a more negative deflection than other conditions was also observed (t=4.59). In the earlier time-window between 0 - 150 ms, a very early

15. Some could argue that the PB condition in Plain context shows a Positive effect, which may presumably be considered a CPS following Kerkhofs et al. (2007), who revealed a context attenuation of the CPS, as it could be argued it is the case in this time window. However, the authors report in their results that they performed an analysis for two ROIs (posterior and frontal) that revealed an effect of Prosodic Break: for the posterior ROI [$F(1, 29) = 5.01$; $p < .05$], but not for the anterior ROI ($F < 1$). Supplementary analyses for the individual electrodes have been revealing a significant effect of Prosodic Break at three posterior sites over the right hemisphere: CP6, P8, and PO8 (all ps $< .05$). The CPS in their Experiment 2 showed a centro-parietal scalp distribution that was restricted to the right hemisphere. But, our statistical analysis revealed no significant lateralized response and no Hemispheric interaction with the others manipulated parameters PB and Context.

effect of prosodic boundary irrespective of the context revealed to be highly significant ($t > 7.$).

In the time-window between 300-650 ms, conditions embedded in Rich context are more negative. Contextual-load also marginally interacted with the presence of violation in 300-650 ms Time-window ($t = 1.91$), with the load-associated negativity again found only after a PB but not after the conjunction violation.

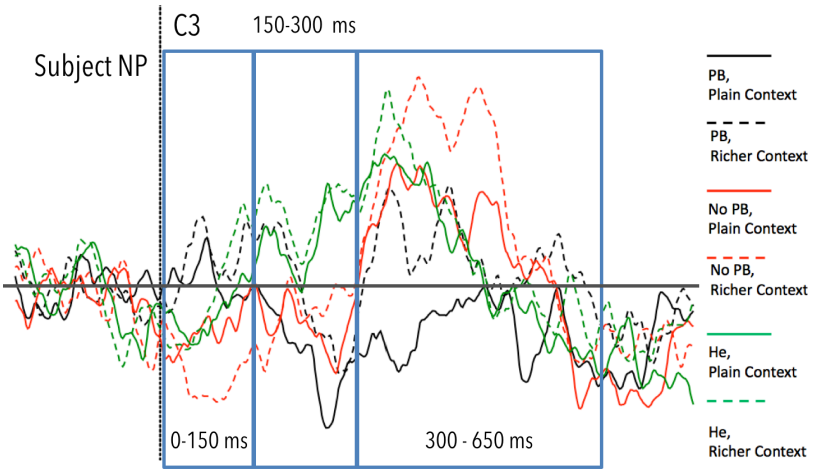


Figure 5.20 – Modulation by context and PB at Subject position in C3. Grand-average waveforms over participants (n = 23), time-locked to the onset of NP2, the Subject of the Comment-clause, for the Plain Subject-related Context (solid lines), and the Rich Topic-related Context (dotted lines), in the three experimental conditions, respectively PB (black), no-PB (red) and Coord-He '*and*' condition (green).

5.4.2.1 Early time-window 150 - 300 ms

noPB: building structure on word-order cues

As subject time-window is immediately following the Prosodic boundary, we could expect to see effects linked to some repair processes linked to the absence of Prosodic boundary. However, as shown by our behavioral results and as predicted, no prosody-syntax mismatch is observable here. The syntactic structure built on word-order ground is avoiding any prosody-syntax mismatch to take place here, and therefore it is understandable that the ERP-components generally reported in the literature for this kind of process are not observable.

Interestingly, observing the earliest time window between 0 - 150 ms, we noted upon ocular inspection a broadly distributed very early effect of prosodic boundary irrespective of the context in which critical sentences appeared (for Grand-average ERPs (n = 23) at nine electrodes see Annexes §E, p.925). This effect, that revealed to be highly significant ($t > 7.$) although descriptively small, is the only element indicating at subject position a possible effect of absence of Prosodic boundary. Importantly, in this time window no PB by Context interaction is attested ($t = 1.7$). We interpret this effect as N100 effect for PB conditions marking the left-hedge boundary of the beginning of the comment, which

is typically described as reflecting the physical features of the stimulus (see Steinhauer et al., 1999 among others), and usually found after an intonational pause (compared to the same words when they ended the sentence, see Strelnikov et al., 2006 or Roll and Horne, 2011 for discussion).

Interface between prosodically cued syntax and Context

Importantly, in this early time-window the interaction between the presence of PB and contextual load (150-300 ms time window, t=3.1), indicated that there was a significant negativity effect for increasing load when a PB was provided (0.44 V, t=3.6) while no significant effect was obtained when the PB was unavailable, which suggests that contextual-load effect is here observed only in the condition where the Topic-Comment sentence-level hierarchy is already built, i.e. the PB condition (PB x Context, t=1.7; noPB x Context, t=5.94).

PB cue triggered the building of the Topic-Comment syntactic structure in previous Time-window, thus establishing the sentence-discourse interface mechanisms that are characteristic of Topic-Comment articulation and to the place occupied by the Topic in the Left-periphery, the lack of PB does not make it impossible for the parser to build a sentence structure on word-order grounds (no violation effects are observed), but it apparently delays the establishment of the sentence-discourse interface. In other words, just relying on word-order cues to parse the sentence slows down the activation of the context bringing mechanisms distinguish Topic-Comment articulation.

Hence, as for the question we raised for the subject time-window – whether the PB located between the Topic and the comment modulates the use of contextual information in comprehending the comment – and the answer is yes, we will see what happens in the next Time-window on the verb.

Ungrammaticality

Additionally, as illustrated in Figure 5.20 the 150-300 ms time-window shows an early effect of un-grammaticality for the Coord-He condition effect, which is more negative than other conditions (t=4.59).

The earlier latency of what we interpret as morpho-syntactic violation effect is mainly to be attributed to the fact this time-window is locked at the onset of the second Np and the conjunction has already been heard. To this should be added that the prosodic contour of Coord-He condition revealed by our phono-acoustic analysis has a significant average pause of more than 100 ms.

Importantly, the violation effect of Coord-He conditions not only precedes noPB effect but also differs for it in this time window, thus indicating that lack of PB after Chinese-style Topic was not processed as a morpho-syntactic violation.

5.4.2.2 Later time-window 300 - 650 ms

Referential interpretation of Subject NP and Context effect

While in the first time-window interaction between the presence of PB and contextual load was indicating a significant negativity effect for increasing load when a PB was provided (0.44 V, t = 3.6), the second time-window shows an increased negativity

for both PB and noPB in Rich context with a stronger effect in absence of Prosodic Boundary, and noPB wave-forms are always more negative.

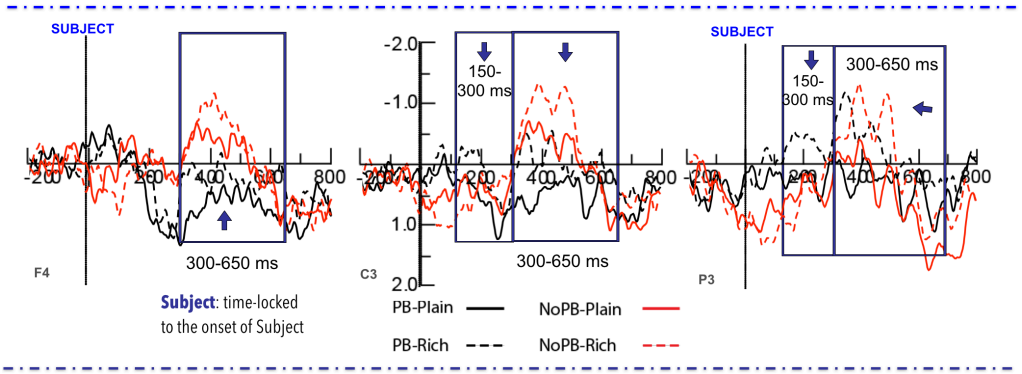


Figure 5.21 – Modulation by context and PB at Subject position in three electrodes F4, C3 and P3, respectively for frontal, cental and posterior scalp regions. Grand-average waveforms over participants (n = 23), time-locked to the onset of NP2, the Subject of the Comment-clause for the Plain Subject-related Context (solid lines), and the Rich Topic-related Context (dotted lines) in the three experimental conditions, respectively PB (black), no-PB (red) and Coord-He '*and*' condition (green).

Our prediction that the mental model constructed for plain context, i.e. giving more saliency and thus a higher degree of accessibility to the Subject referent – which would have made it easier to retrieve the subject referent from context – is actually borne out for PB condition (compare the solid black line with the dotted black line in Fig. 5.21) and through a reduction of the referential negativity (Nref) in noPB conditions (compare the solid red line with the dotted red line in Fig. 5.21). As we saw in our review of the literature, this process of establishing a dependency relation between incoming information in the sentence and information already available in the discourse model generally reflected in modulations of the N400 component in a broadly distributed negative deflection with a peak latency around 400 ms after the onset of the critical input (e.g. Burkhardt, 2006/2007; Streb et al., 2004; van Berkum et al., 1999). The more demanding the access to information in the discourse model is, the more enhanced is the amplitude of the N400.

In line with our predictions, Topic-related contextual effect (i.e. Rich story A) elicited larger sustained negativity on Subject NP, validating our hypothesis about the accessibility of Subject referent in the mental model constructed for Topic-related Context A. The pattern of response observed here clearly shows that this effect is linked to the cognitive accessibility of the subject discourse referent in the mental model.

Both scalp distribution and the peak at 400 ms are in line with previous findings by Petra Bukhardt (Bukhardt et al., 2006; Bukhardt and Roem, 2007) and Petra Schumacher, who calls this a 'Discourse Linking cost'. A study by Hirotani and Schumacher (2011), had namely investigated Topic marking in context in Japanese and revealed an increase in the N400 component as a function of contextual cueing -i.e. the less accessible a referential expression is in the discourse model, the more pronounced is the amplitude of the N400[16].

16. In their study, accessiblility of the referential expression is in the discourse model was modulated

Hence, following these studies on referential assignment of NP in context, we interpret this sustained negativity as reflecting higher processing cost to establish the referential dependency-link with the discourse referent that was less salient. Yet, an alternative account would be linked to the heavier retrieval cost from memory in reason of the [-active] feature of subject NP referent in Topic-related Context A compared to Subject-related Context B, thus interpreting the sustained negativity as reflection working memory load.

Moreover, contextual load also marginally interacted with the presence of violation in 300-650 ms Time-window (t = 1.91), with the load-associated negativity found only after a PB and not after the conjunction violation, showing that we observe contextual processes linked to the sentence-discourse interface property of this sentence-level hierarchy, only when Topic-Comment syntactic articulation is built.

In conclusion, the interplay between absence of PB and contextual saliency of different referents in the mental model of the given contextual information shows in this time-window a patterning that is astonishingly similar to the previous late effects at Topic position, and interestingly the context-related negativity is here stronger for noPB conditions, showing once more that the referential interpretation of a NP, even if facilitated by the saliency of the referent in the discourse-information storage, is made more difficult in the absence of an early sentence-level structure building.

5.4.3 Verb of the Comment

At Verb time-window, we observed a Prosodic Boundary by context interaction (t > 5.0) that was the opposite as the one observed in the previous time-window. It namely unfolded differently: increasing load on the Topic NP (Rich context) made no difference after a PB (PB x Context, t=1.7), but elicited larger negative responses in the noPB condition (0.76 V, t = 6.0).

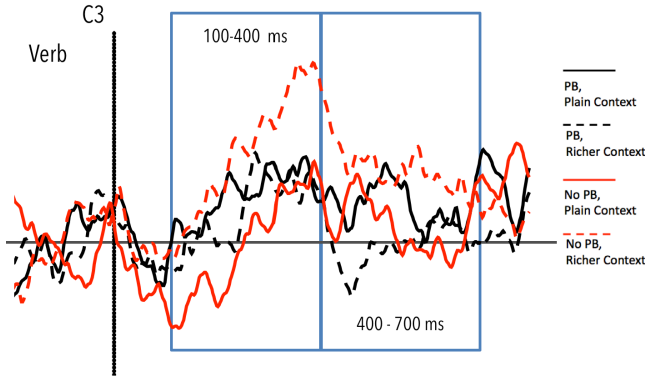


Figure 5.22 – Modulation by context and PB at Verb position at C3 electrode. Grand-average wave-forms over participants (n = 23), time-locked to the onset of the Verb of the Comment-clause for the Plain Context (solid lines), and the Rich Context (dotted lines) in two conditions, respectively PB (black), no-PB (red).

As illustrated in Figure 5.22, the kind of processes going on these two time-windows finally witness a context effect for noPB conditions. Coherently with what argued for

as being a given, an inferred or new referent.

subject position, these context-related effects are now observable in noPB conditions in that the Topic-Comment syntactic structure has been established at subject position. We interpret this as the mark of the simultaneous establishment of syntactic structure and of its sentence-discourse interface mechanisms.

Let us have a closer look to the processes in Early and Late time-window to understand their difference.

5.4.3.1 Earlier time-window 100 - 400 ms

Prosodic Boundary by context interaction

In the 100-400 ms time-window we observe an interaction between context load and Prosodic Boundary marking: rich context yields a more negative effect in noPB conditions than in PB sentences. Namely, increasing load made no difference after a PB (PB x Context, t=1.7), but elicited larger negative responses in the noPB condition (0.76 V, t = 6.0).

In this time window, we can finally see context effects in noPB condition which appears to have been delayed to verbal position compared to PB condition, while it was already present at the subject time-window for PB condition and absent here.

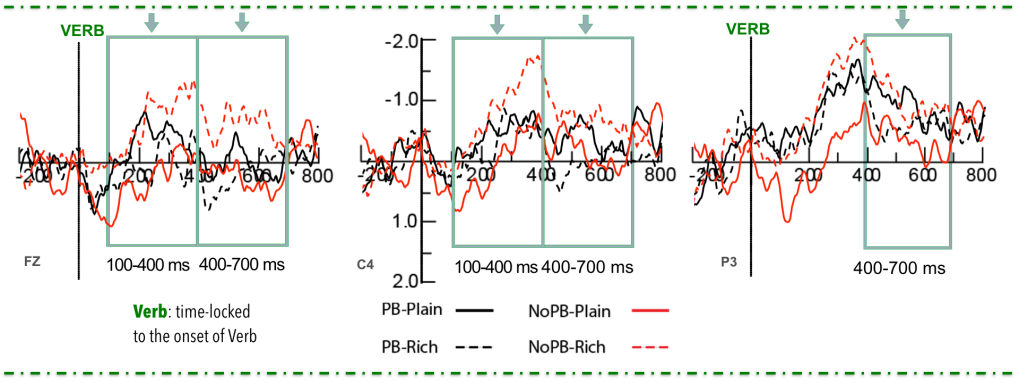


Figure 5.23 – Modulation by context and PB at Verb position in three electrodes FZ, C4 and P3 for respectively frontal, central and posterior scalp region. Grand-average waveforms over participants (n = 23), time-locked to the onset of the verb of the Comment clause for the two conditions, respectively PB (black), noPB (red), embedded in the Plain Context narrative (solid lines), and the Rich Context narrative (dotted lines).

Moreover, the processes going on in this early time-window show some difference with the following time-window, where we observe that the noPB condition is more positive, which is not the case in this early one. Comparing how PB and noPB pattern in this early time-window, we see that, while PB conditions are aligned, only noPB conditions diverge, and only the one embedded rich context shows an effect with negative polarity. We link this context effect to the one we observed in the early time-window between 150 - 300 ms at subject position where the opposite pattern was attested: noPB conditions were aligned and only PB conditions differed with the one embedded rich context showing a negative polarity effect. Hence, while we interpreted the syntax by context interaction as an interface effect between prosodically cued syntax and context, here we interpret this

effect as an interface effect between word-order based syntax and context interaction. In conclusion, this is showing that *context meets the sentence when the syntactic hierarchy between Topic and Comment is established*. We will now turn to the second interaction between our two manipulated factors that are found in the later time-window of Verb position.

5.4.3.2 Late time-window 400 - 700 ms

Task-effect and interaction between Prosodic Boundary and Context

In our model of the processing of Gapless 'Chinese style' Topic-Comment articulations, we predicted that two main processes should be going on at Verb position. A basic one linked to verb function in itself, the 'traditional' on-line Theta-role assignment, and the second linked to the experimental task we used to enhance the core Topic-Comment mechanism, namely the truth-condition evaluation of the Comment in the light of the Topic (cf. Reinhardt, 1982).

Task-related processes were hypothesized to comprise (1) a verification process to check if the action expressed by the verbal synonyms used in critical sentences, actually matched the configuration expressed in the narratives, which could be summarized in the following interrogation: *"Is the verbal content actually taking place in the spatial setting expressed by the Topic with the actor expressed by the subject?*; and (2) a decision process, the actual judgment. Moreover the fact that the verb position is the only position in the sentence where new lexical material appears – subject and Topic NPs are known lexical items from the narratives – would made it probable to observe stronger effects.

Given this configuration, we predicted that context-related effects would have been surely observed in this time window because of our task – a truth-value judgment of the sentence considering its coherence with the context – and because our task can be seen as enhancing the 'natural' truth-conditional evaluation of the Topic over the Comment predication.

We see in this time-window in PB by Context interaction, where Rich context is more negative for noPB, and not for PB sentence. Two explanations can be advanced for this: (1) the absence of PB makes again the sentence-discourse interface anchoring mechanisms more difficult, or, (2) the task-related evaluation happening at the verb is at play.

The pattern of response in this time-window can help us disentangle the two options. Compared to the pattern of response observed in the previous time-window, where the two PB conditions (black, solid and dotted) overlapped, showing by this that the sentence-discourse interface mechanism was ongoing only for noPB conditions. Here the Rich context condition is more positive together with the noPB one in plain context, which is taken to signify that the process happening in the late verb time-window is, as hypothesized, linked to the truth-value judgment. Rich context, compared to plain context in PB conditions, facilitates the evaluation of the truth-conditionality of the Topic on the comment predicate, in that the accessibility of the Topic referent is higher, as predicted. Importantly, the evaluation process for the judgment appears here to be greatly impacted by the presence of PB which inverts this pattern suggesting that the late discovery of the Topic-Comment articulation not only differs the contextual linking processes yielded by its sentence-discourse property, but it also makes the establishment of the framing truth-conditional relation between Topic and Comment predicate more

demanding when the context-load is heavier. These effects converge with the participants' ratings reported for truth-value judgments in previous chapter, where sentences with PB were more likely to be judged as congruent than sentence without PB, probably because of the effects we see in this time-window.In conclusion, the PB aligned with the sentence-level syntactic boundary between Topic and Comment reveals to be once more a crucial factor triggering the relation with context information.

5.5 Summary and Conclusion

Summary of the main findings

This study aimed to unveil the neural processing of Topic-comment sentences marked or not by a Prosodic Boundary marked, and to investigate when and how contextual information interacts with prosodic information during speech comprehension. We examined the different processing strategies under different contextual information load and in presence or absence of a Prosodic information offering an early cue of the Topic-comment articulation of the sentence-unit.

We can summarize our findings as follows:

- A- Building-up a hierarchical structure just with word-order as a syntactic cue (without Prosodic Boundary marking) delays the integration of contextual informational load, we observed context effect only late in the sentence at the verb time-window. Thus, the PB conveying the syntactic articulation of the sentence can be interpreted as an early trigger the sentence-discourse interface and its contextual information integration processes.
- B- A negativity rather than a CPS (Closure Positive Shift) was found for the Prosodic Boundary detection, suggesting that neural response may vary responding to different functions and types of Prosodic Boundaries. Critically, Prosodic Boundary in Topic-comment constructions does not relate to disambiguation because such sentences can only have one possible interpretation yielded by its word-order. Namely, as early as NP2 position, word-order and semantic cues of the Scene-setting Topic construction offer enough information for the parser to engage in the right sentence structure building.
- C- The violation effect of Coord-He conditions preceded PB effect at Subject NP, indicating that lack of Prosodic Boundary in Chinese-style Topic was not processed as a morpho-syntactic violation.
- D- Effects observed for the referential interpretation process of the Subject NP show that the encoding of the referents in the mental model of narratives (i.e. storage strategy) has an impact on the accessibility of the discourse referents, and therfore yields a heavier processing cost when the referent is less salient as it is the case for the subject referent in Topic-related context (i.e. Rich story).
- E- Notably, we observed a modulation on Context integration effects by the presence of PB happening along the Comment sentence. While contextual load effect on Subject NP occurred only for PB-marked sentences and not on the later verb time-window. For noPB Topic-comment sentences the contextual load effect was delayed on the verb. This pattern of results presenting a temporal shift suggested that the presence of prosodic cues anticipates the building of the syntactic hierarchy between Topic and Comment and therefore facilitated the use of contextual information. Namely, in noPB condition the syntactic structure is discovered later on word-order grounds (at subject position), resulting in the observed later sensitivity to contextual embedding.
- F Crucially, the context effects we observed in the PB-condition in the late time-window of the Topic last Syllable and at the subject position in the early time window, are suggesting that during on-line parsing the sentence-discourse interface properties of in Topic-

Comment construction activate a process of reiterately bridging (or living-up) discourse-context information to understand the Comment, as schematically illustrated in Figure 5.24.

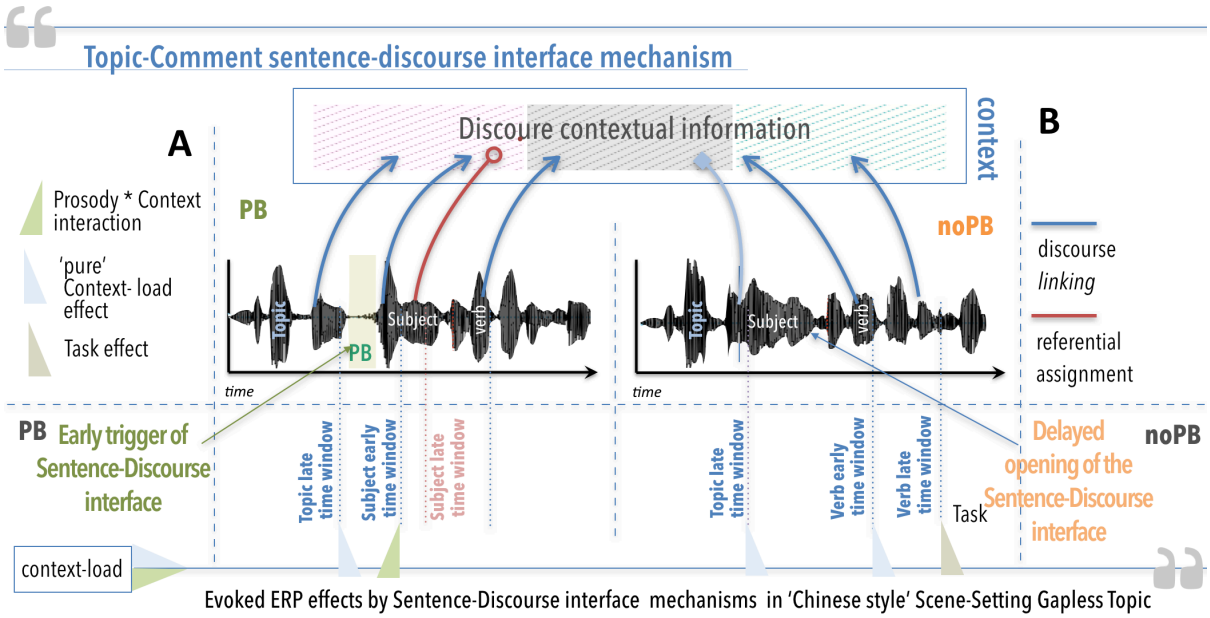


Figure 5.24 – The ERP results contributed to prove that the Topic-comment syntactic articulation structurally activates sentence-discourse mechanisms. And the dependency-link between Topic and Comment in Gapless 'Chinese style' Scene-setting Topics is essentially achieved by a mechanism of online successive integration of information about the Topic element. Moreover, these discourse integration processes of the are reiteratively the taking place at different time-windows in the comment clause according to how early the syntactic articulation between Topic and comment was discovered.

Conclusions

This study offers two pieces of evidence for the incremental processing of prosodic information in general and particularly for Gapless Scene-setting Topic on-line mechanisms.

First, this study demonstrates the immediate use of prosodic information to build sentence structure as soon as pre-boundary cues are available, thus confirming that prosodic and syntactic information interact early in speech perception and take place in the early stages of phrase structure building.

Second, the prosodic information and its early cuing for the syntactic hierarchy between Topic and Comment activates the sentence-discourse interface property of Topic-Comment articulation yielding an interplay between prosody and contextual information as early as at the end of the prosodic boundary effect, where it was observed to facilitate the integration process of discourse context informational-load (in a later time-window).

Critically, we observed a modulation effect of the PB on context-related effects. Contextual effect on Subject NP occurred only for PB sentences, contextual load did not

show difference for PB sentences later on the Verb, whereas it actually did show an effect for noPB sentences. This pattern of results suggests that the presence of prosodic cues anticipates the building of the syntactic hierarchy between Topic and Comment, and therefore facilitates the use of contextual information. Namely, in noPB condition the syntactic structure is discovered later, resulting in the observed later sensitivity to contextual load.

Hence, the first conclusion we get to, *is that context meets the sentence at the prosodic boundary.*

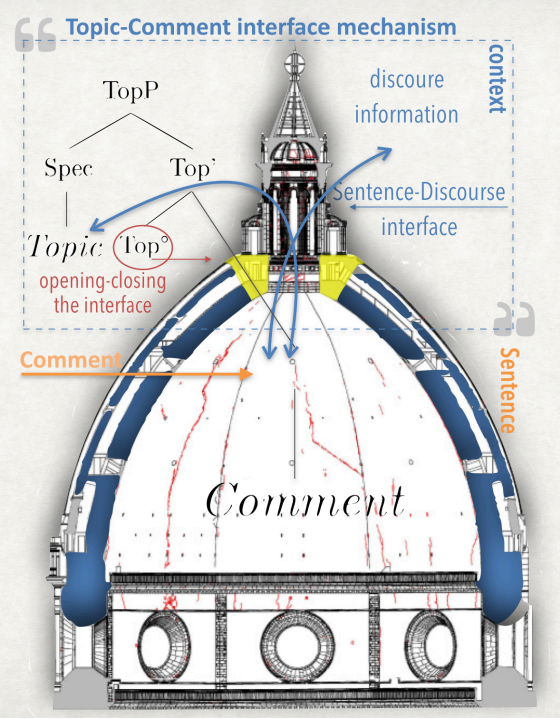


Figure 5.25 – A metaphorical representation of Topic-Comment sentence-discourse interface mechanism. The ERP patterns observed in the experiment suggested that the presence of prosodic cues facilitated the use of contextual information in that it cues earlier for the building of the Topic-Comment sentence-level hierarchical structure. Namely, in noPB condition the syntactic structure is discovered later, resulting in the observed delayed sensitivity to contextual load.

In more general terms, the comparison of our data with previous findings indicate that, in presence of prosodic cues for syntactic structure, the sentence processing system deals with context interface in the later integrative stages of sentence processing (after 550 ms from the beginning of the Prosodic boundary). Specifically, our findings showed that the sentence processing system is sensitive to prosodic cues, and that prosodic cues facilitate sentence comprehension by:

- giving an overt linguistic cue for the building up of the sentence hierarchical structure (in our case into a Topic and a Comment part) and
- facilitating the integration of contextual information in the sentence-unit.

From the point of view of cognitive resources, we speculated that the pause separating the two intonational units of Topic-Comment sentences might be biologically needed to recruit cognitive resources to live-up the context information, which is necessary to understand the relevance of the comment clause in a Topic-Comment 'Chinese style' articulation.

Notably, cerebral-level processing is confirming the behavioral results showing no violation related activity in Topic-Comment sentence embedded in plain context where the Topic referent is less salient.

The comparison with coordination violation and absence of PB, suggested that lack of PB after Chinese-style Topic was not processed as a morpho-syntactic violation and that Chinese Topic-Comment constructions can be actually parsed only on the bases of word-order cues. This might be due either to a default Topic-prominence strategy or to the fact that Chinese syntax generally assigns grammatical roles by positional rules.

In conclusion, sentence-discourse interface mechanisms in Topic-Comment constructions is taking place effortlessly as soon as the syntactic hierarchy between Topic and Comment is built, that is for PB condition as early as at the end of prosodic boundary of the Topic, and in noPB when word-order syntactic evidence in gathered at subject position. This pattern of results suggested that the presence of prosodic cues anticipates

the building of the syntactic hierarchy between Topic and Comment and therefore facilitated the use of contextual information. As in noPB condition the syntactic structure is discovered later, this results in the observed delayed sensitivity to contextual load. ***Taken together, the ERP effects reported here point to a sentence parser that waits the signal of syntax to engage into contextual discourse linking processes***. Further research is needed to confirm this with other sentence structures or other Topic types, that structurally or semantically need more inferential relations to be processed.

In broader terms, these ERP findings are giving neuro-linguistic evidence – from the point of view of incremental on-line sentence processing – that Topic-Comment syntactic articulation triggers a sentence-discourse interface mechanism that is observed during the whole Comment clause. This should remind us of the syntactic structure tree-like representation of this sentence construction, where Topic occupies a position in the sentence-discourse interface CP-layer, namely overarching the whole Comment as in figure 5.25. We reproduce here our "cupola metaphor" incrementing it with the experimental results these last two chapters offer.

Chapter 6

Cortical responses to syntactic transformations in French

"Las de tous ceux qui viennent avec
des mots
Des mots, mais pas de langage,
Je partis pour l'île recouverte de
neige.
L'indomptable n'a pas de mots!

Ses pages blanches s'étalent dans tous
les sens.
Je tombe sur les **traces** de pas
D'un cerf dans la neige
Pas des mots, mais un langage."

*["Weary of all who come with words,
words but no language,
I make my way to the snow-covered
island.*

*The untamed has no words.
The unwritten pages spread out on
every side!*

*I come upon the tracks of deer in the
snow.
Language but no words."]*

Tomas Tranströmer, *Baltiques*,
En mars 79

After having focused in the previous chapter on the on-line processing of sentence hierarchy and the sentence-discourse interface mechanism in absence of intra-sentential dependency-links, we now turn to investigate a crucial and unique property of the sentence-unit: (1) the possibility of establishing a dependency-link between two places in the sentence where only one position is phonologically filled, and (2) the representational and processing complexity effects that this dependency-link yields in cortical responses.

The two main dimensions of syntactic complexity we will address here are fundamen-

tally rooted in the reflection presented in chapter 2 about the neural substrates of the tree-like representation of the sentence. As poetically expressed in the above epigraph by Swedish aphasic poet and 2011 Nobel prize Tomas Tranströmer, this chapter investigates the ***traces*** the language (system) leaves behind to establish dependencies between one place and another in the sentence.

With this chapter, we enter in the experimental investigation based on fMRI methodology of this manuscript. This implies leaving the online incremental processing gaze we cast on the sentence-unit temporal unfolding, to move to a more neurally localized, though less temporal, view on how the mind and the brain represent and process the structural complexity of the sentence-unit as a whole. This change in point of view on the sentence and its structural complexity is not only due to a change of gear in the brain-imaging method we apply to investigate the neural substrates of syntax, but is also motivated by our interest in the psychological reality of the different types of movement theorized in linguistics[1].

Crucially, this chapter will address these two views (i.e. representational and processing-oriented) on syntactic complexity focusing on the final structural syntactic-tree structure of complex sentences irrespective of their different movement derivation and on the possibly different cognitive and cortical resources that diverse types of movement-derived structures can require to be processed.

Why is movement so interesting: two functions on overt item in the sentence

By approaching the issue of syntactic transformations and of the empty syntactic elements that are left when dislocating a sentential element[2], we actually address one of the most crucial, and most probably uniquely human linguistic capacity, that of understanding a single sentential element at two different places in the sentence.

For instance, in the sentence *Who should I call ___?*, the single overt element *who* plays a different role at two different positions in the sentence. It acts within the sentence as if it is occupying two structural positions: in sentence-initial position, where it has the function of clause-typing the sentence, yielding interrogative mode, and in post verbal position, where its basic semantic object function is determined.

In such cases, linguistic theory postulates that complex sentences are constructed out of movement operations which leave syntactic *'empty'* positions vacated by the syntactic operation of displacing a sentential element. Hence, the two different functions are linked to two different positions in the sentence structure: one linked to the syntactic position where the syntactic element lands after displacement, and one in the extraction site from which it was removed (e.g. post-verbal object position).

Positing the so-called *traces* at the original location (i.e. extraction-site) of displaced words, crucially allows the same element to have two roles in the sentence, and the constituent ultimately gets the theme object role by being linked with its *trace* via what is called a "chain". Thus, several local relations can be established at once by forming a dependency-chain between its actual position and the position it was moved from.

1. This will also more or less tangentially lead us to resume to the central issue of the difference between *representation* and *processing* of the sentence's internal structure, we addressed in chapter 1 and 2 (§1.3.5, p.47 or 160).

2. Note that in contrast to next chapter where we will compare gaps and null pronominal (i.e. little pro), here we will only investigate traces.

The *Copy theory of movement* offers an account for the interface with semantic interpretation, in that all the relevant information for sentence interpretation is expressed at each site of the syntactic structure. Hence, the full representation of the above sentence would be having a *silent copy* of the object noted with angled brackets: "Who should I call <who> ?". It is in fact at post-verbal position that the constituent *who* receives from the verb 'read' the theme role that would have been assigned directly to it if it hadn't moved, while in its sentence-inital position (specifically, Spec CP) it has a clause-typing role.

In this regard, the precise theoretical and formal approach to these linguistic structures (cf.2 §2.4.2.1 p.162) analyzes these syntactic configurations as generated by a *syntactic-movement* operation that is actually re-merging an element that is already internal to the sentence thanks to a *Search* operation step that is added to Merge operation in order to first identify the candidate for Merge (i.e. Phrasal Merge). Only after this Search step, the suitable candidate is internally merged with the whole structure[3].

Movement appears then to be a particular type of Merge that instead of taking an external linguistic material (e.g. a lexical array) would move a constituent within the existing structure to a new position. In short, we can then conclude that Movement appears to be a composite operation combining a *Search procedure* and a consequent possible Phrasal *Merge operation*.

This possibility of re-merge allows the operation of syntactic "movement" or displacement, where importantly a single constituent acts within the sentence as if it is occupying two structural positions. Therefore, the gap position in the sentence can be seen as the place where the displaced constituent is re-merged. As we saw in chapter 2 (§2.4.3.2, p.193) this last observation has an important impact on the psycho-linguistic processes happening at trace position.

In sum, focusing on the cerebral representation of syntactic *traces* is a way to approach what is covert and silent in language but structurally needed to build sentence structure and therefore understand sentences.

Hence, one of the main motivations of this study is to investigate the cerebral substrates of the syntactic complexity dimension linked to movement-related processes theorized by derivational syntax. In this way, we will look for what the fine-grained complexity dimensions formalized through syntactic-tree representations of the sentence-unit's structural hierarchy can actually reveal about the neural implementation of the encoding of syntactic information across the different brain regions that constitute the Sentence Network.

By investigating the cortical responses to the different movement transformations in French question formation and declaratives, our aim is twofold:

1. the processing complexity linked to the different types of movement operations that generate the wide variety of sentence structures under examination, in order to assess whether different types of syntactic movements -—wh-movement, Verb-movement, clitic- movement and NP-movement —- are associated to specific brain activation patterns
2. representation increasing syntactic complexity of the syntactic-tree representation of the sentence irrespective of the type of syntactic-movement derivation. Thus, detecting brain areas sensitive to representational complexity of sentence syntactic structures more broadly.

3. Here again surfaces the issue of what are the processes happening in production versus comprehension. We could advance that the building of dependency is happening in production while the establishment of the dependency is what needs to be calculated during comprehension.

In conclusion, this fMRI study aims at revealing the neural signatures of the different linguistic transformations involved in French declaratives and Interrogatives. Specifically we will investigate if different types of syntactic movements and sentence derivations that were theorized by linguistic theory like NP-movement, Verb-movement, wh-movement and clitic-movement actually give rise to different activation patterns in the distributed cortical network for sentence-unit that has emerged in recent years (Pallier et al. 2011, Fedorenko et al. 2012, cf. Sentence Network §1.4.4 and 1.7, see Figures 1.3.4 and 1.11).

Overview of the contents of this chapter

6.1 Experimental goal and hypotheses: Why study movement and *traces* in the brain

As theoretically and experimentally argued in chapter 2 (§2.4.2), linguistic and psycholinguistic evidence have converged in grounding a certain understanding of sentence syntactic complexity as being the result of series of transformations of a basic (i.e. base-generated) sentence pattern.

By moving wh-elements, verbs, nominals, and clitics in French sentences, we aimed on the one hand to increase the derivational complexity of the sentences to uncover the neural substrates of these types of syntactic complexity operations. On the other hand we also aimed at investigating the finer-grained complexity dimensions that is implied by different types of movement operation: Verb-movement, cl-movement, NP-movement and *wh*-movement as illustrated in Figure 6.1.

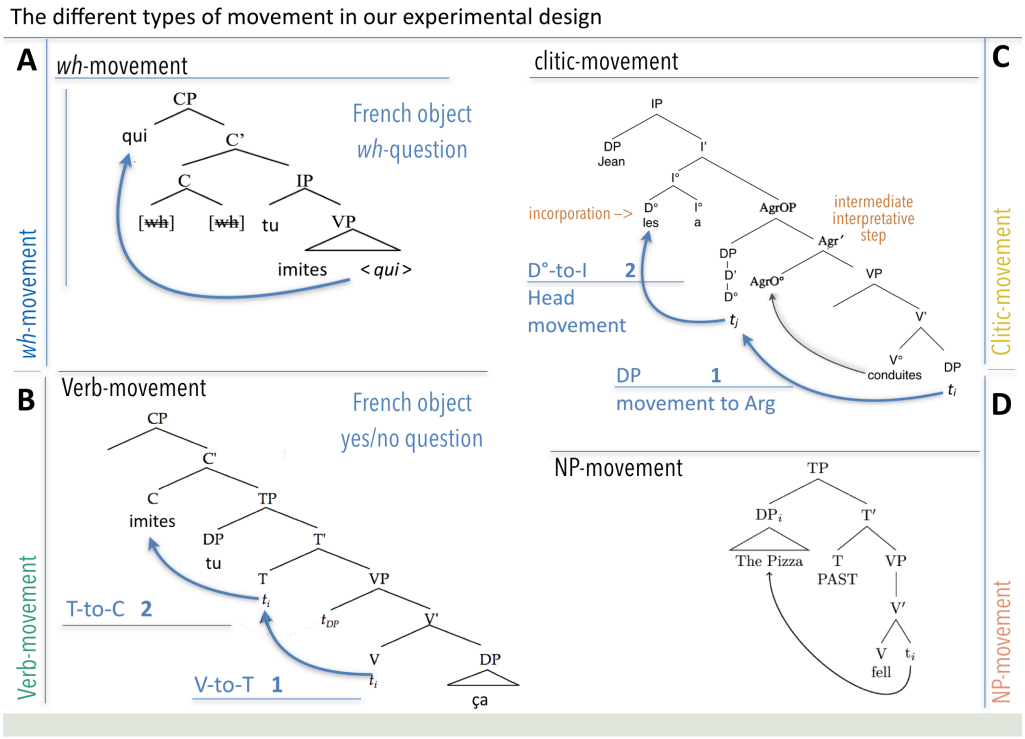


Figure 6.1 – Syntactic-tree diagrams illustrating the syntactic derivation and complexity of different types of movement-derived constructions included in our study: (A) wh-movement, (B) Verb-movement, (C) cl-movement and (D) NP-movement.

Concretely, while studies on linguistic behavior aphasia provide insight into the neural basis of movement-derived complex sentence processing, evidence for the processing difference of sentences where syntactic operations have displaced constituents from their canonical position also comes from the performance of healthy subjects. Thus, *syntactic movement* can be considered as a major contributor to the perceptual complexity of sentences (Fodor et al., 1974; Neville et al., 1991; Cohen and Mehler, 1996).

Chapter 2 reviewed the findings coming from different experimental fields, supporting the psychological reality of syntactic movement (i.e. the structures assigned by syntactic transformations can account for processing behavior during complex sentence understanding). All in all, the facts we reported show the relevance and explanatory potential of movement-based accounts of syntactic complexity offered by syntactic theory. As a brief reminder, we considered:

- (1) Empirical evidences from aphasiology and Specific language Impairments are showing that the impairment patterns of different language disorders can actually be accounted by movement-related dimensions of syntactic complexity. Moreover, the training or gradual recovery patterns of aphasic production can also be explained by the different structural complexity yielded by different types of syntactic movement and the different landing-site they target (cf. Friedmann's or Thompson's research respectively on Agrammatic Aphasia natural recovery and production of wh- and NP-movement constructions after training).
- (2) Behavioral Experimental results in psycho-linguistic studies across different languages, showing stable effects of 'reactivation' of the moved-element at the *trace* position in different types of movement (i.e. wh-movement and NP-movement in Cross-Modal Lexical Priming paradigms, cf. §2.4.3.2)
- (3) Neuro-imaging experimental results showed through adaptation paradigms and simple sentence comprehension that different neural signatures for different syntactic movements are observable in fMRI recordings of brain activity.

The characteristics of the different types of movement *in* our experiment

	A wh-movement in object questions	B Verb-movement in yes/no questions	C clitic-movement	D NP-movement in Unaccusatives	*main characteristics*
	object *wh*-element	tensed verbal element	object DP-element	object NP-element	Nature of the moved element
	A-bar movement	Head-Head movement (A-mov)	NP-movement + Head-movement	NP-movement (A-mov)	Type of movement (A versus A -bar)
	long-distance	long-distance	local-distance	short-distance	Long /Short Distance
	1 step	2 steps (cyclic): [V-to-T] + [T-to-C]	2 steps: [Comp, VP-to-AgrOP]+[D°-to-I°]	1 step	Movement progression
	Spec,CP	C°	I°	Spec,TP	Targeted Landing-site
	wh-criterion Interrogative	Interrogative	Declarative with known discourse referent	Empty subject positon declarative (*case assign*)	Trigger and clause type
	Object wh-questions	**Verb-inverted yes/no questions**	**Object clitic placement**	**Unaccusative verbs**	**French construction**

Figure 6.2 – The syntactic characteristics of (A) wh-movement, (B) Verb-movement, (C) cl-movement and (D) NP-movement.

More specifically, in our linguistic review of movement-related complexity in chapter 2, several differences between these types of moments were pointed out:

1. the nature of the moved element: a NP or a wh-element or a Head,
2. the number of steps different movement types have to achieve to reach the landing-site in the sentence tree-structure,
3. the Sentence Domain targeted, which crucially yields interrogative vs. declarative interpretation when targeting the CP-layer in our stimuli, or a more in more local movement targeting the TP-layer,

4. the structural distance characterizing the dependency-links established by different movement types.

These different complexity dimensions guided our selection of the different types of movements for the experimental design. As illustrated in Figure 6.2 we chose:

1. wh-movement was selected for its characteristic long-distance dependency-link targeting a high position in the CP-layer and its clause-typing properties;
2. Verb-movement was selected for its two-stepped and Head-to-Head movement, its clause-typing properties, and the verbal nature of the moved element.
3. Clitic-movement was selected for its hybrid nature comprising both NP movement and Head-movement, its clause bounded-ness and two-stepped derivational progression;
4. NP-movement was selected for its locality and its particular argument structure

6.1.1 Empirical evidence for different types of movement and our hypotheses

6.1.1.1 *wh*-movement in questions: a particularly difficult movement

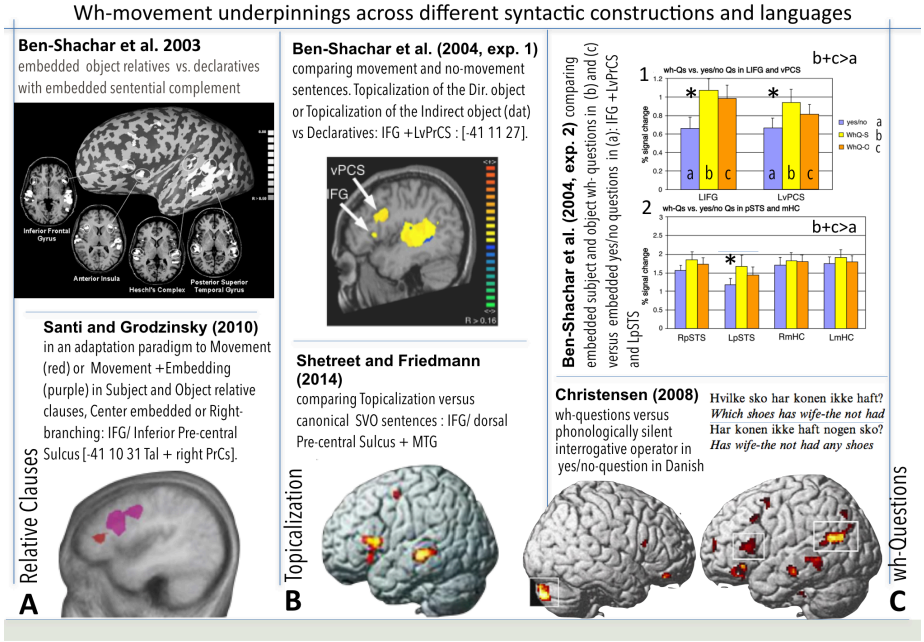


Figure 6.3 – Some representative studies contrasting sentence constructions derived by wh-movement against sentences with no wh-moment. (A) Effect of Relative clauses against sentential embedding (up), Effect of adaptation to wh-movement in red (down). (B) Effects of Topicalization compared to basic SVO in English (up) and in Hebrew (down). (C) Effects for wh-question compared to yes/no questions.

As reported in the fMRI literature and by research on aphasia, *wh*-movement is constantly reported to be recruiting relatively large cognitive and cortical resources. While the reader is invited to refer to chapter 2 for the rich set of evidence accumulated in aphasiology literature (§2.4, p.157), Figure 6.3 summarizes a few representative fMRI

findings contrasting different sentence constructions derived by wh-movement with non-moved baselines, which converge in indicating three main cerebral clusters of increased activation, encompassing Broca Complex (IFG), Precentral Complex (ventral and dorsal) and a temporal Middle to Posterior area. We rely on these results to formulate our predictions for the contrast opposing the object wh-questions present in our design to the declarative baseline.

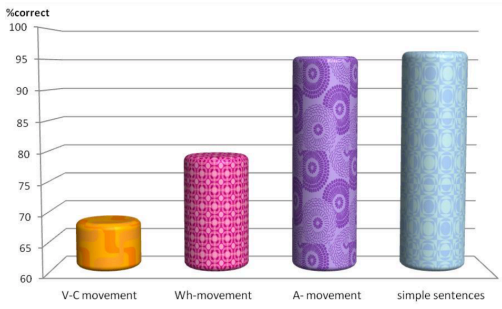


Figure 6.4 – Percentage of correct repetition on the various types of syntactic movement compared with simple sentences without movement in the hearing impaired group. Adapted from Szterman and Friedmann, 2014.

While the majority of the studies investigating wh-movement have focused on relative clauses, their possible embedding and their different linearization of subject and object, , fewer studies selected *wh*-questions as we did. However, a multi-perspective description of *wh*-questions is offered by a series of studies on *wh*-movement impairment in children (Syntactic Specific language impairments - S-SLI), that complete the extensive overview on selective syntactic impairment we delineated through Friedmann and colleagues' work in chapter (p.191, Szterman and Friedmann, 2006). Specifically, S-SLI children with a hearing impairment clearly show different impairment of different types of movements, as illustrated in the percentage of correct repetition of sentences derived through different movement types reported in Figure 6.4 (Szterman and Friedmann, 2014).

As summarized in Table 6.1, Van der Lely and colleagues used various psycho-linguistic and imaging methodologies to have a global approach of the issue of production and comprehension of wh-impairment in syntactically impaired children (S-SLI).

The picture that emerges from production elicitation, grammaticality judgments, online cross-modal priming, and EEG recordings in this Selectively impaired population, is that these children generally fail to compute the syntactic relation the *wh*-word and its extraction-site across all different experimental attempts an interpretation that is confirmed also by studies in other languages (e.g. Italian, German, French, Hebrew and Greek)[4].

As shown in the examples in Figure 6.1A children and adolescent fail to compute filler-gap relations and to mark tense, when eliciting embedded wh-questions from them they omit the auxiliary 'did' and actually move the wh-word only partially to some intermediate position as shown in Table 6.1C. Relevant data on production of object relatives can be found in Friedmann et al. (2015).

Moreover, while they reject semantic violations and recognize grammatical utterances, they again fail to detect errors in filler-gap relations and tense in grammatical judgement task in Table 6.1B. Interestingly, Cross-Modal Priming shows that in children and teenagers with G- SLI, the *wh*-filler does not prime its related word at the gap position (see Table 6.1D). Lastly, the results from an ERP study confirms that while S-SLI teenagers have an appropriate response (i.e. N400) to violations of lexical semantics – compare (i) and (ii) in Table 6.1E –,they show no responses to the violations of expectation based on syntax such as on the word *clown* in (iv) (Table 6.1E).

4. See for German Hamann et al. (1998); for French Jakubowicz and Tuller eds. (2008); for Hebrew Friedmann and Novogrodsky (2007, and more recent findings); for Greek Stavrakaki and Van der Lely (2010); for Italian see Cantiani et al. (2010) and Pizzioli et al. (2007).

Task Van de Lely and Pinker (2014)	Description	Lead-in or prime	Control child's typical response	Child with G-SLI's typical response
Elicitation of 'wh'-questions	'wh'-subject question	Someone saw Mrs White. Ask me who.	Who saw Mrs White?	Who someone see Mrs White?
A	'wh'-object question	Mrs White stole a ring. Ask me which one?	Which ring did Mrs White steal?	Which [....] Mrs White stole the ring?
Sentence Judgments	Semantic judgment	(i) Barbie bakes the cakes (ii) Barbie drinks the cakes	(i). Yes (ii). No	(i). Yes (ii). No
	Syntactic judgment: Tense	(i) Who did Mrs White see? (ii) Who did Mrs White saw?	(i). Yes (ii). No	(i). Yes (ii). Yes
B	Syntactic judgment: 'wh'-movement	(i) Who did Mrs White see? (ii) Who did Mrs White see someone?	(i). Yes (ii). No	(i). Yes (ii). Yes
Elicitation of Embedded 'wh'-questions	Embedded subject 'wh'-question	Joe thought someone hit the man. Ask me who?	Who did Joe think hit the man?	Who did Joe think who hit the man?
C	Embedded object 'wh'-question	Joe thought Mr White hit someone. Ask me who?	Who did Joe think Mr White hit?	Who [...] Joe think Mr White hit someone?
Cross-modal Sentence–Picture Priming	The child is shown a picture of a rabbit or a ladder at one of three points during a spoken sentence and presses a button indicating whether it is animate or inanimate. Response time for position [2], after the adjective, serves as a control. Priming (faster responding) at position [1] shows sensitivity to basic verb semantics (*give* + recipient). Priming at position [3] (the 'wh'-gap) shows sensitivity to syntactic movement: the rabbit matches the 'who' which had been moved from that position.	Baloo gives a long carrot to the rabbit at the farm. Who did Baloo give [1] the long [2] carrots to [3] at the farm?	[1] No priming. [2] No priming. [3] Priming.	[1] Priming. [2] No priming. [3] No priming.
D				
Event Related Potentials (ERP)	Violation of expectation based on semantics (ERP time-locked to the underlined word)	(i). Control: Barbie bakes the cakes in the oven. (ii). Violation: Barbie bakes the people in the oven.	N400	N400
	Violation of expectation based on syntax (ERP time-locked to the underlined word)	(iii). Control: Who did Barbie push the ball into? (iv). Violation: Who did Barbie push the clown into the wall?	ELAN, P600	Right lateralized N400, P600
E				

Table 6.1 – This table summarizes different studies conducted by Van der Lely et al. with children showing a Grammatical specific Language Impairment (G-SLI), in wh-questions (who, what, which) in production elicitation (A and C), grammaticality judgments (B), on-line cross-modal priming (D), and EEG recordings (E). Adapted from Van der Lely and Pinker (2014).

This behavioural and neuro-imaging description of *wh*-movement, leads us to note that *wh*-movement actually displays two types of complexities in wh-questions, both argumental semantics process linked to the identification of *who did what to whom* and Scope-Discourse semantics linked to clause-typing of the utterance as a question. And these two functions are crucially found in different places of the sentence: (i) sentence-initial position for Scope-Discourse semantics, and Argumental semantics at the gap position[5].

6.1.1.2 Evidence for the different processing of different movement types

Given these fMRI results and the clear syntactic complexity effect witnessed by *wh*-questions in syntactically impaired children and teenagers we can move to the neuro-linguistic findings about other movement types. One of the first confirmations for the

5. We reserve for further research, the testing of the possibility to think about movement contrasts:
   - A. in terms of an Attraction to landing-site (i.e. Derivational approach), or
   - B. in terms of Relativized Minimality, a more Representational Approach.

psychological reality of different types of movements comes from the different acquisition patterns that are observed for Verb-movement, precisely V-to-C, compared to wh-movement in children with normal language development.

The results of a simple repetition task in Figure 6.5, show a strong correlation between different structures with the same movement. While sentences with wh- and V-to-C movement are repeated significantly worse than the SVO baseline, there are significant differences in performance between the different movement types, and clearly the repetition of NP-movement was performed with less errors that *wh*-movement, which in turn showed less errors than V-to-C movement.

Another interesting pattern emerging from this study is shown in Figure 6.5B, where we can see that all children who had already acquired V-to-C movement mastered also both *wh*- and NP-movement in Unaccusatives, and all children who already acquired *wh*-movement had also acquired NP-movement.

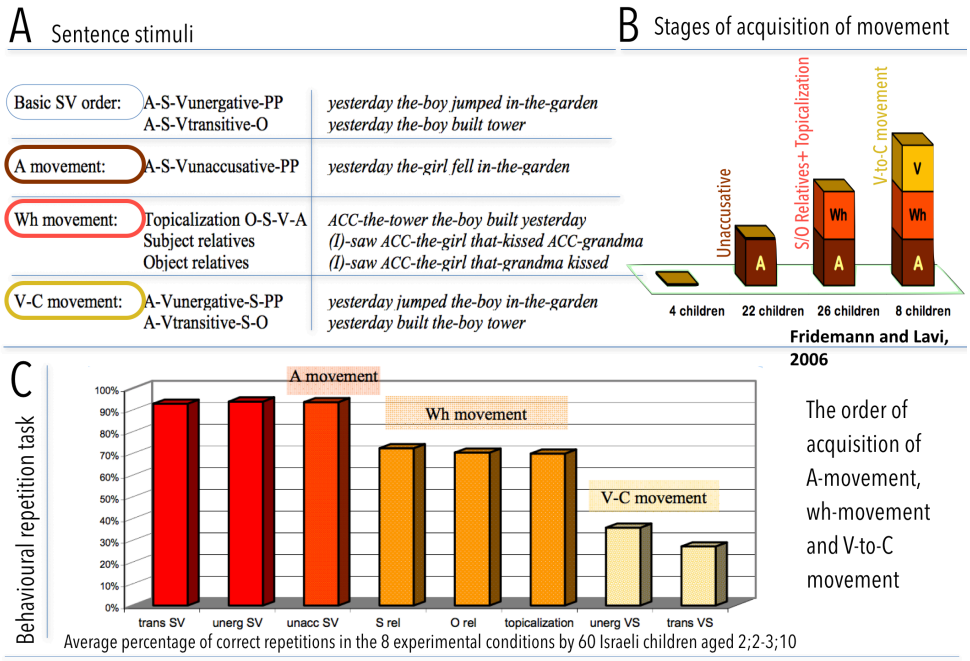


Figure 6.5 – (A) Set of 8 experimental sentences altering different types of movement derivation. (B) The four stages of movement acquisition revealed by this study. Barplots represent the a gradual acquisition of movement types at individual level acorss the 60 children that participated to this study. (C) The order of acquisition of A-movement, wh-movement and V-to-C movement. Average percentage of correct repetitions in the 8 experimental conditions by 60 Israeli children aged 2;2-3;10. Adapted from Friedmann and Lavi (2006).

Further evidence for the neural distinction of different movement types comes from the attested neuro-psychological pertinence of this linguistic distinction. As extensively documented in chapter 2, movement-related sentence complexity is documented in neuropsychology to be impaired in agrammatic populations. For instance, we reviewed evidence that resumptive strategy (i.e. movement-free) compared to syntactic-movement leaving a gap in relative sentences is impaired (cf. 2.39 Friedmann, 2006, Fig. 2.40, p.

192).

To this we can add the agrammatic rehabilitative results obtained through Thompson's treatment/research paradigm (Treatment of Underlying Formes - TUF - Thompson et al., 2010). While generalization to untrained sentences relying on both NP- and *wh*-movement is observed, the results clearly show that generalization patterns are constrained to type of movement: Training wh-movement structures resulted in generalized production of untrained wh-movement structures without influencing production of NP-movement structures, and viceversa (for a detailed discussion see chapter 3, Figs.2.33 and 2.34, p.177).

Not only these data show that derivational linguistic properties of sentences influence sentence production breakdown and recovery in aphasia, but they indicate that movement to an argument position as in NP-movement is neuro-linguistically distinct from movement to a non-argument (A-bar) position as required in *wh*-movement.

It should be added that psycho-linguistic approaches of empty syntactic elements left behind by NP- and *wh*-movement revealed through cross-modal lexical priming (CMLP) paradigms (see §2.4.3.2, p.194) are actually yielding different reactivation effects: for *wh*-movement a pre-trace priming effect is observed (Nicol and Swinney, 1998), while for NP-movement a delayed reactivation is found only after the verb (Osterhout and Swinney, 1993).

Interestingly, a more recent CMLP study in Dutch showed that Verb-movement to second position features a different reactivation pattern compared to *wh*-movement. Namely, the priming effect for related probes was observed at all tested positions, including the intermediate position (de Goede et al., 2009), instead of being only present at the related gap position. This finding brought the authors to conclude that the meaning of moved verbs remains active during the entire sentence, thus showing a processing difference between these two types of movements. Last but not least, it is crucial to note that while agrammatic patients generally show an impaired use of wh-movement derived structures, they have no problem with Head movement (Grodzinsky, 2007).

Notably, delayed reactivation effects have been reported for unaccusative verb structures, like "The leaf$_i$ fell t_i" that are also known to involve NP-movement in Hebrew (Friedmann et al. 2008) and Spanish (Bever and Sanz, 1997) (cf. §2.4.3.2 , p.195)[6]. Hence, we build on the different effects reported for wh-, NP- and Verb-movement traces in CMLP paradigms to hypothesize that different patterns will be observed also in fMRI recordings.

In neuro-imaging, *wh*-movement has largely been investigated in different syntactic structures encompassing relative clauses, *wh* questions, topicalization and clefting (Ben-Shachar et al., 2003 and 2004; Santi and Grodzinsky, 2010; Shetreet and Friedmann, 2014; Den Ouden et al., 2012), as illustrated in previous Figure 6.3.

Only a few isolated fMRI studies investigated the issue of other types of movements. For instance, Den Ouden and colleagues contrasted the **production** in Dutch of sentences with Verb-movement, from final to second position, to sentences in the canonical SOV order (Den Ouden et al., 2008), and observed an increased activation in left Middle and Superior Frontal Gyrus. This study constitutes the first neuro-imaging result demonstrating that the pattern of cortical activation elicited by the production of verb movement to C position appears to be different from the one repeatedly reported in the

6. Note that Friedmann and colleagues compared sentences with unergative verbs to sentences with unaccusatives, both alternating and non-alternating unaccusative.

literature for *wh*-movement comprehension.

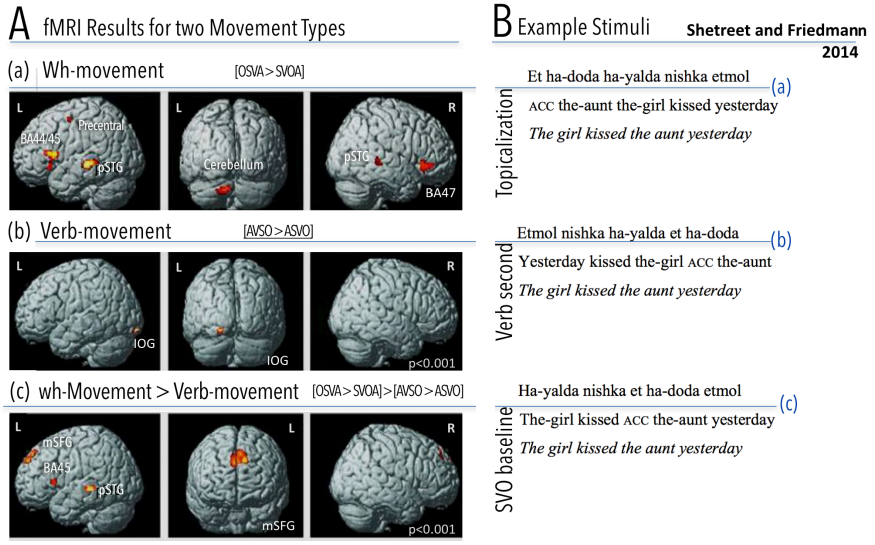


Figure 6.6 – (A) Group-average brain maps for critical contrasts showing (a) wh-movement and (b) V-movement effects against SV0 baseline, and in (c) wh-movemnt versus V-movement contrast mediated by two different baselines, i.e. [OSVA > SVOA] > [AVSO > ASVO]. (B) Experimental stimuli examples. Adapted from Shetreet and Friedmann (2014)Adapted from Shetreet and Friedmann (2014).

Furthermore, the direct comparison between Verb- and *wh*-movement was performed by Shetreet and Friedmann (2014) who reported four main significant clusters of activation for the contrast *wh*-movement versus Verb-movement in Hebrew topicalization and Verb second construction: Broca pars Triangularis, posterior Superior Temporal Gyrus, and bilateral mid-Superior Frontal Gyrus [7].

Although the authors do not discuss the results obtained by directly contrasting the two movement effects, we may say that at least some of the differences observed (e.g. mid-Superior Frontal activation cluster) are linked to the very different interpretative properties of Topicalization (Fig.6.6Ba) and Verb second sentence featuring only an adverbial in sentence-initial position (Fig.6.6Bb).

These studies indicate that these three types of syntactic movement are processed differently, and that further investigation of their neural underpinnings is a promising path in a configuration with minimal interpretative differences between the contrasted clausal-types.

6.1.1.3 Question formation in French and its different movement types

Question formation in French has the advantage of featuring several different syntactic configurations, crucially implying different types of syntactic movements.

Yet, the most peculiar aspect is linked to the possibility to obtain yes-no questions by simply moving the verb in front of subject pronouns in sentence-initial position (i.e. V to C position), a device that is traditionally called Subject-inversion. Moreover, compared

7. Note that the opposite contrast did not yield any significant increased activation cluster.

to other languages, French has the advantage of featuring the reunion of two kinds of movements to derive standard *wh*-questions, like in "Qui imites-tu?" *Who do you imitate?*)[8].

A particular type of y/n question as baseline

To this already optimal configuration, the fact that French also features a type of yes/no question that is obtained by simply adding a question mark at the end of the sentence with no word-order variations, will additionally allow to reveal the neural patterns related to the simple operation of transforming a declarative into an interrogative. A device we presented in chapter 2 through the work of Chen and Rooryck who defined it as an intonational morpheme playing a somehow similar role to Chinese sentence-final interrogative particles (as discussed in §2.4.4.2, p.202).

It is generally acknowledged that French has three types of clauses to express yes-no root questions (Coveney, 2011; Borillo, 1978): (i) y/n Questions introduced by the interrogative complementizer *est-ce que* "Est-ce que il reside là?" *Does he reside here?* (ii) inverted clause by Verb-movement like in "Reside-t-il là?" *Does he reside here?*, and (iii) a syntactic form that is identical to that of a declarative clause but with a final question mark, that can be distinguished from an asserted declarative clause by a rising intonation. However, we should say that when different forms are found with similar uses, a natural question arises about their different use.

In this regard, Mosegaard-Hansen (2001) argues that several interactional factors are favoring one or the other of the three forms, and mainly proposes two factors as playing a crucial role in the choice of the simplest y/n Question form, namely the accessibility of information, and the participation in a dialogical structured situation[9]. These simple y/n Questions with a declarative word-order tend in fact to be about events that are known to the addressee and not to the speaker, while the other two types of y/n Questions tend to be used when referring to other types of events. The particular usage of this simple type of y/n Questions is confirmed through the study of a radio talk show oral corpus. Abaillé and Colleagues (2013) put forward that this type of interrogative – questioning declaratives (Q-declaratives) – actually features a content that is more propositional than the one of a real question (confirming Marandin, 2005; Beyssade-Marandin, 2006). However, pragmatically these utterances appear to be real queries because they call for an explicit response, that strongly favor a confirmation of the query: -74% of Q-declaratives get a confirmation, against only 29% of the *est-ce que* and 26% of Verb-movement y/n Questions. Moreover, the advanced statistical analysis reported by this study suggests that they tend to have two main roles, namely conversation management and topic management[10]. Finally, as they are compatible with epistemic markers and

8. For a thorough review of the patterns of impairment of Verb-movement to C in agrammatic aphasia see Friedmann (2013).

9. Importantly, this study diversified its oral corpus, comprising everyday dialogues, radio talk shows and one school examination (including alternative questions and rhetorical questions). And the author bases her assertion about the centrality of interactional factors in the use of simple y/n Questions by observing that their prevalence crucially depends on the type of the interaction and the distribution in radio debates is more balanced that the other dialogues types.

10. Another interesting aspect emerging from this study is that simple y/n Questions with declarative word-order seem to be more frequent with the second person compared to the other question types, showing that in this oral radio shown they are suited to appeal to the addressee (45% of use a 2nd person subject against only 7% for *est-ce que* and 14% for Verb inverted y/n Questions.

question tags, and given these findings, we can say that French y/n Questions with a declarative word-order can be considered to be in-between interrogatives and declaratives. As interrogatives, they are compatible with several situations where the content of the question is either (i) present in the situation or the dialogue, (ii) directly inferable, (ii) resumed to during dialogue from previous interaction for verification, (iv) part of the knowledge of the locutor or (v) a conjecture from context information.

Predictions for French *wh-* and V-inverted Questions

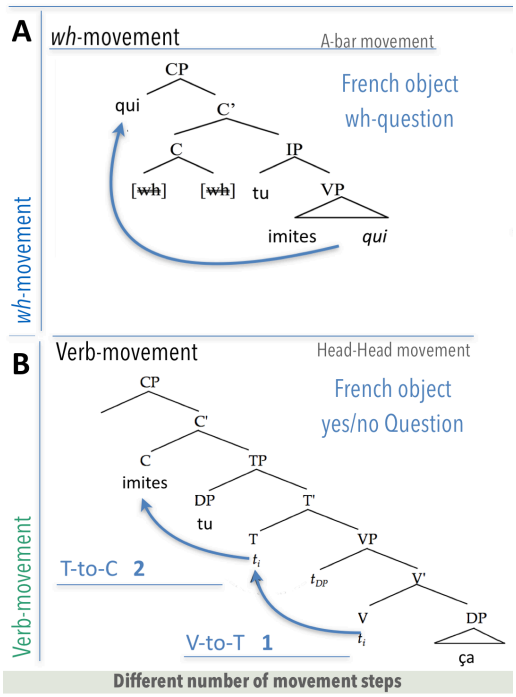


Figure 6.7 – Syntactic-tree diagrams for wh-questions and for V-inverted questions.

These characteristics of question formation in French represent an opportunity to test both for different types of movement maintaining their clausal typing (interrogative) interpretation constant, but also to investigate possible additive effects of these two types of movement inside the same sentence.

While we are aiming for *wh*-movement and Verb-movement to confirm the literature findings, we nonetheless expect to refine them, by distinguishing clause typing and syntactic movement complexity, as we have two different types of baseline: a declarative one and a simple yes/no question one, where interrogative clause-typing is obtained by an intonational morpheme. We will be comparing these two types of movement, which as we saw in chapter 2 are differing in several aspects (§2.4.2.1). The nature of the moved element is a phrasal constituent (wh-word) found in wh-movement, and a Head in the case of Verb-movement. While both movements end up in a position that is located in the CP-layer of the sentence, they do not proceed in the same way. Verb-movement has a step-wise movement progression pattern.

For questions featuring both *wh-* and Verb-movement (henceforth V-inverted wh-questions), our main claim is to be able to experimentally demonstrate that V-inverted wh-questions would reveal the neural patterns of derivationally more complex questions compared to those featuring a single movement (i.e. simple wh-Questions "Qui tu imites?" (*who you imitate?* and Verb-movement one like "Imites-tu ça?" *Imitate you this?*.

6.1.1.4 Clitic movement, a more local *move*

Considering another dimension of movement-related complexity, the targeted sentence domain and the distance covered by syntactic movement, French also offers, like other Romance languages, an interesting case of more local syntactic movement, displacing a particular type of weak pronominal element to preverbal position, the so-called clitic[11]. Contrary to *wh*-phrases, clitic pronouns cannot 'leave' the finite clause to which they belong: "Tu l'imites." *You [him* $_{clitic}$ *imitate.*

11. Clitic languages are languages having clitic pronouns in addition to strong and/or weak pronouns.

(256) Clitics' local movement

a. Clause boundedness

* Je *le* pense [que je verrai].
I him$_{clitic}$ think that [I] will.see

'I think I will see him.'

b. Clitic to pre-verbal positions

Je pense [que je *le* verrai].
I think that I him$_{clitic}$ will.see

'I think I will see him.'

Hence, compared to *wh*-movement, clitic-movement also originates in the Verb Phrase and moves the object upwards, but it is clause-bound, and actually involves a two-step progression: first the movement of a Noun Phrase and then the movement of a Head as discussed in chapter 2 (2.4.2 p. 162). If we compare clitic-movement to Verb-

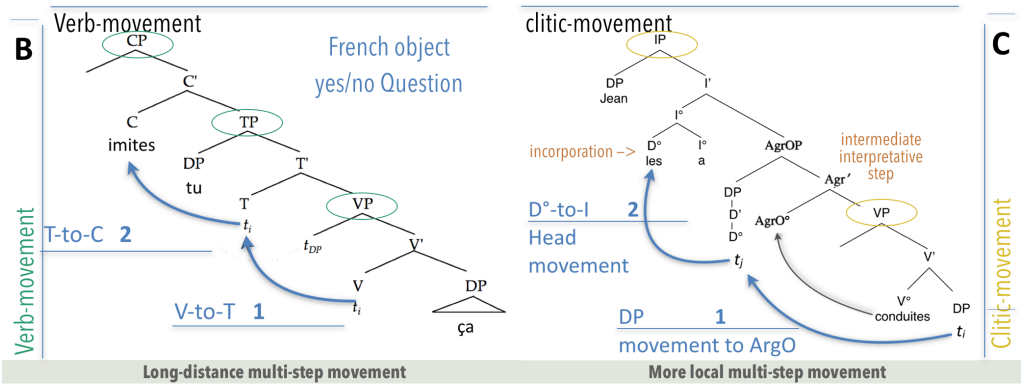


Figure 6.8 – Syntactic-tree diagrams of the multi step progression of (B) Verb-movement to C positions, (C) and clitic-movement to I.

movement, clitics also show a two-step progression, and are additionnally implying an hybrid movement: combining a first step of NP-movement[12]with a second step of Head-movement where the incorporation of the verb takes place (as illustrated in Figure 6.8).

Clitics evidence from acquisition

Evidence for the syntactic complexity of clitics in French and other languages featuring this kind of weak pronominals, comes from acquisition studies. A recent cross-linguistic acquisition study by Varlokosta and colleagues (2016) compared the acquisition of pronominals in languages that lack object clitics ('pronoun languages') with languages

12. Among the proponents of a derivational analysis for clitic placement, the hybrid nature of clitic-movement has been generally adopted to account for both the locality effects typical of A-movement and for the fact that clitic pronouns behave as Heads. In sum, clitic movement is in fact be seen as an instance of A-movement of a maximal projection followed by proper X°-movement (cf. Sportiche, 1989)

that employ clitics in the relevant context ('clitic languages'), by testing the acquisition of third-person pronominal objects in five-year-olds for sixteen languages.

One of the relevant results of this study, for our concern in syntactic complexity, is that French appears to be a language where children do omit accusative clitics (Jakubowicz et al., 1996; Hamann et al., 1996, Jakubowicz and Rigaut, 2000), but do not omit determiners with the exact same phonological form. Although clitics are phonologically similar across languages, yet they are omitted only in some of the clitic languages: the production rate is initially quite low in French and extremely varying rates of clitic production are reported (ranging from 50 to 90%) until the age of 6[13].

6.1.1.5 Unaccusatives

Given our interest for sentence structural complexity, we added to our experimental design an additional movement type that has largely been documented in previous research, namely NP-movement in unaccusatives. This movement type has the peculiarity to be not only a local one, but to be triggered by the need of these verbs having only an internal argument to fill the position of the external argument, the subject .

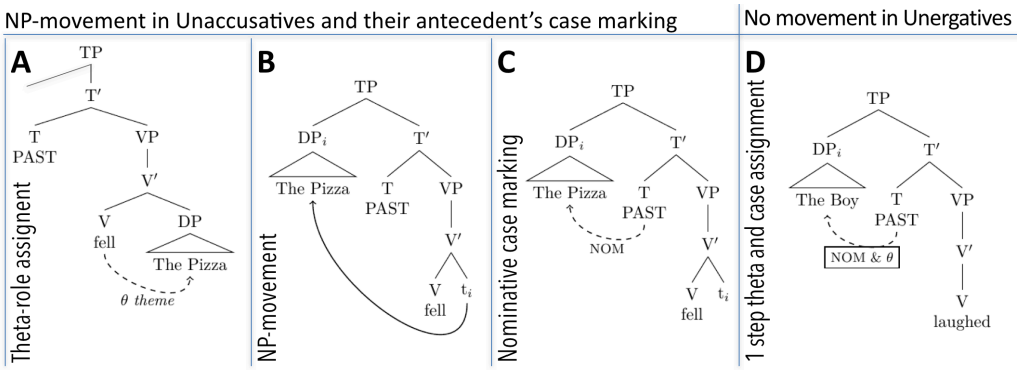


Figure 6.9 – Unaccusatives are intransitive verbs that have an internal argument, but no external one at Deep Structure. Their derivation happens in three steps: (A) First step, Theta-role is assigned. (B) Second step, NP-movement occurs to the empty subject position (spec, TP). (C) Third step, Nominative case assignment after having landed in subject position. (D) Unergatives are intransitive verbs that have an external argument at "deep" structure level and that assign in one step both case and theta-role. While on the surface the single argument of Unergative and Unaccusative verbs gets Nominative case (NOM) in the Specifier of TP (Spec,TP), Unaccusatives feature a Theme Argument generated as a complement of V (cf. UTAH), but cannot assign Accusative case as they lack an external argument (Bruzio's Generalization).

For instance, the class of intransitive verbs called Unaccusatives, presents only one theme-marked argument, which therefore Base-Generated in the verb object position. In this verb class, the theme-marked argument moves to the subject position to satisfy the

13. Thereby authors provided among others an important tool to assess early detection of language impairment : a child who at age 5 is not able to produce any or few pronominals is at risk for language impairment. In this way, pronominal production can be taken as a developmental marker (provided that one takes into account certain cross-linguistic differences discussed in the paper). We can note that omission of third-person clitics or lack of clitic production in obligatory contexts has been claimed to be a clinical marker of SLI in French (Paradis et al., 2003)

requirement that English and other languages have that all sentences must have subjects. Namely NP-movement preserves canonical word-order.

This characteristic aspect of NP-movement is what made psycho-linguists hypothesize and test that the comprehension system engaged into incremental processing of NP-moved sentences would establish NP-*traces* in their original location only after having assigned meaning (McElree and Bever, 1989, using a probe recognition task), rather than in a predictive manner as seen in wh-movement (cf. pre-*trace* priming effect in Nicol and Swinney, 1998). Given these characteristics, we can add that from the point of view of the incremental processing of the sentence, the main difference between the traces left by these two movement types resides in the presence of overt markers for wh-movement, while NP-movement has none.

Unaccusatives involve the movement of the theme to the subject position from the post-verbal position, (Burzio, 1986; Perlmutter, 1978), while VP-internal Subject hypothesis should be considered for Unergatives ; this hypothesis postulates that a Subject position is always present in whatever verb. Although fewer neuro-imaging studies have been examining NP-movement structures, those to be found in the literature concentrate on (1) passive sentence processing, and (2) unaccusatives.

Shetreet et al. (2010b) examined the neural correlates of processing sentences with unaccusative verbs, by comparing unaccusative verbs with verbs that do not undergo NP-movement operation but still have a single agent argument –- i.e. unergatives – and simple transitives with two arguments. Their study is among the first to show that the brain distinguishes between unaccusative and unergative verbs, even when they appear in identical structures (see the comparison in Cross-modal lexical Priming between Unaccusative and Unergative verbs in Friedmann et al., 2008)[14].

Unaccusatives were associated in this study with activation in the left inferior Frontal Gyrus and middle Temporal Gyrus (MTG). While the contrasts between [Transitive > Unergative] did not reveal any significant cluster of activation, Unergatives compared to Transitives activated the [Unergative > Transitive] right MTG.

In this regard, the conjunction analysis performed in this study was particularly revealing. By intersecting the Unaccusativity effect obtained by two different contrasts [unaccusative > unergative and unaccusative > transitive], the comparisons between unaccusatives and the other verb types revealed activations in the left inferior frontal gyrus (IFG BA 45/46) and the left posterior middle temporal gyrus (MTG).

Together with previous neuro-imaging results, this study suggests that the IFG may be involved with the execution of the syntactic operation, whereas the MTG may be responsible for the lexical/morphological operation that derives unaccusative verbs in Hebrew[15].

14. Posterior Middle Temporal Gyrus (pMTG) is also present in optional omitted Transitive verbs compared to full form optional transitive verbs, which seems to confirm the involvement of this area in the (probably lexical) operation of argument omission.

15. To clarify the interpretation that MTG supports morphological operation for unaccusative, it has to be noted that in Hebrew, unaccusatives differ from unergatives in two ways. First, both V-NP and NP-V orders are acceptable for unaccusative verbs, but only NP-V is acceptable for unergatives (Siloni, 2002, 2008). Secondly, there are morphological marking means to transform transitive verbs into unaccusative verbs: unaccusative *hitrasek* "crushed" is actually derived from the transitive *risek* ("smashed", as in Ron smashed the vase). There are also unaccusative verbs that appear in the active pattern (e.g., *nafal* = "fell" or *namas* = "melted"). Observing the experimental materials provided by the authors, 10 out of 14 unaccusative verbs of the experimental design are transitive verbs to which the morphological prefix for unaccusativity is added.

fMRI Studies on the distinction between unaccusatives and unergative

fMRI Study	Language	Task	Verb types	Increased activation clusters (MNI)
Shetreet et al. (2010a)	Hebrew Auditory	Plausibility judgment	Unaccusatives > Unergatives	left IFG [-52 32 11] BA 45/46/47 left post MTG [-63 -20 -16] BA 21 left SFG [-64 -41 -19] BA 37 left Fusif. right cerebellum [35 -76 -57]
Title: The Neural Correlates of Linguistic Distinctions: Unaccusative and Unergative			Conjunction Analysis: Unacc. > Unerg. & Unacc. > Trans.	left IFG [-52 32 14] BA 45 left MTG [-65 -46 12] BA 22
Take home message: Unaccusatives activate left IFG, which can be related to movement from object to subject position			Unaccusatives > Transitivess	left IFG [-49 34 21] BA 46 left IFG [-40 34 -10] BA 47 left IPL [-54 -59 41] BA 39 left Post. MTG [-67 -44 -23] BA 21 right MTG [47 -44 -13] left Fusif.
Shetreet et al. (2012)	Hebrew Auditory	Plausibility judgment	Unaccusatives > Unergatives	left IFG [-40 32 12] BA 45/46/47 left IFG [-26 34 -8] BA 47/11 left MTG [-46 -40 -8] BA 21 right Precuneus [12 -66 34] BA 31/7
Shetreet et al. (2010b)	Hebrew Auditory	Plausibility judgment	Transitives > Unergatives	Med. Precuneus [-4 -69 39] BA 31/7 Ant. Cingulate [3 24 44] BA 32/8 Post. Cingulate [-4 -56 18] BA 23 left Precentral MFG [-40 6 51] BA6
Thompson et al. (2007)	English	Lexical decision	Transitives > Unergatives	left Angular/ Supramarginal Gyrus [-43 -54 30] in BA39/40

Figure 6.10 – IFG = inferior frontal gyrus; pSTG = posterior superior temporal gyrus; AG = angular gyrus; SMG = supramarginal gyrus; pMTG = posterior middle temporal gyrus; aMTG = anterior middle temporal gyrus; STS = superior temporal sulcus; MFG = middle frontal gyrus.

All in all, Shetreet et al. found left IFG activation for unaccusatives, which was also found in English passive structures which are involving NP-movement. However, a different NP-movement structure in German, did not yield IFG activation for NP-movement (Wartenburger et al., 2003).

6.1.2 Interesting left-behinds in syntax

In parallel with the development of the linguistic formal framework postulating the existence of syntactic movement and thus describing complex sentences as the result of precise rule-governed sentential transformations, psycho-linguistic evidence for it has been gathered in the last three decades (§2.3.4). This initial empirical evidence has been confirmed by ever growing converging evidence in patient studies (§2.4.2.2) and in neuro-imaging (§2.4.2). As discussed in §2.4.1 (p.159), the idea that *Derivational complexity* has a psychological reality during sentence processing has emerged in the 70s and is still present in neuro-imaging of syntax in more recent attempts to correlate parsers derivational steps with brain activity during sentence comprehension (A. Bachrach PhD, 2008; Brennan et al., 2012; Nelson et al., 2017) .

As illustrated in Figure 6.11, by varying the number and the combination of different syntactic transformations, we built a rich and varied set of experimental conditions which will enable us to identifyi brain regions whose activation varies with the number of syntactic positions – i.e. the number of words + number of empty categories left at the

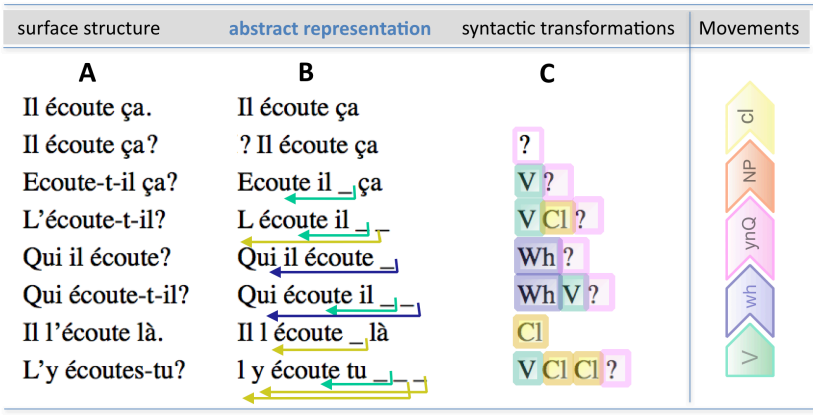


Figure 6.11 – Number of syntactic positions a proxi for syntactic-tree complexity: the number of words + number of empty categories left by different types of syntactic movement, (A) surface structure of our stimuli sentences. (B) Abstract representation of the derivation of the sentence structure. (C) The number and type of syntactic transformations generating the sentences. Searching for brain regions where activations varies with the number of positions in the 'abstract' syntactic representation, and/or with the presence of specific syntactic operations or characteristics.

extraction site by movement – in the 'deep' representation of the sentence, irrespective of the different syntactic movements that derived the single structure.

Thereby, correlating this more abstract and representational measure of the encoding for sentences' syntactic structure in a tree-like format could tentatively reveal the cerebral substrates of the representation of the sentence structures engendered by derivational syntactic operations.

While our first concern is to look for the neural underpinning of specific syntactic transformations/operations and their cerebral processing, the second is to further identify brain areas that are directly implied in the encoding of the hierarchical structures on which syntactic transformations theorized by derivational syntax are operated on[16].

Given the psychological and neuro-psychological evidence for movement related complexity reviewed here and in chapter 3, we can say that the linguistic constructs described above do have important implications for human sentence processing (as well as production). This is what grounded our choice to try and correlate with brain activity the number of syntactic positions. The main rational of this resides in the hypothesis that the set of brain areas underpinning the kind of representation linked to *gaps*, should show an increased activity proportional to the number of empty silent syntactic elements present in the syntactic-tree of our set of movement derived-sentences.

In more general terms, by looking for brain areas correlating with the number of syntactic positions, we sought to investigate *the shape of sentence's structural representation*, the possible tree-like format we discussed in chapter 3.

In sum, given the rich range of experimental results converging towards the psychological dimension of movement-related complexity, we will address the issue of the neural underpinnings of different movement transformations, and take into account as a complexity parameter the number of abstract syntactic object '*gaps* taken together

16. Possibly paralleling what has emerged in the study of vision, we aim at investigating the actual format of encoding of three hierarchies (Pallier, p.c.).

these two ways of looking at our experimental design may reveal important aspects of the organization of syntactic representations in the brain (cf. Figure 6.11).

6.2 Method and Results

6.2.1 Experimental Materials - Stimuli Description

We built an experimental design comprising 35 different sentence structures, identified as (c01) to (c35) condition in Table 6.2.

A total of 525 sentences were constructed, 15 sentences for each of the basic conditions (see the Annexes C.2 (p.866 for the full list of stimuli and the 30 single words added as controls). In order to manipulate the number of syntactic positions, we selected verbs that varied in number of arguments. We therefore used different classes of verbs: unergative, unaccusative, two arguments unaccusative with a locative argument, transitive and distransitive ones. We carefully chose verbs having a fixed number of arguments and a single sub-categorization frame, as Shetreet et al. (2007, 2010), and Meltzer-Asscher et al. (2013) reported effects of numbers of optional complements and syntactic frames.

To classify verbs as unergative and unaccusative (conditions c01 to c06), we used three semantic and syntactic tests. The first test, used to identify unaccusativity, examined the ungrammaticality of nominalization of the verb: unaccusative verbs cannot be nominalized, contrary to unergative ones; i.e., *mourir* ('to die') → **moureur* ('a dier') vs. *ronfler* ('to snore')→ *ronfleur* ('a snorer'). A second test used the grammatically of passivation: unnaccusatives cannot be passivized, whereas unergatives can, i.e., *"*Il a été décédé dans ce lit."* "*It has been died in this bed" versus *"Il a été dormi dans ce lit"* "It has been slept in this bed". The third test used the fact that impersonal constructions and participative forms at sentence beginnings are legal for unaccusative but not for unergative (all the syntactic tests performed to select the verbs for the experimental material are presented in the Annexes §C.2 Table:C.23, p.886).

We created *wh*-object questions by adding to declarative word-order an object *wh*-pronoun in sentence initial position. Interestingly, object interrogative pronoun does not differ from subject interrogative pronoun in French, preventing participants to guess since the beginning of the sentence the grammatical case to attribute to this pronoun.

Verb-movement conditions were constructed thanks to a peculiarity of French interrogative constructions, the so-called inverted questions (broad focus) built by moving the verb in pre-subject position and by adding a dash line between verb and subject pronoun. A third type of movement is found in French and other Romance languages, where the object moves in pre-verbal position.

A particularity of clitics is that they carry topical information and therefore it is presupposed that the hearer already knows about the object referent that is being expressed by the clitic in the sentence.

Finally, a simple yes/no question condition was added to have an interrogative baseline which did not involve the types of movement tested in the other interrogative conditions. For this condition, yes/no question were built in each verb class by simply keeping the same word-order as in the equivalent declarative conditions and adding a question mark at the end of the sentence, which marks in the oral a rising pitch on the final syllables.

Such an inventory of movement types allowed us to situate at each hierarchy-level of the sentence a different movement in order to vary their distance and range across the

nb cond.	verb class	syntactic movement type		Stimuli examples	Abstract Representation	nb Arg.	nb syntactic positions
c01	Unacc	Decl	NP	Il bascule.	he swings _.	1 Argument	3
c02		ynQ	NP	Il bascule?	he swings _?		3
c03		V	NP	Bascule-t-il?	swings he _ _?		4
c04	Uerg	Decl		Il éternue.	he sneezes .		2
c05		ynQ		Il éternue?	he sneezes ?		2
c06		V		Éternue-t-il?	sneezes he _?		3
c07	Unaccusatives	Decl	NP	Il réside là.	he resides _here .	2 Arguments	4
c08		ynQ	NP	Il réside là?	he resides _ here ?		4
c09		V + wh	NP	Où réside-t-il?	where resides _ _ _he?		6
c10		V	NP	Réside-t-il là ?	resides he _ _here ?		5
c11		cl Decl	NP	Il y réside.	he here resides _ _.		5
c12		cl ynQ	NP	Il y réside?	he here resides _ _?		5
c13		cl + V	NP	Y réside-t-il ?	here resides he _ _ _?		6
c14		wh	NP	Où il réside?	where he resides _ _?		5
c15	Transitives	Decl		Il imite ça.	he imitates this.		3
c16		ynQ		Il imite ça?	he imitates this ?		3
c17		V		Qui imite-t-il?	imitates he _ this?		4
c18		V + wh		Imite-t-il ça ?	who imitates he _ _?		5
c19		cl Decl		Il l'imite.	he this imitates _.		4
c20		cl ynQ		Il l'imite?	he this imitates _?		4
c21		cl + V		L'imite-t-il ?	this imitates he _ _ ?		5
c22		wh		Qui il imite?	who he imitates _?		4
c23	Transitives + Locative	Decl		Il envoie ça là.	he sends this there.	3 Arguments	4
c24		ynQ		Il envoie ça là?	he sends this there ?		4
c25		V		Envoie-t-il ça là?	sends he _ this there?		5
c26		V + wh		0ù envoie-t-il ça?	where sends he this _ _?		6
c27		cl Decl		Il l'envoie là.	he this sends _ there.		5
c28		cl ynQ		Il l'envoie?	he this sends _ there ?		5
c29		cl + V		L'envoie-t-il ?	this sends he _ _ there?		6
c30		wh		Où il envoie ça?	where he sends _ this ?		5
c31		cl + wh		Qui il lui envoie?	who he him sends _ _ ?		6
c32		cl + V + wh		Qui lui envoie-t-il?	who him sends he _ _ _?		7
c33		cl + cl Decl		Il l'y envoie.	he this there sends _ _.		6
c34		cl + cl ynQ		Il l'y envoie?	he this there sends _ _?		6
c35		cl + cl + V		L'y envoie-il?	this there sends he _ _ _?		7

Table 6.2 – The 35 experimental sentence conditions. The code of the condition is followed by the types of syntactic transformations and the stimulus examples in French and English literal glosses. The number of syntactic positions is reported in blue.

syntactic-tree. Therefore, experimental sentences contained a varying number of empty syntactc positions from which a given element had been displaced.

6.2.1.1 Statistics on the stimuli

Distributions across conditions of number of words, syllables and letters

As for the length of sentences, we tried to match as much as possible the verbs in terms of lexical frequencies and of lengths considering both the number of characters and syllables across the different verb classes in the experimental conditions. In one-way anovas, the effect of verb class is not significant for length ($F(4, 85) = 0.66$; $p = .62$) and marginal for lexical frequencies ($F(4, 85) = 2.2$; $p = .07$).

The sentences contained 2 to 4 words. Their average character length (comprising spaces and punctuation marks) was 15.2 (min = 9; 1st.quartile = 13; median=15 ; 3rd.quartile = 18 ; max = 25). This correlated significantly with the number of positions (Pearson r = 0.59; $t(523) = 16.7$; $p<.001$).

In addition, we included 30 single words to serve as a potential control condition. We selected them from the Lexique.org database. There were long words (average length of 10.9 char (min = 6; 1st.quartile = 9.25, median = 11; 3rd.quartile = 12 ; max = 15), yet with reasonable lexical frequencies (above 0.1/million according to the Lexique's 'book frequency' variable, for the experimental corpus see the Annexes C.2 (p.866).

6.2.1.2 Argumental structure of the verbs in the stimuli

As we sought to isolate the neural correlates of different types of syntactic movements by varying the transformation factor in an unconfounded manner, we controlled for as many complexity generators as possible.

In our critical comparison, during fMRI recording subjects attended to sentences that required four different movement types (NP/VP/Clitic/ Wh-movement) of transformational analysis and sentences that did not.

In an attempt to set the effect of movement apart from that of other complexity factors, we further manipulated the complexity of verbs like their number of arguments and controlled for the subcategorization options of the selected verbs.

Syntactic tests on Argumental structure: Unaccusativity versus Unergativity

Different tests for the unaccusative/unergative distinction reveal the fundamental differences between these two verb classes. However, different tests tend to apply in different languages. For instance, *Auxiliary selection* is a reliable test for Italian and less in French. There is also a debate about exactly what these tests diagnose - a syntactic difference between the two classes or a semantic difference between two sets of predicates (for a detailed discussion cf. §2.4.2.1 in chapter 2, p.170)

Unergative Tests Unergatives, like *run*, *talk*, *shout*, *lie*, *smile*, have the following characteristics: (i) semantically have a true subject, therefore the subject is animate, (ii) can be nominalized, and go along well with passive verbal morphology, (iii) but they cannot enter impersonal constructions.

Unaccusative Tests In contrast, Unaccusative verbs like *die*, *fall*, *sit*: (i) semantically have just an object (that moves in pre-verbal subject position), (ii) their grammatical-subject can be animate or inanimate and (iii) can enter impersonal constructions. They are typically involuntary verbs, existential verb, movement verbs (e.g. *fall*).

Note that, for all the different verb classes we generally avoided polysemous verbs and verbs that have cognate object (e.g. "courir une course" *run a run*).

Number of Arguments

To control for argumental structure complexity factors in order to discard or control for possible confounds resulting from the variety of argumental structures present in our design, we relayed on the literature on processing of verb argumental structure to avoid some confirmed complexity generators, like the variability in the sub-categorization options of a verb (Shetreet et al. (2007, exp.2 ; Shetreet et al. 2010, e.g. compare discover and loose in 'discover the truth' / 'discover that he is here' versus 'loose the keys').

And based on the complexity effect revealed contrasting *alterning transitives*[17] with simple intransitive revealed by Meltzer-Asscher et al. (2013)[18], we avoided verbal lexical entries corresponding to two different verbal alternates (with multiple thematic grids). To test for these lexical aspects and select the verbs to build experimental materials we used a French Dictionary of verbal valence called *dicovalence* (Van den Eynde and Mertens, 2006)[19], which provided us with the thematic grid of 3700 French verbs [20].

Crucially the study by Shetreet et al. (2010) indicates that verbs with optional complements show similar activation patterns to verbs that have only one frame. Hence we built on this result to be less strict in our search for verb that had no optional complements, as the optionality of an essential complement might not be a relevant complexity factor at the level of brain activation[21].

Hence, the syntactic and semantic tests we used to select unaccusative and unergative verbs in French can be summarized the following (cf. Annexes §C.23, columns *Semantic Test* to *Syntactic Tests*, p.886):

17. The so-called "Alternating transitive" verbs differ from simple intransitives with regard to both the number of thematic grids (two vs. one) and the number of thematic roles (two vs. one)

18. As a side note on the materials of this study, alternating verbs all belonged either to the class of non-agentive verbs of manner or motion, or to the verbal class of externally caused change of state.

However, since verb-noun ambiguity is very widespread in English, many of the verbs in both groups could actually have an alternative nominal reading. The authors underline that the relative frequency of this nominal reading compared to the frequency of the word as a verb was generally small, and did not differ between the two verb groups. A contrasts that still remains to be done would actually compare alternating transitive with real transitives, to make clear if the involvement of posterior activations (i.e. Angular and Supramarginal gyri - BA39 and 40) can be separate from the Middle Superior Frontal (BA 8/9) and Middle Temporal one, as uniquely responding to the complexity dimension of number of arguments as indicated by Thompson et al. (2007 and much related work).

19. Found at the following URL: `http://bach.arts.kuleuven.be/dicovalence/`, and occasionally consulted the following one :`http://sites.univ-provence.fr/delic/lexiques_syntx.html`

20. For the reader willing to have a look at the tests performed on each verbal lexical entry used in the stimuli materials, the column *Argumental grid* of the Table in §C.23 in p.886 is containing 'strange code', it is the thematic grid from the verbal valency dictionary *Dicovalence*. It contains information on the verb's argumental grid, the presence of a question mark before a category signals the optionality of a certain argument or characteristic; the vertical bars signal alternatives and in squared brackets the argument's semantic characteristics are listed (i.e. animacy, etc.).

21. In other words, verbs with optional complements differ from two-frame words in a similar way to one-frame verbs differing from two-frame verbs.

1. the absence of obligatory or frequent Prepositional Phrase (PPs) or Complementizer Phrases (CP) for unergative and Transitive verbs (see column *CP subcategorization option*) (Shetreet et al., 2007 or Thompson et al., 2007 for obligatoriness versus optionality of arguments), the absence of frequent linguistic expressions displaying a different argumental structure from the critical one,
2. the absence of multiple sub-categorization options for the four verbs categories as underlined in Shetreet et al. (2007) who detected activation in left BA9 in response to processing verbs with multiple syntactic realization options (see also Shetreet et al., (2010) (see column *Argumental grid*)
3. for tree arguments verbs : the presence of frequently obligatory dative or locative Prepositonal Phrase (see column *Possible PPs*).

Predictions for number of arguments

fMRI Studies of Verb Argument Structure Processing:
number of arguments, number of complements and subcategorization options

Study	Language	Task	Tested argumeantal complexity dimensions	Increased activation clusters (Tal and MNI)	Remarks
Ben-Shachar et al. (2003)	Hebrew Auditory	Grammaticality judgment	Number of arguments; more > less : 3arg>2arg LpSTS	Left pSTS MNI [-37 -47 20] Tal [-38 -47 23]	In embedded sentences with the confound of the different number of frames !
Newman et al. (2003)	English Visual	Grammaticality judgment	Thematic-role violation	Left IFG (BA45) and pSTS	Un-ecological manipulation implying a violation
Shetreet et al. (2007) Exp 1	Hebrew Auditory	Plausibility judgment	More or less thematic options and subcategorization options	Left STG and BA47 ; IFG, BA9	
Shetreet et al. (2007) Exp 2	Hebrew Auditory	Plausibility judgment	Number of verbal complements	Right Precuneus [16 -64 33] Right antCingulate (BA32) [10 33 30]	Grades effect : no complement < one complement < two complements
Thompson et al. (2007)	English Visual	Lexical decision	Transitive intransitive (transitive + ditransitive) intransitive	Left SMG, AG Bilateral SMG, AG Left Angular and Supra-marginal Gyrus [-43 -54 30] BA 39/40	Lexical decision out of sentences : activation maps : 2Arg> 1Arg verbs (only left) and 2Arg + 3Arg > 1Arg verbs (more bilateral)
Thompson et al. (2009/10)	English Visual	Lexical decision	Transitive intransitive (transitive + ditransitive) intransitive	Bilateral AG, Left AG	Thompson 2010: Parametric, random effects, analysis examining the effects of argument structure with values of one-, two- and three-arguments revealed a significant cluster of 39 voxels ($p < .001$ uncorrected) in the left Angular Gyrus [-57, -66, 18] in BA39.
den Ouden et al. (2009)	English video	Naming action pictures/videos	Transitive > intransitive	Left IFG (BA 44/45), SMG, AG Angular and Supramarginal Gyrus [-30 -54 57] MNI	
Shetreet et al. (2010)	Hebrew Auditory	plausibility judgment	Subcategorization options	Left STG, MFG, SFG	
Meltzer-Aascher et al. (2014)	English Visual	Lexical decision	Alternating transitive > unergative	Bilateral SMG, AG Angular Gyrus, Bilateral MFG, SFG	

Figure 6.12 – IFG = inferior frontal gyrus; pSTG = posterior superior temporal gyrus; AG = angular gyrus; SMG = supramarginal gyrus; pMTG = posterior middle temporal gyrus; aMTG = anterior middle temporal gyrus; STS = superior temporal sulcus; MFG = middle frontal gyrus.

On the basis of the findings of previous studies that we considered in order to build our experimental sentences – summarized in Table 6.12 – we can develop initial predictions concerning the areas that will presumably be involved in the comparison of sentences with different argument structure:

1. Contrast between transitive declaratives (2-argument) and intransitive ones (1-argument): we should expect to observe left IPL (inferior Parietal Lobule) or Angular and Supra-

marginal Gyri as reported by Thompson et al. (2007) for verbs out of sentences in lexical decision task.
2. Contrasting 2- and 3-argument declaratives to 1-argument ones : we should observe bilateral Angulars and Supramarginal Gyri IPL (left > right) (see Thompson, 2007, confirmed by Den Ouden et al. 2012).
3. Gradual increased activation by number of arguments should be observed in bilateral Superior Temmporal Sulcus (STS), following Ben-Shachar et al. (2003) who observed this effect in embedded sentences, and/or bilateral Precuneus as a function of verbal argument complexity as reported in Shetreet et al. (2007) for simple sentences.
4. Having controlled the materials for multiple complementation frames we shouldn't have left STG activations responsible for complementation frames of the verb in the sentence (cf. Shetreet et al., 2010).

More detailed predictions could also be made as for a possible graded effect in right Precuneus and anterior Cingulate cortex, when comparing unergative, transitive and distransitive, as revealed by Shetreet et al. (2007), who kept thematic options and subcategorization options constant in declarative sentences.

6.2.2 Protocol

6.2.2.1 Procedure and Task

Each participant saw the complete set of stimuli, in a different random order. In a rapid-event related design, each trial presented the whole sentence at the center of the screen, for 200ms. The characters were presented in white on a black background, using the Inconsolata fixed-sized font at size 28. The inter-trial interval was 4.5 seconds (+/- 1 second of jitter).

Although briefly flashed, the sentences were fully intelligible and each participant's capacity to read them was tested during a familiarization task before scanning, by asking to read the sentence out loud.

To encourage participants to pay attention to the visual display, an explicit instruction to press a response button ("Appuyez!" *press*, or "Validez!", *validate* or "Cliquez!", *click*) appeared 6 times per session. The participant had been instructed to press a response button with the right thumb when the instructions occurred. The stimulation was controlled by a custom Python script written using expyriment.org.

After the acquisition of an anatomical scan, which lasted for 8 minutes, the stimuli were presented in 5 functional MRI sessions of 7 minutes, comprising 117 trials each. Subjects were instructed to covertly read the sentences and press the button when requested. An additional fMRI session ended the experiment. It consisted in a '*language comprehension localizer*' where complete, 9-10 words long, sentences and lists of consonant strings were displayed using Rapid Serial Visual Presentation (each word was presented for 200ms). For the localizer, a block design was used with 12 blocks where 3 sentences were displayed (inter-sentence blank = 600ms) and 12 blocks of matched consonant-strings. Each block lasted 10 seconds and the inter-block duration was 8 seconds.

6.2.2.2 Participants

Twenty-two native French speakers (10 women), all right-handed (average Edinburg score=0.85, ranging from 0.6 to 1, Oldfield, 1971; average age 32, range 18 to 42.) partic-

ipated in the experiment which was approved by the regional ethic committee (Comité de Protection des Personnes Ile-de-France VII, Protocole de Recherche Biomedicale # 2008-A00241-54/1).

Two participants were excluded because of movements of too large amplitude, we set thresholds within session movement amplitude of 1.5 mm in translation, in any direction, and 1.5 degree of rotation, along any axis.

Acquisition parameters

The acquisition was performed with a 3 Tesla Siemens Tim Trio system equipped with a 32 channels coil. For each participant, an anatomical image was taken, using a 3D Gradient-echo sequence and voxel size of 1x1x1.1 mm; Then, functional EPI scans were acquired using a multiband sequence developed by the Center for Magnetic Resonance Research (Feinberg et al. 2010) and sensitized to the BOLD effect (80 axial slices; MB=4; TR=1.5secs; TE=32msecs; Matrix=128×128; voxel size=1.5×1.5×1.5mm).

6.2.3 Analyses

Given the exceptionally rich set of experimental conditions present in this fMRI study, several analyses were run to leverage its potential to answer the different questions it was built on. The results we will present here literally represent only the tip of the iceberg of the enormous work it constituted to harvest all their potential. This experiment is a gold mine.

Variables of interest for the analyses

In the analyses, we considered the following variables of interest:

- Wh-mov (y/n)
- NP-mov (y/n)
- V-mov (y/n)
- Cl-mov (0,1,2)
- Question (y/n)
- number of positions (2-7)
- number of overt words (2-4). When computing number of words, we ignore euphonic-t and hyphens, that is, we considered that "Gigote-t-il?" contained two words. number of empty categories (0-3). Note that number of positions is number of words plus number of empty categories

The retained confounded variables where mainly two: (i) the size of text, that is number of characters including spaces and punctuation and (ii) the number of syllables, and the relationships between these variables are described in the above section on stimuli statistics (§6.2.1.1).

6.2.3.1 Whole-brain analyses

The functional images were processed with SPM8 (Wellcome Department of Cognitive Neurology, software http://www.fil.ion.ucl. ac.uk/spm), controlled by pypreprocess (https://github.com/neurospin/pypreprocess).

The functional images were realigned to correct for motion. Two participants were rejected at this stage for movements larger than 1.5 mm or 1.5 degree.

The T1 anatomical scan was spatially normalized onto the MNI template using SPM8 default parameters, and the resulting transformation was applied to the functional images. Finally, the images were resampled with 3mm isotropic voxels and smoothed and spatially smoothed with a Gaussian kernel of full width at half maximum (FWHM) = 5 mm.

The individual statistical models comprised a regressor for each of the 35 types of sentences (corresponding to the conditions described in Table 6.2, p. 527), as well as additional regressors for the nouns and the targets trials. These regressors were created by convolving trains of events corresponding to the individual trials from each condition by the impulse hemodynamic response function (iHRF) of SPM.

The localizer data was analyzed as a block design, with blocks defined as epochs of 10s convolved with the iHRF, split into two conditions: Sentences or Consonant-strings.

For the group level analysis, the individual effect maps of the 37 conditions -35 sentences + nouns + targets- were smoothed with an isotropic kernel of FWHM = 8mm and entered into a within-subject Analysis of Variance model, with one regressor per condition and one regressor per participant.

This model was then estimated using a Restricted Maximum Likelihood approach that does not assume the independence of conditions, and contrasts were computed to test for the various hypotheses presented in the Results section. Unless otherwise mentioned, the brain-maps are displayed at a statistical threshold of $p<.001$ voxel-wise uncorrected for multiple comparisons and with a threshold on cluster extent set to *50 voxels*, corresponding to a p-value of .05 given the estimated smoothness.

Both whole-brain and Region of Interest analyses (i.e. brain-maps and barplots) presented in this chapter are all the result of contrasts where the compared conditions were balanced in terms of number of words and syllables.

While for the Unsupervised analyses we run a individual statistical models that comprised in addition to the 35 condition, controls (nouns) and the targets trials, an additional nuisance variables to the usual the estimated movement parameters: a variable modeling the effect of number of characters. This last variable was obtained by creating a second SPM model with a single condition comprising the onsets of all events, and an associated parametric variable specifying number of characters. It was the regressor associated to the latter variable that was reused as a nuisance variable in our main design matrix. Note that this regressor, orthogonalized against the main hrf response, captures the effects of number of characters.

The rational behind this choice was to correct for the effect of number of words and characters in the unsupervised analysis, clustering the patterns of activation inside ROIs (§6.2.3.3).

6.2.3.2 Regions of interest analyses

Previous studies manipulating different aspects of syntactic complexity provided us with several regions of interest (ROI). As shown in Table 6.13, we considered regions that were stably reported to be sensitive to sentence structure building (Pallier et al., 2011; Brennan et al., 2012), syntactic movement (Santi and Grodzinsky, 2010; Ben-Shachar et al., 2004; Shetreet et al., 2010b, 2013), word-order variations (Brokessel-Shelsweky,

2009), and number of verbal arguments (Thompson et al. 2007). Broca Complex main subdivision between (Triangularis and Opercularis) was anatomically defined.

ROI Label	Left Hemisphere Regions	MNI center coordinates	Contrasts or Effect	Study
vPrC	ventral Pre-Central Gyrus	-41 10 31	Adaptation to syntactic movement + embedding	Santi and Grodzinsky 2010 ; Ben-Shachar et al. 2004
pSTS	posterior Superior Temporal Sulcus	-50 -40 3	constituent-size effect	Pallier et al. 2011
dPrC	dorsal Pre-Central Gyrus (BA6)	-48 4 51	Topicalization (wh-movement) > SVO canonical Word Order	Shetreet et al. 2014
aSTS	anterior Superior Temporal Sulcus	-54 -12 -12	constituent-size effect	Pallier et al. 2011
aIns	anterior Insula	-37 24 -6	Merge operation	IFOP : Friederici et al. 2006
TPJ	Temporo-Parietal Junction	-40 -57 19	constituent-size effect	Pallier et al. 2011
TP	Temporal Pole	-47 14 -27	constituent-size effect	Pallier et al. 2011
SMA	Sensory motor Area / pre-SMA	+/-7 23 41	Passive vs. Active sentences	Bornkessel-Schlesewsky et al. 2009 ; Friederici et al. 2006
MTG	middle Temporal Gyrus	-67 -48 -12	Conjunction : [Unacc> and Unerg Unacc> Trans]	Shetreet et al. 2010
IPS	Intra Parietal Sulcus	-42 -52 30	Word-order variations Object-Subject > Subj.-Object	Bornkessel-Schlesewsky et al. , 2009
IFGorb	IFG, Pars orbitalis	-45 33 -6	constituent-size effect	Pallier et al. 2011
Cerebellum	right Cerebellum Lobule VI	36 -62-28	language cluster in meta analysis ALE peak	Stoodley and Schmahmann 2009
BA45 - IFGtri	IFG, Pars triangularis	-52 20 17	Anatomical cyto-architectonical	Amunts et al. 1999 / 2004
BA44 - IFGoper	IFG, Pars opercularis	-51 8 15	Anatomical cyto-architectonical	Amunts et al. 1999 / 2004
AngG	Angular Gyrus	-42 -52 30	Number of arguments effect	Thompson et al. 2007

Figure 6.13 – Apriori Regions of Interest. Light blue for Anatomically defined sub-parts of Broca Complex (Triangularis and Opercularis), Deep blue for Areas responding to structure building Constituent-size manipulations; Deep green for areas reported in syntactic movement manipulations; light green for Areas reported in Dependency-link manipulations; yellow for areas reported as having an linearly increased activation in Word-order manipulations; and finally *water green* ROIs are aposteriori Sub-Cortical areas jointly activated with SMA selected in papers investigating word-order manipulations.

To extract individual data from a given ROI, we first created subject-specific masks in the following way: we intersected the ROI with the statistical T-map obtained from the localizer contrast [Sentences > Consonant-strings], to which we applied a threshold selecting the 10 % of voxels that showed the most robust responses inside the given ROI (*best voxels* method). These individual masks were then used to extract data from individual effect size maps. T-tests were performed in these data for each ROI. Given the number of our ROIs, we applied Bonferroni correction ans retained as statistically

significant the differences whose p-value was inferior to 0.003.

6.2.3.3 Unsupervised Clustering analyses

Encompassing a wide range of syntactic constructions, this study is particularly suited for a more data-driven approach to sentence processing. This is why we decided to perform an unsupervised analysis to possibly uncover (1) some aspects of the modulation of the Sentence Network yielded by our experimental manipulation of sentence's syntactic complexity, and (2) to attempt to functionally characterize brain regions based on their patterns of activation instead of just their response amplitudes as it is the case in the classical linear model (GLM analysis).

Clustering of Brain Regions by the response profiles of conditions

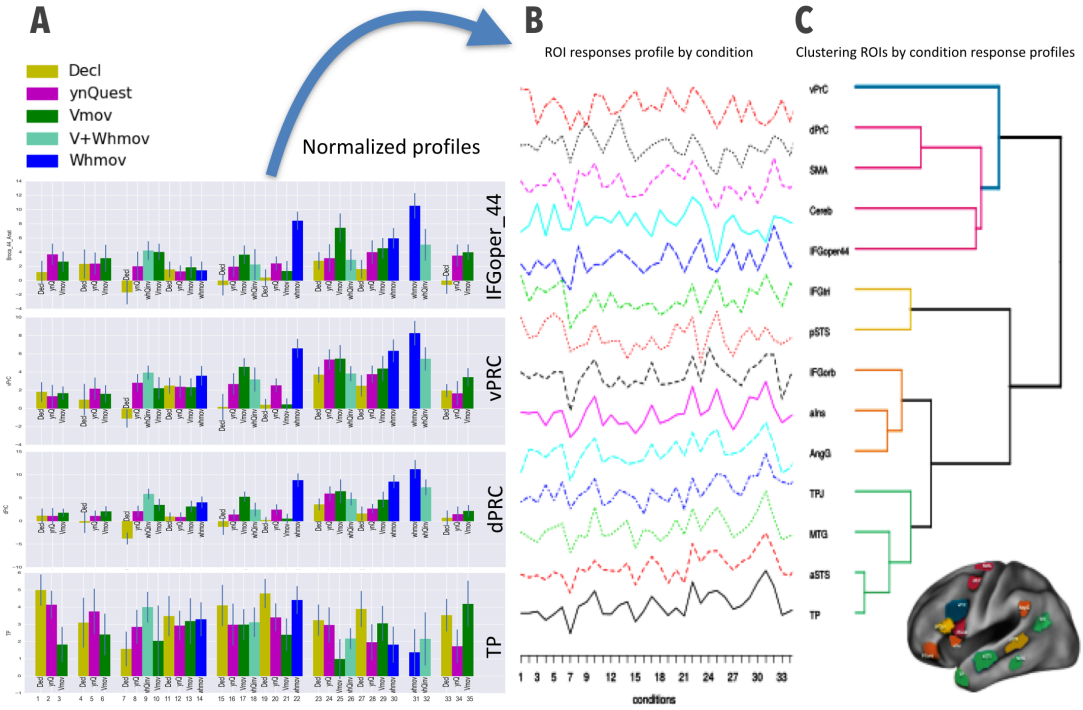


Figure 6.14 – (A) Direct response profiles, (B) Normalized response profiles, (C) Correlation Euclidean distance Conditions profiles by ROIs. Response profiles of the 35 conditions in the 14 Regions of Interest. Color codes for the manipulated variable of movement types: blue = wh-movement ; indigo = yes/no Questions ; yellow = Declaratives ; green = Verb-movement.

Responses to the 35 conditions were estimated using a linear model with a nuisance regressor for number of characters as detailed above. The response profiles across the 35 different sentence structures (our conditions) were then extracted in a set of 14 apriori regions of interest for sentence complexity processing (see Figure 6.14A). The focus being, here, no more on the response amplitudes, but on the brain regions' response

profiles, the direct response patterns in each ROI were normalized (see Figure 6.14B). An unsupervised clustering algorithm, based on Euclidian distance, was then ran on the matrix of correlations between these response profiles.

From these response patterns we were able to obtain the dendrogram on the right part (C) of Figure 6.14, representing a clustering based on the similarity of response profiles of brain regions across conditions. In other words, the closer the ROIs are, the more similar their response patterns to the experimental conditions are (see the similarity of the response profiles in the left part of the figure).

Clustering of conditions by profiles across ROIs

In a second time, a clustering analysis of the pattern of responses for each condition inside each ROI was run to investigate if a given brain area would reveal a similar pattern of response for sentences featuring the same syntactic movement or manipulated syntactic property.

6.2.4 fMRI Results

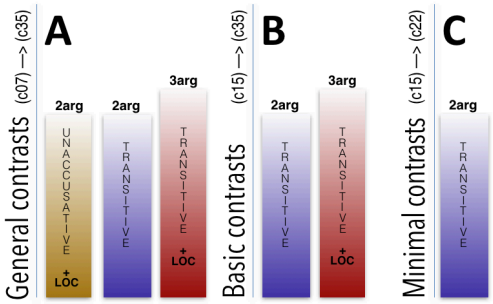


Figure 6.15 – Different types of more or less minimal contrasts. Minimal contrasts include strictly single conditions against another single condition or against two baselines.

This section is organized according to the different experimental questions that prompted the design of this experiment.

Given the rich set of experimental conditions, there are multiple ways to look at specific movement effects, either by combining multiple minimal pairs across different verb classes by restraining contrasts to minimal pairs with the declarative and interrogative baseline. Hence, the several contrasts we performed to answer to the different experimental questions, can be classified into three types according to the number of conditions they encompass:

1. *Minimal contrasts* which reveal more fine-grained distinctions as they oppose minimally different sentences, within the simple transitive verb class (2-argument verbs, the blue category), comprising conditions from (c15) to (c22);
2. *Basic contrasts* within the Transitive + Locative verb class (3-argument verbs, the burgundy category, comprising conditions from (c15) to (c35);
3. *General contrasts* which oppose conditions found in all the different verb classes, encompassing the one argument verbs and 3-argument ones.

4. *Single contrasts*, which strictly compare one condition against another and will be mainly used for ROI analyses.

Minimal, Basic and Single contrasts will be systematically reported when addressing the issue of different types of movements, while General contrasts will be used to answer more broad questions like that of the neural underpinnings of interrogative modality, because they are best suited to reveal effects that are shared by a large number of conditions irrespective of their syntactic derivation.

Note that glass brain figures are never corrected for cluster-size, while inflated brain figures are. ROI results are presented under brain maps as contrast effect size (i.e. signal differences) with a 95% confidence interval. Significance is marked by stars, and a single star (i.e. *) indicates p<.05 level). Lower significance is marked with a point, which is reached as soon as an error bar touches the zero line.

As a short reminder, unless otherwise mentioned, cluster-size correction is set by default at 50 voxels, and the Figures' threshold is set at $p<.001$ (t-value > 3.1).

6.2.4.1 Effects associated to the different types of movement

To look at the specific effects of each movement, we selected the sentences featuring a single type of movement (either *wh*-, Verb-, NP- or Clitic movement) and contrasted them to matched sentences, which were minimally different and were identical in all but the relevant movement. For each comparison we will first present the whole-brain contrasts in General and Minimal contrasts, subtracting our two baselines (i.e Declarative and Interrogative) and then ROI analyses results. Note that as NP-movement and clitic-movement do not imply any interrogative clause-typing, to be able to compare them with *wh*- and Verb-movement, we will observe them in simple yes/no questions.

wh-movement in French *wh*-questions

In a General contrast, *wh*-questions ('Qui elle cite ?', *who she cites?*) were compared with identical declarative SVO sentences with subject pronoun – a verb and an object 'Elle cite ça.' (*She quotes this.*) – and with both declarative and interrogative baseline. As illustrated in Figure 6.16, the Minimal contrast revealed significant activations in the left dorsal Precentral Gyrus (dPrC-BA6), the Supplementary Motor Area (SMA), the left IFG pars opercularis and triangularis, and the inferior Parietal Lobule bilaterally, as shown in Fig. B, while the general contrast revealed a more restrained activation in left dorsal Precentral Gyrus (dPrC-BA6), left Broca Opercularis (BA44) and the SMA at the same threshold (see Fig. A $p> .001$, 50 voxels cluster-size correction).

To factor out the effect of the interrogative Force, differentiating question from declarative, we also compared simple y/n questions [SVO word-order+?] to matched *wh*-questions, including *wh*-movement in Minimal contrasts, as shown in Figure 6.31 (p.552).

wh-movement - ROIs Analyses ROI analysis for the General contrast between *wh*-questions and declaratives confirmed whole-brain results found in the Minimal Contrast and additionally showed a significant *wh*-movement effect in the ventral part of the Precentral Cortex, in the Cerebellum and Broca pars orbitalis.

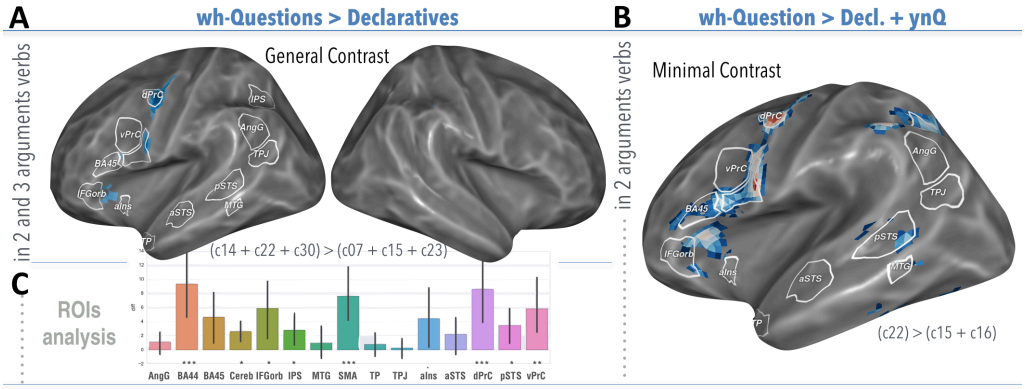
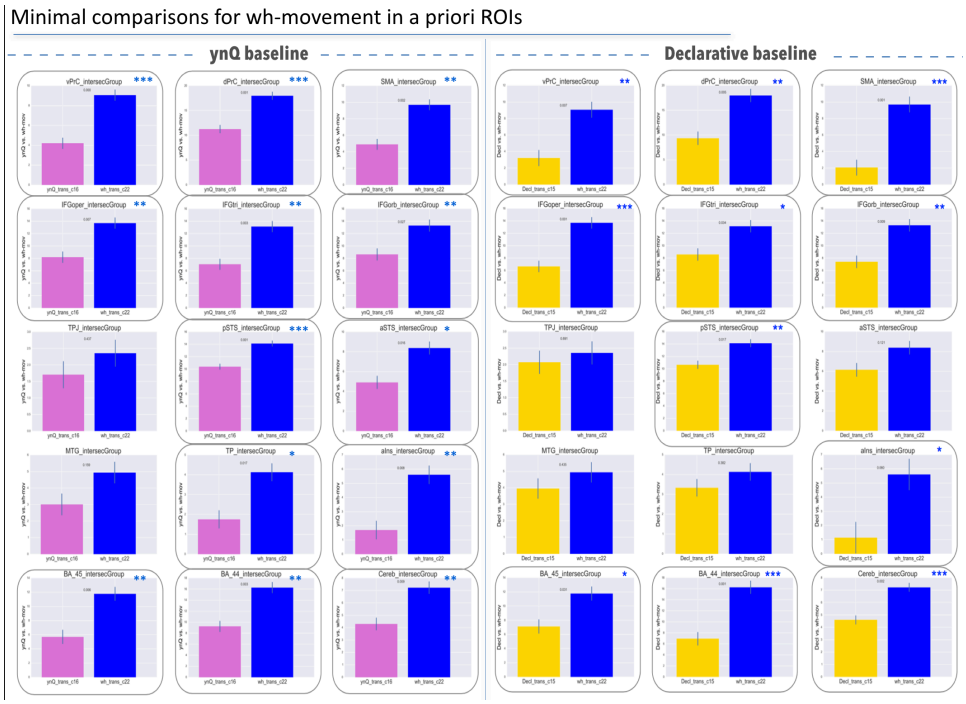
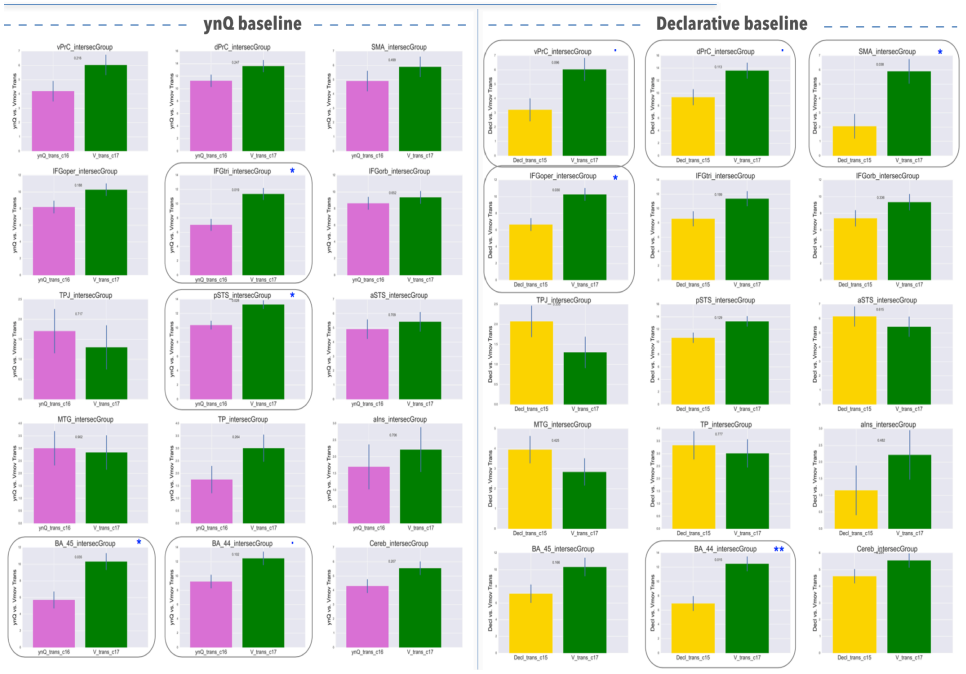


Figure 6.16 – Effect of *wh*-movement. (A) General whole-brain contrast with declarative baseline for *wh*-movement effect, and its ROI results (down). (B) Minimal whole-brain contrast with declarative and y/n questions baselines for *wh*-movement effect.

As for Single contrasts, they showed a nearly identical pattern when a *wh*-question (c22) was compared to a declarative (c15) or to a simple yes/no Question (c16). Significant differences were observed in vPrC, dPrC, SMA, Broca opercularis, triangularis and orbitalis, pSTS, anterior Insula and Cerebellum for the declarative baseline. And when a y/n Question (c16) is subtracted, we can note that an additional significant difference is observed in aSTS and TP (see Figure 6.17a).



(a) ROI barplots of beta averages for *wh*-movement condition (c22), Declarative condition (c15) and simple y/n Question (c16).



(b) ROI barplots of beta averages for Verb-movement condition (c17), Declarative condition (c15) and simple y/n Question (c16).

Figure 6.17

Verb-movement effect against baselines

The General contrast opposing yes/no questions with Verb-inverted y/n questions (e.g. 'Dort-il?' *Sleeps-he?* or 'Critique-t-il ça' *Criticizes-he that?*) to Declarative sentences without Verb-movement (e.g. 'Il dort.' *he sleeps* or 'Il critique ça.' *He criticizes that.*), yielded an activation map with a stronger bilateral activation in the anterior Insula and in a Parietal cluster, and a left lateralized cluster in the mid-temporal lobe, compared to *wh*-movement.

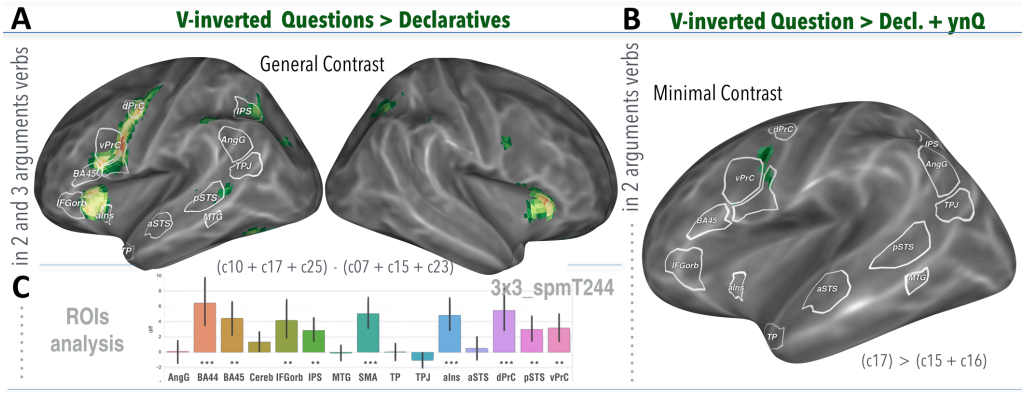


Figure 6.18 – Effect of Verb-movement. (A) General whole-brain contrast with declarative baseline for Verb-movement effect, and its ROI results (down). (B) Minimal whole-brain contrast with declarative and y/n questions baselines for Verb-movement effect.

However, General and Minimal contrasts implying Verb-movement compared to non-movement baselines (i.e. declarative and interrogative baseline) only share a common increased activation in the cluster going from the left Precentral down to the IFG triangularis cluster (compare Figure 6.18 A and B).

This pattern of activation is consistent across a number of sub-analyses (e.g., without *wh*-question; without clitics, without Unaccusatives, see Annexes - Supplementary materials §F.1, p.929), but seems to vary accordingly to the number of augments of the verbal item.

Verb-movement: ROIs Analyses The General contrast at the level of ROIs additionally reveals a significant difference in Broca triangularis and in the pSTS. ROIs Analyses for the Single comparison of yes/no questions with and without Verb-movement, i.e. [(c17) > (c16)], confirm the involvement of these areas. We observed significant differences in Broca triangularis (BA45) and opercularis (BA44), and in the pSTS. The comparison of yes/no questions with Verb-movement to their parallel declarative sentences without movement [(c17) > (c15)], yielded additional significant differences in vPrC, dPrC, SMA, Broca opercularis, and BA44 (see Figure 6.17, compare A and B, p. 540).

Globally, we can note that ROI analyses suggest the areas sensitive to *wh*-movement are also sensitive to Verb-movement, although to a lesser extent: similar trends in these regions are actually observable when significance is not reached. Only TPJ, MTG and TP show inverse patterns, in particular for Verb-movement (see Fig. 540B). Direct contrasts will be more informative about the difference in neural patterns between these

two movements (§6.2.4.2).

Clitic-movement

Sentences with the two other movement types produced different patterns. It has to be noted that, in order to allow a comparison with the above two types of movements, we used simple yes/no question featuring clitic- or NP-movement to asses the neural signatures of these two movements types in an interrogative context, to have the possibility to compare them with the other movements that both imply interrogative clausal typing.

Hence, clitic-movement effect was assessed with Basic and single whole-brain contrasts in interrogative questions and a second type of comparison was made possible by the presence of two clitic movement in three arguments' conditions, thus allowing to highlight the different neural patterns linked to two clitics sentence compared to sentences displaying two strong pronominal like 'ça' *this* and 'là' *there*.

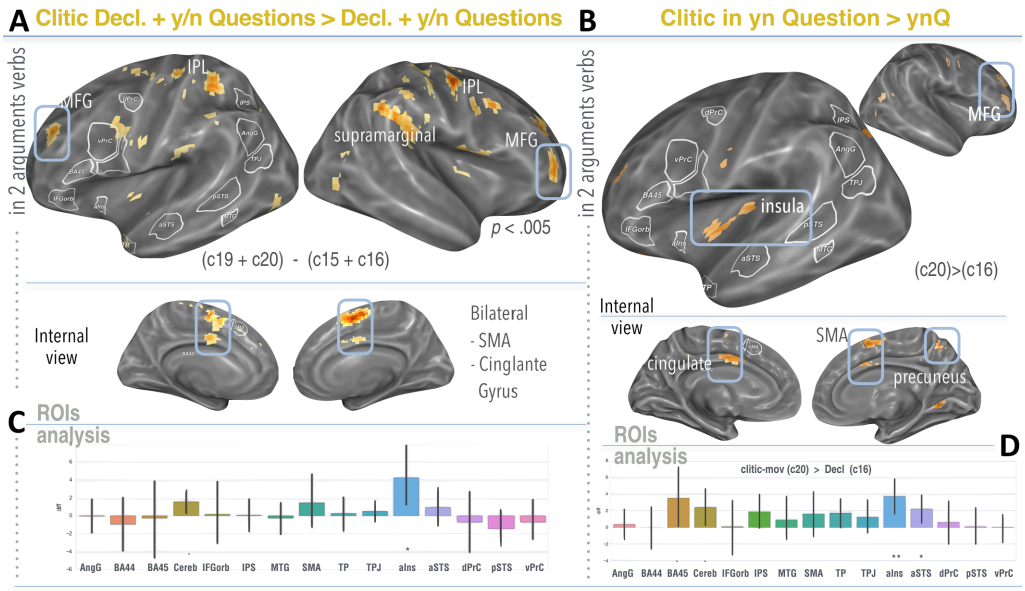


Figure 6.19 – (A) Internal and external view of whole-brain Clitic effect in declaratives and simple yes/no questions with clitics compared to matched sentence containing strong pronominal deictics ('ça' and 'là', *this* and *there*), related ROI contrasts effects are found in (C). (B) Internal and external view of whole-brain Clitic effect only in simple yes/no questions with clitics, related ROI contrasts effects are found in (D).

The Basic contrast in Figure 6.19A revealed only two extended clusters surviving the 50 voxel cluster-correction: one located in the right MFG (BA9) and the global maxima in bilateral SMA (BA6). Lowering the cluster extent correction to 25, only revealed an additional significant cluster in the right Precentral gyrus. Note that the other clusters shown in Figure 6.19A do not resist cluster-extent correction, we just decided to show a brain map at $p<.005$ threshold to offer the reader a clearer idea of the activation patterns obtained at a lower threshold, in order to better understand the ROI analyses and the convergence of the patterns showed by other contrasts for clitics.

The Single contrast (Fig.6.19B) confirms the involvement of three main activation clusters (at 50 voxels cluster-correction), respectively found internally Mid-Cingulate Gyrus (BA32), the upper part of the SMA (BA6) and Cerebellum bilaterally. It additionally confirms the involvement of the two right lateralized extended clusters that was observed in mid-frontal Superior and Middle Frontal Gyrus (SFG / MFG-BA9) and in right Precuneus, while left hemisphere showed two clusters in posterior part of Insular Cortex (adjacent to Putamen) and a very internal one in left Precuneus.

We can note that *anterior* insular cortex adjacent to Broca orbitalis revealed a smaller cluster only at 25 voxels cluster correction.

Moreover, sentences featuring two clitics elicited an increased activation in two clusters, with a global maxima in the left Supramarginal gyrus (adjacent and posterior to Angular Gyrus), an the left SFG (BA10 / MFG) as illustrated in Figure 6.20 (see also p.543).

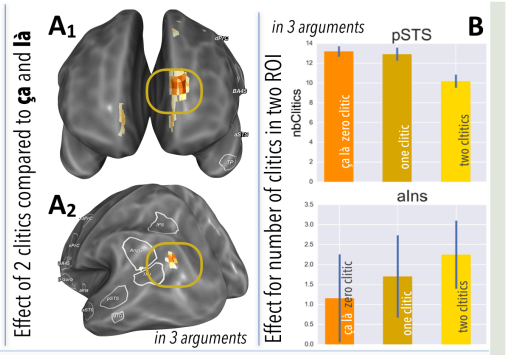


Figure 6.20 – (A) Effect of two clitics versus *ça* and *là* strong pronouns in 3-argument conditions. (B) Two temporal ROIs showing an opposite decreasing and increasing pattern of responses two 0 , 1 and 2 clitics in three arguments declaratives.

cl-movement ROIs Analyses ROIs analyses of the Single comparisons between y/n Questions with and without clitics [(c20) > (c16)] (see 6.19D) significantly implicated the left Insula, left aSTS, and at a more lenient threshold the left TP, left IFG pars triangularis (BA45), and Cerebellum. The increased activation in Insular and cerebellar ROIs is confirmed for the Basic contrast in Fig. 6.19C.

ROI analyses focusing on the number of clitics in three arguments declaratives (i.e. 0, 1, 2) show three different patterns of responses across a priori ROIs (see Figure 6.20B): an increasing one in anterior Insula, a decreasing one in pSTS and a stable one in areas like the MTG and the Angular Gyrus.

As a side note, in a more spurious contrast in Unaccusative declaratives with an additional locative argument, the comparison between one clitic condition (c11) versus (c07) showed a significant difference in dPrC, vPrC, SMA and Broca opercularis (BA44).

In sum, given these different sub-analyses, we can conclude that clitic-movement stably implies: left Insula, left aSTS, bilateral in the mid/anterior Cingulate gyrus, in the left and right MFG (BA9), the Precuneus and left Supramarginal gyrus (TPOJ). All in all, this indicates a greater involvement of temporal areas coupled by internal areas like Cingulate Cortex and the SMA.

NP-movement

Two contrasts were used to assess NP movement, one taking Unergatives as baseline comparing two one arguments verbs, and the second comparing Unaccusatives with a locative argument to transitives.

The first, comparing sentences with unaccusative verbs with SV word-order (e.g., 'Tu maigris.' *You lose weight*) to sentences with unergative verbs with SV word-order (e.g., 'Tu dors.' *You sleep.*), did not yield any significant activation, at the .001 threshold and even lower thresholds (Fig. 6.21B) or in the ROI analysis. The inverse contrast did not

yield any significant contrast either.

The second contrast compared Unaccusative verbs with an additional locative argument to simple transitive verbs, 'Il siege là.' *He sits here* versus 'Il critique ça.' *He criticizes this.* yielded an activation cluster in the MFG going down to the Superior Frontal Gyrus (SFG-BA10) bilaterally, with a cluster on the left in the Superior Parietal Lobule adjacent to the Precuneus, and a cluster in the right Pre-central cortex (BA6) (see Fig. A).

NP-movement ROIs Analyses ROI analyses yielded a single significant difference in the Minimal contrast opposing 2-argument declarative sentences [(c07) > (c15)] in the aSTS (Fig. 6.21C).

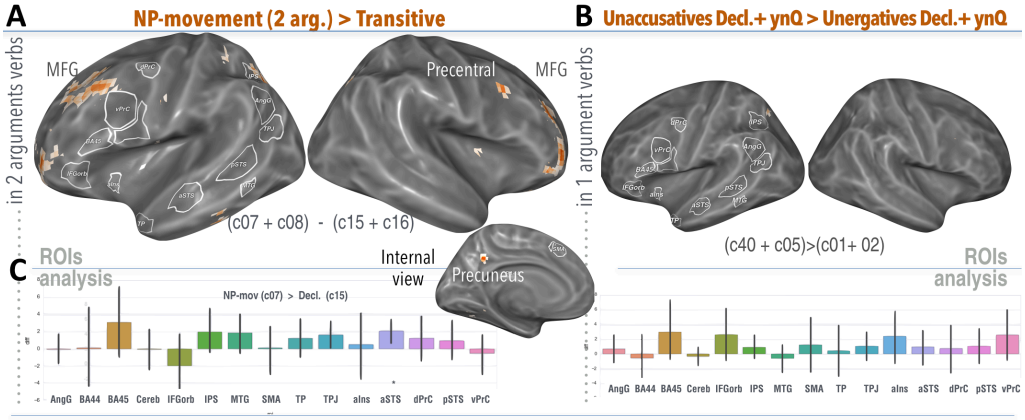


Figure 6.21 – Differences between NP-movement in one and two arguments (A) Internal and external view of whole-brain NP-movement effect in 2-argument unaccusative declaratives and simple yes/no questions compared to matched transitives, related contrasts effects in the ROIs in (C). (B) Whole-brain contrast comparing simple Unaccusatives declaratives and yes/no questions to matched Unergatives, related contrasts effects in the ROIs are found in (D).

Additional contrasts were performed to understand the effect of Unergatives and of Unaccusatives compared to Transitives. Following the rational exposed in previous fMRI study on Unaccusatives (Shetreet et al., 2010), we compared one argument Unaccusatives and Transitives: while the second have both a theme and an agent argument, the first syntactically have both a subject and an object position in that their theme argument is occupying subject position while having been generated in object position. As illustrated by Figure 6.21A, these comparisons showed three significant clusters (uncorrected for cluster-extent) in the right hemisphere, respectively in Inferior Temporal Gyrus (Fursiform Gyrus - BA37), Inferior Parietal Lobule (Supramarginal Gyrus) and Putamen (close to the Insula). The contrast opposing 1- and 2-argument Unaccusatives to matched Transitive declaratives confirms the involvement of right SFG (BA8) in Unaccusatives (Fig. 6.21B) present in both Minimal and Single contrast (Fig.6.21).

Interestingly, the whole-brain contrast opposing 1-argument Unergatives declaratives to Transitives declaratives reveals a unique increased activation cluster in the left IFG pars Orbitalis as illustrated in Figure 6.22C.

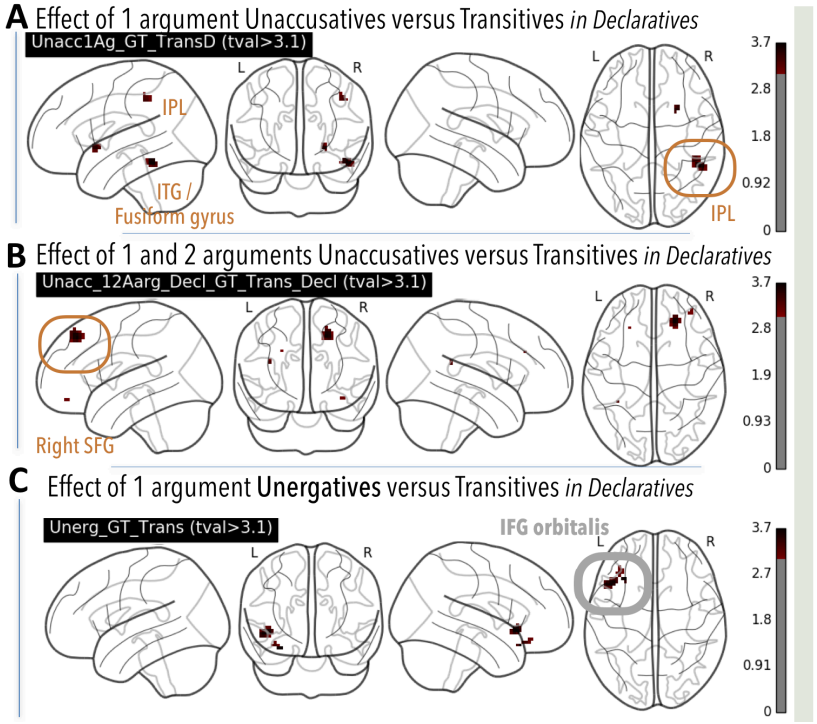


Figure 6.22 – Whole-brain contrasts: (A) Effect of 1-argument Unaccusatives versus Transitives in Declaratives. (B) Effect of 1- and 2-argument Unaccusatives versus Transitives in Declaratives. (C) Effect of 1-argument Unergatives versus Transitives in Declaratives.

All in all, these results indicate that NP-movement preferentially recruits the aSTS, the MFG and SFG bilaterally, the Precuneus and the Inferior Parietal Lobule. Although apparently disparate, when compared with previous findings for unaccustives in Hebrew (see Table 6.10, p.524), our findings replicate and confirm the involvement of:

1. aSTS which was repeatedly reported for Unaccusatives, irrespective of the chosen baseline by Shetreet et al. in two studies (2010 and 2012) with a more posterior distribution leaning towards MTG;
2. right Precuneus, that was also reported by Shetreet and colleagues (2012) in the contrast opposing Unaccusatives versus Unergatives;
3. left MFG and Left SFG, which had already been reported for Unaccusatives in Shetreet et al. (2010), respectively when comparing Unaccusatives to Transitives and to Unergatives in Hebrew.
4. left and right Inferior Parietal Lobule (/Supramarginal Gyrus), which was found with a left laterlized distribution in Hebrew by Shetreet et al. (2010).

We can conclude that both local movements in our design seem to recruit more temporal aSTS and (pre-)frontal pole (MFG) cortical resources, with a more complex response pattern for clitics which additionally elicit more internal areas encompassing the Insula, the SMA and the Cingulate cortex and TPOJ. We will now turn to the direct comparison between different movement types to further elucidate the differences we perceive from their individual signatures.

6.2.4.2 Effects associated to the contrast between different types of movements

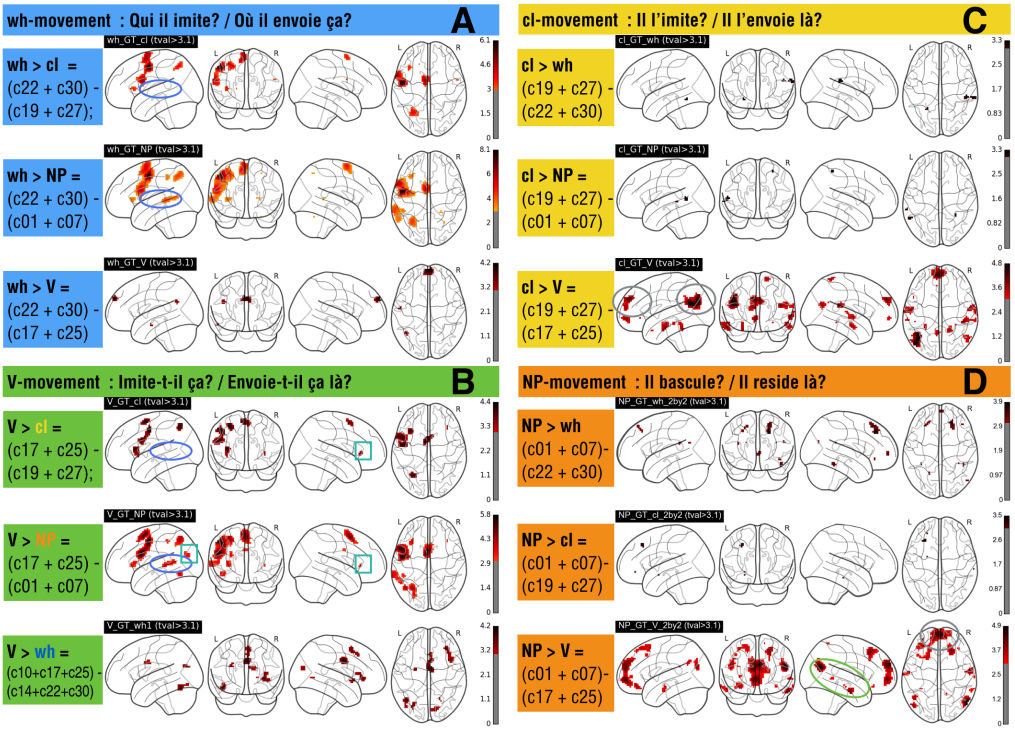


Figure 6.23 – Summary of direct contrasts among movements. Basic whole-brain contrasts in two and three argument verbs. (A) wh-mov compared to all other movements, (B) Verb-Mov compared to all other movements, (C) cl-mov compared to all other movements, (D) NP-Mov compared to all other movements. N.B.: No cluster size-correction has been applied on these maps that are all thresholded at $p<.001$ uncorrected (t-value>3.1).

As can be observed from Figure 6.23 (p.6.23), the direct contrasts between the different types of movement generally confirm the differences seen on the main effects maps: the Precentral/IFG-opercularis cluster is more sensitive to V- and wh-movement than to local movement like Clitic- or NP-movement .

Contrasting Verb and wh-movement: whole-brain and ROIs analyses

Specifically, the direct comparisons between Verb- and wh-movement, reveals for wh-movement four strictly left lateralized clusters: left IFG orbitalis and triangularis, the Inferior parietal region (BA40) and Inferior Temporal Gyrus (BA37). Note that this direct contrast was inclusively masked for positive effect of *wh*-movement in the basic contrasts.

For Verb-movement we observe different areas of increased activation, that are not only more right lateralized but also shifted to more infero posterior regions. Namely, Verb-movement effect compared to *wh*-movement significantly elicited the right IFG triangularis (/insular), the right SMA (BA8), and the left Temporo-Parietal-Occipital junction together with the mid-inferior Occipital Gyrus (adjacent to left Fusiform Gyrus).

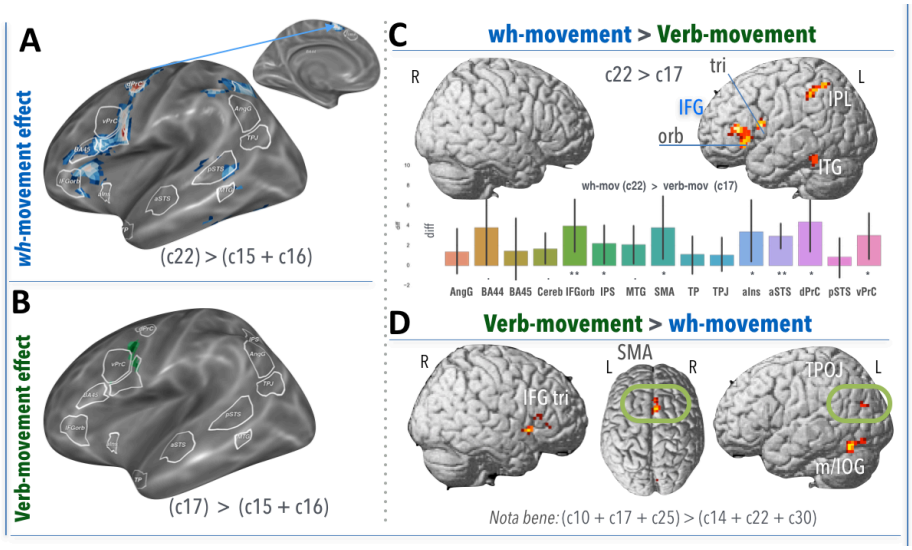


Figure 6.24 – (A) Effect of wh-movement in the Minimal whole-brain contrast with declarative and y/n question baseline. (B) Effect of Verb-movement whole-brain Minimal contrast with declarative an y/n question baseline for Verb-movement effect. (C) Single whole-brain contrast between wh-movement and Verb-movement, ant its contrast effect-size in ROIs. (D) General whole-brain contrast between Verb-movement and wh-movement..

As for wh-movement this direct contrast was inclusively masked for positive effect of Verb-movement in the general contrasts.

Nota bene: to compare Verb-movement to *wh*-movement, we actually selected the General contrast and not the Minimal one, because as shown in the General and Minimal contrasts for Verb-movement in previous section (compare Fig. 6.18A and B and 6.16A and B), an interaction was observed between number of arguments and Verb-movement, which is further confirmed by a post-hoc interaction analysis that we reserve to present in Section §6.4.1.2 (see Figs. 6.58 and 6.57, starting from p.588).

Contrasting Verb- and clitic-movement, whole-brain - ROIs Analyses

The comparison between two movements involving a two-stepped progression up the syntactic-tree, reveals significant increased activation for Verb-movement in the main spots where Verb-movement effect was observed against baselines, namely bilateral SMA, dorsal Precentral Gyrus going down to IFG opercularis and triangularis, a posterior cluster in left IPL and a right lateralized small cluster in the Insula/triangularis. Interestingly, the opposite contrast also confirms what was observed in the clitic whole-brain effect: (1) a bilateral activation of MFG and frontal pole and mid-anterior STS/MTG, (2) a bilateral, but mainly left lateralized broad activation of the Angular/Supramarginal gyrus, together with two right temporal clusters.

ROIs analyses reveal a stronger effect in all frontal ROIs and in Parietal region for Verb-movement, and a single significant increase in activation in the TPJ for Clitic movement.

Importantly, the long-distance movement to the CP-layer does not seem to involve

an increased activation of temporal areas of the Sentence-network compared to the more local movement implied by clitic placement, except for the small effect observed in the pSTS.

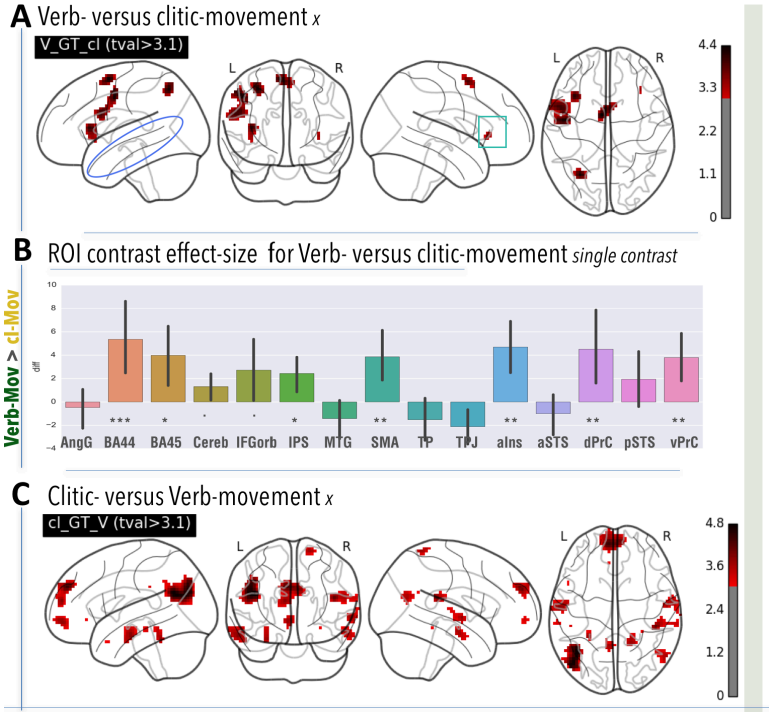


Figure 6.25 – (A)Basic whole-brain contrast opposing Verb-movement versus clitic-movement. (B) ROI contrast size-effects for Verb- versus clitic-movement (C) Basic whole-brain contrast opposing clitic-movement versus Verb-movement.

Contrasting long-distance movements (Verb and Wh-movement) and more local ones (NP- and cl-movement) - whole-brain As for local movements, Figure 6.23 illustrates that NP-movement activates always less than other types of movements, and that the comparison with Verb-movement reveals clusters of activation we already observed in NP-movement main effect. Specifically, NP-movement elicit an increased activation of bilateral MFG, frontal pole (SFG), Angular Gyrus and the mid-anterior STS/MTG.

We can also note that the contrast of clitic- and NP-movement allows the comparison between two local movements displacing object elements from the complement position in the VP to a higher position in the Tense Phrase (TP), that crucially proceed in different ways to reach their landing-site (see syntactic-trees in Fig. 6.1). While this reveals two very small clusters at the level of whole-brain analysis (one in the ventral Precentral gyrus for NP-movement and one in the pSTS for clitic-movement), ROI analysis shows only a significant increased activation for clitic-movement in Broca complex both triangularis and oparcularis, in the SMA, and dorsal Precentral Cortex, TPJ and two lenient effects in the IPS and the ventral Precentral region as illustrated in Fig. 6.26B.

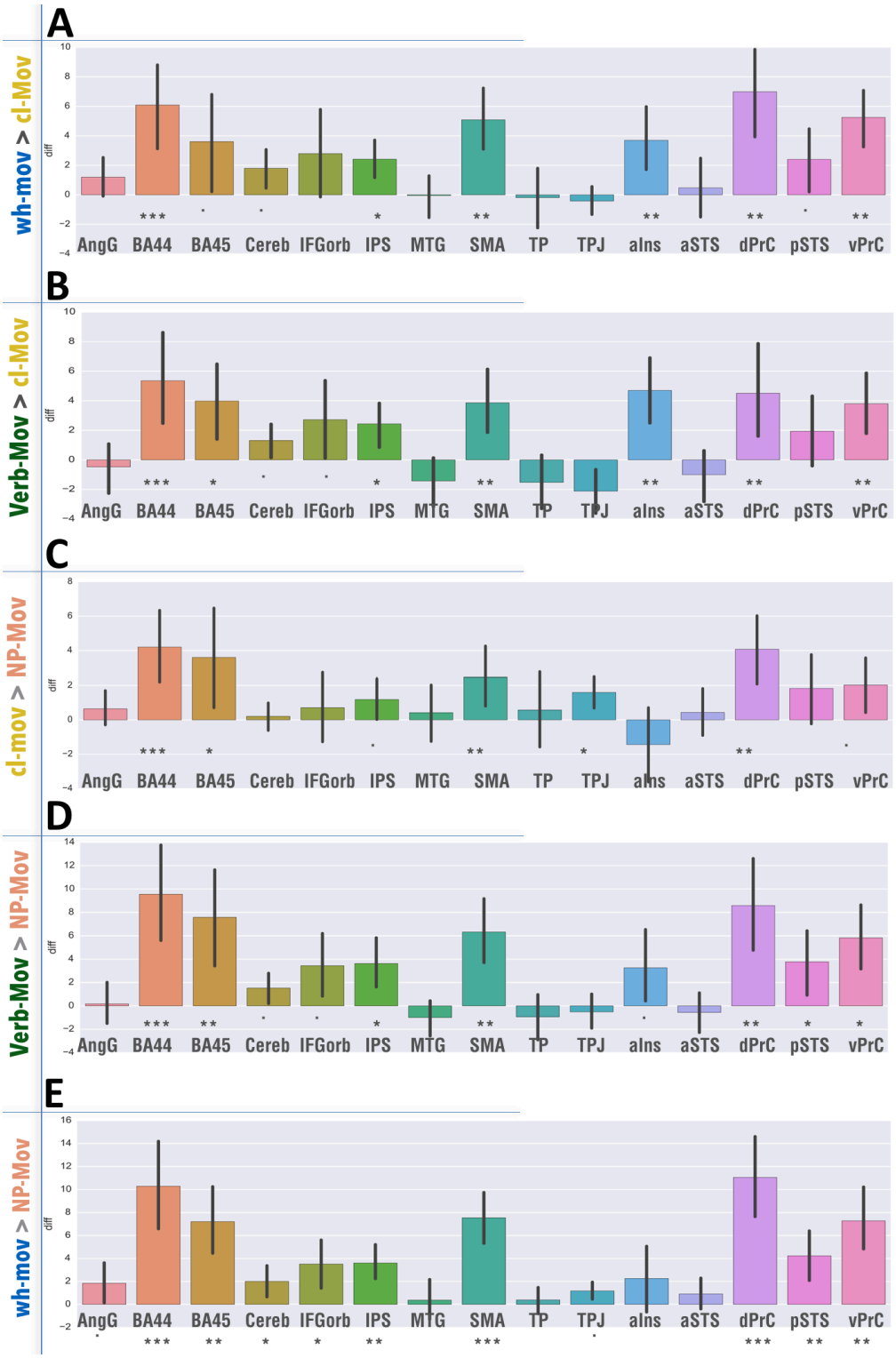


Figure 6.26 – Differences between average betas of single conditions featuring a movement type in the Transitive verb class. (A) wh-mov > cl-Mov, (B) Verb-Mov > cl-Mov, (C) cl-mov > NP-Mov, (D) Verb-Mov > NP-Mov, (E) wh-mov > NP-Mov.

6.2.4.3 Effects associated to the combination of different types of movements

Verb and Wh-movement: whole-brain - ROIs Analyses

As shown in Figure 6.27 contrasting wh-Verb-inverted questions with the declarative baseline reveals a significant increase of activation in dorsal precentral Cortex going down to IFG opercularis and in the SMA. To further investigate the possible additive effect of these two types of movements, we first contrasted wh-Verb inverted questions with the average of wh-questions and Verb inverted ones, which revealed a small activation cluster in a region posterior to Angular Gyrus, in the Supramarginal region (BA39) located posterior and between our regions TPJ and Angular Gyrus. Two finer contrasts were then performed to independently subtract wh-movement (Fig. 6.27B) and Verb-movement to wh-Verb-inverted questions (Fig. 6.27C). While the first did not yield any significant increased activation, the second revealed a broad increase in the same Supramarginal Region as above (the TPOJ).

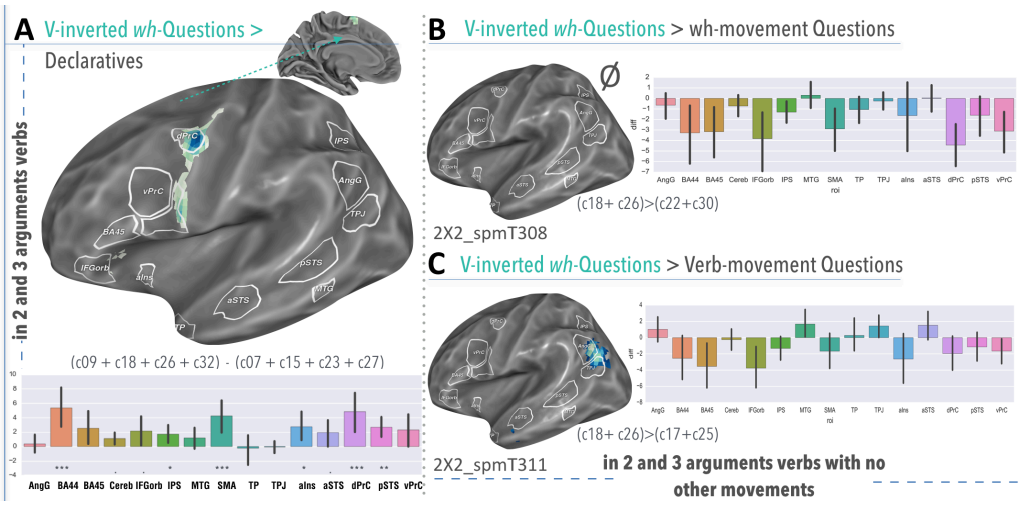


Figure 6.27 – (A) General whole-brain contrast for questions with both wh- and Verb-movement (wh-Verb-inverted questions - whQinv) minus matched Declaratives, relative ROI contrast effect-size (down). (B) Basic contrast for wh-Verb-inverted questions minus matched wh-questions, relative ROI contrast effect-size (down). (C) Basic contrast for wh-Verb-inverted questions minus matched Verb-inverted questions, relative ROI contrast effect-size (down).

The ROI analyses confirmed this pattern in the left TPJ region and MTG (see Fig. 6.43C, and also suggest that the left aSTS may show an additive effect of Verb and *wh*-movement (see Fig. 6.28).

The inverse contrast – the average of Verb-only and *wh*-only questions compared to wh-Verb-inverted (*wh*- + Verb-movement) questions – yielded stronger activation in the SMA at whole-brain level. And the ROI analyses also implicates the left anterior Insula, Precentral regions and the left IFG (pars opercularis, triangularis and orbitalis)[22].

22. As a side note, the event-related analysis that was run to understand the linear effect of movement-related complexity (i.e. effect of number of position in the model where a slice timing correction was performed in the pre-processing of brain images), actually revealed in Precentral cortex where we observed the global maxima of activation in Figure 6.43A (p.569) an earlier peak for questions where both

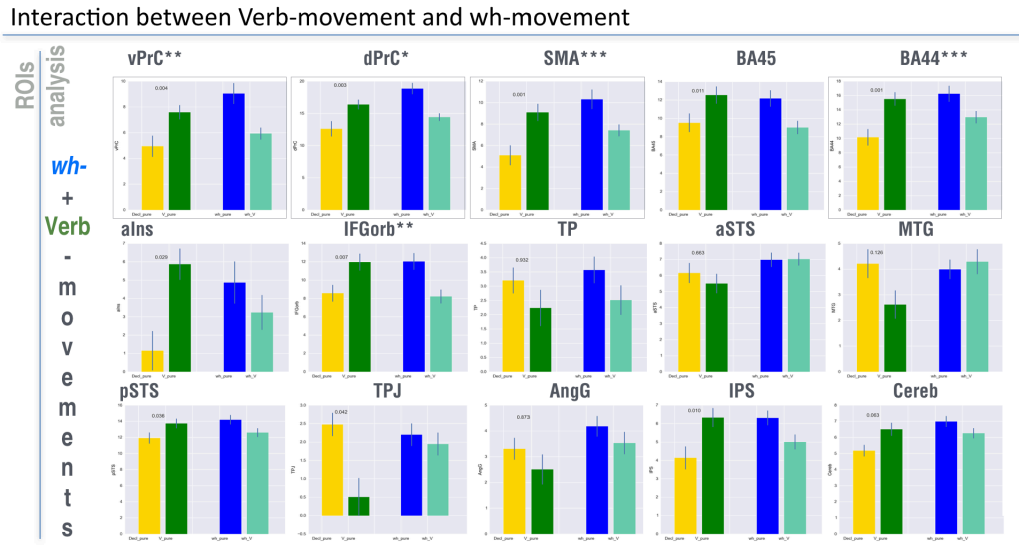


Figure 6.28 – Barplots showing the average betas of 2 and three arguments Declative baseline (yellow), Verb-inverted question (deep green), wh-question (blue) and wh-Verb-inverted questions (acquamarine) (from left to rigth). Interaction wh-movement and Verb-movement in Verb-inverted wh-questions conditions is siginficant in frontal areas encompassing Precentral Complex and Broca Complex.

As we can observe from the ROI analyses reported in Figure 6.28, *wh-* and Verb-movement do not seem to be additive: there is less activation when both types of movements are combined, '*Qui vois-tu?*' 'who see you' than when only wh- or only Verb-movement is present in all the frontal ROIs (i.e. Broca Complex and Precentral Complex) except Pars triangularis where only a trend is observed. Notably, the only ROI where we see an additive effect is the aSTS.

Post-hoc analyses to test for the interaction of different movement types are reported in the Annexes (§F.2.2, p.936).

Verb and clitic-movement: whole-brain - ROIs Analyses

The association between two movements involving a two-stepped progression up the syntactic-tree reveals an increased activation in the whole-brain analysis both when Verb and clitic movement are compared to one clitic sentences (A) and when compared to two clitic sentences (B), as illustrated by Figure 6.29.

wh- and Verb-movement than for wh- or Verb-inverted questions (see Fig. 6.34, p.557).

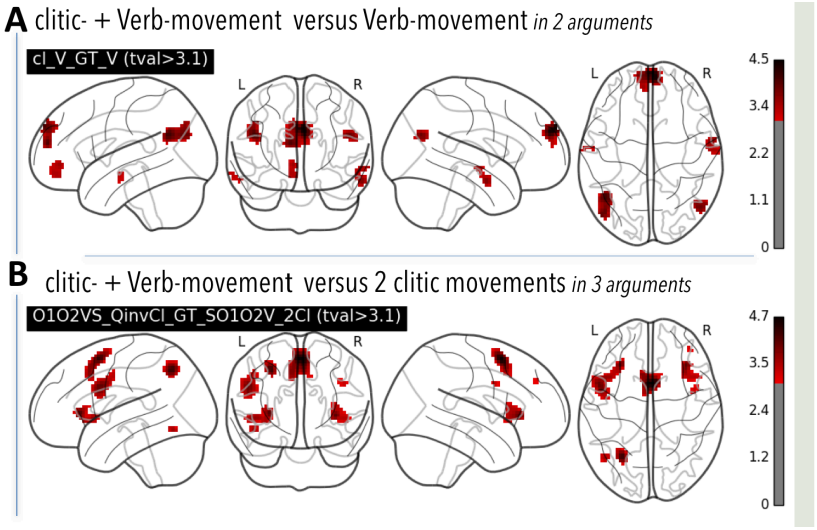


Figure 6.29 – (A) Basic whole-brain contrast opposing questions where both clitic and Verb-movement occurred versus V-inverted questions. (B) Single whole-brain contrast in three arguments conditions opposing Verb- and clitic-movement versus double clitic-movement.

6.2.4.4 Effect of Interrogative mode and question operator

Contrasting all questions vs. all Declaratives yields activations very similar to *wh*- and Verb-movement (i.e. the cluster Precentral-to-IFG Pars opercularis) as shown in Figure 6.30a, while a more refined contrast opposing all the simple yes/no Questions, obtained by adding a question mark to SVO order, to Declaratives yield a small cluster (8 voxels at p<.001 unc., see Figure 6.30aC) in the left anterior Insula/FOP. ROI analyses confirm the involvement of the anterior Insula, but not for all verb classes (cf. Annexes §F.2.3, p.936).

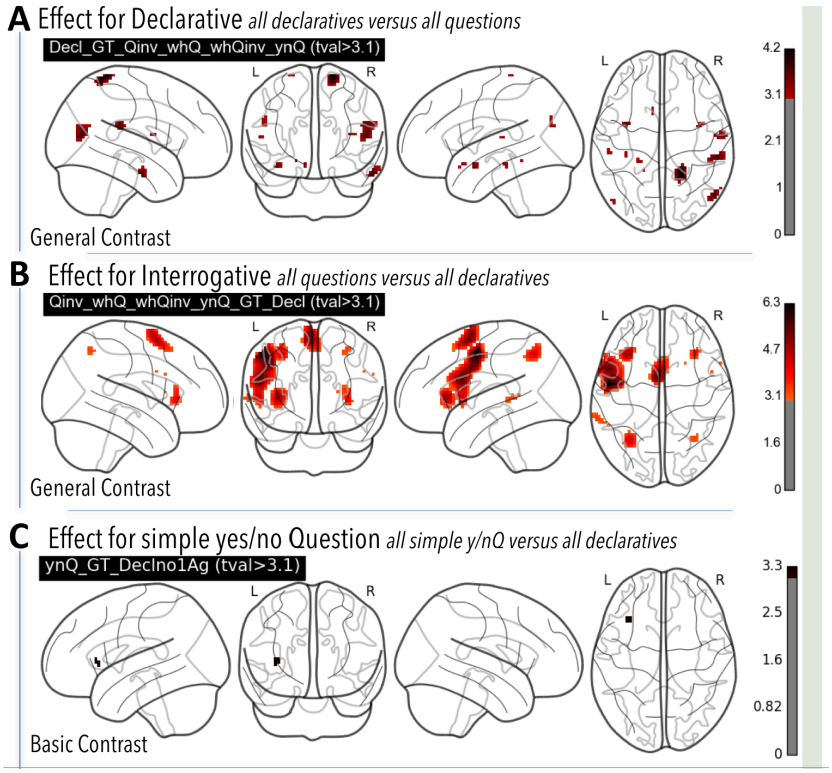
The opposite contrast reveals other right hemisphere clusters, the more extended was located in the TPJ, two other were respectively in a region anterior to Angular Gyrus and the superior Parietal region, a smaller cluster was found in anterior Middle Temporal Gyrus, as show in Figure 6.30aA.

Thus, restraining the contrast to yes/no questions in 3-argument verbs, we observe at the whole-brain level a significant increase in activation in the right Broca Complex (see Fig.6.47).

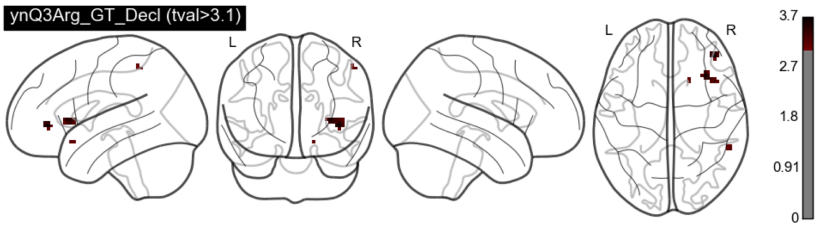
To further investigate the different scope of yes/no questions when the sentences feature a different number of arguments, we performed a minimal contrast opposing the 3-argument and one argument simple yes/no Questions, mediated by their respective declarative counterparts: [(c24 > c23) > (c05 > c04)]. This contrast revealed two right hemisphere clusters, a smaller one in Broca Complex and a more extended one in right Inferior Parietal Region (See Fig. 6.48, p.574).

Subtracting single contrasts with Interrogative or Declarative baseline

An additional interest of having simple yes/no questions in our experimental design was that they actually constitute a second Baseline for the two movement types that asso-



(a) (A) General contrast opposing all the declaratives to their matched questions across all the conditions. (B) General contrast opposing all the Questions across all the conditions to their matched declaratives. (C) Contrast opposing all the simple y/n Question in the 2- and 3-argument conditions to their matched declaratives.



(b) Whole brain contrast opposing 3-argument yes/no questions to their matched declaratives.

Figure 6.30 – Whole-brain contrast for effect of yes/no questions against matched declaratives.

ciate displacement to interrogative interpretation (Illocutory Force). Hence, subtracting simple y/n Questions to Verb-movement and *wh*-movement offers the opportunity to subtract also interrogative effect from movement-operation effects.

As shown in Figure 6.31, Minimal contrasts for the type of movement alternating declarative and interrogative baseline reveal more different responses for *wh*-movement than for Verb-movement. However, comparing the effect of subtracting the two baselines at the level of ROIs analyses leads to the opposite consideration, as shown in Figures 6.17 and 6.17b.

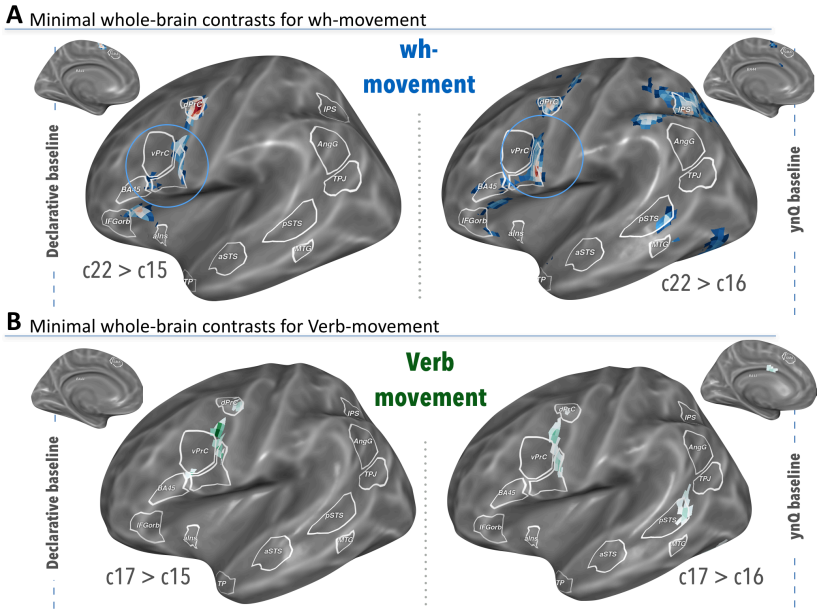


Figure 6.31 – (A) Minimal whole-brain contrasts for wh-movement subtracting the declarative baseline (left) or the y/n question baseline (right). (B) Minimal whole-brain contrasts for Verb-movement subtracting the declarative baseline (left) or the y/n question baseline (right).

6.2.4.5 Effect of number and type of arguments

A characteristic of our experimental design is that we used verbs that varied in number of arguments from 1 to 3 (see relative column of Stimuli table in Figure 6.2). This allowed us to look at the effect of this complexity parameter, as has been done in previous studies.

Figure 6.32B and C report the contrasts respectively opposing declaratives and simple y/n Questions across verb classes - [3arg>2args]; 3args>1arg)[23], comprising all conditions having a certain number of argument regardless of movement considerations (see the Annexes §F, p. 929).

In the ROIs (see Figure 6.32A), various patterns can be observed. Angular gyrus, IPS, vPrC, BA45 show a trend for increasing activation as a function of number of arguments. In the dPrC, activation for 3-argument Declaratives is larger than 2-argument and Unergatives. In the anterior Insula (aIns), a reverse pattern is observed with stronger

23. Note that 2arg>1arg did not yield any significant cluster.

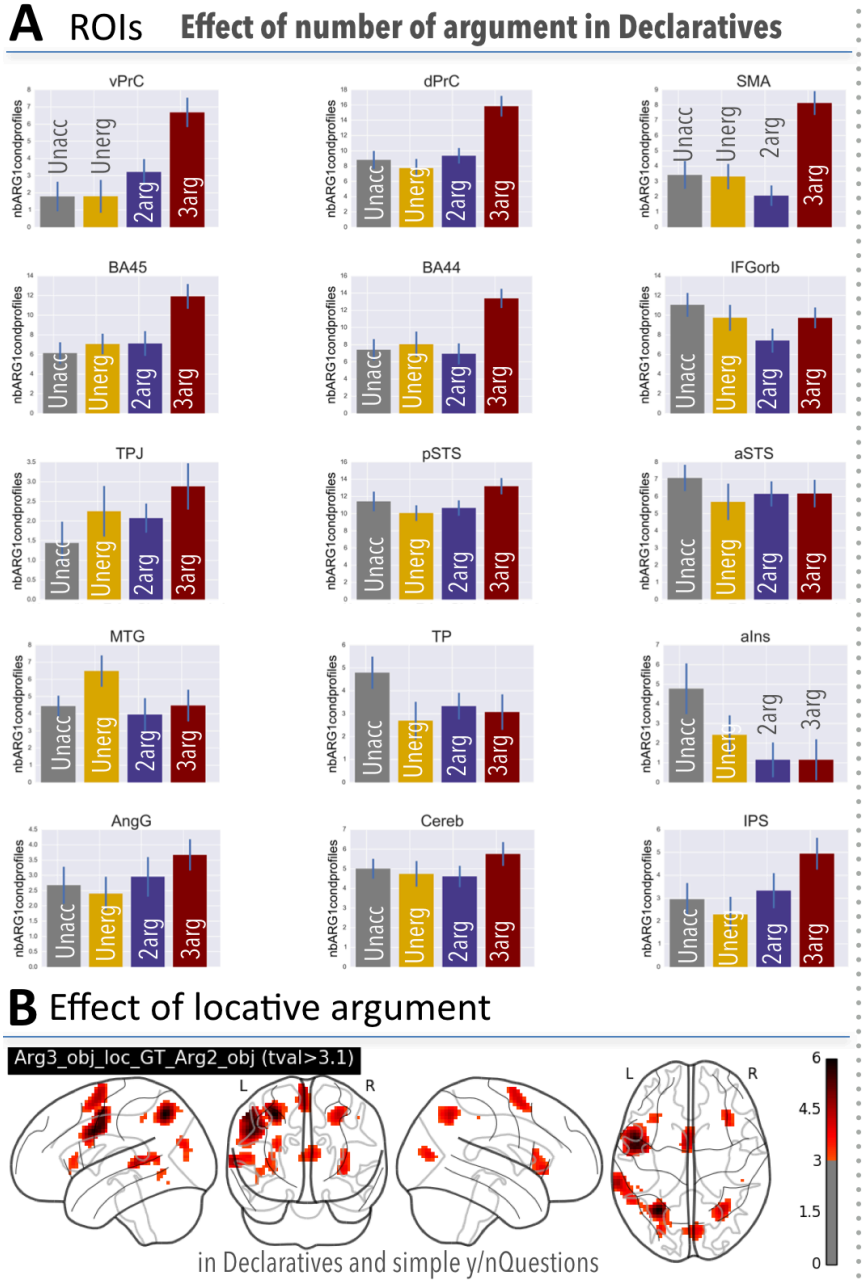


Figure 6.32 – (A) Bar-plots of Beta averages for Declaratives having different numbers and types of arguments in apriori ROIs. Note that 3-argument verbs and 1-argument ones differ in number of syllables and number of characters. (B) Effect of spatial adjunct argument in 3-argument verbs, obtained subtracting simple transitives in a General Contrast.

activation to 1-argument than to 2- or 3-argument. Temporal Pole region shows Unaccusatives with a stronger activation compared to other verb classes.

Effect of two-complement verbs with an additional locative In our design, 3-argument verbs are Transitive verbs that admit in their argumental structure a locative argument. The contrast between 3-argument conditions and 2-argument Transitive conditions allows to observe the effect of an extra locative adjunct argument in Figure 6.32B.

If we compare these activation patterns with those observed for movement complexity in *wh*-movement or Verb-movement effects, we can note that the effect of adding an argument has a huge repercussion on the Sentence Network, that is mainly observable in the three main activation peaks in Precentral complex, parietal region (IPS) and Angular Gyrus.

6.2.4.6 Effects of movement-related syntactic-tree complexity

As our wide range of experimental sentences contained 2 to 4 words and 0 to 3 empty categories (i.e. gaps) left by different movement types, we were able to correlate brain activity during sentence comprehension with the number of syntactic positions each sentence contained. Ranging from 2 to 7 syntactic positions, this syntactic complexity measure was used as a *proxi* measure of the representational complexity of sentences having undergone different types of syntactic-movement.

To find brain areas sensitive to the number of syntactic positions (i.e. the number of words + number of syntactic gaps/empty categories) in the syntactic-tree, we performed a linear contrast, computing a vector of weights set to the number of positions, which was then mean-centered. The map resulting from this contrast is shown on Figure 6.50, and reveals a peak activation in the dorsal Precentral going down the IFG complex.

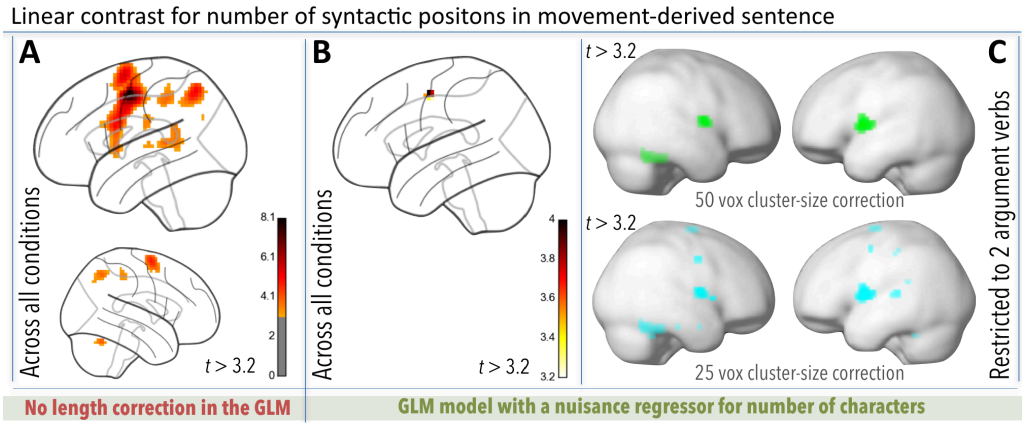


Figure 6.33 – (A) Whole-brain linear contrast across all conditions in the GLM model where no nuisance regressor for number of characters was inserted. (B) Whole-brain linear contrast across all conditions in the GLM model where a nuisance regressor for number of characters was inserted. (C) Whole-brain linear contrast restricted to 2-argument verbs that are minimally differing sentences in terms of length.

It has to be noted that across all the different analyses that were ran[24], a single stable

24. i.e. with and without slice-timing correction, or regressing out the effect for number of characters

cluster of increased activation was detected in the Precentral Sulcus at the following MNI coordinates: [-51 0 46] (see Figure 6.33A and B). Significant linear effect in ROIs is observed in pSTS, dPrC, Cerebellum and vPrC as illustrated by Figure 6.49A[25].

When restricting the whole-brain contrast to conditions with two arguments verbs, minimally differing in terms of length, the effect is even more widespread and encompasses at whole-brain level bilateral Broca complex mostly in its opercular part, and lowering the cluster-size correction it additionally shows bilateral dorsal and ventral Precentral clusters and a left lateralized increased activation in the pSTS as illustrated in Figure 6.33C.

However, as no addivity of *wh*- and Verb-movements is observable in the whole-brain contrasts in Figure 6.27 and in the ROI (Fig. 6.28, we decided to investigate the time-course of the focal activation observed in the Precentral-IFGop cluster in the contrast [V-inverted wh-Questions > Declaratives].

In this way, to better understand the linear effect of movement-related complexity (i.e. effect of number of position) observed in this area, we plotted event-related responses of Verb-movement, *wh*-movement, the combination of both, yes/no questions and declaratives. As shown in Figure 6.34 a stronger response for wh-questions ('Qui tu vois?') than to both Verb-inverted ones ('Vois-tu ça?') and their combination wh+Verb ('Qui vois-tu?') was observed. It is interesting to note that *wh*- and Verb-inverted questions featured a delayed peak: the response to wh+V (whQinv) peaks earlier than the responses to *wh*- and Verb-movement only [26].

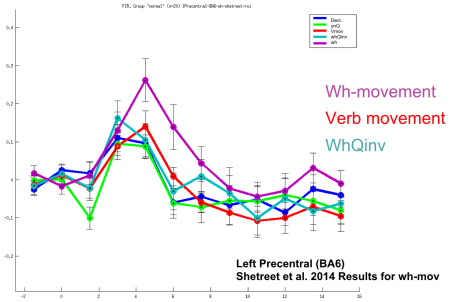


Figure 6.34 – Event-related plot FIR model for types of movements in Precentral apriori ROI.

6.3 Discussion of results, Cortical Responses to syntactic structures in sentences with dislocations

This section will delineate a discussion of the main results exposed in the previous section. While summarizing the specific effects of each movement, we will discuss the effect of the locality of syntactic movement (§6.3.1), the complexity linked to the number of derivational steps taken by different movements and by different combinations of syntactic movements (§6.43) and discuss the actual significance of the linear effect for number of syntactic positions (§6.3.6).

that was correlated to the number of words.

25. This result was also obtained by an exploratory model, that we first performed. At the individual-level, we created categorical models with one regressor for each number of positions (2-7) and two independent regressors for trials with single nouns and targets. The individual effect estimates were entered in a second-level within-subject ANOVA. In this first very basic categorical, we searched for regions where activation increased with the number of positions, neglecting the confounded variables (text size and number of syllables). We first computed the linear contrast search for areas where activation increases from 2 to 7 positions. The right graph shows the amplitudes (Beta-coefficients; Beta 1 to 6 represent number of positions). Three main regions are involved: Left Precentral cortex (BA6 peak/BA9), SMA (SFG, BA6), bilateral intra-parietal sulculs. Note that, when lowering the threshold, the precentral activation goes down to Pars Opercularis.

26. We can also note that whQinv is inferior to *wh*-movement.

6.3.1 Effects associated to the different types of movement

6.3.1.1 Comparing the complexity of A-bar movement and Head-Head movement in wh-movement and Verb-movement

wh-questions Our results for *wh*-questions are in line with all of the previous literature investigating *wh*-movement in other constructions. Comparing *wh*-questions ('Qui elle cite ?' who she cites) to declarative SVO sentences ('Elle cite ça.' *she quotes this*) in different verb classes and with different baselines and both to whole-brain and in ROI analyses, elicited significant activations in the left dorsal Precentral gyrus (BA6), the Supplementary Motor Area (SMA), the left IFG pars opercularis and triangularis, and the inferior parietal lobule bilaterally (see Figure 6.35).

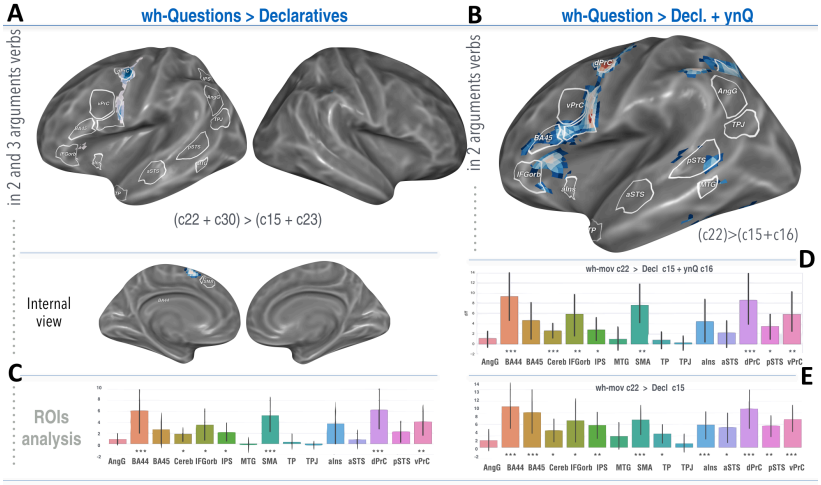


Figure 6.35 – Effect of *wh*-movement. (A) Basic whole-brain contrast with declarative baseline and its effect-size difference in ROIs (C). (B) Minimal whole-brain contrast with declarative and y/n questions baselines, and its effect-size difference in ROIs in (D). In (E) effect-size difference in ROIs for the Single contrast between wh-movement and declarative [(c22) > (c15)].

More concretely we reproduce a very similar activation pattern to that reported in Christensen (2008), who compared *wh*-questions versus phonologically silent interrogative operator in yes/no-question in Danish. Interestingly, as in our ROI analyses, the Danish fMRI experiment shows an increased activation in right cerebellum (see Figure 6.3C, p.513).

Cerebellum and syntactic-movement

Nowadays, the idea that cerebellar processing is required in a variety of cognitive functions is allegedly accepted, even if the exact role in the multiple cognitive domains is still vaguely defined and hardly discussed in neuro-syntax. Our cerebellar ROI is taken in a meta-analysis study by Stoodly and Schmahmann (2009), that addressed the issue of the functional specification of different sub-parts of cerebellum, indicating that the 11 language studies included showed a heavily right-lateralized involvement of cerebellum (right lobule VI, Crus I/Crus II). Note that it is generally accepted that cerebellar input

has a facilitatory effect on the contra-lateral cerebral cortex. Additionally, this study also put forward that language and executive tasks activated regions (i.e. Crus I and lobule VII) which are typically proposed to be involved in Prefrontal-cerebellar connectivity loops.

A lately emerged view on the perceptual and cognitive processing in which the cerebellum is involved puts special emphasis on sequencing processes. Thus, by adducing evidence obtained by neuro-physiological and neuro-psychological lesion studies[27], this trend points to sequence detection across modalities (i.e. temporal and spatial information in sequences) as one of the keystone of cerebellar function (e.g. Molinari et al. 2008). Specifically, Molinari et al. (2008) argue that "the cerebellum intervenes whenever a feed-forward control is needed by identifying patterns that allow a response to be anticipated. This general mechanism would act on simple responses like eye blinking as well as on complex social behavior".

Leggio and colleagues (2008) specifically considered cerebellar involvement in sequencing across verbal, spatial and behavioral cognitive domains in subjects with cerebellar lesions and observed that this population's scores are lower than control group, independent of the nature of the material processed. Interestingly, in case of right cerebellar damage, subjects present worse performances with verbal than with behavioral material, compared to those presenting lesions in left cerebellum.

Hence, if we consider that one of the movement-related complexity dimension resides also in sequence detection and prediction capacity, in that *wh*-movement displays both Argumental semantics process and Scope-Discourse semantics that crucially take place at different places in the sentence: (i) sentence initial position is entailed with Scope-Discourse semantics (in our case interrogative Force), (ii) at the gap position Argumental semantics is assigned[28]. Hence, when encountering sentence-initial *wh*-word, the Q-criterion implies a process of creating an object up-stream to assign the verb's argumental role later in the sequence.

Given this, we can actually better understand why cerebellum shows a linear increased activation with the number of movement operations and why it is observed in *wh*-movement at whole-brain level. Interestingly, as shown in ROI comparisons (Fig. 6.38, p.564), cerebellar increase of activation is always present when a long-distance movement is compared to a more local one, although it is fully significant only for *wh*-movement.

To further ground our speculation on the involvement of cerebellum in *wh*-movement, we can briefly add that one of the first studies investigating Topicalization also reported its increased activation. Notably, the task that Dogil et al. (2002) asked participants to perform is particularly informative in this respect. Two *re-serialization* tasks were performed. The first consisted in asking subjects to reformulate subject-initial sentences so that they started with a different constituent than the subject: a topicalization task, and the second was a list re-serialization task where participants were asked to reorder a list of words so that the second word was always moved to the first position (e.g. ABC-BAC).

27. "Roles of the cerebellum in cognitive functions are suggested by deficits induced by cerebellar lesions in cognitive planning, in practice-related learning and error detection, in judging time intervals, in rapidly shifting attention between sensory modalities, and in cognitive operations in three-dimensional space" (Ito, 2000: 159-160)

28. Note that Dynamic Agreement is only present in wh-movement, the Q is generated in the IP and then rises to CP, e.g. *who John has seen* (Q-Head, Qui)

Crucially, when subtracting the re-serialization task from the topicalization task, Dogil et al. found activation in the left dorso-lateral frontal lobe extending to Broca's area, and at the level of left temporal lobe, encompassing on Wernicke's area, anterior cingulate gyrus, and the cerebellum as well as in the dorsal prefrontal cortex (i.e. Precentral Complex).

As these areas are namely those to which literature on syntactic movement has repeatedly pointed, and the one we observed in our movement-related complexity effects[29], we can conclude with the authors that "the structure-dependency of syntactic operations is controlled by the delineated network" (2010:83), and that cerebellum is indeed involved in the processing of linguistic re-serialization present in Dogil's task as well as in our movement-derived stimuli.

In the light of Dogil's results and following Molinari's proposal and findings, we can conclude on the involvement of the right cerebellum in *wh*-movement and in other movements (cf. linear effect of words and gaps) by proposing that it is here involved in sequencing and detection of movement-derived sentence patterns. This step of sequential pattern detection could indeed constitute the basis for the linguistic system to further formulate predictions about next coming sequential sequence. Or even more specifically about the exact sequential position where the gap is found.

Verb-movement

If we turn to Verb-movement in French yes/no questions, we can first say that both whole brain and ROI analyses suggest that the areas that are sensitive to *wh*-questions are actually also sensitive to Verb-inverted questions, but to a lesser extent (compare Fig. 6.35 and 6.36). To explain this we rely on the fact that agrammatic patients generally do not show an impaired behavior in sentence having undergone Head movement (see Grodzinsky, 2007, for a discussion).

However, direct contrasts between these two types of movement show that the two movements preferentially recruit different sets of regions. IFG triangularis and orbitalis, together with IPL and ITG for *wh*-movement, while Verb-movement shows specific activation of left TPOJ, SMA and IOG, together with right Broca triangularis.

As illustrated in Figure 6.37 (compare B and D), we reproduce the results by Shetreet and Friedmann in the localizing part of Verb-movement complexity in the mIOG (Fusiform)[30], which lead us to follow the interpretation given by the authors about the recruitment of this rarely discussed area. Verb-movement has been analyzed as being motivated by phonological requirements[31].

IOG is a commonly reported area in another domain of language, that of reading

29. See also next chapter on the difference between Base-generated Topics and Movement derived Topics in Mandarin Chinese (§7.4.2 p.657) Figure 7.24.

30. The activation reported by Sh & Fr (2014) is in between IOG and lingual gyrus, while ours is more between mIOG and Fusiform Gyrus.

31. This position resumes to an initial position of Chomsky – which has later been debated – that all Head-Head movements are a 'matter of phonology'. As illustrated in chapter 2 Figure 2.28 (p.2.28), the phonological form (PF) is theorized as being a level of processing where the sentences are assigned with the phonological representation, after the construction of their syntactic structure. Hence, particular movements can happen either covertly at the level of Logical Form (LF) as it was proposed for Chinese *wh*-movement, or certain word-order variations can occur at the phonological stage. As for verb-movement Hebrew, the syntactic component produces an output that is in the canonical (Base-Generated) word-order, and then the verb is placed in pre-subject position only at the phonological stage

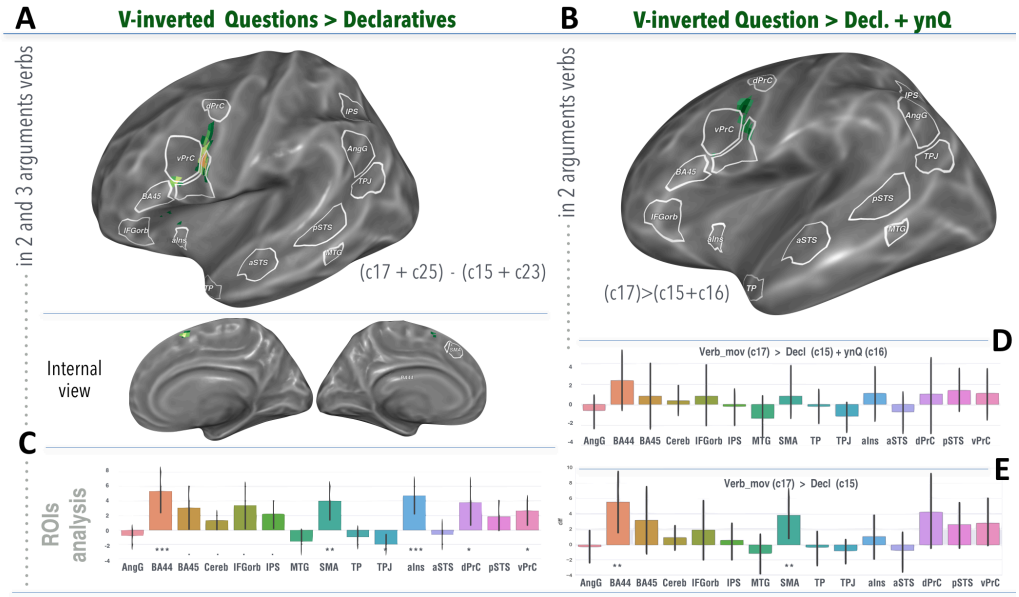


Figure 6.36 – Effect of Verb-movement. (A) Basic whole-brain contrast with declarative baseline and its effect-size difference in ROIs (C). (B) Minimal whole-brain contrast with declarative and y/n questions baselines, and its effect-size difference in ROIs in (D). In (E) effect-size difference in ROIs for the Single contrast between Verb-movement and declarative [(c17) > (c15)].

studies. Several studies have pointed to this region as play a phonological role, in the interface between phonology and orthography, and more particularly in the phonological retrieval from visual input across different scriptural traditions (see Price and Mechelli, 2005; Tan et al., 2005)[32].

Our results are still coherent with Shetreet and Friedmann (2014) conclusions that *wh*- and Verb-movements can be seen as occurring in a different linguistic level narrow syntax for the first and phonological one for Verb-movement, but also show that the movement happening at phonological level still modulates the sentence network yielding increased activation in the SMA, TPOJ and right Broca.

Another activation cluster observed for *wh*-movement in our findings echoes a result from Shetreet and Friedmann (2014). As illustrated in Figure 6.37A and E the Basic contrast between subtracting Verb-movement to wh-movement activated left an right mSFG (BA9), which the authors do not discuss. As they obtained it in Topicalisation and we obtain it in *wh*-question, we may speculate that its increased activation as linked to some discourse properties shared by narrow-focus question and Topicalization in the individuation of referent in discourse. Generally speaking, Frith and Frith (2003)'s review of theory of mind studies indicates that "inferences about the mental states of others activate a number of brain areas, most notably the medial pre-frontal cortex [(mPFC)]

32. The authors stress its involvement in rhyming task, phonological decision, pseudoword reading. This last one is actually a very demanding phonological task requiring requires grapheme-to-phoneme conversion through the phonological route, and phonological assembly in a phonological working memory buffer. Importantly, we should note that in Sh& Fr, the activation of the left IOG, cannot be associated with its role in reading, as all of the stimuli were presented auditorily.

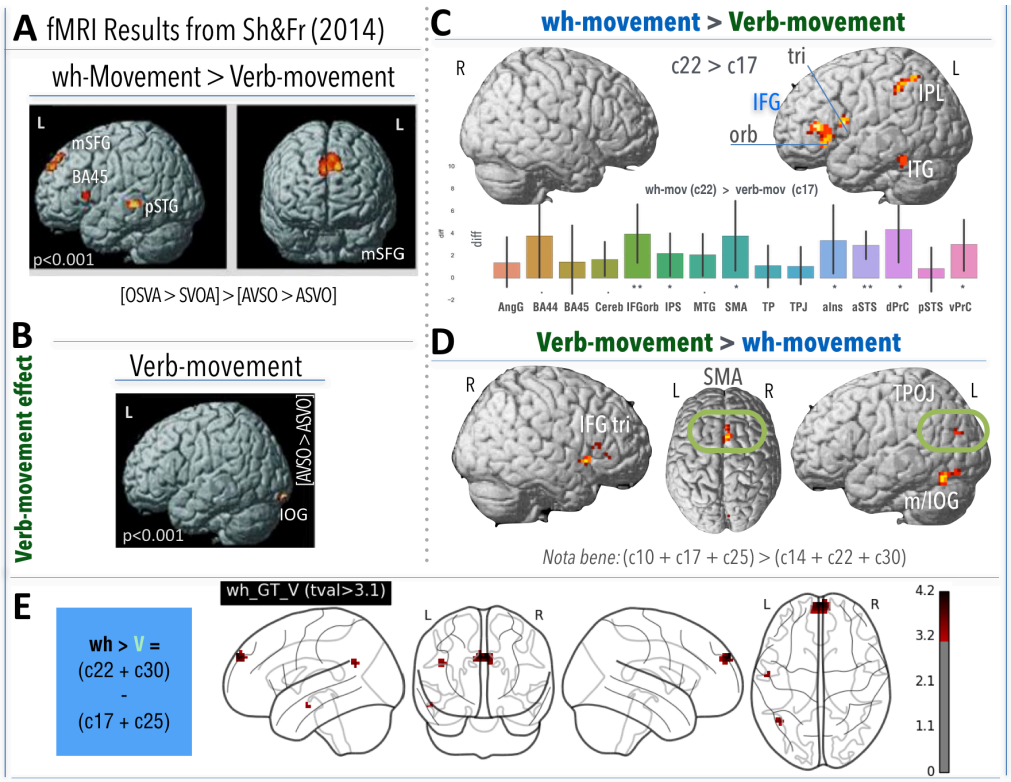


Figure 6.37 – (A) and (B) fMRI result from Shetreet and Friedmann (Sh& Fr) (2014) respecively on Topicalization versus Verb-movement, and Verb-movement effect aginst its baseline in Hebrew (V2). (C) Whole-brain contrast betwenen wh-question and V-inverted questions and its contrast effect-size in ROis. (D) Whole-brain contrast betwenen V-inverted questions and wh-questions in the Genreal Cotnrast. (E) Basic contrast for wh-movement revealing mFG cluster like in Sh& Fr (2014).

and temporo-parietal junction [(TPJ)]. This area, in broader terms (i.e. the dorso-medial Prefrontal Cortex, dmPfC), is reported as increasingly activated when comparing Narratives versus sentences in Xu et al. (2005, dmPfC) and was identified by Whebe et al. as being activated by Protagonist and pronoun linking in a text, in a study that correlated brain activity with a number of different variables present in a text (see Whebe et al. 2015, Leila Wehbe p.c.).

On the opposite, Verb-movement elicit greater activation in right Broca, which is interpreted as a mark of the semantic/pragmatic features of yes/no question interpretation (broad focus question) that is generated through Verb-movement. Right IFG has been namely linked to discourse processing (cf. Kuperberg et al., 2006) and theory of mind and discourse functions. As for the other areas, we will resume to TPOJ briefly, and SMA's role in movement-related syntactic complexity will be discussed more thoroughly in the next chapter.

Moreover, the fact that Verb-movement represents a movement of a Head to a Head position leads us to contrast it with clitic-movement (i.e. Head + NP-movement) and with NP-movement as detailed in the experimental hypotheses. As one can observe in

Figure 6.38 (p.564), ROI results reveal only one difference between these two subtractions in the pSTS, which is significantly more activated by Verb-movement only when compared to NP-movement (see pink circled bars), which we actually don't know how to interpret.

6.3.2 cl-movement neural signature in Precuneus

From the different sub-analyses presented in the Results section, clitic-movement involved the left insula, left aSTS, bilateral in the mid/anterior Cingulate gyrus, in the left and right MFG (BA9), the Precuneus and left Supramarginal gyrus. All in all, this indicates a greater involvement of temporal areas coupled by internal areas like Cingulate Cortex and the SMA.

We understand the increased insular (FoP) activation following Friederici's main distinction between local hierarchy building taking place in the frontal operculum and long distance hierarchical relations recruiting cortical resources in the Broca Complex (Friederici et al. 2006 or Zaccarella et al. 2015)[33].

In fact, the only comparison in ROI analyses yielding no significant activation in the Insula for clitic-movement is when it is compared to NP-movement, which is even more local in the syntactic-tree (see Figure 6.38C and B, p.564). The involvement of aSTS could be linked to the gender and number agreement involved in the first movement step or alternatively to the incorporation to the verb that is one of the main characteristic of clitics occurring in the final step of clitic-movement.

If we further consider the comparatives between movements in the ROI analyses in Figure 6.38, we can find three main confirmations of the point we discussed.

Firstly, the contrast between the two-stepped local movement found in clitic placement against the single-stepped one yields an increased activation in TPOJ. Interestingly, substracting to Head movement , the complex local clitic-movement yields a significant increase in activation in the neighboring ROI to TPOJ area showing that Head movement is indeed derivationnaly more simple probably because of the phonological motivation we discussed following Shetreet and Friedmann above.

Subtracting Clitic-movement or subtracting NP-movement to wh-movement yields different patterns in TPJ, Insula and IFG triangularis, showing that while NP-movement activates more the Insula, clitic-movement activates more the IFG triangularis and Temporo-Parietal junction region (TPOJ), as shown in whole-brain contrasts of these movements compared to baseline conditions.

Importantly, for our discussion on the neural underpinnings of clitic-movement, it should be noted that among the main areas it involves, the posterior part of left and right Precuneus attracted our attention as they were also observed by Bachrach (2008) as the second most extended locus of the unique contribution of derivation steps in the left hemisphere (see Figure 6.39B, p.566). We will partly rely on his discussion on the role of this area in movement-related complexity in the following.

33. While the Broca's area (BA44/45) can be seen as increasingly activated whenever the internal re-construction of a hierarchical structure from a sequential input is necessary, the Frontal operculum (FoP) is reported to be involved in the processing of local structural dependencies, following Angela Friederici's interpretation that is differentiating between local syntactic structure building and more long-range sentence structure building.

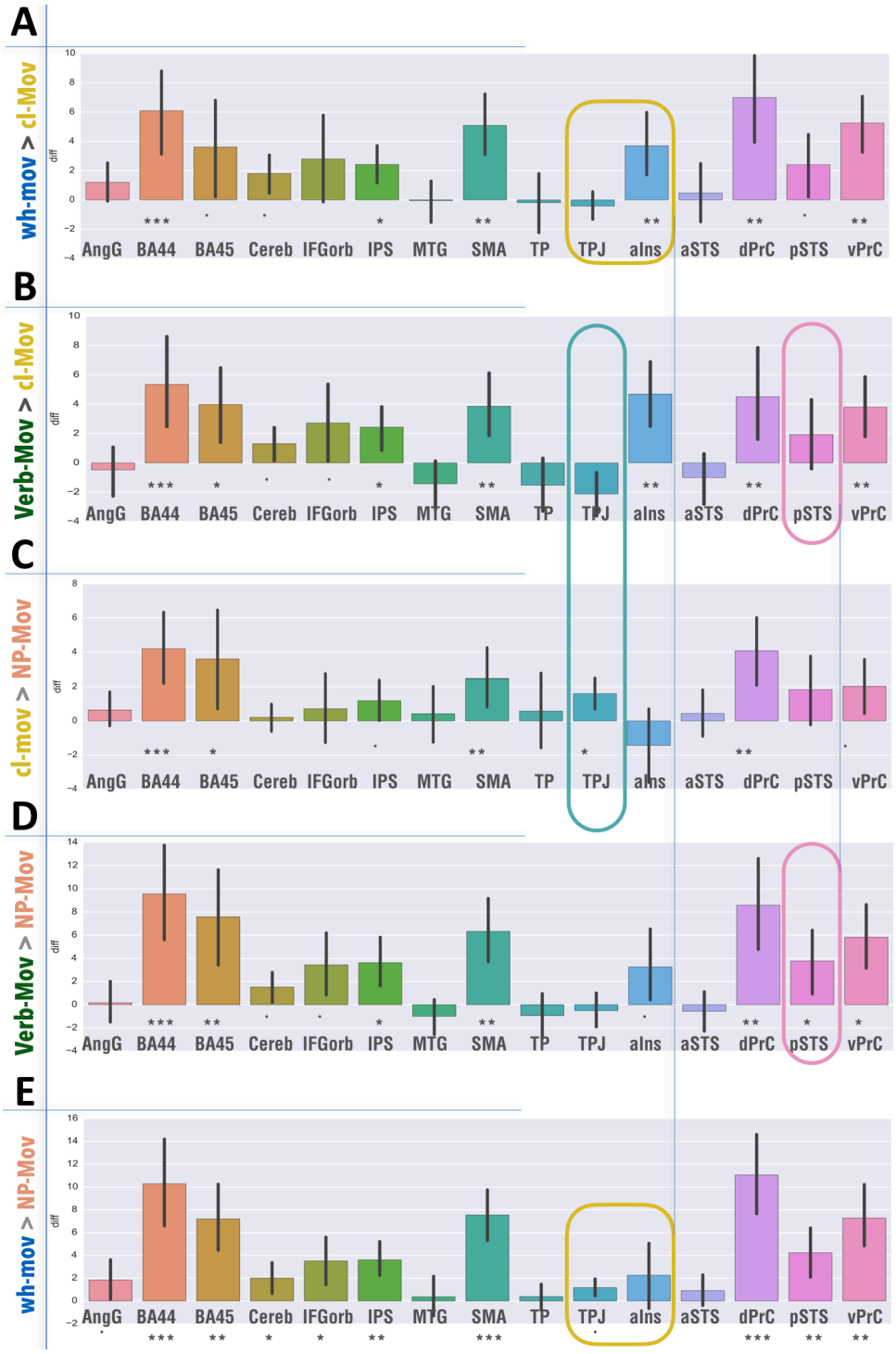


Figure 6.38 – Differences between average betas of single conditions featuring a movement type in the Transitive verb class. (A) wh-mov > cl-Mov, (B) Verb-Mov > cl-Mov, (C) cl-mov > NP-Mov, (D) Verb-Mov > NP-Mov, (E) wh-mov > NP-Mov.

Precuneus and syntactic complexity

Given the involvement of Precuneus in clitic movement, we looked for its role in the literature, as it is hardly included in the Sentence network. Although sometimes reported in contrast tables of sentence processing studies, it is generally less discussed or explored in the literature. An exception to this is to be found in Shetreet and colleagues (2007), who revealed graded verbal complexity effects in this area by manipulating the number of verbal complements, yet several other cognitive functions have been attributed to this area (i.e. the postero-medial portion of the parietal lobe), such as in the higher-order cognitive functions like episodic memory successful retrieval (more posterior), self-centered imagery (more anterior), theory of mind and mental tasks like navigation (for a review see Cavanna and Trimble, 2006).

To corroborate this gross division between an anterior portion, part of the 'default network' and self-oriented reflection, and a posterior section, involved in episodic memory, some studies found that the Precuneus is active in non-imagery related episodic memory of musical sequences and abstract words (Cavanna and Trimble, 2006).

Importantly, a very interesting[34] study by Cabeza et al. (2003) investigated the neural correlates of different stages of episodic memory retrieval and found the posterior part of Precuneus to be specifically engaged in the post-recall stage, which made the authors propose that this area plays an active role in the "processing of internally generated stimuli" (p. 390).

Elaborating on Cabeza's and colleagues' proposal, we partly follow Bachrach's discussion of this study, by affirming that "one could view the posterior precuneus as supporting the representation or processing of hierarchically complex internal representations. Music, theory of mind, navigation and associative or episodic memory all require complex internal representations and so we can identify a unified role for the posterior precuneus throughout these diverse domains" (Bachrach 2008:103). Bachrach namely reported precuneus activation correlating brain activity to the number of parser steps (see Figure 6.45, p.571).

Given this view on Precuneus, it is not hazardous to advance that it could actually play a role in syntactic processing, and particularly in the aspect we are interested in: the internally generated hierarchical representations and their format. However, our results taken together with Shetreet et al. (2007)'s ones on verbal complexity seem to indicate a narrower contribution of Precuneus in argumental complexity.

The graded activation of right Precuneus and anterior Cingulate in the results by Shetreet et al. (2007) can be reinterpreted as indicating a gradual increase of the hierarchical complexity resulting from the increasing number of complements (i.e. number of branches for the complements).Similarly, the recruitment of Precuneus and Cingulate Gyrus in clitic-movement compared to the baseline, can be interpreted as linked to the hierarchical (number of branches for the clitic complements) complexity that this movement operation yields in the IP and VP-layers (cf. AgroP projection Figure 6.8C, p.521). Note that NP-movement (in 2-argument Unaccusatives) yields right Precuneus increased activation at whole-brain level (see Figure 6.21A, p.544).

34. We will resume to it when discussing the function of Precentral cortex in section §6.3.6.

6.3.3 Number of derivational steps of different movement and different movement's combinations

To explain the increased activation we observed for clitic-movement in the region posterior to our Angular Gyrus and TPJ ROIs, two different interpretations can be advanced. The first would be considering the fact that this region, in very broad terms, has often been reported as the locus of theory of mind or pragmatic processes, which could indeed fit with the necessary topicality of the discourse referents of clitic pronouns. While the second would actually consider the results from a study from Bachrach's PhD (2008) we presented in chapter 2 (cf. §2.4.1.1, p.162), that investigated several syntactic-oriented computational measures, and identified this precise region as being correlated to a local measure of the structural complexity of the sentence, the number of derivational steps used during the integration of a given word into incrementally constructed syntactic structures[35] (see Figure 6.39B). Although our paradigm focuses on the global represen-

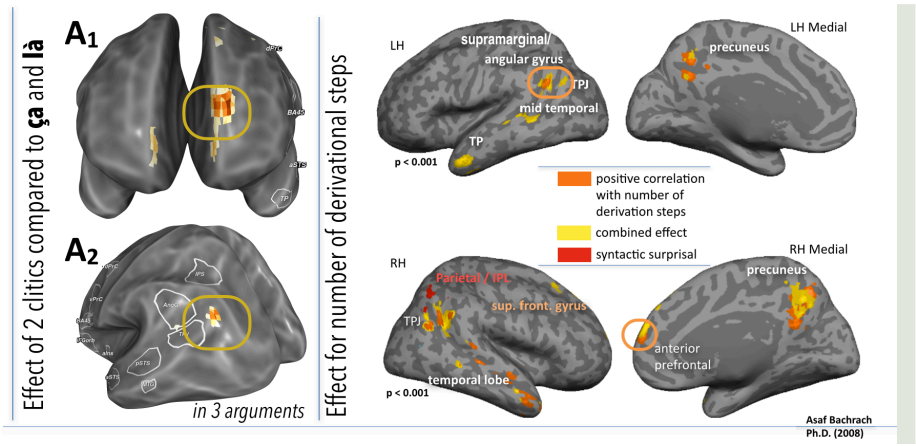


Figure 6.39 – (A1/2) Effect of two clitics versus *ça* and *là* strong pronouns in 3-argument conditions. (B) Effect of number of derivational steps in short narratives adapted from Bachrach (2008, p. 95).

tational complexity of the sentence and not on its incremental aspect, these findings indicate nonetheless that a complexity measure linked to the transformational aspects of sentence structure building, like the number of derivational steps, is indeed observable in the brain in this area. Moreover it has to be noted that the author also correlated brain activity with an average measure of these parser's derivational steps, which he calls syntactic complexity, in this way he offered some results that are closer to the syntactic complexity measures under focus in our study. This measure replicates the finding for Supramarginal and Precuneus left-hemisphere involvement in syntactic derivational complexity and crucially shows an additional frontal cluster precisely located in dorsal Precentral gyrus, where we actually observed a linear effect of number of syntactic positions[36] (cf. §6.2.4.6).

35. Nota bene, this corresponds to the number of Rorack's parser steps, being equal to the number of derivational rules applied at each word.

36. We can also add that the sentence model proposed by Bachrach is attributing to this Angular/Supramarginal Gyrus area the role of Phase integration, which he defines in the following terms:

We can link the kind of sentence structure complexity measures we investigate by considering syntactic movement and those correlated by Bachrach with brain activity during the comprehension of short narratives, through the prism of derivational theory of complexity (cf. §2.4.1) and its general underlying assumption about syntactic complexity. According to this theory, the more complex a representation is – the longer and more complex are the linguistic computations necessary to generate a representation – the longer a linguistic task should take, or the more the brain should activate resources, or put to work areas in order to build or access this complex representation, while performing sentence understanding.

Alternatively, we could simply consider the increased activation in this posterior Temporo-parietal area as reflecting the stronger theory of mind demands (cf. the role of bilateral TPJ in Saxe and Kanwisher, 2003), that are linked to the topicality of clitic referent. In French, as in other languages, the use of clitics is licensed by the salient discourse status of its referential antecedent (i.e [+accessible] and [+topical]). Notably, the contrast opposing two clitic sentences to sentences presenting two *in-situ* strong pronouns 'çà' and 'là' (Figure 6.39A) can help us decide between these two hypotheses. In fact, both clitcs (i.e. weak pronouns) and deictic strong pronouns are pronominal elements that actually share similar discursive licensing conditions: both require the search for a salient antecedent's referent in discourse carrying the [+accessible] and [+topical] features. Thus, the increased activation in Supramarginal Gyrus observed for clitics cannot be accounted by complexity effects linked to Topicality.

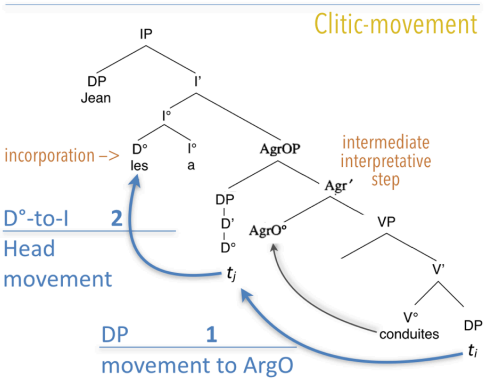


Figure 6.40 – Clitic-movement progression and its two-stepped derivation.

Moreover, as we will more thoroughly discuss in the next chapter dedicated to the neural underpinnings of different Topic-Comment constructions, studies investigating the discourse effects of Topicality (e.g. Caplan and Dapretto, 2001, discussed in next chapter) converge in indicating the central role of right IFG, which will be confirmed by our next fMRI study and by the absence of increased activation in this area in the contrast under discussion.

To further verify the functional attribution of this area, we focused on the contrasts opposing conditions that feature more or less derivational-steps (multiple combinations of different movement) or movement featuring more or less intermediate progression steps. We observed that for a great number of them an increased activation in supramarginal area is verified.

For instance, by considering the difference between different movements implying more or less intermediate steps in their derivation, we can note that the contrast be-

"[...]*the cyclic application of Spell-out to the syntactic structure.* This iterative interpretation of the syntactic structure at each Phase brings about the question of Phase integration. That is, how is the output of a smaller Phase integrated into the output of the Phase containing it? We propose that this integration takes place in the angular gyrus." The author grounds his claim on both the literature and his results, adducing that this area has been found to be particularly engaged by sentences containing syntactic displacement. Dronkers et al. (2004)'s lesion study or fMRI experiments, such as Cooke et al. (2002); Constable et al. (2004); Ben-Shachar et al. (2004); Christensen (2008), have repeatedly indicated it as a locus where increased activation is observed over stimulating sentential Displacement complexity.

tween two more local movements, namely differing in their movement progression – clitic-movement (2-stepped) and NP-movement (see Fig. 6.40) – actually show an increased activation in the TPJ ROI that partly overlaps with the TPOJ/Supramarginal cluster under discussion. Moreover, as illustrated in the contrasts presented in Figure 6.41, the association of clitic and Verb-movement shows an increased activation of TPOJ when compared to Verb-movement only sentences (see 6.41A), while it does not when compared to two clitic-movements ones (see 6.41B). Note that 2 clitic movements compared to baseline do show an increased activation in this area (see 6.41D).

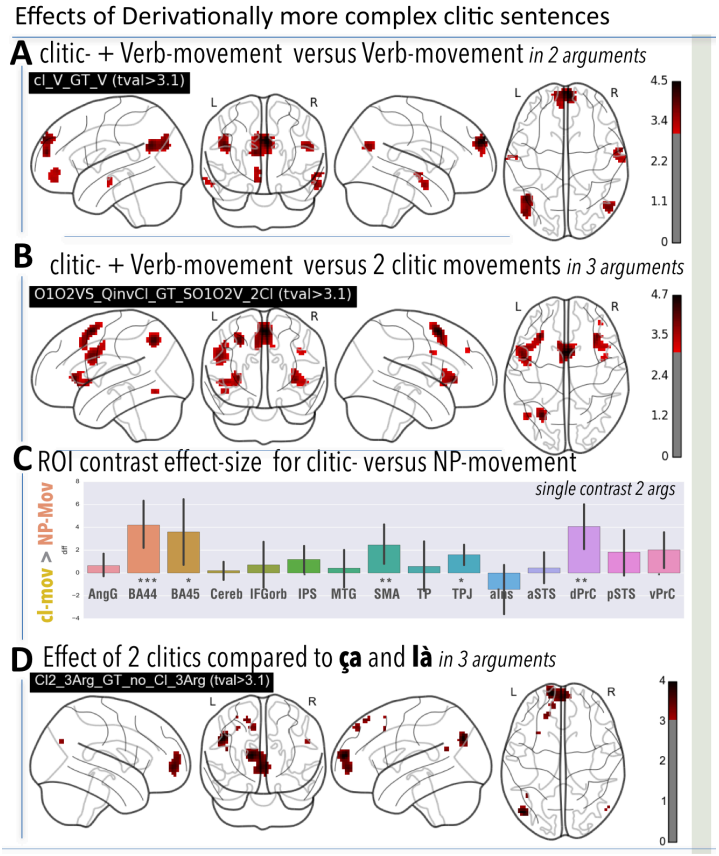


Figure 6.41 – (A) clitic + Verb-movement versus Verb-movement in 2 arguments (B) clitic + Verb-movement versus 2 clitic movements in 3 arguments. (C) ROI contrast effect-size for clitic- versus NP-movement. (D) Effect of 2 clitics compared to 'ça' and 'là' in 3 arguments.

Increase in derivational steps by combining *wh*-movement and Verb-movement

Wh-questions compared to declaratives and Verb-movement compared declarative sentences, overlap to a large extent, but their combination in the same question does not yield an increased activation or the expected additional effect, compared to the average of the two movements in isolation, and the two movement together in *wh*-Verb-inverted questions interact in several areas (see Bar-plots in the ROIs in Figure 6.28). In other

words, there is less activation when both types of movements are combined, '*Qui vois-tu?*' *who see-you* than when only *wh-* or only Verb-movement is present. One could explain this interaction by simply adducing the fact that in French, *wh*-questions are canonically implying Verb-movement. For example, 'Qui vois-tu?', *who see you*, is the normal way of asking 'Who do you see?', while the question with only *wh*-movement 'Qui tu vois?' *who you see* is a more colloquial way of asking the same question, although the pragmatic implication can be different.

However, the non-additive complexity effect of *wh*-movement and Verb-movement has already been observed for Hebrew in a study on aphasic behavior by Friedmann and Shapiro (2003). Agrammatic comprehension of object-subject-verb (OSV) and object-verb-subject (OVS) structures was investigated through a sentence–picture matching task. Importantly, while the first structure involves focalization by *wh*-movement, the second OVS involves two movements: a movement of the verb to the second sentential position, following the displaced object and preceding the subject (movement to C). Note that contrary to the French construction we investigated, in Hebrew this type of verb movement is optional and generally does not yield any change in the sentence meaning of the sentence. The results obtained show a lack of significant difference between OSV and OVS, suggesting that in situation where an argument has been moved (*wh*-movement in OSV structure), verb-movement does not have an additive contribution to the impairment.

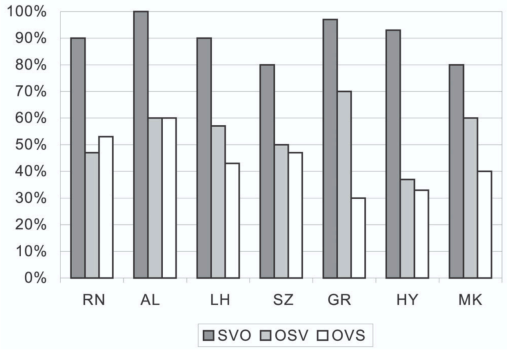


Figure 6.42 – Comprehension of SVO, OSV and OVS structures in the agrammatic aphasia group. Adapted from Friedmann and Shapiro (2003).

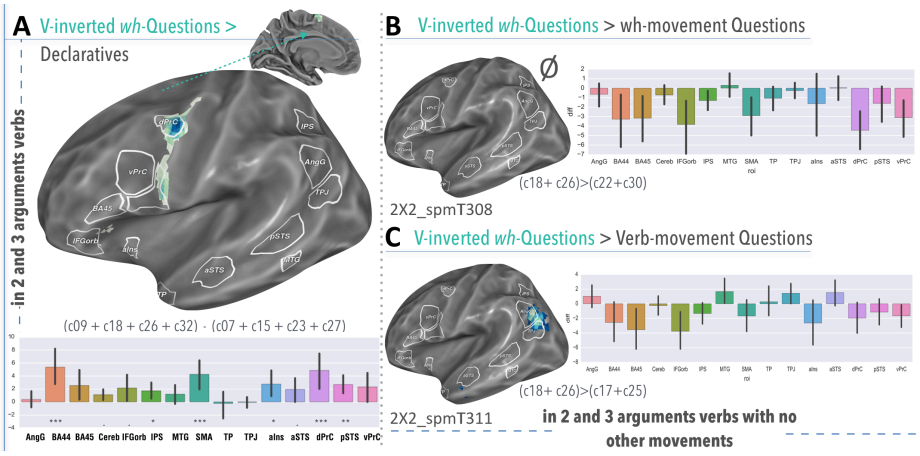


Figure 6.43 – Maps associated to the combination of both movements ('Qui vois-tu?'), the average of single movements ('Qui tu vois?' and 'Vois tu ça?'), and the contrasts between these two maps. (A) General whole-brain contrast for questions with both wh- and Verb-movement (wh-Verb-inverted questions - WhQinv) minus matched Declaratives, relative ROI contrast effect-size. (B) Basic contrast for wh-Verb-inverted questions minus matched wh-questions, relative ROI contrast effect-size. (C) Basic contrast for wh-Verb-inverted questions minus matched Verb-inverted questions, relative ROI contrast effect-size.

Following the interpretation given by these authors, we can say that the non-additive effect observed in agrammatic linguistic behaviour corroborates for the claim that verb-movement is not impaired in agrammatic comprehension (Grodzinsky, 1995), and further suggest in our case that the syntactic complexity implied in Verb-movement is not taxing into the resources that are impaired in this agrammatic population. We can observe in Figure 6.43 that the only region where the combination of *wh*- and Verb-movements generated significantly larger activity than Verb-movements, is a region posterior to Angular gyrus BA39/40 (TPOJ), the same area is present but to a lesser extent when subtracting the average of *wh*-movement and Verb-movement.

The ROI analyses confirm this pattern in the left TPJ (the ROI that overlaps most the activation cluster shown in Figure 6.43C), and also suggest that the left aSTS may show an additive effect of Verb and Wh-movement (see Fig. 6.28, p.551).

As we saw in comparisons involving clitic-movement, the posterior area in the Temporo-occipital-parietal Junction (TPOJ) is observed once more in a syntactic configuration where more derivational steps are required to achieve the displacement of one or more elements in the sentence. Following this interpretative direction, we verified if this was actually the case in other contrasts involving other movement combinations or in other movement comparisons. This is the case for the comparison between clitic movement and NP-movement, as already put forward previously (see Figure 6.41, but also in other contrasts that we summarize in Figure 6.44. Note that TPOJ also appears to be more activated, although to a lesser extent, when subtracting both NP-movement and wh-movement to Verb-movement (cf. 6.44D and E). Before leaving the topic of the deriva-

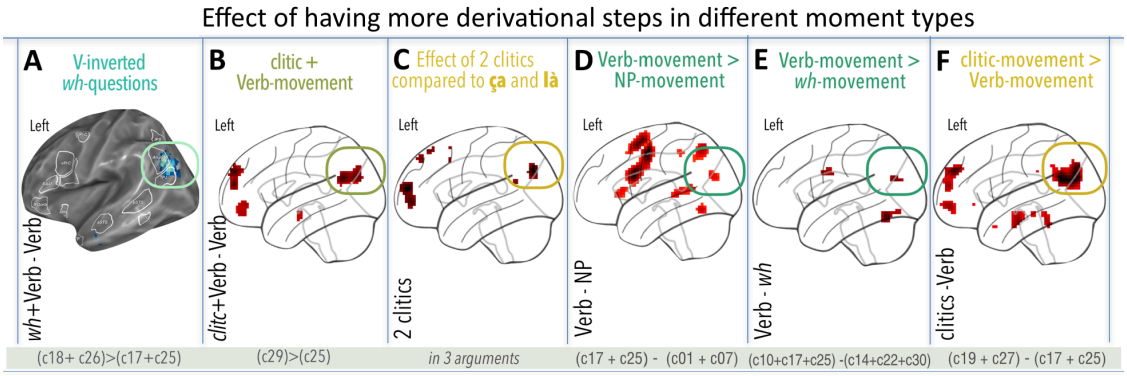


Figure 6.44 – Whole-brain contrasts for the contrast as described in the figure.

tional complexity, we briefly return to the comparison with the findings from Bachrach (2008), and rely on this to interpret the involvement of TPOJ in derivational complexity, as already noted in the above discussion about the derivational complexity implied in the two-stepped progression of clitic-movement (6.44F) compared to NP-movement (6.41C), or in double clitic-movement (6.44C), or the additive effect of Verb- and clitic-movement (6.44B).

Figure 6.45 illustrates how more or less fine-grained measures of syntactic complexity converge in showing that the increased activation of the area posterior to Angular Gyrus and TPOJ in the left hemisphere is correlated to displacement together with dorsal precentral small cluster (see6.45C). As noted above for the results in Figure 6.45A,

while Angular Gyrus shows a combined effect of syntactic entropy and of the average number of derivational steps, the more posterior region under discussion (TPOJ) shows an activation uniquely driven by the number of derivational steps.

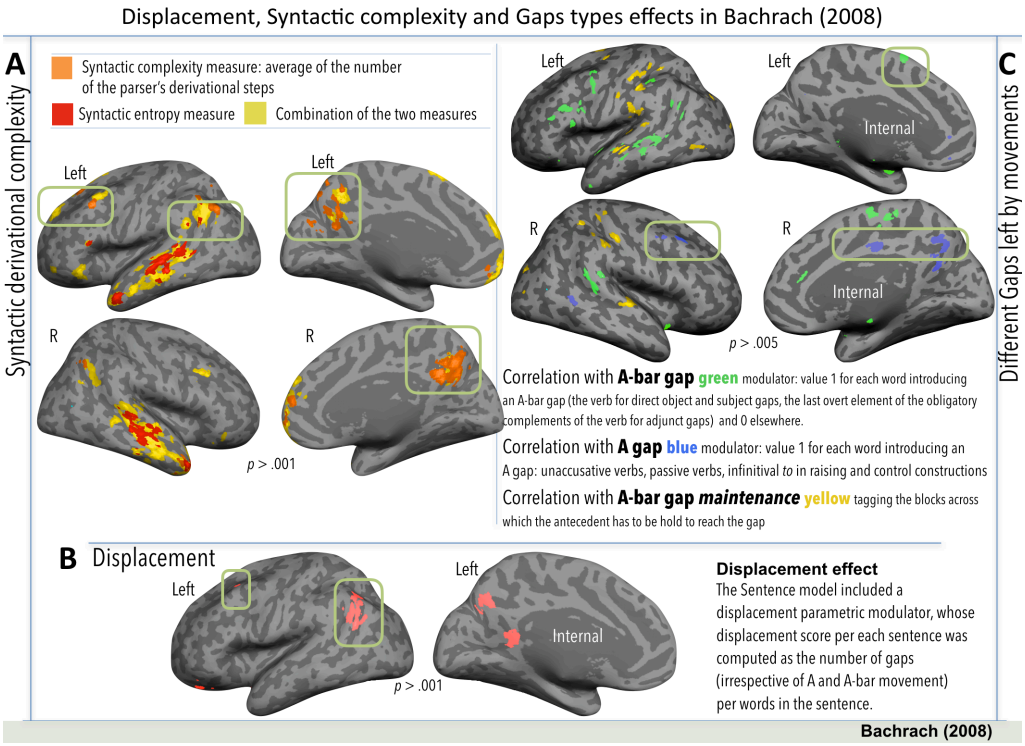


Figure 6.45 – Results from Bachrach's PhD (2008) for Displacement, Syntactic complexity and Gaps types, obtained while participants where reading short narratives. (A) Effect of Syntactic derivational complexity (orange), i.e. average of number of the parser's derivational steps. (B) fMRI Results for three different modulators having a value one when an A-bar gap (green) or A gap (blue) is encountered in the sentence, and for fMRI blocks where the maintenance of the antecedent of the A-bar dependency is to be held before reaching the gap. (C) Displacement effect obtained thanks to a displacement parametric modulator, whose score per each sentence was computed as the number of gaps (irrespective of A and A-bar movement) per words in the sentence. Note that the circled areas are the ones we reproduce in our study across different movement effects.

Following this interpretation, the fact that we observed in Precentral Gyrus an earlier peak for *wh*-Verb-inverted questions (see event-related plots in Figure 6.34, could eventually implicate that the processing of double movement derivation is happening elsewhere, namely in TPOJ as are now arguing for. Given this discussion, the TPOJ seems to be a good candidate for an area that shows an increased activation for additional syntactic movement operations or their derivational 'multi-step' properties.

6.3.4 NP-movement complexity and Local movement effects

Unergatives and Unaccusatives in French

Our results for NP-movement seem to indicate that this local syntactic movement preferentially recruits the aSTS, the MFG and SFG bilaterally, the Precuneus and the Inferior Parietal Lobule. Compared with previous findings for unaccustives in Hebrew (see Table 6.10, p.524), our findings replicate and confirm the involvement of:

1. aSTS, which was repeatedly reported for unaccusative irrespective of the chosen baseline by Shetreet et al. in two studies (2010 and 2012) with a more posterior distribution leaning towards MTG;
2. right Precuneus, that was also reported by Shetreet and colleagues (2012) in the contrast opposing Unaccusative versus Unergatives;
3. left MFG and Left SFG, which had already been reported for Unaccusatives in Shetreet et al. (2010), respectively when comparing Unaccusatives to transitives and to unergatives in Hebrew.
4. left and right Inferior Parietal Lobule (/Supramarginal Gyrus), which was found with a left lateralized distribution in Hebrew by Shetreet et al. (2010).

While these activation patterns reproduce in several aspects of two neuro-imaging studies on Hebrew Unaccusative by Shetreet and colleagues (2010/2012), we did not observe any increased activation in Broca complex IFG both when comparing Unaccusatives to Unergatives and to Transitives, as was expected given previous neuro-imaging results and aphasic linguistic behavior in unaccusatives.

However, when comparing the results for French Unaccusatives to Hebrew ones, it has to be noted that Unaccusatives in Hebrew are obtained by morpho-syntactic marking, we may advance that this particular linguistic means to achieve Unaccusativity (verbal construction) was actually the main determinant of the complexity effect reported by Shetreet et al. (2007 and 2012), and that the activation observed in Broca Complex is possibly reflecting increasing demands linked to the morpho-syntactic marking operation to obtain unaccusative verbs.

Considering the absence of difference with Unergatives, it can be explained by the fact Unergatives do show an increased activation cluster in Broca when compared to Transitives (cf. Fig. 6.22C, p.545). One may speculate that Broca activation is linked to the licensing or the creation of the null object gap in Unergatives.

Reproducing several Local-movement effects

Concerning the opposition between long-distance movement targeting the CP-layer and more local movements, we can conclude that both local movements in our design seem to recruit more anterior temporal (aSTS) and (pre-)frontal pole (MFG) cortical resources, with a more complex response pattern for clitics which additionally elicit more internal areas encompassing the Insula, the SMA and the Precuneus and Cingulate cortex. Comparing our results for more local movements with those from the literature summarized in Figure 6.46, we can underline several commonalities.

Firstly, the increased activation for one-argument Unaccusatives compared to Transitives in the right anterior Insula can be linked to those obtained by Ben-Shachar et al. (2004, see Figure 6.46A) for the so-called *dative shift* (e.g. “John gave the red book to the professor from Oxford” vs. “John gave the professor from Oxford the red book”), a

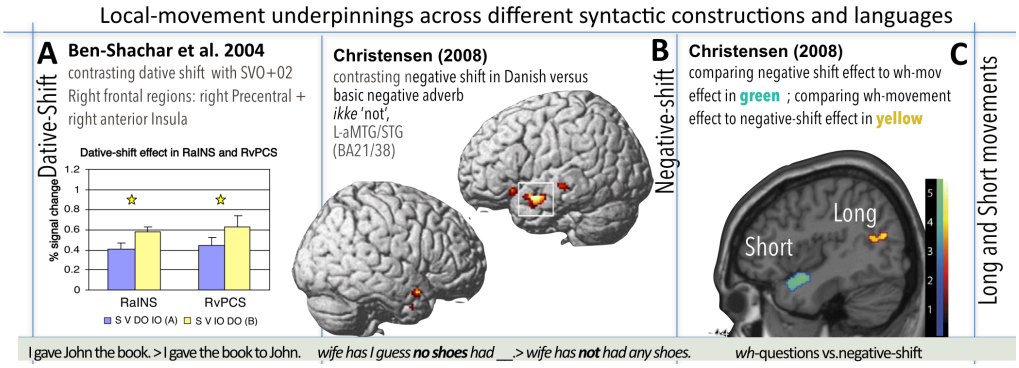


Figure 6.46 – The results of two studies investigating local movement effect (A) Dative-shift effect in Ben-Shachar (2004), (B) Negative-shift effect in Christensen (2008). (C) Comparing local movement targeting IP-layer positions in Negative shift (green) versus CP-layer position in the case of wh-question and inverse contrast in yellow.

displacement that like NP-movement does not target the CP-layer and is concerned with case assignment.

Secondly, the authors reported no increased activation in Broca's area, for this VP-internal alternation, but observed the activation of right anterior Insula (R-aINS) and right ventral Precentral Gyrus (R-vPrCG), which we actually observed in our contrast for NP-movement effect, opposing 2-argument Unaccusatives (declarative and yes/no questions) against matched Transitives (cf. Fig. 6.21, p.544).

Thirdly, the involvement of aSTS in local-movement like negative-shift in Dutch illustrated in Figure 6.46B is also observed in our ROI analyses for clitic-movement (Fig. 6.19D) and NP-movement (Fig. 6.21C), although we do not know how to interpret its involvement. Interestingly, comparing local movements against long ones targetting the CP-layer Christenen (2008, see Figure 6.46C) found again an increased activation of aSTS, which he unfortunately did not discuss.

In this regard, we can further note that subtracting NP-movement to *wh*-movement, or to Verb-movement, yields an activation cluster in the posterior temporal region – actually more anterior than the cluster reported by Christensen in Figure C – which is always absent when to these two movements targeting CP-layer is subtracted the activation to clitic movement (see blue circled regions in Figure 6.23D, p.546).

6.3.5 Question formation and interrogative scope

Discussion about simple yes/no Questions and declaratives: Interrogative operator effects

As argued in chapter 2, y/n-Question generated by simply adding a question mark at the end of an SVO Declarative sentence are considered as featuring an interrogative intonational morpheme (see Rooryck and Chen, 2006) to achieve clause-typing and interrogative interpretation.

The most general contrast [Il méprise ça? *He disdains this?* > Il méprise ça. *He disdains this.*], including all the declarative conditions minus all the yes/no questions conditions yielded activation in the left anterior Insula / Frontal Operculum (FoP).

Although the cluster doesn't resist correction, the more detailed ROIs analyses reveal

a more complex pattern of activation for the yes/no question condition that seems to be driven by consideration of the number of verbal arguments and interrogative scope. Hence if we consider for example the contrast opposing yes/no questions to their matched declaratives limited to three arguments conditions in Figure 6.47 we see that only right hemisphere in a region that is commonly reported to be involved in discourse-level semantics, the right IFG.

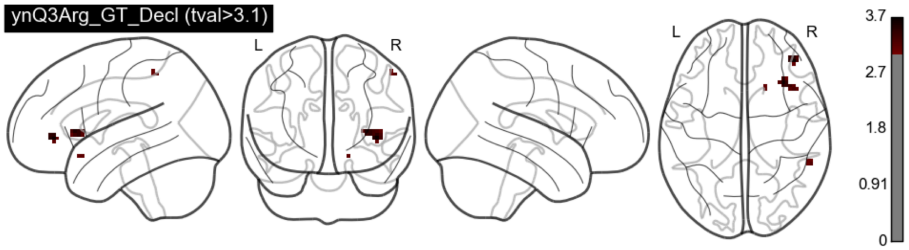


Figure 6.47 – Whole brain contrast opposing 3-argument yes/no questions to their matched declaratives.

An increased activation that we interpret as reflecting the interrogative force interpretation that the question mark is assigning to the Declarative SVO word order or the sentence. The presence of the question mark determines not only clause-type assignment information, but also assignment of illocutory Force interpretation by a Q operator. Namely, the fact y/n Questions compared to Declaratives feature a question scope calculation process, should imply some discourse-level calculation.

As a more detailed consideration of simple yes/no question conditions in our experimental design, we can note that with the increase in number of argument, the number of possible options for the scope of the question increases too. For example in a question like "Il envoit ça là ?" *He puts this here?* the scope of the question marker could be either bear on *this* or on *here*, or on both and even to operate at the level of the whole sentence predication, according to the intonational information. In other words one could utter this question to ask about 'ça' *this* or about 'là' *here* or about the whole sentence according to variations in the prosodic contours. On the contrary when uttering a question like "Il éternue?" *He sneezes?* the interrogative scope of the question ark is more likely to bear on the whole sentence.

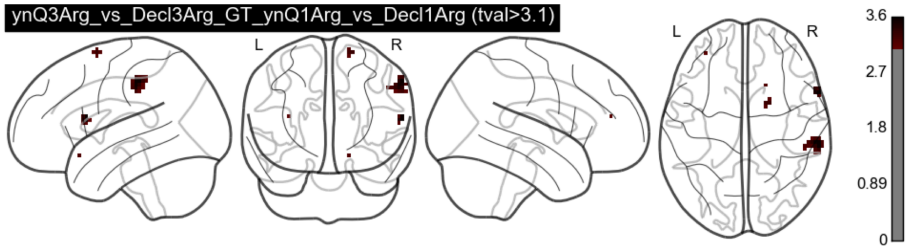


Figure 6.48 – Question scope effect obtained by contrasting the effect of y/n questions in three arguments against that of y/n questions in one argument sentences: [(c24)*Il envoit ça là?* > (c23)*Il envoit ça là.*] and [(c05) *Il eternue?* > (c04) *Il eternue*] .

Give these remarks on the different interrogative scope of y/n questions across number

of arguments, it is interesting to observe that the contrast opposing the effect of y/n questions in three arguments against that of y/n questions in one argument sentences, actually yields right hemisphere activation in as illustrated by Figure 6.48, confirming the involvement of right IFG in the discourse-semantics aspect of interrogative Force, and additionally pointing to an other region, the left Supramarginal/Parietal gyrus, that is commonly acknowledged as being part of the phonological network.

We interpret the additional involvement of the Supramarginal/Parietal region as linked to the fact that the question mark assigns a [+ marked] prosodic contour to the two last strong deictic pronouns 'ça' *this* and 'là' *there*, following studies on prosodic deixis in French in-situ prosodic focus by Lœvenbruck et al. (2005) link in example (257):

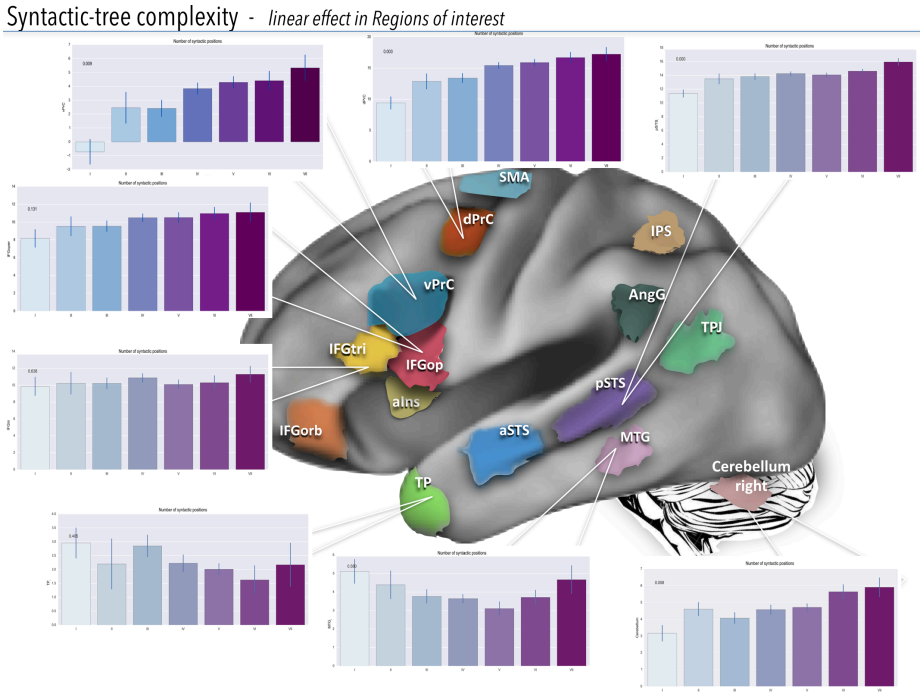
(257) "Madeleine accompanied me"
  a. Baseline condition: *Madeleine m'amena*
  b. Prosodic deixis condition: MADELEINE m'amena

6.3.6 Approaching what is covert in language : the number of traces and overt words in the sentence

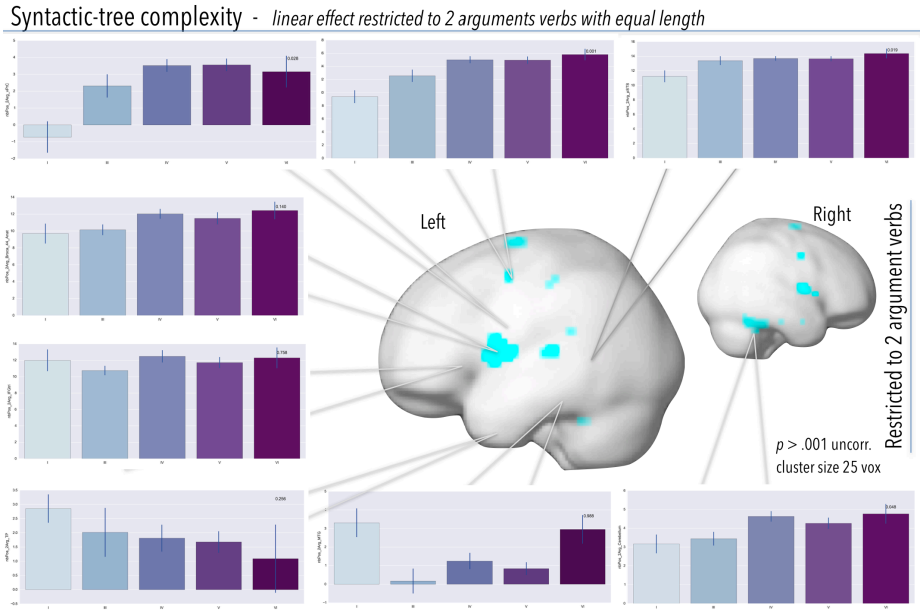
We now change gear and turn to the linear effect we observed both in whole-brain contrast and in ROI analyses for the number of syntactic positions — i.e. the number of syntactic position obtained by summing the number of words in the sentence and the number of empty abstract syntactic elements left by movement.

Table 6.3 – Peaks of activation in Precentral Cortex in fMRI studies on syntactic complexity.

Study	Coordinates	Contrast
Shetreet and Friedmann 2014	[-48 -1 50] Tal.	wh-movement in Topicalization > Declaratives
Ben-Shachar et al. 2004	[-41 11 27] Tal. [-45 8 25] Tal. [44 12 32] Tal. [42 10 27] Tal.	Exp. 1: Topicalization of Direct or Indirect object > Declaratives Exp. 2: embedded wh- questions > embedded yes/no questions
Santi et al. 2010	[-41 10 31] Tal. [48 20 36] Tal.	Adaptation movement and embedding
Santi et al. 2012	[-49 3 37] Tal. [-40 0 39] Tal.	object wh-quest. > subject wh-quest. Pronoun > object wh-quest.
Mack et al. 2012	[-38 -2 54] MNI	Passive > Active
Röder et al. 2002	[-41 6 31] Tal. [-44 3 36] Tal.	Syntax effect
Den Ouden et al. 2012	[-39 6 45] MNI	Object cleft>Subject cleft



(A) Number of syntactic positions across all conditons (NB:corrected for number of characters).



(B) Number of syntactic positions restricted to 2 arguments conditions where no difference in number characters is present.

Figure 6.49 – Number of syntactic positions , i.e. overt nouns in the sentence plus the empty positions left by movement, raging from 0 the noun conditions to 7 syntactic positions in the ROIs. Barplot graphs show the response amplitudes (Beta-coefficients) of I to VII number of positions, and condition I corresponds to isolated Nouns (control condition). Significant linear effect in four ROIs: pSTS, dorsal and ventral Precentral and Cerebellum.

What does this complexity measure mean? Assuming a more representational perspective taking, the size of syntactic-trees being directly related to their number of positions, or terminal nodes, we assumed that this linear increase in activation could reflect a global measure of the size of the syntactic-tree and therefore a precise measure of the complexity of the underlying sentence structure. It should be noted that from a processing perspective, the areas that would witness such a linear increase may encode the syntactic-tree complexity irrespective of the specificity of the different movement types, as all *traces* have in common the fact of being syntagmatic.

We computed the linear contrast in search for areas where activation increases from 2 to 7 syntactic positions, yielding the brain map in 6.50, and further explored the linear effect observed in the ROI analysis, as illustrated in Figure 6.49A. Four main regions are involved: Left Precentral (peak in BA6 expanding to BA9), SMA (SFG/BA6), bilateral Inferior Parietal Sulculs. Note that, when lowering the threshold the Precentral activation goes down and reaches Pars Opercularis.

Interestingly, the first two of these regions have been reported in the literature as involved in syntactic complexity manipulations. The left Precentral Gyrus (BA6) is reported in Shetreet and Friedmann (2014) as being involved in *wh*-movement compared with canonical cord-order (in Hebrew Topicalization (OSVA>SVOA). Although the cluster of activation reported in this study is relatively small the peak coordinate are astonishingly similar: [-48 -1 50], while our peak is [-49 -1 51] (MNI). Additionally, left Precentral Sulcus was also reported by Santi and Grodzinsky (2010) in an adaptation paradigm to syntactic-movement and embedding (IFG/ Inferior Precentral Sulcus with a peak at [-41 10 31]). A similar result was also reported with a more ventral distribution by Ben-Shachar et al. (2004) comparing *wh*-movement and no-movement sentences – Topicalization of the object or Topicalization of the indirect object (dat.) versus Declaratives yielded increased activation of the Left ventral Precentral Sulcus (LvPrCS's peak at [-41 11 27]) (see Table 6.3).

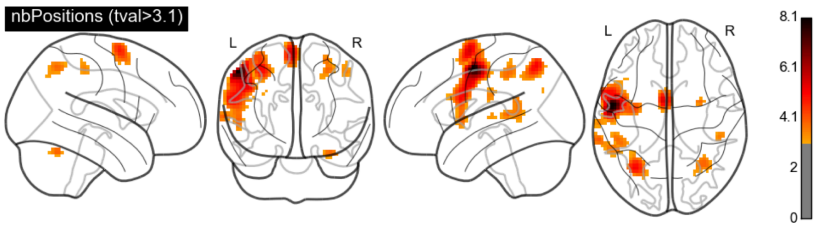


Figure 6.50 – Whole brain linear contrast for number of syntactic positions, i.e. overt nouns in the sentence plus the empty positions left by movement, raging from 0 the noun conditions to 7 syntactic positions.

6.3.6.1 Discussion about the syntactic complexity effect of number of syntactic positions

The first question that arises considering the linear effect of number of syntactic positions is about the role of less discussed peripheral areas like Precentral Gyrus. This area is where we observed the most stable focal activation across different analyses (i.e. correcting or not for number of characters, across different verb classes or restricted to 2-argument verbs, see Fig. 6.33, p. 556). The other area whose linear increased activation was confirmed both by whole-brain at different cluster-extent corrections and in the

linear patterns observed in the ROI (Fig. 6.50, p.577) was the right cerebellum. While we already discussed the possible implication of right Cerebellum in sequential pattern detection, we will here discuss the our result in light of the relatively developed literature reporting dorsal and ventral activation of Precentral cortex in sequence processing across different modalities and cognitive tasks.

The syntactic function of Precentral: Prediction or structural skeleton building?

As the bulk of frontal activation for our movement-related syntactic-tree complexity measure was observed in dorsal Precentral Complex, and not located in Broca Complex (IFG), we engaged in the literature to understand the possible syntactic processes this area appears to be implied in.

At first sight this result might seem surprising, however if we consider as reported above that the involvement of Precentral Gyrus has been convergently indicated by many studies on *wh*-movement, we have a first confirmation of its stable implication in movement-related syntactic complexity. Despite the exceptional reproductibility[37] of its increased activation in studies manipulating syntactic complexity in reading and listening task, its role has mainly remained un-discussed.

Although the functional role of this region has been seemingly left on the side, by reviewing its involvement in different cognitive tasks, we can note that its activation is also observed in several rule-based mental-operations and tasks (Dogil et al., 2002), and that it is this is not uniquely involved in syntactic computation as we reviewed (Ben-Shachar et al., 2003, 2004; Christensen, 2008, 2009; Röder et al., 2002).

A- Mapping sequences onto linguistic structural templates

Research on working memory has commonly associated working memory processes to the Precentral cortex (e.g. Smith and Jonides, 1999; LaBar et al., 1999). Specifically, stimuli and tasks that require to re-code stimulus as structured chunks, have collectively pointed to dorso-lateral prefrontal cortex and inferior parietal cortex, as well as in the premotor cortex in both spatial and non-spatial domain (Bor and Owen, 2007) and more linguistic tasks (Dogil et al., 2002 cf. page 6.3.1.1 , §6.3.1.1).

More recently, ventrally distributed activation of premotor cortex was reported to be activated by the *mapping of sequential input onto structural linguistic templates* (Fiebach and Schubotz, 2006)[38].

37. As proof of this the reader may want to verify this affirmation by just visually check all the brain maps illustrating our neuro-linguistic literature review in chapter 3. It will be a challenge not to have to circle out the left hemisphere Precentral cortex of every figure! For evidence of the lesion of this area in agrammatic population in different languages see Table D.2 in the Annexes (p. 909).

38. These authors organized a special issue of *Cortex* around the idea that Broca's area (BA44/45) and the left ventral premotor cortex (BA6) together with the frontal operculum support different functions during language processing. Note that the functional differentiation between Broca and the ventral premotor cortex is discussed in the context of the neuro-cyto-architectonical differentiation between agranular structure characterizing BA6 and a dysgranular cortical structure characterizing BA44/45. As a side note, the fact that motor and premotor cortex in the dorsal Precentral Gyrus can be engaged also skeletal motor movement is compatible with the fact this also requires rule-governed programming in the sequencing of hierarchical representations (see Schubotz and von Cramon, 2002, 2003). These authors (cf. discussion in chapter 1 , p.1.4.5.2). An interesting discussion on this point and the difference between the role of Pars opercularis versus ventral premotor cortex (vPMC) in sequencing function is found in Bornkessel et al. (2009), the author makes the distinction between sequences that involve

Further evidence in support of the claim that precentral Cortex's activation is linked to the manipulation of rule-governed hierarchical representations comes from a study by Hanakawa and colleagues (2002). This study found premotor cortex to be activated in non-motor tasks that involved *rule-based* non-motor 'mental-operation tasks'. As illustrated in Figure 6.51 a numerical, a verbal, and a spatial task were tested and crucially, none of these tasks involved any motor activity.

The overlap of activations to these three different cognitive rule-based manipulations of patterns of representations in working memory yielded an increased activation in the Pre-SMA, the posterior parietal cortex, the Right Cerebellum and the dorsal Precentral Gyrus ('dMdr' in the Figure).

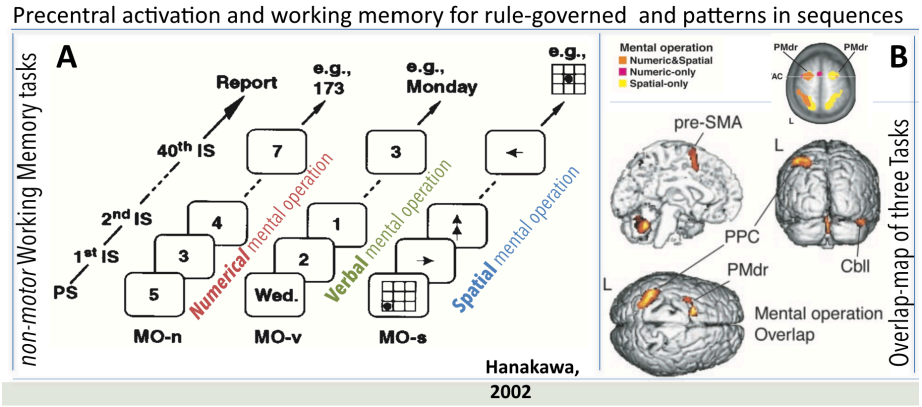


Figure 6.51 – (A) Experimental design from Hanakawa et al. (2002) (B) fMRI results of the overlap-map between the areas exhibiting significant activity during all three mental operations tasks in (A) relative to visual fixation task. PMdr= dorso-lateral premotor Cortex.

Moreover, in a study investigating the link between working memory and syntactic complexity, Prat and Just (2011) tested individuals that had varying working memory capacities and vocabularies. The participants read syntactically simple or complex sentences under conditions of varying extrinsic working memory demands (i.e. (a) sentences alone, (b) preceded by to-be-remembered words or (c) non-words). The result reveal that the areas of the Sentence network showing a grater adaptability to syntactic complexity with increasing extrinsic working memory demands are the prefrontal cortex with a major involvement of dorsal Precentral cortex (i.e. the most extended activation cluster in the no working-memory condition) and the striatum.

A more syntactically-oriented study by Christensen (2010) found Premotor cortex (BA6) increased activation by investigating the cortical activation triggered by syntactic reconstruction in ellipsis and re-analysis of garden-path sentences. Importantly, ellipsis is a syntactic operation that involves neither structural ambiguity nor a change in word-order, but crucially involves the restructuring of syntactic representations in working memory. We follow the author's description of the manipulation in these terms, as just a change in relation between linear-order and hierarchical structure that require to hold in working memory part of the structure representations under analysis.

Importantly, in order to counter the arguments of the tenants of sub-vocal rehearsal

binary hierarchical trees structures and those that don't.

as happening in this area, we stress here that Christensen's results for ellipsis in clausal comparative construction – such as *More people have been to Paris than to Oslo.* versus Pseudo-elliptical clausal comparatives like *More people have been to Paris than Mary has.* – definitely show that while this structure features a way of reducing the load on phonological working memory, they do show increased activation in in BA6[39]. Namely, as argued by the author without ellipsis, these sentences may be 'heavier' or longer, and would therefore require more working memory, just consider "Susan read that crazy book about the pyramids twice last year, and John did [read that crazy book about the pyramids twice last year] too".

Moreover, syntactic priming experiments like Menenti et al.(2011) (cf. §2.3.4, Fig. 2.20 revealed distinct neural networks adapting to the three main linguistic processes involving semantic, lexical, and syntactic information in both production and comprehension modality[40]. Syntactic adaptation effects across modalities was found in only three areas in the left hemisphere, namely Pre-central cortex (BA6), Inferior frontal cortex (Broca/IFG) and Middle Temporal Gyrus (MTG), which were equally strong within and between processing modalities. Importantly, this study was also able to bring to light a network adapting to peripheral low-level processes involved in sentence comprehension, like number of syllables contrasting them between modalities and observed the activation of other areas than those that adapted to syntactic complexity, (cf. 2.20B, p.154).

From what we have said so far, and having put forward its fundamental amodal involvement in sentence processing it is plausible to conclude that this area is involved in structure-dependent computations, such as ***sequential ordering of hierarchical structures in working memory***. Hence, taking into account these results, their interpretations, and considering the results we obtained for the increasing number of syntactic positions our sentences, we have some elements to affirm that when syntactic-movement occurs, this area plays the role of holding in working memory the hierarchical structure of the sentence.

However, what still remains to be accounted for is why the *wh*-movement seems to activate more than other movements this area, as will be revealed by our unsupervised analysis of each conditions across brain regions (ROIs) in Figure 6.64 in page 599 (see also Movement types ROI analysis in Annexes F.2.1, p.936). This leads us to consider further possible interpretation of the functional role of this area, and to refine our argumentation in the following.

B- Syntactic prediction in sentence and Precentral activation

Only recently, a series of neuro-syntax fMRI studies have focused on sentence prediction processes, which crucially revealed the involvement of dorsal and ventral Precentral Gyrus (BA6) in sentence-level prediction of processes. Although it is widely agreed that

39. The actual contrast was ellipsis like *More men have lived in tent than in hotel.* versus *More men have lived in tent than Mary has.*

40. Menenti et al. (2011) run a speech comprehension and production design where subject had either to produce a sentence according to a transitive verb and a subsequent picture presented on a screen (see Figure 2.20A for an example). Pictures presented a color coding for the participants of the action (green, for grammatical subject, and red for grammatical object) that was cuing for the production of passive or active syntactic structure. Alternatively, during comprehension trials, a sentence-picture matching paradigm was used and participants were presented with a photograph in gray color scale and an auditory sentence describing the picture.

linguistic predictions are part of language comprehension, these processes were less studied in the fMRI literature than in the EEG one until recently. The first fMRI study focusing on prediction mechanisms, did it by focusing on this processes in broader cognitive terms, and showed that ventral premotor cortex is sensitive to entropy of stimuli sequences – a measure of the predicatbility of upcoming elements in a more or less ordered sequence.

For instance, Nastase and colleagues presented series of auditory (i.e. pure tones) and visual stimuli (simple colored shapes), which crucially differed in entropy, going from totally random sequences to highly ordered ones. The authors reported the left ventral premotor cortex and SMA (among other regions) to be sensitive to the level of entropy in the series in both modalities (Nastase et al., 2014), these two regions feature a linear increase in the ROI analysis (cf. Figure 6.50).

Importantly, Schubotz and Von Cramon (2004) found the left ventral premotor cortex to be sensitive to predictability across different modalities. In their paradigm, participants were presented with sequences of actions, like someone putting a paper into a post box, and with sequences of abstract shapes, in this study again the left ventral premotor cortex was found to be sensitive to predictability of both types of stimuli. These studies indicate that particularly ventral Precentral is involved in prediction-related mechanisms.

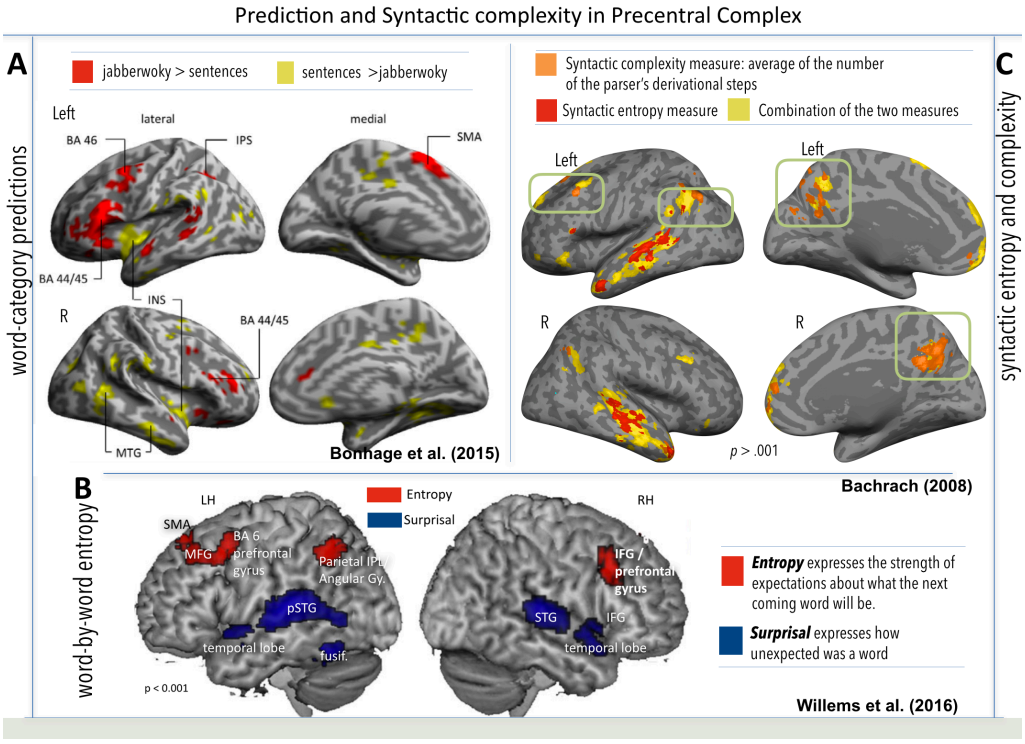


Figure 6.52 – (A) Activation clusters and increased activation peaks during prediction, in Bonhage et al. 2015. (B) syntactic entropy an its combination with Syntactic complexity measure (i.e. average number of parser's steps per sentence) in Bachrach PhD (2008). (C) Willems et al. 2016 Prediction, effect of Entropy in Precentral gyrus.

As for studies investigating the predictability of linguistic elements in the sentence-unit, we can cite Bachrach PhD (2008) and Willems et al. (2016). These two studies correlated brain activity with two different measures of entropy, one calculated as the word by word measure of the strength of expectation of the next coming word in a sentence, and as for syntactic entropy calculated by Bachrach (2008) it was obtained averaging syntactic and lexical surprisal (i.e. a measure of the surprise or the extent of compatibility of an input with prior expectations), as well as word frequency per each sentence. The results are quite different as illustrated in Figure 6.52B and C, but both find dorsal Precentral activation, and Willems et al. also reported SMA increased activation for more predictable sequences.

Notably, the fact that Willems et al.'s results show dorsal Precentral increased activation in a word-by-word measure of syntactic predictability is actually compatible with the working-memory account presented above, thus showing that the syntactic patterns that are possibly held in working memory in this area are incrementally updated. Another result from one of these two studies brings us back to attribute to this area a role in the encoding of syntactic-tree complexity. The dorsal Precentral cluster in Bachrach's study shows a combined effect of syntactic entropy measure and a syntactic complexity metric consisting in the averaged number of parser steps per sentence.

Moreover, the dorsal part of premotor cortex is not only systematically present in fMRI results from all kinds sentence localizer tasks and syntactic movement-related complexity manipulations, but also present in the lately developing intra-cranial studies investigating the neural implementation of syntactic information in the sentence (see also Nelson et al., 2017) like Fedorenko et al. (2016). This last study used the traditional paradigm comparing non-words' lists, jabberwocky (i.e. de-lexicalized sentences) and real sentences, and identified the Precentral Gyrus among the main hot spots showing both a significant [Sentence versus > Noun] effect, and a monotonic increase across word positions in the sentence condition[41].

Another recent fMRI study investigated sentential prediction in the same linguistic stimuli —Jabberwocky sentences and Normal sentences— and designed a 'Predictive eye-gaze reading task' combining eye tracking and and fMRI recording in order to be able to track syntactic and semantic aspects of linguistic prediction via anticipatory eye-movements. The two types of stimuli Bonhage et al. (2015) used fundamentally differ according to the type of linguistic prediction they allow: (i) Jabberwocky condition, where only function works are available, maximizes the mechanisms of predicting syntactic categories in a sentence, and therefore probably also the building of the sentence's functional structural skeleton in absence of lexical semantic information, (ii) whereas regular sentences allow readers to perform predictions at each specific lexical or function word.

As both regular and jabberwocky sentences share the important property of building syntactic phrase structure based on the availability of function words as well as morpho-syntactic cues, their conjunction should reveal the brain areas that are involved in prediction of the word-category of the target word based on the on these linguistic cues. As illustrated in Figure 6.52A, the results from fMRI recording, show that the conjunction analysis for word-category prediction [Sentences > Non-Words-Lists & Jabberwocky > Non-Words-Lists], reveals a large clusters of increased activation in areas we are now familiar with: bilateral Precentral gyrus (coordinates: [-63 -13 7]), Medial

41. See in Figure 1.14A the blue circled in white electrodes (p. 64).

temporal lobe and basal ganglia, with the additional implication of right Cerebellum (coordinates: [15 79 32]).

Hence, we can conclude from this study that Premotor cortex, thalamus, caudate nucleus, hippocampus, and cerebellum, can be interpreted as playing a crucial role for word category-based linguistic predictions and the consequent syntactic phrase structure building.

All in all, these studies in multiple cognitive domains indicate a positive relationship with predictability and activation in Precentral cortex, SMA and Cerebellum, in other words when predictability of the upcoming word was high (i.e. entropy is low), activation in these areas was also high.

Given these findings indicating the role of both dorsal and ventral Precentral cortex in sentential prediction mechanisms, and going back to the predominance of *wh*-movement in this effect, we could advance that among the different movements we included in our stimuli *wh*-movement implies a stronger prediction cost given its long-distance property: the process of creating an object up-stream to assign the verb's argumental role, later in the sequence, as was indeed repeatedly demonstrated by Cross-modal Lexical priming paradigms we reviewed. Note Verb-movement showed a different priming pattern in Dutch (de Goede et al., 2009)[42].

Alternatively, one of the possible hypotheses we can emit is that Precentral (BA6) activation may be functionally mapped into the linguistic *Search process* of scanning the syntactic-tree to go and search for the candidates for Internal Merge (i.e. one of the formal steps to obtain movement derivation, cf. definition in p.207).

Specifically, looking into the sentence for candidates to find an element bearing its same features (as presented in chapter 2 (§2.27 p.167) could be understood as a composite process where "big brother eye" would be screening over the whole sentence's trees structure constructed so far, to find a suitable candidate to perform the internal move step. While the big-brother role could be played by Broca's area the working memory buffer where the syntactic-tree structure would be hold to perform the *Search* operation could be played by the Precentral cortex.

Alternatively, in a simplified understanding of syntactic movement, a predictive mechanism could be hypothesized to posit a gap in the structure[43]. This issue on the predictive mechanisms linked to the detection of displaced/pre-posed elements in the sentence leads us to resume to a phenomenon obliquely addressed in our discussion about the involvement of right cerebellum in *wh*-movement, a particular sort of linguistic prediction process is engaged during the comprehension of sentence that have undergone syntactic movement, and more concretely in our *wh*-questions by encountering the sentence-initial wh-object one has to posit an up-stream object.

42. The priming effect for related probes was observed at all tested positions, including the intermediate position, instead of being only present at the related gap position as it is the case in wh-movement(de Goede et al., 2009).

43. Theoretically finer hypothesis and differences between wh-movement and Verb-movement could actually be the following, according to these steps :

- Verb-movement in sentences like "As-tu vu?" *Have-you seen.* can imply a process of creating a configuration and a null element
- *wh*-movement in "Qui tu as vu? ", implies an other process linked to Q-criterion that creates the object up-stream in the Comp.
- wh-verb-inverted (whQinv) sentence like "Qui as-tu vu?" involve Spec-Head wh- and local configuration.

Alternative views on Precentral Complex and syntactic-tree complexity

The results from this wide range of neuro-imaging studies point to attribute to this region a central role in *amodal* syntactic structure representation and building, and collectively support our claim of attributing it a central role in holding syntactic structure in working memory. However, such a crucial function at the cross-road between sentence structure building, prediction (both sentence-level and word-by-word) and working-memory, could hardly be supported by an area that would be peripheral one in the language network. For this reason and to further ground our claim, we searched for studies investigating its connectivity.

Interestingly, the results of a tractography and parcellation study by Raetting (PhD, 2010) aliment our discussion and shows that this area is highly connected by the fiber tracts coming from three temporal areas. This study on the connectivity of anterior, middle and Posterior STG/STS demonstrates that the lager proportion of fibers of these three temporal language areas do target the ventral portion of Precentral cortex (BA6), as illustrated in Figure 6.53, thus giving to Precentral complex a non-peripheral status.

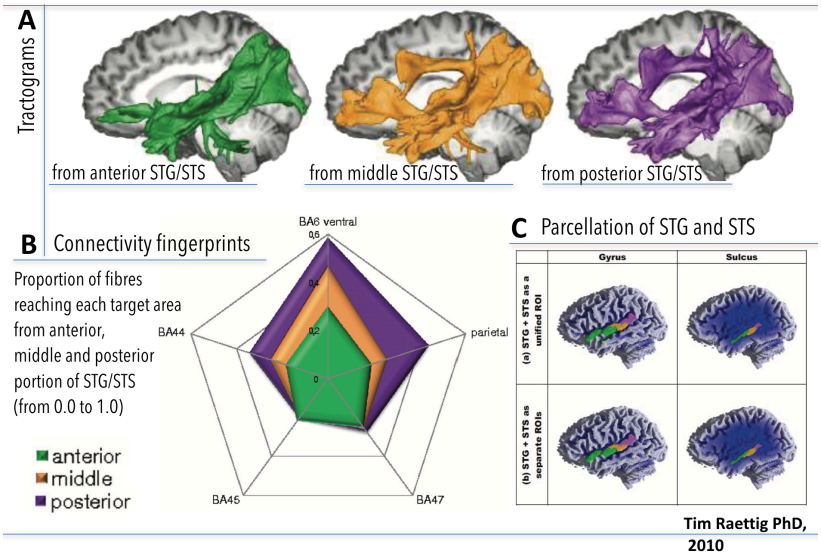


Figure 6.53 – (A) Whole-brain tractigrams for anterior, middle and posterior subregions discovered in the Parcellation in (C). (B) Connectivity fingerprints, the proportion of fibres reaching each target area from anterior, middle and posterior portion of STG/STS (from 0.0 to 1.0) (C) Anterior, middle and posterior subregions as uncovered by computing the parcellation of STG and STS as a uniform ROI for the averaged DTI data. Adapted from Tim Raettig PhD (2010, from figure 5.4 p. 55-56).

All in all, from this lengthy but, we hope, still enjoyably demonstration, one of the possible conclusion to which this discussion could lead are the following three alternatives:

1. housing the structural prediction process of positing gaps in coordination with Cerebellum whose role would be a more low-level one of detecting detecting sentential patterns
2. involvement in the screening process of the sentence structure that the Search mechanism may require for selecting candidates for Internal Merge

3. reflecting the increased syntactic working-memory load of holding hierarchical patterns that are also including the higher syntactic layers where the wh-objects have been moved to (i.e the height of the sentence's syntactic-tree).

Yet, whatever option one would choose, the three remain actually compatible with the working memory account proposed by the reviewed literature, hence we could momentarily conclude by speculating that Precentral cortex constantly keeps in working memory the sentence's abstract structural patterns (i.e. the syntactic-tree structure), while Broca's area, in its three different sub-parts, orchestrates the different aspects of their incremental building.

6.4 Other complexity measures in movement-derived sentences

6.4.1 Verbal Argument structure and movement

Addressing the issue of Verb argument structure processing and its complexity, we can first say that increasing argument structure complexity implies different dimensions, the literature has emphasized the role of verbal lexical information and its associated syntactic information:

- subcategorization frames, the number of different syntactic phrase types e.g. verbs that take complements introduced by "that/for");
- the number of arguments (i.e. verb argument structure or event structure: the participant roles entailed within the verb's representation);
- and their varying thematic-roles (e.g., agent, patient/theme) that are its thematic grid, i.e. the set of arguments associated with the verb in terms of their possible thematic roles.

6.4.1.1 Number of verbal arguments: Discussion

As we varied the number of arguments so that more elements should actually undergo syntactic movement – notably our 3-argument conditions and our 2-argument unaccusative conditions. For this reason, verbs were selected when their thematic grid could be enriched by an additional locative argument, that was optional but categorized as an essential argument being structurally part of the thematic grid offered by the valence dictionary we build our stimuli on (cf. §6.2.1.2, p. 529). Thus, these verbs may show a complexity effect linked to their number of thematic options compared for example to simple transitives.

Hence, performing a posteriori a contrast between 3-argument (Transitive + locative) and simple transitives, we expected to find increased complexity mainly in angular gyrus, a ROI we had selected to observe number of arguments related increase, and some of the areas that have been reported to be linked to the number of complements or thematic options in verb processing as summarized in Figure 6.54.

Several neuro-imaging studies have found that neural activation in posterior brain regions – including the left posterior superior temporal sulcus, supramarginal and angular gyrus – are associated with verb argument structure processing, but not only, MFG and other temporal regions are also stably reported (see Table 6.12 for a summary of experimental studies, p.531).

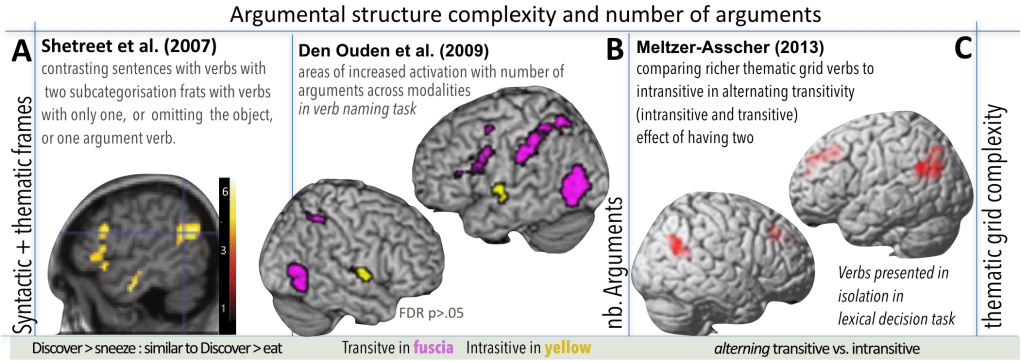


Figure 6.54 – Three representative studies revealing the neural correlates of different aspects of verb argument structure (A) Subcategorization frames: Shetreet et al. (2007). (B) Number of arguments Den Ouden et al. (2009). (C) Thematic grid complexity: Meltzer-Asscher et al. (2013).

Specifically, Shetreet et al. (2007) compared verbs like discover and loose in 'discover the truth' / 'discover that he is here' versus 'loose the keys', and observed related complexity effects in a predominantly left lateralized circuit, encompassing Angular/Supramarginal Gyrus and anterior cingulate, left MFG (BA6) and SFG (BA9), ITG and MTG,and IFG, as illustrated in Figure A. These results were confirmed by a later study showing that thematic options and subategorization options do show a graded effect in STG and BA9 (Shetreet et al., 2010). Similar results were found by Meltzer-Aascher et al. (2013) (see Figure C) for alternating transitive verbs like 'melt' (i.e. the ice melt, or 'she melt the chocolate') compared with unergative intransitive verbs, like 'sleep'. These findings indicate that the MFG bilaterally and the posterior peri-sylvian region are crucially engaged during processing of information related to verb argument structure and its thematic grid, showing that the listeners encountering the verbal root '*break*' will actually engage in building two syntactic structures, one for the intransitive verb *break* and another for the transitive one.

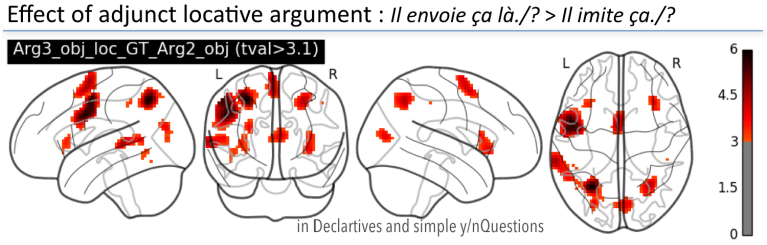


Figure 6.55 – Whole-brain contrast comparing two complement [agent+ theme + location] verbs in our 3-argument verb class to simple transitive verbs[agent+ theme] in declaratives en y/n questions.

The pattern of responses illustrated in Figure 6.54 B is particularly interesting for us because it narrowly mirrors the result we obtained for the effect of adding a locative argument. This multi-modal visual and auditory study by Den Ouden et al. (2009) replicates earlier studies examining verb argument structure complexity (Thompson et

al., 2007/2009)[44], and also found increased activation in Precuneus, which was bilaterally activated, in processing of the number of verb complements in Shetreet et al. (2007). It generally confirms the involvement of the following areas in argumental structure complexity:

1. bilateral posterior middle temporal and angular and supramarginal gyrus (BA39/40),
2. bilateral parietal superior activation (BA7),
3. bilateral precuneus,
4. bilateral activation in inferior occipital gyrus/ fusiform gyrus (BA37), middle occipital gyrus (BA19),
5. left precentral gyri (BA 6),
6. left middle frontal gyrus (BAs 6, 9, 46) and
7. left in the triangular and opercular parts of the inferior frontal gyrus (BAs 44 and 45).

As we can observe in Figure 6.55, the above pattern of responses actually directly mirrors the effect we found contrasting 3-argument against 2-argument transitives (in declaratives and y/n questions), except that we additionally observed the involvement of the SMA.

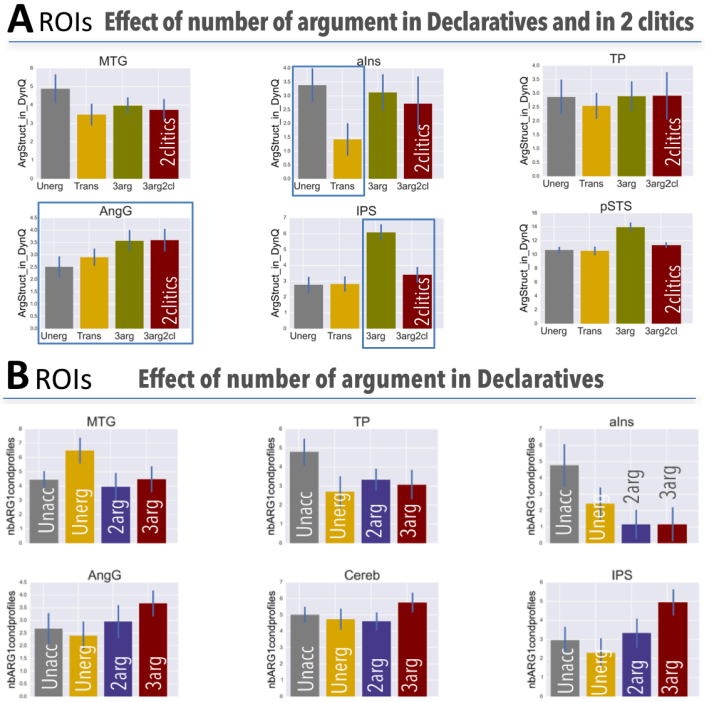


Figure 6.56 – (B) Barplots of Beta-averages for Declaratives having different numbers and types of arguments and (A) for three argument declaratives having two clitics in a selection of apriori ROIs showing different response patterns..

44. Den Ouden namely reproduces results obtained in healthy participants using a lexical decision task (Thompson et al. ,2007) where a graded activation in left supramarginal and angular gyri as a function of the number of arguments selected by the verb was found (e.g. two versus one argument verbs: "follow" vs. "whistle"), whereas bilateral activation in these regions was found for processing three-argument verbs like "send" versus one-argument verbs. not that this patter seems extremely stable and was replicated by Thompson et al. (2009) too.

As for the ROI analysis for a graded effect of number of arguments, as expected, we observe only a gradual increase in number of arguments in Angular Gyrus ROI and not in Broca complex, which is indeed confirmed by our results (cf. Fig 6.56A).

The fact we did not observe a graded effect in Broca confirms neuro-psychological findings showing that verb processing in Broca's aphasics is preserved and they can detect verb argument structure violations (e.g. Kim and Thompson, 2004). This result is also in line with Bonakdarpour et al. (2008), who tested argument structure processing and found a positive correlation between wider lesions in Angular and Supramarginal gyri and the ability to process argument structure.

Moreover, in Figure 6.56B we observe a increased activation in TP and Insula for Unaccusative verbs and MTG shows a different pattern where Unergative are more active. Importantly, as illustrated in Figure 6.56A, Angular Gyrus shows indeed an effect of increased activation to the number of argument irrespective of their status as clitics.

6.4.1.2 Interaction between Verb-movement and number of arguments

By observing the difference between Verb-movement and wh-movement main effects with both baselines (declarative and y/n question) across the different verb classes present in our experimental design, we noted in previous section §6.3.3 the presence of a possible interaction between Verb-movement and number of arguments as summarized by Figure 6.57. We can namely observe in Figure 6.58A that General contrast (i.e. encompassing 3 verbal classes) for verb-movement effect shows broader recruitment of the bilateral language areas that Minimal contrast irrespective of the subtracted baseline. Yet, we observe the exact opposite for *wh*-movement effects in Fig. B, with additional constant activation when subtracting declarative baseline. Note that it is namely for this reason that when directly comparing Verb- and *wh*-movement to reveal their specific neural signature we decided to compare them in the three argument class.

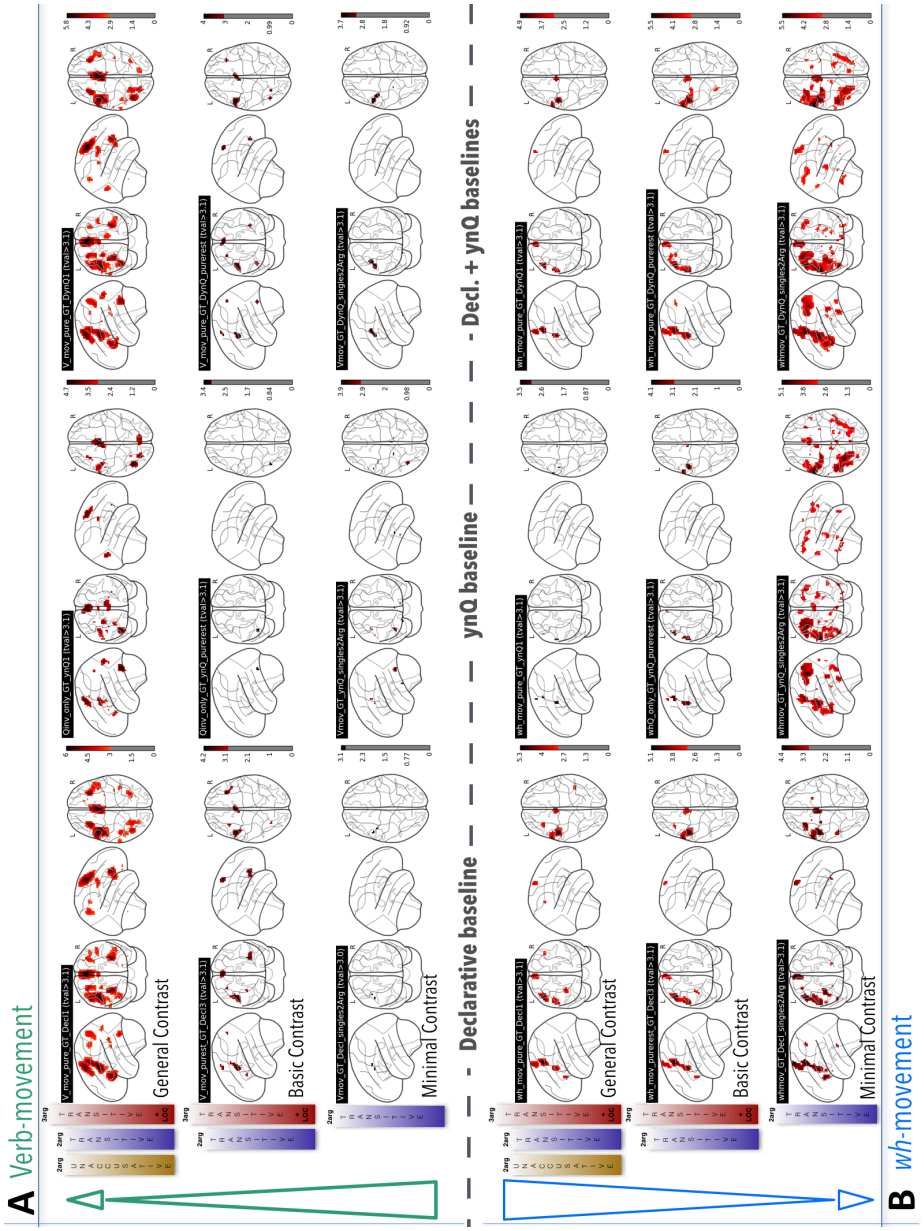


Figure 6.57 – (A) Verb-movement and (B) wh-movement effects across different number of arguments and different base-lines. These patterns are globally indicating an interaction between Verb-movement and number of arguments.

To confirm this interaction, we performed a post-hoc whole-brain contrast featuring the following difference between Verb-movement and *wh*-movement in three different argument classes and the same contrast, but only in Transitives. Figure 6.58 shows the result masked inclusively for the positive effect of Verb-movement when compared to declarative and simple y/n question baseline.

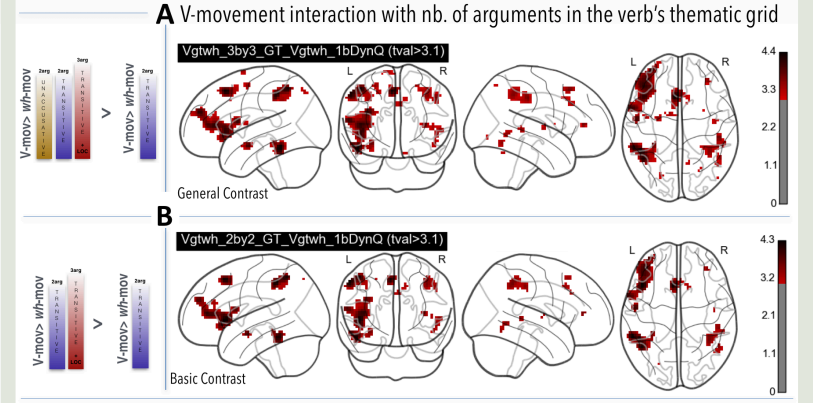


Figure 6.58 – Interaction contrast between Verb-movement and number of arguments at whole-brain level. (A) General contrast: [(V-mov > *wh*-mov)in the 3-argument classes – (V-mov > *wh*-mov) only in Transitive] (B) Basic contrast [(V-mov > *wh*-mov) in 2 and 3 arg – (V-mov > *wh*-mov) only in Transitives].

The interaction effect under consideration is observed in a large part of the sentence network and in some right hemisphere clusters, but crucially avoiding Angular Gyrus which is allegedly acknowledge as an area involved in number of arguments. Instead, we observe a bilateral activation of the inferior parietal lobule (IPL) which was repeatedly reported (e.g. Shetreet et al., 2010) as involved in contrast where verbs with additional complements. Moreover, vPRC and pSTS, two areas that we systematically saw in wh-movement effect in the literature, are not involved in this interaction effect.

Notably, the temporal activations in this interaction (Fig. 6.58) are focally involving only a small cluster in the aSTS and in MTG/ITG area that are often reported in studies on argumental complexity linked to the number of thematic options and the number of complements (Shetreet et al., 2007; Den Ouden et al., 2009), as illustrated in previous section (cf. Figure 6.54).

As for frontal clusters, we can note that MFG was reported by Meltzer-Asscher et al. (2013) to be involved in verbs having richer thematic grids and that Broca and more particularity pars opercularis (BA47) was reported by Shetreet and colleagues (2007/2010) as being elicited by denser thematic grids.

A possible syntactic interpretation for this interaction between Verb-movement and number of argument would consider the presence of the so-called VP-shells in the case where our verbs display an additional locative argument in their thematic grid. It is namely the case that the verbs we selected for the 3-argument verbs class are double-complement verbs like "envoyer" *send*, "adresser" *address* or "mettre" *put*. These verbs are associated with the argument array agent, location, theme and crucially feature an additional dimension of complexity in the structure of their VP-layer: they feature a

VP-shell like distransitive do[45].

As illustrated in Figure 6.59A, the presence of an essential locative argument, like in "Envoie-t-il ça là?" (*Do you send that there?*), implies an additional node in the VP structure, a kind of double VP structure, thus, yielding a more complex structure than the one where the only complement is the direct object in 6.59B.

Depending on the linguistic analysis, this type of double structure is either found under the *light verb* (vP) or is analyzed as a small clause (Rooryck p.c.). Hence, the supplementary complexity of the VP-shell hosting an additional complement is linked to its configuration, where the main verb (V) moves to the "little v"(i.e. light verb) that assigns the Agent theta-role, while the lower VP assigns the theta-roles to the OBJ and the PP/loc[46].

Hence, we can speculate that the presence of more heads at the level of the VP, or the movement of the verb to v', are likely to yield additional complexity to Verb-movement operation.

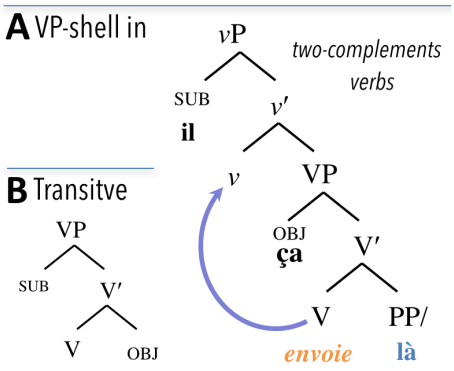


Figure 6.59 – Syntactic-tree diagram of the VP shell of double complement verbs in our 3-argument conditions, e.g. "Il envoie ça là." *He sends this there.*.

I

To conclude, while the minimal contrasts in the whole-brain and ROI analysis revealed relatively similar response profiles (§6.2.4.1), this interaction analysis -irrespective of the linguistic interpretation we gave- actually points to a fundamental processing difference between *wh*-movement and Verb-movement related to the nature of the moved element. Crucially, the complexity linked to the lexical properties of verbs, like their complementation options and the number of arguments in the verb thematic grid yields a greater recruitment of cortical resources in Verb movement compared to wh-movement

6.4.1.3 Argumental structure linearization

Another important aspect of sentence processing pertains to verb argument structure processing, considered as an interface between the semantics of Verbs (i.e. *who did what to whom*) and syntax, which can be manifested in word-order linearization.

Thematic reanalysis has been the focus of several studies we saw in chapter 2 (§2.2.2) (e.g. Grewe et al. 2002) and considerations about word-order of verbal arguments have been the focus of many studies in German, going from agent-first processing strategies to agrammatic theories linked to the linear assignment of roles in presence of sentence-comprehension deficits. These approaches to sentence comprehension and its complexity predict a crucial role for the surface linearization of arguments.

Given the rich panel of diverse word-orders we have in our experimental stimuli, we decided to temporary abstract away from movement-related understanding of our stimuli

45. One should of course not make the parallel between recipient and location too straight forward and say that double-complement verbs like in our 3-argument sentences are the same as double-object distransitive structures, the only thing we say here is that double-complement verb like *send* or *put* are associated with the argument array agent-location-theme and feature a VP-shell.

46. As side note for those interested in the issue of binary branching trees, this proposition about VP-shells was initially done by Larson (1988) to account for distransitve verbs in binary branching trees.

to try and check a possible analysis of the complexity of the experimental sentences in terms of "argumental structure linearization".

In this way, we will very naively consider only the order of argumental roles and classify sentences only according to their distance from canonical SVO linearization. As shown in the box below, this leads to abstract away also from the fact that certain argumental roles are actually instantiated by *wh*-words[47]:

SVO	>	OVS (whQinv)	>	OVS (clV)
c07 Elle réside ici.		c09 Où réside-t-elle?		c13 Y réside-t-elle ?
c15 Il méprise ça.		c18 Qui méprise-t-il?		c21 Le méprise-t-il?
c23 Tu colles ça là.		c26 Où colles-tu ça?		c29 Le colles-tu là?

The question of the linearization of arguments could be addressed by looking to several aspects of argumental word-order[48]. For instance, looking at what syntactic position (pre-/post-verbal) receives what argumental-role (i.e. theta-role), would imply considering Verb-moved question as displaying VSO word-order, wh-moved ones OSV. Thus Clitic + Verb-movement sentences would feature the surface configuration OVS compared to Declaratives showing the SVO canonical Argumental assignment configuration.

Three arguments conditions give us the opportunity to observe sentences with different word-order configurations. The very naive hypothesis was the following: the more the linear word-order of the sentence diverges from the canonical one the more complexity effect in certain areas should increase. More specifically an area in the brain that is responsive to this dimension of sentence complexity would therefore be a candidate for an increased activation for the complexity linked to argumental structure linearization.

Alternatively, a simplified movement-related understanding of these linearization patterns would explain them as being more complex when a "non-local movement" of arguments takes place, that is when a movement of an argument over another argument of the same verb occur[49].

An analysis restricted to three arguments' conditions (subject + direct object + locative) offered us an unexpected observation. We were able to identify three different activation patterns across ROIs: Angular Gyrus, Precentral Gyrus and anterior Insula show increasing activation according to three different factors. Interestingly, Angular Gyrus and anterior Insula show a mirror pattern activation which we will try to understand.

47. In this perspective it would be actually interesting to study a language like Romanian, which gives the occasion to observe movements that do not determine any different argumental linearization. For instance, the question *who what does?* "Cine ce face?" is licit while *what who does?* "*Ce cine face?" is not (Rizzi p.c.).

48. These are different complexity measures that can be taken into account when considering the linearization aspect of the argumental structure of a sentence having undergone different syntactic movements:

- Relative order between Direct object and Indirect object
- Verb-initial complexity
- Contiguity or distance between Verb and its complements
- Number of distortions from the canonical argumental: (i) Cl. + V > Le vois-tu? = 2 linearization distortions, otherwise Clitic could move with the verb; (ii) V only > Vois-tu ça ? = 1 linearization distortion linearization.

49. As proposed by Friedmann and Shapiro (2003, footnote 4) when discussing patterns of impairment in agrammatic patients in Hebrew.

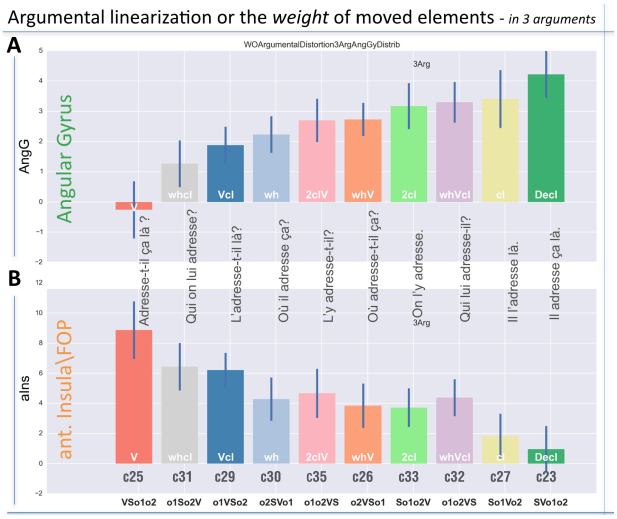


Figure 6.60 – Barplot of the ten conditions having three arguments in our experimental design. Each bar represents the average activation (beta) for subject-specific ROI analyses, in Angular Gyrus and anterior Insula/FOP.

As shown in Figure 6.60, Angular Gyrus shows the lowest activation for Verb-movement (e.g. "Adresse-t-il ça là?", *Addresses-he that there?*), and an increasing level of activation the more the sentence mirrors the canonical surface word-order, while the anterior Insula shows the exact opposite pattern. Notably, pSTS is nearly equally activated by these conditions and seems not to be involved in this complexity measure. Finally, Precentral Gyrus shows a different ordering pattern that seems to be more sensitive to long distance movements targeting higher landing sites in the CP-layer.

According to the literature on argumental structure, number of verbal arguments and semantic compositionnality, Angular Gyrus was consistently reported to be more activated by greater number of arguments (Thompson et al. 2007, 2009 and 2010), and by more frequent and canonical semantic combination (Price et al. 2015). However, this can only explain one of the two patterns we observe here, namely the one present in Angular Gyrus.

To explain the activation pattern in the Insula is showing we have to go back to movement related complexity conditions that have an increased activation are those presenting a greater number of arguments having moved over the subject. However, this alternative hypothesis considering that an object moving over a subject is more cost-full, cannot totally explain why both conditions featuring two sentence-initial objects (c33 and c35) activate less than condition (c31) showing a [O_1SO_2V] pattern. This cannot explain why (c33 and c35) activate less than Verb-movement only (c25) or why the condition showing three movements displacing the two object and the verb, yielding a sentence-final subject, is in the middle of the slope.

Yet, a possible interpretation or the observed pattern would be to understand it as linked to the weight of the moved element: the verb being the heaviest and a simple clitic the lightest.

Hence, if we consider the insular pattern, we can indeed observe that the combination of clit and Verb-movement are always more active. This can linguistically be motivated

by the fact that in clitic-movement the last step implies a incorporation to the verb which crucially yields a configuration where the moved object is a larger unit and thus "heavier". Notably, wh- and Verb-movement activate less the anterior Insula (FOP) than clitic and Verb-movement although the final surface word-order is a very similar one (i.e. OVSO). This further corroborates that the clitic incorporation to the verb creates a heavier object to displace. To this should be added that two clitics appear indeed to be heavier objects than one clitic, and wh-object and clitic movement are heavier than a simple wh-object movement.

We can conclude that the Insula could actually be activated by the movement of heavier syntactic elements, and that Angular Gyrus is showing a mirror increased activation for configurations where the word-order is more similar to the canonical one.

This could generally mean that Angular Gyrus performs a retrieval process of thematic configurations more than a building process, which would be in-line with the results pointing to this area as being more activated by more "lexicalized combinations" (e.g. Price et al., 2015, e.g. "plaid jacket" > "moss pony").

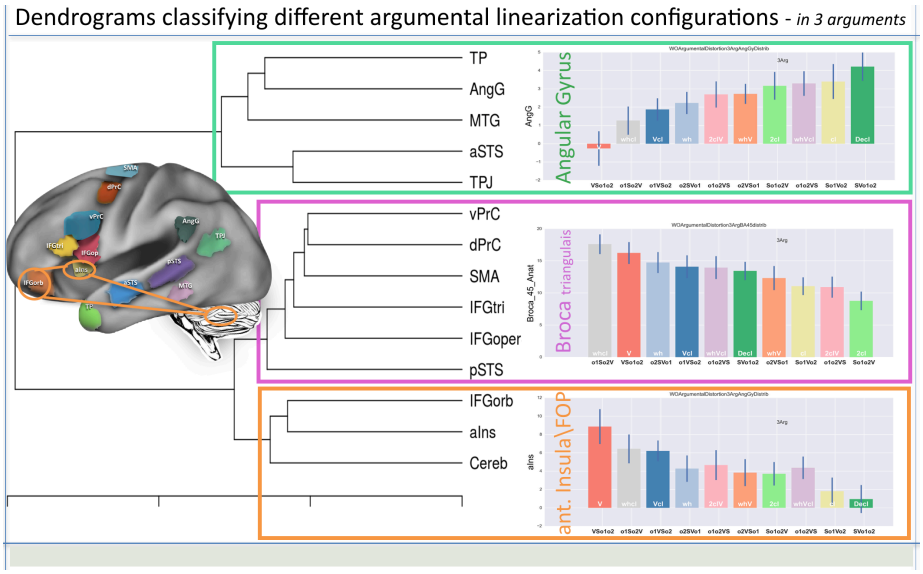


Figure 6.61 – Dendrogram of clustering classification based on the correlation between activation pattern across conditions in each ROIs for three arguments conditions.

To test for the generalization of this post-hoc observation, we applied an unsupervised clustering analysis (see Methods section and next section §6.2.3.3). The clustering classification was based on the similarity of patterns of activation found in each ROIs confirms the above observation.

As illustrated in the dendrogram in Figure 6.61, a tripartite division of labor is observed. The Insula pattern is reproduced also in IFG pars orbitalis and in the Cerebellum. Angular Gyrus pattern is shared with greater variability by other temporal regions and Inferior frontal regions. Yet, a totally different linear ordering is observed in the pSTS-frontal cluster. These regions appear to be driven by other factors: simple declaratives are found in the middle on the slope (see Figure 6.61B) and two clitic sentences are the least activated.

6.4.2 A possible Intervention effect

Further observations linked to the different set of arguments that are found between the 2-argument group and the 3-argument one lead us to consider another parameter of syntactic complexity to explain a unforcasted complexity effect we observed when comparing wh-questions with an object wh-word to those featuring a locative one.

Formalized in the frame of the generative approach to linguistic theory, the *principled* account of movement-related complexity effects offered by Relativised Minimality (Rizzi, 1990 and 2004; Starke, 2001), has already provided insightful explanation to agrammatic sentence comprehension patterns in Broca's aphasics (Grillo, 2009) and child grammar (Friedmann et al. , 2009 and Hu Shen'ai, 2015 for Chinese). We already introduced the different calculations and consequent behavioral complexity effect yielded by sentences whose syntactic configuration implies a featural configurations linked to intervention effects (see §3.4.4.4, p.399, reproduced here for convenience in 6.62). Notably, a particular set of syntactic complexity effects can be captured by the locality principles expressed by the RM approach, particularly interpreting the RM principle in terms of features.

As illustrate in Figure 6.62, in this framework, movement is allowed only when specific configurations of featural specification are met. The features characterizing the moved item and other possible items found in-between the extraction-site and the landing-site (the so-called *interveners* in orange) can be in an (i) identity, (ii) a subset relation or (iii) be distinct.

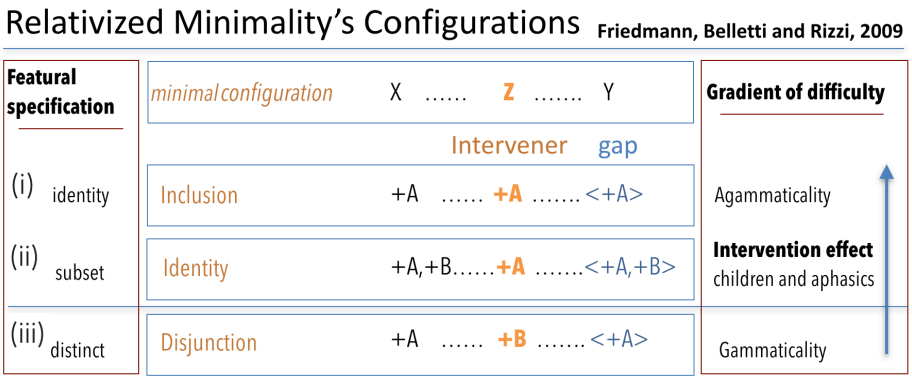


Figure 6.62 – A and B stand for abstract morpho-syntactic features triggering the syntactic movement of an item Z. According to the RM principle, a local relation between X and Y cannot be established if Z, has the same feature as Y, and acts therefore as a potential candidate for the same relation. Adapted from Fiedmann et al. (2009).

In short, the fundamental idea of applying these locality principles to understand linguistic behavior (in normal or impaired populations) is that (i) the *representation* of the full array of morpho-syntactic features is needed to then be able (ii) to perform the task of distinguishing for example a moved-object from the intervening-subject.

Many investigations into movement processing have been based on the object vs. subject movement asymmetry especially in relative clauses (see Table Annexes studies on object versus subject relatives) and lately in object versus subject cleft (e.g. Den Ouden, 2012). In fact, object movement (unlike subject displacement) presents a configuration where lexical material is intervening between the phonologically realized position where

the object lands and the silent position where it gets thematically interpreted (i.e. the *gap*).

Across different frameworks and psycho-linguistic approaches, difficulty associated with processing object (compared to subject movement) has been largely attributed to the degree of referential similarity between the intervening argument and the moved one. Behavioral studies (e.g. Gordon et al., 2001) and fMRI studies (Chen et al. 2006) showed that the processing cost asymmetry between subject and object extracted relative clauses could be modified by manipulating the featural characterization of the intervening head of the relative clause[50]. For instance the reading times measures in Gordon et al. (2001) are longer for object-extracted relative clauses compared to subject-extracted ones, when the NP within the relative clause was of the same type as the filler. In contrast, when a proper name was used, little to no difference between object and subject extracted relative clauses was observed. Another behavioral study by Van Dyke (2007) showed that the syntactic position of the intervening material has an impact on the "interference effects" [51] in subject-verb agreement dependency: the subject of a complement clause creates more interference within a subject-verb dependency than does the same NP within an object PP (see example (258) for a similar configuration).

Our results could be put in parallel with those obtained by a recent fMRI study by Glaser et al. (2013), who showed in *subject-verb agreement* dependency (not a movement dependency) that the 'similarity' of an intervener (an NP) to the head of the dependency drives the activation in Broca Complex (BA44 and 45).

Notably, the high interference condition had an intervening subject NP *visitor* within a complement clause (258b), whereas the low interference condition had an intervening NP within a PP (a). They observed an increased activation within Broca Complex for (b) compared to (a), that was understood as reflecting similarity-based interference effects in Broca area during what the authors interpret as a cue-based retrieval.

(258) Broca activation in high versus low interference configuration

a. Low: The client [who had arrived [***after*** the important **visitor** that day]] was complaining about the investigation.

b. High: The client [who implied [that the **visitor** was important that day]] was complaining about the investigation.

Taken together, these results demonstrate that the parser is sensitive to the syntactic and/or semantic similarity of features (e.g. [+sing], [+animate], [+definite]) between referential items be they in a verb agreement relation or in a filler-gap one.

Conversely, we tentatively adopt this approach to account for the complexity observed in the whole-brain contrast in Figure 6.4.2.

Although [+Animate][52] feature has not been considered among the features playing a main role in feature Relativized Minimality, we could consider the following featural configuration for conditions (c22), (c14) and (c30) in (259):

(259) Featural specification of (c22), (c14) and (c30)

50. In Chen et al. the complexity effect was found when the head noun of the relative clause was animate and the subject noun of the relative clause was inanimate.

51. Term used by the authors that we keep as it indicates an effect in a subject-verb agreement dependency.

52. We should note here that probably the notion of phi-features could be more adequate than the the one of animacy in our case (Bocci p.c.).

a. (c22) Qui $_i$ il imite ___$_i$? '*who he imitates?*'

wh-Pronoun,+ANI	+Pronoun +ANI	<wh-Pronoun,+ANI>
wh-object	Subject	<*qui*Object>

b. (c14) Où$_i$ il reside ___$_i$? '*Where he resides?*'

wh-Pronoun+LOC	+Pronoun +ANI	<wh-Pronoun,+LOC>
wh-Loc (PPloc)	Subject	<wh-Loc (PPloc)>

c. (c30) Où $_i$ il envoie ça ___$_i$? '*where he sends this?*'

wh-Pronoun+LOC	+Pronoun +ANI	+Pronoun+Obj +ANI .....
wh-Loc	Subject	Object
<+LOC wh-Pronoun>		
<wh-Loc>		

We can try and understand the results of the comparisons between (c14) and (c22) as linked to the fact that only in (c22) the silent object (<whObject>) is attracted to sentence-initial position and has to cross an intervener-subject that carries the same [+Animate] feature, thus yielding a so-called intervention effect. In (259b and c) the locative interrogative element "*Où*" has moved from its first merged position where it left a copy (indicated here by < > in the structure and underscore ___in the sentence).

In (259a) *Qui* and *il* share some features that define them, they are namely in a configuration where a subset of the moved item is shared by the intervening one. The direct object crossed over the similarly semantically restricted position of the subject, a configuration that was reported to be either more difficult or inaccessible to certain populations like children and aphasics.

However, one could ague that the fact that in (c14) *où* does not corresponds to an NP and a complement of the verb, but an adverbial, actually makes the contrast between (c14) and (c22) an imperfect minimal pair. This is why we performed an additional contrast where both (c14 and c30) are subtracted to (c22) (see Figure 6.4.2B. Namely if one could argue that *où* may not be a complement in (c14), it certainly is one in (c30). The fact that the results in contrast reported in 6.4.2B remain fundamentally the same actually reinforce the above exposed interpretation.

Considering sentences' complexity accounts related to intervention and those reported for non-movement derived 'interference' effects (e.g. Glaser et al., 2013 in verb-agreement dependencies), we may speculate that the four main activation's clusters revealed by the contrasts in Figure 6.4.2, point to areas that are implied in the calculation of the locality within which syntactic relations have to be satisfied. We could interpret the increased activation of temporal areas as involved in the *representation* of the full array of morpho-syntactic features. While frontal areas (i.e. Broca and Precentral) may be involved in the calculations of subset relation characterizing the featural specification of the intervener element and the linguistic item crossing it. The fact we observe Broca activation in this intervention effect is in line with the graded activation that Santi and Grodzinsky (2007b) observed in a fMRI study manipulating the increased number of additional NP interveners within a movement dependency.

We may cautiously conclude that from our results, that animacy feature may possibly contribute to similarity-based *interference* during sentence-internal dependency resolution. This could suggest that the semantic similarity along the [+/- animate] feature between the linguistic items found in a movement dependency has and impact on the amount of neural resources recruited to solve a filler-gap dependency.

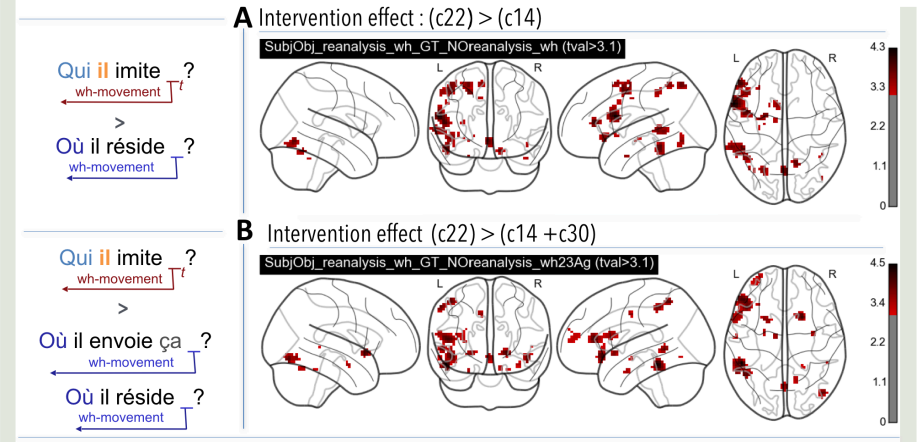


Figure 6.63 – Contrast between wh-question (c22) with object wh-Pronoun "qui" *who* and wh-question (c14) with locative wh-Pronoun "où" *where*. Peak activation clusters ($p < .001$ unc.) in IFG triangularis, IFG opercularis, IFG orbitalis, posterior Superior Temporal Sulcus, etc.

All in all, this contrast may be taken as pointing to the neural underpinnings of a complexity effects that can be cautiously link to the kind of intervention effects theorized within the RM framework, although we should remain cautious about the linguistic theory. Further discussion on this tentative interpretation is surely needed.

6.5 Unsupervised Analyses

It is more and more frequently hypothesized that language functions do not reside in single brain regions. This trend, for example, brought to the emergence of concepts like that of the "*network of interest*" (cf. Fedorenko and Thompson-Shill, 2014) or of the "*distributed network*" (cf. Pallier et al. 2011), and tangentially also to the emphasis given nowadays to *Localizers* paradigms, as we had the occasion to discuss presenting what we called the Sentence Network in chapter 1 (§1.4.4 and 1.3.4, p.55)[53].

More specifically, we can say that this trend is rooted in the conception that "*the mapping between neurons and cognition relies less on what individual nodes can do and more on the topology of their connectivity*" as discussed by Sporns (2011:184). Hence, we decided to perform an analysis bringing to light the brain regions that cluster together, based on the similarity of their patterns of responses to the manipulated variables in our study, in order to identify sub-networks within the Sentence Network.

As one of our main aims is to give empirical evidence for the cerebral division/distribution of labor between different fine-grained syntactic processes within the Sentence Network, we also pursued an analysis within each ROI of the similarity of patterns of activation to the 35 experimental conditions. This to explore if a given brain area would code for a particular syntactic complexity manipulation – an approach that would tentatively

53. For the brain maps of our French and Chinese localizers see Figure 1.11 (p.59). Section §H (p.949) in the Annexes is dedicated to the whole-brain and ROI results of the direct comparison of the 'Sentence network' of these typologically distant two languages. The reader will also find individual brain-maps of the contrast between Sentence and pseudo-characters or pseudo-words (i.e. strings of consonants) showing how reproducible are the hot-spots of this network across individuals.

contribute to sketch the *big picture* of the functional characterization of the Sentence Network's sub-parts.

A clear advantage to obtain results in this direction is that our rich experimental design allows to perform multivariate analysis, as a complementary investigation of the ROI analysis presented in the previous sections, on a very wide range of syntactic constructions. To our knowledge, this is the first brain imaging study on sentence processing that uses such a wide range of syntactic constructions, characterized by a relatively large amount of different syntactic transformations and a complete spectrum of their possible combinations.

One of the goals of an unsupervised analysis is to uncover patterns of responses that were initially not belonging to our experimental hypothesis, and in our case to gain a deeper understanding of how the manipulated variables of syntactic complexity modulate cerebral activation inside the Sentence Network – similar patterns across brain areas will be taken as evidence for collaboration in the processing, encoding or exchange of *similar* information.

Before leaving the *traditional approach* concerned with brain regions response amplitudes and to normalize the profiles to perform clustering analyses, we can still briefly consider the sensitivity of different ROIs to our syntactic manipulation. As shown in Figure 6.64, the different effect-size of the Conditions profiles across ROIs identifies the brain regions where the manipulated syntactic variable of the different types of Question formation in French have a wider effect range. Figure 6.64 shows that the areas show-

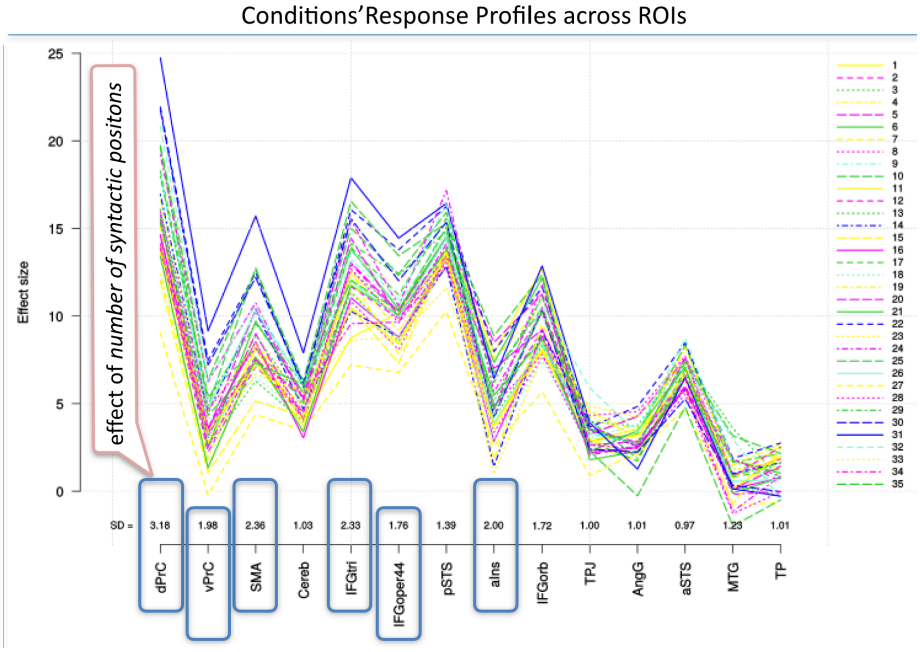


Figure 6.64 – Conditions' profiles across ROIs. The spatial response profiles of the 35 conditions (profile = line) in the 14 Regions of Interest. Color codes for the manipulated variable of movement types: *blue*: wh-movement; *indigo*: yes/no Questions; *yellow*: Declaratives; *green*: Verb-movement.

ing a greater range in amplitude in response to our movement-related manipulation are

mainly six, in decreasing order: dPrC, SMA, IFG-tri, a-Insula, vPrC and IFG-op. Interestingly, the significant effect of number of abstract positions which was detected in the left Precentral Gyrus, is also observed here. The conditions' profiles across ROIs show that this is the region showing a wider range in amplitude to the different conditions and that question types seem to present a graded distribution, with wh-movement being at the top followed by Verb-movement and simple yes/no questions, declarative being the least activated. This graded pattern is no more observable in more temporal areas (to the left of the graph) and crucially in pSTS and anterior Insula, which can therefore be taken as being functionally apart, as already mentioned in previous discussions.

6.5.1 Clustering brain regions patterns across conditions

The first unsupervised clustering analysis we performed on the fMRI resposes to the wide range of syntactic constructions, will concretely offer a measure of the distance between the conditions' response patterns in the ROIs illustrated in Figure 6.65A.

A brief reminder of how we proceeded to obtain the profiles in Figure 6.65A: after having estimated the responses to the 35 conditions using a linear model, the response profiles of the conditions were then extracted in our set of 14 regions of interest (ROIs) and normalized – the focus being here no more on the response amplitudes, but on the brain regions' response profiles. While the dendrogram in 6.65B was obtained by an unsupervised clustering algorithm, based on euclidian distance, that was then run on the matrix of correlations between the response profiles in (A). The results of this correlation revealed the 4 sub-sets in Figure 6.65C[54]. Put it simple, the closer the ROIs are, the more similar their response patterns to the experimental conditions are (see similarity of the response profiles in the left part of the figure).

A few remarks can be made about these four different clusters. The first and most basic is that these clusters group-up together non-adjacent areas, which is a first cue showing their functional pertinence.

Secondly, we can discuss these findings in the light of connectivity studies that performed both functional and anatomical connectivity analyses of sentence comprehension (e.g. Saur et al., 2010 and Xiang et al., 2010). By combining functional and anatomical connectivity, studies like Saur et al. 2010 started to describe how specific cognitive operations emerge through the interaction between anatomically interconnected brain areas. Hence, leaning on their findings on the functional and anatomical connectivity of auditory sentence comprehension, can give an additional anatomic grounding to the similarity of patterns observed in the three sub-networks emerged from our clustering analysis.

The posterior temporal region in proximity of the Angular gyrus ROI is reported to be directly linked to BA47-anterior temporal network via the inferior longitudinal fasciculus (or the Superior Longituninal Fasciculus III, see Marguiles et al., 2008), thus confirming the physical connection between Angular Gyrus and pars opercularis (orange network). This converges with the clustering we observed in the previous section for 3-arguments only conditions, when we investigated the linearization complexity of verbal arguments. This convergence could lead us to say that this sub-network is preferentially engaged in argument linearization complexity, thus explaining mirror response profile

54. For a graphical representation of these steps, cf. 6.14, p. 536

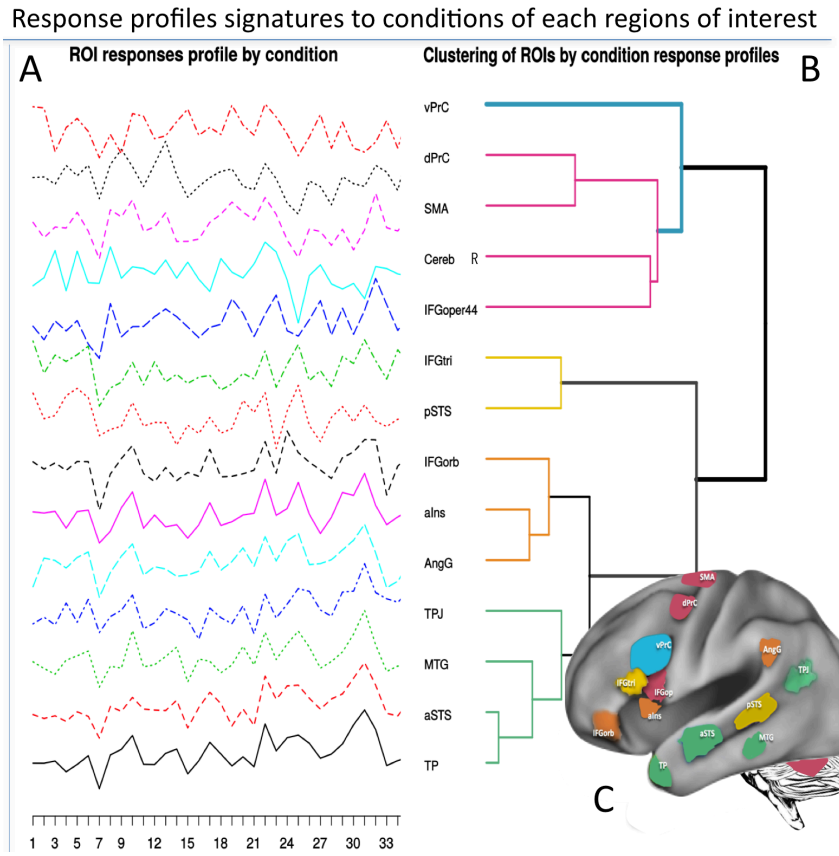


Figure 6.65 – Response profiles in 14 regions, and the associated correlation-based dendrogram. (A) Normalized response profiles. (B) Correlation Euclidean distance. (C) Sub-networks identified by the clustering on a brain map.

between anterior Insula and Angular Gyrus (cf. Figure 6.60, p.593)[55].

Moreover, the posterior part of Temporal Sulcus/Gyrus is connected to the MTG and superior temporal regions via the posterior section of the Arcuate Fasciculus (Catani et al., 2005), which confirms that the clustering of TPJ, MTG, aSTS and TP, is also related to an anatomical connectivity too.

Thirdly, we can note that although the centrality of Broca and Wernicke area in the Language and Sentence network is indisputable, our clustering reveals that the three different anatomically defined sub-parts of Broca, although contiguous, do not share similar patterns: each one belongs to a different sub-network and only Pars Triangularis (BA45) shows a similar pattern of responses with the pSTS (i.e. Wernicke).

Interestingly, Pars Opercularis (BA44) is clustered together with the Precentral complex and the right Cerebellum, while Ventral Precentral Cortex shows to be part of the Precentral-Cerebellum network, but still 'isolated' from it (see Figure 6.60).

The dorsal Precentral-Cerebellar similarity could be put in parallel with our previous discussion on sequencing processes and working memory – the right Cerebellum would be implied in the low level detection of patterns and the Precentral in holding them in a more hierarchical format, both encoding a rule-governed sequential ordering of hierarchical structures in working memory (Schubotz and Cramon, 2002, 2003; Christensen, 2010) and possibly being sensitive to hierarchical sequences predictability (Schubotz and Cramon, 2004). In this context Pars Opercularis would namely play the essential role it has often been attributed of building long-range sentence structure hierarchies *à la Friederici*.

Given the cardinality of this area in both its ventral and dorsal part in syntactic-movement results in the literature (e.g. adaptation paradigm by Santi and Grodzinsky, 2010) we can speculate that its relative isolation from the Precentral-Opercular cluster can attribute to it the role of a hub in movement-related complexity. The connectivity study by Raetting (2010) presented in the previous section discussing the central role of Precentral in the linear effect of number of syntactic positions, can be viewed as an additional argument grounding this speculation. Ventral Precentral area represents the most connected area to the Temporal regions, and it turns up to be the confluence point of more fiber bundles as those that reach Broca Opercularis and Triangularis (BA44 and BA45, cf. Figure 6.53, p.584).

We remain cautious on the functional characterization of each of these sub-networks, but still can say that they show a main functional separation between frontal and temporal areas, with nonetheless three long-range similar pattern/connections, (1) a close connection between Triangularis and the pSTS, (2) a similarity between Precentral-Opercular complex and right Cerebellum, and (3) between Angular Gyrus - anterior Insula and Pars orbitalis. Last but not least, the very interesting relative 'isolation' of ventral Precentral response patterns confirms previous results indicating that this area plays a central role in syntactic-movement operation and may constitute a hub in the frontal region.

6.5.2 Clustering conditions' patterns inside ROIs

As a second unsupervised analysis, we performed in each ROI a clustering of our experimental conditions based on the similarity of their spatial patterns of activation across

55. Angular Gyrus could be dedicated to the retrieval of thematic arguments as indicated by previous literature, while Insula would show an increasing activation with the distance from canonical word-order.

voxels. This revealed several interesting groupings. We limit ourselves here to three areas that showed clear-cut results as illustrated by Figure 6.66.

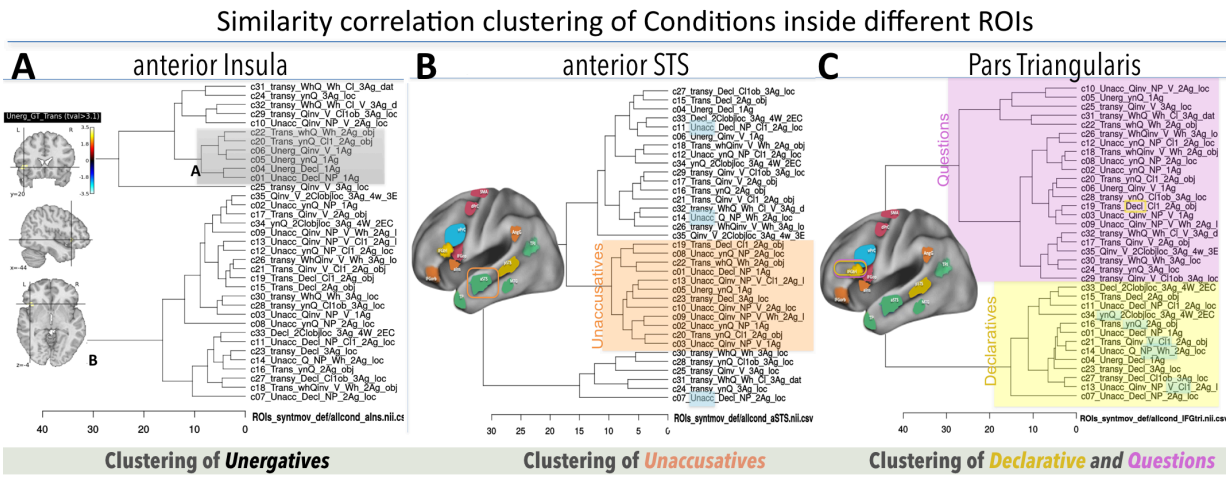


Figure 6.66 – (A) Correlation-based dendrogram in **anterior Insula**, and whole-brain contrast T-map for Unergatives vs. Transitives. (B) Correlation-based dendrogram in **aSTS**. (C) Correlation-based dendrogram in **Pars Triangularis**.

The classification observed in anterior Insula confirms what we observed in the whole-brain analysis when comparing Unergatives to Transitives' conditions, we found an increased activation cluster in IFG Triangularis, a result we interpreted as reflecting the creation or the licensing of a null object in Unergatives having only an agent argument (i.e. legitimation of a little *pro*). Here, all Unergative conditions group-up in the correlation-based dendrogram of anterior Insula/FoP (Fig. A).

A second set of results obtained by this analysis concerns the Unaccusative conditions, shown to have similar response patterns in anterior Superior temporal Sulcus (aSTS), thus confirming the results we obtained in the ROI analysis contrasting Declaratives and yes/no questions in two-arguments Unaccusatives. Here, all Unaccusative conditions group-up in a branch of the similarity-based dendrogram in Figure 6.66B irrespective of their number of arguments. This result allows us to extend our claim about the central role played by this anterior temporal area in Local movements like clitic- and NP-movement in our experiment, and two other local movements from the literature in Danish (i.e. negative-shift in Christensen, 2008) and in Hebrew (i.e. dative-shift in Ben-Shachar et al., 2004).

Thirdly, in Figure 6.66C, we can observe that all questions group-up in one of the two main branches of the dendrogram, letting all declaratives on the other side in the correlation-based dendrogram in IFG Triangularis.

Similarity between response patterns across brain areas and condition patterns inside each brain area: Correlation matrix

Crossing the information from these two clustering analyses – i.e. clustering conditions in (A) and clustering ROI's response profiles in (B) – we obtain the following Correlation matrix in Figure 6.67.

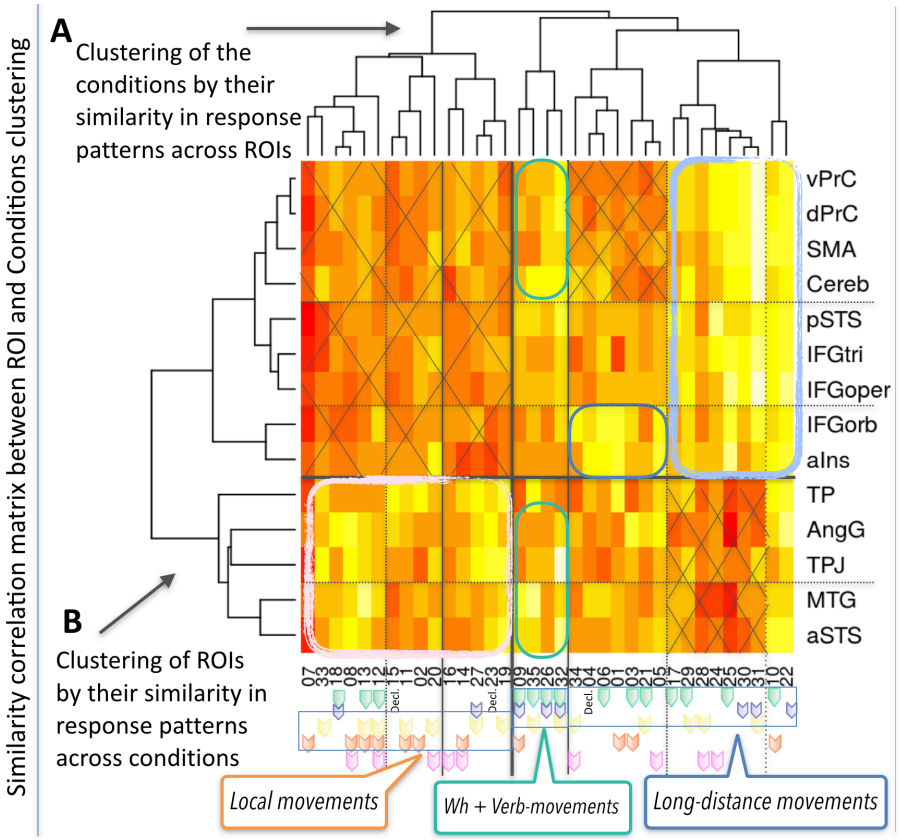


Figure 6.67 – (A) Correlation by Euclidean distance clustering conditions. (B) Correlation by Euclidean distance clustering ROI's response profiles. Color codes for the manipulated variable of movement types for the different conditions: *blue* = *wh*-movement ; *indigo* = yes/no questions ; *yellow* = declaratives ; *green* = Verb-movement *orange* = Unaccusatives. Color code for correlation matrix: Yellow higher correlation, Red lower correlation.

Three main observations can be made. Starting from the right, Long-distance movements targeting the CP-layer are mainly clustered on the right side and clearly show more similar patterns in the frontal network encompassing IFG-Precentral complex, Cerebellum and pSTS. On the left-hand side, we observe an opposite pattern for more local movements (clitic and NP-movement) showing more similar activation patterns in Temporal areas.

In the middle of the matrix, we can observe the combination of *wh*- and Verb-movement yields a third and mixed pattern where both frontal Precentral Complex (i.e. vPrC, dPrC, SMA and Cerebellum) and temporal areas (Angular Gyrsu, TPJ, MTG and aSTS) show higher similarity of activation patterns, thus revealing both a local- and long-distance complexity. This last more temporal distribution is coherent with the fact that our ROI analysis revealed an additive effect of *wh*- and Verb-movement in the aSTS (cf. Figure 6.28, p.551).

These findings, sketching a temporal-frontal division of labor, can be related to results reported by Friederici and colleagues (Friederici et al., 2009) and Bornkessel and colleagues (Bornkessel et al., 2005), who observed a correlation between left IFG activa-

tion and the computation of a syntactic frame during sentence processing, whereas posterior superior temporal areas were recruited when the interpretation of thematic-roles and verb argument hierarchies was needed. This can also account for the whole-bain and ROI results we observed for clitic and NP-movement, which imply an increased syntactic complexity of the inferior sentence layers, namely the VP and IP domains that are dedicated to Thematic argument structure and grammatical functions.

Further evidence going in this direction can be seen in fMRI studies using violations paradigms like Newman et al. (2003), that reported greater activation in the IFG Opercularis for sentences containing purely syntactic violations, while increased activation of left posterior temporal regions was observed when processing both syntactic violations *and* thematic errors.

We conclude this discussion by letting it momentously suspended until the end of next chapter. The results of the fMRI study on Chinese will indeed contribute to further tease apart the actual contribution of the sub-networks identified by this unsupervised approach and the initial fronto-temporal syntactic task-sharing emerging from our results and confirmed by the above correlation matrix.

6.6 Summary and Perspectives

This section will retrace the main findings of this experiment and their broader implications for the functional characterization of the Sentence Network. Future research, follow-up questions, and comments on some possible critiques and objections, will be then put forth.

This study brought together linguistic formal theory of sentential displacement phenomena with the neuro-imaging investigation of syntactic complexity. We searched for the neural correlates of theoretically postulated syntactic derivations of French questions, clitic placement and unaccusatives. Uni-variate and multi-variate approach, allowed us to observe several fine-grained effects related to different complexity dimensions of syntactic-movement, and some broader movement-related modulations of the Sentence network, allowing us to enlarging our perspective and draw some conclusions on the neural underpinnings of syntactic complexity in the brain.

Searching for the neural correlates of movement-related syntactic complexity in different types of movements, we observed several distinct effects:

1. The effect observed for *wh*-movement in French *wh*-questions fully replicates previous findings in Danish and Hebrew *wh*-questions (Christensen, 2008 and Ben-Shachar, 2004) and also for other *wh*-movement constructions like Topicalization (Shetreet and Friedmann, 2014) or Relative clauses (Ben-Shachar et al., 2003). Our results further confirm the involvement of Cerebellum and Precentral cortex in this syntactic transformation.
2. As for the difference between *wh*- and Verb-movement in French questions, we observed that comparing these two movement to their respective baselines yielded two similar whole-brain effects. While their direct contrast revealed for *wh*-movement three main clusters of increased activation in Broca Triangularis /Opercularis, in Inferior Parietal Lobule and Inferior Temporal Gyrus. The opposite contrast for Verb-movement revealed increased activation in right Broca Complex, the SMA, inferior Occipital Gyrus and Temporal-Parietal

Occipital junction (TPOJ). However, the main proof for the fundamental difference of these two movement operations is to be found in the interaction with verbal argument complexity, which was observed only for Verb-movement.

3. Clitic-movement complexity yielded more temporally and internally distributed activations. We observed the increased activation of aSTS and left posterior Supramarginal Gyrus (i.e. TPOJ), together with more internal areas like the anterior Insula, the anterior Cingulate, the SMA and Precuneus. The only frontal activation was elicited in the MFG (BA9).
4. NP-movement showed together with clitic-movement the preferential recruitment of MFG (/SFG), aSTS and Precuneus (right), and additionally showed two right lateralized clusters in dorsal Precentral and inferior Parietal region. Although no significant difference was observed in the contrasts between Unaccusatives versus Unergatives, our results for this movement type in 2-argument, and the ones obtained by comparing one argument Unaccusatives versus Transitives, fully replicate previous finding in Hebrew by Shetreet and colleagues (2010 and 2012). Moreover, Unergatives showed an increased activation of Broca area.
5. More generally, the two local movements we included (NP- and clitic-movement) reproduced the temporal activation of anterior/mid-temporal lobe that were reported for other local movements in the literature, like dative-shift and negative-shift (Ben-Shachar et al. 2004 and Christensen, 2008), specifically in aSTS.
6. The derivational complexity linked to the multi-step progression of clitic- and Verb-movements, or the one characterizing the combination of *wh*- and Verb-movement, and the combination of clitic and Verb-movement or in double clitic sentences, were all observed in a very posterior temporal area, the Temporo-Parieto-Occipital Junction (TPOJ). A result we put in parallel with the increased activation in the overlapping posterior Supramarginal area obtained by Bachrach (2008), when correlating brain activity with a measure of the number of derivational steps of a parser, or even more simply by tagging syntactic displacement in narratives.
7. Simple yes/no questions (obtained by adding a question mark to the SVO declarative word-order) revealed a small focal activation in anterior Insula (FoP) in a General contrast, while in a more Minimal contrast, right Broca showed an increased activation, reflecting the discourse properties of interrogative mode.
8. The linear effect for the increasing number of syntactic positions – our proxy-measure for syntactic-tree complexity – actually revealed a contribution of each movement operation to a general complexity effect in a large part of the Sentence network, and more consistently in the dorsal part of Precentral Gyrus. Interestingly, interaction between movements were observed in several ROIs for different movement combinations, notably for wh- and Verb-movement, and for clitic- and Verb-movement.
9. An effect of number of argument was observed in Angular Gyrus irrespective of the grammatical nature of the argument (i.e. clitic or strong pronouns). Notably, this area also showed an intriguing mirror response compared to anterior Insula along with the degree of distance from the canonical SVO surface linearization of arguments. While the Insula appears to respond to the heaviness of the displaced syntactic element, the Angular Gyrus shows an inverse pattern and responded more to the canonical word-order and the lightest displaced elements. This mirror response could indicate that Angular Gyrus is fundamentally involved more in a retrieval processes of arguments than in movement-related operations applied to verbal arguments.

6.6.1 Broader Implications for the functional characterization of the Sentence Network

Big picture: What do we learn from our unsupervised approach

The main point of performing unsupervised analyses was to discover what kind of modulation of the Sentence Network (reduced to our ROIs selection) was induced by the movement-related syntactic manipulations characterizing our stimuli materials. Crucially, it allowed us to have a broader overview on the results we discussed in details using the traditional contrast-based approach. We can thus conclude this chapter by pointing out to some '*big picture*' observations on the Sentence Network organization.

The areas that showed a greater range in amplitude in response to our movement-related manipulation are mainly six in decreasing order: dPrC, vPrC, SMA, IFG-tri, IFG-op and a-Insula. And the significant effect of number of abstract positions that was detected in the left Precentral Gyrus, is also observed in the conditions' profiles across ROIs. Notably, we can observe that this region is not only the one showing a wider range in amplitude to the different conditions, but also that the different question types (*wh*-questions, verb-inverted and simple y/n questions) appear to show a graded distribution, where *wh*-movement is at the top.

The clustering of ROIs, based on the similarity of their responses, showed a clear functional tri-partition of Broca's area with each of its sub-parts clustering with a different sub-network. Importantly, Pars Opercularis groups up with the Precentral Complex and Cerebellum, which indicates that the role of right Cerebellum is tightly linked to the processes happening in the Precentral-Opercular network. Note that the ROI analyses also revealed the involvement of this network in different types of movements, as also showed by its linear increase in activation with the number of syntactic positions. We discussed the possible implication of Cerebellum in the detection of sequential patterns in movement-derived sentences, while Precentral Gyrus was interpreted as housing working-memory processes of holding sequential ordering of sentence's hierarchical structure having undergone movement.

From the correlation matrix, one can conclude that Verb- and *wh*-movements activate similar frontal regions (IFG and Precentral Complex) when found in isolation, while their combination shows a pattern encroaching both frontal and temporal regions. Crucially, more local movements like clitic- and NP-movement appear neurally distinct and modulate activation mostly in the temporo-parietal regions, as they show in this analysis more similar activation patterns in temporal ROIs. We advanced that this temporal /frontal distribution could be associated to a distinction between Argumental dependency characterizing syntactic movements targeting Argumental positions (A-movement) and non-argumental involved in wh-movement, or the hierarchy dimension of targeting a higher syntactic layer present in both Verb- and *wh*-movements.

We can conclude by saying that the convergence of this analysis with the classic contrast based results of this study contributed to tease apart the contribution of the different sub-networks identified by this unsupervised clustering approach. The fronto-temporal syntactic *task-sharing* emerging from our results seems to indicate that frontal areas are more involved in long syntactic-movement related processes. We could speculate that this is linked to the additional portion of syntactic-tree to be built when movement is targeting a higher position in the syntactic-tree. The clustering analysis correlated this syntactic complexity dimension to the observed Precentral-Opercularis Cerebellum

sub-network, where we argued that an interplay between a sequential pattern detection process in the Cerebellum, and a working-memory retention process of hierarchical tree-structure in the Precentral-Opercularis complex can take place. On the other side, a temporal network appears to be more concerned with the kind of transformations happening in the lower layers of the Sentence Domains's hierarchy yielded by more local movement (clitic- and NP-movement) targeting argumental positions.

6.6.2 Caveats and Future research

One of the most urgent step for future research is, in our opinion, to replicate the rich set of results obtained in this experiment. This is actually why the idea came to implement an oral version of this study slightly modified by adding or replacing five conditions, and precisely modifying the declarative baseline conditions.

The question of having a good baseline in neuro-imaging contrasts is classically a very important one. Yet, some possible critiques can be raised about the Declarative baselines we used, would be namely pointing to the contrastive or at least deictic use of the strong pronouns "ça" *this* and "là" *there* in object position. A better controlled baseline would allow, for example, to perform some conjunction analysis which would reveal the network shared by the different types of movements (e.g. clitic and NP movement, Verb and clitic movement, etc.).

In the perspective of replicating in the auditory modality the results we obtained, it would be even more essential to avoid the possible contrastive of the strong pronouns "ça" *this* and "là" *there* – the oral recording of "ça" *this* declarative sentence cannot but put emphasis on this contrastiveness feature linked to its deictic use.

While the use of such strong pronouns surely represents from the linguistic point of view a caveat of our experimental conditions (we essentially chose them to constrain the length of the sentences), it should be noted however that they also crucially allowed us to confirm that the posterior temporal activation observed in two clitics sentences – when compared to their declarative counterparts – was linked to their derivational complexity and not to the features of their discourse antecedent, as both clitics and strong pronouns need to have a [+ accessible, + active] discourse referent.

A possible solution would be to have baseline conditions with monosyllabic full NP objects, like proper names (e.g. Zoe, Jean, ...), which would probably serve as an optimal baseline for clitics. This would permit to observe the brain areas responsive for clitic pronoun, while we had a baseline that also involved strong pronominal elements (e.g. "ça" *this* and "là" *there*.

Moreover, French Unaccusatives could be worth some further investigation. The questioning raised by the null result we obtained for the comparisons between Unergative and Unaccusative verbs when compared to previous studies on Hebrew, may find an answer by comparing the type of strictly semantic Unaccusatives we selected and another type of morphologically obtained Unaccusatives in French, i.e. verbs introduced by the morpheme *se* like "se salir" *to dirt oneself* or "se coller" *to stick oneself*, which actually constituted a pretty important proportion of the Unaccusatives verbs present in the valence dictionary we consulted for our verb choice. Using this type of Unaccusatives may reveal similar brain responses as Hebrew ones and additionally permit to contrast the use of morphological means to obtain unaccusativity to purely Unaccusative verbs[56].

56. Note however that in this case, a way to control for the number of thematic grids should be found.

At last, the role of Precental Gyrus remains open to discussion between three main movement-related complexity processes, i.e. (i) gap prediction, (ii) syntactic working-memory for holding sentence structure's hierarchical patterns, and (iii) encoding of sequential ordering of hierarchical sentence's structure. The next chapter will deliver some new elements on this issue, we thus move on to the next and last study of this manuscript.

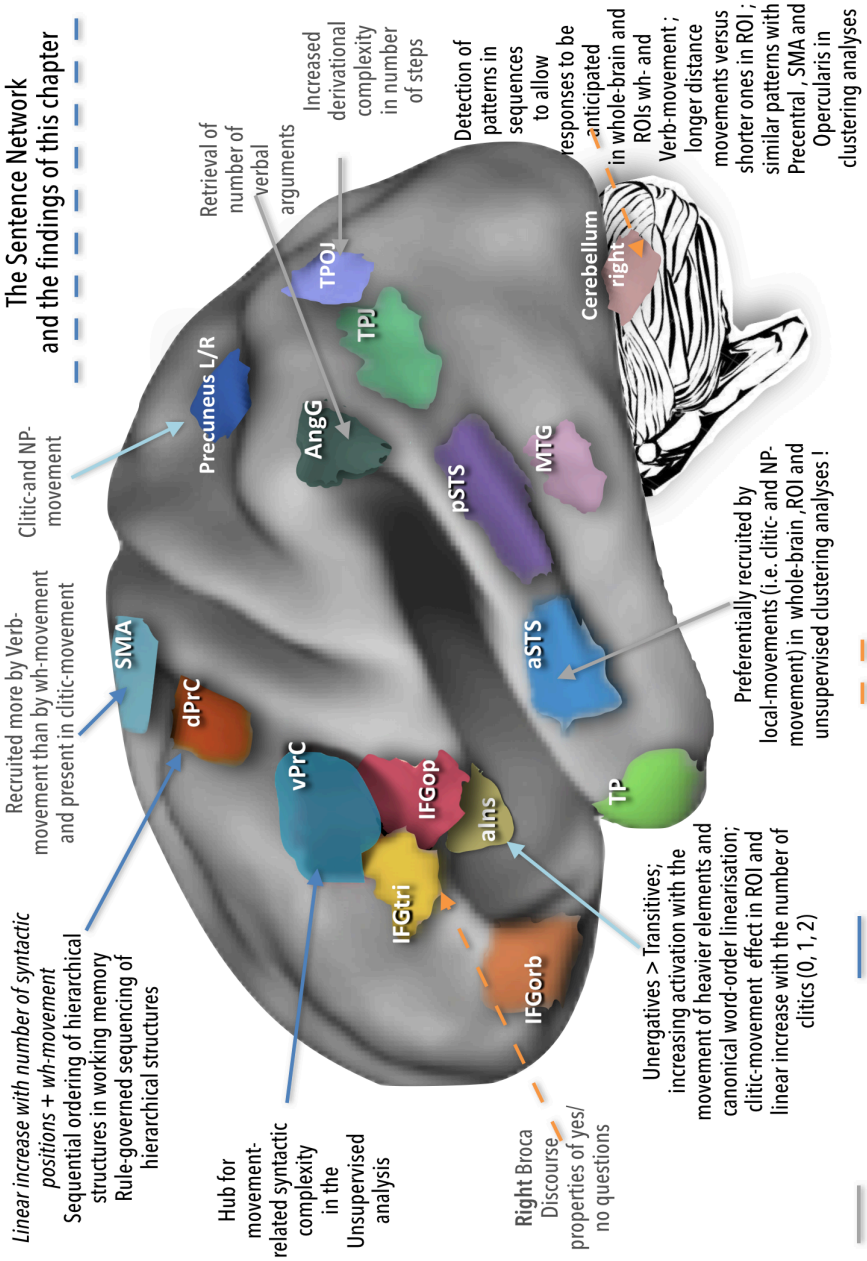


Figure 6.68 – Mapping results of current chapter onto the Sentence Network, to functionally characterize the areas and the different sub-processes that it houses.

Chapter 7

Cerebral encoding of syntactic-layers in the Chinese Left-Periphery

> Dans le passé, en effet, à travers plusieurs problèmes, nous avons valorisé l'extériorité [...] [or] l'enveloppe mérite d'être relevée et réfléchie, car elle enferme en elle de riches significations à découvrir, ce qu'elle doit d'ailleurs à son statut, celui de l'interfacialité ; elle se situe entre le dedans et le dehors, elle traduit l'un et occupe l'autre.
>
> *[In the past, in fact, across different issues, we valorized exteriority [...] but the* ***envelop*** *merits to be considered and thought about, because it contains rich significations to be discovered, that the envelop acquires thanks to its status:* ***inter-faciality****. As it is situated between the inside and the outside, it translates the first one and occupies the second.]*
>
> François Dagognet

[1]

In this last chapter, typological perspective and formal approaches to Chinese syntax chorally unite, to contribute several testing hypotheses to experimentally investigate the representation of the hierarchical structure of the sentence-unit by the human mind and brain. As stated in the above epigraph by a French poet, we will argue that focusing on the sentence-discourse interfacial phenomena — the sentence's "*envelop*" of the Left-Periphery —, can reveal several aspects of the cerebral representation of syntactic structure and of intra-sentential dependency-links. As was the case when considering French question formation in the previous chapter, we will once again leverage some crucial characteristics gathered by the linguistic system of a given language to shed light

1. *Les noms et les mots*, Les Belles lettres, coll. Encre marins (2008).

on the functional specialization of certain brain areas in the different movement-related processes yielding syntactic complexity effects.

By addressing the issue of the internal structure of the sentence across languages in chapter 2 (§2.2.4), Topic-Comment articulations were identified as carrying interesting discourse properties, which enriched our questioning about the neural underpinnings of the sentence-unit. We will, focus now on sentence-discourse interfacial phenomena, like Topic or Focus in Chinese, to investigate how they modify the articulation of the sentence, in order to reveal the sentence-unit structural organization. We illustrated this strategy by the now familiar architectonic metaphor of the cupola, which could be reformulated referring back to the french Poet in the epigraph, as a strategy targeting the interfaciality of the 'envelop' of the sentence to understand the inside, the internal structure.

In this chapter like in previous one, the linguistic formal approach to the sentence structural representation that will prove again to be a crucial tool to study the sentence-unit from the perspective of cerebral processing. Namely, the theoretical formalization step of the sentence structure it took some time to introduce (see chapters 2 and 3), becomes now the center of our experimental hypotheses on the cerebral representation of (i) sentence's domains, (ii) syntactic derivations of sentence structure and (iii) the ordering of functional projections in the Left-periphery. Moreover, the comparison with French fMRI results will be possible only through these theoretical tools.

This fMRI study we now going to describe, re-proposes under a different light the general question of this doctoral work about how the brain represents the hierarchical structures of sentences during language comprehension, and how the cerebral representation of the sentence structure can actually be aligned with the hierarchical structures, syntactic analyses and mechanisms postulated in linguistics. Specifically, we will here investigate three fundamental properties of the sentence-unit:

1. the processes that determine its basic and complex structures, comparing canonical SVO sentences, Topic-Comment articulations and movement-derived constructions;
2. the intra-sentential dependency relations, comparing different anaphoric strategies between Topic and Comment, encompassing resumption, gap and null pronouns;
3. the encoding of the complexity dimension linked to the presence of elements in the sentence-discourse interface CP layer, the so-called Left-Periphery.

Overview of the contents of this chapter

7.1 Dissociating different syntactic complexity dimensions: movement, Left-Periphery and dependency-links

7.1.1 An answer the state of the art in neuro-syntax

The actual state of the art in neuro-syntax is compelling researchers to deepen their understanding of the often intermingled factors that compose the sentence's complexity effects reported in the neuro-imaging literature, to be able to break down broad complexity effects, like movement related processes, into finer neuro-linguistic sub-processes.

Syntactic complexity: from word-order to movement related effect In psycho-linguistics and neuro-linguistics, it is acknowledged that sentences with non-canonical word-order involve greater processing loads compared with canonical word-orders (cf. §1.4.4, Fig.1.10 in aphasics and §2.2.2.1, in neuro-imaging)(cf.2.2.2.3).

As delineated in the last two sections of chapter 2 (cf. §2.4, p.157) several dimensions of syntactic complexity have been brought to light, mainly manipulating syntactic parameters, going from (1) simple word-order variations as we saw in several German studies (Grewe et al., 2006; Obleser et al., 2011 (Fig. 2.4) and Bornkessel et al., 2009 (Fig.2.6), or (2) sentence embedding as we saw in Ben-Shachar et al. (2004 exp.2), to more theoretically-driven manipulations of syntactic complexity, implying syntactic movement in different sentence structures in Hebrew (Shetreet et al. 2014), English (Ben-Shachar et al. 2004), and Danish (Christensen, 2008).

However, if we consider the different dimensions that are involved in syntactic movement, we can at least highlight five aspects that are intermingled in this movement operation:

1. Gap: A moved element leaves a covert empty element at it extraction site, which can be of different type (cf. §2.4.3);
2. Embedding: Syntactic movement can imply or not embedding;
3. Movement Type: Different movement types exist, they can follow different derivational steps and be triggered by different elements, or leave different kind of traces behind yielding different dependency chains, as we saw in previous chapters.
4. Landing site: Movement can target different positions in the syntactic-tree structure and crucially different sentence domains, and the height of the position in the syntactic-tree where the moved element lands holds for a complexity dimension as attested in aphasic recovery patterns (cf. fig. 7.1)
5. Filler-gap dependency-link: Movement establishes a chain between the moved element and the extraction site and referential assignment is done based on these kind of dependency links.

To these purely syntactic dimensions one could also add a psycho-linguistic one, such as the predictably of a trace (Santi and Grodzinsky, 2012), which can facilitate cerebral processing[2]. Alternatively, displacement effects have been studied under different perspectives, and we reviewed the Relativized Minimality's complexity effects observed in children and aphasic populations when a dependency is crossing an intervener of the same type as the head of the dependency.

2. The prevalence inside a give language of certain sentence constructions like object-first constructions in Swedish (see Platzak et al., 2001, p.210) has also been reported as a parameter facilitating the processing.

Our review of the literature in chapter 2, delineated how the exploratory studies of the beginnings of sentence neuro-imaging were followed by more linguistically-based ones, which started investigating some of the complexity parameters linked to the syntactic-tree structure representation of the sentence-unit (§2.4 and 2.3).

Among the most theoretically-oriented studies some have been attempting to disentangle (i) EMBEDDING FROM MOVEMENT, others explored the neural underpinnings of (ii) different aspects linked to the presence of gaps left by syntactic movement, and yet others (iii) sought to characterize the neural patterns for different types of movements:

(i) For instance, Ben-Shachar et al. (2003) contrasted embeddings with transformations (object-relatives) against sentences that contained embeddings without transformations (embedded declaratives with sentential complement), and Santi and Grodzinsky (2007) found a brain area adapting to movement only and one adapting to movement and embedding.

(ii) Different syntactic aspects linked to the presence of gaps left by syntactic movement were investigated by Santi and Grodzinsky (2012), who attempted to distinguish between the activation pattern linked to the establishment of intra-sentential dependency from of that associated to displacement.

(iii) As for different types of movements Shetreet et al. (2014) investigated wh-movement in Topicalization versus Verb-movement to sentence second position, while Ben-Shachar et al. (2004) contrasted wh-movement versus Dative-shift.

Hence, what we propose here is to continue in this direction and find a syntactic configuration allowing to observe the cerebral *encoding* of the ***different layers of the sentence representation*** (i.e. the height in the syntactic-tree) as distinct from syntactic movement operation itself.

7.1.2 Dissociating complexity dimensions: position in the syntactic-tree and movement

Since the seventies a large number of studies have repeatedly shown that the comprehension and production of non-canonical sentences derived by syntactic movement of the object can be impaired in individuals with syntactic disorders (e.g. agrammatic aphasia, see Caramazza and Zurif, 1976; for a review see Grodzinsky, 2006). Interestingly, the impairment of *wh*-movement-derived sentences in agrammatism has been interpreted as the result of the degradation of the cerebral representation of the landing-site of this type of movement — a position that is crucially found at the top of the syntactic-tree skeleton in the CP-layer as illustrated in Figure 7.1 (cf. the *Tree-Pruning Hypothesis* by Friedmann et al., 2006a, and earlier work, §2.4.4.4).

As argued in our review in Chapter 2 about neuro-syntax (§2.4.4), not only the neuro-imaging findings on which is built the *Sentence Domain hypothesis* by Christensen (2008)[3], but the reproducible neuro-psychological effects reported from the linguistic

3. The division of labor proposed by such an hypothesis attributes different activation patterns or sub-components of the language network to the Complemetizer Phrase, this particular sentence domain that we previously presented closing the IP-domain and linking the propositional information of IP and VP to discourse.

More precisely, this hypothesis posits that differential activations in the sub-components of the sentence cortical network could reflects the computation of different syntactic domains: the neural sub-component of the VP would encode the Thematic Argument Structure, those of the IP for grammatical relations, and those of the CP for pragmatic functions and quantification.

behavior of aphasics (§2.4.4.4) converge in pointing to the height in the syntactic-tree representation where moved constituents land, as a central complexity parameter to be considered in order to understand the cerebral encoding of the sentence-unit. In other words, according to these hypotheses on the representation of syntactic complexity in the brain, the syntactic-tree structure is taken as an interiorized representation, which yields brain activity (i.e. in Broca's area) to depend on the height in the tree-hierarchy of functional nodes where the elements having undergone movement land.

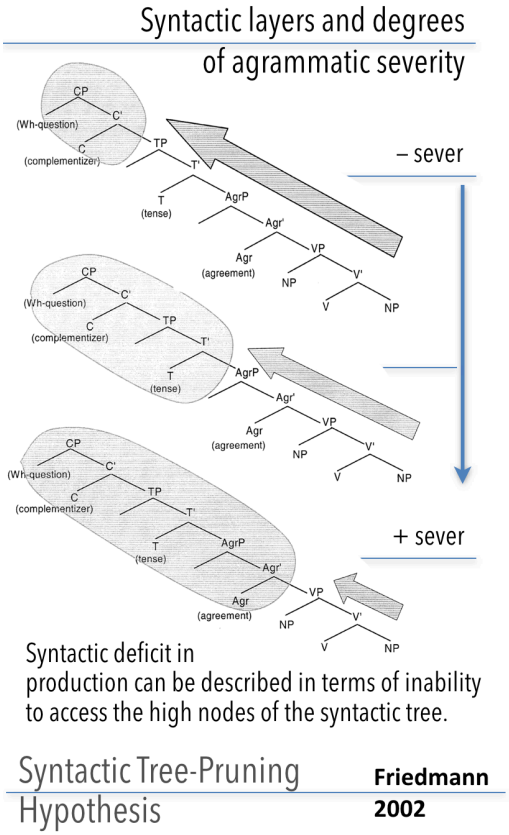


Figure 7.1 – Tree-Pruning Hypothesis. Schematic representation of how the production syntactic deficit in agrammatic patients can be described in terms of inability to access to the higher nodes of the sentence's syntactic-tree skeleton. Syntactic domains circled in grey represent the extent of inaccessible tree-nodes according to the degree of severity of the syntactic impairment. Adapted from Friedmann (2002).

This complexity dimension linked to the cerebral representation of the sentence's syntactic domains is now at the center of the current experimental design. One of the central hypotheses will be that the complexity linked to hosting an element in the CP layer is observable in Chinese as being distinct from other movement-related processes. In this, we attempt to go beyond the actual state of the art in neuro-linguistics: leveraging on the characteristics of Mandarin Chinese to dissociate two syntactic complexity dimensions that are hardly dissociable in other linguistic systems, namely syntactic movement and the presence (/height) of a constituent in the Left-Periphery.

Our fMRI study is the first to compare cerebral activity related to the processing of very similar syntactic constructions that differ in their syntactic derivation: Base-Generated *in-situ* Topics versus A-bar moved ones. This particular configuration where the Topic is found in the CP-layer without having been moved to it — the so-called Base-Generated or in-situ Topic—will allow us to observe their individual contribution to syntactic complexity effects reported in the literature, and to nail down the cerebral representation of the CP syntactic-tree layer, in absence of embedding, movement, gaps and dependency-links.

The second main issue we address in this study regards the cases where the presence of a moved element in the CP layer also implies the creation of a dependency-link (i.e. a movement chain) with the extraction site. I n this again Chinese offers a favorable syntactic configuration allowing to compare with minimal lexical difference the brain activation patterns for the same syntactic structure where a gap, a null pronoun or Resumptive pronoun can be found. This will permit to observe different

effects yielded by the different nature of dependency-links established between the Topic and these different syntactic elements.

In sum, given these characteristics, we will be able to theoretically disentangle three complexity parameters present in moved Topics: (1) the complexity of the syntactic derivation by movement versus the presence of an *in-situ* Topic in the CP layer; (2) the complexity of establishing a dependency-link between the Topic and Comment versus its absence; (3) the complexity parameter linked to the presence of a covert syntactic element in in the Comment clause (i.e. gap or *pro* versus Resumptive).

We will recapitulate in the following how Chinese Left-Periphery presents several formal characteristics that offer an ideal configuration to investigate how the brain represents Sentence-domains.In this, both Tree-pruning hypothesis and Sentence Domains Hypothesis framework will be grounding our experimental approach to Left-Periphery phenomena in Mandarin Chinese, specifically when addressing the issue of the fMRI activation patterns for different ordinal positions in the Left-Periphery to search for a brain area that would possibly encode for the relative height in the syntactic-tree representation where different Topic elements are hosted according to the Cartographic approach.

7.1.3 Why is Chinese an ideal testing ground

The thorough overview into the syntactic characteristics of Mandarin Topic-comment structures provided in chapter 3 (p.225) offered a full-fledged linguistics grounding for our interest in Chinese Left-periphery. We observed that Mandarin Chinese gathers an astonishing number of syntactic characteristics allowing us to disentangle, isolate and test the following syntactic complexity parameters that are otherwise often correlated: (a) syntactic hierarchy, (b) syntactic movement transformations, (c) presence of resumption versus empty syntactic positions (i.e. gap or null pronouns), and (d) the representation of the sentence-discourse syntactic layer without other complexity parameters.

Topic-comment an essential syntactic construct to 'say something about something' Addressing the issue of *sentence as a cognitive object*, Topic-comment articulations also provided the opportunity to test the most extreme and at the same time essential configuration capable of creating propositional meaning in human language: the possibility of 'saying something about something', by simply partitioning the sentence into a two step predication: first positing a Topic and then uttering a Comment about it.

However, this very basic predicative relationship —saying something about something— that Topic-Comment sentences embody, requires a well defined structural architecture of the sentence-unit. Namely, this minimal predication configuration implies the establishment of a relationship between the Topic and the Comment-clause, that over-arches the subject-verb agreement one, and sets a sentence-level hierarchy between the Topic and the whole Comment sentence[4].

Linearity, hierarchy and syntactic derivation Hence, besides sitting on the throne of Sinology as the discipline of *alterity* to explain the presence of Chinese in the experimental

4. The reader might remember we had defined this way of structuring the sentence-unit as a sentential Copernican Revolution in chapter 3 p. 228 (An Initial Intuition), where the verb and subject are excentered from the the 'sentence system' and the tree-diagram representing its syntactic encoding.

part of this work - the classical '*Chinese is different*' argument-, our choice was motivated by the wish to investigate the biological basis of a language system, where articulating sentences according to the Topic-Comment pattern is the predominant sentential form in every-day ordinary usage, and where linear position and a pause prosodic-marking (cf. chapt. 4 and 5) are the minimal cues required to maximally change the hierarchical relations in the sentence into a Topic-Comment articulation.[5]

Minimal morpho-syntactic marking The second advantage of choosing Mandarin as a neuro-linguistic testing ground is its relatively scarce morpho-syntactic marking. Specifically, in Topic-Comment articulations morpho-syntactic marking is fairly optional, thus *permitting to study how the mind and the brain manage this incredible equilibrium the sentence achieves between linearity and hierarchy* without morph-syntactic cues.

The idea is that different types of *linguistic cues* to build sentence structure, in our case word-order compared to morpho-syntactic marking of sentences constructions, are likely to recruit different cognitive resources to be represented and processed by the brain. A basic example could illustrate this point: a language like German marks argumental roles using case marking inflectional morphology, while a language like English or Chinese uses word-order to achieve the same argumental role encoding in the sentence. Hence, Mandarin Topic-Comment structures allow to investigate sentence hierarchical structure based on word-order variations without the necessary presence of overt functional elements, like Topic heads.

Two different syntactic derivations for Chinese Topics Thirdly, the issue of Topic-prominence and the consequent basic status in the Mandarin linguistic 'eco-system' of Topic-comment sentences, is actually linked to their particular syntactic derivation.

Following Paul and Whithmann (2015) we reinterpret the traditionally labeled Topic-prominence parameter in the following terms: the Topic Projection's functional head can be filled via External Merge and not exclusively be filled by movement (i.e. Internal Merge). We will namely leverage this property to uncover the neural patters associated to the different syntactic derivations of Chinese Topic-comment articulations, i.e. A-bar movement versus Base-Generation, as demonstrated in chapter 3 (§3.4.2).

Overt-covert dependency links The second main issue we will address is linked to how the brain achieves and represent the different dependency relations within the sentence-unit. Comparing three different 'anaphoric' strategies between Topic and Comment, we sought to uncover how the different syntactic properties of these +/- overt and +/- pronominal linguistic means to settle a dependency relation between two syntactic elements for reference assignment are represented by the brain. In this regard, the fact that Mandarin easily shows object-drop and an animacy-based rule to block resumption strategy in case of Topicalized inanimate object (while it allows optional resumptives in case of [+animate] object Topics), are two essential characteristics to test for the neural underpinnings of the dependency links respectively achieved by resumptives, gaps or null pronouns.

5. As we saw in chapter 5, to achieve an online structure building the receiver of the linguistic stimulus has to build a syntactic hierarchy and to assign different roles to the Subject and the Topic of the sentence, by essentially relying on linear word-order cues.

Moreover, Chinese allows to test for the absence of grammatical dependency-links between Topic and Comment, in that 'Chinese style' Topics feature a particular syntactic configuration where the relation between Topic and Comment is established in absence of selectional relationship with the main verb, and in absence of syntactic (and sometimes even direct semantic) bounds between the Topic and Comment-clause.

Chinese Topic: a dangling element in sentence-initial position and its link to Discourse information This interesting property of Mandarin Topic-Comment allowing the first element to have no grammatical explicit link with the Comment, and no selectional relationship with the main verb, is shown in the following English example -"*Speaking of David, what has Jannie been up to, lately?*". These types of sentences find its extreme expression in Mandarin Chinese, where the very minimal overt grammatical marking of word-order linearization is the sole responsible not only for being a syntactic cue of the hierarchy between Topic and Comment clause (i.e. the previous sentence is licit even without 'speaking of' introducing the topic), but it also guides the on-line building of an interface with the discourse information, to achieve the understanding of the link between "*David and Jannie*".[6]

The Sentence-Discourse Interface and its syntactic encoding In this regard, our questioning about sentence structure was enriched by taking into consideration the linguistic phenomena that witness of an interface between sentence and discourse linguistic levels of analysis. Far from any hasty pragmatic interpretation of these interfacial linguistic phenomena, what we identified in chapter 2 (cf. cartographic project) as *syntactically encoded* property establishing a link between the sentence-unit and discourse, is now at the heart of our experimental research on the cerebral representation of syntactic complexity of the Chinese Left-Periphery.

We expect the various levels of the sentence structure representations to be neurally represented in the Sentence Network[7]. And anticipating on the next section presenting our Experimental conditions, we can say that we will capitalized on the fact the grammatical system of Chinese offers a unique opportunity to directly contrast two sentence constructions sharing (1) the same content words, (2) the same propositional meaning, (3) a comparable degree of surface complexity, and (4) are both considered to be basic and unmarked structures in the linguistic system, but crucially differ in their syntactic hierarchy: one activating the sentence-discourse interface by the presence of a Topic and the other not.

In sum, the fundamental question of this doctoral work about '*how the mind and the brain achieve sentence hierarchical structure building and the representation of its complexity*', is interpreted in the present fMRI study on Mandarin Chinese by testing for the cerebral encoding of syntactic-layers in Mandarin Left-Periphery, along the following four axes:

6. This aspect, was shown in chapter 5 to imply during on-line sentence comprehension the reiterated *reviving* of all the contextual information about the topic ('David') to understand the content of the Comment.

7. For a comparison of Chinese and French Sentence Networks see the Annexes on Localizers H, p.949

- The issue of the neural underpinning of Sentences domains, to pinpoint the syntactic complexity parameter linked to the presence of linguistic elements in the layer dedicated to the sentence interface with discourse (cf. §2.4.4)
- The question about neuronal correlates of the linguistic distinction between Base-Generated Topics versus Left-dislocated Topics though wh-movement?
- The question relative to the presence of brain areas that selectively respond to syntactic gaps -empty categories trace and *pro*- compared to a resumptive to achieve different kinds of dependency-links with the the Comment-clause.
- and the issue of the existence of separable sub-systems in the Sentence Network dedicated to when the syntactic information is conveyed by word-order strategy or by explicit morpho-syntactic marking

In the following section, we detail the syntactic parameters we manipulated in our fMRI experimental design, and present how our experimental stimuli permit to answer to the state of the art by trying to disentangle the determinants of activation of the sentence's cerebral network to the different dimension of movement-related complexity effects.

7.2 Conditions and Experimental Hypotheses

Overview of the Experimental Design

Building on fine-grained linguistic analyses presented in chapters 2 and 3 Both chapter 2 and chapter 3 are essential to ground the present fMRI study. While chapter 2 discusses and describes the theoretical issues at stake when investigating the cerebral bases of the different syntactic phenomena we selected as the object of our neuro-linguistic research (i.e. syntactic movement transformations in French and Topic-comment articulations in Chinese), chapter 3 clarified the detailed syntactic account of the experimental conditions we retained for our fMRI design on Chinese.

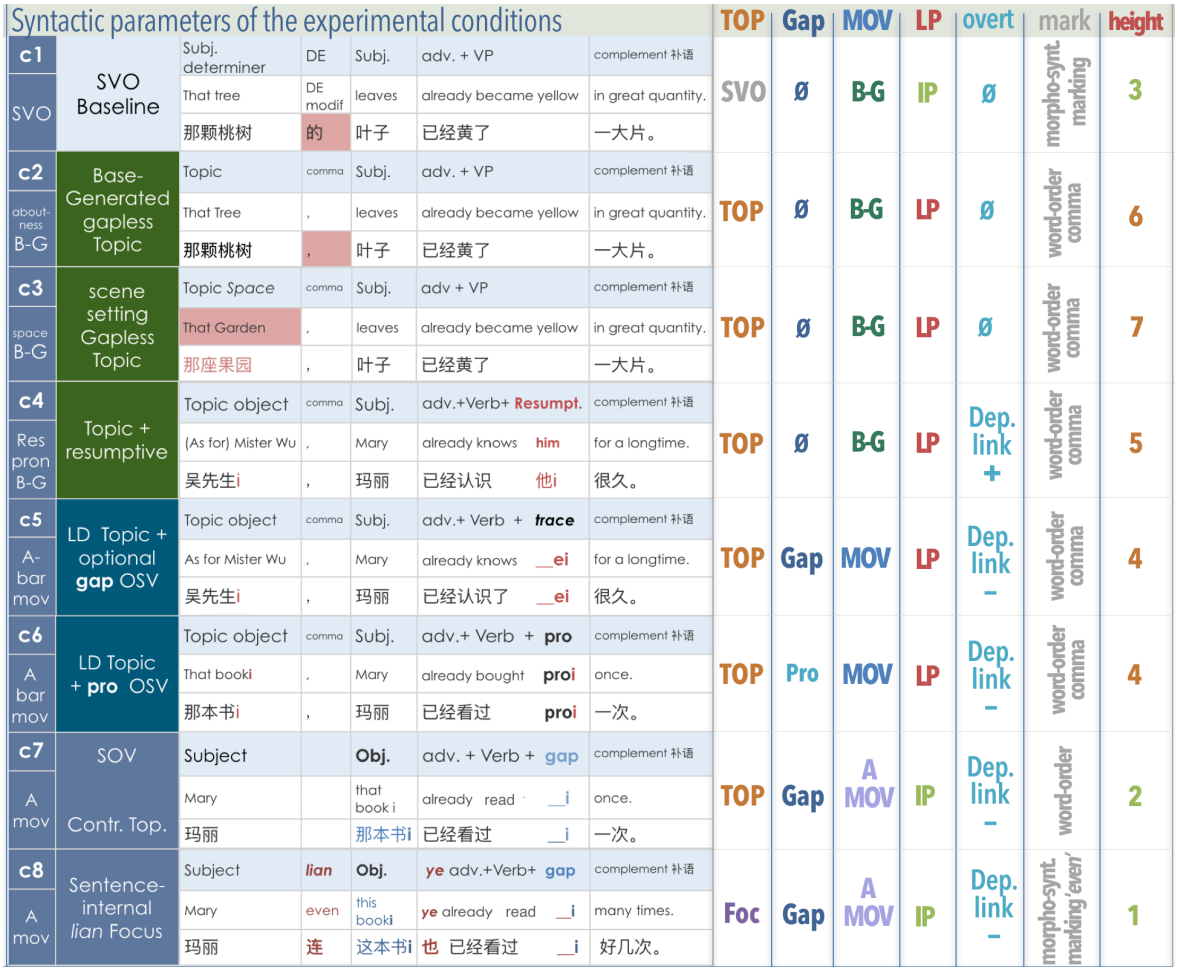
Previous typological and formal research on the syntax of Mandarin Chinese Left-Periphery (see Topic and Focus literature: Badan 2008; Badan and Del Gobbo, 2011 and 2015; Paul, 2002, 2005, 2015; Shyu, 1995 and 2001, and others) offers rich testing hypothesis to focus on the cerebral processing and encoding of the syntactic layers present in different constructions: Topic-Comment articulation, Focus and contrastive pre-posed objects.

Leveraging these fine-grained linguistic analyses, this fMRI experimental design aims at discovering how the brain represents different syntactic structures using linguistic minimal pairs with very constrained interpretative and semantic difference.

The manipulated parameters presented in the right columns of Figure 7.2, will enable us to observe the brain activations related to: (i) Topic-Comment syntactic structures in sentences with no movement analysis (i.e. conditions (c2), (c3), (c4)), (ii) Left-Dislocation through A-bar Movement targeting the Left-Periphery of the sentence (i.e. conditions (c6) and (c5): Topicalization), and (iii) cases of A-Movement to a sentence internal position for pre-posed object sentences (i.e. (c7) condition or *even*-Focus in condition (c8)).

In this section (7.2), we will first present the experimental contrasts and briefly recapitulate the linguistic analyses that have preempted to the choice of the set of experimental conditions. Then, Experimental Method and Procedure, Analyses and Results

will be reported in section 7.3. Finally, brain imaging results are discussed one by one in section 7.4, followed by some brief conclusive remarks and further research ideas (7.5).



Syntactic parameters of the experimental conditions							TOP	Gap	MOV	LP	overt	mark	height
c1 SVO	SVO Baseline	Subj. determiner	DE	Subj.	adv. + VP	complement 补语	SVO	ø	B-G	IP	ø	morpho-synt. marking	3
		That tree	DE modif	leaves	already became yellow	in great quantity.							
		那颗桃树	的	叶子	已经黄了	一大片。							
c2 about-ness B-G	Base-Generated gapless Topic	Topic	comma	Subj.	adv. + VP	complement 补语	TOP	ø	B-G	LP	ø	word-order comma	6
		That Tree	,	leaves	already became yellow	in great quantity.							
		那颗桃树	,	叶子	已经黄了	一大片。							
c3 space B-G	scene setting Gapless Topic	Topic *Space*	comma	Subj.	adv + VP	complement 补语	TOP	ø	B-G	LP	ø	word-order comma	7
		That Garden	,	leaves	already became yellow	in great quantity.							
		那座果园	,	叶子	已经黄了	一大片。							
c4 Res pron B-G	Topic + resumptive	Topic object	comma	Subj.	adv.+Verb+ **Resumpt.**	complement 补语	TOP	ø	B-G	LP	Dep. link +	word-order comma	5
		(As for) Mister Wu	,	Mary	already knows **him**	for a longtime.							
		吴先生i	,	玛丽	已经认识 **他i**	很久。							
c5 A-bar mov	LD Topic + optional **gap** OSV	Topic object	comma	Subj.	adv.+ Verb + ***trace***	complement 补语	TOP	Gap	MOV	LP	Dep. link –	word-order comma	4
		As for Mister Wu	,	Mary	already knows **__ei**	for a longtime.							
		吴先生i	,	玛丽	已经认识了 **__ei**	很久。							
c6 A bar mov	LD Topic + **pro** OSV	Topic object	comma	Subj.	adv.+ Verb + **pro**	complement 补语	TOP	Pro	MOV	LP	Dep. link –	word-order comma	4
		That booki	,	Mary	already bought **proi**	once.							
		那本书i	,	玛丽	已经看过 **proi**	一次。							
c7 A mov	SOV Contr. Top.	Subject		**Obj.**	adv. + Verb + **gap**	complement 补语	TOP	Gap	A MOV	IP	Dep. link –	word-order	2
		Mary		that book i	already read **__i**	once.							
		玛丽		那本书i	已经看过 **__i**	一次。							
c8 A mov	Sentence-internal *lian* Focus	Subject	***lian***	**Obj.**	***ye*** adv.+Verb+ **gap**	complement 补语	Foc	Gap	A MOV	IP	Dep. link –	morpho-synt. marking *'even'*	1
		Mary	even	this booki	***ye*** already read **__i**	many times.							
		玛丽	连	这本书i	**也** 已经看过 **__i**	好几次。							

Figure 7.2 – Experimental conditions of the fMRI Experiment on Chinese Left-Periphery, directly contrasting the brain activation related to these 8 conditions. The *Right columns* present the manipulated parameters: TOP: stands for presence or absence of a Topic; GAP: stands for presence or absence of gap as marking a dependency-link between Topic and Comment; MOV: characterizes the syntactic derivation of each condition LP: stands for presence or absence of an element in the Left-Periphery; and OVERT column: notes the +/- overt feature of dependency-links between sentential elements or their absence. Mark: notes the cues on which sentence hierarchy is built. Height: quantifies the relative height (from Low-Periphery to topmost position in the CP-layer) in the mapping of the Left-periphery of the realized functional projection of each condition.

7.2.1 'Chinese-style' Topics height in the Tree and derivation: Contrasts and predictions

7.2.1.1 a- Aboutness Topic and SVO: two ways to structure the Chinese sentence

							TOP	Gap	MOV	LP	overt	mark	height
c1	SVO Baseline	Subj. determiner	DE	Subj.	adv. + VP	complement 补语	SVO	ø	B-G	IP	ø	morpho-synt. marking	3
SVO		That tree	DE modif	leaves	already became yellow	in great quantity.							
		那颗桃树	的	叶子	已经黄了	一大片。							
c2	Base-Generated gapless Topic	Topic	comma	Subj.	adv. + VP	complement 补语	TOP	ø	B-G	LP	ø	word-order comma	6
about-ness B-G		That Tree	,	leaves	already became yellow	in great quantity.							
		那颗桃树	,	叶子	已经黄了	一大片。							

Figure 7.3 – Comparing Aboutness Topic-Comment articulation in (c2) and (c1) SVO baseline.

Chinese-style Topic and their basicness The very central feature of Aboutness Topics we built on, is that the nominal in Topic position is not related to any position in the sentence as there is no gap or pronoun linked to it, in that the Topic is not sub-categorized by the verb. Additionally, the part-whole semantic relation that can be achieved between Topic (NP1) and Subject (NP2) makes it possible to transform the first NP into the determiner of the second, just by inserting the particle *de* 'of' beween the two, as illustrated by conditions (c1) and (c2) in figure 7.3.

Contrasting (c1) and (c2) conditions, we can appreciate how Mandarin Chinese offers the possibility to contrast sentences constructions that minimally differ in their surface structure and that additionally convey the same propositional content. We will thus be able to compare the processing of sentences with and without a Topic element in the CP layer[8], as illustrated by the Tree-diagrams in Figure 7.4.

Moreover, as highlighted above, from a more formal point of view, in this T-C articulation, the Topic projection's functional head is not filled by movement, but via another syntactic process, namely External Merge (i.e. Base-Generation). Hence, both constructions feature the External Merge of the first NP, but in two different positions in the sentence hierarchy, respectively the determiner of the subject in clause-internal position for (c1) condition and the Specifier of the Topic projection in the CP layer for (c2) condition.

To this *syntactic basicness* of the Aboutness Topic in (c2), we should add that the prevalence of Topic-Comment structures is a typological distinctive feature of Mandarin Chinese which makes such sentence articulation a basic and unmarked sentence structure. This has the advantage of making (c1) and (c2) two equally basic utterances. It has namely been shown that young Chinese children in the early stages of syntax acquisition have the ability to distinguish between the notions of Subject and Topic (Chien, 1983, cf. §3.2.4).

We assume therefore that the comparison with canonical SVO sentence will reveal the

8. In Aboutness Topics (c2) the Topic role is attributed to a sentence-initial constituent on word-order grounds, and NP1- the Topic of the sentence- is not reducible to an implicit modifier of the subject (where *de* 'of' would have been dropped) because of the presence of the comma (cf. discussion on Double subject constructions' derivation in Chinese §3.2.3.3, p. 315).

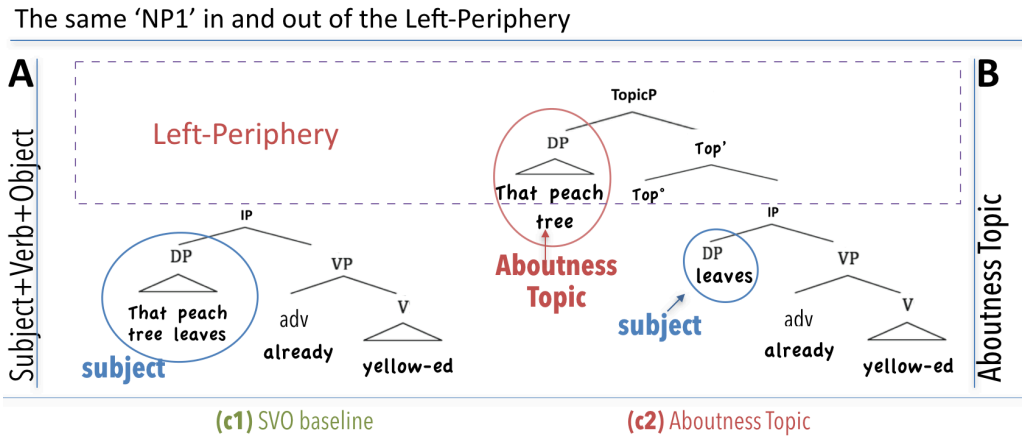


Figure 7.4 – Syntactic-tree representation of SVO-baseline and Aboutenss Topic, the same 'NP1' is in the Left-Periphery or in clause internal position. (A) Experimental baseline condition (c1) Subject+Verb+Object. (B) Experimental condition (c2) for Aboutness Base-generated Topic.

syntactic architecture of Topic-comment in absence of movement and of any contrastive interpretation, but that some effect reflecting that the sentence discourse interface layer is activated by the presence of the Topic element would nonetheless still be observable. However, it should be said that the basic status that Aboutness Topics have in the Chinese linguistic system, might also have the drawback of eliciting only weak activation.

Predictions Aboutness B-G Topics c2 > c1 SVO Given the basic character of Topic-comment constructions in the linguistics system of Chinese and the absence of Movement related-processes, one of our first predictions for the contrast [(c2) > (c1)], is to find only a focal activation and not to find a very broad effect.

What we mainly expect to observe is an activation linked to the representation of the syntactic position in the Left-Periphery where the Topic element *nà-kē táoshù* 那棵桃树 'this peach tree' is hosted (i.e. Base-Generated), as (c1) does not have any element in the sentence-discourse inter-facial layer as illustrated in the above Tree-diagrams (figure 7.4).

For this complexity dimension, we predict the involvement of Broca complex, as expected from neuro-psychological and neuro-imaging. And as predicted by the two hypotheses we lean on – the Tree-pruning Hypothesis and Sentence-Domain Hypothesis –, which state that Broca Complex is linked to the complexity of representing an element in the CP-layer.

For our prediction, we partly relay on the findings of Shetreet and Friedmann (2014), who opposed Hebrew Topicalization (A-bar movement) to SVO declaratives, and reported bilateral Broca Complex (BA45 left, BA47 right), the left Pre-Central Gyrus (BA6), bilateral pSTG and cerebellar activation. We expect to find a sub-part of the regions reported by Shetreet and Collegues because of the absence of movement-related processes in our contrast.

Although Aboutness Topic is free of any contrastive interpretation, we still expect to find some effects reflecting the activation of the sentence-discourse interface layer.

We predict, following Shetreet and Friedmann, that the right IFG[9], which is recurrently reported in wh-movement contrasts – and was also present in the study of Ben-Shachar et al., (2004) on Topicalization – would be observed to sub-serve the discourse properties of Topic-Comment articulation.

Importantly, thanks to Topic-comment linearization rule and the head-final parameter (i.e. imposing the order Derterminer-Determined) distinguishing Chinese[10], we obtain the same word-order in the two conditions, thus avoiding any possible Working-memory related interpretation of the activation patterns that may be observed in Broca Complex. In fact, as put forward in much psycho-linguistic and ERP literature, and in an interesting article on working-memory related issues in Binding of reflexives and Movement process by Santi and Grodzinsky (2007), the on-line linking of *fillers* and *gaps* taxes Working Memory resources[11].

7.2.1.2 b- Frame-Topic versus Aboutness Topic: two different Topic functions

Aboutness Topics differ remarkably from other Topic types, in that the Comment as a whole 'is about' the Topic, and their absence of gap is not only interesting from the point of view of establishing a syntactic derivation, but it also shows that Topic and Comment can establish an 'aboutness' or a Fame-setting relation rather than an antecedent-gap relation. Specifically, from the point of view of cognitive processing of the sentence the Aboutness or Frame-setting relations encompass the whole Comment-clause and are not resolved by encountering a given element in the Comment-clause and no direct intra-sentential dependency-link is in fact to be established.

Syntactic parameters of the experimental conditions							TOP	Gap	MOV	LP	overt	mark	height
c2	Base-Generated gapless Topic	Topic	comma	Subj.	adv. + VP	complement 补语							
about-ness B-G		That Tree	,	leaves	already became yellow	in great quantity.	TOP	ø	B-G	LP	ø	word-order comma	6
		那颗桃树	,	叶子	已经黄了	一大片。							
c3	scene setting Gapless Topic	Topic *Space*	comma	Subj.	adv + VP	complement 补语							
space B-G		That Garden	,	leaves	already became yellow	in great quantity.	TOP	ø	B-G	LP	ø	word-order comma	7
		那座果园	,	叶子	已经黄了	一大片。							

Figure 7.5 – Experimental condition (c2) Aboutness Base-generated Topic and (c3) Spatial Scene-Setting Base-generated Topic.

9. Note that Shetreet et al. argue for the involvement of this area because of its involvement in theory of mind and discourse functions (citing Kobayashi, 2008; Kuperberg et al., 2006; Langleben et al., 2005), and underline that the formulation of Topic-comment articulations imply taking into account the mental state of the hearer, and consequently their understanding should imply the processing of the communicative intention of the speakers.

10. Note that this configuration is unique to the so-called double subject constructions, because of the part-whole relations the two sentence-initial Noun Phases feature. Yet, Topicalization contrasts present in the literature like Ben-Shachar et al. (2004) and Shetreet et al. (2014) cannot avid the confound of working-memory interpretation.

11. The authors investigated namely filler-gap dependency relations by varying the distance between the dependent elements within each of the constructions to test for an effect of WM demands in Broca area. Their findings revealed a positive linear effect of distance for all types of dependencies (Binding and Movement ones) only in bilateral parahippocampal/ fusiform gyri, while Broca area (BA45) showed a positive linear effect only for Filler–Gaps relations derived by movement

Contrasting the first minimal pair (c1) and (c2) (7.3), we aim at giving experimental evidence for the difference between Topic-Comment and Subject-Predicate relation, whereas the contrast between (c3) and (c2) illustrated in 7.5 will possibly reveal the difference between Aboutness and Frame-setting function of the Topic.

The issue of the semantic content of the Topic element and its the Frame-setting function is addressed by comparing brain activation for Scene-Setting Topic condition (c3) where the first DP plays the role of a Frame localizing the following Comment clause, with (c2) condition, which differs minimally in terms of lexical material.

Predictions Abountess Topic c3 > c2 Frame-Setting Topic We expect that the lexical difference between the two conditions would yield more mid-temporal activation, possibly reflecting either the simple lexical difference between the two sentence initial Noun Phrase or reflecting the increased semantic distance between Topic and Subject. While this low-level lexical predictions are quite evident, precise predictions for the locus of activation of the Frame-setting function are more difficult to come about.

Theoretically another effect could be observed in this contrast, and it is linked to the fact these two types of Topics occupy two different positions in the Left-Periphery. As illustrated by the tree-diagrams in 7.6, Aboutness Topics are hosted in a lower Functional Projection in the syntactic-tree skeleton.

Hence, if our experimental design is sensitive enough to allow us to detect such subtle differences, we may find an effect linked to the complexity of representing the higher position where Scene-setting Topics are realized compared to Aboutness ones. As For the localization of this effect we remain cautious by actually not advancing any, if not the classical Broca Complex activation observed in fast every complexity manipulation.

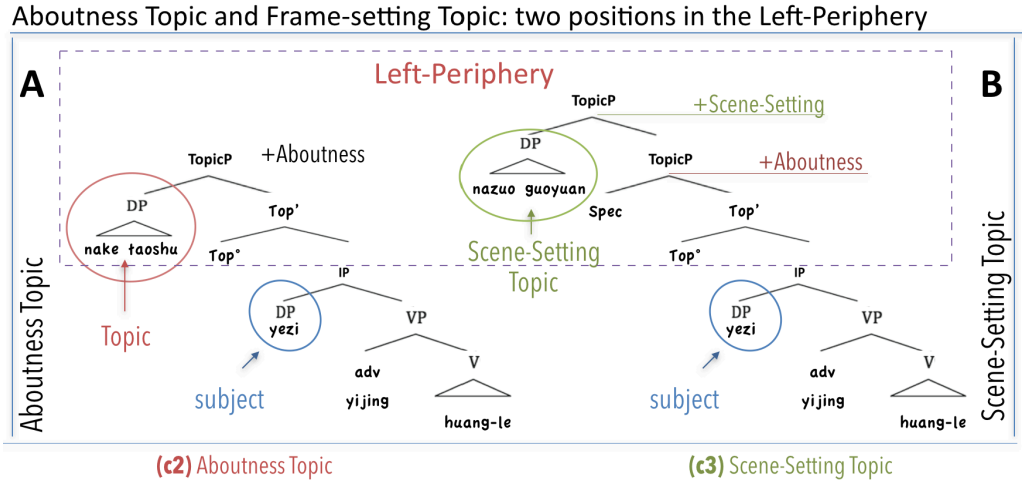


Figure 7.6 – Syntactic-tree representation of Aboutenss Topic and Frame-setting Topic.(A) Experimental condition (c2) Aboutness Base-generated Topic. (B) Experimental condition (c3) Spatial Scene-Setting Base-generated Topic.

7.2.2 Different dependency-links in the sentence

In addition to the previous contrasts, this experiment also attempts to tackle the broader question of the cerebral encoding of different types of dependency-links inside the sentence-unit.

While in chapter 5, the online incremental mechanisms of Gapless Scene-setting Topic-Comment sentences have been investigated using a high-temporal resolution imaging method (EEG-Electro–Encephalography), here the spatial resolution offered by fMRI imaging technique will help us to uncover the cerebral underpinnings for expressing the dependency between words in Gapped and Gapless Topic constructions and in Topic-comment articulations with a resumptive referring back to the Topic.

In this regard, Topic-comment constructions in Chinese appear once more to be an ideal testing-ground to study the neural implementation of different dependency-links in sentence constructions that have or not undergone syntactic movement transformations.

Importantly, dependency relations can be achieved by overt or covert linguistic means, and Chinese resumption strategies offer a particular configuration where animacy gives rise to a more or less overt realization of the object constituent in the Comment clause (cf. §3.4.4), simultaneously preserving the same surface word-order, as illustrated in the three experimental conditions we selected to this effect in Figure 7.7.

Syntactic parameters of the (c4), (c5) and (c6) experimental conditions								TOP	Gap	MOV	sem	overt
c4	Topic + resumptive	Topic object	comma	Subj.	adv.+Verb+ Resumpt.	complement 补语	Overt Resumptive Pronoun	TOP LP	ø RP	B-G	ANI	Dep. link +
Res pron B-G		(As for) Mister Wu	,	Mary	already knows him	for a longtime.	Base-Generated Topic + RP					
		吴先生i	,	玛丽	已经认识 他i	很久。	Hanging Topic					
c5	LD Topic + optional **gap** OSV	Topic object	comma	Subj.	adv.+ Verb + ***trace***	complement 补语	Optional Resumptive Pronoun	TOP LP	Gap	MOV	ANI	Dep. link –
A-bar mov		As for Mister Wu	,	Mary	already knows __ei	for a longtime.	Left-Dislocated Topic Abar Movement					
		吴先生i	,	玛丽	已经认识了 __ei	很久。	Island test					
c6	LD Topic + **pro** OSV	Topic object	comma	Subj.	adv.+ Verb + **pro**	complement 补语	Topicalization targeting CP	TOP LP	*pro*	MOV	INA	Dep. link –
A bar mov		That booki	,	Mary	already read proi	once.	Left-Dislocated Topic Abar Movement					
		那本书i	,	玛丽	已经看过 proi	一次。	Resumption is agrammatical					

Figure 7.7 – The three experimental conditions where dependency relations are achieved by Overt-Covert linguistic means between Topic-Comment.

The above examples show that, in the overt and covert realization of Topics anaphoras, Mandarin Chinese differs from French or Italian, where resumption is overtly realized as a clitic. In fact, when there is no island, a gap and a Resumtive Pronoun can be in free alternation in Topic-comment structures, as in our experimental conditions (c4) and (c5). In (c6), the inanimate lexical feature of the Topic imposes an obligatory null pronominal forms, and the inanimate object pronoun remains phonetically covert as a A-bar bound *pro*[12].

As illustrated in chapter 2 (§2.4.3), all human languages encode syntactic information in the form of overt and covert elements, whose interpretation relies on previously encountered elements in the sentence. However, a crucial syntactic distinction should be made between the dependency relations which result from movement operations (i.e. gaps

12. For the series of diagnostic tests on which this syntactic analysis is based please refer to chapter 3, §3.4.4 and 3.4.4.2, p.390.

and covert-dependencies, like null pronouns), and the dependency operations involving pure reference (pronouns, overt-dependencies), like in pronominal reference assignment.

In the selected conditions, two dependency types are to be distinguished, in (c5) and (c6) the dependency-link is established between the Topic and an empty or phonologically covert syntactic element yielding a filler-gap dependency, while in (c4) another linking process is establish between the Resumptive pronoun and its antecedent, the Base-generated Topic in sentence-initial position[13].

Finally, the fact that 'Chinese style' Topics feature a particular syntactic configuration where the relation between Topic and Comment can be established in absence of selectional relation with the main verb like in the Aboutness (c2) and Scene-setting Topic (c3) present above, which will actually constitute an additional baseline for the above three conditions (c4), (c5) and (c6), and will allow to test for the effects of the absence of grammatical dependency-links between Topic and Comment.

Given this ideal configuration, the different contrasts we plan to test for are summarized in Figure 7.8 and will tentatively allow to observe the seven different effects illustrated. In the following we will detail our prediction for each of them.

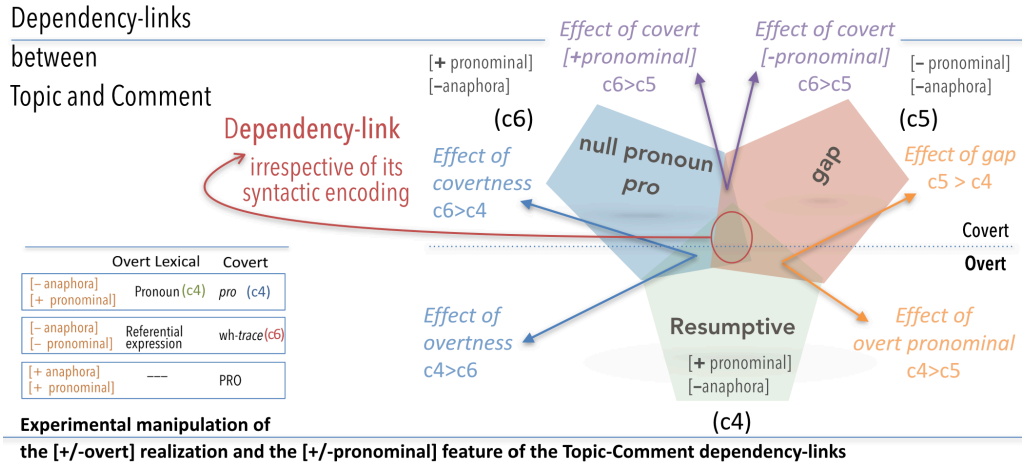


Figure 7.8 – Experimental manipulation of the [+/-overt] realization and the [+/-pronominal] feature of the Topic-Comment dependency-links.

7.2.2.1 Predictions on different dependency-links effects

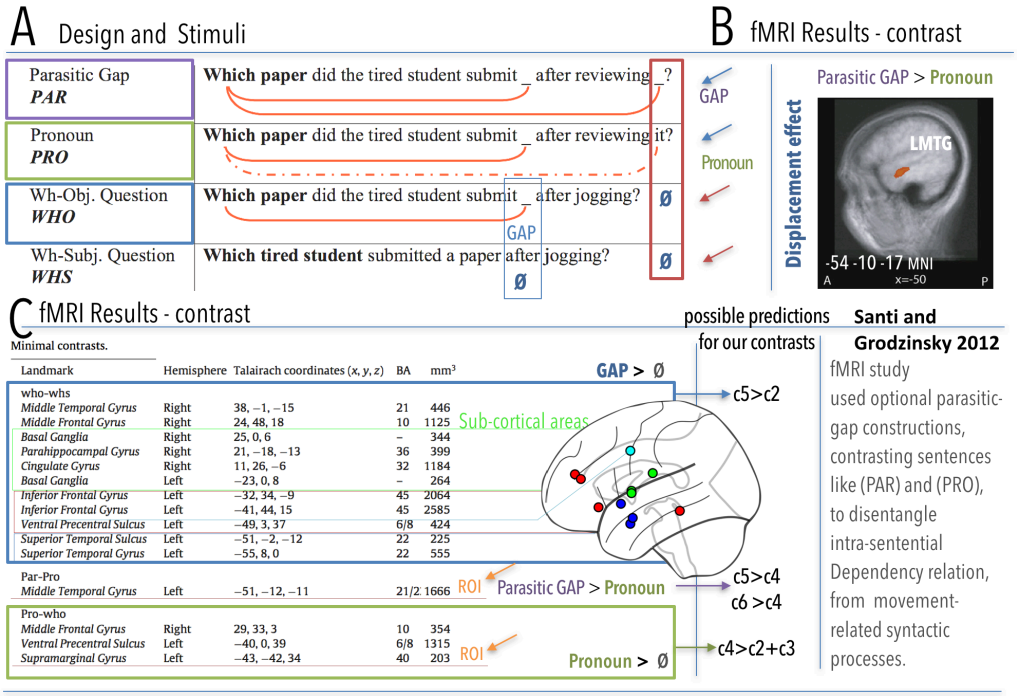
The direct contrast between these three conditions will possibly reveal what are the neural underpinnings of the rules that were revealed by the series of syntactic tests that

13. For instance, as attested in section §3.4.4.2 (p.392), Left-Dislocated structures with a gap like (c5) give rise to island and crossover effects, while Topic-comment structures with a Resumptive Pronouns like (c4) do not give rise to any of these effects, a pattern that is supporting the syntactic analysis stating that these two sentence structures are derivationally distinct from one another. Importantly, in Mandarin Chinese, when there is no island, a gap and a Resumptive Pronoun can be in free alternation in Topic-comment structures, as in our experimental conditions (c4) and (c5). In chapter 3, we saw that Topics that are moved and leave a gap are in fact sensitive to island conditions, they do show reconstruction effects, which are generally taken as a diagnostic for movement. Moreover, as argued through the diagnostic tests illustrated in chapter 3 (see p.392) and since Huang (1987), the gap in Topic-comment sentences like condition (c6) is a resumption that remains phonetically covert.

allowed us to distinguish three experimental conditions on the basis of their different syntactic behavior (e.g. island sensitivity, Reconstruction or Crossover effect, etc.).

Although such subtle syntactic differences have been scarcely addressed in the literature, we can rely on a few studies that investigated different sentential dependency-links to formulate some predictions.

The two types of [+/- overt] anaphoric relations have mainly been investigated in their temporal unfolding, but some studies have also attempted to look for the localization of these dependency mechanisms in the fMRI literature like Santi and Grodzinsky (2012), who focused on syntactic aspects linked to the presence of gaps left by syntactic movement, to try to distinguish between the activation patterns linked to the establishment of intra-sentential dependency from that associated to displacement through wh-movement.



Landmark	Hemisphere	Talairach coordinates (x, y, z)	BA	mm³
who-whs				
Middle Temporal Gyrus	Right	38, −1, −15	21	446
Middle Frontal Gyrus	Right	24, 48, 18	10	1125
Basal Ganglia	Right	25, 0, 6	–	344
Parahippocampal Gyrus	Right	21, −18, −13	36	399
Cingulate Gyrus	Right	11, 26, −6	32	1184
Basal Ganglia	Left	−23, 0, 8	–	264
Inferior Frontal Gyrus	Left	−32, 34, −9	45	2064
Inferior Frontal Gyrus	Left	−41, 44, 15	45	2585
Ventral Precentral Sulcus	Left	−49, 3, 37	6/8	424
Superior Temporal Sulcus	Left	−51, −2, −12	22	225
Superior Temporal Gyrus	Left	−55, 8, 0	22	555
Par-Pro				
Middle Temporal Gyrus	Left	−51, −12, −11	21/2	1666
Pro-who				
Middle Frontal Gyrus	Right	29, 33, 3	10	354
Ventral Precentral Sulcus	Left	−40, 0, 39	6/8	1315
Supramarginal Gyrus	Left	−43, −42, 34	40	203

Figure 7.9 – (A)Examples of experimental conditions. (B) Whole-brain effect for the contrasts opposing [Parasitic gap > Overt Optional Pronoun]. (C) Table of results of the minimal contrast in the study and possible parallels with our experimental contrasts. The glass-brain in the middle represents the location of the underlined peaks on the table. Adapted from Santi and Grodzinsky (2012).

As illustrated in figure 7.9, this study contrasted the optional presence of a parasitic pronoun and revealed a single cluster of activation for the presence of a Parasitic Gap in the left Middle Temporal Gyrus, that the authors call 'displacement effect', because of the different derivation of gapped structure compared to the Resumptive one.

Yet, a second fMRI study offers some different hypotheses about the locus of the effects we could observe by contrasting gapped structures versus resumptive ones. The fMRI study by Matchin et al. (2014) offers namely some insights on the distinction

between the establishment of dependency-link based on empty versus lexical categories in the sentence. The authors contrasted English sentences featuring a gap in a wh-question, and a pronoun in Backward Anaphora. As shown in figure 7.10(A) these contrasts present an interesting temporal/frontal division of labor between gap strategy and presence of an anaphoric pronoun. Backward Anaphora compared to *wh*-gap shows a bilateral temporal activation with an anterior and a posterior cluster, while the reverse contrast reveals a unique cluster in Precentral Gyrus for wh-questions gaps effect.

Comparing Resumptives and Gaps: predictions

Hence, given these two different findings, the very minimal contrast between [c5 > c4] – allowing to observe the difference between a co-reference link achieved by means of a covert linguistic mean like the gap in (c5) versus an overt Resumptive in a Base-Generated Topic like (c4) – , will alternatively reveal a cluster of activation in left Precentral Gyrus following Matchin et al. (2014) or a cluster in left Middle Temporal Gyrus as reported by Santi and Grodzinsky (2012).

Predictions for the opposite contrast [c4 > c5] revealing a Resumptive Pronoun effect are more straightforward, as both the pronoun effect reported by Santi and Grodzinsky (2012), and the Backward Anaphora effect by Matchin et al. (2014) indicate the involvement of a posterior cluster encompassing the Angular Gyrus for the second and the Supramarginal Gyrus for the first.

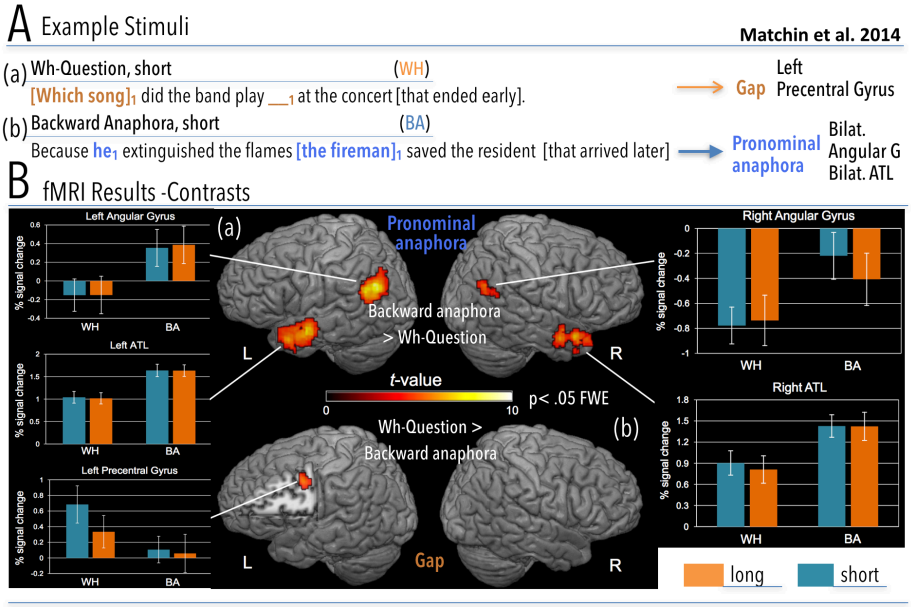


Figure 7.10 – (A) Experimental stimuli examples. (B) Group averaged brain maps results for the contrast between Backward anaphora > Wh-Question in (a), and Wh-Question > Backward anaphora in (b). Barplots report average percent signal change for each condition within selected clusters of activation. Adapted from Matchin et al. (2014).

However, based on Matchin et al. (2014) and on other studies, a more general pattern could emerge from the series of contrasts we planned to do to explore the issue intra-

sentential dependency links. We may expect to find a more antero-/posterio-temporal activation for the establishment of a dependency-link between overt syntactic elements in an antecedent-pronoun relation, and a stronger activation in Broca complex (and the adjacent Precentral Gyrus) for the establishment of a co-referential link in a syntactic configuration, where a silent gap has to be posited.

This last prediction is coherent with other findings from Santi and Grodzinsky's work, who reported (1) activation in Broca's area (pars triangularis/BA45) for filler-gap dependencies generated by movement and not for canonical anaphoric dependencies (2007a), and (2) the involvement of precentral Gyrus (BA6) and Broca BA44 in presence of syntactic gaps (as the result of adaptation to both movement and embedding, cf. page 165, in Santi and Grodzinsky, 2010).

Comparing Null pronoun and Gap, c5 versus c6

As for the particular configuration where animacy gives rise to a [+/-overt] realization of the object constituent in the Comment clause, preserving the same surface word-order, our predictions could either concentrate on the difference in animacy between (c5) and (c6), or on the subtle difference between these two covert [+/- pronominal] means of achieving co-referential assignment.

Syntactic parameters of the (c5) and (c6) experimental conditions								TOP	Gap	MOV	sem	overt
c5	LD Topic + optional **gap** OSV	Topic object	comma	Subj.	adv.+ Verb + ***trace***	complement 补语	**Optional Resumptive Pronoun**	TOP LP	Gap	MOV	ANI	Dep. link –
A-bar mov		As for Mister Wu	,	Mary	already knows __ei	for a longtime.	Left-Dislocated Topic Abar Movement					
		吴先生i	,	玛丽	已经认识了 __ei	很久。	Island test					
c6	LD Topic + **pro** OSV	Topic object	comma	Subj.	adv.+ Verb + **pro**	complement 补语	**Topicalization targeting CP**	TOP LP	*pro*	MOV	INA	Dep. link –
A bar mov		That booki	,	Mary	already read proi	once.	Left-Dislocated Topic Abar Movement					
		那本书i	,	玛丽	已经看过 proi	一次。	Resumption is agrammatical					

As for animacy, in the contrast [c5 > c6], we could follow Hammer and colleagues (2007/2011) who reported an effect of animacy linearization, specifically located in the anterior Superior Temporal Gyrus (aSTG), where Animate antecedent compared to Inanimate ones yielded an increased activation (see Figure 7.11).

For the second aspect concerning the nature of null pronoun in (c6), the predictions are more difficult, as we can only rely on an ERP study by Yang and Liu (2014), showing that a negativity component is elicited at the verb reflecting the retrieval of the Topic from working memory in a Topic-Comment structure like the one in our condition (c6):

(260) 桌子$_i$ 经理踢了 *pro*$_i$ 两脚。

Zhuōzi$_i$ jīnglǐ tī-le ___$_i$ liǎng-jiǎo.
table$_i$ manager kick-ASP. *pro*$_i$ two-CL.foot

'The table$_i$, the manager kicked *pro*$_i$ twice.' Yang and Liu, (2014, ex.4)

While for our contrast [c6 *pro* > c5 gap] an increased activation in a brain area that also responds to Pronoun binding could be observed as the *pro* carries the [-anaphora, +pronominal] features, while gaps features are [-anaphora, -pronominal]. Possible candidates could be the brain areas reported in the Backward Anaphora effect by Matchin et al. (2014): Bilateral Angular Gyrus and Bilateral anterior Temporal Lobe.

To recapitulate some of the findings on which we can base our predictions, Figure 7.11 illustrates the main brain areas that were reported in literature as being involved in various kinds of intra-sentential dependency-links.

Given these ERP findings, we expect that a comparable sensitivity to this grammatical restriction linked to animacy would be observable in the contrasts [c6 *pro* > c4 Resumptive], probably in similar area as that reported by Santi and Grodzinsky (2012) for their comparison between Parasitic Gap and Pronoun, namely the Middle Temporal Gyrus.

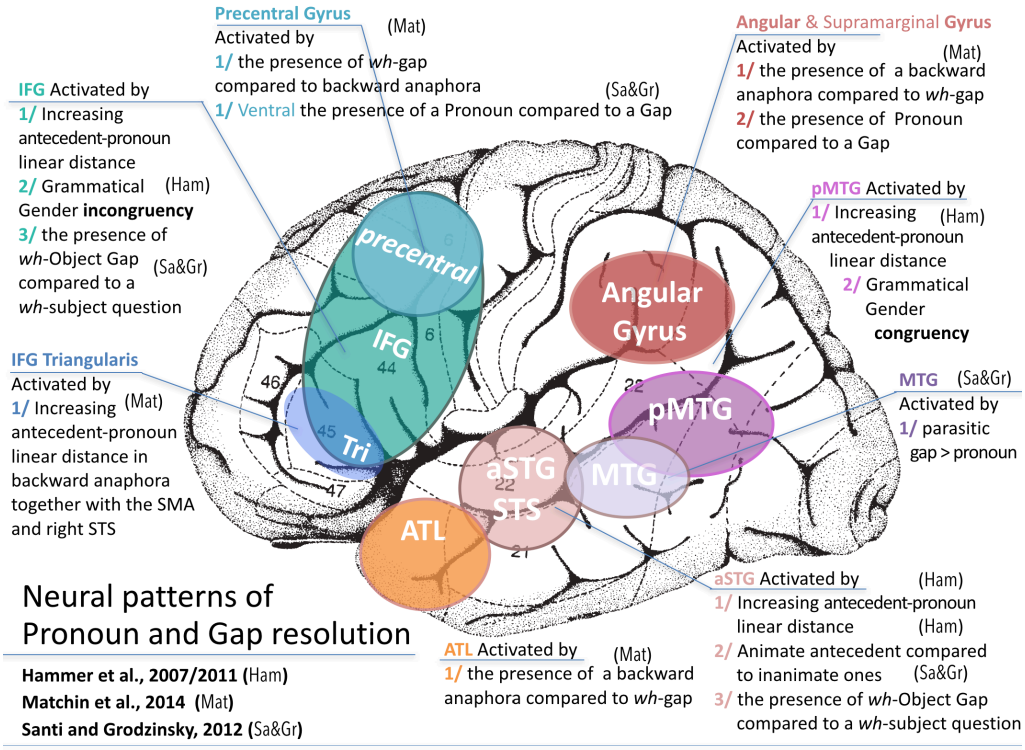


Figure 7.11 – Neural patterns of Pronoun and Gap resolution. Recapitulation of the fMRI findings of Hammer et al., 2007/2011 (i.e. Ham); Matchin et al., 2014 (i.e. Mat); Santi and Grodzinsky, 2012 (i.e. Sa& Gr).

7.2.3 Focus and contrastive Topic conditions

Contrastiveness and Focus

Chinese Left-Periphery properties present also the opportunity to obtain some discourse related interpretations like contrastivness and focus without moving constituents to sentence-initial position and to the CP layer, like in our conditions (c8) sand (c7).

Our aim in comparing the contrastive interpretation obtained by simple word-order in bare pre-posed object (c7)[14] with the Focus interpretation obtained by *lian even*-Focus

14. In Mandarin Chinese the direct object moved to a preverbal position can be preceded, obligatorily

in the same pre-verbal position (c8) is threefold.

Our first (1) interest is linked to the presence of an overt syntactic functional head marking compared to word-order strategy to achieve pragmatic markedness. We selected this configuration to help us uncover the neural substrates of sentence structure building based on morpho-syntactic information versus word-order cues. The second interest (2) is to bring to light the pragmatic effects of these two conditions that do not move to Left-Periphery[15]. And thirdly, (3) we want to compare the effects of landing-site in Topicalized object (c6 and c5 A-bar moved to CP) compared to pre-verbal objects in (c7), which is moved clause-internally through A-moved in (i.e. [c6 > c7] and [c5 > c6])[16].

Syntactic parameters of the (c7) and (c8) experimental conditions								TOP	Gap	MOV	LP	overt
c7	SOV	Subject		Obj.	adv. + Verb + gap	complement 补语	Movement targeting the IP	TOP contr.	Gap	A MOV	IP	word-order
A mov	Contr. Top.	Mary		that book i	already read __i	once.	Bare preposed-object no movement to the LP					
		玛丽		那本书i	已经看过 __i	一次。	A Movement => no Resumptives					
c8	Sentence-internal *lian* Focus	Subject	*lian*	Obj.	*ye* adv.+Verb+ gap	complement 补语	sentence-internal *lian* Focus	Foc scalar	Gap	A MOV	IP	morpho-synt. marking *ever*
A mov		Mary	even	this booki	*ye* already read __i	many times.	A-Movement					
		玛丽	连	这本书i	也 已经看过 __i	好几次。	Movement targeting sentence-internal position					

Figure 7.12 – *lián*-focus compared to preposed object i.e. Contrastive Topic.

7.2.3.1 Predictions for Morpho-syntactic marking versus Word-order

As illustrated by Figure 7.12 (for tree-diagrams see 7.13), the presence of an overt syntactic functional head *dou* in (c8)[17] and of the function word *lian* 'even', will allow to observe at the brain processing level the effect morpho-syntactic marking versus word-order marking.

Importantly, for sentences like our condition (c7) we assume following Shyu (1995) and Badan (2007) (cf. §3.4.3) an A-movement analysis[18]. Hence, the fact that condition (c7) and (c8) are both analyzed as A-movements, further allows the contrast between them in balanced in terms (for detailed syntactic properties of *lian*-XP within IP see Paris 1979, 1998, 1999; Shyu 1995 and Badan and Del Gobbo, 2015).

Our first prediction for the contrasts [c8 > c7] relays on a study by Allen et al. (2012) who compared the two ways in which English can realize double object verbs like give:

(261) Contrast of Double object constructions

or optionally, by the morpheme *-ba*. The exact function of *-ba* in disposal structures is a very discussed issue among linguists. It has been treated either as a verb (Hashimoto, 1971), as a preposition (Li, 2001), as a Case marker (Huang, 1982) or as a higher verbal head by Paul and Whitman (2001). For an analysis of functions and optionality/obligatoriness of *-ba* see also Li (2006).

15. We can note that Shyu (1995) made a structural distinction between *focused* OSV and unmarked OSV, by stating that the former is in IP-adjoined position, while the latter occupies the Spec, TopicP. Yet, within the cartography framework, several authors actually propose that every kind of Topic in the CP area can optionally have a contrastive reading.

16. In Mandarin the Contrastive Focus cannot (overtly) move up to the Left-Periphery, and it is always in-situ (see Gao 1994; Badan 2007; Badan and Del Gobbo 2010).

17. Note that Badan and Del Gobbo (2015) treat *lian*, following Bayer's (1996) proposal, as focus particle or focusing adverb, and understand it as a *minor heads* taking its modifee as complement.

18. As this movement is clause-bound, and no functional category can intervene between *lian*-XP and *dou* (i.e. they are in a spec-head relation) it is an instance of A-movement.

a. Sally gave Joe a book.
b. Sally gave the book to Joe.

In (261), the two grammatical constructions share the same content words, propositional meaning, and degree of surface complexity, but in the first sentence the roles of the post-verbal NPs are attributed on word-order grounds, while in (b) the object is in its canonical post-verbal position and dative case assignment is realized through overt morpho-syntactic marking, thanks to the preposition '*to*'. The authors report results from their Multi-Voxels Patterns Analysis (MVPA), showing that the activation patterns of two brain areas – anterior BA22 and BA47 – are sufficient to distinguish the two constructions better than the controls and better than chance. Given these findings we predict that Broca pars Orbitalis and/or the Temporal Pole will be observed in the contrast opposing (c8)>(c7).

7.2.3.2 Discourse-related Brain Activations

To these predictions linked to the presence of function words in the sentence, we add that this comparison will allow to observe the different discursive/pragmatic interpretations that are assigned to morpho-syntactically marked and unmarked pre-verbal objects that crucially are not hosted in the Left-Periphery and have both undergone Argumental Movement (A-movement).

An aspect we left aside during our review of the fMRI literature in chapter 2 is directly linked to Discourse semantics effects. This aspect could actually reveal to be one of the main determinants of activation in our contrastive Topic condition (c7) and Focus condition (c8), therefore we will briefly present two studies that brought to light some candidate brain areas for the kind of pragmatic calculations involved in discourse manipulations.

In order to investigate the brain areas sub-serving discourse-related information processing, Menenti et al. (2009) embedded in two types of context, sentences containing or not some world knowledge anomalies, e.g. "Dutch trains are yellow/white and blue." (yellow:correct/white :anomaly). As the authors wanted to test if a sentence that departs from our world knowledge like *The elephant flies* is preceded by a discourse context that supports it like in a circumstance where a circus traveling by airplane, they embedded correct and anomalous sentence in a Neural context that was introducing the topic of the sentence and was in line with the world knowledge as in (262a), Local contexts presented some information that making the deviation from the default world knowledge more acceptable as in (262b).

(262) a. *Neural context:* The Netherlands are famous for their designers. In addition, the Dutch railways have chosen a very conspicuous color scheme, which makes them recognizable everywhere.
   i. "Dutch trains are yellow and blue."
   ii. "Dutch trains are white and blue."
   b. *Local contexts:* The coming world championships are one big national spectacle. The Dutch railways have painted the Dutch flag on their trains.
   i. "Dutch trains are yellow and blue."
   ii. "Dutch trains are white and blue."

As different sensitivity of Right and Left Inferior Frontal Gyri to Discourse and World Knowledge manipulation was observed, the authors concluded that these different contextual calculations recruit partly different sets of brain areas: the right Inferior Frontal Gyrus was identified as more sensitive to the discourse manipulation than the left inferior frontal gyrus[19].

Similarly Kuperberg et al. (2006) observed stronger responses to sentences unrelated to the preceding contexts (i.e. two preceding sentences) than to related ones, in the right inferior frontal gyrus. Yet, the general involvement of the right hemisphere in discourse integration had already been put forward by St George et al. (1999), who reported that cerebral activity for untitled and therefore less coherent paragraphs was more right lateralized compared to titled paragraphs.

Interestingly for our manipulation of Topicality and its cerebral activation patterns, Caplan and Dapretto (2001) studied different types of discourse coherence violation effects in question and answers pairs, contrasting off-topic and illogical anomalies, exemplified in (263).

(263) a. On- or off-topic: *Do you believe in angels?*
   i. on-topic answer: Yes, I have my own special angel.
   ii. off-topic answer: Yes, I like to go to camp.
  b. Logical or illogical: *Do you like having fun?*
   i. logical: Yes, because it makes me happy.
   ii. illogical: No, because it makes me happy.

They found that the effect of discourse anomalies was more left-lateralized for logicality and more right-lateralized for topicality.

Predictions for discourse processes

Given the implication of right IFG in integration of discourse-level information during sentence comprehension reported in these studies and generally in the literature, we hypothesize to observe right Inferior Frontal Gyrus activation in the contrasts involving more pronounced sentence-discourse calculations like (c7) and (c8). In the (c8)>(c7) contrast we expect to find a increased activation in right IFG (rIFG), as an effect of the complex semantic calculation linked to Scalar Focus compared to the simpler one involving only contrastiveness. This being said we also could expect to tap into areas linked to mathematical calculation, as to understand scalar focus one has to realize a complex set-subset operation like in "Even this apartment, we cannot buy".

Moreover, Caplan and Dapretto (2001) findings about off-Topic answers could actually also lead us to predict the involvement of right hemisphere activation in Broca Complex, not only in the contrast between the Left-Dislocated Topicalized sentences and Gapless Topics, but also in the minimal contrast between Aboutness Topic and SVO baseline (c2>c1) we already discussed.

19. Note that the left Angular Gyrus reacted strongly to the degree of discourse coherence between the context and critical sentence.

7.2.3.3 Comparing A-bar movement and A-movement

As shown in Figure 7.13, (c7) and (c8) have a strong pragmatic and contrastive interpretation without being moved to the LP.

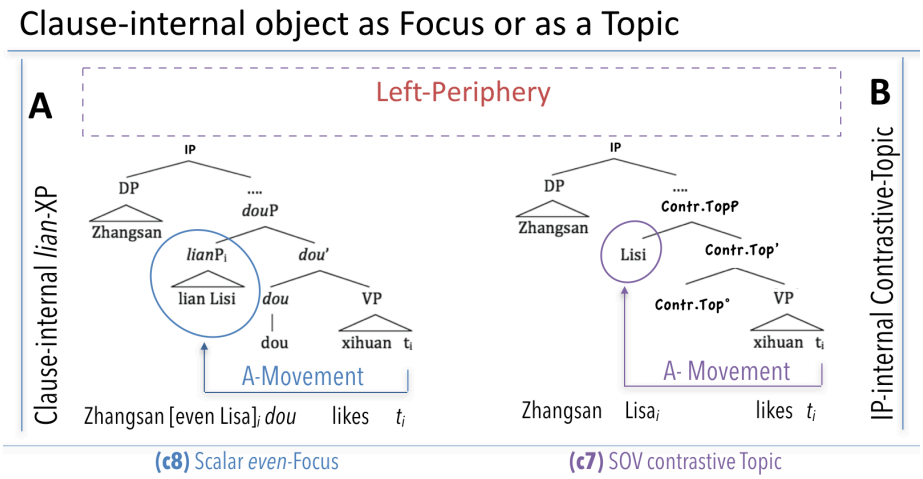


Figure 7.13 – Syntactic-tree representation of clause-internal object as Focus. The object is in the same linear clause internal position, but has two different roles, TopicSOV Contrastive Topic and even-Focus Condition respectively attributed on word-order grounds or through morpho-syntactic marking (*lian…dou*). (A) Experimental baseline condition (c8) Subject+Verb+Object. (B) Experimental condition (c7) SOV Contrastive Topic.

And the second is linked to the fact this property allows the de-correlation of pragmatic and contrastive interpretations usually linked to a position in the sentence-discourse interfacial layer from the complexity associated to occupying a position in the Left-Periphery.

Given this different movement derivation, the comparison between [c6 or c5 > c7] will feature two Topics targeting respectively the CP layer of the clause internal pre-verbal position, but also two different syntactic derivations: A-Movement to sentence-internal position and A-bar Movement to the CP layer.

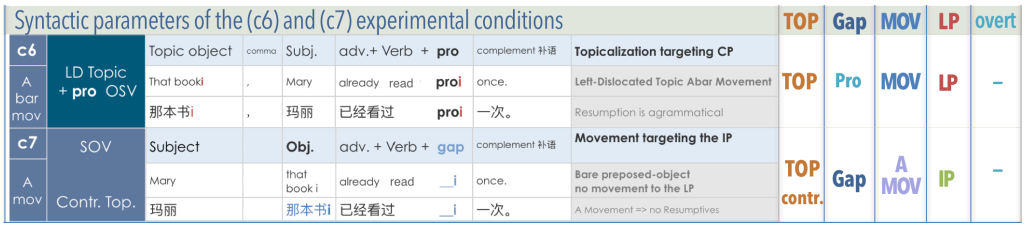


Figure 7.14 – Marked and unmarked sentences having non-truth conditional differences like our experimental conditions (c6) and (c7) in chapter 7.

In this contrast we expect to observe in the left Broca Complex (possibly in pars triangularis) an effect of displacement distance that is recurrently reported in the literature (Matchin et al., 2014; Santi and Grodzinsky 2007; Fiebach et al., 2001; Cooke et

al.,2002; Fiebach et al., 2008).

Additionally we also expect to observe some effects reported in studies investigating word-order variations, like the one reported by Bornkessel et al. (2009) that we reviewed in section §2.2.2. For this reason we selected several ROIs that are generally considered at the margins of the language network but that have been recurrently reported in word-order related manipulations, like the SMA (sensory motor area) and the Inferior Parietal Sulcus. Importantly, inferior parietal lobe activation has already been identified as reflecting thematic re-analysis processes involved in understanding object-before-subject linearization (see Bornkessel et al. 2009, Figure 2.6 p.111, or Europa and Thompson, unpublished, p. 184).

In general we expect wh-movement (A-bar movement) in Left-Dislocated Topics (c5 and c6) to yield an increased activation of the language network compared to A-movement. In fact, as we saw from the findings presented in chapter 2, several studies suggest that processing wh-movement requires greater neural resources than a more local movement like NP-movement, especially in Broca complex.

In contrast, for the opposite contrast [c7 > c6] we mainly expect to observe increased activation of the right Inferior Frontal Gyrus, as an effect linked to its reported sensitivity to the discourse manipulations and to the contrastiveness of pre-posed object construction.

7.2.4 Movement derivation: Contrasts and predictions

Topic-prominence and base-generated Topics Topic-prominence classification in linguistic Typology has been successively refined, and, since the pioneer work of Li and Thompson in 1976, it has also been recast in the framework of Principle and Parameters.

In this way, the properties of Topic-prominent languages that were isolated thanks to Typology and functional approaches have been reinterpreted, for example by stating that a Topic-prominent language is characterized by the tendency to activate the Topic Phrase (TP) (Xu Liejiong, 2001). Most recently, Paul and Whitmann (2015), stated that: "the salient feature of languages traditionally labeled as Topic-prominent is that they do not fill the functional head of the Topic projection by movement; if at all, this head is filled via External Merge (Base-Generation) of Topic particles."

Assuming this framework – through the argumented set of diagnostics presented in chapter 3 (cf. §3.4.2) –, also implies that the comparison between Base-Generated Topics and Moved ones corresponds to a contrast between External Merge and Internal-Merge operations. It is difficult to draw predictions for this brand and abstract formal distinction. We can only speculate that Internal Merge also implies an operation of Search which might reveal an area that has access to the screening of the whole sentence to perform the search operation of identifying the candidate to be Internally Merged[20]

c- Predictions for A-bar Movement versus Base-generation

As we already argued, Chinese Topic-Comment articulations are linguistic phenomena where the complexity of the syntactic-tree, linked to elements hosted in the topmost

20. As observed in chapter 2, in Phrasal Movement a *Search* operation is added within this primary buffer to identify the candidate for Phrasal Merge, and only after this Search step the suitable candidate is internally merged with the whole structure. We can then conclude that Movement appears to be a composite operation combining a *Search procedure* and a consequent possible Phrasal *Merge operation*.

Sentence domain, can be observed without movement and without embedding, so that testing Base-Generated Topics against Moved ones will constitute a different baseline compared to the ones that have been previously selected in the literature. We will namely compare two different syntactic derivations in practically the same syntactic structure having a similar linearization of words.

Hence, our first prediction comparing Moved Topics against Base-generated Topics is to reproduce the findings of several fMRI studiesç for *wh*-movement constructions compared to basic declaratives or sentence with sententential complements. More specifically, as Shetreet and Friedmann (2014) who opposed Hebrew Topicalization (achieved through A-bar movement) to declaratives, we expect to find similar patters as those reported for their contrasts: bilateral Broca Complex (BA45 left, BA47 right), the left Pre-Central Gyrus (BA6), bilateral pSTG and cerebellar activation as shown in Figure 7.15.

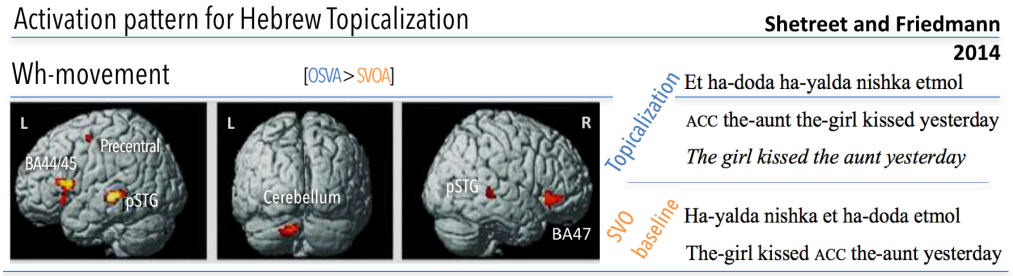


Figure 7.15 – Adapted from Shetreet and Friedmann (2014).

Syntactic parameters of B-G Topics versus Moved Topics

								height	Gap	MOV	LP	overt
c2	Base-Generated gapless Topic	Topic	comma	Subj.	adv. + VP	complement 补语	Aboutness Topic (AT) position in the Left Periphery > lower in CP	6	ø	B-G	LP	ø
aboutness B-G		That Tree	,	leaves	already became yellow	in great quantity.	Base-Generated Topic in LeftPeriphery					
		那颗桃树	,	叶子	已经黄了	一大片。	Pause marking					
c3	scene setting Gapless Topic	Topic *Space*	comma	Subj.	adv + VP	complement 补语	Hanging space Topic position in the Left Periphery > higher in CP	7	ø	B-G	LP	ø
space B-G		That Garden	,	leaves	already became yellow	in great quantity.	Base-Generated scene-setting Topic					
		那座果园	,	叶子	已经黄了	一大片。	Topic localizing the Comment					
c4	Topic + resumptive	Topic object	comma	Subj.	adv.+Verb+ Resumpt.	complement 补语	Overt Resumptive Pronoun	5	ø RP	B-G	LP	Dep. link +
Res pron B-G		(As for) Mister Wu	,	Mary	already knows him	for a longtime.	Base-Generated Topic + RP					
		吴先生i	,	玛丽	已经认识 他i	很久。	Hanging Topic					
c5	LD Topic + optional **gap** OSV	Topic object	comma	Subj.	adv.+ Verb + *trace*	complement 补语	Optional Resumptive Pronoun	4	Gap	MOV	LP	Dep. link –
A-bar mov		As for Mister Wu	,	Mary	already knows __ei	for a longtime.	Left-Dislocated Topic Abar Movement					
		吴先生i	,	玛丽	已经认识了 __ei	很久。	Island test					
c6	LD Topic + **pro** OSV	Topic object	comma	Subj.	adv.+ Verb + **pro**	complement 补语	Topicalization targeting CP	4	Pro	MOV	LP	Dep. link –
A bar mov		That booki	,	Mary	already read proi	once.	Left-Dislocated Topic Abar Movement					
		那本书i	,	玛丽	已经看过 proi	一次。	Resumption is agrammatical					

Importantly, although (c2) and (c3) differ in lexical terms from (c5) and (c6) (i.e. only one third of the sentences share the same verb), the fact that the Base-generated (c4) condition has identical lexical material as Left Dislocated Topics (c5) and (c6), will balance our substraction from the moved conditions, and further allow to subtract the

complexity factor of the establishment of a thematic chain between Topic and Comment that is also intermingled with the effect of movement operation.

Another prediction can be done regarding the possible discourse effects that characterize Left-dislocated Topics compared to Gapless ones. Given the implication of right IFG in integration of discourse-level information during sentence comprehension reported in the literature (e.g. Menenti et al., 2009), we predicted a more bilateral activation of language areas in moved Topics compared to Base-Generated ones.

7.2.5 Effect of the order of Functional Heads in the fine structure of the CP layer

Considering the overall representation of Chinese Left-Periphery in Figure 7.16, several experimental hypotheses can be drawn. The first is related to possible cerebral representation of the ordering constrains we observed in chapter 3 (§3.4.5). And the second bears on the direct contrast between the different portion of syntactic-tree that is represent when a Topic that is higher in the Left-periphery hierarchy is realized compared to one that is lower in the functional skeleton of the sentence. As already pointed in

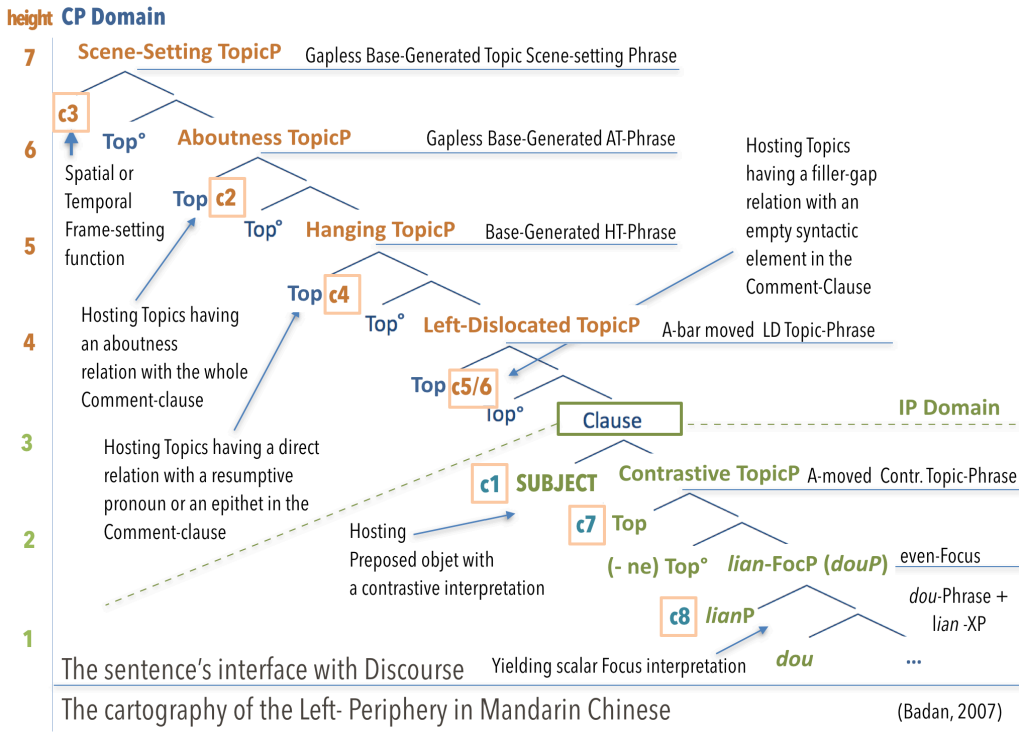


Figure 7.16 – The Cartography of the Left- and Low-Periphery of Mandarin Chinese and its ordering constraints, as developed by Badan (2007 and much related work). To each functional projection corresponds one of our experimental conditions (circled in orange).

chapter 2 (§2.4.4.1), each of the Functional Heads in Figure 7.16have a specific semantic interpretation, and the CP layer appears to be split in different hierarchically organized Projections, each encoding different semantic properties.

Hence, pragmatic interpretations like Contrastive Topic, Focus, Aboutness Topic, etc., corresponds to a syntactic position in a one-to-one relation between hierarchical position (i.e. height in the functional skeleton of the sentence) and function.

7.2.5.1 The hierarchical pointer hypothesis

By hypothesizing the cerebral encoding of this one-to-one mapping between ordinal position in the sentence's hierarchical representation and function —that was theorized following the split-CP hypothesis (Rizzi, 1997)—, we actually further develop the two sentence structure encoding hypotheses on which we based our experimental design.Our testing hypothesis namely pushes a step further the *Sentence Domains Hypothesis* by Christensen (2008), which is theorizing in broad terms the neuro-psychological relevance of the three main sentence domains (i.e. VP, IP and CP), and gives a more representational and 'static' (i.e. non movement-related) interpretation to the complexity dimension related to the height in the syntactic-tree that is theorized by the *Tree-Pruning Hypothesis* by Friedmann (2003 and much related work). Hence, as presented in chapter 3 (cf. 3.4.5, ex. 264, p.406), the complexity dimension we want to put forward is related to the complexity of representing a greater portion of syntactic-tree structure when an element is hosted in a high position in the sentence structural skeleton, it could be formulated in the following terms:

(264) ***The hierarchical pointer hypothesis*** is the hypothesis according to which there exist in the cerebral organization of syntactic knowledge, a brain area playing the role of a pointer in charge for encoding the hierarchical organization of the different Functional Projections (each one encoding different semantic properties) put forward by the cartographic analysis.

Concretely, by testing if we can tackle the cerebral implementation of the hierarchical and ordinal rules yielding the organization of the different functional projections' in the Chinese Left-Periphery, we actually implicitly hypothesize that the organization of the sentence cartographic skeleton may approximates to some extent the way the brain manages the hierarchical structural representation of the sentence-unit. If this hypothesis is born out, we should observe a brain area encoding for the ordering constrains between the different specialized positions assigned to different Topic types, that is linearly increasingly activated by the presence experimental conditions of elements in increasingly higher positions in the sentence hierarchy, for example hosting Topic elements in higher and higher syntactic positions. Note that by this we automatically hypothesize that once one position is occupied by a element all the one that are found below are automatically represented, thus giving rise to an increased representational complexity.

An explicit prediction for this exploratory complexity dimension related to syntactic-tree representation is difficult to come about. The simplest prediction we could advance is evidently that Broca complex might be involved in the encoding of this complexity dimension, which yields agramaticality when the ordering constraints it conveys are not respected. Alternatively, we may also say the linear manipulation of the positions in the Left-Periphery present in our experimental design, will enable us to discover a *new brain area* that shows a linear activation according to the height in the syntactic-tree and specifically by the height in the Left and Low-Periphery that our different experimental conditions realize.

In conclusion, we are well aware that the large number of hypotheses presented in this section might be overwhelming, however this level of detail is needed when building on fine-grained linguistic analyses to formulate hypotheses on how the brain represents the complex structure of the sentence-unit[21]. We will now move to present the Method section of the fMRI experiment.

7.3 fMRI Experiment

7.3.1 Protocol and experimental materials: corpus of experimental conditions

For the stimuli list see Appendix C.3.

Participants 21 Mandarin Chinese native speakers (14 women; 7 men) who arrived in France after high-school to complete their university training (mean age=29) participated in the experiment which was approved by the regional ethic committee (Comité de Protection des Personnes Ile-de-France VII, Prot. de Recherche Biomed. # 2008-A00241-54/1). All were right-handed (Edinburg-score ranging from 0.4 to 1, average score=0.81, Oldfield, 1971).

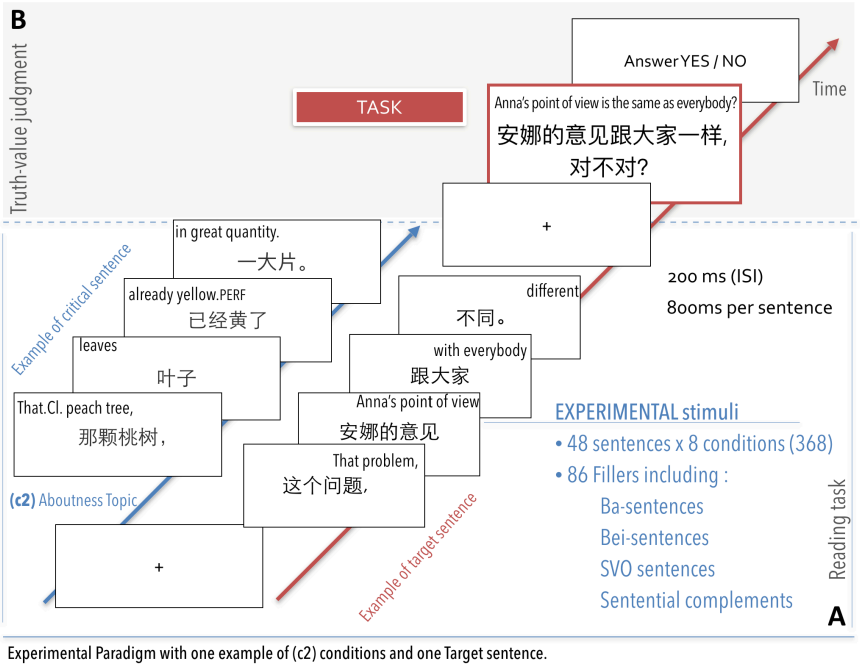


Figure 7.17 – Experimental Paradigm with one example of (c2) conditions and one Target sentence.

21. We could ironically call this the *linguist effect in front of the brain*.

Procedure

Each participant saw half of each set of the critical sentences conditions. This was done to prevent the same participant to read the two sentences of a minimal pair (like in pairs c1 and c2; c4 and c5; c7 and c8). An event-related design was used, sentences were presented in different random order for each subject. For each trial, a fixation cross appeared on the center of the screen for 1 second and the sentence stimulus was presented constituent by constituent –each consisting of 1 to 5 characters, sometimes ending with a comma –for 200ms at the center of the screen (see Fig. 7.17), each sentence lasted 800ms. The inter-trial interval was 4.5 sec (+/- 500 ms of jitter). The characters were presented in black on a white background, using the UKaiCN font at size 88. Constituents were from one to 5 characters, commas did not appear in isolation.

Task

Throughout the experiment, participants performed a truth value-judgment task on some occasional stimulus to ensure that they attended to the sentences and processed them entirely. The question related to previous sentence appeared after the stimulus sentence, as illustrated in Figure 7.17. Three to four trials per session were followed by a target comprehension sentence.

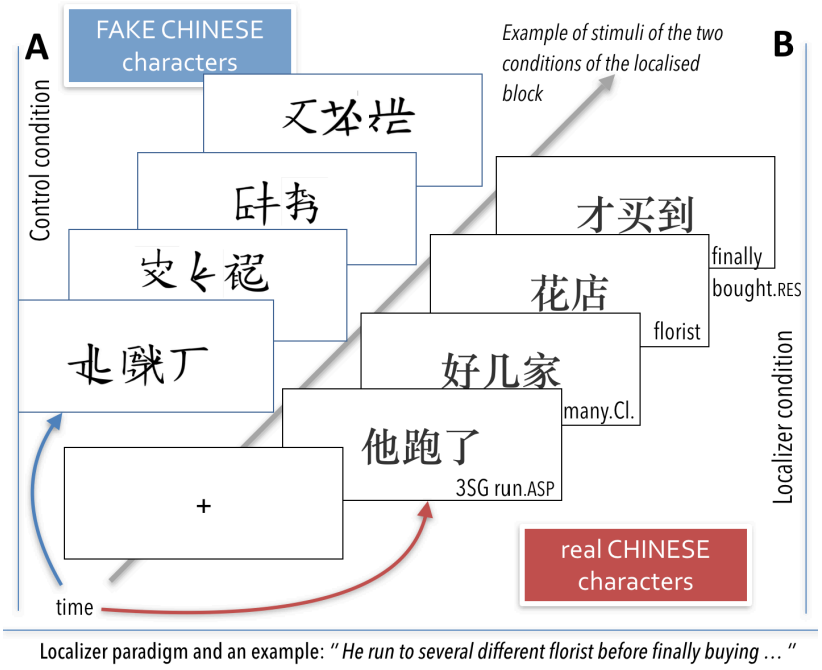


Figure 7.18 – Localizer Paradigm. Example of two stimuli conditions in the localizer block. (A) To the left in blue, an example of a trial of *control* conditions involving fake characters paired in number of charahcters to the real sentence on the right with regular Chinese sentences (B).

The participant had been instructed to press a response button with his right thumb

when 'correct' 对 *duì* and with the left thumb when 'incorrect' 不对 *búduì*. Before scanning session participants underwent a training session where they were familiarized with the pace of presentation of the sentence on the screen and with the task.

After 8 minutes of acquisition of an anatomical scan (T1), the stimuli were presented in 4 functional MRI sessions of 10 minutes, comprising 73 trials each. Subjects were instructed to covertly read the sentences and to answer to the occasional questions displayed on the scree. An initial fMRI session was a '*language comprehension localizer*' where 12 blocks of 3 real sentences (10 sec.) were displayed in alternation with 12 blocks of matched pseudo-sentences containing strokes pseudo-characters. Blocks were separated by 8 seconds.

Acquisition parameters

The acquisition was performed with a 3 Tesla Siemens Tim Trio system equipped with a 32 channels coil. For each participant, an anatomical image was taken, using a 3D Gradient-echo sequence and voxel size of 1x1x1.1 mm; Then, functional EPI scans were acquired using a multiband sequence developed by the Center for Magnetic Resonance Research (Feinberg et al. 2010) and sensitized to the BOLD effect (80 axial slices; MB=4; TR=1.5secs; TE=32msecs; Matrix=128×128; voxel size=1.5×1.5×1.5mm).

7.3.1.1 Material

Experimental conditions

A total of 384 sentences were constructed, 48 for each of the 8 experimental conditions described in Table 7.18. The entire experimental corpus is available in the Annexes (§C.3, p. 887).

The order of the constituents was manipulated across the different conditions to test for syntactic effects linked to word-order and syntactic movement in Mandarin Chinese. For this conditions were selected in order to have different derivational analyses, different constituents occupying the Left-Periphery, and different resumption strategies (see §3.4.4). As for the other linguistic characteristics of the experimental conditions that we built, we can recapitulate the following.

Pause marking and comma marking

As presented in section 3.2.3.2 (cf. discussion in §3.2.2.3, p.298), comma marking was preferred to other morphological markers to ensure full acceptability of the sentences by Mandarin speakers across regional variations, and to avoid possible emphasis and consequent contrastiveness or pragmatic saliency given to the Topic element by markers like ne, a (ya), me (Paul, 2005; Li, 2006).

Comma marking Although the literature on Mandarin Chinese identified several topic markers as *a* (*ya*), *me* , *ne* (Shi, 2000; Paul, 2005 ; Li, 2006), we decided to use the comma marking in our reading paradigm. As previously argued, the marking native speakers most agree on across dialectal areas is pause marking, which can be is transcribed in the written style by a simple comma. Our ERP study in chapter 5 showed evidence for the fact that speakers process the presence of absence of the pause, and the phono-acoustic

analysis of a corpus of 580 Gapless Topic-Comment sentences we performed, revealed a the prosodic signature of Topic comment Gapless topics as three folds: (1) Average Pause length 68 ms; (2) Significant Topic DP last syllable lengthening (78 ms); (3) 20 Hz of average pitch difference (F0) between Topic and Comment.

Animacy When building the experimental stimuli, we were particularly careful in the choice of the verbal entries. For instance, because some studies reported that matrix verb choice also affects the interpretation of embedded null objects in Chinese[22].

Secondly, as animacy is a semantic parameter that plays an important role in psycholinguistic investigation. We established that in one third of the sentences, the verb that were chosen could fit either with animate or with inanimate subjects and objects, as in the followign examples:

(265) 蒙娜，　林怡 曾经　　采访过　　她 几次。
Méngnà Línyí céngjīng cǎifǎng-guo tā jǐcì.
'Mengna, Linyi has already interviewed her several times.'

(266) 蒙娜，　林怡 曾经　　深度　采访过　　几次。
Méngnà, Línyí céngjīng shēndù cǎifǎng-guo jǐcì.
'Mengna, Linyi has already thoroughly interviewed [her] several times.'

(267) 雾霾　的 问题，林怡 采访过　　几次。
wùmái de wèntí, Línyí cǎifǎng-guo jǐcì.
'The smog problem, Lin Yi has inquired [it] several times.'

This was also meant to control for a potential impact of animacy on Topichood. Namely an ERP study by Hung and Schumacher (2014) put forward that among the two factors – animacy and givenness – that contribute to an element's potential to form an optimal topic (i.e. topic-worthiness), animate topics in Chinese showed to interacts givenness feature, modulating the effort for less prominent entities to compete for topichood.

Hence, in order to control for the potential interfering factor of Animacy. Having human [+animate] Noun Phrases in (c4) and (c5) and [+inanimate] Noun Phrase in (c1), (c2) and (c3), we varied as much as possible in (c1), (c2) and (c3) conditions sets the animacy feature of the two sentence initial NPs Topic and Subject, having 1/3 of the experimental conditions with animate NPs, showing a kinship relation with two animate NPs or other relations with one animate NP and one inanimate NP (i.e. human/body parts, human/feelings).

Post-verbal complements We tried to match as much as possible the length of the experimental conditions (average number of 13,55 characters, ranging from 15 to 12). Adverbs with different number of characters, but same meaning were systematically added in order to equate the different conditions length in terms of number of characters (and consequently number of syllables) and to have the same overall structure and ending across conditions, for this an additional post-verbal complement was also added to all conditions.

22. See the difference point out by Hsieh (2009) between *shuo* 'say' and *yiwei* 'assume'.

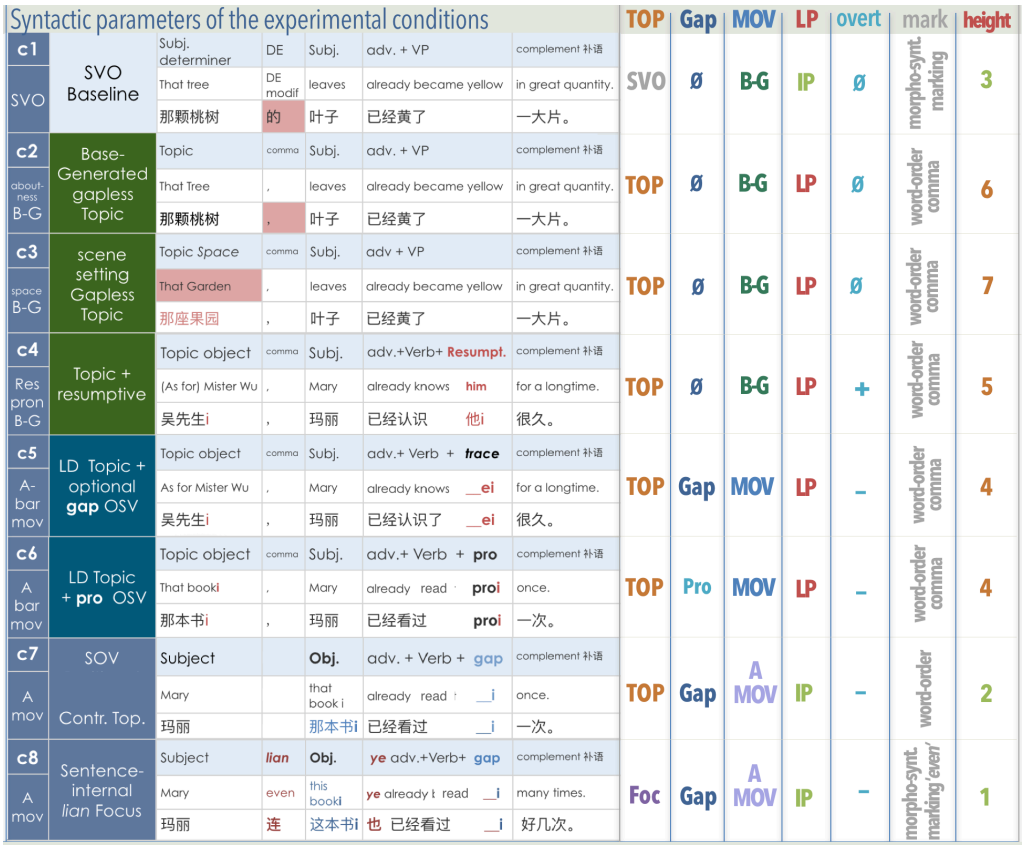


Figure 7.19 – Experimental Design.

Why we did not use the particle *le* at the end of the sentence

Fillers In addition, we included 88 fillers sentences, to distract the participant from de experimental manipulation. A variety of syntactic constructions were selected. Some of them presented a canonical word-order modified through overt morpho-syntactic marking like Ba-disposal structures, where the object is pre-posed to pre-verbal position and marked by BA, or Bei-passives sentences. But, also more simple structures like long sentential complements, and basic SVO sentences were used.

By building fillers we also considered the results from the literature showing certain trends in the co-occurence of linguistic features in the different constructions we selected. For instance, we considered the results from Lin (2004), who reported the tendency for the *bèi* construction to be associated with a negative connotation[23], and those from Murphy (2007), who found (in a multi-language) study some graded effects of animacy and telicity on the acceptability of the bèi and bǎ constructions.

In addition to the linguistic characteristics of the stimuli we already discussed, we avoided sentence final particle 了 *le* because it is known as being subject to movement

23. While verb semantics had no effect. For this the author suggested that the syntax-pragmatic interaction could occur without lexical mediation.

to the Left-Periphery (Paul, 2015b; among others see Li B.'s PhD, 2000).

Since Tang (1988/1989), Chinese final particles have often been analyzed as complementizers occupying the C position in the CP sentence domain. However, as noted in §2.1.2.4 (p. 90) it seems not plausible that all final particles, with their rich sample of different contributions to the interpretation of the sentence, are uniformly generated in one position in the sentence hierarchy. Moreover, final particles can actually co-occur, and in this case they obey a certain order, which shows that in Chinese final particles are heads of distinct functional projections in the C-domain as analyzed by Li (Li B., 2006). Already in Gasde and Paul (1996) Mandarin final particles were analyzed as complementizers occupying the head position of CP, and most crucially for us having scope over Topics, which suggested that their position in the CP layer dominates that of the Topic functional projection TopP[24]. In sum, following these analyses, we decided no to include any final particle in our stimuli.

7.3.2 Analysis

Preprocessing parameters

The functional images were first corrected for slice-timing differences, using FSL's slice timer (https://fsl.fmrib.ox.ac.uk/fsl) which handles the multiband parallel acquisition scheme. Henceforth, processing was performed with SPM8 (Wellcome Department of Cognitive Neurology, software http://www.fil.ion.ucl. ac.uk/spm), controlled by pypreprocess (https://github.com/neurospin/pypreprocess).

The functional images were realigned in order to correct for motion, one participant was rejected at this stage for movements larger than 1.5 mm or 1.5 degree. We discarded also a certain number of runs (11 out of 80 for 20 subjects) applying the same criterion. The T1 anatomical scan was spatially normalized onto the MNI template using SPM8 default parameters, and the resulting transformation was applied to the functional images. Finally, the images were smoothed and spatially smoothed with a Gaussian kernel of full width at half maximum (FWHM)=3mm.

Whole-Brain analyses

The individual statistical models comprised a regressor for each of the 8 types of sentences (corresponding to the conditions described in Table 1), as well as additional regressors for the 4 types of fillers and the targets trials. These regressors were created by convolving trains of events corresponding to the individual trials from each condition by the impulse hemodynamic response function (iHRF) of SPM. In addition, we added the estimated 6 movement parameters as nuisance variables.

The localizer data was analyzed as a block design, with blocks defined as epochs of 10 seconds convolved with the iHRF, split into two conditions (sentences or pseudo-characters). For the group level analysis, the individual effect maps of the 14 conditions (8 critical sentences + 5 filler sentences + 1 question targets) were smoothed with an isotropic kernel of FWHM=8mm and entered into a within-subject Analysis of Variance

24. Note that LI B. (2006) considers different final particles as corresponding to different functional categories, and assumes that all the functional categories headed by the final particles have scope over the topics, which is not the case of Gasde and Paul (1996). Hence, she proposes the following hierarchy: Discourse > Degree > Evaluative > Top* (from Chapt. 2 page 64 (ex.112) in LI B. PhD, 2006).

model (with one regressor per experimental condition including the fillers and one regressor per participant). This model was then estimated using a Restricted Maximum Likelihood approach that does not assume the independence of conditions, and contrasts were computed to test for various hypotheses. Unless otherwise mentioned, the maps are displayed at a statistical threshold of $p<.001$ voxel-wise uncorrected for multiple comparisons and a threshold on cluster extent set to 50 voxels, corresponding to a p-value of .05 given the estimated smoothness. To localize activations' areas and identify related Brodmann areas (BA) we used xjView (`www.alivelearn.net/xjview`).

Region of interest analyses (ROI)

Previous works manipulating different aspects of syntactic complexity provided us with several regions of interest (ROI). As show in Table 7.20, we considered regions that were reported to be sensitive to sentence structure building (in Deep blue) and syntactic movement (in Deep green), word-order variations in yellow, and intra-sentential Dependency-link manipulations (light green). Broca Complex main subdivision between (Triangularis and Opercularis) was anatomically defined (Light blue).

The the apriori regions of interest we added three aposteriori Regions to better elucidate the role of SMA activation in the different contrasts characterizing intra-sentential dependency-link establishment and animacy of the Topic NP. For this, we decided to focus on some areas that are often jointly activated in the studies we had based our predictions on. Three Sub-cortical aposteriori Region of interest to our analysis (cf. aquamarine green ROI Table, p.647).

To extract individual data from a given ROI, we first created subject-specific masks in the following way : we intersected the ROI with the statistical T-map obtained from the localizer contrast [Sentences > Pseudo-characters sentences] to which we applied a threshold selecting the 10% of voxels that showed the most robust responses inside the given ROI. These individual masks were then used to extract data from individual effect-size maps.

ROI Label	Hemisphere and Region	MNI center coordinates	Contrasts or Effect	Study
IFGorb	IFG, Pars orbitalis	-45 33 -6	constituent-size effect	Pallier et al. , 2011
IFGtri - BA45	IFG, Pars triangularis	-52 20 17	Broca Triangualris Anatomical cyto-architectonical definiton	Amunts et al.,1999/2004
aIns	anterior Insula	-37 24 -6	Merge operation	IFOP : Friederici et al., 2006
IFGtop - BA44	IFG, Pars opercularis	-51 8 15	Broca Opercularis Anatomical cyto-architectonical defintion	Amunts et al., 1999/2004
vPrC	ventral Pre-Central Gyrus	-41 10 31	Adaptation to syntactic movement + embedding	Santi and Grodzinsky 2010/2012 ; Ben-Shachar et al. 2004
dPrC	dorsal Pre-Central Gyrus (BA6)	-48 4 51	Topicalization (wh-movement) > SVO canonical Word Order	Shetreet et al., 2014
SMA	Sensory motor Area / pre-SMA	+/-7 23 41	Word-order changes // word-order making> Inflectional morph.	Bornkessel-Schlesewsky et al., 2009; Newman et al. ,2015
TP	Temporal Pole	-47 14 -27	constituent-size effect	Pallier et al., 2011
aSTS	anterior Superior Temporal Sulcus	-54 -12 -12	constituent-size effect	Pallier et al., 2011
pSTS	posterior Superior Temporal Sulcus	-50 -40 3	constituent-size effect	Pallier et al., 2011
MTG	middle Temporal Gyrus	-67 -48 -12 -54 -10 -17	Morpho-syntactic marking (unacc.) // Establishment of dependency-link: Parasitic Gap > Pronoun	Shetreet et al., 2010 ; Santi et Grodzinsky, 2012
IPS	Intra Parietal Sulcus / Supramarginal Gyrus	-42 -52 30 (-44 -44 35)	Word-order variations Object-Subject > Subject-Object // Pronoun > Gap Dependency-link in the sentence	Bornkessel-Schlesewsky et al. , 2009 Santi et Grodzinsky, 2012
Cerebellum R	right Cerebellum Lobule VI	36 -62-28	language cluster in ALE peaks meta-analysis	Stoodley and Schmahmann, 2009
Subcortical ROI label	**Hemisphere**	**MNI center coordinates**	**Contrasts or Effect**	**Study**
Thalamus	left Ventro-Lateral Thalamus	-8 -14 0	word-order: Object-Subject > Subject-Object	Xu et al. 2005 ; Bornkessel-Schlesewsky et al., 2009
Basal Ganglia	Basal Ganglia	-23 0 8 Tal	wh-Movemnt Gap: in Object Wh > Subject Wh	Santi et Grodzinsky, 2012
Putamen	left and right	-24 -1 6 29 6 -5	ANIMACY effects on linearization	Grewe et al., 2006

Figure 7.20 – Apriori Regions of Interest and aposteriori sub-cortical Regions of Interest. Light blue for Anatomically defined sub-parts of Broca Complex (Triangularis and Opercularis), Deep blue for Areas responding to structure building Constituent-size manipulations; Deep green for areas reported in syntactic movement manipulations; light green for Areas reported in Dependency-link manipulations; yellow for areas reported as having an linearly increased activation in Word-order manipulations; and finally water green ROIs are aposteriori Sub-Cortical areas jointly activated with SMA in selected in papers investigating word-order manipulations.

7.3.3 Results

This subsection reports the results of our fMRI experiment following the five main questions presented in section 7.2:

1. the brain activations linked to the so-called 'Chinese-style' Gapless Topic (i.e. Aboutness and Scene-Setting Topics);
2. the neural underpinning of the different syntactic derivations of Topic-Comment constructions in Mandarin;
3. the effects of the different types of dependency links [+/- overt] in Left-Dislocated Topics;
4. the sentence-internal word-order variation and pragmatic interpretation in 连... 也 *lián...yě* *even*-Focus compared to Contrastive Topic in sentence-internal position;
5. the effect of the order of Functional Heads in the fine structure of the CP-layer.

We will restrict to report results and leave the figures for the discussion section to facilitated the discussion and linguist reader's approach to neuro-imaging results.

7.3.3.1 Group-level and ROI analyses

Group-level analyses reveal different brain maps for the above cited different contrasts. In the following, we will report the whole-brain contrasts results and ROI analyses. The tables reproducing the cluster statistics and MNI peaks are all reported in the Annexes (§G.1, p.941).

7.3.3.2 Topic-Comment vs. SVO baseline c2>c1

The whole-brain contrast between Topic-Comment against the SVO baseline [c2 > c1] evoked three activation clusters at a very low threshold of $p<0.01$ with a cluster correction size of 10 voxels: two are located in the left hemisphere, one in the Broca Complex (between Pars triangularis/opercularis) and the other in Precentral Cortex (BA6), between the two Ventral and Dorsal regions that are classically reported in the fMRI literature on syntactic Movement (cf. Table G.4 in Annexes §G.1, p.942). The third and larger one was located in the right hemisphere in Broca's area (BA45 and BA44).

ROI analyses reported in Table 7.2, confirmed the involvement of the above clusters in the left hemisphere and additionally revealed increased activation of the SMA and right dorsal Precentral region in Aboutness Topic condition (c2).

The opposite contrast did not yield any significant cluster of activation, at whole-brain level or in the apriori ROIs.

	Comparing B-G Topic and SVO		
Cond.	**ROI**	**diff.**	**T-test *p*-values**
c2>c1	Broca Pars Triangularis (BA45 Anat.)	1.725	0.0838 .
	Broca Pars Opercularis (BA44 Anat.)	1.684	0.0992 .
	left dPRC	1.985	0.0472 *
	right dPRC	1.887	0.0591 *
	SMA	1.898	0.577 *
c1>c2			

Table 7.2 – [c2 > c1] and [c1 > c2], statistically significant t-tests in the ROIs.

To further understand the nature of this effect we investigated the temporal modulation of the percent signal change of these two conditions. The event-related plots confirm a significant B-G Topic effect in Pars Triangularis (ROI from Pallier et al., 2011) in the late time-window between approximately 5 seconds and 14 seconds after the onset of the sentence.

7.3.3.3 Effect of Scene-setting Topic (c3>c2)

The whole-brain contrast between [Scene-setting Topic (c3) > (c2) Aboutness Topic sentence] yielded an increased activation of left aSTS, left TP and of two cluster, namely parahippocampal place area (PPA) and retrosplenial cortex (RSC) at an uncorrected threshold of p<0.001, with a cluster correction size of 25 voxels (cf. Table G.3 in Annexes).

The last two cluster of activation were compared with allegedly acknowledge reference MNI coordinates for the Place Area [−27–46–15] (PPA) and Retro-Splenial Cortex (RSC) [−30–39–13] (see Epstein and Kanwisher, *Nature* 1998; Park and Chun, 2009)[25], and the peak references for these two scene-sensitive areas were found to be comprised inside our activation clusters.

Additionally, ROI analyses revealed a significant difference between the two conditions in the bilateral Precentral Cortex in favor of Scene-setting Topic, while no significant region was observed for Aboutness Topic when compared to Scene-setting one.

Comparing B-G Topic and Scene-Setting Topic			
Cond.	**ROI**	**diff.**	**T-test *p*-values**
c3>c2	left dPRC	1.895	0,0581 *
	right dPRC	1.887	0.0591 *

Table 7.4 – c2 and c3, statistically significant ROIs results.

7.3.3.4 Syntactic Movement or Base-generation (Effect of A-bar movement)

Base-generated Topics against baseline (c2+c3+c4)> (c1) Whole-brain contrast between Base-generated Topics against the SVO Baseline, [(c2+c3+c4)> (c1)] shows two significant activation clusters at an uncorrected (voxel-wise) threshold of p<.001: one in pars Opercularis and the other in the dorsal Precentral Cortex (BA6), with a cluster-extent correction of 50 voxels.

Moved-Topics against baseline (c5+c6) > (c1) Whole-brain contrast between Moved Topics against the SVO Baseline, [(c5+c6) > (c1)] yields five main clusters, located bilaterally in pars Triangularis/ Opercularis, bilaterally in the Precentral Cortex (BA6) and in left Anterior Temporal Lobe (ATL) at an uncorrected threshold of p<.001, with a cluster-extent correction of 50 voxels.

Activation in the left pSTS do not survive cluster-size correction of 50 voxels, but its involvement in Moved-Topics effect is confirmed by ROI analyses. pSTS activation level for both Moved Left-Dislocated Topics -(c5) and (c6)- is significantly superior to percent signal change of the SVO baseline and to the two Gapless Base-generated Topics (c2 and

25. Park and Chun (2009) Neuroimage report the following reference MNI coordinates: -27 −46 −15; 30 −44 -14; 24, -33, -23.

c3), while when compared to the Base-Generated Topic featuring a Resumptive pronoun (c4), it is not significantly different.

Comparing two syntactic derivations of Topics A-bar movement versus Base-generation

Un-direct comparison of Base-generation vs. A-bar movement effects shows that Base-generation elicits a sub-set of brain areas present in A-bar movement effect, while the direct comparison between the effect of A-bar moved Topics versus Base-generated Topics however presents only a single significant cluster of activation localized in the right MTG [51 -13 -8] ($p<0.001$, vox-size=25). The conjunction of the two effects yields two clusters of activation, one in Broca pars Triangularis and dorsal Pre-Central cortex ($p<0.005$, cluster-size correction=25 voxels).

7.3.3.5 Dependency-links in the sentence

For the following contrast we present the results for both apriori and aposterori ROIs.

Comparing Resumptive Topic-Comment c4 and c5 Gapped Topic-Comment Whole-brain analyses reveal no significant activation at a threshold of $p<0.001$ (cluster-size correction = 25 voxels) comparing Hanging Topic with a Resumptive (c4) versus to Left-Dislocated Topic with a gap (c5) in both directions. Lowering the voxel-extent correction we observe one small cluster in favor of (c4) Resumptive condition in the medial orbito-frontal cortex (BA11), and another in favor of Gapped Condition (c5) in the right cerebellum ($p<0.001$, voxel-wise). The following Table presents ROI results showing an increased activation in the left Putamen for gapped Topic and in dorsal Precentral Cortex for Resumptive Topic condition.

	Comparing Resumptive and Gapped Topic		
Cond.	**ROI**	**diff.**	**T-test *p*-values**
c5-c4	left Putamen (Grewe)	-2.099	0.0358 *
c4-c5	Precentral	3.979	6.93e-05 ***

Table 7.6 – c4 and c5, statistically significant ROIs results.

c4 and c6 Whole-brain contrasts subtracting null pronoun (*pro*) condition (c6) from Resumptive (c4) revealed an extended activation pattern at $p<0.001$ uncorrected threshold (cluster-size correction = 50 voxels), encompassing right Broca opercularis, a left hemisphere broad cluster going from the top-most part of Middle Frontal Gyrus through the dorsal Precentral region (BA6) and ventral Precentral down to Broca pars Triangularis and Opercularis. Additional clusters show an increased activation in bilateral superior Parietal Lobule and left Inferior Parietal region. No activation was found in the whole-brain for the opposite contrast, while ROI analysis revealed a significant increase in activation in left Middle Temporal Gyrus selected from Santi and Grodzinsky study on filler-gap dependencies.

Comparing Resumptive and Null pronoun			
Cond.	**ROI**	**diff.**	**T-test *p*-values**
c4-c6	Broca 44 (anat)	4.103	4.08e-05 ***
	Broca 45 (anat)	3.297	9.76e-04 ***
	right vPRC	3.555	3.77e-04 ***
	dPRC	3.801	1.44e-04 ***
	aSTS Pallier	2.292	0.0219 *
	TPJ Pallier	2.070	0.0384 *
	pSTS Pallier	4.039	5.36e-05 ***
	IPS	3.996	6.45e-05 ***
c6-c4	MTG	-2.112	0.0347 *
	left Basal Ganglia	-2.191	0.0285 *

Table 7.8 – c4 and c6, statistically significant ROIs results.

c5 and c6 Whole-brain contrasts subtracting null pronoun (*pro*) condition (c6) from the Gap condition (c5) revealed an extended activation pattern very similar to the one reported for [c4>c6] at the same threshold. The activation clusters were located in right Broca opercularis, the same broad cluster going from the top-most part of Middle Frontal Gyrus through the dorsal Precentral region (BA6) and ventral Precentral down to Broca pars Triangularis and Opercularis was observed, and a bilateral increased activation was found only in Inferior Parietal region. An additional broad cluster (compared to previous contrast) was present in the SMA region.

No activation was found in the whole-brain for the opposite contrast, and here again the ROI analyses revealed a significant increase in activation in left Middle Temporal Gyrus for (c6) conditions featuring a null pronominal (*pro*).

Comparing Gapped Topics and Null Pronoun Topics (*pro*			
Cond.	**ROI**	**diff.**	**T-test *p*-values**
c5-c6	Broca 44 (anat)	-3.917	8.96e-05 ***
	Broca 45 (anat)	-3.763	1.68e-04 ***
	right vPRC	-3.079	2.08e-03 **
	dPRC	-3.232	1.23e-03 **
	SMA	-3.213	8.14e-03 **
	pSTS Pallier	-2.499	0.0124 *
	IPS	-4.163	3.14e-05 ***
	Cerebellum	-2.427	0.0152 *
	Left Putamen Grewe	-2.893	3.82e-03 **
c6-c5	MTG	1.970	0.0488 *

Table 7.10 – c5 and c6, statistically significant ROIs results.

Conjunction analysis of Dependency-links The conjunction analysis of the different Dependency-links between Topic and Comment was obtained by intersecting three contrasts: i.e. [c4 > c2] ∩ [c5 > c2] ∩ [c6 > c2]. Each Topic condition featuring a dependency-link (i.e. (c4), (c5) and (c6)) was first contrasted against condition (c2), where the Aboutness Topic has no dependency with the Comment. This revealed three clusters of activation in the left hemisphere at an uncorrected threshold of $p < 0.005$ (cluster-extent correction 50

voxels) located in the anterior Temporal region, in the Middle temporal Gyrus and in the Insula. Interestingly, lowering the voxel-extent correction to 25 the posterior Superior Temporal Sulcus was also observed.

Contrastiveness and Focus

c7 and c8 Whole-Brain contrast opposing morpho-syntactically marked *even*-Focus and pre-posed object contrastive Topic condition (i.e. [c8>c7]) elicited a broad cluster of activation in the left Temporal Pole, two close activation peaks in the Precentral Region, one in the right SMA, a very strong activation of the Thalamus, and a bilateral activation of the inferior occipital region, probably due to the average character difference between the two conditions.

The inverse contrast did not yield any significant cluster even at a very permissive threshold, at whole-brian level and in the ROIs.

	Comparing Contrastive Topic (SOV) and *even*-Focus (S+*lian*+O+*dou*+V)		
Cond.	**ROI**	**diff.**	**T-test *p*-values**
c8-c7	Broca Orbitalis (IFG orb)	-2.892	3.83e-03 **
	aSTS	-2.998	2.72e-03 **
	lMTG displace	-2.978	2.91e-03 **
	rBasal Ganglia	-3.183	1.46e-03 **
	Thalamus	-2.013	0.0441 **
	TP Pallier	-4.219	2.46e-05 ***
c7-c8			

Table 7.12 – c7 and c8, statistically significant t-test ROIs results.

Order of Functional Heads in the fine structure of the CP-layer

The linear contrast assigning to the different conditions the value of their ordinal position in the Left-Periphery and Low-Periphery revealed at an uncorrected threshold of $p < 0.005$ (cluster-extent correction of 50 voxels) a unique cluster in the dorsal Precentral top-most part at the limit with Middle Frontal Gyrus, while the $p < 0{,}001$ (voxel-wise) threshold reveals two bilateral smaller clusters in the same area the Precentral (aal) Brodmann area 6.

7.4 Discussion

7.4.1 To-pic or no to-pic, that is the question

The first patterns of results we will discuss are those linked the so-called 'Chinese-style' Gapless Topic. We will discuss (a) the effects elicited by the presence in the sentence of an element in the CP layer, and the (b) effects observed for the semantic content of the Topic an the Frame-setting relation between Topic and Comment-clause will be discussed, and at last (c) we will focus on the neural underpinning of the different syntactic derivations of Topic-Comment constructions in Mandarin.

7.4.1.1 Effect of Topic-Comment vs. SVO sentence in Base-Generated Topics

While previous neuro-linguistics literature has identified cerebral activations for Topicalisation by Left-Dislocation, which crucially imply (i) syntactic movement operation, (ii) filler-gap dependency processes and (iii) the representation of an extra sentential layer (e.g. wh-movement effects in Shetreet et al. 2014; Ben-Shachar et al. 2004; Santi and Grodzinsky, 2007), the comparison between Gapless Base-generated Topic Sentences like the Aboutness Topic in (c2) with the SVO baseline (c1) (see Figure 7.21) was meant to isolate the effect linked to the presence of the Topic in the Left-Periphery from that of movement-related processes and those for establishing filler-gap dependencies between the Topic constituent and the place in the Comment-clause where it has been extracted from.

The whole-brain contrasts between (c2) B-G Gapless Topic and (c1) SVO baseline reveals three clusters of activation only at a very low threshold ($p < .01$): one in left Pars Triangualris/0percularis, one in the left dorsal Pre-Central area and the main one in right Pars Triangularis (BA45).

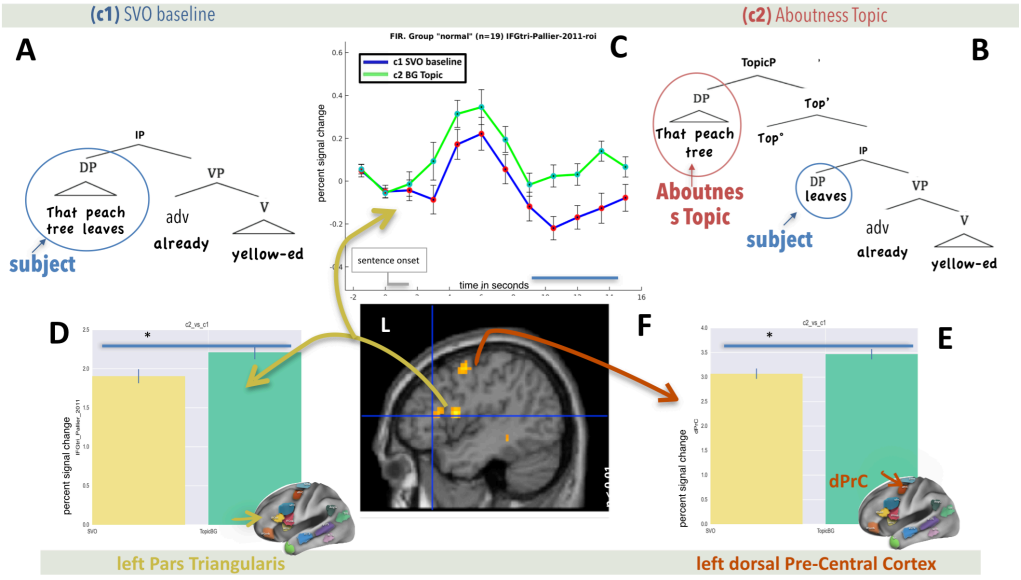


Figure 7.21 – Effects of having an NP in the CP and of representing the height in the CP-layer of the sentence-unit. (A) Syntactic tree-diagram for SVO Baseline (c1). (B) Syntactic tree-diagram for Aboutness Topic (c2) condition. (C) Event-relted plot (FIR model) in Pars Triangularis Pallier et al. (2011). (D) Signal Percent change in Pars triangularis ROI. (E) Signal Percent change in dorsal Precentral ROI. (F) Whole-brain contrasts between [*in situ* Topics externally merged (c2) > SVO Baseline (c1)], p<0.01 uncorr., with a voxel extent correction of 10.

Although the whole-brain contrast reveals a very lenient effect ROI analyses confirm the involvement of left hemisphere activation in Pars Triangularis (BA45) ($p<0.05$) and Precentral Sulcus bilaterally ($p<0.01$) and additionally show the involvement of the SMA ($p<0.05$) for Aboutness Topic[26].

26. ROI analysis additionally shows the involvement of a sub-cortical region - the Thalamus and Putamen regions ($p<0.01$)- for SVO baseline compared to (c2).

To further understand the nature of this effect, we investigated the temporal modulation of the percent signal change of these two conditions. The event-related plots reported in Figure 7.21, left) confirm a significant effect of Aboutness Topic in Pars Triangularis (ROI form Pallier et al., 2011), but mainly in the late time-window between approximately 5 seconds and 14 seconds after the onset of the sentence.

This result shows that left Pars Triangularis of Broca's Area is activated by the presence of a Base-Generated Topic with its extra sentential Left-periphery position for the Topic NP.

Note that from a derivational point of view, the first constituent 'that peach tree' *na-ke taoshu* 那棵桃树 is externally merged in both conditions (c1) and (c2). Hence, as the two externally merged constituents mainly differ only for their position the sentence syntactic-tree, we may advance that Broca's Area Triangularis activation reflects together with that of the Precentral Gyrus to the representation of the Topic element and its position in the Left-Periphery of the sentence.

Importantly, the fact that the main cluster of activation is found in Broca region in the right hemisphere, confirms our predictions that the activation of the sentence-discourse interface layer by the presence of the Topic would yield right Broca increased activation as reported in Shetreet et al. (2014) and Ben-Shachar et al. (2004). Following these studies, and studies that manipulated discourse-related processes in sentences like Caplan and Dapretto (2001) and Menenti et al. (2009) (cf. Prediction section), we interpret the involvement of Right Broca's area as reflection in the sentence-discourse interface properties associated to Topicality.

In sum, although the whole-brain results for the comparisons between Aboutness Topic and Declarative SVO baseline are presented here using uncorrected statistics for multiple comparison[27] and at a very low threshold, we are confident that these results are not spurious. Not only they fit with our expectations, but the areas where we observed an effect are all areas where an increased activation was expected based on previous fMRI experiments investigating similar syntactic structures, and on the predictions of the two models we rely on for this experiment, namely the Tree-pruning Hypothesis (2.4.4.5) and the Sentence Domain Hypothesis, which have both been supported by rich fMRI and neuro-psychological evidence (cf. §2.4.4.4, p.211 and 2.4.4.3, see Table 2.6).

7.4.1.2 Topic as a Spatial Frame

We turn now to the issue of function of 'Chinese-style' Topics and their semantic content. In chapter 3, we saw, through our linguistic literature review, that a characterization of the different possible contents that Topics can host has proven difficult to establish. In fact, no one-to-one correspondence between this syntactic position and a particular semantic content has been found, which led Paul (2014) to propose that the interpretation of a constituent in Topic position is mainly the result of the interaction between its syntactic and semantic properties. Namely, this is what happens in Scene-setting Topics, where the interplay between the semantics of the Topic NP and the Topic-Comment articulation determines its function as a Frame-setter. In our condition (c3), the Topic plays the role of a spatial frame-setter localizing the Comment-clause predication. Importantly, the shift in function from "aboutness" to "frame-setting" is here obtained by the simple semantic change of the content of the Topic as shown in Figure 7.22. This makes

27. This will be the case in the following too. In the following cluster-size correction will be 25 to 50 voxels to ensure that the areas identified constitute real focal activations.

the two conditions (c2) and (c3) differ minimally in terms of lexical material, while the role of their Topics and the relation they establish with the Comment are fundamentally different.

As we saw in §3.1.3.3 and §4.1, the frame-like function of Scene-setting Topics is allegedly established across different linguistic frameworks and its first formulation by Chafe defined it as *a way to limit the applicability of the main predication to a certain restricted domain, setting for example a spatial frame within which the main predication holds* (1976:51). This is indeed the case in our condition (c3).

Hence to address this issue linked to the interpretation of the two main functions of Topics – 'aboutness' or 'frame-setting' –, we compared the Scene-Setting Topic in condition (c3), with (c2) Aboutness Topic condition.

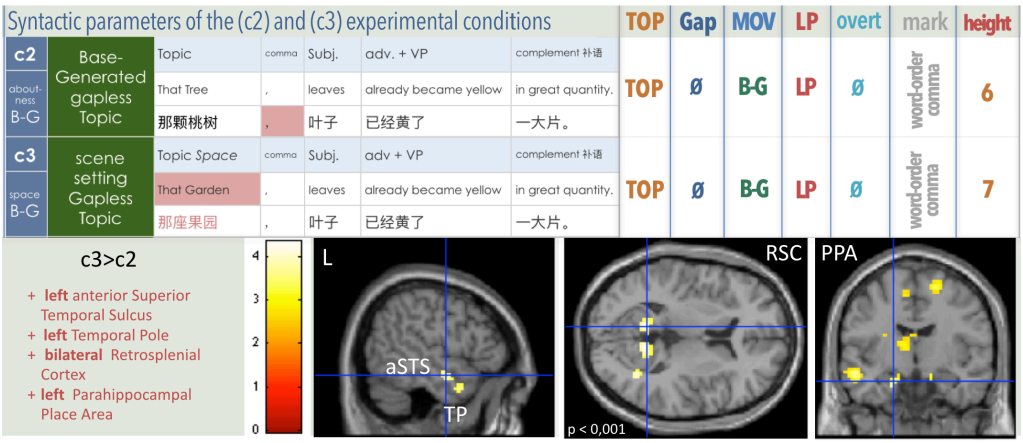


Figure 7.22 – Whole-brain contrast between [(c3) Space DP B-G Topic sentence > (c2) B-G Topic sentence], p<0.001 uncorr. voxel-wise. Reference MNI coordinates for the Place Area reported in the literature [−27–46–15] are comprised inside the activation clusters reported in the above figure (left), [−30 − 39 − 13]. PPA stands for Parahippocampus Place Area ; RSC stands for Retro-Splenial Cortex.

As shown in Figure 7.22, the contrast (c3 > c2) detects an increased cerebral activity in two regions of the sentence network -the anterior STS and the Temporal Pole (TP)-, but also yields a strong effect in two scene-sensitive regions that are not part of the language network: the parahippocampal place area (PPA) and retrosplenial cortex (RSC). Interestingly, the parahippocampal place area (PPA) is a sub-region of the para-hippocampal cortex that lies medially in the inferior temporo-occipital cortex, and plays a central role in the encoding and recognition of environmental scenes (compared to faces, see Epstein and Kanwisher, 1998; Park and Chun, 2009).

This last unexpected result provides some evidence for the difference between (i) a Topic in an aboutness relation with the Comment, and (ii) a Topic which sets a spatial framework for the Comment's predication. We interpret this result as being in line with the rich linguistic literature on Topic framing role and understand the activation of scene sensitive regions as reflecting the sentence-discourse interface mechanism yielded by the Topic as revealed by the ERP results presented in chapter 5.

Effect of the semantic content of the Topic Noun-Phrase While the increased activation in the aSTS can be interpreted as caused by the lexical difference between the two

Topic NPs, the activation in the anterior part of the Temporal Lobe (TP apriori regions of interest) could reflect semantic composition and integration processes (Bemis and Pylkkanen, 2013) linked to effort of integrating the two sentence initial NPs: the subject and the Topic feature in the (c3) condition have a greater semantic distance: *garden–leaves* in (c3) condition compared to *tree-leaves* part-whole relation in (c2).

Effect of the height in the syntactic-tree Importantly, the ROI analysis reveals a significant increase of activation in bilateral dorsal Precentral area for Scene-setting Topic compared to Aboutness Topic (cf. Table 7.4 in the Results section, p.649). As we saw in §7.2.1.2, these two Topics crucially differ in the height of their respective position in the Left-Periphery. We interpret the involvement of this area, which is systematically reported in studies on syntactic movement[28], as linked to the representation of the height of the functional projection where Scene-setting Topics are hosted in the sentence's structural skeleton, complexity measure reflecting the additional portions of the syntactic-tree that are needed to be represented under the position where a given Topic is hosted, as illustrated below (Figure 7.23).

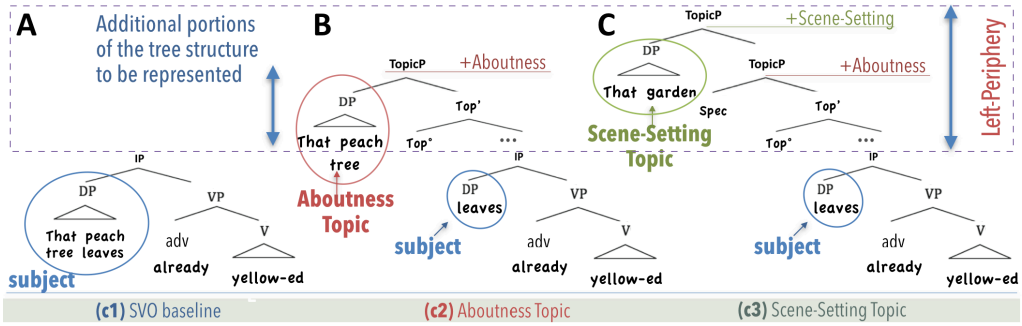


Figure 7.23 – Additional portions of the syntactic-tree structure to be represented in the Left-Periphery for (c1), (c2) and (c3) conditions..

Note that this interpretation is also coherent with the observed increased activation in Precentral cortex when comparing Aboutness Topic with the SVO baseline (cf. figure 7.21, ROI results Table 7.2, p.648). This also contributes to differentiate the role of Broca Pars Triangularis from the one played by dorsal Precentral Cortex. In fact, if we consider the height of the syntactic-tree diagrams in Figure 7.23 as the being the determinant of activation of the dorsal Precentral region, we can note that Broca Complex increased activation is observed only when subtracting the SVO baseline, with its empty Left-Periphery, to the conditions (c2) in (A) and (c3) in (B) – where a Topic is hosted in the left-Periphery –, but not in the [(c3) > (c2)] contrast, as both conditions have a element in the CP-layer. Hence, this suggests that while Broca complex is activated by the presence of an element in the Left-Periphery, the Precentral regions (dPrC) seems to be involved in the complexity dimension encoding the height in the syntactic-tree and the consequent additional portion of tree to be represented.

28. cf. §2.4.2, p.162, and Annexes §D.1 (p.908) gathering several evidence for the cardinality of Precentral Cortex (dorsal and ventral portion) in Syntactic processes across different neuro-imaging studies in healthy subjects and patients.

Sentence-discourse interface processes Moreover, we can advance that the significant increase in activation observed in two scene-sensitive regions that are outside the language network seems to converge with the online context-load effects we observed during the online ERP recordings of Chinese Scene-setting Topics in chapter 5. Namely, our ERP findings on the temporal unfolding of the very same Scene-setting Topic constructions had revealed a reiterated access to context information during the auditory comprehension of Scene-setting Topics in context. Put together, fMRI results and ERP results bring to light the cerebral effects of the sentence-discourse interface properties of Topic-Comment syntactic articulations that are theorized in linguistics. These sentence-discourse interface processes, associated to hosting a constituent in the Left-Periphery of the sentence structure, were revealed both by the context-load effects we observed during online sentence processing of Scene-setting Topics and by the increased activation of scene-sensitive regions outside the Sentence Network we observe here.

7.4.2 Syntactic Movement or Base-generation (A-bar movement) of Topics

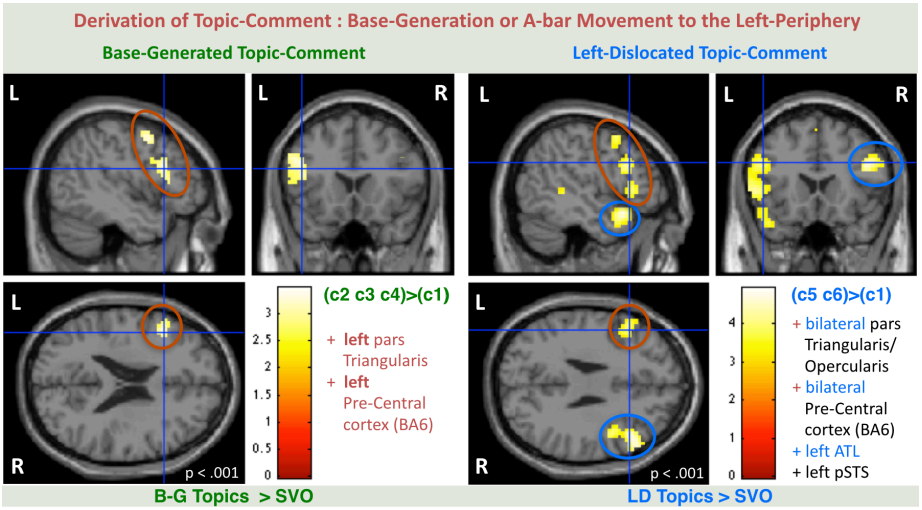


Figure 7.24 – Different effects of the syntactic derivation.
LEFT: Base-Generated Topic effect compared to baseline, whole-brain contrast between [(c2+c3+c4) > (c1)]. RIGHT: Left-Dislocated Topics (A-bar) compared to the same baseline on the right, whole-brain contrast between [(c5+c6) > (c1)], p<.001 uncorr. voxel-wise.].

As presented in section §7.2.4 (cf. also §3.4.2), our design casts a neuro-linguistic look on the debated linguistic issue of the syntactic derivation of 'Chinese style' Gapless Topics.

The different activation patterns for Base-Generated Topics compared to Left-Dislocated Topics (A-bar Movement) presented in Figure 7.24, show, as predicted, a broader extent of activation for Moved-Topics compared to Base-Genrerated ones. It has to be noted that on both sides some conditions present a dependency-link between Topic and Comment: in (c4) the Topic is Base-generated (externally merged) in the Left-Periphery but has to establish co-referential binding with the resumptive in the Comment. It has to be noted that this condition also shares the same lexical material as the moved Topic conditions (c5 and c6).

Base-Generated Topic-Comments (c2+c3+c4) compared to SVO baseline (c1) elicited two activation clusters, the same two brain areas where we found an effect for Gapless Aboutness Topic (c2): Broca pars Triangularis and dorsal Pre-Central cortex (circled in red in Figure 7.24).

Left-Dislocated Topics activated a wider and more bilateral network, comprising the same areas as Base-Generated Topics, and the following additional areas (circled in blue, see Fig. 7.24): left anterior Temporal Region, Broca pars Triangularis/Opercularis and the Pre-Central cortex in the right hemisphere.

However, the direct comparison of Left-Dislocated Topics (A-bar Movement) minus Base-Generated Topics yielded a single significant cluster of activation in the right MTG [51 -13 -8] ($p<0.001$, vox-size=25).

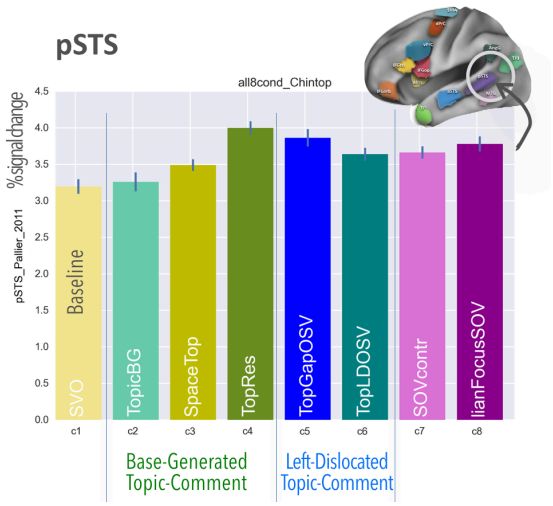


Figure 7.25 – Percent signal change for the 8 experimental condition in posterior Superior Temporal Sulcus (pSTS). For the profiles of activation of all our apriori and aposteriori (sub-cortical) ROIs see the Annexes §G.2 (p.945). The pattern of response observed in this region seem to be related to presence or absence of dependency-link between Topic and comment, rather than determined by movement related processes.

Notice that the cluster activation in the left pSTS shown in Figure 7.24 does not survive cluster-size correction of 50 voxels, but its involvement in Moved-Topics effect is nonetheless confirmed by ROI analyses.

As shown in Figure 7.25, the activation level of the posterior Superior Temporal Sulcus ROI for the two Left-Dislocated Topics (c5 and c6) is significantly superior to both SVO baseline and to the two Gapless Base-generated Topics (c2) and (c3). On the contrary the response of Base-Generated Topic featuring a Resumptive pronoun (c4) appears to patterns with (c5) and (c6) conditions. This could indicate that pSTS plays a role in the establishment of the referential dependency-link between Topic and Comment that is present in conditions (c5), (c6) and (c4), independently from their different movement derivation.

The conjunction of the two effects yields two clusters of activation, in Broca pars Triangularis and dorsal Pre-Central cortex ($p<0.005$, vox-size=25). We can therefore conclude that the common neural substrates of Topic-Comment sentences regardless of their syntactic derivation – Base-Generation implying only External Merge or A-bar wh-movement implying also Internal Merge – are these two clusters situated in the left Inferior Frontal cortex.

As these are the same two clusters that were observed in the contrast between Aboutnesse Topic and SVO baseline, we can advance that they may sub-serve the complexity of representing elements in the Left-Periphery.

In sum, the neural signature linked to the presence of a Topic element in the Sentence's Left-Periphery regardless of its syntactic derivation involves an increased activation of Broca's Pars and Triangularis (BA 45) and of the dorsal Pre-Central region (BA6), and that movement-related processes are also involving a right hemisphere region situated in the Middle Temporal Gyrus.

7.4.3 Different intra-sentential dependency-links between Topic and Comment

We now move to discuss the observed effects of the different [+/- overt] and [+/- pronominal] dependency-links present in Base-Generated Hanging Topics with a Resumptive in (c4) and in the two Left-Dislocated Topics, containing a gap in (c5), or a null pronominal *pro* in (c6).

7.4.3.1 Gap or Resumption (c4) vs. (c5)

The whole-brain contrasts between gap (c5) and resumptive (c4) revealed no significant clusters of activation at whole brain level in either directions (c5 > c4 and c4 > c5), even at very permissive statistical thresholds. From this, we can say that the optionality of the gap is not a major source of complexity effects at the level of cerebral processing.

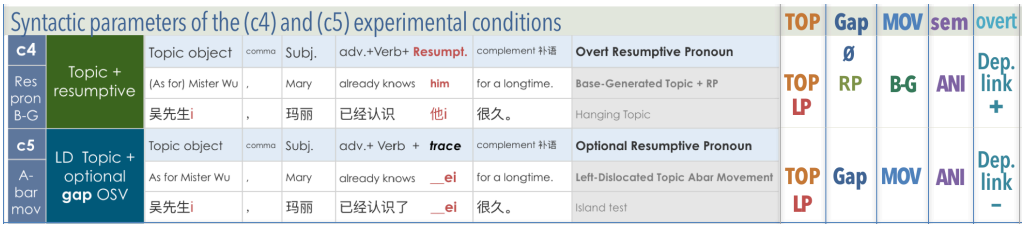


Figure 7.26 – Gap vs. Resumptive Event-Related Plot of (c4) Resumptive B-G Topic condition (blue) and (c5) A-bar Left-Dislocated gaps (green) activation time-course in SMA apriori Region of Interest.

However, as argued in the linguistic introduction to the experimental conditions (§7.2.2 and chapter 3, §3.4.4), the two sentences actually also differ in their syntactic derivation and in the height of the syntactic position where the Topic is hosted in Left-Periphery, see figure 7.26 and figure 7.16 (p.638).

In this regard, the ROI analyses reveal an effect that can be associated to this last syntactic complexity parameter: a significant increase in activation of the dorsal Precentral cortex is found for the Hanging Topic condition (c4), which is Base-Generated in a higher portion of the Left-periphery compared to Left-dislocated topics. Importantly, this further confirms the previously argued function of this region. The activation for HT topics compared to LD ones in dorsal Precentral Cortex is indeed in line with the fact that this area is involved in representing the height of the syntactic tree or the additions portion of sentence structural skeleton to be represented when a sentential element is hosted in a higher position in the Left-Periphery.

7.4.3.2 Null Pronoun, Gap, Resumptive and Animacy

As shown in section §7.2.2, Chinese grammar features an asymmetry between Left-Dislocation of animate object compared to that of inanimate objects: Resumption strategy is ungrammatical in case of inanimate objects. The linguistic analysis of this configuration presented in chapter 3, assigns different lexically empty syntactic elements to the original position of the fronted elements in (c5) and (c6), respectively a gap [-anaphoric, -pronominal] and a phonologically non-realized pronoun [-anaphoric, +pronominal], the so-called little *pro*.

Interestingly, the whole-brain contrast [(c5) vs. (c6)] reveals that four brain regions are more active when the fronted object leaves a gap at its original position than when a null pronominal is found: encompassing dorsal Pre-central Cortex down to Broca Pars Triangularis/Opercularis, Inferior parietal Lobule and the Sensory Motor Area, as illustrated in Figure 7.27.

Figure 7.27 – Effect of [+/- pronominal] empty syntactic elements at gap position to establish a dependency-link between Topic and Comment. Right Top: Whole-brain contrast between [(c5) Abar Left-Dislocated Topic > (c6) A-bar Left-Dislocated], p<0.001 uncorr. voxel-wise. On the right bottom : activation time-course in SMA Region of Interest of (c6) null pronoun condition (green) and (c5) gapped Topic (blue).

Animacy and the SMA Several studies report stronger effects for person [+animate] objects compared to object [–animate] objects in cases of anaphora resolution (Hammer, 2007 and 2011), and in cases of animacy effects on linearization (Grewe et al. 2005).

Specifically, the selection of the object in (c5) and (c4) compared to (c6) should be more difficult possibly because of the competition of two sentence-initial animate NPs, or because in its movement to sentence-initial position the animate object has crossed the [+animate] subject.

Contrary to what we had predicted following Hammer and Colleagues (2007/2011) findings of animacy linearization, (see Figure 7.11), we do not observe an effect of animacy in the temporal lobe (aSTG), but an increased activation in two areas we had selected in our apriori ROIs because recurrently reported in the processing of word-order variations and more specifically Object-before-Subject word orders: the IPL and the SMA.

While the ROI analyses show a clear animacy effect in the IPL region where both animate Topic conditions (c5) and (c4) are more activated compared to the inanimate Topic condition (c6) (see figure 7.31B, ROI Tables 7.8 and 7.10 in the Results section §7.3.3.5), the pattern of responses in the SMA are more complex.

Namely, as the increased activation of the SMA is shown in the whole-brain analysis only for [c5 > c6] and not for [c4 > c6], the role of the SMA cannot reflect the effort of retrieving the object referent in cases where the sentence-initial NPs are both sharing the same Animate and Human features.

As the temporal dimension is important in the resolution of filler-gap dependencies (i.e. the filler has indeed to be kept in working memory until the gap is found in the sentence), we investigated the activation time-course of the SMA Region of Interest for (c6) inanimate topic obligatory null pronoun (green) and (c5) optional gaps (blue). As shown in Figure 7.27, no temporal difference or delay in the response pattern is observed.

To further clarify the possible link between animacy and the SMA, we additionally

compared the event-related time course of conditions (c5) with a gap and condition (c4) with a resumptive (see Figure 7.28), and found that the presence of a gap compared to that of Resumptive Pronoun in the Comment clause shows a stronger activation of the Sensory Motor Area (SMA), which definitely invalidates the hypothesis that the SMA response pattern are to be related to purely animacy based considerations, possibly related to the competition of two sentence-initial NPs that are both sharing the same [+animate, +human] features.

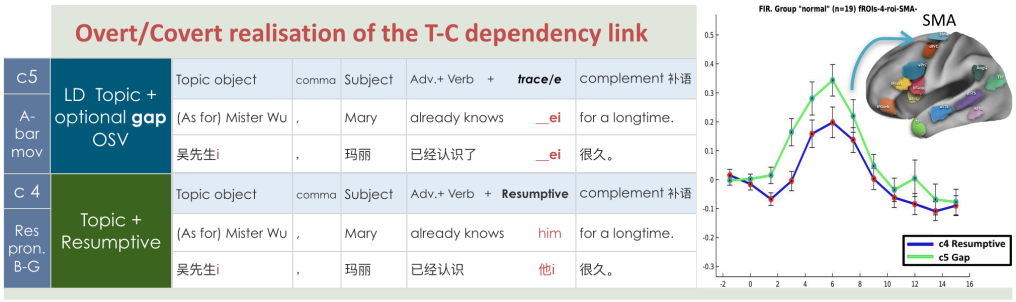


Figure 7.28 – Gap vs. Resumptive. Event-Related Plot of (c4) Resumptive B-G Topic condition (in blue) and (c5) A-bar Left-Dislocated gap (green) activation time-course in SMA apriori Region of Interest.

In fact, while previous contrast [(c5) gap > (c6) *pro*] features a difference in animacy of the Topic element, the [(c5) gap > (c4) Resumptive] presents two sentences with identical linear word-order and lexical material.

We can, thus, refine the initial interpretation, and say that the effect we observe in the SMA can only be due to the increased difficulty of assigning the object role in presence of two sentence-intial animate NP, but only when a gap is found in the Comment-clause compared to the presence of pronominal, be it lexically overt as in (c4) or covert as in (c6).

A short overview of the neuro-imaging literature shows that the increased activation of the SMA region is consistently reported in case of non-canonical word-orders. For instance, Röder et al. (2002) showed that more difficult object-before-subject German word-order (e.g. [Indir.Obj.+ Dir.Obj.+ Subj.] or [DirObj. + IndirObj. + Subj.]) involve Pre-Central area and SMA brain regions. These findings were reproduced by Friederici et al. (2006), who reported a graded activation of Broca's pars Opercularis and SMA as increasing with the number of successive object frontings.

Additionally, another study manipulating German direct Object linear position reports increased activation of the SMA and Precentral Gyrus for the contrast Object-initial versus Subject-initial sentences (Bornkessel-Schlesewsky et al., 2009), thus, corroborating the hypothesis that the SMA is involved in the processing of word-order variations and more specifically Object-before-Subject word orders. Yet, another study reports bilateral activation of the SMA region in a syntactically more informative contrast involving another type of syntactic movement (i.e. NP-movement) in Passive sentences compared to Active sentences in German (Mack et al., 2012).

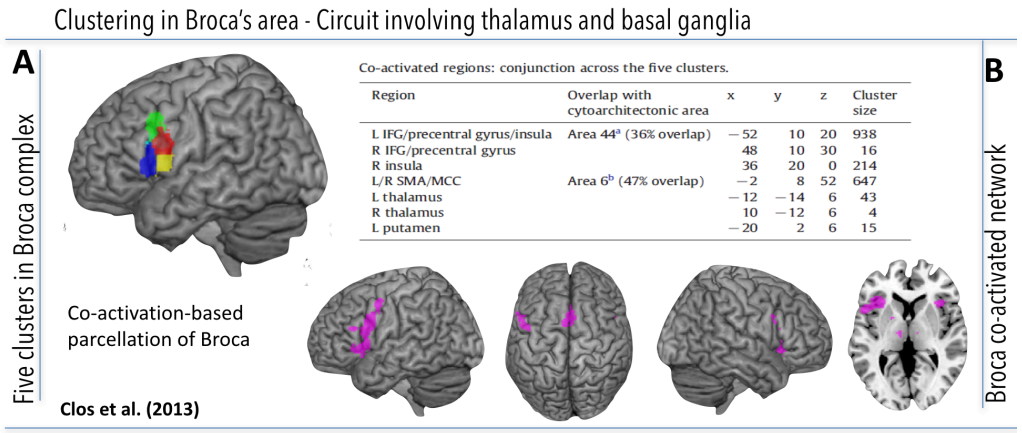
Overall, this finding suggests that the SMA activation patterns can be interpreted as linked to the cost of assigning object thematic-role in a non-canonical word-order, where the object is fronted, and crucially leaves a gap behind, irrespective of the type of

syntactic movement yielding these word-orders variations.

In sum, the SMA response pattern to a [± overt] realization of the dependency-link and to a [± pronominal] feature of the covert element found in object position seem to indicate what is at stake here is an effect of the difficulty in the *retrievability* of the object when a [-anaphoric and -pronominal] empty syntactic element is present: only the gap in (c5), when compared to a pronominal – be it lexically overt like in (c4) or covert like in (c6) – yields this effect.

Sub-cortical post-hoc ROI analyses

To better elucidate the role of this more peripheral area of the Sentence Network, we decided to focus on some areas that are often jointly activated in the studies we reviewed above. This brought us to add three Sub-cortical aposteriori Region of interest to our analysis (cf. ROI Table, p.647). Several fMRI studies report sub-cortical activations in word-order manipulations where the object is fronted, respectively thalamic activation in nucleus Anterior and Medial thalamus (Bornkessel-Schlesewsky et al., 2011) and Left caudate nucleus (Bornkessel-Schlesewsky et al., 2009). Particularly, manipulating feature animacy in German ditransitive passive constructions, Grewe et al. (2006) reported bilateral Basal Ganglia activation in sentences with two animate arguments.



Co-activated regions: conjunction across the five clusters.

Region	Overlap with cytoarchitectonic area	x	y	z	Cluster size
L IFG/precentral gyrus/insula	Area 44[a] (36% overlap)	−52	10	20	938
R IFG/precentral gyrus		48	10	30	16
R insula		36	20	0	214
L/R SMA/MCC	Area 6[b] (47% overlap)	−2	8	52	647
L thalamus		−12	−14	6	43
R thalamus		10	−12	6	4
L putamen		−20	2	6	15

Figure 7.29 – Adapted from Clos et al. (2013).

Moreover, clear evidence exists that Broca in its different sub-parts co-activates with a circuit involving Thalamus and Basal Ganglia. As illustrated by Figure 7.29 a study of coactivation-based parcellation of Broca Complex shows that the intersection of the five different sub-circuits of co-activated areas with the different portions of Broca (A) yields the purple circuit in (B) of areas encompassing IFG, Precentral, Insula, SMA and crucially also Thalamus and Putamen.

This results encouraged us to further investigating the possible function of these co-activated regions in the manipulation of word-order, we ended up in finding that the (pre-) Sensory Motor Region is reported to play a central role in lexical selection together with several sub-cortical areas.

Interestingly, recent models of lexical selection, like Crosson (2013), feature a (pre-)SMA-Thalamus network for the step of lexical selection in naming (see Figure 7.30). The involvement of Subcortical regions in processes of lexical selection is generally acknowledged in the literature. Particularly, the involvement of Basal Ganglia in mediating lexical selection processes has been successively reported in patient studies, for instance when patients have problems at inhibiting competing alternatives while selecting among semantically appropriate alternatives compared to novel word forms (Longworth et al., 2005).

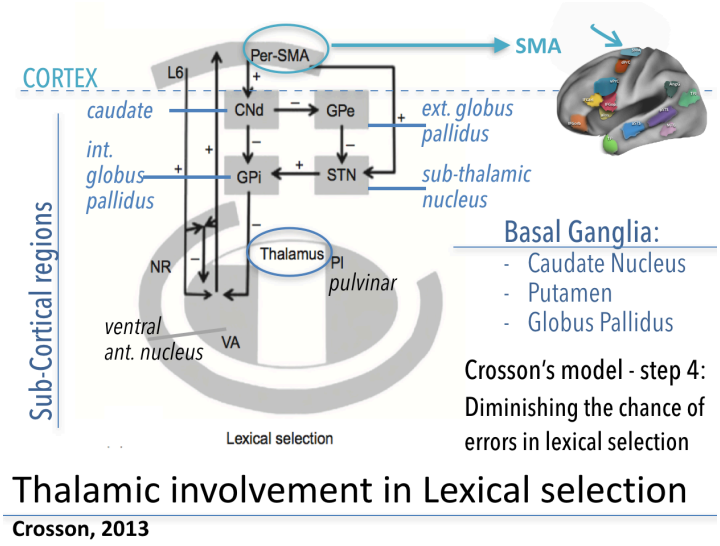


Figure 7.30 – Crosson's model of lexical selection involving Sub-Cortical areas and SMA region. Lexical selection is the fourth and last step of his Naming process model.

While speech fluency deficits are commonly associated with lesions to the Putamen and/or to the Globus pallidus (for speech motor-control function of Basal Ganglia see Kotz et al., 2009), damage to the head of the caudate nucleus is associated to lexico-semantics problems, encompassing the use of rules that govern the construction of sequences of morphemes, like in the production of regular English past tense (verb + *-ed*) (Ullman et al., 1997).

Furthermore, a study using electrical stimulation of the head of Caudate Nucleus showed an effect of perseveration of previously named picture (Gil Robles et al., 2005), thus contributing to confirm the role in selection of linguistic items, that this SMA-Subcortical network may underpin[29]. In sum, we can say that these studies consistently support the a role of Basal Ganglia in inhibiting competing alternatives during selection.

Hence, given this clearly established link of the SMA with word-order manipulations and object /subject animacy feature on the one hand, and the recurrent joint activation of SMA (pre-SMA) and sub-cortical areas (basal ganglia and thalamus) on the other hand, we decided to preform a post-hoc ROI analysis to establish if the condition activating the SMA (c5) also activate the sub-cortical lexical selection network, when contrasting two Left-dislocated sentences differing in terms of animacy of the object and also crucially featuring a gap or a null pronominal.

ROI analyses show that both (c5>c4) and (c5>c6) present a significant effect in the left Putamen reported in Grewe et al. (2006) in the contrast between sentences with two animate arguments (SAO vs. OSA)[30]. This confirms that the SMA contribution is linked to the lexical selection of the object among competing sentence-initial animate NPs when a [-pronominal] empty element is present at the original object position, i.e. a gap.

In the barplots in Figures 7.31B and 7.31C) we namely observe two fundamentally different patterns across the language network and that Sub-cortical regions pattern only

29. Note that Thompson-Shill et al. (2009) co-localizes stroop-effects and syntactic ambiguity in the SMA.

30. In an analysis including the mean reaction times per condition and per participant as regressors.

with the SMA activation profile[31].

If we consider the difference between the overt and covert realization of the pronominal we can note that a significant effect is observed in the pSTS (see Fig. 7.31B). This activation could reflect the additional gender agreement process that is needed when the referential link is established based on the presence of an overt pronoun[32].

7.4.3.3 Null Pronoun (c6) compared to resumptive and Gap

As for the comparison between (c6) condition with a null pronoun and the other two conditions (c5) and (c4), the whole-brain analyses did not show any cluster of activation, indicating that this condition with an inanimate fronted object evidently recruits less cognitive resources compared to the ones where an animate object is found in sentence initial position either as the result of Topicalization or as an *in-situ* Topic.

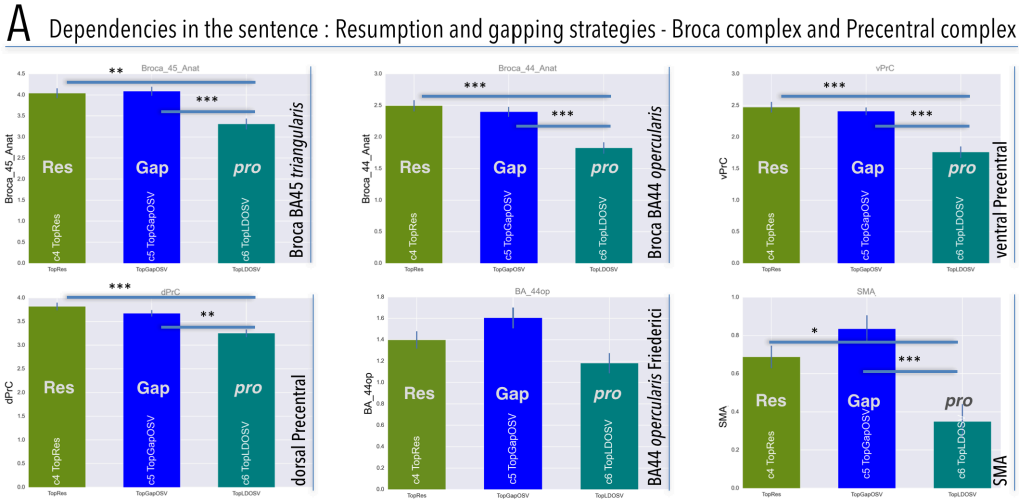
However, ROI analyses revealed a significant increase in a single Region, the left Middle Temporal Gyrus (see Figure 7.31B), thus confirming our predictions based on the findings of Santi and Grodzinsky study on intra-sentential dependency-links (2012). Specifically, this region was reported as the only cluster of activation in the contrast opposing parasitic gap to the optional realization of an overt pronoun in wh-questions like ‘‘*Which paper did the tired student submit ___after reviewing ___/ it ?*” Given that we observe lMTG increased activation both when null-pronominal is compared to Resumptive (c4) and Gap conditions (c5)[33], we follow previous findings demonstrating that the lMTG is involved in lexical and semantic processes (Humphries, et al., 2006; Kotz et al., 2002), and interpret its activation as being linked to the covert lexical dimension of the ‘little *pro*’ left by movement[34].

31. Note that Middleton and Strick (2000) bring evidence for the existence of a Basal ganglia and cerebellar loops in higher cognitive functions (i.e. not only in motor circuits), which could explain why our Cerebellar ROI shows and increased activation only for Gapped Topic-comment construction in (c5) against null pronominal *pro* (c6). This actually shows that cerebellar ROI where we also found a linear increase in number of syntactic positions in the French fMRI experiment, is indeed involved in the representation or processing of empty syntactic gaps.

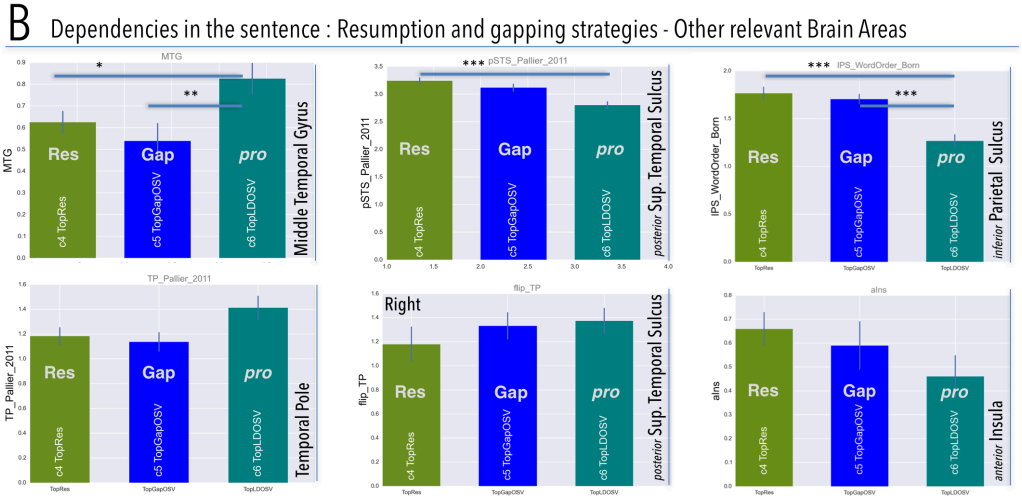
32. As shown in the Stimuli presented in the Annexes (§C.3 p. 887) we purposely assigned clear gender to the proper nouns in Topic position to facilitate the co-referential link.

33. The neuro-image reader can note that this apriori ROI by Santi and Grodzinsky is a few voxels apart from that we use in previous experiment from the study of Shetreet and Colleagues on NP-movement, for which the authors argued a possible implication in the morphological change implied by the realization of unaccusatives by affixes determining verb classes in Hebrew. We could maybe speculate that this areas achieves some lexico-morphological processes as the one that could be at stake in the lexically covert realization of a pronoun.

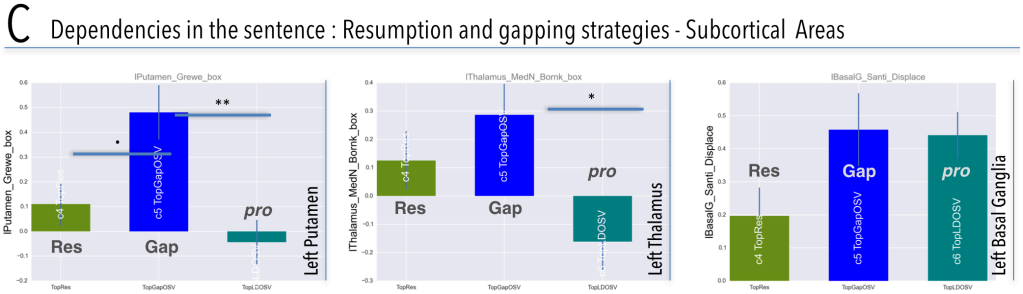
34. This actually contradicts the prediction we made and the results we observed are not linked to the [-anaphoric, +pronominal] features but of the covert realization of the pronominal in (c6)



(A) Experimental manipulation of the [+/-overt] realization and the [+/-pronominal] feature of the Topic-Comment dependency-links.



(B) Experimental manipulation of the [+/-overt] realization and the [+/-pronominal] feature of the Topic-Comment dependency-links.



(C) Experimental manipulation of the [+/-overt] realization and the [+/-pronominal] feature of the Topic-Comment dependency-links.

Figure 7.31

7.4.3.4 Recapitulation of the Dependecy links - ROI results

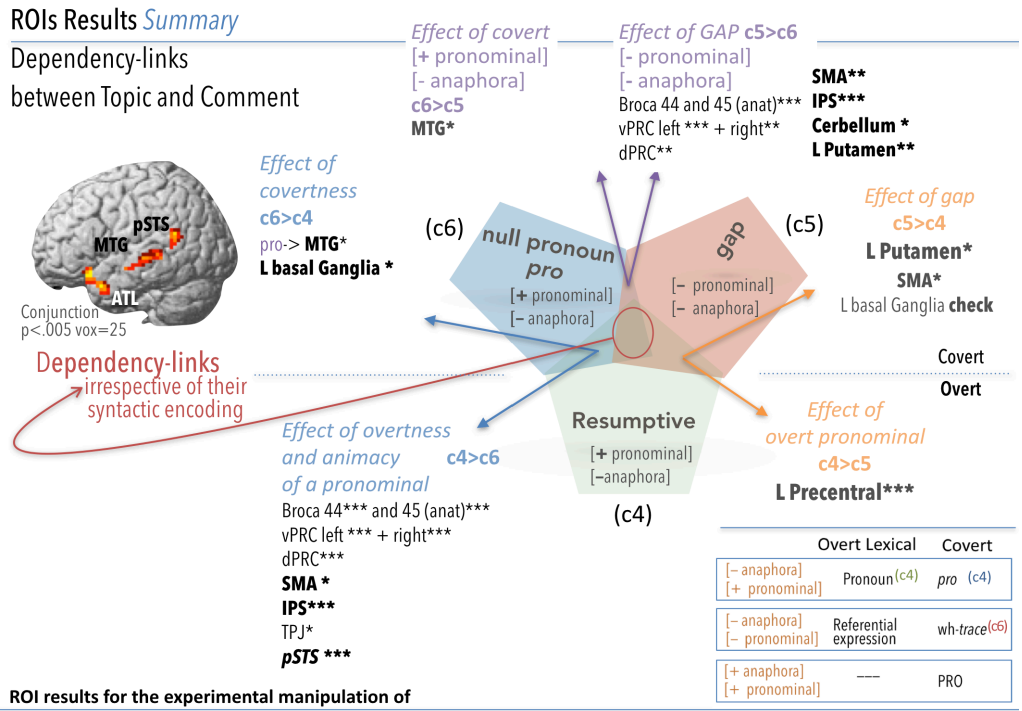


Figure 7.32 – Summary of the ROIs analysis for the experimental manipulation of the [+/-overt] realization and the [+/-pronominal] feature of the Topic-Comment dependency-links. See in bold the discussed results.

7.4.4 Contrastiveness and Focus out of the Left-periphery

One of our aims in contrasting (c7) and (c8) conditions was to determine whether separable subsystems in the Sentence Network are observable when pragmatic effects are obtained through morpho-syntactic marking, like in (c8), compared to a word-order strategy to achieve pragmatic markedness, like in (c7).

Secondly, contrasting these two conditions we will be able to observe which brain areas that are involved in pragmatic processes. Especially because in this configuration, discourse-related processes are not obtained through movement to the Left-Periphery, but the displaced constituents target the clause internal pre-verbal position (Fig. 7.33).

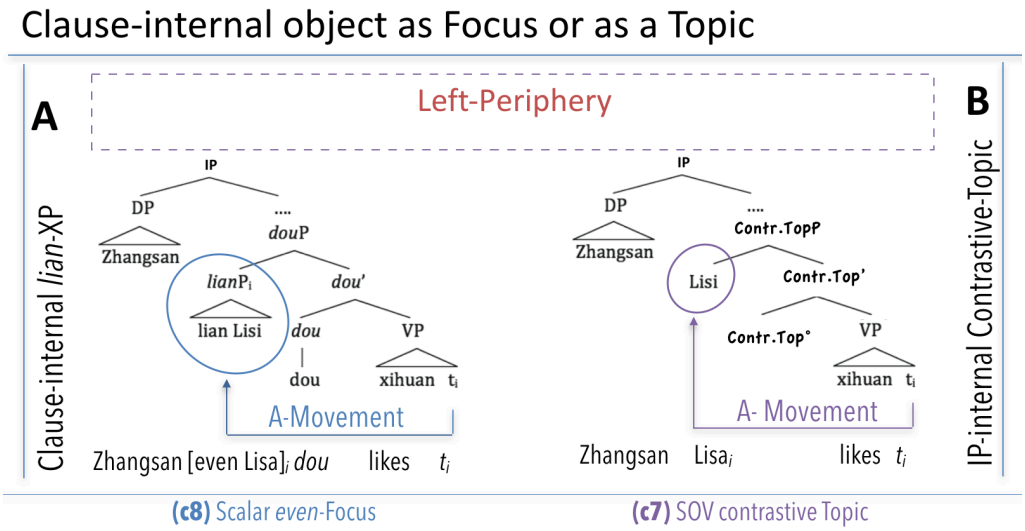


Figure 7.33 – Syntactic-tree representation of clause-internally displaces objects in case of Contrastive Topic on the right and even-Focus on the left. object as Focus.

When comparing morpho-syntactically marked Focus with contrastive Topic obtained by object preposing (SOV), our predictions, based on Allen et al. (2012)[35], are borne out. We observe in the whole-brain contrast a broad focal activation in the left anterior Temporal Region, thus confirming the involvement of this region in achieving syntactic structure building by morphological marking compared to simple sequential word-order[36].

As we already had the occasion to underline (cf. §1.4.5.1, p. 62), the role of this anterior temporal region is debated in literature.

Although it has been generally identified in aphasia and neuro-imaging studies as playing a role in semantic integration processes, it has been associated to structure building

35. This study found increase activation of the aSTS/TP when comparing "*Give this book to Joe*" where the object is in its canonical post-verbal position and dative case assignment is realized through overt morpho-syntactic marking, thanks to the preposition '*to*', to "*Give Joe this book*" where reliance on word order cue achieve the argument role assignment.

36. Feng et al. (2015) found that the contrast between preposed objects introduces by disposal particle ba with SVO baseline [(S + *ba*0 + V) > (SVO)], elicited left and right superior and medial frontal gyrus (BA 9/10). This contrast involves however both word-order change and presence of morpho-syntactc material.

and syntactic processes by a number of converging studies comparing regular sentences to scrambled version having the same lexical material (e.g. Vanderberghe et al. 2002), and by studies investigating the processing of morpho-syntax during sentence comprehension (cf. Dronkers et al., 2004, and also Hagoort, 2005; Humphries et al., 2006)[37]. Specifically, Humphries et al. (2006) reported stronger activations in the left ATL to sentences than to lexically identical lists of words which did not contain grammatical words.

Figure 7.34 – Whole-brain contrast between (c8) lian Focus sentence> (c7) Positional Focus SOV sentence, p<0.001 uncorr. Effect of having morphosyntactic marking lian ...ye to focalize the constituent this book zhe-yi ben shu 这一本书.

Moreover, syntactic priming[38] effects reported in the neuro-imaging literature reveal activation of lATL too (e.g. Noppeney and Price, 2004). Crucially, these experimental designs compared to other syntactic priming designs typically obtain priming adaptation effect by repetition of function words with different lexical material (cf. 2.3.4, p.154 for a comprehensive discussion). Thus, if we consider that the main syntactic difference between condition (c7) and condition (c8) is linked to the presence of two extra function words of the 连... 也 *lián...yě* construction, the above findings appear to be consistent with our observation.

A lately emerged account for this area is to be found in Brennan et al. (2012), who report that lATL activation is also correlated to a syntactic-tree node count measure providing an estimate of the amount of structure built word-by-word[39] (see Section §2.3.3.3, p.151). This finding can also be linked to what was tested by our contrast, in that as we see in Figure 7.33 the presence of morpho-syntactic marking adds complexity to the syntactic-tree increasing the number of its final syntactic nodes (cf. Badan and DelGobbo, 2015).

In sum, the focal activation of anterior Temporal region in this contrast is coherent with the proposition that lATL plays a role in the morpho-syntactic construction of the sentence structure. Although the configuration of our two experimental conditions does not allow the separation of pragmatico-semantic properties from morpho-syntactic marking, we follow Newman et al. (2015), who reported anterior temporal activation

37. Holland, R. and Lambon Ralph, M. A. (2010). "The anterior temporal lobe semantic hub is a part of the language neural network: Selective disruption of irregular past tense verbs by rTMS". *Cerebral Cortex*, 20, 2771–2775.

38. An adaptation effect observed in the brain when presenting repeatedly the same syntactic structure with different lexical material.

39. Not entering into the details of the parses analyses represented in the syntactic trees we can say that the authors correlated with brain activation the number of terminal nodes in a fragment of an auditory presented story.

comparing morphologically complex inflected word forms to word-order compositions in ASL (American Sign Language), and advance that a possible interpretation of the present finding is that the TP/aSTS is recruited more heavily when morpho-syntactic cues are present for the building of syntactic structure compared to a condition where pragmatic markedness is obtained through word-order information alone.

We can therefore conclude that our findings replicate and confirm both Newman et al. (2015) and Allen et al. (2012) results, and further speculate that the crucial step of *feature matching* (cf. discussion in §2.3.2.2, p. 138) that is to be performed in presence of function words to merge them into Phrases, could be processed in this area; this would also explain the fact that it has been reported both for semantic and syntactic integration processes in the literature.

7.4.5 Fine-structure of the Left-Periphery in the brain

The linear contrast assigning to the different conditions the the value of their ordinal position in the Left-Periphery and Low-Periphery – [c3> c2 > c4 > c5 and c6 > c7 > c8] – revealed at an uncorrected threshold of $p < 0.005$ (cluster-extent correction of 50 voxels) a unique cluster in the dorsal Precentral top-most part at the limit with Middle Frontal Gyrus, while at a $p < 0{,}001$ (voxel-wise) threshold we observe two bilateral smaller clusters in the same area the Precentral (aal) Brodmann area 6.

First of all, this activation confirms what three of our minimal clusters had already revealed about the implication of left Precentral Gyrus in conditions where the portion of syntactic tree to be represented is more important given the higher position of the functional projection where the sentence-initial element is hosted (cf. [c1>c2], [c3>c2] and [c4>c5]).

Secondly, we can link this activation to the one that was revealed by the linear contrast of the number of syntactic positions (i.e. overt words + gaps) in French Questions and declaratives in the previous chapter. The peak activation of these two effects are not too far apart in the dorsal Precentral area, although for the linear increase fort he height in the tree is leaning more onto the middle frontal gyrus (mFG). The fMRI literature overview presented in chapter 2 already demonstrated the strong correlation between activation in the dorsal/ventral Precentral and Syntactic complexity manipulations including syntactic priming (for additional evidence refer to the Annexes on precentral gyrus - BA6, §D.1.0.2, p.908). As for the interpretation of the activation in this –often reported, but hardly discussed– area, we relay on studies showing that it is also implicated in the deployment of strategic processes to encode sequentially ordered of information (see Marshuetz and Smith, 2006; or Rypma and d'Esposito, 1999).

This interpretation is further supported by the fact left medial frontal gyrus and Precentral (BA6), is stably associated with working memory processes since the nineties (e.g. Jonides et al., 1993; Schumacher et al., 1996; LaBar et al., 1999)[40].

Crucially, in a very thorough review[41], Cabeza and Nyberg (2000) reported reliable activation of Precentral/GFm for *spatial* and verbal tasks that require the active maintenance of abstract representations, while no reliable activation in htis area is observed

40. Importantly the two first are PET studies that converge in showing the amodal recruitment of the Precentral complex both verbal working memory and spatial working memory, whereas the last MRI study reports an overlap in this area between the networks respectively engaged in spatial attention and working memory.

41. A review 275 PET and fMRI studies, and 421 contrasts.

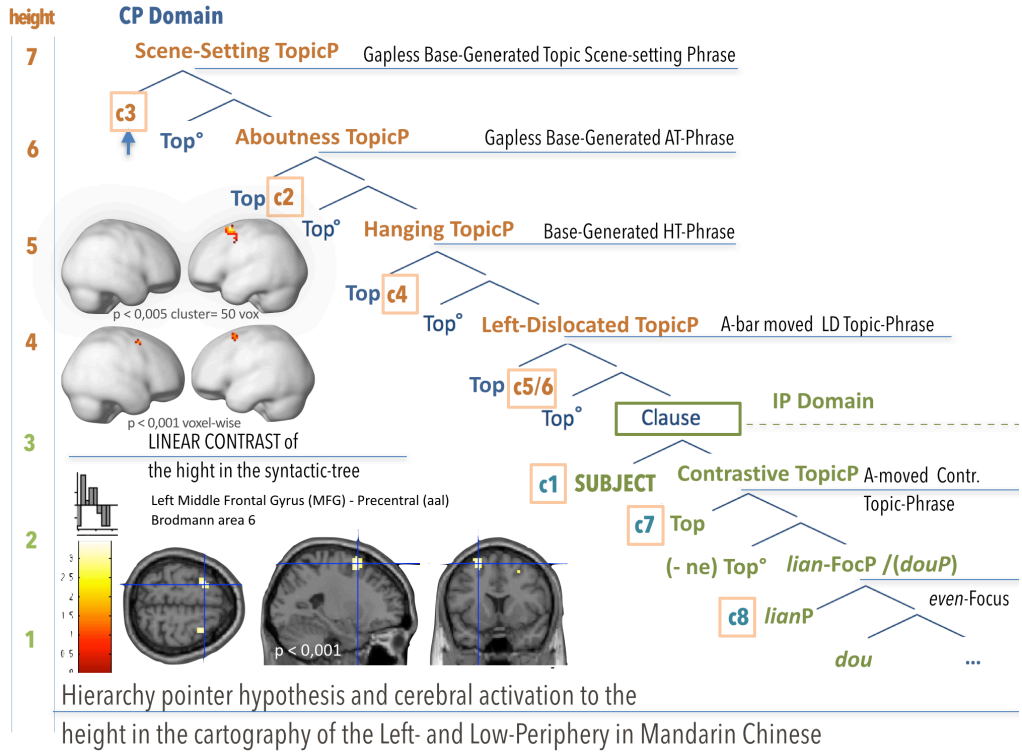


Figure 7.35 – Whole-brain linear contrast of the highest position occupied by each experimetnal conditions in the syntactic structure skeleton on the sentence as delineated by the carthogapphic analysis of Badan (2007) and refiend by us. p<0.001 uncorr voxel-wise.

for tasks involving iconic stimuli like objects and faces.

Interestingly, Newman et al. (2015) report the dorso-lateral Precentral cluster as the most extensively activated area for their contrast opposing word-order based syntactic structure building against inflectional morphology cued structures. Interpreting this result the authors put forward the implication of this area in the increased working memory demands that the absence of morpho-syntactic information is causing, when word-order is the sole device on which to resolve syntactic structure building.

Hence, we continue the argumentation we put forward in our interpretation of the linear increase for syntactic-tree complexity in Precentral complex for French in previous chapter. As more dorsal (/rostral) Premotor cortex (BA 6) has been identified to be involved in structure-dependent computation, such as sequential ordering of hierarchical structures in working memory (e.g. Christensen, 2010; Hanakawa et al., 2002), and in Stroop tasks (Marshuetz, 2005), we propose that representing a larger portion of syntactic hierarchy to process syntactically more complex sentences is actually tapping into the kind of cognitive resources for the active maintenance of abstract representations that reported to be underpinned by Precentral (mFG) Complex.

As we had the occasion to argue in previous chapter, this activation in Precentral Complex is compatible with the fact this region was reported to be involved in both rule-based mental-operation tasks and syntactic computation linked to movement trans-

formations (Ben-Shachar et al., 2003 and 2004; Christensen, 2008 and 2009; Dogil et al., 2002; Röder et al., 2002).

For instance, Hanakawa and Colleagues (2002) PET study revealed premotor cortex (Precentral Gyrus, BA 4/6) and SMA are constantly activated across three different non-motor tasks that involved rule-based and non-motor mental-operation like numerical (i.e.serial mental addition), verbal (i.e. mentally advancing the day of the week according to a numerical cue), and spatial tasks (i.e. memorize the successive location of a marker in a grid see Figure 6.51, p.579). Crucially, these tasks which uniquely involved cognitive, rule-based manipulation of representations in working memory (see also Stowe et al., 2005 for a review), show that Precentral Complex is involved in the rule-base association of symbolic cues and their processing[42].

The characterization we offer here is in line with studies observing its activation in mental arithmetic processes (Dehaene et al., 1996), but also in studies comparing mathematical operation and language -related activations to find neural dissociations between math and language. For instance, Monti and colleagues (2012), report an overlap of the network elicited by the processing the hierarchical structure of algebraic expressions and by natural language sentence processing in the dorsal precentral (middle frontal gyrus - BA6), and the medial segments of precuneus (BA7). In this way our functional characterization of this area could actually also shed a new light on some of results from the latest study by Amalric et al. (2016), where the dorsal portion of Precentral cortex was recruited both by mathematical and linguistic activity/reflection (A. Amalric p.c., 2016:p.4).

While in chapter 6, we had advanced that it might be linked to the Search step of Internal Merge operation involved in syntactic movement and the modification between linear-order and hierarchical structure that syntactic transformation yield in the sentence-unit, we can here refine this interpretation by saying that this area is actually involved in a more abstract process of holding in working memory the representation of the *rule-based ordinal hierarchical patterns* of the additional portion of syntactic-tree that are built to host higher elements in the structural sentence skeleton that has been proposed by cartographic approach.

All in all, these findings seem to converge and indeed explain why the linear increase related to higher and higher positions in the syntactic-tree skeleton is observed in an area also dedicated to (amodal) serial working memory of abstract patterns, and why fMRI studies on syntactic movement and word-order variations in the last twenty years nearly systemically reported an increased activation of this area for the more complex sentence conditions.

7.5 Summary and Perspectives

This study was an attempt to unify into a pluri-disciplinary approach the results from fine-grained linguistic formal theory, the syntactic analyses offered by the literature on Mandarin Chinese together with the results from cognitive neuro-imaging of language and hypotheses driven form aphasic linguistic behavior.

We searched for neural correlates of linguistics analyses and theoretically postulated

42. Importantly, Hanakawa et al. added also a verbal rehearsal task to their experiment which revealed a focal activation of distinct and non overlapping areas namely the frontal operculum and bi-latereral cerebellar hemisphere

syntactic derivations of the sentence structure at the brain processing level, and observed several effects related to the representation of nominal elements hosted in the sentence-discourse interface layer, that we can summarize in the following:

1. The representation of elements hosted in the sentence-discourse interface layer (i.e. CP or Left-Periphery) involves Broca complex and particularly Pars Triangularis as hypothesized by the Tree-pruning Hypothesis (Friedmann, 2002/2006) and the Sentence Domain hypothesis (Christensen, 2008).
2. The complexity parameter linked to the height of the syntactic position – the functional projection – where a Topic element is represented in the sentence's structural skeleton involves the dorsal part of Precentral Cortex.
3. The presence of a NP in the sentence-discourse interface layer triggers contextual integration processes that are show in the activation of scene-sensitive brain regions. As already revealed by our ERP study in chapter 5 the function of setting the spatial frame of the Comment-clause characterizing Scene-setting Topics was reflected by context integration processes that were observable during the online incremental processing of this sentence structure in context.

In sum, we can conclude that Broca's area and Precentral Cortex are the two brain regions underpinning the representation of Topic-comment articulation irrespective of movement-derivation and of the presence of dependency-links between Topic and Comment. On the contrary, an very important effect on the extent of activation of the Sentence Network is caused by syntactic-movement related processes, which is also comprising a mid-Temporal regions in the right hemisphere.

As for the establishment of filler-gap dependencies and the referential assignment of the object role in the different Topic-comment constructions under analysis, we observed:

1. an increased activation of the *SMA-Subcortical* network dedicated to lexical selection only when a gap left by movement is found at the original object positions. This effect that probably also linked to a possible "competition" of the two sentence initial animate NPs (respectively, object Topic and Subject).
2. When intersecting the different effects for the three Topic-Comment articulation where the dependency is achieved by means of a Resumptive, gap or null pronoun, the common substrates for the process of establishing intra-sentential dependency-links show a strict temporal distribution (irrespective of their syntactic derivation).
3. In the comparison between achieving syntactic structure building through simple word-order linearization or by means of function words, the anterior Temporal region (TP) shows to be preferentially recruited for the building of syntactic-phrase structure in presence of function words.

Broader Implications for the functional characterization of the Sentence Network

More generally, the findings we discussed suggest a temporal-frontal division of labor between three main dimensions distinguishing the sentence-unit: its Linearity and Hierarchy on one side, and the dependencies that are achieved inside the sentence on the other. Resuming to a discussion in chapter 2 (§2.3.3, p.143) on the advantages to represent the linearity, hierarchy and dependency relations in the sentence-unit through the syntactic-trees '*format*', we can draw the following conclusions.

Firstly, we can say that in this fMRI experiment, the conjunction analysis intersecting the three different effects for Topic-comment articulations, where the dependency is

respectively achieved by means of a Resumptive, a gap or null pronoun, reveals that the establishment of intra-sentential dependency-links elicit an increased activation with a strict temporal distribution, irrespective of its syntactic derivation.

Secondly, the contrasts between Base-generated Aboutness Topics with no dependency-links against basic SVO Declarative sentence with identical word-order but no Topic, reveals the recruitment of cortical resources that exclusively frontally distributed.

These two findings put together, inform us about the existence of separable sub-systems in the Sentence Network, respectively dedicated to (i) the hierarchy dimension of sentence-unit, and (ii) to the establishment of non-local dependency-links across the sentence structure.

As illustrated by Figure 7.36,this fMRI experiment revealed that, while the representation of dependency-links mainly involves temporal areas with sub-cortical connections to the SMA in case of gaps (B), the complexity of syntactic hierarchy in absence of movement and of intra-sentential dependency-links uniquely activates frontal areas (A). The division of labor emerging from our findings is in line with the one high-

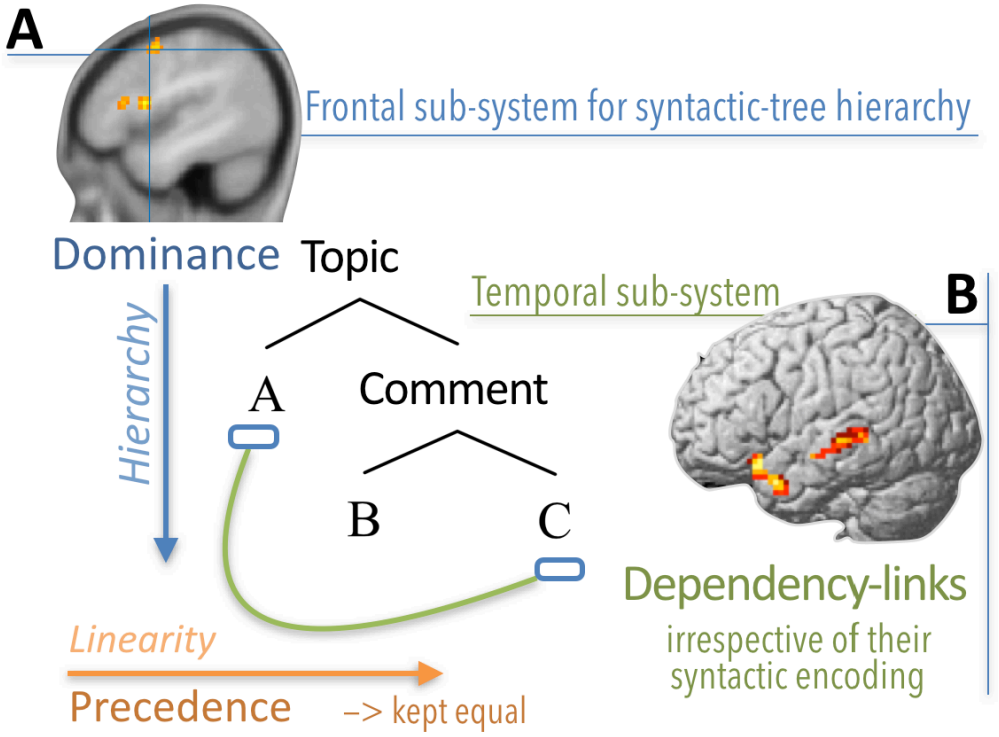


Figure 7.36 – Linearity and Hierarchy dimensions of the sentence-unit and their Fronto-Temporal task-sharing. In this experiment, (A) the representation of dependency-links mainly involves temporal areas with sub-cortical connections to the SMA, while (B) the syntactic hierarchy in absence of movement and of intra-sentential dependency-links uniquely activates frontal areas.

lighted by the extended Argument Dependency Model (eADM) model by Ina Bornkessel-Schlesewsky and Matthias Schlesewsky, one of the rare models in the literature on the

neuro-implementation of sentence-level mechanisms. As illustrated in Figure 7.37, this model focuses on modeling, in a dual stream fashion, two different dimensions of the grammatical and argumental dependency relationships present in the sentence-unit. It mainly distinguishes the relationships based on sequential information reflected in word-order encoding, and the dependencies achieved through non-sequential encoding in the sentence-unit. Hence, drawing the link between the cerebral localization of the distinc-

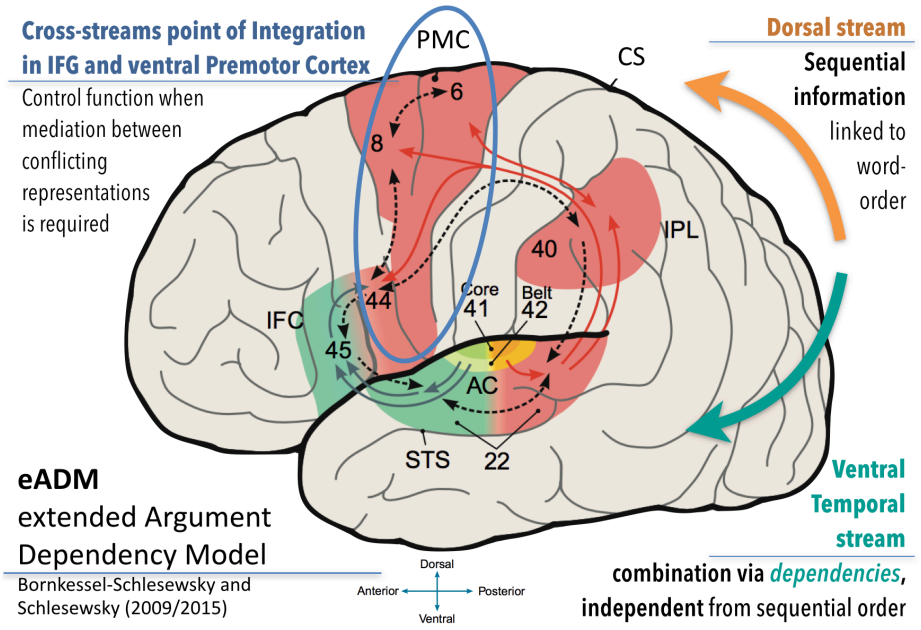


Figure 7.37 – eADM - extended Argument Dependency Model - A computational division of labor between the two streams. Dual stream model of the dependencies between arguments. Adapted from Bornkessel-Schlesewsky and Schlesewsky (2009 and 2015).

tion proposed by the eADM model and our findings an unexpected convergence appears. Notably, this model assigns to its temporal stream the function of processing combinations via non-contiguous dependencies, independent from sequential order, which is the kind of effect that was indeed observed in our conjunction analysis reveling the shared network for the establishment of dependency-links.

On the other hand, the model's second stream deals with sequential information reflected in word-order encoding, and encompasses dorso- frontal areas, which are namely the one that we observed by investigating the syntactic hierarchy dimension of Aboutness Topic conditions versus our minimally differing baseline. The hierarchy of our Topic-comment constructions is achieved by word-order syntactic encoding, and reveals to be dealt by the dorsal-frontal stream as predicted by the eADM model.

To conclude, we recapitulate the main findings of this chapter by mapping them onto the Sentence Network in Figure 7.38 to push further the functional characterization we promised to gradually achieve.

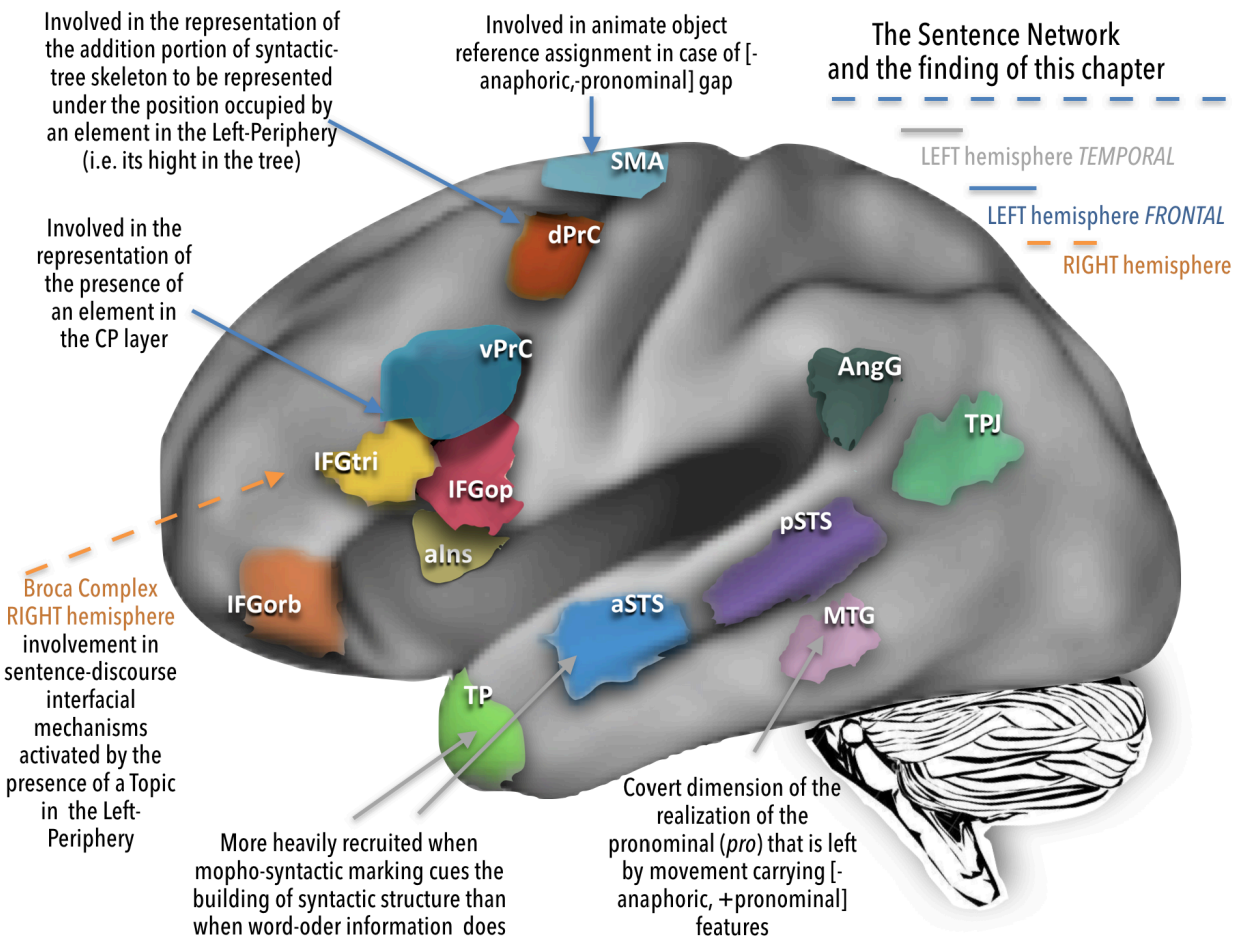


Figure 7.38 – Mapping results onto the Sentence Network, to functionally characterize the areas and the different sub-processes that it houses.

7.5.1 Implications for Mandarin Chinese research

Topic-prominence parameter, one of the most distinctive typological features of Mandarin Chinese, is here studied using fMRI brain imaging technique, to answer the question about what it means from the brain point of view to build a 'Chinese-style'Topic (B-G Topics), structuring the sentence around the notion of Topic rather than that of Subject.

Concretely, our results show that one of the core areas of the language network, namely Broca's Pars Triangularis is implied in the building of Gapless Topic-Comment sentences. Moreover, we were able to observe different neural signatures for Base-Generated Topics –the most salient feature of a topic-prominent language like Mandarin Chinese –and for Left-Dislocated Topics.

The so-called 'Chinese-style Topic' (Chafe, 1976:50) activates a sub-part of Broca's area pars Triangularis and the Pre-Central cortex, while Left-Dislocated Topics activate the Sentence Network to a wider extent and bilaterally (comprising anterior Temporal Lobe activation), and elicit an additional right hemisphere focal activation of the Middle Temporal Gyrus.

Secondly, in Left Dislocated Topic-Comment sentences, establishing a filler-gap dependency in absence of a Resumptive Pronoun in gap position yields more activation in the SMA, even when compared with sentences with inanimate fronted objects where resumption is ungrammatical.

Lastly, morpho-syntactic marking of Focus in Mandarin Chinese 连... 也 *lián...yě* *even*-constructions compared to preposed object contrastive Word-Order (SOV) reveals focal activation in the left anterior Temporal region, a broad area presented in the literature as hosting in its different sub-parts semantic integration processes and syntactic complexity effects such as syntactic-tree node built at each word.

To conclude our brain-imaging findings are not only compatible with a syntactic movement analysis for Left-Dislocated Topics (A-bar movement), but they also provide neural evidence for the different syntactic derivations for 'Chinese style' Gapless Topic (Base-Generation) and Left-Dislocated ones.

To our knowledge this is the first experimental evidence that the brain makes the difference between syntactic structure and syntactic derivation. Last but not least, it shows neuro-linguistic observations can provide additional experimental evidence for linguistics questions and debates.

7.5.2 Future research

We will now move to some remarks on possible future research, highlighting some follow-up questions.

Several questions remain to be addressed in future research. The first is surely linked to confirming the effects we observed for Base-Generated elements in the CP-layer in Broca Pars Triangularis. One strategy could be either to choose a more sensitive brain-imaging technique like intra-cranial recordings or by building a refined experimental design where the intersection of several condition with adequate baselines for each would allow to observe again the effect of hosting an element in the Left-Periphery, irrespective of movement derivation and of pragmatic interpretation. One idea could actually be to test the situation where even-Focus targets the Left-Periphery and see id the brain maps obtained for this condition actually differ from those we obtained when it targets a clause internal position. This would balance our design that avoided strong pragmatic

interpretations to be yielded by movement to sentence-initial position.

The second aspect that could be further investigated is linked to the difference between overt marking of Topic-comment hierarchy by Topic heads compared to the word-order based strategy we decided to investigate. Both Chinese and Japanese are Topic-prominent languages, but Topic markers are obligatory in Japanese, while they are optional in Chinese (Xu, 2006).

In the current study, we decided not to manipulate the experimental sentences along the dimension of the overt marking of topicality because of the variability of the informers judgments on the contrastive role of some of them. However, for future research it would be essential to clarify this point and to test for Topic-Comment structures where overt marking of Topic (heads) adds to the word-order positional cues. This would namely help confirming the role of anterior temporal region and to better understand the contribution to the building of the syntactic structure that overt morphological means feature at the brain processing level.

A third open issue is linked to a complexity dimension linked to movement we did not investigate, although Mandarin actually present some good configurations to test the complexity effects linked Topic's moving across several clauses and the presence of multiple gaps.

Namely, following Santi et al. (2015), it would be interesting to discover how the crossing of several sentence boundaries can actually constitute a complexity dimension implied in movement-related processes.

Particularly in light of the results present in the literature about Broca Complex (left and right) showing together with basal ganglia a linear effect of interveners[43] present between the two tails of a filler-gap dependency[44], it would be interesting to more clearly differentiate the complexity dimension liked to interveners from that of clausal boundaries.

For examples, as pointed out in chapter 3, to establish the relation between the Topic element (*zhe-ge haizi* 'this child') and the empty category left at the extraction site, multiple clause boundaries may be crossed as illustrated in the following example:

(268) 这个孩子$_i$，[$_{CP}$ 张三知道 [$_{CP}$ 李四看见 [$_{IP}$ 外婆在画 t$_i$]]]。

Zhège háizi$_i$, [$_{CP}$ Zhāngsān zhīdào [$_{CP}$ Lǐsì kànjiàn [$_{IP}$ wàipó zài huà t$_i$]]].
this-CL. child [$_{CP}$ Zhangsan know [$_{CP}$ Lisi see [$_{IP}$ grandma PROG. draw t$_i$.]]]

'As for this child$_i$, Zhangsan knows that Lisi saw the grandma drawing [him]$_i$.' (Hu, 2014)

An alternative idea could be to test for parasitic gaps in Chinese which can reach the number of three as in (269), and look for an increasing complexity effects in the SMA-Subcortical network related to the presence of multiple gaps linked to the same Topic element.

(269) Multiple gaps linked to the same Topic element

a. Parasitic Gap + adverbial clause: 2 gaps binded by the Topic:
那个员工 indi[在老板见过之后立刻就被开除了。

43. For a definition of interveners, see section §3.4.4.4 and Figure 3.24, p.400.
44. Already Makuuchi et al. (2013) reported a linear effect in basal ganglia (globus pallidus) of the interveners found within a movement dependency.

Nà-ge yuángōngindi [zài lǎobǎn jiàn-guò ___ PG_i zhīhòu] $\_\_\__i$ lìkè
That-CL. employee [at boss meet-ASP. ___ PG_i after] $\_\_\__i$ immediately
jiù bèi-kāichú le.
JIU was-fired PART.

'That employee, after the boss met [him], [he] was immediately fired.' from Ting and Huang (2008)

b. Parasitic Gap: 3 gaps binded by the Topic:
那个员工 indi[在老板见过之后立刻就被开除了，就很难过。

Nà-ge yuángōngindi [zài lǎobǎn jiàn-guò ___ PG_i zhīhòu] $\_\_\__i$ lìkè
That-CL. employee [at boss meet-ASP. ___ PG_i after] $\_\_\__i$ immediately
jiù bèi-kāichú le, $\_\_\__i$ jiù hěn nánguò.
JIU was-fired PART., $\_\_\__i$ JIU very depressed

'That employee, after the boss met [him], [he] was immediately fired.' from Ting and Huang (2008)

All in all, as we reviewed a variety of studies reported the basal ganglia and SMA to be sensitive to syntactic complexity (see also Prat and Just, 2011), the results of our study are encouraging us to pursue a more thorough investigation of the Cortico-Subcortical circuits involving Thalamus, Basal Ganglia, SMA and Broca Complex (Clos et al., 2013) in syntactic complexity processes, and specifically by focusing on the modulation of this circuits in presence of multiple gaps, multiple interveners or multiple clausal boundaries.

PART

Conclusive remarks and Further directions

Bornons ici cette carrière,
Les longs ouvrages me font peur ;
Loin d'épuiser une matière,
On n'en doit prendre que la fleur.

[Our peregrination must end there.
One's skin creeps when poets persevere.
Don't press pith from core to perimeter;
Take the ower of the subject, the thing that is rare.
Here let us stop, no further jaunting,
As I find lengthy volumes daunting,
That too well illustrate a theme,
When all that's wanted is the cream.
]

La Fontaine (1621-1695), translation by J. Hall and C. Hill, *Fables*, vol. VI.

We summarize here the main findings and consider the limitations of the current research and its implications for future linguistic and neuro-linguistic investigation of the sentence-unit.

Unconventionally, for a PhD -but conventionally for our Greek Drama structure- the reader will find here an *Epilogue* (695), whose purpose will be to add a little insight to what we leaned from our pluridisiplinary and theory-oriented research approach. This last chapter will tentatively make the reader grasp some epistemological insight of the current research by briefly developing a series of issues our reflection was literally 'dragged' to consider doing research.

Conclusion

山不在高，有仙则名。
　　水不在深，有龙则灵。[...]

[The value of a mountain is not given by its height but by the spirits that dwell in it.
The value of river is not given by its depth but by the dragons that live in it.[...]]

刘禹锡的《陋室铭》Liu Yuxi, "The Scholar's Humble Dwelling"

As the reader can see from the above epigraph, ancient Chinese poets and thinkers were already convinced about what a long-standing desert addict, like Antoine de Saint-Exupéry, had formulated in these terms: the essential is invisible to the eye. I hope the reader or the jury will agree that this neuro-linguistic *river* of sentence-units, brain-imaging data, literature findings, interpretations and reflections is also worth for its immaterial "dragon ideas" and for the non-tangible intellectual itinerary it represents. The amount of experimental results it contains is difficult to condensate in a few final pages, although this is what academic tradition and the theatrical narrative form we chose for structuring this work require. These conclusive pages are thus dedicated to a recapitualtive discourse offering a hint of a sequel. The fate of the characters, housed in each chapter of this disciplinary theatrical play, will be outlined.

Inscribing this work in the long-standing effort of philosophers, logicians, linguists and cognitive psychologists to provide new knowledge about the extraordinary linguistic capacity of man, we decided to root it in the reflection offered by Baroque times on what is unique in the human mind to favor its linguistic capacity for structured sentence-units, that not so trivially "say something about something". Thereby we sketched in chapter 1 a definition of man centered around his ability to utter syntactically structured sentences-units – the *homo phraseologicus*. This allowed us to critically delineate different aspects that current research usually takes for granted, such as the fact that mental objects like sentences are computations, or that there exist universal linguistic properties that make some grammars possible and other impossible.

Leveraging on the work linguists have been developing since the 1930s, we presented and followed the gradual settling of the theoretical and analytic tools that allow us today to build on the variety of experimental approaches using brain-imaging methods that have sought in the three last decades to **characterize the regions implicated in syntactic processes** and to reveal part of the **neural organization of language** – a key issue when one undertakes to unveil some of the crucial properties of the **sentence as a cognitive object**, and more broadly to understand what is uniquely human in language. Hence, using different **brain-imaging techniques** (fMRI and ERPs), this work embodies the attempt to narrow the gap between the fine understanding of sentence structures **across languages** revealed by linguistics, and the still relatively coarse-grained description of the neural circuitry that underpins sentence's cerebral representation and processing. Based on the rich experimental work in psycho-linguistics and in aphasiology presented in chapter 2, we estimated enough evidence had been gathered to consider that both cross-linguistic generalizations and the theory of syntactic-movement transformations – revealed by typological and formal syntactic theories – could have a cognitive relevance, and not only be represented in the mind/brain, but also be observable in cerebral activations patterns which can be recorded through actual brain-imaging techniques during sentence comprehension.

Pursuing the "holy grail" of grasping additional pieces of the complex puzzle of the neural underpinnings of the cognitive representation of sentence-unit, this dissertation was devoted to meeting the challenge of the integration of a cross-linguistic approach together with a fine-grained theoretical analysis of sentence structural representations to explore the neural underpinnings of the *sentence-unit* (chapter 2).

Our first linguistic and experimental focus examined two among the most essential properties of the sentence-unit: (i) its most general predicative function to "say something about something" and relative syntactic encoding, (ii) its structural openness to discourse-level interface, and more specifically the interplay between syntactic hierarchy and discourse context. Chapter 3 investigated how these properties are gathered by a particular bi-parted sentence articulation – Topic-Comment construction –, and to the special basic status this sentential articulation has in Chinese, given its typological characterization as a Topic-prominent language.

The psycho-linguistic behavioral investigation in chapter 4 revealed that Topic-Comment sentences in Chinese are equally well understood when the Topic's discourse referent is salient [+active, +accessible] and when it is not. Secondly, we showed that Gapless Scene-setting Topic-Comment articulations in a Topic-prominent language like Mandarin are comparatively acceptable in absence of the pause and Prosodic Boundary (final Topic syllable lengthening and pitch variation) usually characterizing them, a finding that is revealing how basic and fundamental Topic-Comment articulation is in Mandarin Chinese. We concluded that the Prosodic Boundary helps, but is not crucial for the comprehension of short sentences like the one we tested, and that in its absence the comprehension relays on word-order syntactic cues.

After having revealed, through phono-acoustic measures, the prosodic signature of the sentence's hierarchical articulation between Topic and Comment, and having confirmed the absence of behavioral contextual licensing (i.e. its basicness) in for Chinese, we approached the incremental processing mechanisms of a particular type of Gapless

"Chinese style" Topic in narrative context. Scene-setting Topic allowed us to track the online Sentence-Discourse interface strategy that is activated to understand sentences where the Topic has no selectional relationship with the verb (chapter 5).

Our research question about how the sentence's interface with discourse is processed by the brain, found a first answer in the ERP study in chapter 5. By manipulating the availability of early prosodic cues marking the Topic-Comment syntactic structure, we observed during online processing that sentence-discourse interface mechanisms like context information integration effects are activated earlier when the syntactic hierarchy between Topic and Comment is cued by Prosodic Boundary, and later when the sentence hierarchy is discovered and built on word-order grounds.

We concluded that building the sentence's hierarchical structure just on word-order grounds (without Prosodic Boundary marking) delays the integration of contextual informational load. The Prosodic Boundary conveying the syntactic articulation of the sentence thus represents an early trigger of the sentence-discourse interface and therefore activates its contextual information integration processes. Taken together, the ERP effects reported here, point to a sentence parser that waits the signal of syntax to engage into contextual discourse linking processes, which can be seen as offering some experimentally based arguments in favor of the "syntacticization of scope-discourse semantics" proposed by the cartographic approach (Cinque and Rizzi, 2010).

Moreover, the contextual interface mechanisms triggered by Topic-Comment syntactic articulation yields a multi-step integration strategy during sentence parsing, that is namely reiterately observed at Topic and Subject's time-windows. We interpreted these recurrent contextual bridging mechanisms as featuring a process of accessing or living-up discourse-context information to incrementally understand the link between the Topic and the Comment, when the Topic is not sub-categorized by the Verb, as it is the case in "Chinese style" Topics.

In a second phase, in order to continue investigating the syntactic encoding of the sentence-discourse interface, we followed the impulsion of fMRI research programs having emerged in the last ten years on the mapping onto the Brain of syntactic complexity and of sequence hierarchies processing.

Specifically, we pursued a direction that has increasingly proven to be felicitous and investigated the dislocation of sentential elements by syntactic-movement operations. As argued in chapter 2, we relayed on the complexity measures issued from syntactic-tree representation of sentential hierarchies and relations. We then explored the path of hypothesizing even more fine-grained mappings between syntactic knowledge about the sentence structure, and therefore took seriously several aspects of the linguistic description offered by theoretical linguistics.

Hence, under a cross-linguistic light and with a fine-grained set of syntactic descriptive tools, we addressed the general issue of the representation of hierarchical linguistic structures by the brain, thus questioning to what extent are the hierarchical sentence's structures postulated in linguistics corresponding to actual representations formats manipulated by the brain during sentence comprehension. To assess such a questioning, we investigate three fundamental characteristics of the sentence-unit:

– 1 the representation of the sentence's architecture into the three sentence domains determining basic sentence structure – namely VP, IP and CP,

– 2 the processes and operations that generate complex syntactic structures, like syntactic-movement transformations, and
– 3 the intra-sentential dependency relations achieved by overt or covert linguistic means.

One of the major difficulties in this regard is to find linguistic configurations enabling to separate these different complexity parameters, which are most of the time jointly present in a sentence. This is the point where our cross-linguistic approach became crucial and Chinese-style Topic constructions and French question formation with its different syntactic transformations, came into play offering two particularly interesting and optimal testing grounds.

On one side, French interrogatives and clitics represented an ideal testing ground to cast a light on the sentence complexity engendered by multiple and different syntactic derivations, thus allowing us to test for the complexity effects yielded by diverse kinds of syntactic-movements. On the other side, Chinese Topic-Comment constructions permitted us to observe the sentence's hierarchical organization of a sentence's articulation where the syntactic-tree complexity related to the presence of elements in the highest Sentence Domain (the CP-layer) is observable without movement, without embedding and without dependency-links, thanks to Topic-prominence parameter determining not only the basic status of Topic-Comment sentences, but also their Base-Generation in Chinese.

Concretely, searching for neural correlates of movement-related syntactic complexity in different types of movements, we observed several distinct effects. The effect observed for *wh*-movement in French *wh*-questions fully replicates previous findings in Danish and Hebrew *wh*-questions (Christensen 2008 and Ben-Shachar, 2004), and effects reported for other *wh*-movement constructions like Topicalization (Shetreet and Friedmann, 2014) or Relative clauses (Ben-Shachar et al., 2003). Our results also further confirm the involvement of Cerebellum and Precentral cortex in this syntactic transformation.

The other three syntactic-movement operations generating sentences having different complexity dimensions showed to be neurally distinct. Notably, the main evidence for the fundamental difference of *wh*- and Verb-movement in French questions is to be found in the interaction with verbal argument complexity, which was observed only for Verb-movement.

Clitic-movement together with NP-movement showed to have a more temporal distribution with an increased activation of aSTS, and Precuneus, and a unique frontal activation located in the MFG/SFG (BA9). Clitic-movement showed to preferentially elicit Temporal-Parietal-Occipital Junction (left posterior Supramarginal Gyrus) and more internal areas like the anterior Insula, the anterior Cingulate, the SMA and Precuneus. While NP-movement additionally showed two right lateralized clusters in dorsal Precentral and inferior Parietal region.

More generally, those two more local movements reproduced the temporal activation of anterior/mid-temporal lobe that were reported for other local movements in the literature, like dative-shift and negative-shift (Ben-Shachar et al. 2004 and Christensen, 2008), specifically in the aSTS.

An interesting element that was brought to light by this experimental investigation, is the recruitment of Temporal-Parietal-Occipital Junction (TPOJ) along with increasing derivational complexity: (i) the multi-step progression of clitic- and Verb-movements, (ii) the derivational complexity characterizing the combination of wh- and Verb-movement,

(iii) the combination of clitic and Verb-movement, and (iv) the presence of double clitics, all yielded the increased activation of this posterior temporal area. We interpreted these results by putting them in parallel with the increased activation in the overlapping posterior Supramarginal area obtained by Bachrach (2008), when correlating brain activity with a measure of the number of derivational steps of a parser, or even more simply by tagging syntactic displacement in narratives.

Our fMRI study on Chinese (chapter 7) focalized on two essential elements of the sentence-unit, its hierarchy and dependency-links. The sentence hierarchy dimension was addressed by contrasting Base-Generated Aboutness Topics and a minimally differing SVO sentences. This revealed, together with other contrasts, the implication of Precentral Gyrus and Broca complex in the encoding of Topic-Comment sentential hierarchy. All in all, we observed that the representation of elements hosted in the sentence-discourse interface layer (i.e. the CP) involves Broca complex and particularly Pars Triangularis, offering additional evidence to the Tree-pruning Hypothesis (Friedmann, 2002/2006) and the Sentence Domain hypothesis (Christensen, 2008). As for the sentence-discourse interface mechanisms activated by the presence of a NP in the CP-layer, we could observe that Scene-Setting Topics trigger contextual integration processes that we observed through the activation of scene-sensitive brain regions, like the Parahippocampal Place Area. These results converged with what already revealed by our ERP study in chapter 5. The function of setting the spatial frame for the Comment-clause, that characterizes Scene-setting Topics, was observable through the context integration processes taking place during the online incremental processing in context. In this regard, another aspect of convergence between our two fMRI studies is that right IFG appeared to be activated by the sentence-discourse interface semantic properties (i.e. Scope Discourse properties) of Interrogative Force in questions, and more precisely in yes/no questions in French (both Verb-inverted and simple declaratives with a question mark), and also by discourse interface properties of Topic-Comment constructions in Chinese.

One of the most original findings of this study was that the complexity parameter linked to the height of the syntactic position where a Topic element is represented in the sentence's structural skeleton (i.e. the functional projection) involves the dorsal part of Precentral Cortex, thus confirming the involvement of this area in syntactic-tree complexity.

Considering the issue of syntactic derivation of Topic-comment constructions in Chinese, the observed difference between Aboutness, Scene-setting and Hanging Topic on one side, and Left-Dislocated Topics on the other, offered some neurally-based arguments for the linguistic controversy on Base-Generated topics and Moved Topics in Chinese.

As for how the brain represents sentential dependency-links, the comparison of A-bar gaps with cases of overt resumption strategy, revealed the activation of an unexpected SMA-subcortical network for gaped structures. Specifically, we observed that the SMA-Subcortical network is recruited in case of reference assignment in presence of a gap, but not in presence of a phonologically covert pronouns (*pro*). Moreover, the conjunction analysis intersecting the three different effects for Topic-Comment articulations where the dependency-links are achieved by means of a Resumptive, gap or null pronoun, reveals that the establishment of intra-sentential dependency-links elicit an increased activation

of a strict temporal lobe distribution.

All in all, the neuro-imaging investigation of Chinese Left-Periphery clearly revealed a broad distinction between temporal areas involved in the establishment of dependency links between Topic and Comment, and frontal areas involved in the representation and building of Topic-comment hierarchy, in case of absence of movement and of intra-sentential dependency-links. Once more, these findings for Mandarin suggest a temporal-frontal division of labor between three main dimensions distinguishing the sentence-unit: its Linearity and Hierarchy on one side, and the dependencies that are achieved inside the sentence on the other. We interpreted this "task sharing" as being in line with the lADM model by Borkessel and Schleswesky (2009), which points to separable sub-systems in the Sentence Network, respectively dedicated to (i) the hierarchy dimension of sentence-unit, and (ii) to the establishment of non-local dependency-links across the sentence structure.

Given these results, we can conclude that both fMRI chapters 6 and 7 contributed to refining the study of syntactic knowledge in the brain by examining (1) the cerebral sentence structure representation, (2) that of intra-sentential dependencies (gaps or resumptives), and (3) the processing of syntactic displacement and related processes during sentence comprehension. Taken together, these findings offer some elements to functionally characterize the role of several areas of the distributed language network (see Figure 7.39, p. 690), but also highlight a broad fronto-temporal syntactic task-sharing, that was actually confirmed by the unsupervized analysis we performed in chapter 6.

In our unsupervized analysis we were able to observe the modulation of the distributed sentence processing network induced by the movement-related complexity of French stimuli. This indicates a general division of labor between temporal areas and frontal ones.

The broad picture view offered by the clustering analyses of regions and conditions confirmed the patterns reveled by the detailed contrast-based approach of our two fMRI studies. We can thus conclude that Verb- and *wh*-movements activate similar frontal regions (IFG and Precentral Complex) when found in isolation, while their combination shows an activation pattern encroaching both frontal and temporal regions. Crucially, more local movements, like clitic- and NP-movement, appeared neurally distinct and modulated activation mostly in the temporo-parietal regions, as they show in this analysis more similar activation patterns in temporal ROIs.

Hence, the fronto-temporal syntactic task-sharing emerging from our French results indicate that frontal areas are involved in long-range movement-related processes, with an interplay between a process of sequential patterns detection in the Cerebellum, and a process of working-memory retention of hierarchical tree-structure in the Precentral-opercularis Complex. On the other side, the temporal network appears to be more concerned with the kind of transformation taking place in lower Sentence Domains, namely through syntactic transformations yielded by more local movements like clitic- and NP-movement. This pattern can be taken as an empirical result in favor of the representation of the sentential's architecture into Sentence Domains, where syntactic transformations like clitic and NP-movement that target IP-internal and argumental positions elicit more temporal, while wh-movement and Verb-movement in French interrogative target higher positions in the CP-layer. This point was also confirmed by the result we obtained for

Chinese even-Focus (Argumental movement), that showed a significant increased activation of the aSTS, the only area with anterior Insula, where our ROI analyses revealed a significant effect for NP- and clitic-movement against matched baselines.

A second broad consideration about the distributed Sentence Network can be made based on similarity of activation patterns between regions and their clustering. We were able to identify three sub-circuits and a clear functional tri-partition of Broca's area where each one of its sub-parts was clustered with one of the three different sub-networks.

Importantly, it also revealed an interesting network encompassing Precentral Cortex, the SMA, Opercularis and the Cerebellum, thus confirming the co-activation of Cerebellum and Precentral Complex in different movements we had observed in ROIs and whole-brain contrasts, and in the linear effect for syntactic positions.

In this regard, a further element of convergence between the investigation of sentence hierarchy structure in Chinese and French can be highlighted. The two syntactic-tree complexity measures tested (i.e number of syntactic positions and height of a sentential element in the cartographic mapping of the sentence structural skeleton) in these two typologically distant languages pointed to Precentral Gyrus (BA6 - dorsal and ventral portion) as a locus representing and, probably, storing in working memory the syntactic-tree hierarchical complexity or at least its highest layers.

can conclude that the risky challenge of bringing more linguistics and more syntactic theory in neuro-imaging experimental hypotheses actually payed back. The highly pluri-disciplinary gaze adopted to understand the determinants of activation to syntactic complexity revealed and confirmed several possible characterizations of the role of both classically discussed, and less considered, brain areas in the Sentence network. We iconically summarize the functional characterization that can be sketched from our findings in Figure 7.39 and we hope this highly pluri-disciplinary work will pave the way for future research engaged in these directions. In order to pursue the *grail trail* and obtain a detailed view of which brain region codes for the different aspects of the sentence structure complexity into a comprehensive neuro-linguistic model.

The way human brain computes and represents syntactic sentence structures has been nonetheless locally unveiled to an extent that actually allows – thank to interpretation that was made possible by the ever-growing literature on these topics – to constitute a draft of such a model. The broad division of labor between temporal and frontal circuits, for processes implied in establishing intra-sentential dependency-links (and local and Argumental movements), versus syntactic hierarchy representation (and of the CP-layer), could constitute one of the main contribution to a model of the neural implementation of the sentence-unit.

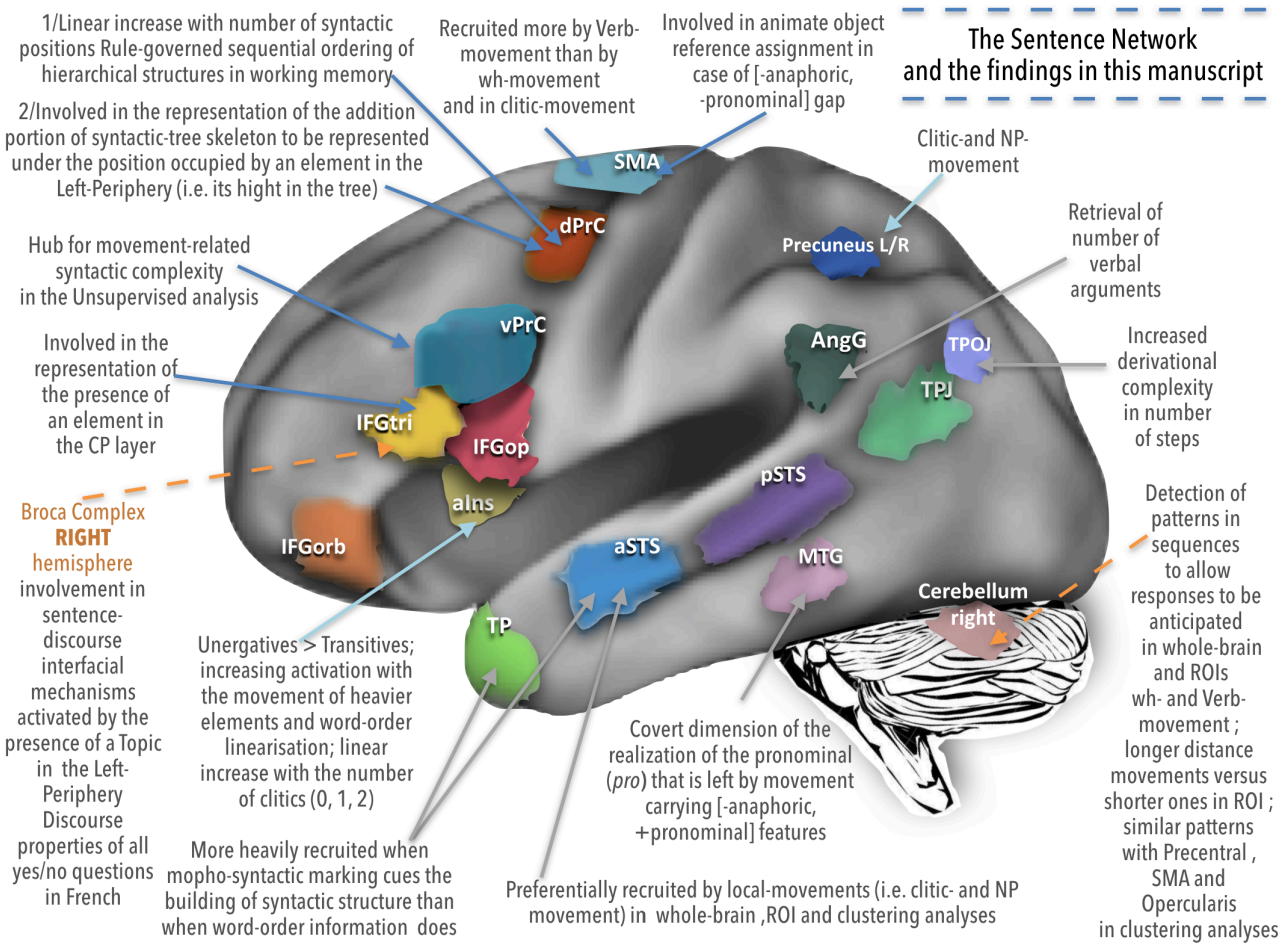


Figure 7.39 – Iconic recapitulation of the fMRI findings in this manuscript. To summarize our contribution to the larger issue of the representation of linguistic knowledge in the brain, this Figure offers an overview of the detailed results obtained by investigating the different sentence structure and complexity parameters we focused on, and thereby the functional characterization of the sentence hierarchical structures network.

Perspectives

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

Through the unknown, remembered gate
When the last of earth left to discover
Is that which was the beginning;
At the source of the longest river

The voice of the hidden waterfall
And the children in the apple-tree
Not known, because not looked for
But heard, half heard, in the stillness

Between the two waves of the sea.
Quick now, here, now, always–
A condition of complete simplicity
(Costing not less than everything)
And all shall be well and

All manner of things shall be well
When the tongues of flame are in-folded
Into the crowned knot of fire
And the fire and the rose are one.

Little Gidding V, Four Quartets. T.S. Eliot (1943)

As any scientific work and analysis, the studies presented in this manuscript are provisional and structurally open to further research as attested by the conclusive remarks on future research that were presented at the end of each experimental chapter. We will here resume some of them and delineate a few additional conceivable areas for future research.

The first linguistic perspective, would be to further develop the study of the sen-

tence's interface with discourse in the oral modality represents another important aspect to confirm in normal adult sentence comprehension the importance of the rich set of syntactic cues that is conveyed by prosodic intonational patterns. We would also like to extend our phono-acoustic analysis to other Gapless Topic constructions in Mandarin Chinese.

Another linguistic perspective of this work would be as highlighted in chapter 3, to complete the analysis of the different linguistic facts we brought to light in our cross-linguistics study of Scene-setting Topics.

Specifically, the relation between different argumental structures and argument linearization preferences in the Comment clause, in Chinese Topic-comment articulations, would need further analysis. Hence, drawing on the parallel with Neapolitan Double subject structures and French Scene-setting Topics yielding both subject-inversion in the Comment-clause, a deeper analysis may reveal all the factors determining the preference of native speakers for post-verbal subject linearization in the Comment-clause where an unaccusative verb is found. In this regard, the parallel with Belletti's proposal for Italian unaccusatives Subject inversion is a promising direction we leave here for future research.

Considering Topic-Comment articulation, another issue worth further investigation could be the formal modeling of Topic-Comment predication. As we saw in Reinhart's (1982) model, information is modeled as a set of file cards identifying an entity and listing the properties of that entity (and its relations to other entities). Thus, an interesting future development would be to investigate the neural underpinning of Topicality in short stories. This would imply testing for the fact that Topic expressions can be modeled by file cards which only contribute to the naming of an entity it collects information about, while the Comments would be interpreted as an expression adding information to the Topic file card. Correlating this two different predicative dimensions with brain activity should help understand the very general predicative mechanism that is syntactically realized in the bi-partition of the sentence into a Topic and a Comment.

From a more psycho-linguistic point of view, as an exact model of how all the processes involved in understanding Topic-Comment sentences interact in online sentence comprehension remains to be built, a perspective to address experimentally in further research would be to observe during online sentence comprehension the difference between gap and resumption strategy in the minimally differing Topic sentences we tested in Mandarin Chinese.

Concretely, (a) an ERP study nearly following the design of the fMRI study presented in Chapter 7 and (b) an fMRI study testing the replicability in the oral modality of results for French syntax presented in chapter 6, would confirm the evidence gathered about the neural underpinnings of different movement types and syntactic derivation our two fMRI studies converged on. Furthermore, it would be essential to replicate Precentral Gyrus activation as proportional to height in the syntactic-tree where different sentential elements are hosted, and to further investigate the possibility that Precentral, IFG Opercularis and Cerebellum are involved in processes linked to the complexity of hierarchical pattern found in the sentence-unit[1].

The intriguing results linked to cerebellar involvement in syntactic complexity representation and processing, beg for a deeper understanding of the role of cerebellum in

1. Another way to have a deeper understanding of this last point, would include testing different parser's quantitative outputs on naturalistic texts to search for the sub-system or circuity encoding for a dimension that would be near to the syntactic-tree "*Cartographic height*" of elements in the Left-Periphery of the sentence, that we detailed in chapter 7.

sentence-processing and of the functional involvement in sentence complexity comprehension of the SMA; without forgetting the important role Sub-cortical regions probably play in sentence word-order variation during sentence comprehension.

As we have been grounding much of our research on findings from the literature on aphasic linguistic behavior, it could be interesting to extend some of our research issues to experimental settings involving these populations. In order, for example, to test the difference between sentences with Base-generated constituents in the Complementizer Phrase and movement-derived sentences with the similar syntactic-tree configuration as we were able to do thanks to Mandarin Chinese Topic-prominence parameter (allowing to oppose external merge and A-bar Movement). This could provide additional evidence that what is complex is the representation of the highest Sentence domain in sentences, and not the access to it by a movement operation that is targeting a landing-site in CP-layer. In aphasics this would correspond to say that what is '*pruned*' is the Sentence domain , and not the access to it that is damaged. Chinese or another language with the suitable syntactic configuration and typological parameters would be an ideal candidate to test for this.

All in all, there is a large venue for future research, and this consciousness is hopefully accompanied by knowing that the goals one forecasts will be collectively accomplished through time and probably generations. Broadly speaking, the future goal of neuro-syntax is to further elucidate the syntactic processing or sub-processes that are sub-served by the different brain areas of the Sentence network.

Moreover, the imperious need of more detailed theories and models in neuro-linguistics may in the future drive our interest into building a model of syntactic knowledge implementation in the brain. To go in this direction, we would advocate for future studies to take more into account the neural underpinnings of different languages. What we observed experimentally and sketched in our conclusive Figure 7.39, was made possible because of the particular linguistic characteristics of Chinese and French.Hence, as a result of our doctoral work, we will add a cross-linguistic accent to this functional characterization enterprise, taking into account the rich typological variation that is observed across languages.

Pushing further in the direction of directly contrasting typologically distant languages in a framework of *parametric* linguistic variation among languages. An interesting path to take – in order to understand the possible modulation of the Sentence network by parametric cross-linguistic variation – would be to compare the sentence network of languages that realize in their grammar the two opposite versions of a given parameter. This may be relevant for cognitive neuro-science of language to gain insights on which brain area(s) in the Sentence network processes linguistic information that is subject to parameter setting and variation, to then also identify the areas encoding for linguistic aspects which are not subject to parametric variation, such as the universal principles that are more tightly related to the core mechanisms of the syntactic capacity of the so-called *homo phraseologicus*. Thereby we may speculate that this would lead to discovering which part of the Sentence network has to be plastic during the period of language acquisition and parameter setting, and what part is instead probably innately underpinning a certain set of universal syntactic rules.

Given this last consideration – as a last perspective of this Doctoral work –, we may share with the reader that one future orientation of our research on syntactic processes would like to be much more concerned with multilingualism in both normal and impaired

populations: speech errors of advanced bilingual and second-language learners, or the patterns of impairment found in skilled bilinguals are among the issues I would like to dig in.

Epilogue

Ph.D.: Experimenting experience

> "Descartes a été un mauvais cartésien, il a trahit lui-même ce qu'il avait découvert, c'est-à-dire qu'il faut de la méthode pour connaitre la réalité et cette méthode signifie qu'il faut suivre toujours la voie de l'**expérience**."
>
> *["Descartes became a bad Cartesian, he betrayed his own discovery: one needs method to get to know reality. And this method consists in always following the path of experience."]*
>
> Charles Péguy, 1914

The two major goals of this doctoral work where (1) to show how sentence unit is a pertinent level of linguistic representation for the brain, and (2) to give a pluri-disciplinary depiction of the cerebral underpinnings of syntactic complexity in two typologically distant languages like French and Chinese. However, after having attacked in depth its two scientific challenges, we can now spend a few words to qualify the kind of work the reader went through, defining this *piece of work as a piece of work*.

A PhD, actually, can hardly be considered as just being a written text or a book. It has the peculiarity, compared to other written texts, of being at the same time a manuscript, an intellectual human experience, a period in life, and the output of a certain research experience. Thus, if the majority of its pages is dedicated to report research results – as a regular PhD manuscript ought to be – this Epilogue is dedicated to considerations on the knowledge experience a PhD essentially is.

This is why these pages will be dedicated to our personal research experience and to how our human understanding benefited from it. They might seem superfluous, but can actually qualify as the "Philosophy pages" to obtain the grade of Philosophiae Doctor. After all, it may well be that some of these considerations will possibly be among the best inheritance we received from these years, and we should definitely bequeath them to actual or future generations.

Moreover, as these years of research activity have been featuring a close acquaintance to different disciplinary fields, they brought us to develop a certain theoretical reflection about what it means ***to know*** in nowadays scientific and intellectual context and spirit. More specifically, the pluri-disciplinary research topic we chose urged some fundamental questions: what do we know when the object of study approached is multi-disciplinary? And, how do we get to know it when we continuously have to switch point of view and analysis? Or, what it means to implement a certain experimental method in this configuration? In brief, the issue of what it means to "make pluri-disciplinary science" today continuously came back on the stage.

This part could have been placed at the beginning of this manuscript as it has a fundamental methodological flavor, but it was better placed at the end of it, as it presents part of the experience we have been harvesting in these years, as an epistemological take-home message. The epistemological seed we are planting here will hopefully grow, and allow the reader to grasp, or at least make him sympathetic with part of the *harshness* of the pluri-disciplinary spot where this PhD is rooted – a mountain ridge from which is surely possible to enjoy the panorama, but after a difficult ascension. In this ascension, our intellectual cohesion (or equilibrium) has consisted in being able to orchestrate dialogues between the two slopes of the crest on which we were making our path, in order to investigate the sentence-unit. On one side of the mountain ridge, Linguistics was offering different descriptive and analytic tools, and, on the other, Cognitive neuro-imaging was imposing its exact quantitative methods to measure brain activity.

Hoping not to offer a cacophony, these Epistemological Considerations were written as a personal contribution to express what it meant for us to do experimental science. They will elliptically sketch some of the epistemological apriori and difficulties of (1) experimental, (2) pluri-disciplinary and (3) theory-driven research.

Overview of the contents of this chapter

Delineating a Research Approach: some epistemological notes

> *"As Bergson taught me a certain way of saying the word 'reality': reality as a happening, as something that is happening, and that is asking something back from me,* ***asking back the freedom of my thinking to recognize it****, asking to give space to it, asking to tell 'yes come here, make me have a look at you', awaiting from me to acknowledge it."*
>
> Charles Péguy, 1914.

One of the important outcomes one could harvest from this doctoral work resides in the research approach it delineates and adopts. Research approach issues are traditionally classified under the label of a branch of philosophy called epistemology, which takes issues related to the knowledge of the world as investigation object. Although this conclusive section does not have the overt ambition to be called philosophical, it is nonetheless dedicated to what we learned and how our thinking evolved during the *research experience* this dissertation gives a voice to.

It has to be considered as a humble attempt to express some epistemological reflections that dwelled in our mind during these years of cross-disciplinary odyssey. It surely needs strengthening at several places, and some readers might have also enjoyed more concrete examples to illustrate some of the assertions it contains, but, in the interest of brevity and to avoid any moralistic tone, we deliberately decided not to exceed in this direction, by obliging the reader to tame his curiosity and directly read the 'raw materials' -i.e. the *texts*- that represented important turning points for our reflection on what it means to get to know something about reality, when carrying on a pluri-disciplinary research project, that is crucially orchestrating theory and experimentation to investigate the human mind and its linguistic capacity.

We somehow obviously started by asking the very basic question, that every PhD candidate or researcher asks himself at a certain point: what am I doing? Or, ultimately, what is research? The answer we gave was very concise: It is a though-at-work or a work-in-progress-thinking. In other words, we can say that it was mainly a ***thinking experience***.

These two words have the advantage of bridging together what essentially defines research: the interplay of human reason – *thinking* – and reality through *experience*. However, saying 'experience' implies a third and essential element that is *method*.

In fact, a thinking-at-work has a certain amount of *a priori* and methodological aspects that are worth describing in this section which will summarize how we personally realized their importance. Hence, among all the above-cited purposes of this chapter, it is worthwhile to explicitly point out that for the linguist reader it will put forward some central difficulties linked to empirical research and recall the reader in cognitive science that some perspective taking during experimental work could actually be salutary. By analyzing some of the difficulties or post-modern trends that are nowadays to be found in pluri-disciplinary work, we hope to make the reader sympathetic with the scientific and, from time to time, social hardship of trans-disciplinary research.

In sum, you will find here, some words well spent about the backstage of the scientific content of this manuscript, and about this PhD research as a knowledge experience. The broad aim of the epistemological patchwork reported here, is to give the reader a flavor of what could be called having a bunch of thoughts at work in your head – in French *une pensée au travail* –, and how this set of thoughts could end up in enabling to know something about reality.

Hence, in this conclusive chapter, I take the responsibility to share with the reader the fact that I have been dragged by my research experience into a previously unknown field - that of epistemological reflection. It has to be said that this meta-reflexive process[2] has been even more necessary and urgent given the pluri-disciplinary approach that was adopted.

Three fundamental problems had to be addressed for progress to be made in this field of inquiry, and the following three sections respectively treat them one by one. The first is dedicated to the kind of relationship to reality experimentation imposes (section *Eperimentation and Reality*), the second one is devoted to pluri-disciplinarity (section *Pluridisciplinarity or Fragmentarity*), and the third (section *Theory and human Knowledge: long story short*) is more simply dedicated to some aspects we discovered to be true during theses years of research experience.

To give some acknowledgments to what made this kind of reflection possible, we can say that this last part reflects the kind of work we were able to develop in daily silent internal talks, in long and articulated talks with friends[3], and in occasional synthetic and incisive car talks[4]. However, our own sketchy epistemological considerations transcended understanding into aesthetics, when our thoughts occasionally met with the texts of some philosophers we will share with the reader by simply transcribing and briefly commenting them.

In conclusion, this chapter will have a relatively wide scope spanning from (i) tentatively explain, or at least give a suggestive depiction, of the difficulties of empirical research to readers coming from linguistic fields, (ii) argument for the centrality of theory in my approach, and last but not least (ii) show how my thinking evolved.

2. The reader may find it probably a bit too *'meta-'* (too much looking at oneself while thinking) and I apologize before hand. Skip it as soon as you are fed up with it. I could not get rid of these epistemological interrogations before I ended up writing them clearly on paper, and would not have thought of having such a chapter in my PhD before I wrote it.

3. Often following the pipeline: Baptist talks followed by Christian's debriefing, validated by Tristan's "Apéros" and Servane's lunch chat, and crowned by Maria Silvia's quadrennial check points.

4. A Citroen Picasso that I thank for the recurrent hospitality given to my tired body on the left passenger seat.

Experimentation and Reality

The Primacy of Hosting Reality

> "La réalité, en chacun de ses points, est comme une ville bloquée. L'armée royale est partie au secours. Mais l'armée royale ne peut parvenir elle-même et il faut qu'une sortie de la place même vienne au devant d'elle et lui donne la main.
> En ce point intermédiaire entre l'homme et le monde, en ce point intermédiaire entre l'esprit et la réalité, en ce point intermédiaire où s'établit la liaison entre l'armée de secours et littéralement le secours propre de la place, en ce point s'opère pour Descartes la connaissance de la vérité."
>
> *["Reality at each of its points is like a besieged city. The Royal Army came to rescue it. However, the Royal Army cannot succeed by its own means. There must be someone inside the citadel making an outing, someone who would come in front of the Royal Army to give it a hand. In this intermediary point between man and the world, in this intermediary point between the mind and reality, in this intermediary point the link between the rescuing army, and literally the rescuing of reality, is established. For Descartes, it is from this point that proceeds the knowledge of truth."]*
>
> Charles Péguy, 1914.

As a first consideration on the kind of experience a PhD is, and on what it means to have a knowledge experience, we would say that it surprisingly resembles to the configuration Charles Péguy is exposing in his comment to bergsonian philosophy in the above epigraph (we suggest the reader to go through it twice, as it is expressed in a convoluted way)[5]. Paraphrasing it, the reality we want to know through research, is in his metaphor a *besieged citadel*[6], and the army approaching to deliver it, is our reason with all its tools.

5. Excerpt from "Notes sur M. Bergson et la philosophie bergsonienne", La Pléiade, p. 58.

6. This 20th century author has a long reflection on Cartesianism and reality in modern times in two not oft-quoted essays he wrote the year of his death: "Note conjointe sur M. Descartes et la philosophie cartésienne" and "Note sur M. Bergson et la philosophie bergsonienne".

However, from this simple configuration, what this metaphor actually reveals is that to 'rescue' the *citadel of reality*, the army of reason and knowledge should, first, and somehow counter-intuitively, count more on the unexpected outing of "*someone who would come in front of the Royal Army to give it a hand*", on an intermediate point between the vase reality and his the mind, rather than on its own strength or on the sharpening of its rational tools. Using the two philosophers' words, this is a particular attitude where "***facts became the inauguration of the work of thoughts***" (Péguy on Bergson, 1914)[7]. As illustrated by Figure 7.40, the exit of a *fact* from inside the

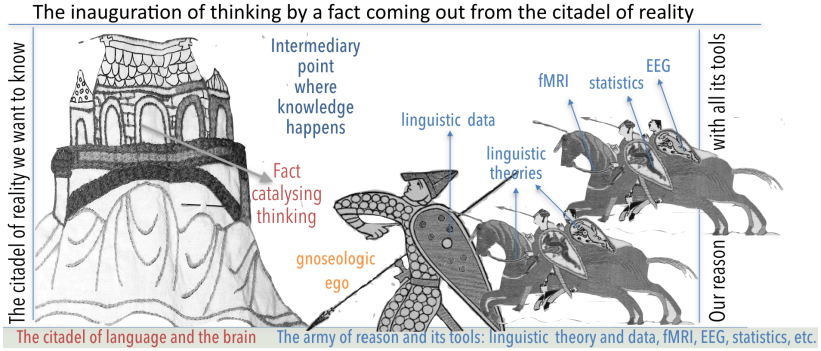


Figure 7.40 – Iconic representation of Descartes point where the knowledge of truth proceeds. The intermediary point between man and the world, between the mind and reality..

citadel of reality inaugurates the work of thoughts and constitutes the very first step of knowledge, and the fundamental priority that certain facts have in the knowledge experience of a researcher, can easilly attest for this. Reaching "this intermediary point between man and the world", "this intermediary point between the mind and reality", can be thus understood as one of the main concerns of researchers or individuals engaged in knowledge activity. One may attack the citadel to rescue it, with all rational means, but if a single real fact from inside reality is not making an outing, the "link between the rescuing army" of reason and the world is not established.

The accent we put on this prior step to *active* research activity is due to what we happen to understand during thes last years of experimental neuro-linguistic approach. As long as one do not acknowledge the facts that we have in front our eyes, and if we doubt of their reality too much, it is then impossible to engage in a factual knowledge experience, because nothing, in the end, seems to be real. Concretely, if we think the experimental data in front our eyes are not real – e.g. they are imperfect recordings of what the brain really does, they may be just linked to the variability of our data, they may be too 'noisy' or they may just sometimes do not mean anything to us –, one remains epistemologically blocked, and it is thus impossible to gain any knowledge from these data. It is as if it became impossible to actually *see what we see*, and the inauguration of thinking that proceeds from the facts coming out of the citadel of reality results in being structurally blocked.

In this regard, we can read and probably better understand the previous epigraph, saying:

7. In French: *"Les faits sont comme une inauguration du travail de la pensée".*

> "As Bergson taught me a certain way of saying the word 'reality': reality as a happening, as something that is happening, and that is asking something back from me, asking back the freedom of my thinking to recognize it, asking to give space to it, asking to tell 'yes come here, make me have a look at you', awaiting from me to acknowledge it."
>
> Charles Péguy, 1914

Although, reaching the intermediate point between reality and reason, that Descartes and Péguy point to, happens in a fairly passive manner, this requires an essential step of accepting facts, and we could say *hosting* them, which is impossible if systematic doubt is actually cast on the reality of facts and observations. We learned personally, that only by first taking an extremely realistic position of *seeing what we see*, we were then able to interpret experimental results and data, or just take seriously literature findings. It is this step that brought us to end-up discovering that our own experimental results actually replicated largely previous studies, and were not so 'un-understandable' as initially thought.

Thereby, we discovered that to be fundamentally put in the constitutive passive position of first *seeing what we see*, was the only way to engage in our research work seriously. This is what we would call the crucial step of '*deliberately hosting reality*', we want to put forward here. Contrary to what one could think, knowledge experience is not started-up by our research activity, nor by active observation of the world, and not even by the sharpening of methodological tools, and even less by structural doubt, but by a fact that stands out from the citadel of reality, and bring us to engage with it. As it is usually said for photographers – "It is not how photographers look at the world that is important. It is their intimate relationship with it." And the relationship described in the above metaphor could be summarized in fundamental tendency to be open towards facts – an awkwardly passive position that precedes any intellectual activity.

We can conclude from this experience, that before any analytic action or doubt, the real epistemological effort is to first recognize and acknowledge facts and to see what we see. Crucially, giving the priority to facts, prevents us from considering reality as something that is merely outside, or something to which we have to 'add' our thinking to, and crucially it prevents us to doubt about reality and think the facts we observe could be something that is ultimately unreal, which would inevitably block observation and therefore knowledge discovery[8], and research activity.

After this factual inauguration of the work of thoughts, and only after, comes the reunion between human reason, theory and experimentation, in what is classically called the sacred cycle of empirical sciences: "observation-theory-experimentation".

A second example of this can be found in the first fact that inaugurated my thinking, which happened quite a while a ago, when I started studying Chinese. In order to utter a sentence (that could *sound* Chinese) I had to commit to a small Copernican revolution in my head and to structure it, according to an extremely productive and frequent sentence structure in Mandarin Chinese: a Topic-Comment articulation.

This change in perspective on the sentence-unit catalyzed my attention in such an unforeseen way that it started my work of thoughts, so that later I continued in studying Chinese Linguistics. Then, driven by the intuition that such a sentence articulation should correspond to a particular mental construal (or representation)[9], I approached Cognitive Sciences, where I met the brain – this wonderfully complex biological object –

8. By this I mean *gnoseological activity.*
9. We went into the details of this initial intuition in Part I (page 77).

that inaugurated a second side of my thinking. In this way, and particularly in front of the complexity of the brain, I happened to understand that one needs, in first instance, to keep a certain attitude of openness, letting reality have the first contact – for instance through astonishment – letting reality first meet us up in front of the deployment of methodological tools and disciplinary knowledge that the 'troops' of scientific reason usually orchestrate[10].

We can, then, say that we have been *hosting* two elements in these years. The first are the experimental results obtained in brain-imaging – keeping them in my mind by just acknowledging them, which crucially brought us to simply observe their convergence with the literature and understand them. Hence, what we shared is surely not the final word on these findings, and we hope they will be housed by other minds and further understood. The second fact and research object we have been hosting for years, is the issue of structuring a sentence according to the Topic-Comment articulation. This fact was omnipresent to our mind, as if we were besieged. When listening to the radio, taking a shower or listening to linguistic interactions in the bus, the question of brain activation linked to French clitic or the pervasiveness of Topic-Comment articulations in everyday life was recurrent. Moreover, as we had the occasion to note in chapter 3 (§*Topic is in the air*, p.231), Topic-comment utterances where highly on fashion in the last five years in France, because the former president could hardly express himself without this syntactic structure!

The 'work' of Experimentation

> "[L'homme] au lieu d'observer les phénomènes naturels tels qu'il lui sont naturellement donnés, place la nature dans les conditions de son entendement."
>
> *["Man, instead of observing the natural phenomena as they are naturally given, places nature in the conditions of his understanding."]*
>
> Hannah Arendt, *Condition de l'homme moderne, 1983, p.299.*

One could argue that the mechanism described in the above sections could more simply be described as the everlasting play between human reason and reality. However, the need for a receptive relationship with facts to inaugurate thinking, and the previously described passivity, still constitute two discriminant aspects that can revert – at least partly – the classical paradigm of unveiling reality by an active initiative that is characterizing modernity and positive thinking. In such a framework, the role of experimentation could seem fundamentally contradictory. The next step will be then to qualify in detail what it means to perform experimentation, while preserving the two steps presented above.

Experimentation, in the broad sense, could be understood as being imposed by a natural inquisitive inclination of human mind, finding an acute expression in PhD candi-

10. Something one could express in French as *"se laisser rejoindre par la realité."*

dates and researchers[11]. Its pervasiveness in history shows that it did not wait the Age of Enlightenment or Modernity to appear. As cited above by Hannah Arendt, it expresses the indispensable step of putting reality in the condition of human understanding.

However, this inquisitive attitude and its modern and unavoidable act of experimental conditioning and formatting of facts, would actually set a snare to our understanding, if the experimental elaboration would not be preceded by the three attitudes/steps previously explained. Hence, meeting-up with facts outside the "citadel", letting them inaugurate thinking, and the passivity of hosting reality, need indeed to be understood as happening sequentially in the knowledge process.

To this basic definition of experimentation, the modern experimental dialogue that science entertains with reality adds the complexity of the instruments of measure and statistics. It consists in manipulating reality to shape it in a way that it would become measurable, calculable and reproducible. By isolating experimental parameters in a sometimes extreme fashion, this scientific dialogue with reality manipulates facts to put them not only in the condition of human understanding, but also in the conditions that are necessary to fit instruments of experimental measure and statistical methods. This extra step is unavoidably a reduction of the observed phenomena to quantification.

In my case, this kind of scientific dialogue between facts, methods, measures and stimuli initially found a silent expression on the stage of my mind. There, the very classical distinction between qualitative and quantitative considerations came to surface: the disciplinary location of this project urged me to clear my mind theoretically and epistemologically on what I was doing. I literally had to 'try linguistic descriptions – by essence qualitative – on brain activation data – by essence quantitative.' This point is indeed essential to understand one of the choices I made by shifting towards theoretical linguistic models that could more easily offer some quantitative predictions about syntactic complexity. In fact, when phenomena are not quantitative in nature like sentences, scientific experimentation requires a theoretical apparatus that would allow to consider a complex sentence to be quantitatively more complex than another. For instance, having linguistic complexity matrices that could be correlated with measurable quantitative brain activation data has been essential. Although, one should say, to temperate these arguments, that Neuro-imaging with its methodology of contrasting brain activity between two experimental conditions still allows to test for some kind of qualitative differences. This, nonetheless, always presupposing that the processes at stake in the two contrasted sentences take place in different parts of the brain, which in the end is often difficult to predict.We went into further details about what kind of assumptions brain-imaging techniques entail in experimental reasoning in chapter 1, §1.5.

Furthermore, operationally, not taking for granted that we know reality and that we have observed it enough could help preventing arbitrary manipulations[12]. I personally observed that the danger of doing experimental science while leaving reality outside the door is always latently present, especially if one considers the unreal cosmos that experimental manipulations necessarily cast on linguistic phenomena: controlling lexical frequency of each word, number of letters, pragmatic plausibility of sentences outside

11. This shows to what extent an intellectual adventure like a PhD is not a sedentary form of reverie of a man that would fear of the world and flee from it, giving free rein to thoughts, nor a lazy version of the cynical misanthropic attitude.

12. These would have the draw back of limiting the validity of experimental results to and wouldn't inform on the real object of research.

discourse context, or selecting an extremely reduced experimental corpus, without forgetting noise in the scanner. These elements, just to quote a few, are experimentally controlled measures that heavily weight on the naturalness of linguistic experimental stimuli. Hence, building experimental designs can often be such a perilous activity: *orthogonalizing* an experimental design in order to be able to observe the single contribution of a linguistic element, compared to another, can quickly lead to selecting absurd linguistic materials. It should be said this isn't much more than old wine in new bottles: this issue is known under the name of "ecological validity", and it is the crux of experimental psychology (Cutler, 1981).

In sum, a long-standing attitude of hosting reality further on, and at every steps of the experimental procedure, seems urgent in order not to loose reality, for instance in the intricacies of isolating experimental parameters. Moreover, what could seem at first sight a somehow obvious step, appears even more essential in order not to limit our knowledge to the realm of the experimental manipulation, but to cast a light on reality itself. More concretely, it can happen that a badly designed experiment would impede experimental data and results from being the basis for understanding progress beyond the immediate linguistic question addressed in the given study. .

To avoid this, I particularly worked out the naturalness of the experimental stimuli. As a matter of facts, it took me months to validate the sentences constituting the materials of the experiments presented in Part II (see Annexes with Experimental Corpora C) with five different informers each time, both being in Beijing and in Paris[13]. The main problem imposed by experimental setting was residing in the semantic plausibility of sentences outside an enunciation context (for more details on the Informers see the Notice to the reader , p.xxiii. This is actually the main reason why I chose to run a study on Topic-Comment sentences in context, reported in chapters 4 and 5.

In order to prevent the risk of experimental results not constituting a generalization of real linguistic world and language faculty (cf. 'ecological validity'), a recent emerged experimental tendency selects texts instead of sentence as experimental materials. If the very first attempt to use texts is to be found in the PhD Thesis of A. Bachrach (2008) – who actually deeply worked out the structure and many lexical aspects of his texts – the actual trend is to increasingly use naturalistic texts and stories to test for different linguistic parameters (see Brennan et al. 2012, Wehbe et al. 2014, Hale et al. 2015). It should be said that while this represents an initial answer to the above described difficulties, it opens up other methodological issues, but this is beyond the scope of this section.

Given these considerations, the preliminary step – hosting reality acknowledging and recognizing it – before experimentation is even more urgent when we are conscious of what it means to submit it to the heavy load of experimental methodology. In fact, always having in the back of my mind real linguistic utterances overheard in the news or in the bus really helped me at every step of the experimental process[14]. One could therefore conclude that the very initial step of knowledge described here is more than initial, it spans over the whole process of experimental Science from observation to theory and experimentation. We will conclude this chapter by saying that this is an attitude towards reality, but let us first consider how the pluri-disciplinarity of this research project transforms the knowledge configuration and setting described until here.

13. I should thank here my Doctoral School and the ANR project of my Supervisor (Space and Languages) that allowed me to go to China with three scholarships.

14. In fact, everybody speaks with Topic-Comment articulations.

Pluridisciplinarity or Fragmentarity?

Pluridisciplinarity, or the best way to be a philosopher

> "Diese Griechen waren oberflächlich – aus Tiefe!"
>
> *["Those Greeks were superficial - out of profundity!"]*
>
> FRIEDRICH NIETZSCHE, *Die fröhliche Wissenschaft*, CHAP. 2 §4

One of the main interests of the research work presented here as my doctoral dissertation, lies in the fundamental claim of requiring a multidisciplinary approach to address the question of sentence, of its structures, and of its neural underpinnings in Chinese and French.

Pluri-disciplinarity is necessarily adding a supplementary complexity to the three steps detailed in the previous section - (1) passive inauguration of thinking by facts, (2) receptive hosting of reality and (3) successive experimentation. Going back to the now familiar *citadel* metaphor, one would add the image of intervention of multiple armies arriving from different disciplinary kingdoms. And, we might wonder how multiplying the rescuing armies would impact the very first step: would reality need to lend different hands to help the armies rescue the citadel, in order to inaugurate thinking?

This multidisciplinary approach inherently posits some philosophical questions to which this epistemological sub-section is dedicated.

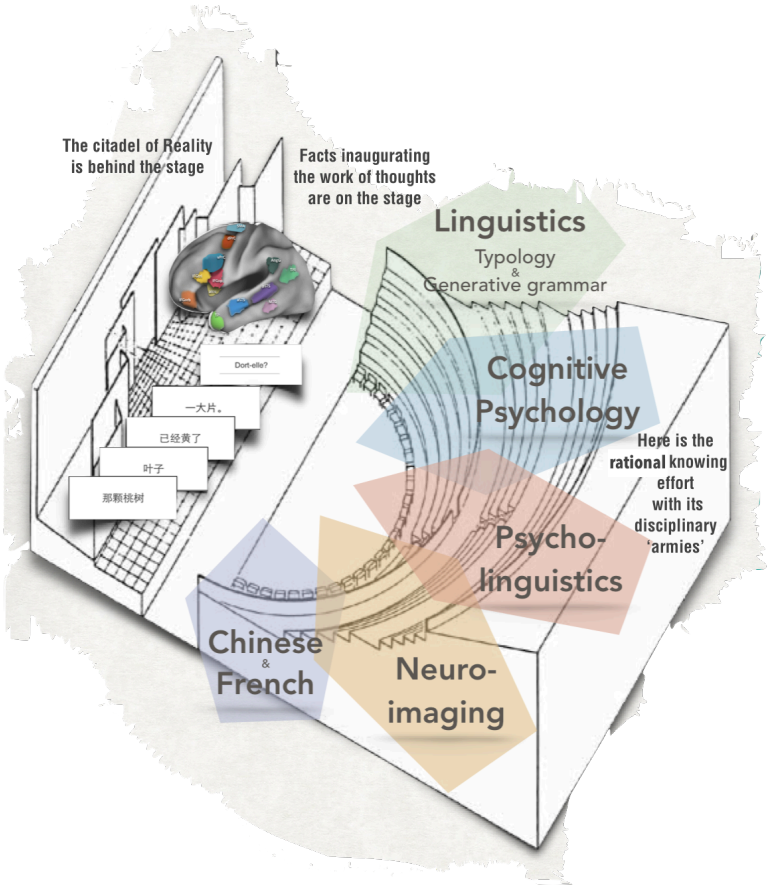
There is no doubt that by nature a theme, a fact or a problem transcends the classical disciplinary boundaries; it exists as an object per se[15]. While facts do not belong to a given discipline, they can be viewed with a different focus and this focus is necessarily disciplinary both from an epistemological point of view – given the different methodologies and questions asked to a given fact, – and from a sociological point of view[16]. This disciplinary question has been further addressed in the next section *Theory and humand knowledge: long story short* (p. 712), which presents how some epistemological obstacles, comprising disciplinary ones, can be nonetheless useful for research.

Coming back to what happens in our pluri-disciplinary configuration, we can first note that the inauguration of thinking comes from different elements all getting out of the 'citadel' by opening its door. These multiple elements allow *different* and *interconnected* questions to be raised: why are basic sentence structures different in Mandarin Chinese and French? What does it mean for the mind and brain to structure sentences with different hierarchical syntactic constructions? How can the language network in the brain achieve building up dependencies among words inside a sentence in absence of overt marking? Here are the initial inaugural facts - expressed into questions - that entered the theater stage (cf. Figure) of this PhD project to constitute its multiple objects of investigation.To tell the truth, I have been recurrently doubting about the feasibility of such a pluri-disciplinary enterprise, so that, as a starting point of this intellectual

15. One could indeed say that some questions stand by themselves and are not discipline-specific, take for example the question about how do children learn language.

16. in fact, peoples and intellectuals organize into groups sharing a terminology and other signs showing their belonging to this group that has the advantage of setting them out from the crowd.

itinerary, I would like to invite the reader to add in his field of vision what is called in Kantian terms *criticism*: an activity where human consciousness puts itself into question in order to know and evaluate what is at reach for its knowing capacity and effort.



The fact of being permanently balancing between two outlooks on sentences as experimental stimuli - the one of the linguist and that of the experimental cognitive psychologist - forced me to develop the reflection, tinted with philosophy, that I share here with the reader. The fact of oscillating between situations of linguistics precision accompanied by experimental inadequacy, and situations of linguistic under-specification accompanied by experimental adequacy, made me fear at some point that the *diplomacy* between disciplines would result in an impossible armistice, dooming my research project in an unsatisfactory superficiality.

It is at that moment that I came across, in my personal readings, one of Nietzsche's exciting and controversial exclamations in The Joyful Science: *"Those Greeks [philosophers] were superficial - out of profundity!"*. I suddenly understood that pluri-disciplinarity was one of the best ways to be philosopher, being deep in superficiality: (1) choosing the right level of analysis of the linguistic facts under study, (2) understanding the right balance or the right priority to give to the different disciplines was an intellectual matter, that was coinciding with the step of (a) thoroughly grasping the interrogation at stake, and (b) strongly fixating the theoretical background framing my research inquiry. All this, always putting oneself into question in order to know and evaluate what is at reach for one's knowing capacity and effort.

Science without philosophy?

To the reader that still would not be convinced that this part of the manuscript is to be viewed as essential to obtain a PhD degree, I suggest to skip it. However, the speedy 'line rider' should nevertheless acknowledge, as previously said, that it constitutes the Philosophy part of the name PhD, and that it is consequently opportune and not off-topic. Given the philosophical reflections I had drifted to by pluri-disciplinarity, a naive question naturally raised from my tentative to chorally link together different disciplines: Is it possible to do pluri-disciplinary science without philosophy? Jurgen Habermas – one of the most eminent philosophers of the second generation of Frankfurt school – is one of the numerous thinkers that have been warning about this danger. He

was reminding the contemporary scientific field, and especially the sciences that take man for an investigation object, that to do good science one needs a good philosophical gaze[17]. We can read some sharp criticisms in one of his central works: "The scientist faith [*szientistische Glaube*] in a science that one day will be able not only to complete man's personal consciousness [auto-awareness] by an objectivating description, but also at the same time will dissolve man in it, is not science but bad philosophy." (Habermas 1968, in *Erkenntnis und Interesse*).

Pluridisciplinarity or against fragmentarity

In the history of thoughts, it seems that we got to the point [18] where the progress of different branches of science, that was obtained through the specificity and the differentiation of their terminology and methodologies, dooms to failure every tentative discourse on totality, as if the only possible position would be the extremes of rupture, or nostalgia of a *total* gaze on reality.

Nonetheless, when I started understanding that the fragmentarity distinguishing the gaze we – post-modern men and women – cast on the world is far from being just a disciplinary or academic problem, but is a general intellectual posture (as exemplified by the philosophical Divorce Court example in the next quote by Chesterton), I decided once more to resist and continue, despite the harshness of this pluri-disciplinary adventure, which was sometimes very discouraging. I chose to finish this work also to try to curb this tendency for fragmentarity. Furthermore, part of my determination was given by the fact that one of my objects of study – the brain – was intrinsically the biological place where complexity and totality are not only naturally taken into account, but also processed. In other words, the brain is a complex biological system than can process complexity, it seemed therefore totally reasonable to study it as such, and to cast a total and pluri-disciplinarity gaze on it.

Thus, before continuing in the description of the challenges of pluri-disciplinarity, the reader will be offered a enjoyable pause thanks to an example of the nicest British prose from the beginning of the 20th century. This excerpt is one of the best illustrations I could find in literature of what I define here as fragmentarity and the intellectual posture that distinguishes it. The text could seem at first sight a real indictment, but ends up in being an ode to experimentation giving priority to reality, winning over the intellectualism of casting *"an intelligent look at things"*:

> As I have said before I am a believer in staring blankly at things; if you do it something always happens. For instance, I am staring blankly at this sheet of paper and I firmly believe that something more or less intelligible will happen soon. Men stared at the blank blue sky and invented a million mythologies. Staring stupidly at live people is more dangerous; but even this has its fascination; and if you ever see your companion's face turned towards you with the rounded and complete expression of a congenital idiot, you may be certain again that he is nearer at that moment than at any other to knowing what you really are.

17. I would add, this is especially true in order not to get lost in technicality, trying all the methods and data analysis.

18. According to philosophers and historians this point has already been reached after Hegel (see J. Hersch in *L'étonnement philosophique*, 1993), but I do not have any real insight on this point to allow myself to affirm more than my personal experience of the scientific world nowadays.

When we cast 'an intelligent look' – as they say in books – at a thing, it only means that we stamp our own significance upon it. When we look wisely at a post we see what we mean by a post. But when we look stupidly at a post we see what a post means.

In such a trance of divine imbecility I remember once staring at the paving-stones under my feet, until I went off into a sort of dream of paving-stones. They passed perpetually under my feet like flat and silent waves of stones, and **all the time I was asking myself what they were**. Street after street I passed, looking at the ground like a cow.

And then it suddenly seemed to me that they were all gravestones; the gravestones of innumerable and utterly forgotten men. For under every one of them, almost certainly, there was human dust. I seemed to see fantastic epitaphs on them, commemorating the deeds of heroes who are too old and too great to be remembered. There, for instance, was the man who found fire and the man who made the first wheel; men too necessary to be ever named.

And among those imaginary benefactors in all ages I seemed to see one class especially predominant. I mean the people who in the dim beginning of time **united one thing artificially, but permanently, with another.**

What primeval priest, for instance, married bread and cheese? Who was the wild visionary (of later times) who, after ransacking all the forests, and counting all the fruits of the earth, discovered that almonds and raisins had been looking for each other since the world began? Who, above all, discovered such a thing as the happy marriage between music and literature? The men who are least known from the past are certainly the men who made this combination. And the men who are best known at the present day are certainly those who are tearing such combinations in pieces.

This is the worst element in our anarchic world of today. The whole is one vast system of separation - *an enormous philosophical Divorce Court*. [...]

The theory of art for art's sake, for instance, as applied to painting, was a proposal to separate a picture from the subject of the picture. Sentiment would be better without art, art would be better without sentiment. In other words, a picture would be a better picture if it were not a picture of anything. And a subject would be all the better subject if you did not paint it. Such moderns easily might, I think some moderns really have, applied the same principle to that ancient combination called a song. A very modern poet might easily say that the words would convey their own natural rhythms much better without a tune. A very modern musician might easily say that the only perfectly musical songs would be songs without words.

No one has yet had the star-defying audacity to hint at a separation between bread and cheese. But we must be prepared to have it said before long by some profligate aesthete that bread would be more breadish without cheese, and that cheese would be more exquisitely and penetratingly cheesy without bread. We must be prepared, I say, for a perpetual tendency towards such cleavages; and we must be prepared to answer them by insisting on the immemorial right of mankind to perpetuate such alliances. Man has from the beginning joined spoken words to an air, and the two have grown old and wise

> together. Those whom man hath joined let no man sunder.
> G. K. Chesterton, On Fragments, in *The spice of life*, 1906.

Although, the *"intelligent look"* of *"art for art sake"* has been echoing predominantly in the auricular pavilions of artist and literary people, it nonetheless ended up reaching academic ears in a slightly deformed fashion as usually happens with echoing sounds. It transformed into "discipline for discipline sake". For this reason, someone could ask why not writing two parts: a linguistic one and a neuro-imaging one. This would have been, indeed, much easier and safer, but not so pluri-disciplinary. In other words, it could have embodied a usual post-modern utilitarian schizophrenia I try to curb: dividing clearly linguistics as being involved in the formulation of experimental hypothesis and in the designing of stimuli, relegating technical neuro-imaging aspects to mere technical issues, and switching off the linguistic gaze on sentences while reading neuro-imaging studies, and so forth. It would have been an easy, but not realistic research approach and process.

At each stage on this research project, the two disciplines (or more) were always intimately exchanging even during data analysis. The linguist in me was never sleeping, and this permanent dialogue made this research project a *polyphonic* one. If I succeed in conveying the process that generated this research work, then the reader will hopefully not listen to the linguistic partition, and then to the neuro-imaging one separately. This should tentatively reflect the spirit of mutual interdependence of disciplines that was the one at the beginning of University - the *Universitas scientiarum* - and, as we will seen, it reinforces the pluri-disciplinary project of the early years of Cognitive Science, that seems evermore and gradually dawning because of post-modern disciplinary fragmentarity.

The intrinsic pluri-disciplinarity of cognitive science

Inter-disciplinarity is not only echoing the intellectual spirit of the beginning of University, but is also deeply linked to the nature of the second research object being studied here: the brain. If there is a place in nature where everything is naturally done outside of disciplinary fragmentarity, that place is the brain[19]. Choosing the brain, and, neuro-imaging as a method of investigation inevitably imposes to consider language and all the linguistic levels that distinguish it as part of human faculties, as a competence. This constitutes indeed an important focus shift that has been addressed in chapter 1 (see §1.3).

Ever since the beginning of the Cognitive Science project, Linguistics has been considered as a reasonable source of hypotheses about the representation and processing of language in the mind and brain. And, in the emerging research field of Cognitive Neuroscience the study of language was conceived as an integrated interdisciplinary project, but the 70's and 80's were the theater of a general loss of currency for such a total-integrative view, and the cognitive project rapidly underwent disciplinary fragmentation: Linguistics became the study of the speakers' knowledge of language, and more specifically of his grammatical competence for generativists ; Psycho-linguistics specialized as a branch of psychology, dedicated to the study of the cognitive processes underlying verbal behavior ; and Computational Linguists oriented towards the development of engineering solutions for automatic processing of text.

This stage of fragmentation is at the origin of a certain skepticism about the positive issue of the combined investigation of brain and language to concretely generate progress

19. Here I do not consider modularity to be linked to disciplines, see sections §1.3.4 and §1.4.4.

in at least one of the two disciplines. In this context the intrinsic problem of granularity in Neuro-linguistics could just grow wider: the distinctions made in the neurological study of language are coarse in comparison to with the distinctions made by linguistics. Historically in fact, while on the linguistic side, the work within the framework of theoretical linguistics provided deep insights into some possible architecture of grammar, involving a number of fine-grained distinctions[20], on the other side, neuro-scientific approaches to language continued to operate in terms of broader conceptual distinctions. Thus, the problem of the *'linking hypothesis*' to map linguistic distinctions onto brain functions grew wider.

Nonetheless, on a brighter note, according to Marantz (2005), the first decade of the 21st century could have been a golden era for the cognitive neuro-science of language[21]. In that the advent of in-vivo non-invasive neuro-imaging techniques (fMRI, ERP, MEG, PET) has provided novel empirical data, that are now starting to give stable reproducible results in linguistic domains ranging from phonetics to discourse processing (cf. chapter 2 for a review of the current literature, p. 81). As already pointed out by Poeppel and Embick (2005) among others, the central challenge resides in formulating explicit interdisciplinary linking hypotheses that can help brain function data to bear on theoretical debates, and conversely, show how theoretical linguistic constructs and insights can serve to deepen our understanding of brain function, even though the units of linguistic computation and the units of neurological computation are still in-commensurable (cf. §1.5).

In this context, my research approach can be viewed as building on Cognitive Science's interdisciplinary project, in the attempt to realize the full potential of an integrative view. Making use of linguistic theory, combined with psycho-linguistic measures and neuro-imaging tools to investigate empirically the multiple dimensions of a sentence processing task and of its cerebral underpinnings.

Gradually shifting from Competence to Cognition, we get to the third development of the Cognitive Science's project, that of Computation[22]. A increasing number of researchers build research projects where the critical link between disciplines comes from computation, and ideally, from computational models that should map the appropriate level of analysis and abstraction to allow an interface between linguistics and neurobiology. We went into further detail about these models, and the characteristics they should gather to have a neuro-physiological grounding, in chapter 1. However, at the level of neural computation, it should be said that the field is still lacking appropriate animal models, so that we still know very little about the neural computation underlying symbolic and linguistic cognition. Nonetheless, language is arguably the highest cognitive domain for which detailed computational models are now available – thanks to grammar development in computational linguistics and in Generative Grammar. This fact allows researchers in the domain of language to investigate quantitatively complex internal representations and processes to an extent that is not currently possible in many other higher order cognitive domains. A number of proposals – reaffirming the connec-

20. A discussion on its reasons and its effect on the study of language in the last five decades goes beyond the scope of this general introduction (cf. Phillips, 1996; Marantz, 2005 for an argumented discussion).

21. Marantz, (2005). Generative linguistics within the cognitive neuro-science of language. in *The Linguistic Review* (22:429–445).

22. An articulate discussion about these three pillars of Cognitive Science, that are Competence, Cognition and Computation has been in chapter 2 in §1.3.

tion between the two disciplines – have been made both in computational models of parsing inspired on human behavior and cognition and for psycho-linguistically inspired computational models (Roark, 2001)[23]. Just to quote a few of them, we can consider: (i) Vosse and Kempen (2000), who subsequently inspired Peter Hagoort's Unification process; (ii) Hale (2001), who proposed the computational notion of Surprisal considering cognitive load in terms of the total probability of structural options that have been dis-confirmed at some point in a sentence, and a measure of Entropy Reduction in 2006, and finally (iii) Lewis and Vasishth (2005), who have applied principles linked to working memory and cognitive skills to sentence parsing. This neuro-computational new wave, however, is hiding or may be answering to the crucial lake of *theory* that neuro-imaging has been experiencing for sometimes.

Theory and human Knowledge: long story short

Cardinality of theory: an ode to complete thinking in a fragmentary skyline

After having presented the fragmentary vision that disciplines may cast on knowledge, and the resulting hardship in communicating across disciplinary fields - even when they are concerned by the same research object -, we can turn to more methodological considerations lying at the very core of this research work in that they are likely to resolve fragmentarity.

The first of these considerations concerns the fundamental role of theory in the development of our experimental research project. Here, in fact, it is something more than pluri-disciplinarity that holds everything together, it is *theory*: a fundamental methodological element of Science which happens to be frequently removed in today's post-modern world. Referring back to the previous metaphor of the philosophical Divorce Court, the same should be said for science and theory (or should we dare saying for knowledge and theory). The centrality of theory exemplified by this work, featuring theoretically-driven experimental hypothesis on linguistic structures of the sentence, should be seen as an ode to **complete thinking** in a world of fragmentation.

Through this consideration we reach the point in this reflection where, in a very elliptical fashion, we touch one of the most fascinating epistemological issues of the XVIIth century: how to conjugate theory and experimentation. We will just indulge in the following succinct and superficial remark on such a cardinal philosophical topic. If one would commit the fallacy of entering the neuro-linguistics' kingdom with the secret aim of paying heed to the Popperian obsession that science is all about falsifying existing theories, he would be largely disappointed. Why? Because, nowadays, any complete explanatory theory identifying the nature of linguistic computation or representation in the brain hasn't really emerged yet (cf. Granularity mapping problem, §1.5). One hardly has model sketches of the implementation of syntax in brain. Moreover, the only models that are on the stage are so general that one could difficultly falsify them by experimentation.I should temper this assertion by acknowledging that in the last fifteen years of neuro-imaging somethings have been learned even without having adequate theories, by simply falsifying naive models (see section D.5, p.914). Thus, given the actual state of research on language models, the only ready-made theories, to be falsified, are probably linguistics' ones, and for the moment they seem to resist the impact with

23. The reader will find an example thoroughly illustrated in §2.3.3.3.

neural implementation especially if we consider neuro-psychological pieces of evidence (see section 2.4.2.2).

Obstacles or the Theory of Knowledge

After having developed this tropism to step back from everyday scientific practice and question epistemologically the shortcuts of our daily **automatic reasoning**, I realized in this *critical* stage of my thinking process, that neuro-imaging of sentence structures need more than intuitions. What I was doing in fact was Cognitive Neuro-science, and not merely neuro-imaging of a human activity where an intuitive glimpse to qualify or observe phenomena could be sufficient. This apparently tenuous distinction between neuro-imaging and Cognitive Neuro-science is in reality fundamental. The difference lies in the need of theory to advance, and in the resolution not to choose the facility (or pride) of erecting one's own intuition into an experimental paradigm. I personally learned to beware of *Intuitions*, both my own ones and the ones of others, because through experience they turned out to be either highly misleading (e.g. the fact that Precentral Gyrus, Broadmann Area 6, is only linked to sub-articulatory activity while reading), or to work without really knowing why.However, I really understood why one should keep intuitions at distance when I read the following excerpt on *opinion* from a French theoretician of scientific knowledge and founder of the so-called 'new French epistemology', Gaston Bachelard (1884 - 1962).

> When we start looking for the psychological conditions in which scientific progress is made, we are soon convinced that the problem of scientific knowledge must be posed in terms of obstacles. This is not a matter of considering external obstacles, such as the complexity and transience of phenomena, or indeed of incriminating the weakness of the senses or of the human mind. It is at the very heart of the act of cognition that, by some kind of functional necessity, sluggishness and disturbances arise. It is in the act of cognition that we shall show causes of stagnation and even of regression; there too we shall discern causes of inertia that we shall call *epistemological obstacles*.
>
> Knowledge of reality is a light that always casts a shadow in some nook or cranny. It is never immediate, never complete. *Revelations of reality are always recurrent.* ***Reality is never 'what we might believe or think it to be': it is always what we ought to have thought***[24]***. Empirical thought is clear in retrospect, when the apparatus of reason has been developed.*** *Whenever we look back and see the errors of our past, we discover truth through a real intellectual repentance.* Indeed, we know against previous knowledge, when we destroy knowledge that was badly made and surmount all those obstacles to spiritualisation that lie in the mind itself.
>
> The idea that we start from scratch when creating and increasing our possessions could only arise in cultural systems based on simple juxtaposition, where as soon as something is known it immediately becomes something that enriches. Yet, when our soul confronts all the mystery of reality, it cannot make itself ingenuous just by decree. It is impossible then

24. Le réel n'est jamais "ce qu'on pourrait croire" mais il est toujours ce qu'on aurait dû penser.

> to erase every single trace of our ordinary, everyday knowledge once and for all. ***When we contemplate reality, what we think we know very well casts its shadow over what we ought to know.*** Even when it first approaches scientific knowledge, the mind is never young. It is very old, in fact, as old as its prejudices. When we enter the realms of science, we grow younger in mind and spirit and we submit to a sudden mutation that must contradict the past.
>
> SCIENCE IS TOTALLY OPPOSED TO OPINION, not just in principle but equally in its need to come to full fruition. If it happens to justify opinion on a particular point, it is for reasons other than those that are the basis of opinion; opinion's right is therefore always to be wrong. ***Opinion thinks badly; it does not think but instead translates needs into knowledge. By referring to objects in terms of their use, it prevents itself from knowing them.*** Nothing can be founded on opinion: we must start by destroying it. Opinion is the first obstacle that has to be surmounted. It is not enough for example to rectify opinion on specific points, so maintaining provisional common knowledge like some kind of provisional morality. The scientific mind forbids us to have an opinion on questions we do not understand and cannot formulate clearly. Before all else, we have to be able to pose problems. And in scientific life, whatever people may say, problems do not pose themselves. ***It is indeed having this sense of the problem that marks out the true scientific mind. For a scientific mind, all knowledge is an answer to a question.*** If there has been no question, there can be no scientific knowledge. Nothing is self-evident. Nothing is given. Everything is constructed. [...]
>
> AN EPISTEMOLOGICAL OBSTACLE WILL ENCRUST any knowledge that is not questioned. Intellectual habits that were once useful and healthy can, in the long run, hamper research. As Bergson has so rightly said, *"our minds have the irresistible tendency to regard the idea most often of use to them as being the clearest"*. Ideas will thus acquire far too much intrinsic clarity. And with use, ideas take on unwarranted value. A value in itself impedes the circulation of values. ***It is a factor of inertia for the mind. On occasion, a dominant idea will polarize the mind in its totality.*** An irreverent epistemologist said, some twenty years ago, that great men are useful to science in the first half of their lives and harmful in the second. [...] The *formative* instinct is so persistent in some thinkers that this witticism should not alarm us. Yet this *formative* instinct will in the end yield to one that is *conservative*. There comes a time when the mind's preference is for what confirms its knowledge rather than what contradicts it, for answers rather than questions. The conservative instinct then dominates and intellectual growth stops.
>
> G. BACHELARD, *La formation de l'esprit scientifique*, CHAP. 1, 1938.

In Bachelardian terms, the itinerary of this scientific project enjoyed obstacles and delays on many grounds. The first was chronological and medical, and following the above definition, it could qualify as an "external obstacle". However, this medical pause allowed me to dive in what is here described as "contemplate reality" and have the time to feel to what extent "*what we think we know very well casts its shadow over what we*

ought to know". This period also had the advantage of favoring my acquaintance with philosophical texts on epistemology. Although, I should admit that I found this text exactly one week before my accident. Moreover, *time* - this great fellow of clarity of judgments - gave me an occasion to clarify the question I was asking about Sentence both as a syntactic unit and as a cognitive object. Hence, the *reason d'être* of chapter 1 is precisely to tame the "*inertia of the mind*" described by Bachelard, in order to posit all the questions that are at stake in this research project. Those pages are representative, I hope, of a kind of "*formative instinct*" that the epistimologist described in the above quote.

The analysis of the inevitable scientific routine proposed in this text allowed me to understand many big and small difficulties that have recurrently occurred during this research project. I will recall briefly only two of them. First, I encountered some difficulty in explaining to some colleagues of the Psychology Department hosting me in China what was a Topic, despite the fact that the grammar of their mother-tongue language had one, and an even greater difficulty explaining what was the syntactic hierarchy it generates in the sentence, as this construction is not formally taught in school. The interrogation about this difficulty has been long tickling my mind until I understood that among the obstacles Bachelard is describing here, this difficulties could probably be accounted as "opinion" or "usefulness": "our minds have the irresistible tendency to regard the idea most often of use to them as being the clearest". One should consider the 'editorial utility' of the "*opinion*" that Chinese language is in everything different from English or French, to the point of not having syntax and realizing every sentence-level dependency by simple semantic composition. However, we can say that this difficulty was useful to make "*the apparatus of reason clearer*" as Bachlard's underlines.

Secondly, Bachelard's initial remark on stagnation and on recurrent revelations in the discovery of reality has proven particularly adequate in explaining the difficulty of interpreting the results of the rich experimental designs we dared running. All this, without forgetting, of course, the problem of encrusted opinions about where in the brain certain processes should have happened.

The strongly stated position of the philosopher against intuitions and in favor of the construction of definitions of one's object of study – "Opinion thinks badly; it does not think but instead translates needs into knowledge. By referring to objects in terms of their use, it prevents itself from knowing them. Nothing can be founded on opinion" –, motivated chapter 2 and chapter 3 that deal indeed with the two fairly intuitive notions of Sentence and Topic. On these two objects of investigation we tried to put into practice the above "*Nothing is self-evident. Nothing is given. Everything is constructed*" by gathering both linguistic and experimental evidence to introduce them.

One last insight into my research experience offered by this text comes from the system of "knowledge by afterthoughts" described by Bachelard. I must say it really was the engine of a great deal of reflections about the decisions that were taken during the construction of the different experimental designs presented in Part II. This is why some part of the experimental discussion sound, using Bachelard's expression, like "intellectual repentance", hoping this can make the reader "discover truth" afterwards.

In conclusion, the very diverse "mandatory difficulties" this project suffered represented, all in all, an opportunity for perfecting the understanding of the initial question, with the words of the philosopher they definitely made "*the apparatus of reason clearer*".

An *apriori* to human knowing capacity

In this sketchy philosophical itinerary, the last stop is dedicated to what was handled at the beginning of it: the question of reality. This could be considered a recurrent question that research activity always make researchers go back to. I realized, indeed, that many colleagues shared with me this concern. Once, one of my supervisors confessed having written a long part of his manuscript on what is reality and the knowledge we have of it, but he ended deciding to erased it because 'he deemed it unsatisfactory'. If you can now read these lines, it means I haven't erased this Part, letting the reader be the judge of my foolish decision.

This chapter may surely contain trivialities or redundancies, but this should be essentially viewed as the result of the decision I made to share the long and short quotes by the intellectuals that alimented my gradual understanding of the epistemological problem I was facing in Neuro-linguistic research. The understanding of reality contained in these quotes went far beyond mine, and the ability of these authors to express it, is by far clearer than my prose. Reading them, in fact, I felt their incredible intellectual proximity, and, most of all, I felt I was not lonely in my questioning. This last point might be one of the deepest pleasures an intellectual work like a PhD can give.

The problem of reality is a matter every researcher will face one day or another, that should be, in my understanding, dealt even before starting to think or act. Once dealt with it might probably result in changing one's outlook towards the modern way of understanding science, at least this is what happened to me. The Spanish philosopher and poet María Zambrano (1904-1991) is one of those intellectuals or 'distant friends' that opened my eyes on the problem of reality we are facing nowadays. She draws a deep and structured analysis of what it means in our epoch to cast our gaze on reality. I reproduce here part of one of her writings because it summarizes probably part of the most long-lasting mark this doctoral work will leave in my personal conception of the world[25].

> In this modern times, that we can define as the epoch of reality crisis, the attitude towards reality hasn't been taken into account.
>
> And the attitude **towards reality is a different thing compared to the conditions that are required to know something starting from the mere reality perception.** We want, in this way, say in a radical way - as far as the conditions to perceive reality are concerned - that there exist in the human being a disposition toward metaphysical and practical reality that is unitary. This necessity is more of a vocation, that is to say a total necessity. A vocation by virtue of which the potentialities of the human being can uniquely be realized. And therefore among the many definitions of man that have been formulated, one could add the following one too:
>
> Man is the creature that has to realize its being through reality, a creature that was *cast* for reality. Therefore, in this sens, this vocation implies

25. It is important to note that to date, only one of Zambrano's texts has been translated from Spanish to English, and that the translation reported here is my own starting from an Italian translation: *Per l'amore e per la libertà: scritti sulla filosofia e sull'educazione* (Genova: Marietti 1820, 2008) Trans. Annarosa Buttarelli. I recommend to the reader the following essay "L'agonie de l'Europe". Trans. Maria Poumier. Valencia: Editorial UPV, 2004., which has the privilege of being the only book non-directly linked to my scientific work I read in the last six years. Another: M. Zambrano, Filosofía y Educación (Manuscritos), Editorial Ágora, Málaga 2007 which I have not had the time to read yet.

all the sensible, intellectual and of whatever order, that perception and even the mere "contact" with reality - like the sensation that we feel in front of it- requires. Things being in this way, giving a the "theory of knowledge" does not mean offering the first and primary considerations on what modern times have defined as the problem of reality.

BUT SUCH A "THEORY OF KNOWLEDGE" has to come after the **a priori** knowledge about the situation,
after the *attitude* of the man in front of reality, and after what is congenital to man in its relation to reality, and after the fact that reality is, in whatever historical circumstance, in whatever personal situation what what matters more, what matters inexorably more. That is to say, that in its specific situation in front of reality - in front, in, with - the human being discovers its strictly human and personal condition. And by modulating it, the concrete situation in which the man of particular epoch, and event the particular individual, discovers himself, in a inevitably revealing way. [...]

THE ATTITUDE that corresponds to the relationship and the comprehension of reality is more radical and deep than the intellectual operations needed to grasp it, which are considerable only as tools, instruments, methods and ways that are inevitably conditioned by the attitude towards reality.

THE ATTITUDE towards reality conditions human knowledge end even his effective presence, because human freedom manifests itself, as in everything, or, should we say even in this aspect, as the possibility of saying no or yes in front of it [reality]. Which means among others that we need to discover reality, and that,

BEFORE DISCOVERING It, we have to search it. Reality in a certain way stands by itself, overwhelmingly, inexorably - given the human condition - it requires to be searched.

HUMAN LIFE IS A TRIP TO REALITY, a trip to knowledge.

MARIA ZAMBRANO, *Filosofía y Educación*, 2007.

This far reaching insights into the dynamics of knowledge process have been clarifying many of the epistimological hunches that spontaneously rose during these years of experimental work. The problem of reality is the problem of the post-modern men and women we are, knowing reality has never, in the history of man, been so accessible given the proliferation of scientific methods and technical facilities[26]. Nonetheless, as we are living in the "epoch of *post-truth*", the emphasis we wanted to give to the attitude towards reality described by the philosopher Zambrano is that it crucially represents the most critical step of the knowledge process.

As already stipulated in the drizzling richness of European history of thought, there exists a "pure" state of thinking before experience comes into play, that conditions it *a priori*. Emmanuel Kant in his complex philosophical system postulated it as an a priori before knowledge, that he named *transcendental apperception*: an anticipative knowledge of the unity of the object that is posited before the consciousness of the subject emerges. At this stage of knowledge, consciousness names as object something it does not know

26. Although these are in fact only ways to measure it.

yet, but that it presupposes as a *unity*. What I understood is that to do research, we need not only to presuppose the object of knowledge as a unity, as it is expressed in Kantian *transcendental appreception* - and consequently reality as a unity -, but also to establish the certain attitude towards reality the philosopher is describing. The setting of this attitude is full of consequences at every step of the research process.

In conclusion and on a more personal level, I clearly see now that a radical change in attitude occurred in me during this research work, I shifted from thinking knowledge was essentially only a rational *activity*, to understanding it was ultimately a nonetheless rational *receptive attitude* towards reality. I ended up being convinced that this attitude towards reality constitutes the very core of the value of a person, and should we say the most essential and personal "work" in life?

Exodos - One last word

This house is my own and here I dwell,
I've never aped nothing from no one
and – laugh at each master, mark me well,
who at himself has not poked fun.
Over my front door.

FRIEDRICH NIETZSCHE, *Die fröhliche Wissenschaft*, FOREWORD, 1882

As a last word, we would like to propose to the reader a joyful and Greek ending to give this manuscript the theatrical form of a Comedy and hopefully not that of a Tragedy. In the *Exodos*, the final chorus chant is traditionally discussing the moral of the play, to this aim we adopt Nietzsche's paradigmatic '*joyful*' scientific spirit, and let the choir sing an excerpt from the Foreword of his *fröhliche Wissenschaft*.

As a joyful ending, two main aspects of its scientific *joy* should be underlined. While the second is present in the below quotation, the first is already contained in the original provençal title Nietzsche gave to his work – *Gaya scienza [The Gay Science]'*. The title *Die frölische Wissenshaft* is indeed a translation into German of the Provençal term *gaya*, meaning joyful or cheerful, and it was actually chosen because it was a term used by the troubadours in the XII$^{\text{th}}$ to XIV$^{\text{th}}$ century to refer to the art of *poetry*. In his *Ecce Homo*, Nietzsche wrote that he had purposely used the term *'gaya'* in *Gaya scienza* to designate the specific unity of "*singer, knight, and free spirit*" which was characteristic of early Provençal culture. After this doctoral work we namely think that this apparently heterogeneous unity of "*singer, knight, and free spirit*" is necessary in today's neuro-linguistic research to meet the challenges of a polyhedral gaze on linguistic facts.

> "Finally, lest what is most important remain unsaid: from such abysses, from such severe illness, also from the illness of severe suspicion, one returns newborn, having shed one's skin, more ticklish and malicious, with a more delicate taste for joy, with a more tender tongue for all good things, with merrier senses, joyful with a more dangerous second innocence, more childlike, and at the same time a hundred times subtler than one had ever been before.
>
> How repulsive enjoyment is to us now, that crude, muggy, brown enjoyment as understood by those who enjoy it, our 'educated', our rich, and our rulers!

How maliciously we nowadays listen to the great fairground boom-boom with which the 'educated person' and urbanite today allows art, books and music - aided by spirituous beverages - to rape him for 'forms of spiritual enjoyment'!

How the theatrical cry of passion now hurts our ears; that whole romantic uproar and tumult of the senses that is loved by the educated mob together with its aspirations towards the sublime, the elevated, the distorted, how foreign it has become to our taste!

No, if we convalescents still need art, it is another kind of art - a mocking, light, fleeting, divinely untroubled, divinely artificial art that, like a bright flame, blazes into an unclouded sky!

Above all: an art for artists, only for artists! In addition we will know better what is first and foremost needed for that: cheerfulness - any cheerfulness, my friends! As artists, too, we will know this - I would like to prove it.

There are some things we now know too well, we knowing ones: oh, how we nowadays learn as artists to forget well, to be good at not knowing! And as for our future, one will hardly find us again on the paths of those Egyptian youths who make temples unsafe at night, embrace statues, and want by all means to unveil, uncover, and put into a bright light whatever is kept concealed for good reasons.

No, we have grown sick of this bad taste, this will to truth, to 'truth at any price', this youthful madness in the love of truth: we are too experienced, too serious, too jovial, too burned, too deep for that... We no longer believe that truth remains truth when one pulls off the veil; we have lived too much to believe this.

Today we consider it a matter of decency not to wish to see everything naked, to be present everywhere, to understand and 'know' everything.

'Is it true that God is everywhere?' a little girl asked her mother; 'I find that indecent!' – a hint for philosophers! One should have more respect for the bashfulness with which nature has hidden behind riddles and iridescent uncertainties. Perhaps truth is a woman who has grounds for not showing her grounds? Perhaps her name is – to speak Greek – Baubo?... Oh, those Greeks! They knew how to live: what is needed for that is to stop bravely at the surface, the fold, the skin; to worship appearance, to believe in shapes, tones, words – in the whole Olympus of appearance! **Those Greeks were superficial - out of profundity!**

> And is not this precisely what we are coming back to, we daredevils of the spirit who have climbed the highest and most dangerous peak of current thought and looked around from up there, looked down from up there? Are we not just in this respect – Greeks? **Worshippers of shapes, tones, words?** And therefore – artists?" Friedrich Nietzsche, *Die fröhliche Wissenschaft*, Foreword, 1882.

Thanks to Nietzsche' words, we can conclude that after this doctoral study, we are indeed like Greeks, and we are left to adore the concrete factuality of the shapes, tones and the sounds of words that we perceive. We are left with this incredible mystery of the *Homo Phraseologicus*: *language*.

The next step will be persevering in considering language as a fact, what Humboldt expressed so clearly, by stating that "Language is not a construct, but it is a contingent matter of fact":

"*Spache ist nicht ein Werk aber eine Tätigkeit*".

Linguistic facts, as all facts, will hopefully remain inexhaustible, like the *contemplation* we make of them to discover their beauty, mysteries and perfection.

PART

Bibliography & Index

General index

Authors index

Thanks to this index you can enjoy another trip in this manuscript following your primary or unique interests.

Languages index

Thanks to this index you can enjoy encountering nice and inspiring people, another trip in this manuscript to enlarge your acquaintances, friends or enemies.

Bibliography

Bibliographic essay

Galeotto fu il Libro e qui lo scrisse.
Charming was the book and who wrote it.

Dantes, Divine Comedy, Hell, canto V

Bibliographic essay on central entries ordered by disciplinary and thematic area.

By nature a research question should transcends the classical disciplinary boundaries, and the interest of my Doctoral Research work lies in the fundamental claim of requiring a multidisciplinary approach to address the question of Chinese and French syntax and the representation of its structures in the brain.

This bibliographic essay traces my evolving understanding of Topic-Comment constructions as a cross-linguistic phenomenon, under multiples points of view: theoretical, cognitive, typological, formal and experimental. It has to be noticed that hardly any experimental perspective had never been addressed before by scholars in all the fields I cover in my dissertation.

A few book have accompanied me on my desks (the one in the linguistic Lab and the one in the Cognitive science Lab) for all the years of this Doctoral research. The red book (Xu Liejion and Liu Danqing), the orange book (Auray Li's Grammar) and the blue (Li Wendan, 2005) book will always be on my shelves, together with some inspirational PhD manuscripts I read entirely (2 out of the 23 manuscripts) or when back to quite frequently. Either because often consulted or because they inaugurated my reflection, they deserve mention here and to them is dedicated this Bibliographic essay.

This multidisciplinary approach helped me in providing a wider and fully rounded account of the *prosodic signature* of Topic-Comment structures in Chinese and their acceptability and processing *in context*, two linguistic aspects that offer some new insights to the understanding of the relation between Scene-Setting *in-situ* Topics and contextual information background in the field of Chinese Linguistics and psycho-lingusitics. Showing how the behavioral judgments of Native speakers are not determined by the salience of the Topic in previous context, we contribute to the experimental grounding of the basicness claimed for *gapless* Topics in Chinese.

My Dissertation on the Time-course and Functional activations of syntactic processing of Chinese Topic-Comment constructions, although investigating a novel subject that

had never been studied before, has greatly benefited of studies from the three main discipline that are interleaved in my Doctoral research, namely Linguistics, Psycho-linguistics and Neuro-linguistics.

Topic-Comment structures in Mandarin Chinese have been the object of a certain amount of studies in linguistic typology since the end of the seventies (cf. Linguistic literature on Topic Chinese –Linguistic Typology section).

Later came more formal approaches to this linguistic phenomenon in the Functional and Generative Grammar framework. Both Chinese and English publications constitute the basis of my linguistic analysis presented in the first chapter of my dissertation.

The pioneering study by Li and Thompson (1976) on Topic-prominent languages offered me the very first research hypothesis that would have been later grounding the whole experimental part of my Thesis. The architecture of my reflection has greatly benefited from the rich formal syntax literature on Topic-Comment constructions (cf. Linguistic literature on Topic Chinese – Formal Syntax section, here-under). Formal approach offered me testing hypothesis for experimental investigation. In particular, Rizzi's recent book in the Oxford Studies in Comparative Syntax collection on the Cartography of Syntactic structures (2010), and James Huang and Audrey Li's recent book, "The Syntax of Chinese "(2009), provided me not only with a useful methodological framework to describe Topic-Comment constructions, but also helped me developing the hypothesis and analysis carried out in the second part of my dissertation, chapters 5, 6 and 7.

Moreover, my linguistics chapters greatly capitalize on two previous thesis works, namely Badan (2007), Shyu (1995), who already tried to unite typological and formal approach doing Chinese linguistics.

In the second part of my dissertation I introduce the study of online Topic-Comment constructions and therefore lean on the rich literature of ERP studies realized on German Language and English.

As for the literature on the online sentence parsing mechanisms, I benefited from Colin Phillips's important contribution to the ERP sentence processing literature; and particularly two of his experiments on Chinese sentence structures (e.g. 2010 and Hsu et al. 2009) represented a key reference for my ERP research (cf. Online Sentence Processing ERP literature section, here-under). His approach helped me understand the complex and often ambivalent ERP online sentence processing, so that, only reliable and well designed studies have been selected in the ERP literature section. Other studies on the processing of Chinese sentence, from Beijing CBCS laboratory where I performed my two ERP studies, such Zhou et al. (2010), Ye et al. (2006) and Luo et al. (2010) helped me interpret the ERP results presented in chapter 5 and 4.

In the third and last part of my dissertation, it is undoubtedly the work by Pallier et al. (2011) that has been a critical starting point for the reflection carried out in my fMRI research, providing me with the key issues I have been trying to answer with the two fMRI studies I undertook. Along with the results it offers, it helped me shape the idea that sentence linguistic structures might be represented in the brain by a distributed cell assembly. From a *theoretical* point of view, the reading of works from Naama Friedmann's lab on the representation of syntactic structure in the brain, represented a critical turning point in my understanding of the final scope of my work on the relationship between syntax and the brain.

Since my dissertation directly concerns sentence-structure brain processing, I realized two preliminary state-of-the-art studies on the brain processing of verbal argument-

structure, echoing some of my linguistic work on Chinese (see chapter 2 and syntactic movement in the neuro-imaging literature (approximately 135 papers). This review allowed me to go through a lot of articles and studies, and to discover some young scholars' works that guided some of my methodological choices. Recent scholarship on "The neural correlates of linguistic distinctions", by Shetreet and Friedmann (2007 and 2012), has proven to be another essential support to my research work: especially for the approach they embody by trying to introduce the fine-grained linguistics distinctions, offered by linguistic theory, in the neuro-imaging research field. In this very same perspective, I have to quote Asaf Bachrach's Thesis (2008) too, which also provided me with a thoughtful analysis of the implications of linguistic theory for the neuro-linguistics research.

References

Aarons, Debra. 1994. "Aspects of the Syntax of American Sign Language." PhD diss., Boston University.

Abe, Kentaro, and Dai Watanabe. 2011. "Songbirds possess the spontaneous ability to discriminate syntactic rules." *Nature Neuroscience* 14 (8): 1067–1074.

Abeillé, Anne, Benoît Crabbe, Danièle Godard, and Jean-Marie Marandin. 2012. "French questioning declaratives : a corpus study." Paris, September 19.

———. 2013. "French Polar Questions : a corpus study." July 22.

Aboh, Enoch Oladé. 1998. "From the syntax of Gungbe to the grammar of Gbe." Doctoral Thesis, Université de Genève.

———. 2010. "The P Route." In *Mapping Spatial PPs,* edited by Guglielmo Cinque and Luigi Rizzi, 225–260. New York: Oxford University Press.

Aboh, Enoch Oladé, Leonie Cornips, and Jenny Doetjes. 2004. "Topic and focus within D." In *Linguistics in the Netherlands 2004,* 1–12. 21. August 24.

Aboh, Enoch Oladé, and Roland Pfau. 2011. "What's a Wh-Word Got to Do with It?" In *Mapping the Left Periphery: the Cartography of Syntactic Structures,* edited by Paola Benincà and Nicola Munaro, 5:91–124. Oxford University Press, January 31.

Abrams, Kenneth, and Thomas G. Bever. 1969. "Syntactic structure modifies attention during speech perception and recognition." *The Quarterly Journal of Experimental Psychology* 21 (3): 280–290.

Abutalebi, Jubin, Matteo Canini, Pasquale A. Della Rosa, David W. Green, and Brendan S. Weekes. 2015. "The neuroprotective effects of bilingualism upon the inferior parietal lobule: A Structural Neuroimaging Study in Aging Chinese Bilinguals." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:3–13.

Agnew, Zarinah K., Hans van de Koot, Carolyn McGettigan, and Sophie K. Scott. 2014. "Do sentences with unaccusative verbs involve syntactic movement? Evidence from neuroimaging." *Language, Cognition and Neuroscience* 29 (9): 1035–1045.

Allen, Kachina, Francisco Pereira, Matthew Botvinick, and Adele Goldberg. 2012. "Distinguishing grammatical constructions with fMRI pattern analysis." *Brain and language* 123.

Alleton, Viviane. 1973. *Grammaire du chinois.* Que sais-je ? 1519. Paris: Presses universitaires de France.

Almeida, Roberto G. de, Levi Riven, Christina Manouilidou, Ovidiu Lungu, Veena D. Dwivedi, Gonia Jarema, and Brendan Gillon. 2016. "The Neuronal Correlates of Indeterminate Sentence Comprehension: An fMRI Study." *Frontiers in Human Neuroscience* 10.

Ameka, Felix K. 1998. "Particules énonciatives en Ewe." *Faits de langues* 6 (11): 179–204.

Amunts, Katrin, Axel Schleicher, Uli Bürgel, Hartmut Mohlberg, Harry B.M. Uylings, and Karl Zilles. 1999. "Broca's region revisited: Cytoarchitecture and intersubject variability." *The Journal of Comparative Neurology* 412 (2): 319–341.

Amunts, Katrin, Peter H. Weiss, Hartmut Mohlberg, Peter Pieperhoff, Simon Eickhoff, Jennifer M. Gurd, John C. Marshall, Nadim J. Shah, Gereon R. Fink, and Karl Zilles. 2004. "Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space—The roles of Brodmann areas 44 and 45." *NeuroImage* 22 (1): 42–56.

Andler, Daniel, ed. 2004. *Introduction aux sciences cognitives.* Folio essais 179. Paris: Gallimard.

Ansaldo, Umberto, Jackie Lai, Fanlu Jia, Wai Ting Siok, Li Hai Tan, and Stephen Matthews. 2015. "Neural basis for processing hidden complexity indexed by small and finite clauses in Mandarin Chinese." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:118–127.

Aoshima, Sachiko, Colin Phillips, and Amy Weinberg. 2004. "Processing filler-gap dependencies in a head-final language." *Journal of Memory and Language* 51:23–54.

Aoshima, Sachiko, Masaya Yoshida, and Colin Phillips. 2009. "Incremental Processing of Coreference and Binding in Japanese." *Syntax* 12 (2): 93–134.

Aoun, Joseph, Lina Choueiri, and Norbert Hornstein. 2001. "Resumption, Movement, and Derivational Economy." *Linguistic Inquiry* 32 (3): 371–403.

Axelrod, Vadim, Moshe Bar, Geraint Rees, and Galit Yovel. 2014. "Neural Correlates of Subliminal Language Processing." *Cerebral Cortex* 25 (8): 2160–2169.

Ayoun, Dalila. 1999. "Verb movement in French L2 acquisition." *Bilingualism: Language and Cognition* 2:103–125.

Bachrach, Asaf. 2008. "Imaging neural correlates of syntactic complexity in a naturalistic context." PhD diss., Massachusetts Institute of Technology, Dept. of Linguistics and Philosophy.

Badan, Linda. 2007. "High and Low Periphery: a Comparison between Italian and Chinese." PhD diss., Università degli Studi di Padova.

———. 2008. "The even-construction in Mandarin Chinese." In *Chinese linguistics in Leipzig,* edited by Redouane Djamouri and Rint Sybesma, 101–116. Paris: EHESS.

Badan, Linda, and Francesca Del Gobbo. 2011. "On the Syntax of Topic and Focus in Chinese." In *Mapping the Left Periphery,* edited by Paola Benincà and Nicola Munaro, 63–90. New York: Oxford University Press.

Bahlmann, Jörg, Antoni Rodriguez-Fornells, Michael Rotte, and Thomas F. Münte. 2007. "An fMRI study of canonical and noncanonical word order in German." *Human Brain Mapping* 28 (10): 940–949.

Bai, Chen, Shushan Cai, and Petra B. Schumacher. 2011. "Reversibility in Chinese word formation influences target identification." *Neuroscience Letters* 499 (1): 14–18.

Baker, Mark C. 2001. *The Atoms Of Language: The Mind's Hidden Rules Of Grammar.* New York: Basic Books.

———. 2013. "On agreement and its relationship to case: Some generative ideas and results." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:14–32.

Balaban, Noga, Naama Friedmann, and Mira Ariel. 2016. "The effect of theory of mind impairment on language: Referring after right-hemisphere damage." *Aphasiology* 30 (12): 1424–1460.

Balaban, Noga, Naama Friedmann, and Margalit Ziv. 2016. "Theory of mind impairment after right-hemisphere damage." *Aphasiology* 30:1399–1423.

Balas, Benjamin, and Alyson Saville. 2015. "N170 face specificity and face memory depend on hometown size." *Neuropsychologia* 69:211–217.

Baldwin, R. Scott, and James M. Coady. 1978. "Psycholinguistic Approaches to a Theory of Punctuation." *Journal of Reading Behavior* 10 (4): 363–375.

Ballard, Kirrie J., and Cynthia K. Thompson. 1999. "Treatment and generalization of complex sentence production in agrammatism." *Journal of speech, language, and hearing research: JSLHR* 42 (3): 690–707.

Bally, Charles. 1944. *Linguistique générale et linguistique française.* Paris: P.U.F.

Baltin, Mark, and Chris Collins, eds. 2003. "Handbook of Contemporary Syntactic Theory." In *Handbook of Contemporary Syntactic Theory,* 670–696. Oxford, UK: Blackwell Publishing Ltd.

Bambini, Valentina, and Jacopo Torregrossa. 2010. "Cognitive categories behind early Topic/Comment structures." In *Topic, struttura dell'informazione e acquisizione linguistica,* edited by Marina Chini, 35–58. Milano: Franco Angeli.

Bastiaansen, Marcel, Lilla Magyari, and Peter Hagoort. 2010. "Syntactic Unification Operations Are Reflected in Oscillatory Dynamics during On-line Sentence Comprehension." *Journal of Cognitive Neuroscience* 22 (7): 1333–1347.

Battistella, Edwin, and Yonghui Xu. 2009. "Remarks on the reflexive in Chinese." *Linguistics* 28 (2): 205–240.

Baumann, Stefan, and Petra B. Schumacher. 2012. "(De-)accentuation and the process of information status: evidence from event-related brain potentials." *Language and Speech* 55 (Pt 3): 361–381.

Bazillon, Thierry, Benjamin Maza, Michael Rouvier, Frédéric Béchet, and Alexis Nasr. 2011. “Qui êtes vous ? Catégoriser les questions pour déterminer le rôle des locuteurs dans des conversations orales.”

Bechtel, William, and Robert C. Richardson. 2010. *Neuroimaging as a tool for functionally decomposing cognitive processes. in Foundational issues in human brain mapping.*

Beeke, Suzanne, Ray Wilkinson, and Jane Maxim. 2003a. “Exploring aphasic grammar. 1: A single case analysis of conversation.” *Clinical Linguistics & Phonetics* 17 (2): 81–107.

———. 2003b. “Exploring aphasic grammar 2: do language testing and conversation tell a similar story?” *Clinical Linguistics & Phonetics* 17 (2): 109–134.

———. 2007. “Grammar without sentence structure: A conversation analytic investigation of agrammatism.” *Aphasiology* 21 (3): 256–282.

Belletti, Adriana. 2004. “Aspects of the low IP area.” In *The structure of CP and IP. The cartography of syntactic structures,* edited by Luigi Rizzi, 2:16–51. New York: Oxford University Press.

Benincà, Paola, and Nicola Munaro, eds. 2011. *Mapping the Left Periphery.* The Cartography of Syntactic Structures 5. Oxford: Oxford University Press.

Benincà, Paola, and Cecilia Poletto. 2004. “Topic, Focus and V2: defining the CP sublayers.” In *The Structure of CP and IP: The Cartography of Syntactic Structures: Volume 2,* edited by Luigi Rizzi, 52–76. Oxford studies in comparative syntax. New York: Oxford University Press.

Ben-Shachar, Michal, Talma Hendler, Itamar Kahn, Dafna Ben-Bashat, and Yosef Grodzinsky. 2003. “The Neural Reality of Syntactic Transformations: Evidence from Functional Magnetic Resonance Imaging.” *Psychological Science* 14 (5): 433–440.

Ben-Shachar, Michal, Dafna Palti, and Yosef Grodzinsky. 2004. “Neural correlates of syntactic movement: converging evidence from two fMRI experiments.” *NeuroImage* 21 (4): 1320–1336.

Benveniste, \{{\textbackslash}’E\}mile, Noam Chomsky, Ivan Fónagy, Roman Jacobson, Jerzy Kurylowicz, Maurice Leroy, André Martinet, et al. 1966. *Problèmes du langage.* Diogène. Paris: Gallimard.

Beretta, A. 2006. “Agrammatism II: Linguistic Approaches.” In *Encyclopedia of Language & Linguistics (Second Edition),* edited by Keith Brown, 126–131. Oxford: Elsevier.

Beretta, A., and A. Munn. 1998. “Double-agents and Trace-Deletion in agrammatism.” *Brain and language* 65 (3): 404–421.

Beretta, Alan. 2008. “Disorders of Syntax.” In *Handbook of the Neuroscience of Language,* edited by Brigitte Stemmer and Harry A. Whitaker, 155–163. San Diego: Elsevier.

Berkum, Jos J. A. van, Colin M. Brown, and Peter Hagoort. 1999. “Early Referential Context Effects in Sentence Processing: Evidence from Event-Related Brain Potentials.” *Journal of Memory and Language* 41 (2): 147–182.

Berkum, Jos J. A. van, Peter Hagoort, and Colin M. Brown. 1999. "Semantic integration in sentences and discourse: evidence from the N400." *Journal of Cognitive Neuroscience* 11 (6): 657–671.

Berkum, Jos J. A. van, Arnout W. Koornneef, Marte Otten, and Mante S. Nieuwland. 2007. "Establishing reference in language comprehension: an electrophysiological perspective." *Brain Research* 1146:158–171.

Berns, M., and K. Matsuda. 2006. "Applied Linguistics: Overview and History." In *Encyclopedia of Language & Linguistics (Second Edition),* edited by Keith Brown, 394–405. Oxford: Elsevier.

Bever, Thomas G. 1970. "The cognitive basis for linguistic structures." In *Cognition and the Development of Language,* edited by John R. Hayes, 279–362. New York: Wiley / Sons.

Beyssade, Claire. 2007. "The Prosody of French Interrogatives." *Nouveaux Cahiers de Linguistique Française,* no. 28: 163–175.

Bialystok, Ellen, Fergus I. M. Craik, and Gigi Luk. 2012. "Bilingualism: consequences for mind and brain." *Trends in Cognitive Sciences* 16 (4): 240–250.

Bianchi, Valentina, and Mara Frascarelli. 2010. "Is topic a root phenomenon?" *Iberia* 2 (1): 43–88.

Biberauer, Theresa, Anders Holmberg, Ian Roberts, and Michelle Sheehan. 2014. "Complexity in comparative syntax: the view from modern parametric theory." In *Measuring Grammatical Complexity,* edited by Frederick J. Newmeyer and Laurel B. Preston, 103–127. Oxford: Oxford University Press, October 30.

Binder, Jeffrey R., Rutvik H. Desai, William W. Graves, and Lisa L. Conant. 2009. "Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies." *Cerebral Cortex* 19 (12): 2767–2796.

Biran, Michal, and Naama Friedmann. 2012. "The representation of lexical-syntactic information: Evidence from syntactic and lexical retrieval impairments in aphasia." *Cortex* 48 (9): 1103–1127.

Bloomfield, Leonard. 1933. *Language.* University of Chicago Press.

Blumstein, Sheila E., Gary Byma, Kathleen Kurowski, Jennifer Hourihan, Todd Brown, and Adele Hutchinson. 1998. "On-Line Processing of Filler–Gap Constructions in Aphasia." *Brain and Language* 61 (2): 149–168.

Bocci, Giuliano. 2013. *The Syntax–Prosody Interface: A cartographic perspective with evidence from Italian.* Vol. 204. Linguistik Aktuell/Linguistics Today. Amsterdam: John Benjamins Publishing Company, August 13.

Bock, Kathryn, Gary S. Dell, Franklin Chang, and Kristine H. Onishi. 2007. "Persistent structural priming from language comprehension to language production." *Cognition* 104 (3): 437–458.

Boeckx, Cedric, Anna Martinez-Alvarez, and Evelina Leivada. 2014. "The functional neuroanatomy of serial order in language." *Journal of Neurolinguistics* 32:1–15.

Bonnot, Christine. 1999. “Pour une définition formelle et fonctionnelle de la notion de thème (sur l’exemple du russe moderne).” In *La thématisation dans les langues, actes du colloque de Caen (9-11 octobre 1997),* edited by Claude Guimier, 15–31. Bern: Peter Lang.

Bor, Daniel, and Adrian M. Owen. 2007. “A common prefrontal-parietal network for mnemonic and mathematical recoding strategies within working memory.” *Cerebral Cortex (New York, N.Y.: 1991)* 17 (4): 778–786.

Borillo, Andrée. 1978. “Structure et valeur énonciative de l’interrogation totale en français.” Thèse d’Etat, Université d’Aix-en-Provence.

Bornkessel, Ina, and Matthias Schlesewsky. 2006a. “The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages.” *Psychological Review* 113 (4): 787–821.

———. 2006b. “The Role of Contrast in the Local Licensing of Scrambling in German: Evidence from Online Comprehension.” *Journal of Germanic Linguistics* 18 (1): 1–43.

Bornkessel, Ina, Matthias Schlesewsky, and Angela D. Friederici. 2002a. “Beyond syntax: language-related positivities reflect the revision of hierarchies.” *Neuroreport* 13 (3): 361–364.

———. 2002b. “Grammar overrides frequency: evidence from the online processing of flexible word order.” *Cognition* 85 (2): B21–30.

———. 2003a. “Contextual information modulates initial processes of syntactic integration: the role of inter- versus intrasentential predictions.” *Journal of Experimental Psychology. Learning, Memory, and Cognition* 29 (5): 871–882.

———. 2003b. “Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing.” *Language and Cognitive Processes* 18 (3): 269–298.

Bornkessel, Ina, Stefan Zysset, Angela D. Friederici, D. Yves von Cramon, and Matthias Schlesewsky. 2005. “Who did what to whom? The neural basis of argument hierarchies during language comprehension.” *NeuroImage* 26 (1): 221–233.

Bornkessel-Schlesewsky, Ina, Tanja Grewe, and Matthias Schlesewsky. 2012. “Prominence vs. aboutness in sequencing: a functional distinction within the left inferior frontal gyrus.” *Brain and Language* 120 (2): 96–107.

Bornkessel-Schlesewsky, Ina, and Matthias Schlesewsky. 2008. “An alternative perspective on ”semantic P600” effects in language comprehension.” *Brain Research Reviews* 59 (1): 55–73.

———. 2009a. “Minimality as vacuous distinctness: Evidence from cross-linguistic sentence comprehension.” *Lingua,* Experimental Evidence for minimal structure, 119 (10): 1541–1559.

———. 2009b. “The Role of Prominence Information in the Real-Time Comprehension of Transitive Constructions: A Cross-Linguistic Approach.” *Language and Linguistics Compass* 3 (1): 19–58.

Bornkessel-Schlesewsky, Ina, Matthias Schlesewsky, and D. Yves von Cramon. 2009. "Word order and Broca's region: evidence for a supra-syntactic perspective." *Brain and Language* 111 (3): 125–139.

Bornkessel-Schlesewsky, Ina, and Petra B. Schumacher. 2016. "Towards a Neurobiology of Information Structure."

Bowerman, Melissa. 1983. "Structural relationships in children's utterances: Semantic or syntactic?" In *Cognitive development and the acquisition of language,* edited by Timothy E. Moore, 197–213. New York: Academic Press.

Bradley, Lynette, and Peter E. Bryant. 1983. "Categorizing sounds and learning to read —a causal connection." *Nature* 301 (5899): 419–421.

Brandl, Johannes. 2014. "Brentano's Theory of Judgement." In *The Stanford Encyclopedia of Philosophy,* Summer 2014, edited by Edward N. Zalta.

Branigan, Holly P., Martin J. Pickering, and A. A. Cleland. 2000. "Syntactic co-ordination in dialogue." *Cognition* 75 (2): B13–B25.

Brennan, Jonathan R., and Liina Pylkkänen. 2017. "MEG Evidence for Incremental Sentence Composition in the Anterior Temporal Lobe." *Cognitive Science* 41:1515–1531.

Brennan, Jonathan, Yuval Nir, Uri Hasson, Rafael Malach, David J. Heeger, and Liina Pylkkänen. 2012. "Syntactic structure building in the anterior temporal lobe during natural story listening." *Brain and Language* 120 (2): 163–173.

Bresnan, Joan, and Jonni M. Kanerva. 1989. "Locative Inversion in Chichewa: A Case Study of Factorization in Grammar." *Linguistic Inquiry* 20 (1): 1–50.

Brown, C. M., and Peter Hagoort. 1993. "The processing nature of the N400: evidence from masked priming." *Journal of Cognitive Neuroscience* 5 (1): 34–44.

Bruening, Benjamin. 2010. "Language-particular syntactic rules and constraints: English locative inversion and do-support." *Language* 86:43–84.

Buiatti, Marco, Marcela Peña, and Ghislaine Dehaene-Lambertz. 2009. "Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses." *NeuroImage* 44 (2): 509–519.

Burchert, Frank, Maria Swoboda-Moll, and Ria De Bleser. 2005. "Tense and Agreement dissociations in German agrammatic speakers: Underspecification vs. hierarchy." *Brain and Language* 94 (2): 188–199.

Büring, Daniel. 2003. "On D-Trees, Beans, and B-Accents." *Linguistics and Philosophy* 26 (5): 511–545.

———. 2007. "Topic and Comment." In *The Cambridge Encyclopedia of the Language Sciences,* edited by Patrick Colm Hogan. Cambridge: Cambridge University Press.

Burkhardt, Petra. 2006. "Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials." *Brain and Language* 98 (2): 159–168.

———. 2007. "The P600 reflects cost of new information in discourse memory." *Neuroreport* 18 (17): 1851–1854.

Burkhardt, Petra, Maria Mercedes Piñango, and Keng Wong. 2003. "The role of the anterior left hemisphere in real-time sentence comprehension: Evidence from split intransitivity." *Brain and Language,* Understanding Language, 86 (1): 9–22.

Burkhardt, Petra, and Dietmar Roehm. 2007. "Differential effects of saliency: An event-related brain potential study." *Neuroscience Letters* 413 (2): 115–120.

———. 2008. "Syntactic function and definiteness determine the prominence of an antecedent: Revisiting the Repeated Name Penalty." Chapel Hill, NH, March 15.

Burmester, Juliane, Katharina Spalek, and Isabell Wartenburger. 2014. "Context updating during sentence comprehension: The effect of aboutness topic." *Brain and Language* 137 (Supplement C): 62–76.

Bursill-Hall, Geoffrey L., Sten Ebbesen, and E. F. K. Koerner. 1990. *De Ortu Grammaticae: Studies in medieval grammar and linguistic theory in memory of Jan Pinborg.* Amsterdam: John Benjamins Publishing, January 1.

Cabeza, Roberto, and Lars Nyberg. 2000. "Imaging cognition II: An empirical review of 275 PET and fMRI studies." *Journal of Cognitive Neuroscience* 12 (1): 1–47.

Cabredo-Hofherr, Patricia. 2004. "Les clitiques sujets du français et le paramètre du sujet nul." *Langue française* 141 (1): 99–109.

Cairns, Helen Smith. 1999. *Psycholinguistics: An Introduction.* Austin, Texas: Pro-Ed.

Calvo, Alejandra, and Ellen Bialystok. 2014. "Independent effects of bilingualism and socioeconomic status on language ability and executive functioning." *Cognition* 130 (3): 278–288.

Canseco-Gonzalez, Enriqueta. 2000. "Using the Recording of Event-Related Brain Potentials in the Study of Sentence Processing." In *Language and the Brain, Representation and Processing,* edited by Yosef Grodzinsky, Lewis P. Shapiro, and David Swinney, 229–266. Foundations of Neuropsychology. San Diego, CA: Academic Press.

Cantiani, Chiara, Maria Luisa Lorusso, Paolo Perego, and Maria Teresa Guasti. 2010. "ERP correlates of anomalous morphosyntactic processing in adults and children with developmental dyslexia." In *BULCD 35 Online Proceedings Supplement,* edited by Nick Danis, Kate Mesh, and Hyunsuk Sung. Somerville, MA: Cascadilla Press, November 5.

Caplan, D. 2015. "Grammar and Syntax." In *Brain Mapping,* edited by Arthur W. Toga, 667–670. Waltham: Academic Press.

Caplan, David, Nathaniel Alpert, Gloria Waters, and Anthony Olivieri. 2000. "Activation of Broca's area by syntactic processing under conditions of concurrent articulation." *Human Brain Mapping* 9 (2): 65–71.

Caramazza, Alfonso, Costanza Papagno, and Wheeler Ruml. 2000. "The Selective Impairment of Phonological Processing in Speech Production." *Brain and Language* 75 (3): 428–450.

Caramazza, Alfonso, and Edgar B. Zurif. 1976. "Dissociation of algorithmic and heuristic processes in language comprehension: evidence from aphasia." *Brain and Language* 3 (4): 572–582.

Cardinaletti, Anna. 2004. “Towards a cartography of subject positions.” In *The structure of CP and IP. The cartography of syntactic structures,* edited by Luigi Rizzi, 2:115–165. New York: Oxford University Press.

Carreiras, Manuel, Blair C. Armstrong, Manuel Perea, and Ram Frost. 2014. “The what, when, where, and how of visual word recognition.” *Trends in Cognitive Sciences* 18 (2): 90–98.

Carreiras, Manuel, Jon Andoni Duñabeitia, Marta Vergara, Irene de la Cruz-Pavía, and Itziar Laka. 2010. “Subject relative clauses are not universally easier to process: Evidence from Basque.” *Cognition* 115 (1): 79–92.

Carreiras, Manuel, Chotiga Pattamadilok, Enrique Meseguer, Horacio Barber, and Joseph T. Devlin. 2012. “Broca’s area plays a causal role in morphosyntactic processing.” *Neuropsychologia* 50 (5): 816–820.

Carvalho, Alex de, Isabelle Dautriche, Isabelle Lin, and Anne Christophe. 2017. “Phrasal prosody constrains syntactic analysis in toddlers.” *Cognition* 163 (Supplement C): 67–79.

Casado, Pilar, Manuel Martín-Loeches, Francisco Muñoz, and Carlos Fernández-Frías. 2005. “Are semantic and syntactic cues inducing the same processes in the identification of word order?” *Brain Research. Cognitive Brain Research* 24 (3): 526–543.

Cat, Cécile de. 2007. “French dislocation without movement.” *Natural Language & Linguistic Theory* 25 (3): 485–534.

Cattinelli, Isabella, N. Alberto Borghese, Marcello Gallucci, and Eraldo Paulesu. 2013. “Reading the reading brain: A new meta-analysis of functional imaging data on reading.” *Journal of Neurolinguistics* 26 (1): 214–238.

Cavanna, Andrea E., and Michael R. Trimble. 2006. “The precuneus: a review of its functional anatomy and behavioural correlates.” *Brain: A Journal of Neurology* 129 (Pt 3): 564–583.

Chacón, Dustin Alfonso, Shota Momma, and Colin Phillips. 2016. “Linguistic representations and memory architectures: The devil is in the details.” *Behavioral and Brain Sciences* 39.

Chafe, Wallace L. 1974. “Language and Consciousness.” *Language* 50 (1): 111–133.

———. 1976. “Givenness, contrastiveness, definiteness, subjects, topics, and point of view.” In *Subject and Topic,* edited by Charles N. Li, 25–55. New York: Academic Press.

———. 1985. “Linguistic Differences Produced by Differences Between Speaking and Writing.” In *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing,* edited by David R. Olson, Nancy Torrance, and Angela Hildyard, 105–122. Cambridge: Cambridge University Press.

Chafe, Wallace L. 1987. “Cognitive constraints on information flow.” In *Coherence and Grounding in Discourse: Outcome of a Symposium, Eugene, Oregon, June 1984,* edited by Russell S. Tomlin, 11:21–51. Typological Studies in Language. Amsterdam: John Benjamins Publishing Company.

Chang 張, Yueh-chin 月琴. 2001. “Phonetic realization of ambiguous phrases in Taiwan mandarin [台灣華語歧義句的語音體現].” *Journal of Chinese Linguistics Monograph Series,* no. 17: 168–191.

Chang, Hsing-Wu. 1992. “The Acquisition of Chinese Syntax.” In *Language Processing in Chinese,* edited by Hsuan-Chih Chen and Ovid J. -L. Tzeng, 90:277–311. Advances in Psychology. Amsterdam: North-Holland, January 1.

Chang, Lisa. 1997. “Wh-in-situ phenomena in French.” Master of Arts dissertation, University of British Columbia.

Chang, Li-Yun, Joseph Z. Stafura, Ben Rickles, Hsueh-Chih Chen, and Charles A. Perfetti. 2015. “Incremental learning of Chinese orthography: ERP indicators of animated and static stroke displays on character form and meaning acquisition.” *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:78–95.

Chao, L. L., and A. Martin. 1999. “Cortical regions associated with perceiving, naming, and knowing about colors.” *Journal of Cognitive Neuroscience* 11 (1): 25–35.

Chao, Yuanren. 1968. *A Grammar of Spoken Chinese.* Berkeley: Univeristy of California Press.

Chappell, Hilary. 1991. “Strategies for the assertion of obviousness and disagreement in mandarin: A semantic study of the modal particle me.” *Australian Journal of Linguistics* 11 (1): 39–65.

———. 1994. “Mandarin Semantic Primitives.” In *Semantic and Lexical Universals: Theory and empirical findings,* edited by Cliff Goddard and Anna Wierzbicka, 25:109. Amsterdam: John Benjamins Publishing Company.

Charolles, Michel. 2003. “De la topicalité des adverbiaux détachés en tête de phrase.” *Travaux de linguistique* 47 (2): 11–49.

Charolles, Michel, and Sophie Prévost. 2003. *Adverbiaux et topiques.* Bruxelles: Duculot.

Chee, Michael W. L., Edsel W. L. Tan, and Thorsten Thiel. 1999. “Mandarin and English Single Word Processing Studied with Functional Magnetic Resonance Imaging.” *Journal of Neuroscience* 19 (8): 3050–3056.

Chen, Caicai. 2011. *The Influence of Chinese Topic Prominence Construction on English Acquisition —A Study on Non-English Majors and English Majors at a Chinese College.* Kristianstad: Kristianstad University Sweden.

Chen, Chun-Yin Doris. 1995. “Topics in Chinese: A or A’?” *Studies in English Literature and Linguistics* 21:79–106.

Chen, Evan, Edward Gibson, and Florian Wolf. 2005. “Online syntactic storage costs in sentence comprehension.” *Journal of Memory and Language* 52 (1): 144–169.

Chen, Evan, W. Caroline West, Gloria Waters, and David Caplan. 2006. "Determinants of BOLD signal correlates of processing object-extracted relative clauses." *Cortex* 42 (4): 591–604.

Chen, Ping. 1996a. "Pragmatic interpretation of structural topic and relativization in Chinese." *Journal of Pragmatics* 26:389–406.

———. 1996b. "Pragmatic interpretations of structural topics and relativization in Chinese." *Journal of Pragmatics* 26 (3): 389–406.

Chen, Yiya. 2010. "Post-focus F0 compression – Now you see it, now you don't." *Journal of Phonetics* 38 (4): 517–525.

Cheng, Lisa Lai Shen, Norbert Corver, and Luigi Rizzi, eds. 2006. "On the Form of Chains: Criterial Positions and ECP Effects." In *Wh-movement : moving on,* 97–133. Cambridge, Mass. : MIT Press.

Cheng, Lisa Lai-Shen. 1991. "On the typology of wh-questions." PhD diss., Massachusetts Institute of Technology.

Cheng, Lisa Lai-Shen, and Johan Rooryck. 2000. "Licensing Wh-in-situ." *Syntax* 3 (1): 1–19.

Cheung, Candice Chi-Hang. 2015. "On the Fine Structure of the Left Periphery." In *The Cartography of Chinese Syntax,* edited by Wei-Tien Dylan Tsai, 11:75–130. Oxford: Oxford University Press.

Chien, Yu-Chin, and Barbara Lust. 1983. "Topic-Comment Structure and Grammatical Subject in First Language Acquisition of Mandarin Chinese: A Study of Equi-Constructions." In *Papers and Reports on Child Language Development, Volume 22,* 22:74–82. Papers and Reports on Child Language Development. Stanford, CA: PRCLD, Department of Linguistics, March.

Chiss, Jean-Louis. 1985. "La stylistique de Charles Bally : de la notion de « sujet parlant » à la théorie de renonciation." *Langages* 19 (77): 85–94.

Choe, Hyon Sook. 1995. "Focus and Topic Movement in Korean and Licensing." In *Discourse configurational languages,* edited by Katalin É. Kiss, 269–334. New York: Oxford University Press.

Chomsky, Noam. 1959. "Review of Verbal behavior." In collaboration with B. F. Skinner. *Language* 35 (1): 26–58.

———. 1981. *Lectures on Government and Binding.* Dordrecht: Foris Publication.

———. 1993. *Lectures on Government and Binding: The Pisa Lectures.* 7 Sub. Studies in Generative Grammar 9. Mouton de Gruyter.

———. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

———. 2013. "Problems of projection." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:33–49.

Cho-Reyes, Soojin, Jennifer E. Mack, and Cynthia K. Thompson. 2016. "Grammatical encoding and learning in agrammatic aphasia: Evidence from structural priming." *Journal of Memory and Language,* New Approaches to Structural Priming, 91 (Supplement C): 202–218.

Chou, Tai-Li, Shu-Hui Lee, Shao-Min Hung, and Hsuan-Chih Chen. 2012. "The role of inferior frontal gyrus in processing Chinese classifiers." *Neuropsychologia* 50 (7): 1408–1415.

Christensen, Ken Ramshøj. 2008. "Interfaces, syntactic movement, and neural activation: A new perspective on the implementation of language in the brain." *Journal of Neurolinguistics* 21 (2): 73–103.

———. 2010. "Syntactic reconstruction and reanalysis, semantic dead ends, and prefrontal cortex." *Brain and Cognition* 73 (1): 41–50.

Christensen, Ken Ramshøj, and Mikkel Wallentin. 2011. "The locative alternation: distinguishing linguistic processing cost from error signals in Broca's region." *NeuroImage* 56 (3): 1622–1631.

Christophe, Anne, Séverine Millotte, Savita Bernal, and Jeffrey Lidz. 2008. "Bootstrapping lexical and syntactic acquisition." *Language and Speech* 51 (Pt 1-2): 61–75.

Christophe, Anne, Sharon Peperkamp, Christophe Pallier, Eliza Block, and Jacques Mehler. 2004. "Phonological phrase boundaries constrain lexical access I. Adult data." *Journal of Memory and Language* 51 (4): 523–547.

Chu, Chauncey C. 1989. "The interplay of syntax, semantics, and pragmatics in Mandarin Chinese." In *Functionalism and Chinese Grammar,* edited by James H.-Y. Tai and Frank F.S. Hsueh, 19–43. Chinese Teachers Association Monograph Series 1. South Orange, NJ: Chinese Language Teachers Association.

———. 1993. "The Prototypicality of Topic in Mandarin Chinese." *Journal of the Chinese Language Teachers Association* 28 (1): 25–48.

———. 1998. *A Discourse Grammar of Mandarin Chinese.* Berkeley Models of Grammars 6. New York: Peter Lang.

———. n.d. "A Contrastive Approach to Discourse Particles-A Case Study of the Mandarin UFP ne 呢." *外国语 Waiguoyu [Journal of Foreign Languages = journal of the Shanghai Foreign Language University].*

Chwilla, Dorothee J., Colin M. Brown, and Peter Hagoort. 1995. "The N400 as a function of the level of processing." *Psychophysiology* 32 (3): 274–285.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective.* Oxford Studies in Comparative Syntax. Oxford, New York: Oxford University Press, March 18.

———. 2005. "Deriving Greenberg's Universal 20 and Its Exceptions." *Linguistic Inquiry* 36 (3): 315–332.

———. 2013. "Cognition, universal grammar, and typological generalizations." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:50–65.

Cinque, Guglielmo, and Luigi Rizzi. 2009. "The Cartography of Syntactic Structures." In *The Oxford Handbook of Linguistic Analysis,* edited by Bernd Heine and Heiko Narrog. New York: Oxford University Press.

———. 2010. *Mapping Spatial PPs: The Cartography of Syntactic Structures.* New York: Oxford University Press, August 13.

Clark, Eve V. 2016. *First Language Acquisition.* Cambridge: Cambridge University Press.

Clarke, Edwin, Kenneth Dewhurst, and Michael Jeffrey Aminoff. 1996. *An Illustrated History of Brain Function: Imaging the Brain from Antiquity to the Present.* Norman Publishing.

Colin, Phillips. 2013. "Parser-grammar relations: We don't understand everything twice." In *Language down the garden path: the cognitive and biological basis for linguistic structures,* edited by Montserrat Sanz, Itziar Laka, and Michael K Tanenhaus, 294–315. Oxford, New York: Oxford University Press.

Combettes, Bernard, and Sophie Prévost. 2001. "Évolution des marqueurs de topicalisation." *Cahiers de praxématique,* no. 37: 103–124.

Comrie, Bernard. 1981. *Language Universals and Linguistic Typology.* Chicago: University of Chicago Press.

———. 1988. "Topics, Grammaticalized Topics, and Subjects." *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 14:265–279.

Constant, Noah. 2011. "On the Independence of Mandarin Aspectual and Contrastive Sentence-Final \textit{ne}." In *Proceedings of the 23rd North American Conference on Chinese Lingusitics (NACCL-23),* edited by Jing-Schmidt Zhuo, 15–29. Ohio State University: NACCL Proceedings Online.

———. 2012. "Topic Abstraction as the Source for Nested Alternatives: A Conservative Semantics for Contrastive Topic." In *Proceedings of the 30th West Coast Conference on Formal Linguistics,* edited by Nathan Arnett and Ryan Benett, 120–130. Somerville, MA: Cascadilla Proceedings Project.

———. 2014. "Contrastive Topic: Meanings and Realizations." PhD diss., University of Massachusetts, Amherst.

Constant, Noah, and Chloe C. Gu. 2010. "Mandarin 'even', 'all' and the Trigger of Focus Movement." In *Proceedings of the 33rd Annual Penn Linguistics Colloquium,* 16:21–30.

Contemori, Carla, and Maria Garraffa. 2010. "Comparison of modalities in SLI syntax: A study on the comprehension and production of non-canonical sentences." *Lingua* 120 (8): 1940–1955.

Cooke, Ayanna, Murray Grossman, Christian DeVita, Julio Gonzalez-Atavales, Peachie Moore, Willis Chen, James Gee, and John Detre. 2006. "Large-scale neural network for sentence processing." *Brain and Language* 96 (1): 14–36.

Coopmans, Peter. 1989. "Where Stylistic and Syntactic Processes Meet: Locative Inversion in English." *Language* 65 (4): 728–751.

Coppens, Patrick, and Janet Patterson. 2017. *Aphasia Rehabilitation: Clinical Challenges.* Jones & Bartlett Learning, January 6.

Courage, Mary L., Aishah Bakhtiar, Cheryll Fitzpatrick, Sophie Kenny, and Katie Brandeau. 2015. "Growing up multitasking: The costs and benefits for cognitive development." *Developmental Review,* Special Issue: Living in the "Net"Generation: Multitasking, Learning, and Development, 35:5–41.

Coveney, Aidan. 2012. "L'interrogation directe." *Travaux de linguistique,* no. 63: 112–145.

Crocker, Matthew W. 1994. "On the nature of the principle-based sentence processor." In *Perspectives on Sentence Processing,* edited by Charles Clifton Jr., Lyn Frazier, and Keith Rayner, 245–266. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cui 崔, Xiliang 希亮. 1990. "试论关联形式 "连... 也/都..." 的多重语言信息 Shilun guanlian xingshi "lian ... ye/dou ... " de duozhong yuyqn xinxi [The multiple information of conjunctive form lian...ye/dou]." *世界汉语教学 Shijie hanyu jiaoxue [Chinese Teaching in the world],* no. 3: 139–147.

Culbertson, Jenny, and Géraldine Legendre. 2008. "Qu'en est-il des clitiques sujet en français oral contemporain ?" In *Actes du 1er Congrès Mondial de Linguistique Française,* edited by Jacques Durand, Benoît Habert, and Bernard Laks, 2651–2662. Paris: EDP Sciences.

Culioli, Antoine. 1990. *Pour une linguistique de l'énonciation: opérations et représentations.* Vol. 1. Homme dans la langue. Paris: Ophrys.

Cutler, Anne. 1981. "Making up materials is a confounded nuisance, or: Will we able to run any psycholinguistic experiments at all in 1990?" *Cognition* 10 (1): 65–70.

Cutler, Anne, and D. Robert Ladd, eds. 1983. *Prosody: Models and Measurements.* Redacted by Willem J. M. Levelt. Vol. 14. Springer Series in Language and Communication. Berlin: Springer.

Cutler, Anne, Delphine Oahan, and Wilma van Donselaar. 1997. "Prosody in the comprehension of spoken language: a literature review." *Language and Speech* 40 (Pt 2):141–201.

Cyrino, Sonia. 1997. *O objeto nulo no português do Brasil: um estudo sintático-diacrônico.* London: UEL.

Cyrino, Sonia Maria. 2004. "Null Complement Anaphora and Null Objects in Brazilian Portuguese."

Cyrino, Sonia, and Uli Reich. 2001. "Uma visão integrada do objeto nulo no Português Brasileiro." *Romanistisches Jahrbuch* 52 (1): 360.

Dai 戴, Qingsha 庆厦. 2001. "景颇语的话题 Jingpoyu de huati." *语言研究 Yuyan yanjiu* 1:100–105.

Danet, Brenda, and Susan C. Herring. 2007. *The Multilingual Internet: Language, Culture, and Communication Online.* Oxford University Press.

De Benedictis, Alessandro, Hugues Duffau, Beatrice Paradiso, Enrico Grandi, Sergio Balbi, Enrico Granieri, Enzo Colarusso, Franco Chioffi, Carlo Efisio Marras, and Silvio Sarubbo. 2014. "Anatomo-functional study of the temporo-parieto-occipital region: dissection, tractographic and brain mapping evidence from a neurosurgical perspective." *Journal of Anatomy* 225 (2): 132–151.

Dechauvelle, Anne-Dominique. 2008. "Exploration du réseau cérébral impliqué dans les traitements syntaxiques et lexico-sémantiques des phrases." PhD diss., Université de Paris VI.

Demiral, Şükrü Barış, Matthias Schlesewsky, and Ina Bornkessel-Schlesewsky. 2008. "On the universality of language comprehension strategies: Evidence from Turkish." *Cognition* 106 (1): 484–500.

Demuth, Katherine. 1990. "Subject, topic and Sesotho passive." *Journal of Child Language* 17 (1): 67–84.

Desikan, Rahul S., Florent Ségonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, et al. 2006. "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest." *NeuroImage* 31 (3): 968–980.

Devauchelle, Anne-Dominique, Catherine Oppenheim, Luigi Rizzi, Stanislas Dehaene, and Christophe Pallier. 2009. "Sentence syntax and content in the human temporal lobe: an fMRI adaptation study in auditory and visual modalities." *Journal of Cognitive Neuroscience* 21 (5): 1000–1012.

Diaz, Michele T., and Gregory McCarthy. 2009. "A comparison of brain activity evoked by single content and function words: An fMRI investigation of implicit word processing." *Brain research* 1282:38–49.

Díaz-Santos, Mirella, Bo Cao, Arash Yazdanbakhsh, Daniel J. Norton, Sandy Neargarder, and Alice Cronin-Golomb. 2015. "Perceptual, cognitive, and personality rigidity in Parkinson's disease." *Neuropsychologia* 69:183–193.

Dickey, Michael Walsh, JungWon Janet Choy, and Cynthia K. Thompson. 2007. "Real-time comprehension of wh- movement in aphasia: Evidence from eyetracking while listening." *Brain and language* 100 (1): 1–22.

Dickey, Michael Walsh, and Cynthia K. Thompson. 2009. "Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking." *Journal of neurolinguistics* 22 (6): 563–583.

Dik, Simon C. 1989. *The Theory of Functional Grammar*. Functional Grammar Series 20. Berlin: De Gruyter.

Dimitrova, Diana V. 2011. "The neurobiology of linguistic prosody." Lunteren.

Dimitrova, Diana V., Laurie A. Stowe, Gisela Redeker, and John C. J. Hoeks. 2012. "Less Is Not More: Neural Responses to Missing and Superfluous Accents in Context." *Journal of Cognitive Neuroscience* 24 (12): 2400–2418.

Ding, Guosheng, Danling Peng, and Marcus Taft. 2004. "The Nature of the Mental Representation of Radicals in Chinese: A Priming Study." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30 (2): 530–539.

Dominey, Peter Ford, and Michel Hoen. 2006. "Structure mapping and semantic integration in a construction-based neurolinguistic model of sentence processing." *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 42 (4): 476–479.

Dowty, David. 1991. "Thematic Proto-Roles and Argument Selection." *Language* 67 (3): 547–619.

Drijbooms, Elise, and Roelien Bastiaanse. 2011. "Production of pronouns and reflexives by agrammatic French-Dutch bilingual speakers." In *Paper presented at Science of Aphasia XII*. Barcelona.

Dronkers, Nina F., Odile Plaisant, Marie-Thérèse Iba-Zizen, and Emmanuel Alain Cabanis. 2007. "Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong." *Brain: A Journal of Neurology* 130 (Pt 5): 1432–1441.

Dronkers, Nina F, David P Wilkins, Robert D Van Valin, Brenda B Redfern, and Jeri J Jaeger. 2004. "Lesion analysis of the brain areas involved in language comprehension." *Cognition,* Towards a New Functional Anatomy of Language, 92 (1): 145–177.

Duncan, John. 2010. "The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour." *Trends in Cognitive Sciences* 14 (4): 172–179.

Duncan, Stewart. 2013. "Thomas Hobbes." In *The Stanford Encyclopedia of Philosophy,* Summer 2013, edited by Edward N. Zalta.

Ebert, Christian, Cornelia Ebert, and Stefan Hinterwimmer. 2013. "The Interpretation of the German Specificity Markers Bestimmt and Gewiss." In *Different Kinds of Specificity Across Languages,* 31–74. Studies in Linguistics and Philosophy. Dordrecht: Springer.

Ebert, Christian, Cornelia Endriss, and Stefan Hinterwimmer. 2008. "A Unified Analysis of Indicative and Biscuit Conditionals as Topics." *Semantics and Linguistic Theory* 18 (0): 266–283.

Ebert, Cornelia, and Stefan Hinterwimmer. 2009. "The interpretation of topical indefinites as direct and indirect aboutness topics." In *Information Structure,* edited by Malte Zimmermann and Caroline Féry, 89–114. Oxford, New York: Oxford University Press, October.

Eckstein, Korinna, and Angela D. Friederici. 2006. "It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension." *Journal of Cognitive Neuroscience* 18 (10): 1696–1711.

Embick, David, Alec Marantz, Yasushi Miyashita, Wayne O'Neil, and Kuniyoshi L. Sakai. 2000. "A syntactic specialization for Broca's area." *Proceedings of the National Academy of Sciences* 97 (11): 6150–6154.

Endriss, Cornelia. 2009. *Quantificational topics: a scopal treatment of exceptional wide scope phenomena.* Dordrecht: Springer.

Epstein, Russell, and Nancy Kanwisher. 1998. “A cortical representation of the local visual environment.” *Nature* 392 (6676): 598–601.

Erbaugh, Mary S. 1983. “Why Chinese children’s acquisition of Mandarin predicates should be “exactly like English”.” *Papers and Reports on Child Language Development* 22:49–57.

Erdocia, Kepa, Itziar Laka, Anna Mestres-Missé, and Antoni Rodriguez-Fornells. 2009. “Syntactic complexity and ambiguity resolution in a free word order language: behavioral and electrophysiological evidences from Basque.” *Brain and Language* 109 (1): 1–17.

Ernst, Thomas. 2004. “Domain adverbs and the syntax of adjuncts.” In *Adverbials: The interplay between meaning, context, and syntactic structure,* edited by Jennifer R. Austin, Stefan Engelberg, and Gisa Rauh, 103–129. Linguistik Aktuell/Linguistics Today 70. Amsterdam: John Benjamins.

Ernst, Thomas, and Chengchi Wang. 1995. “Object preposing in Mandarin Chinese.” *Journal of East Asian Linguistics* 4 (3): 235–260.

Erteschik-Shir, Nomi. 1999. “Focus Structure and Scope.” In *Grammar of Focus,* edited by Georges Rebuschi and Laurice Tuller, 119. Linguistik Aktuell/Linguistics Today 24. Amsterdam: John Benjamins Publishing Company.

———. 2006. *The Dynamics of Focus Structure.* Cambridge Studies in Linguistics 84. Cambridge: Cambridge University Press.

Erteschik-Shir, Nomi, Lena Ibnbari, and Sharon Taube. 2013. “Missing objects as Topic Drop.” *Lingua,* SI: Information Structure Triggers, 136 (Supplement C): 145–169.

Etard, Olivier, and Nathalie Tzourio-Mazoyer, eds. 2003. *Cerveau et langage.* Paris: Hermès Science Publications.

Evans, Vyvyan, and Paul A. Chilton, eds. 2010. *Language, Cognition and Space: the State of the Art and New Directions.* Advances in cognitive linguistics. London: Equinox Publushing.

Everaert, Martin B. H., Marinus A. C. Huybregts, Noam Chomsky, Robert C. Berwick, and Johan J. Bolhuis. 2015. “Structures, Not Strings: Linguistics as Part of the Cognitive Sciences.” *Trends in Cognitive Sciences* 19 (12): 729–743.

Fabiani, Monica, Gabriele Gratton, and Kara D. Federmeier. 2007. “Event-Related Brain Potentials: Methods, Theory, and Applications.” In *Handbook of Psychophysiology,* Third edition, edited by John T. Cacioppo, Louis G. Tassinary, and Gary Berntson, 85–119. Cambridge: Cambridge University Press.

Fabre, Murielle. 2009. “Dyslexie en langue chinoise anjourd’hui.” Mémoire de Master 1 sous la direction de Franck Ramus et Dan Xu, Inalco.

Fan 范, Jiyan 继淹. 1984. “多项 NP 句 Duo xiang NP ju.” 中国语文 *Zhongguo Yuwen [Studies of the Chinese Language]* 1.

Fang 方, Mei 梅. 1994. “北京话句中语气词的功能研究 Beijingyu juzhong yuqici de gongneng yanjiu.” 中国语文 *Zhongguo Yuwen [Studies of the Chinese Language]* 2.

Fattal, Iris, Naama Friedmann, and Aviva Fattal-Valevski. 2011. "The crucial role of thiamine in the development of syntax and lexical retrieval: a study of infantile thiamine deficiency." *Brain* 134 (6): 1720–1739.

Fedorenko, Evelina, John Duncan, and Nancy Kanwisher. 2012. "Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area." *Current Biology* 22 (21): 2059–2062.

Fedorenko, Evelina, and Nancy Kanwisher. 2009. "Neuroimaging of Language: Why Hasn't a Clearer Picture Emerged?" *Language and Linguistics Compass* 3:839–865.

Fedorenko, Evelina, Alfonso Nieto-Castañon, and Nancy Kanwisher. 2012. "Lexical and syntactic representations in the brain: an fMRI investigation with multi-voxel pattern analyses." *Neuropsychologia* 50 (4): 499–513.

Fedorenko, Evelina, Alfonso Nieto-Castañón, and Nancy Kanwisher. 2012. "Syntactic processing in the human brain: what we know, what we don't know, and a suggestion for how to proceed." *Brain and Language* 120 (2): 187–207.

Fedorenko, Evelina, and Sharon L. Thompson-Schill. 2014. "Reworking the language network." *Trends in Cognitive Sciences* 18 (3): 120–126.

Felser, Claudia, Harald Clahsen, and Thomas F. Münte. 2003. "Storage and integration in the processing of filler-gap dependencies: an ERP study of topicalization and wh-movement in German." *Brain and Language* 87 (3): 345–354.

Felser, Claudia, Colin Phillips, and Matthew Wagers. 2017. "Editorial: Encoding and Navigating Linguistic Representations in Memory." *Frontiers in Psychology* 8.

Feng, Shiwen, Jennifer Legault, Long Yang, Junwei Zhu, Keqing Shao, and Yiming Yang. 2015. "Differences in grammatical processing strategies for active and passive sentences: An fMRI study." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:104–117.

Ferreira, Victor S., and Kathryn Bock. 2006. "The functions of structural priming." *Language and cognitive processes* 21 (7): 1011–1029.

Ferstl, Evelyn C., Jane Neumann, Carsten Bogler, and D. Yves von Cramon. 2008. "The extended language network: a meta-analysis of neuroimaging studies on text comprehension." *Human Brain Mapping* 29 (5): 581–593.

Fiebach, Christian J., Matthias Schlesewsky, Ina D. Bornkessel, and Angela D. Friederici. 2004. "Distinct neural correlates of legal and illegal word-order variations in German: How can fMRI inform cognitive models of sentence processing?" In *The Online Study of Sentence Comprehension: Eyetracking, ERPs and Beyond,* edited by Manuel Carreiras and Charles Clifton Jr., 357–370. New York: Psychology Press.

Fiebach, Christian J., Matthias Schlesewsky, and Angela D. Friederici. 2002. "Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions." *Journal of Memory and Language* 47 (2): 250–272.

Fiebach, Christian J., and Ricarda I. Schubotz. 2006. "Dynamic Anticipatory Processing of Hierarchical Sequential Events: a Common Role for Broca's Area and Ventral Premotor Cortex Across Domains?" *Cortex* 42 (4): 499–502.

Fillmore, Charles J. 1982. “Toward a descriptive framework for spatial deixis.” In *Speech, Place and Action: Studies in Deixis and Related Topics,* John Wiley, edited by Robert J. Jarvella and Wolfgang Klein, 31–59. London.

Finkbeiner, Rita, Jörg Meibauer, and Petra B. Schumacher, eds. 2012. *What is a Context? Linguistic approaches and challenges.* Vol. 196. Linguistik Aktuell/Linguistics Today. Amsterdam: John Benjamins, October 23.

Firbas, Jan. 1964. “On defining the theme in functional sentence analysis.” *Travaux Linguistique de Prague* 1:267–280.

———. 1992. *Functional Sentence Perspective in Written and Spoken Communication.* Cambridge University Press. Studies in English language. Cambridge.

Fitch, W. Tecumseh. 2014. “Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition.” *Physics of Life Reviews* 11 (3): 329–364.

Fitch, W. Tecumseh, and Marc D. Hauser. 2004. “Computational constraints on syntactic processing in a nonhuman primate.” *Science* 303 (5656): 377–380.

Fodor, Jerry Alan, Thomas G Bever, and Merrill F Garrett. 1974. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar.* New York: McGraw-Hill.

Ford, Marilyn, Joan Bresnan, and Ronald M. Kaplan. 1982. “A competence-based theory of syntactic closure.” In *The Mental Representation of Grammatical Relations,* edited by Joan Bresnan, 727–796. Cambridge, MA: MIT Press.

Fournier, Nathalie, and Catherine Fuchs. 2004. “Place du sujet nominal et opérations de thématisation.” In *La Linguistique cognitive,* edited by Catherine Fuchs, 55–88. Cogniprisme. Paris: Ophrys.

François, Jacques. 2004. “Le fonctionnalisme linguistique et les enjeux cognitifs.” In *La Linguistique cognitive,* edited by Catherine Fuchs, 99–134. Cogniprisme. Paris: Ophrys.

Frascarelli, Mara, and Roland Hinterhölzl. 2007. “Types of topics in German and Italian.” In *On Information Structure, Meaning and Form: Generalizations across languages,* edited by Kerstin Schwabe and Susanne Winkler, 87–116. Linguistik Aktuell/Linguistics Today 100. Amsterdam: John Benjamins Publishing Company.

Frazier, Lyn. 1988. “Sentence Processing: A Tutorial Review.” In *Attention and performance XII: the psychology of reading,* edited by M Coltheart, 559–586. Hove: Lawrence Erlbaum Associates.

Frazier, Lyn, Katy Carlson, and Charles Clifton. 2006. “Prosodic phrasing is central to language comprehension.” *Trends in Cognitive Sciences* 10 (6): 244–249.

Frey, Werner. 2004. “The grammar-pragmatics interface and the German prefield.” *Sprache une Pragmatik* 52:1–39.

———. n.d. “A medial topic position for German. Linguistische Berichte.” *Linguistische Berichte,* no. 198: 153–190.

Friederici, A. D., B. Opitz, and D. Y. von Cramon. 2000. "Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types." *Cerebral Cortex (New York, N.Y.: 1991)* 10 (7): 698–705.

Friederici, Angela D. 1995. "The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data." *Brain and Language* 50 (3): 259–281.

———. 2002. "Towards a neural basis of auditory sentence processing." *Trends in Cognitive Sciences* 6 (2): 78–84.

———. 2004. "Processing local transitions versus long-distance syntactic hierarchies." *Trends in Cognitive Sciences* 8 (6): 245–247.

Friederici, Angela D., Christian J. Fiebach, Matthias Schlesewsky, Ina D. Bornkessel, and D. Yves von Cramon. 2006. "Processing linguistic complexity and grammaticality in the left frontal cortex." *Cerebral Cortex* 16 (12): 1709–1717.

Friederici, Angela D., Anja Hahne, and Douglas Saddy. 2002. "Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair." *Journal of Psycholinguistic Research* 31 (1): 45–63.

Friederici, Angela D., Martin Meyer, and D. Yves von Cramon. 2000. "Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information." *Brain and Language* 74 (2): 289–300.

Friederici, Angela D., Shirley-Ann Rüschemeyer, Anja Hahne, and Christian J. Fiebach. 2003. "The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes." *Cerebral Cortex* 13 (2): 170–177.

Friederici, Angela D., Karsten Steinhauer, Axel Mecklinger, and Martin Meyer. 1998. "Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses." *Biological Psychology* 47 (3): 193–221.

Friederici, Angela D., and Jürgen Weissenborn. 2007. "Mapping sentence form onto meaning: the syntax-semantic interface." *Brain Research* 1146:50–58.

Friedmann, N. 2001. "Agrammatism and the psychological reality of the syntactic tree." *Journal of Psycholinguistic Research* 30 (1): 71–90.

Friedmann, Naama. 2002. "Question Production in Agrammatism: The Tree Pruning Hypothesis." *Brain and Language* 80 (2): 160–187.

———. 2005. "Degrees of severity and recovery in agrammatism: Climbing up the syntactic tree." *Aphasiology* 19 (10): 1037–1051.

———. 2006a. "Generalizations on variations in comprehension and production: a further source of variation and a possible account." *Brain and Language* 96 (2): 151–170.

———. 2006b. "Speech Production in Broca's Agrammatic Aphasia: Syntactic Tree Pruning." In *Broca's Region,* edited by Yosef Grodzinsky and Katrin Amunts, 63–82. Oxford University Press, May.

———. 2008. "Traceless relatives: Agrammatic comprehension of relative clauses with resumptive pronouns." *Journal of Neurolinguistics,* The Left Periphery of Sentences, 21 (2): 138–149.

———. 2013. "Verb Movement to C: From Agrammatic Aphasia to Syntactic Analysis." In *Diagnosing Syntax,* edited by Lisa Lai-Shen Cheng and Norbert Corver, 75–86. Oxford, New York: Oxford University Press, July 25.

Friedmann, Naama, and João Costa. 2011. "Last resort and no resort: resumptive pronouns in Hebrew and Palestinian Arabic hearing impairment." In *Language Faculty and Beyond,* edited by Alain Rouveret, 223–240. Amsterdam: John Benjamins.

Friedmann, Na'ama, and Yosef Grodzinsky. 1997. "Tense and Agreement in Agrammatic Production: Pruning the Syntactic Tree." *Brain and Language* 56 (3): 397–425.

Friedmann, Naama, and Rama Novogrodsky. 2007. "Is the movement deficit in syntactic SLI related to traces or to thematic role transfer?" *Brain and Language* 101 (1): 50–63.

———. 2008. "Subtypes of SLI: SySLI, PhoSLI, LeSLI, and PraSLI." In *Language acquisition and development: Proceedings of GALA 2007,* edited by Anna Gavarró and João Freitas, 205–217. Newcastle: Cambridge Scholars Press.

———. 2011. "Which questions are most difficult to understand?: The comprehension of Wh questions in three subtypes of SLI." *Lingua,* Specific Language Impairment, 121 (3): 367–382.

Friedmann, Naama, Rama Novogrodsky, Ronit Szterman, and Omer Preminger. 2008. "Resumptive pronouns as a last resort when movement is impaired: Relative clauses in hearing impairment." In *Linguistik Aktuell/Linguistics Today,* edited by Sharon Armon-Lotem, Gabi Danon, and Susan D. Rothstein, 134:267–290. Amsterdam: John Benjamins.

Friedmann, Naama, Dafna Olenik, and M Gil. 2000. "From theory to practice: Treatment of agrammatic production in Hebrew based on the Tree Pruning Hypothesis." *Journal of Neurolinguistics* 13:250–254.

Friedmann, Naama, and Dana Rusou. 2015. "Critical period for first language: the crucial role of language input during the first year of life." *Current Opinion in Neurobiology* 35:27–34.

Friedmann, Naama, Gina Taranto, Lewis P. Shapiro, and David Swinney. 2008. "The Leaf Fell (the Leaf): The Online Processing of Unaccusatives." *Linguistic inquiry (Online)* 39 (3): 355–377.

Friedmann, Naama, Maya Yachini, and Ronit Szterman. 2015. "Relatively easy relatives: Children with syntactic SLI avoid intervention." In *Structures, Strategies and Beyond: Studies in honour of Adriana Belletti,* edited by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 303–320. Linguistik Aktuell/Linguistics Today 223. Amsterdam: John Benjamins.

Frisch, Stefan, and Matthias Schlesewsky. 2001. "The N400 reflects problems of thematic hierarchizing." *Neuroreport* 12 (15): 3391–3394.

Frisch, Stefan, Matthias Schlesewsky, Douglas Saddy, and Annegret Alpermann. 2002. "The P600 as an indicator of syntactic ambiguity." *Cognition* 85 (3): B83–92.

Frith, Uta, and Chris Frith. 2003. "Development and Neurophysiology of mentalizing." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 358:459–73.

Fu 傅, Jingqi 京起, and Dan 丹 Xu 徐. 2009. "SVO 语言里的宾语置 SVO Yuyan li de bingyu qianzhi." 民族语文 *Mingzu Yuwen* 3:3–15.

Fu, Jingqi. 1994. "SVO Word Oder in Chinese and IP specifier." University of Southern California, Los Angeles.

Fuchs, Catherine, ed. 2004. *La Linguistique cognitive.* Cogniprisme. Paris: Ophrys.

Fuchs, Catherine, and Christiane Marchello-Nizia. 1998. "Les opérations de thématisation en Français." *Cahiers de praxématique* 30:113–131.

Fuller, Judith W., and Jeanette K. Gundel. 1987. "Topic-Prominence in Interlanguage." *Language Learning* 37 (1): 1–18.

Galmiche, Michel. 1992. "Au carrefour des malentendus : le thème." *L'Information Grammaticale* 54:3–10.

Garde, Paul. 1977. "Ordre linéaire et dépendance syntaxique : contribution à une typologie." *Bulletin de la Société de Linguistique de Paris* 72:1–26.

Gasde, Horst-Dieter, and Waltraud Paul. 1996. "Functional categories, topic prominence, and complex sentences in Mandarin Chinese." *Linguistics* 34 (2): 263–294.

Gibson, Edward. 1998. "Linguistic complexity: locality of syntactic dependencies." *Cognition* 68 (1): 1–76.

———. 2000. "The dependency locality theory: A distance-based theory of linguistic complexity." In *Image, language, brain,* edited by Alec Marantz, Yasushi Miyashita, and Wayne O'Neil, 95–126. Cambridge, MA: MIT Press.

Gibson, Edward, and Neal J. Pearlmutter. 1998. "Constraints on Sentence Comprehension." *Trends in Cognitive Sciences* 2 (7): 262–268.

———. 2000. "Distinguishing serial and parallel parsing." *Journal of Psycholinguistic Research* 29:231–240.

Giraud, Anne-Lise, and David Poeppel. 2012. "Cortical oscillations and speech processing: emerging computational principles and operations." *Nature Neuroscience* 15 (4): 511–517.

Givón, Talmy. 1976. "Topic, Pronoun, and Grammatical Agreement." In *Subject and Topic,* edited by Charles N. Li, 149–188. New York: Academic Press.

———. 1978. "Definiteness and referentiality." In *Universals of Human Language: Syntax,* edited by Joseph H. Greenberg. Stanford: Stanford University Press.

———. 1983a. *Topic Continuity in Discourse: A Quantitative Cross-language Study.* Amsterdam: John Benjamins.

———. 1983b. “Topic continuity in discourse: An introduction.” In *Topic Continuity in Discourse: A Quantitative Cross-language Study,* John Benjamins, 3–41. Amsterdam.

———. 1984. *Syntax: A functional-typological introduction. Volume I.* Vol. 1. Amsterdam: John Benjamins, January 1.

———. 1992. “On interpreting text-distributional correlations: some methodological issues.” In *Pragmatics of Word Order Flexibility,* edited by Doris Payne, 231–240. Amsterdam: John Benjamins Publishing Company.

Givón, Talmy, and Masayoshi Shibatani. 2009. *Syntactic Complexity: Diachrony, Acquisition, Neuro-cognition, Evolution.* John Benjamins Publishing.

Gold, Brian T. 2015. “Lifelong bilingualism and neural reserve against Alzheimer’s disease: A review of findings and potential mechanisms.” *Behavioural Brain Research* 281:9–15.

Goldrick, Matthew, Victor Ferreira, and Michele Miozzo. 2014. *The Oxford Handbook of Language Production.* Oxford University Press, April 11.

Golestani, Narly, F.-Xavier Alario, Sébastien Meriaux, Denis Le Bihan, Stanislas Dehaene, and Christophe Pallier. 2006. “Syntax production in bilinguals.” *Neuropsychologia* 44 (7): 1029–1040.

Gouvea, Ana C., Colin Phillips, Nina Kazanina, and David Poeppel. 2010. “The Linguistic Processes Underlying the P600.” *Language and Cognitive Processes* 25 (2): 149–188.

Grady, Cheryl L., Gigi Luk, Fergus I. M. Craik, and Ellen Bialystok. 2015. “Brain network activity in monolingual and bilingual older adults.” *Neuropsychologia* 66:170–181.

Grant, Angela M., Shin-Yi Fang, and Ping Li. 2015. “Second language lexical development and cognitive control: A longitudinal fMRI study.” *Brain and Language* 144:35–47.

Greenberg, Joseph H. 1963. “Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements.” In *Universals of Human Language,* edited by Joseph H. Greenberg, 73–113. Cambridge, Mass: MIT Press.

Grewe, Tanja, Ina Bornkessel, Stefan Zysset, Richard Wiese, D. Yves von Cramon, and Matthias Schlesewsky. 2005. “The emergence of the unmarked: A new perspective on the language-specific function of Broca’s area.” *Human Brain Mapping* 26 (3): 178–190.

———. 2006. “Linguistic prominence and Broca’s area: the influence of animacy as a linearization principle.” *NeuroImage* 32 (3): 1395–1402.

Grewe, Tanja, Ina Bornkessel-Schlesewsky, Stefan Zysset, Richard Wiese, D. Yves von Cramon, and Matthias Schlesewsky. 2007. “The role of the posterior superior temporal sulcus in the processing of unmarked transitivity.” *NeuroImage* 35 (1): 343–352.

Grillo, Nino. 2008. “Generalized Minimality: Syntactic underspecification in Broca’s aphasia.” PhD diss., Utrecht University.

———. 2009. “Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism.” *Lingua,* Experimental Evidence for minimal structure, 119 (10): 1426–1443.

Grill-Spector, Kalanit, Richard Henson, and Alex Martin. 2006. “Repetition and the brain: neural models of stimulus-specific effects.” *Trends in Cognitive Sciences* 10 (1): 14–23.

Grill-Spector, Kalanit, and Rafael Malach. 2001. “fMR-adaptation: a tool for studying the functional properties of human cortical neurons.” *Acta Psychologica,* Beyond the decade of the brain: Towards functional neuronanatomy of the mind, 107 (1): 293–321.

Grimshaw, Jane. 1990. *Argument Structure.* Linguistic inquiry monographs. Cambridge, MA: MIT Press.

Grodner, Daniel, Edward Gibson, and Duane Watson. 2005. “The influence of contextual contrast on syntactic processing: evidence for strong-interaction in sentence comprehension.” *Cognition* 95 (3): 275–296.

Grodzinsky, Yosef. 2000. “The neurology of syntax: Language use without Broca’s area.” *The Behavioral and Brain Sciences* 23 (1): 1–21, discussion 21–71.

Grodzinsky, Yosef, and Katrin Amunts. 2006. *Broca’s Region.* Oxford, New York: Oxford University Press, May 25.

Grodzinsky, Yosef, and Lisa Finkel. 1998. “The neurology of empty categories: aphasics’ failure to detect ungrammaticality.” *Journal of Cognitive Neuroscience* 10 (2): 281–292.

Grodzinsky, Yosef, and Angela D. Friederici. 2006. “Neuroimaging of syntax and syntactic processing.” *Current Opinion in Neurobiology* 16 (2): 240–246.

Grodzinsky, Yosef, and Andrea Santi. 2008. “The battle for Broca’s region.” *Trends in Cognitive Sciences* 12 (12): 474–480.

Grossman, Murray, Ayanna Cooke, Chris DeVita, David Alsop, John Detre, Willis Chen, and James Gee. 2002. “Age-related changes in working memory during sentence comprehension: an fMRI study.” *NeuroImage* 15 (2): 302–317.

Grossman, Murray, Phyllis Koenig, Chris DeVita, Guila Glosser, David Alsop, John Detre, and James Gee. 2002. “The neural basis for category-specific knowledge: an fMRI study.” *NeuroImage* 15 (4): 936–948.

Gruber, Jeffrey S. 1967. “Topicalization in Child Language.” *Foundations of Language* 3 (1): 37–65.

Guimier, Claude. 1996. *Les adverbes du français. Le cas des adverbes en –ment.* L’essentiel. Paris: Ophrys.

———, ed. 1999. *La thématisation dans les langues, actes du colloque de Caen (9-11 octobre 1997).* Bern: Peter Lang.

Gundel, Jeanette K. 1974. “Role of topic and comment in linguistic theory.” PhD diss., University of Texas.

———. 1977. *Role of topic and comment in linguistic theory.* Bloomington: Indiana University Linguistics Club.

———. 1985. “‘Shared knowledge’ and topicality.” *Journal of Pragmatics* 9 (1): 83–107.

———. 1988. “Universals of topic-comment structure.” In *Studies in Synctatic Typology,* edited by Michael Hammond, Edith A. Moravcsik, and Jessica Wirth, 209–293. Amsterdam: John Benjamins Publishing Company.

———. 1998. “On different kinds of focus.” In *Focus in Natural Language Processing,* edited by Peter Bosch and Robert van der Sandt. Cambridge: Cambridge University Press.

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 1993. “Cognitive Status and the Form of Referring Expressions in Discourse.” *Language* 69 (2): 274–307.

Gunter, Thomas C., and Angela D. Friederici. 1999. “Concerning the automaticity of syntactic processing.” *Psychophysiology* 36:126–137.

Gunter, Thomas C., Angela D. Friederici, and Herbert Schriefers. 2000. “Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction.” *Journal of Cognitive Neuroscience* 12 (4): 556–568.

Guo, Yuqing, Haifeng Wang, and Josef van Genabith. 2007. “Recovering Non-Local Dependencies for Chinese.” In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning,* 257–266. Prague, Czech Republic.

Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation.* Cambridge University Press.

Gutiérrez-Bravo, Rodrigo, Line Mikkelsen, and Eric Potsdam. 2011. “Representing Language: Essays in Honor of Judith Aissen.” *Linguistics Research Center,* no. 0.

Hackl, Martin. 2013. “The syntax–semantics interface.” *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:66–87.

Haegeman, Liliane. 1997a. “Elements of Grammar.” In *Elements of Grammar: Handbook in Generative Syntax,* 1–71. Dordrecht: Kluwer Academic.

———. 1997b. *Elements of Grammar: Handbook in Generative Syntax.* Edited by Liliane Haegeman. Dordrecht: Kluwer Academic, July 31.

———. 2003. “Notes on Long Adverbial Fronting in English and the Left Periphery.” *Linguistic Inquiry* 34 (4): 640–649.

———. 2013. “The syntax of registers: Diary subject omission and the privilege of the root.” *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:88–110.

Hagège, Claude. 1974. “Les pronoms logophoriques.” *Bulletin de la Société de Linguistique de Paris* 1 (69): 287–310.

Hagiwara, H. 1995. “The Breakdown of Functional Categories and the Economy of Derivation.” *Brain and Language* 50 (1): 92–116.

Hagoort, Peter. 2003a. “How the brain solves the binding problem for language: a neurocomputational model of syntactic processing.” *NeuroImage* 20 (Suppl 1): S18–29.

———. 2003b. “Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations.” *Journal of Cognitive Neuroscience* 15 (6): 883–899.

———. 2005. “On Broca, brain, and binding: a new framework.” *Trends in Cognitive Sciences* 9 (9): 416–423.

Hagoort, Peter, and Colin M. Brown. 1994. “Brain responses to lexical ambiguity resolution and parsing.” In *Perspectives on Sentence Processing,* edited by Charles Clifton Jr., Lyn Frazier, and Keith Rayner, 45–80. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hagoort, Peter, Lea Hald, Marcel Bastiaansen, and Karl Magnus Petersson. 2004. “Integration of word meaning and world knowledge in language comprehension.” *Science* 304 (5669): 438–441.

Hahne, Anja, and Angela D. Friederici. 1999. “Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes.” *Journal of Cognitive Neuroscience* 11 (2): 194–205.

Hakulinen, Soili, and Meri Larjavaara. 2012. “Circonstants initiaux : perspectives en français et en finnois.” Oral Communication, INALCO, Paris, September.

Hale, John T. 2014. *Automaton Theories of Human Sentence Comprehension.* Studies in Computational Linguistics. Chicago: CSLI Publications/Center for the Study of Language & Information.

Halliday, Michael A. K. 1967a. *Intonation and grammar in British English.* The Hague, Paris: Mouton.

———. 1967b. “Notes on transitivity and theme in English: Part 2.” *Journal of Linguistics* 3 (2): 199–244.

Halliday, Michael A. K., and Christian M. I. M. Matthiessen. 1985. *An introduction to functional grammar.* London: Edward Arnold.

Hamann, Cornelia, Zvi Penner, and Katrin Lindner. 1998. “German Impaired Grammar: The Clause Structure Revisited.” *Language Acquisition* 7 (2): 193–245.

Hammer, Anke, Bernadette M. Jansma, Claus Tempelmann, and Thomas F. Münte. 2011. “Neural Mechanisms of Anaphoric Reference Revealed by fMRI.” *Frontiers in Psychology* 2.

Hammond, Michael, Edith A. Moravcsik, and Jessica Wirth, eds. 1988. *Studies in Syntactic Typology.* Typological Studies in Language 17. Amsterdam: John Benjamins Publishing Company, January 1.

Hanakawa, Takashi, Manabu Honda, Nobukatsu Sawamoto, Tomohisa Okada, Yoshiharu Yonekura, Hidena Fukuyama, and Hiroshi Shibasaki. 2002. "The role of rostral Brodmann area 6 in mental-operation tasks: an integrative neuroimaging approach." *Cerebral Cortex (New York, N.Y.: 1991)* 12 (11): 1157–1170.

Hanson, Stephen José, Toshihiko Matsuka, and James V. Haxby. 2004. "Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area?" *NeuroImage* 23 (1): 156–166.

Hara, Yurie. 2006. "Japanese discourse items at interfaces." PhD diss., University of Delaware.

Hasegawa, Nobuko. 1981. "A lexical interpretive theory with emphasis on the role of subject." PhD diss., University of Washigton.

Haupt, Friederike S., Matthias Schlesewsky, Dietmar Roehm, Angela D. Friederici, and Ina Bornkessel-Schlesewsky. 2008. "The status of subject–object reanalyses in the language comprehension architecture." *Journal of Memory and Language* 59 (1): 54–96.

Hauser, Marc D., Noam Chomsky, and W. Tecumseh Fitch. 2002. "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science* 298 (5598): 1569–1579.

Hauser, Marc D., and Jeffrey Watumull. 2017. "The Universal Generative Faculty: The source of our expressive power in language, mathematics, morality, and music." *Journal of Neurolinguistics,* Language Evolution: On the Origin of Lexical and Syntactic Structures, 43 (Part B): 78–94.

Hawkins, John A. 1995. *A Performance Theory of Order and Constituency.* Cambridge: Cambridge University Press.

———. 2002. "Symmetries and asymmetries: their grammar, typology and parsing." *Theoretical Linguistics* 28 (2): 95–150.

Haxby, James V., M. Ida Gobbini, Maura L. Furey, Alumit Ishai, Jennifer L. Schouten, and Pietro Pietrini. 2001. "Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex." *Science* 293 (5539): 2425–2430.

Head, Lauren M., Melissa Baralt, and Ashley E. Darcy Mahoney. 2015. "Bilingualism as a Potential Strategy to Improve Executive Function in Preterm Infants: A Review." *Journal of Pediatric Health Care* 29 (2): 126–136.

Heeschen, Claus, and Emanuel A. Schegloff. 2003. "Aphasic Agrammatism as Interactional Artifact and Achievement." In *Conversation and Brain Damage,* edited by Charles Goodwin, 231–282. New York: Oxford University Press.

Heim, Irene. 1982. "The Semantics of Definite and Indefinite Noun Phrases." PhD diss., UMass Amherst.

Heim, Stefan, Kai Alter, Anja K. Ischebeck, Katrin Amunts, Simon B. Eickhoff, Hartmut Mohlberg, Karl Zilles, D. Yves von Cramon, and Angela D. Friederici. 2005. "The role of the left Brodmann's areas 44 and 45 in reading words and pseudowords." *Cognitive Brain Research* 25 (3): 982–993.

Hemforth, Barbara. 1993. *Kognitives Parsing: Repräsentation und Verarbeitung sprachlichen Wissens.*

Henry, Paul. 1977. *Le mauvais outil : langue, sujet et discours.* Paris: Klincksieck.

Her 何, One-Soon 万顺. 2003. "Chinese inversion constructions within a simplified LMT [汉语句法倒置结构在简化词汇映照理论中之解析]." In *Lexical-functional grammar analysis of Chinese,* edited by Adams B. Bodomo and Kang Kwong Luke, 1–31. Journal of Chinese Linguistics Monograph Series 19.

Her, One-Soon. 1991. "Topic as a grammatical function in Chinese." *Lingua,* no. 1: 1–23.

———. 2007. "Argument-function mismatches in Mandarin resultatives: A lexical mapping account." *Lingua* 117 (1): 221–246.

———. 2009. "Apparent subject-object inversion in Chinese." *Linguistics* 47 (5): 1143–1181.

Hernández, Mireia, Clara D. Martin, Francisco Barceló, and Albert Costa. 2013. "Where is the bilingual advantage in task-switching?" *Journal of Memory and Language* 69 (3): 257–276.

Hestvik, Arild, Nathan Maxfield, Richard G. Schwartz, and Valerie Shafer. 2007. "Brain responses to filled gaps." *Brain and language* 100 (3): 301–316.

Heycock, Caroline. 1993. "Syntactic predication in Japanese." *Journal of East Asian Linguistics* 2 (2): 167–211.

Heycock, Caroline, and Anthony Kroch. 2002. "Topic, Focus, and Syntactic Representations." In *WCCFL 21: Proceedings of the 21st West Coast Conference on Formal Linguistics,* edited by Line Mikkelsen and Christopher Potts, 101–125. Somerville, MA: Cascadilla Press.

Hickmann, Maya. 2003. *Children's Discourse: Person, Space and Time across Languages.* Cambridge Studies in Linguistics 98. Cambridge: Cambridge University Press.

Hickmann, Maya, and Henriëtte Hendriks. 1999. "Cohesion and anaphora in children's narratives: a comparison of English, French German, and Mandarin Chinese." *Journal of Child Language* 26 (2): 419–452.

Hickmann, Maya, and Stéphane Robert, eds. 2006. *Space in Languages: Linguistic Systems and Cognitive Categories.* Vol. 66. Typological Studies in Language. Amsterdam: John Benjamins.

Hickok, Gregory, and Steven L. Small, eds. 2016. *Neurobiology of Language.* San Diego: Academic Press.

Hillis, Argye E., Brenda Rapp, Cristina Romani, and Alfonso Caramazza. 1990. "Selective impairment of semantics in lexical processing." *Cognitive Neuropsychology* 7 (3): 191–243.

Hinterwimmer, Stefan, and Sophie Repp. 2010. "Topics(?) in the left periphery." Potsdam, July 10.

Hirotani, Masako. 2005. “Constraints on prosodic structures in grammar and parser: Scrambled and unscrambled sentences in Japanese.” In *Papers on prosody,* edited by Shigeto Kawahara, 53–90. Amherst, MA: GLSA, Dept. of Linguistics, University of Massachusetts at Amherst.

Hirotani, Masako, Michiru Makuuchi, Shirley-Ann Rüschemeyer, and Angela D. Friederici. 2011. “Who was the agent? The neural correlates of reanalysis processes during sentence comprehension.” *Human Brain Mapping* 32 (11): 1775–1787.

Hirotani, Masako, and Petra B. Schumacher. 2011. “Context and topic marking affect distinct processes during discourse comprehension in Japanese.” *Journal of Neurolinguistics* 24 (3): 276–292.

Hirshorn, Elizabeth A., Matthew W. G. Dye, Peter C. Hauser, Ted R. Supalla, and Daphne Bavelier. 2014. “Neural networks mediating sentence reading in the deaf.” *Frontiers in Human Neuroscience* 8:394.

Hirst, Daniel. 1983a. “Interpreting Intonation: A Modular Approach.” *Journal of Semantics* 2 (2): 171–182.

———. 1983b. “Structures and Categories in Prosodic Representations.” In *Prosody: Models and Measurements,* edited by Anne Cutler and D. Robert Ladd, 93–109. Springer Series in Language and Communication. Berlin: Springer.

Hirst, Daniel, Albert Di Cristo, and Robert Espesser. 2000. “Levels of Representation and Levels of Analysis for the Description of Intonation Systems.” In *Prosody: Theory and Experiment - Studies Presented to Gösta Bruce,* edited by Merle Horne, 51–87. Text, Speech and Language Technology. Dordrecht: Kluwer Academic.

Ho, Connie Suk-Han, and Peter Bryant. 1997. “Learning to Read Chinese beyond the Logographic Phase.” *Reading Research Quarterly* 32 (3): 276–289.

Ho, Connie Suk-Han, and Peter E. Bryant. 1997a. “Development of phonological awareness of Chinese children in Hong Kong.” *Journal of Psycholinguistic Research* 26 (1): 109–126.

———. 1997b. “Phonological skills are important in learning to read Chinese.” *Developmental Psychology* 33 (6): 946–951.

Ho, Connie Suk-Han, David Wai-Ock Chan, Kevin K.H. Chung, Suk-Han Lee, and Suk-Man Tsang. 2007. “In search of subtypes of Chinese developmental dyslexia.” *Journal of Experimental Child Psychology* 97 (1): 61–83.

Ho, Connie Suk-Han, David Wai-Ock Chan, Suk-Han Lee, Suk-Man Tsang, and Vivian Hui Luan. 2004. “Cognitive profiling and preliminary subtyping in Chinese developmental dyslexia.” *Cognition* 91 (1): 43–75.

Ho, Connie Suk-Han, Teresa Pui-Sze Law, and Penny Man Ng. 2000. “The phonological deficit hypothesis in Chinese developmental dyslexia.” *Reading and Writing* 13 (1): 57–79.

Hockett, Charles Francis. 1958. *A course in modern linguistics.* New York: Macmillan.

Holmberg, Anders, and Ian Roberts. 2013. “The syntax–morphology relation.” *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:111–131.

Holzgrefe, Julia, Caroline Wellmann, Caterina Petrone, Hubert Truckenbrodt, Barbara Höhle, and Isabell Wartenburger. 2013. "Brain response to prosodic boundary cues depends on boundary position." *Frontiers in Psychology* 4.

Hovdhaugen, Even. 1990. "Una et Eadem: Some observations on Roger Bacon's Greek grammar." In *De Ortu Grammaticae: Studies in medieval grammar and linguistic theory in memory of Jan Pinborg,* edited by Geoffrey L. Bursill-Hall, Sten Ebbesen, and E.F.K. Koerner, 43:117. Amsterdam: John Benjamins.

Hsiao, Franny, and Edward Gibson. 2003. "Processing relative clauses in Chinese." *Cognition* 90 (1): 3–27.

Hsieh, Fu-Tsai. 2017. "Shuo 'say' vs. Yiwei 'assume' in Chinese." In *Language at the University of Essex (LangUE) 2008 Proceedings,* edited by Kholoud Al-Thubaiti, C Batterham, Kakia Chatsiou, Lena Papadopoulou, and Keisuke Yoshimoto, 67–76. Department of Language and Linguistics, University of Essex.

Hsu, Chun-chieh Natalie, Felicia Hurewitz, and Colin Phillips. 2009. "Context influences structure generation: Evidence from Chinese." *Submitted.*

Hsu, Chun-Hsien, Chia-Ying Lee, and Ovid J. L. Tzeng. 2014. "Early MEG markers for reading Chinese phonograms: Evidence from radical combinability and consistency effects." *Brain and Language* 139:1–9.

Hsu, Yu-Yin. 2008. "The Sentence-Internal Topic and Focus in Chinese." In *Proceedings of the 20th North American Conference on Chinese Linguistics,* 635–652. Colombus, Ohio: The Ohio State University.

Hu, Chieh-Fang, and Hugh W. Catts. 1998. "The Role of Phonological Processing in Early Reading Ability: What We Can Learn From Chinese." *Scientific Studies of Reading* 2 (1): 55–79.

Hu, Jianhua, and Haihua Pan. 2008. "Mandarin Intransitive Verbs and Their Objects at the Syntax-Information Structure Interface1." In *Proceedings of the 20th North American Conference on Chinese Linguistics,* 653–663. Colombus, Ohio.

Hu, Shenai. 2014. "Intervention effects and the acquisition of relativization and topicalization in Chinese." PhD diss., Universitat Autònoma de Barcelona.

Huang 黄, C-T James 正德. 2007. "[汉语动词的题元结构与其句法表现] Thematic structures of verbs in Chinese and their syntactic projections." *[语言科学] Linguistic Sciences* 4:3–21.

Huang, Cheng-Teh James. 1982. "Logical Relation in Chinese and the Theory of Grammar." PhD diss., MIT.

———. 1984. "On the distribution and reference of empty pronouns." *Linguistic Inquiry* 15 (4): 531–574.

———. 2005. *On syntactic analyticity and the other end of the parameter. Lecture notes from LSA 2005.* Harvard University.

———. 2014. "Syntactic analyticity and parametric theory." In *Chinese Syntax in a Cross-Linguistic Perspective,* edited by Audrey Li, Andrew Simpson, and Wei-Tien Dylan Tsai, 446. Oxford: Oxford University Press.

Huang, Cheng-Teh James, and Yen-Hui Audrey Li. 1996. "Recent Generative Studies in Chinese Syntax." In *New Horizons in Chinese Linguistics,* edited by Cheng-Teh James Huang and Yen-hui Audrey Li, 49–95. Dordrecht: Springer.

Huang, Cheng-Teh James, Yen-Hui Audrey Li, and Yafei Li. 2009. *The Syntax of Chinese.* Cambridge syntax guides. New York: Cambridge University Press.

Huang, Cheng-Teh James, and C.-S. Luther Liu. 2000. "Logophoricitv, Attitudes, And Ziji At The Interface." In *Long Distance Reflexives,* 141–195. Syntax and semantics 33. London: BRILL, October 17.

Huang, Yu-Chi, and Elsi Kaiser. 2008. "Investigating filler-gap dependencies in Chinese topicalization." In *Proceedings of the 20th North American Conference on Chinese Linguistics,* 927–941. Colombus, Ohio: The Ohio State University.

Huang, Chu-Ren, Keh-Jiann Chen, and Keh-Jiann T'sou, eds. 1996. *Readings in Chinese natural language processing.* Journal of Chinese linguistics. Monograph series 9. Berkeley: Project on Linguistic Analysis, University of California.

Huang, H.S., and J.Richard Hanley. 1995. "Phonological awareness and visual skills in learning to read Chinese and English." *Cognition* 54 (1): 73–98.

Huang, Hui-ting, and One-Soon Her. 1998. "Mandarin locative inversion and relation-changing rules." In *Selected papers from the second international symposium on languages in Taiwan,* 287–304. Tapei: Crane Publishing.

Huang, Ray Rui-heng, and Jen Ting. 2006. "Are There Dangling Topics in Mandarin Chinese?" *Concentric: Studies in Linguistics* 32 (1): 119–146.

Huddleston, Rodney. 1984. *Introduction to the Grammar of English.* Cambridge Textbooks in Linguistics. Cambridge University Press.

Huddleston, Rodney, and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press.

Huemer, Wolfgang. 2015. "Franz Brentano." In *The Stanford Encyclopedia of Philosophy,* Fall 2015, edited by Edward N. Zalta.

Hulk, Aafke, and Jean-Yves Pollock. 2001. *Subject Inversion in Romance and the Theory of Universal Grammar.* Oxford Studies in Comparative Syntax. Oxford, New York: Oxford University Press.

Hulst, Yves d', Johan Rooryck, and Jan Schroten. 2001. *Romance Languages and Linguistic Theory 1999: Selected Papers from 'Going Romance' 1999, Leiden, 9-11 December.* John Benjamins Publishing.

Humphries, Colin, Jeffrey R. Binder, David A. Medler, and Einat Liebenthal. 2006. "Syntactic and semantic modulation of neural activity during auditory sentence comprehension." *Journal of Cognitive Neuroscience* 18 (4): 665–679.

Humphries, Colin, Tracy Love, David Swinney, and Gregory Hickok. 2005. "Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing." *Human Brain Mapping* 26 (2): 128–138.

Hung, Yu-Chen, and Petra B. Schumacher. 2012. “Topicality matters: position-specific demands on Chinese discourse processing.” *Neuroscience Letters* 511 (2): 59–64.

———. 2014. “Animacy matters: ERP evidence for the multi-dimensionality of topic-worthiness in Chinese.” *Brain Research* 1555 (Supplement C): 36–47.

Inui, T., Y. Otsu, S. Tanaka, T. Okada, S. Nishizawa, and J. Konishi. 1998. “A functional MRI analysis of comprehension processes of Japanese sentences.” *Neuroreport* 9 (14): 3325–3328.

Inui, Toshio, Yukio Otsu, Shigeki Tanaka, Tomohisa Okada, Sadahiko Nishizawa, and Jyunji Konishi. 1998. “A functional MRI analysis of comprehension processes of Japanese sentences. [Miscellaneous Article].” *Neuroreport* 9 (14): 3325–3328.

Isel, Frédéric, Anja Hahne, Burkhard Maess, and Angela D. Friederici. 2007. “Neurodynamics of sentence interpretation: ERP evidence from French.” *Biological Psychology* 74 (3): 337–346.

Jackendoff, Ray. 2003. *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford, New York: Oxford University Press, September 4.

Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. Current Studies in Linguistics. Cambridge, MA: MIT Press.

———. 1996. “The architecture of the linguistic-spatial interface.” In *Language and space,* edited by Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, 1–30. Language, speech, and communication. Cambridge, MA: MIT Press.

———. 2002. “The Mind Doesn’t Work That Way: The Scope and Limits of Computational Psychology (review).” *Language* 78:164–170.

———. 2007. *Language, Consciousness, Culture: Essays on Mental Structure.* Cambridge, MA: MIT Press.

Jacobs, Beverly J., and Cynthia K. Thompson. 2000. “Cross-modal generalization effects of training noncanonical sentence comprehension and production in agrammatic aphasia.” *Journal of speech, language, and hearing research: JSLHR* 43 (1): 5–20.

Jacobs, Joachim. 1984. “Funktionale Satzperspektive und Illokutionssemantik.” *Linguistische Berichte* 91:25–58.

———. 1997. “I-Topikalisierung.” *Linguistische Berichte* 168:91–133.

———. 2001. “The dimensions of topic-comment.” *Linguistics* 39 (4): 641–681.

Jacobs, Joachim, Arnim von Stechow, Wolfgang Sternefeld, Theo Vennemann, Noam Chomsky, and Howard Lasnik, eds. 1993. “Principles and Parameters Theory.” In *Syntax: an international hanbook of contemporary research,* 506–559. Berlin; New York: W. de Gruyter.

Jaeggli, Osvaldo, Carmen Silva-Corvalán, and Eduardo Raposo, eds. 1986. “On the null object in European Portuguese.” In *Studies in Romance linguistics,* 373–390. Dordrecht: Foris, December.

Jakubovitz, Celia, and Laurice Tuller. 2008. “Specific language impairment in French.” In *Language Learning & Language Teaching,* edited by Dalila Ayoun, 21:97–133. Amsterdam: John Benjamins Publishing Company.

Jiang, Xiaoming, and Xiaolin Zhou. 2009. “Processing different levels of syntactic hierarchy: an ERP study on Chinese.” *Neuropsychologia* 47 (5): 1282–1293.

Jiang, Zixin. 1990. “A constraint on topic in Chinese.” *Journal of Chinese Linguistics* 18:231–260.

———. 1991. “Some Aspects of the Syntax of Topic and Subject in Chinese.” PhD diss., University of Chicago, Department of Linguistics.

Jobard, Gaël, Fabrice Crivello, and Nathalie Tzourio-Mazoyer. 2003. “Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies.” *NeuroImage* 20 (2): 693–712.

Jonare, Birgitta. 1976. *L’Inversion Dans La Principale Non-Interrogative En Français Contemporain.* Stockholm: Almqvist och Wiksell.

Jonides, John, Eric H. Schumacher, Edward E. Smith, Robert A. Koeppe, Edward Awh, Patricia A. Reuter-Lorenz, Christy Marshuetz, and Christopher R. Willis. 1998. “The Role of Parietal Cortex in Verbal Working Memory.” *Journal of Neuroscience* 18 (13): 5026–5034.

Jonides, John, Edward E. Smith, Robert A. Koeppe, Edward Awh, Satoshi Minoshima, and Mark A. Mintun. 1993. “Spatial working memory in humans as revealed by PET.” *Nature* 363 (6430): 623–625.

Jung, Euen Hyuk (Sarah). 2004. “Topic and Subject Prominence in Interlanguage Development.” *Language Learning* 54 (4): 713–738.

Just, Marcel Adam, Patricia A. Carpenter, Timothy A. Keller, William F. Eddy, and Keith R. Thulborn. 1996. “Brain activation modulated by sentence comprehension.” *Science* 274 (5284): 114–116.

Kaan, Edith, Andrea C. Dallas, and Christopher M. Barkley. 2007. “Processing bare quantifiers in discourse.” *Brain Research* 1146:199–209.

Kaan, Edith, Anthony Harris, Edward Gibson, and Phillip Holcomb. 2000. “The P600 as an index of syntactic integration difficulty.” *Language and Cognitive Processes* 15 (2): 159–201.

Kaan, Edith, and Tamara Y. Swaab. 2002. “The brain circuitry of syntactic comprehension.” *Trends in Cognitive Sciences* 6 (8): 350–356.

Kalbertodt, Janina, Beatrice Primus, and Petra B. Schumacher. 2015. “Punctuation, Prosody, and Discourse: Afterthought Vs. Right Dislocation.” *Frontiers in Psychology* 6:1803.

Kambara, Toshimune, Takashi Tsukiura, Satoru Yokoyama, Kei Takahashi, Yayoi Shigemune, Tadao Miyamoto, Daiko Takahashi, Shigeru Sato, and Ryuta Kawashima. 2013. “Differential contributions of the inferior parietal and inferior frontal regions to the processing of grammatical and semantic relationships in wh-questions.” *Language Sciences* 37:14–21.

Kamide, Yuki, Gerry T. M. Altmann, and Sarah L. Haywood. 2003. "The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements." *Journal of Memory and Language* 49 (1): 133–156.

Kampen, Jacqueline van. 2005. "Language specific bootstraps for UG categories." *International Journal of Bilingualism* 9 (2): 253–275.

Kayne, Richard S. 2004. "Here and There." In *Lexique, Syntaxe et Lexique-Grammaire / Syntax, Lexis & Lexicon-Grammar: Papers in honour of Maurice Gross,* edited by Christian Leclère, Éric Laporte, Mireille Piot, and Max Silberztein, 253–273. Amsterdam: John Benjamins Publishing Company.

———. 2013. "Comparative syntax." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:132–151.

Kayne, Richard S., and Jean-Yves Pollock. 2001. "New thoughts on stylistic inversion." In *Subject Inversion in Romance and the Theory of Universal Grammar,* 107–162. Oxford Studies in Comparative Syntax. Oxford, New York: Oxford University Press.

Keenan, Edward L. 1976. "Remarkable Subjects in Malagasy." In *Subject and Topic,* edited by Charles N. Li, 247–301. New York: Academic Press.

Keller, Timothy A., Patricia A. Carpenter, and Marcel Adam Just. 2001. "The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing." *Cerebral Cortex* 11 (3): 223–237.

Kemmerer, David. 2005. "The spatial and temporal meanings of English prepositions can be independently impaired." *Neuropsychologia* 43 (5): 797–806.

Kenstowicz, Michael J., and Noam Chomsky, eds. 2001. "Derivation by phase." In *Ken Hale: a life in language,* 1–52. Cambridge, Mass.: MIT Press.

Kerkhofs, Roel, Wietske Vonk, Herbert Schriefers, and Dorothee J. Chwilla. 2007. "Discourse, syntax, and prosody: the brain reveals an immediate interaction." *Journal of Cognitive Neuroscience* 19 (9): 1421–1434.

Kimmelman, Vadim. 2012a. "Doubling in RSL and NGT: Towards a unified explanation." In *Proceedings of IATL 27,* edited by Evan Cohen, 57–81. Cambridge, MA: MITWPL.

———. 2012b. "Information structure categories: evidence from sign languages." Max Planck Institute for Psycholinguistics, Nijmegen, November 10.

King, Jonathan W., and Marta Kutas. 1995. "Who did what and when? Using word- and clause-jevel ERPs to monitor working memory usage in reading." *Journal of Cognitive Neuroscience* 7 (3): 376–395.

Kinno, Ryuta, Mitsuru Kawamura, Seiji Shioda, and Kuniyoshi L. Sakai. 2008. "Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task." *Human Brain Mapping* 29 (9): 1015–1027.

Kiss, Katalin É. 1995. *Discourse Configurational Languages.* New York: Oxford University Press.

Kliegl, Reinhold, Antje Nuthmann, and Ralf Engbert. 2006. "Tracking the mind during reading: the influence of past, present, and future words on fixation durations." *Journal of Experimental Psychology: General* 135 (1): 12–35.

Kolk, Herman, and Dorothee Chwilla. 2007. "Late positivities in unusual situations." *Brain and Language* 100 (3): 257–261.

Koopman, Hilda, and Dominique Sportiche. 1991. "The position of subjects." *Lingua* 85 (2): 211–258.

Koring, Loes, Pim Mak, and Eric Reuland. 2012. "The time course of argument reactivation revealed: Using the visual world paradigm." *Cognition* 123 (3): 361–379.

Kotchoubey, Boris. 2002. "Do event-related brain potentials reflect mental (cognitive) operations?" *Journal of Psychophysiology* 16 (3): 129–149.

Kranjec, Alexander, Eileen R. Cardillo, Gwenda L. Schmidt, and Anjan Chatterjee. 2010. "Prescribed spatial prepositions influence how we think about time." *Cognition* 114 (1): 111–116.

Kreiner, Hamutal, and Tamar Degani. 2015. "Tip-of-the-tongue in a second language: The effects of brief first-language exposure and long-term use." *Cognition* 137:106–114.

Krifka, Manfred. 2008. "Functional similarities between bimanual coordination and topic/comment structure." In *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS),* edited by Shinichiro Ishihara, Stefanie Jannedy, and Anne Schwarz, 61–96. Potsdam: Universitätsverlag Potsdam.

Krishnan, Ananthanarayan, Jackson T. Gandour, Saradha Ananthakrishnan, and Venkatakrishnan Vijayaraghavan. 2015. "Language experience enhances early cortical pitch-dependent responses." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:128–148.

Kristensen, Line Burholt, Elisabeth Engberg-Pedersen, and Mads Poulsen. 2014. "Context improves comprehension of fronted objects." *Journal of Psycholinguistic Research* 43 (2): 125–140.

Kuno, Susumo. 1972. "Functional sentence perspective: A case study from Japanese and English." *Linguistic Inquiry* 3 (3): 269–320.

———. 1976. "Subject raising." In *Japanese Generative Grammar: Syntax and Semantics, vol. 5,* edited by Masayoshi Shibatani, 17–49. Academic Press.

Kuno, Susumo, and Etsuko Kaburaki. 1977. "Empathy and Syntax." *Linguistic Inquiry* 8 (4): 627–672.

Kuno, Susumu. 1973. *The Structure of the Japanese Language.* Cambridge, MA: MIT Press.

Kuo, Wen-Jui, Tzu-Chen Yeh, Jun-Ren Lee, Li-Fen Chen, Po-Lei Lee, Shyan-Shiou Chen, Low-Tone Ho, Daisy L. Hung, Ovid J. L. Tzeng, and Jen-Chuen Hsieh. 2004. "Orthographic and phonological processing of Chinese characters: an fMRI study." *NeuroImage* 21 (4): 1721–1731.

Kuperberg, Gina R. 2007. “Neural mechanisms of language comprehension: challenges to syntax.” *Brain Research* 1146:23–49.

Kuperberg, Gina R., Eric Halgren, Douglas Greve, Bruce Fischl, Rahul Desikan, Gloria Waters, Sujith Vijayan, et al. 2000. “Event-related fMRI reveals distinct patterns of neural modulation during semantic and syntactic processing of sentences.” *NeuroImage* 11 (5): S263.

Kuperberg, Gina R., and T. Florian Jaeger. 2016. “What do we mean by prediction in language comprehension?” *Language, Cognition and Neuroscience* 31 (1): 32–59.

Kuperberg, Gina R., Donna A. Kreher, Tatiana Sitnikova, David N. Caplan, and Phillip J. Holcomb. 2007. “The role of animacy and thematic relationships in processing active English sentences: evidence from event-related potentials.” *Brain and Language* 100 (3): 223–237.

Kuperberg, Gina R., Balaji M. Lakshmanan, David N. Caplan, and Phillip J. Holcomb. 2006. “Making sense of discourse: an fMRI study of causal inferencing across sentences.” *NeuroImage* 33 (1): 343–361.

Kuperberg, Gina R., Tatiana Sitnikova, David Caplan, and Phillip J. Holcomb. 2003. “Electrophysiological distinctions in processing conceptual relationships within simple sentences.” *Brain Research. Cognitive Brain Research* 17 (1): 117–129.

Kuperberg, Gina R., Tatiana Sitnikova, and Balaji M. Lakshmanan. 2008. “Neuroanatomical distinctions within the semantic system during sentence comprehension: evidence from functional magnetic resonance imaging.” *NeuroImage* 40 (1): 367–388.

Kuroda, Sige-Yuki. 1972. “The Categorical and the Thetic Judgment: Evidence from Japanese Syntax.” *Foundations of Language* 9 (2): 153–185.

———. 1973. “Le jugement catégorique et le jugement thétique : exemples tirés de la syntaxe japonaise.” *Langages* 8 (30): 81–110.

———. 2005. “Focusing On The Matter of Topic: A Study of Wa and Ga in Japanese.” *Journal of East Asian Linguistics* 14 (1): 1–58.

Kutas, Marta, and Steven A. Hillyard. 1980. “Reading senseless sentences: brain potentials reflect semantic incongruity.” *Science* 207 (4427): 203–205.

———. 1984. “Brain potentials during reading reflect word expectancy and semantic association.” *Nature* 307 (5947): 161–163.

Kutas, Marta, and Cyma K. Van Petten. 1988. “Event-Related Brain Potential Studies of Language.” In *Advances in Psychophysiology, Volume 3,* edited by Patrick K. Ackles, J. Richard Jennings, and Michael G. H. Coles, 139–187. Greenwich, CT: JAI Press.

Kutas, Marta, Cyma K. Van Petten, and Robert Kluender. 2006. “Psycholinguistics Electrified II (1994–2005).” In *Handbook of Psycholinguistics (Second Edition),* edited by Matthew J. Traxler and Morton A. Gernsbacher, 659–724. London: Academic Press.

LaBar, Kevin S., Darren R. Gitelman, Todd B. Parrish, and M.-Marsel Mesulam. 1999. "Neuroanatomic Overlap of Working Memory and Spatial Attention Networks: A Functional MRI Comparison within Subjects." *NeuroImage* 10 (6): 695–704.

Lacheret, Anne, Sabine Ploux, and Bernard Victorri. 1998. "Prosodie et thématisation en français parlé." *Cahiers de praxématique* 30:89–111.

Lahousse, Karen. 2003a. "La complexité de la notion de topique et l'inversion du sujet nominal." *Travaux de linguistique* 47 (2): 111–136.

———. 2003b. "La distribution de l'inversion nominale en français dans les principales non interrogatives et les subordonnées circonstan." *Lingvisticæ Investigationes* 26 (1): 123–158.

———. 2003c. "The distribution of postverbal nominal subjects in French : a syntactic, semantic and pragmatic analysis." PhD diss., Université de Paris VIII.

———. 2006. "NP subject inversion in French: two types, two configurations." *Lingua* 116 (4): 424–461.

Lambrecht, Knud. 1981. *Topic, Antitopic and Verb Agreement in Non-Standard French.* Pragmatics & Beyond. John Benjamins Publishing Company.

———. 1987. "On the status of SVO sentences in French discourse." In *Coherence and Grounding in Discourse: Outcome of a Symposium, Eugene, Oregon, June 1984,* edited by Russell S. Tomlin, vol. 11. Typological Studies in Language. Amsterdam: John Benjamins Publishing Company, January 1.

———. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents.* Cambridge: Cambridge University Press.

———. 1998. "Sur la relation formelle et fonctionnelle entre topiques et vocatifs." *Langues* 1 (1): 34–45.

———. 2001. "Dislocation." In *Language Typology and Language Universals: An International Handbook,* edited by Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, 2:1050–1078. Berlin: De Gruyter.

LaPolla, Randy J. 1990. "Grammatical relation in Chinese, synchronic and diachronic considerations." PhD diss., University of California.

———. 1995. "Pragmatic relations and word order in Chinese." In *Word order in Discourse,* edited by Pamela A. Downing and Michael Noonan, 30:297. Typological Studies in Language. Amsterdam: John Benjamins Publishing Company.

Larson, Richard K. 1985. "Bare-NP Adverbs." *Linguistic Inquiry* 16 (4): 595–621.

———. 1988. "On the Double Object Construction." *Linguistic Inquiry* 19 (3): 335–391.

Lau, Ellen, Clare Stroud, Silke Plesch, and Colin Phillips. 2006. "The role of structural prediction in rapid syntactic analysis." *Brain and Language* 98 (1): 74–88.

Le Bidois, Robert. 1952. *L'inversion du sujet dans la prose contemporaine (1900-1950): Étudiée plus spécialement dans l'oeuvre de Marcel Proust.* Paris: Éditions d'Artrey.

Le Ny, Jean-François. 2004. “Eléments de psycholinguistique cognitive : des représentations à la compréhension.” In *La Linguistique cognitive,* edited by Catherine Fuchs, 155–171. Cogniprisme. Paris: Ophrys.

Ledgeway, Adam. 2011. “Subject Licensing in CP: The Neapolitan Double-Subject Construction.” In *Mapping the Left Periphery,* edited by Paola Benincà and Nicola Munaro. The Cartography of Syntactic Structures 5. Oxford: Oxford University Press.

Ledoux, Kerry, Peter C. Gordon, C. Christine Camblin, and Tamara Y. Swaab. 2007. “Coreference and lexical repetition: Mechanisms of discourse integration.” *Memory & Cognition* 35 (4): 801–815.

Lee, Chia-Ying, Hsu-Wen Huang, Wen-Jui Kuo, Jie-Li Tsai, and Ovid J. L. Tzeng. 2010. “Cognitive and neural basis of the consistency and lexicality effects in reading Chinese.” *Journal of Neurolinguistics* 23 (1): 10–27.

Lee, Chia-Ying, Jie-Li Tsai, Erica Chung, I Su, Ovid J. L. Tzeng, and Daisy L. Hung. 2005. “Consistency, regularity and frequency effects in naming Chinese characters.” *Language and Linguistics,* no. 6: 175–197.

Lee, Chia-Ying, Jie-Li Tsai, Hsu-Wen Huang, Daisy L. Hung, and Ovid J. L. Tzeng. 2006. “The temporal signatures of semantic and phonological activations for Chinese sublexical processing: An event-related potential study.” *Brain Research* 1121 (1): 150–159.

Lee, Chungmin. 2003. “Contrastive Topic and proposition structure.” In *Asymmetry in Grammar: Volume 1: Syntax and semantics,* edited by Anna Maria Di Sciullo, 1:345–371. Linguistik Aktuell/Linguistics Today 57. Amsterdam: John Benjamins.

Lee-Wong, Song Mei. 2001. “Coherence, Focus And Structure: The Role Of Discourse Particle Ne.” *Pragmatics* 11 (2): 139–153.

Lei, Miaomei, Hiroyuki Akama, and Brian Murphy. 2014. “Neural basis of language switching in the brain: fMRI evidence from Korean–Chinese early bilinguals.” *Brain and Language* 138:12–18.

Lely, Heather K. J. van der, and Jackie Battell. 2003. “Wh-Movement in Children with Grammatical SLI: A Test of the RDDR Hypothesis.” *Language* 79 (1): 153–181.

Lely, Heather K. J. van der, Melanie Jones, and Chloë R. Marshall. 2011. “Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI.” *Lingua. International Review of General Linguistics. Revue Internationale De Linguistique Generale* 121 (3): 408–422.

Lely, Heather K. J. van der, and Steven Pinker. 2014. “The biological basis of language: insight from developmental grammatical impairments.” *Trends in Cognitive Sciences* 18 (11): 586–595.

Levelt, Willem J. M. 1989. *Speaking: From Intention to Articulation.* ACL-MIT Press series in natural-language processing. Cambridge, MA: MIT Press.

———. 1996. “Perspective taking and ellipsis in spatial descriptions.” In *Language and space,* edited by Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, 77–107. Cambridge, MA: MIT Press.

Levinson, Stephen C. 2003. *Space in Language and Cognition: Explorations in Cognitive Diversity.* Cambridge: Cambridge University Press.

Levy, Roger. 2008. “Expectation-based syntactic comprehension.” *Cognition* 106 (3): 1126–1177.

Li, Charles N., and Sandra A. Thompson. 1976. “Subjet and Topic: A new Typology of Language.” In *Subject and Topic,* edited by Charles N. Li, 457–489. New York: Academic Press.

———. 1981. *Mandarin Chinese: a Functional Reference Grammar.* Berkeley: University of California Press.

Li, Le, Jubin Abutalebi, Lijuan Zou, Xin Yan, Lanfang Liu, Xiaoxia Feng, Ruiming Wang, Taomei Guo, and Guosheng Ding. 2015. “Bilingualism alters brain functional connectivity between “control”regions and “language”regions: Evidence from bimodal bilinguals.” *Neuropsychologia* 71:236–247.

Li, Weijun, and Yaling Yang. 2009. “Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences.” *Neuroscience* 158 (4): 1416–1425.

Li, Wendan. 1996. “Second language acquisition of Topic-Comment structures in Mandarin Chinese.” PhD diss., University of Alberta.

———. 2005. *Topic chains in Chinese : a discourse analysis and applications in language teaching.* München: Lincom.

Li, Xiao-qing, and Gui-qin Ren. 2012. “How and when accentuation influences temporally selective attention and subsequent semantic processing during on-line spoken language comprehension: An ERP study.” *Neuropsychologia* 50:1882–1894.

Li, Xiao-qing, Yu-fang Yang, and Yong Lu. 2010. “How and when prosodic boundaries influence syntactic parsing under different discourse contexts: an ERP study.” *Biological Psychology* 83 (3): 250–259.

Li, Yen-Hui Audrey. 1990. *Order and Constituency in Mandarin Chinese.* Vol. 19. Studies in Natural Language and Linguistic Theory. Dordrecht: Kluwer Academic.

———. 1998. “Argument Determiner Phrases and Number Phrases.” *Linguistic Inquiry* 29 (4): 693–702.

———. 2000. “Topic Structures and Minimal Effort.” *ZAS Papers in Linguistics* 20:1–20.

———. 2003. “NP as argument.” In *Functional Structure(s), Form and Interpretation: Perspectives from East Asian Languages,* edited by Yen-Hui Audrey Li and Andrew Simpson, 3–27. London: RoutledgeCurzon.

———. 2014. “Born empty.” *Lingua,* Structural Approaches to Ellipsis, 151, Part A:43–68.

Li, Yen-Hui Audrey, Andrew Simpson, and Wei-Tien Dylan Tsai, eds. 2014. *Chinese Syntax in a Cross-linguistic Perspective.* Oxford: Oxford University Press.

Liao 廖, Huaibao 怀宝. 2004. “英汉语逗号的比较研究 Yinghanyu douhao de bijiao yanjiu [A Comparative Study of the Comma in English and Chinese].” PhD diss., Shanghai: East China Normal University.

Lin 林, Maocan 茂灿. 2006. “疑问和陈述语气与边界调 Yiwen he chenshu yuqi yu bianjiediao [Interrogative vs. declarative and the boundary tone in Standard Chinese].” *中国语文*, no. 4: 364–376.

———. 2012. “汉语焦点重音和功能语气及其特征 Hanyu jiaodian chongyin he gongneng yuqi ji qi tezheng.” *中国语音学报 Zhongguo Yuyan Xuebao* 3.

Lindell, Annukka K. 2006. “In Your Right Mind: Right Hemisphere Contributions to Language Processing and Production.” *Neuropsychology Review* 16 (3): 131–148.

Liu 刘, Danqing 丹青. 1995. “语义先还是语用优先? 汉语语法学体系建设断想 Yuyi xian haishi yuyong youxian? Hanyu yufaxue tixi jianshe duanxiang.” *语言研究 Yuyan yanjiu,* no. 2.

———. 1998. “普通话与上海话中的拷贝式话题结构 Putonghua yu Shanghaihua zhong de kaobeishi huati jiegou.” *语言教学与研究 Yuyan Jiaoxue yu Yanjiu* 2.

———. 2001. “论元分裂式话题结构初探 Lun yuan fenlieshi huati jiegou chutan.” In *Yuyan wenti zai renshi [语言问题再认识],* edited by Kaitai 开泰 Fan 范 and Huyang 沪扬 Qi 齐. Shanghai: Shanghai Jiaoyu [上海教育出版社].

———. 2008. “汉语名词性短语的句法类型特征 [Typological features of nominal phrases in Chinese].” *中国语文 Zhongguo Yuwen [Studies of the Chinese Language]* 322 (1): 3–20.

Liu 刘, Danqing 丹青, and Liejong 烈炯 Xu 徐. 1998. “焦点与背景、话题机及汉语 “连” 字句 Jiaodian yu beijing, huati ji Hanyu ‘lian’ziju [Focus, background, and lian-clause in Chinese language].” *中国语文 Zhongguo Yuwen [Studies of the Chinese Language],* no. 4: 243–252.

Liu 刘, Ningsheng 宁生. 1983. “汉语口语中的双主谓结构句 Hanyu kouyuzhong de shuang zhuwei jiegou.” *中国语文 Zhongguo Yuwen [Studies of the Chinese Language]* 2 (2): 97–98.

Liu 刘, Xiao 晓. 2005. “英语标点的元功能研究 Yingyu biaodian de yuan gongneng yanjiu[A metafunctional study of English punctuation].” PhD diss., Hunan Normal University.

Liu, Baolin, Zhongning Wang, and Zhixing Jin. 2010. “The effects of punctuations in Chinese sentence comprehension: An ERP study.” *Journal of Neurolinguistics* 23 (1): 66–80.

Liu, Chao, Wu-Tian Zhang, Yi-Yuan Tang, Xiao-Qin Mai, Hsuan-Chih Chen, Twila Tardif, and Yue-Jia Luo. 2008. “The Visual Word Form Area: Evidence from an fMRI study of implicit processing of Chinese characters.” *NeuroImage* 40 (3): 1350–1361.

Liu, Danqing. 2000. “Identical topics in Mandarin Chinese and Shanghainese.” In *Issues on topics,* edited by Kerstin Schwabe, 20:43–69. ZAS papers in linguistics. Berlin: Universitätsbibliothek Johann Christian Senckenberg.

Lœvenbruck, Hélène, Monica Baciu, Christoph Segebarth, and Christian Abry. 2005. "The left inferior frontal gyrus under focus: an fMRI study of the production of deixis via syntactic extraction and prosodic focus." *Journal of Neurolinguistics* 18 (3): 237–258.

Logothetis, Nikos K. 2008. "What we can do and what we cannot do with fMRI." *Nature* 453 (7197): 869–878.

Logothetis, Nikos K., and Josef Pfeuffer. 2004. "On the nature of the BOLD fMRI contrast mechanism." *Magnetic Resonance Imaging* 22 (10): 1517–1531.

Longobardi, Giuseppe, Andrea Ceolin, Aaron Ecay, Silvia Ghirotto, Cristina Guardiano, Monica-Alexandrina Irimia, Dimitris Michelioudakis, et al. 2016. "Formal linguistics as a cue to demographic history." *Journal of anthropological sciences = Rivista di antropologia: JASS* 94:147–155.

Longworth, Catherine E., Lorraine K. Tyler, and William D. Marslen-Wilson. 2003. "Language deficits and basal ganglia lesions: The past tense." *Brain and Language,* Academy of Aphasia Abstract Issue 2003, 87 (1): 7–8.

Love, Tracy, David Swinney, Matthew Walenski, and Edgar Zurif. 2008. "How left inferior frontal cortex participates in syntactic processing: Evidence from aphasia." *Brain and Language* 107 (3): 203–219.

Lowder, Matthew W., and Peter C. Gordon. 2014. "Effects of Animacy and Noun-Phrase Relatedness on the Processing of Complex Sentences." *Memory & cognition* 42 (5): 794–805.

Lü 吕, Shuxiang 叔湘. 1980. 现代汉语八百词 *Xiandai Hanyu babai ci.* Shanghai: Commercial Press [商务印书馆].

———. 1986a. "主谓谓语句举例 Zhuwei weiyuju juli [Samples of sentences with subject-predicate predicates]." 中国语文, no. 5: 334–340.

———. 1986b. "句法的灵活性 [The Flexibility of Chinese syntax]." *Studies of the Chinese Language [*中国语文 - *Zhongguo Yuwen]* 1:40–47.

Lu, Bingfu, and Yuzhi Shi. 2009. "The placement of Chinese adverbials revisited: what differs Chinese word order from other SVO languages." In *Studies of Chinese Linguistics,* edited by Janet Zhiqun Xing, 23–44. Hong Kong: Hong Kong University Press.

Lu, Peng. 2003. "Subordination adverbiale en chinois contemporain." PhD diss., Paris 7.

———. 2004. "La Subordination Adverbiale en Chinois Contemporain." PhD diss., Université de Paris VII.

Luo, Yingyi, and Xiaolin Zhou. 2010. "ERP evidence for the online processing of rhythmic pattern during Chinese sentence reading." *NeuroImage* 49 (3): 2836–2849.

Mack, Jennifer E., Aya Meltzer-Asscher, Elena Barbieri, and Cynthia K. Thompson. 2013. "Neural Correlates of Processing Passive Sentences." *Brain Sciences* 3 (3): 1198–1214.

Mack, Jennifer E., Michaela Nerantzini, and Cynthia K. Thompson. 2017. "Recovery of Sentence Production Processes Following Language Treatment in Aphasia: Evidence from Eyetracking." *Frontiers in Human Neuroscience* 11.

MacWhinney, Brian. 2000. *The Childes Project: Tools for Analyzing Talk.* 3rd. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, Brian, and Elizabeth Bates, eds. 1989. "The learner's problem of arranging words." In *The Crosslinguistic Study of Sentence Processing,* 292–327. Cambridge: Cambridge University Press.

Majid, Asifa, Melissa Bowerman, Sotaro Kita, Daniel B. M. Haun, and Stephen C. Levinson. 2004. "Can language restructure cognition? The case for space." *Trends in Cognitive Sciences* 8 (3): 108–114.

Maki, Hideki, Lizanne Kaiser, and Masao Ochi. 1999. "Embedded topicalization in English and Japanese." *Lingua* 109 (1): 1–14.

Makuuchi, Michiru, Jörg Bahlmann, Alfred Anwander, and Angela D. Friederici. 2009. "Segregating the core computational faculty of human language from working memory." *Proceedings of the National Academy of Sciences* 106 (20): 8362–8367.

Makuuchi, Michiru, Yosef Grodzinsky, Katrin Amunts, Andrea Santi, and Angela D. Friederici. 2013. "Processing noncanonical sentences in broca's region: reflections of movement distance and type." *Cerebral Cortex* 23 (3): 694–702.

Mann, William C., and Sandra A. Thompson, eds. 1992. *Discourse Description: Diverse linguistic analyses of a fund-raising text.* Vol. 16. Pragmatics & Beyond New Series. Amsterdam: John Benjamins Publishing Company, April 8.

Marandin, Jean-Marie. 2001. "Unaccusative inversion in French." In *Romance Languages and Linguistic Theory 1999,* edited by Frank Drijkoningen, 195–222. Amsterdam: John Benjamins.

———. 2005. "Formatage de l'information: focus et contexte." In *Interpréter en contexte,* edited by Francis Corblin and Claire Gardent, 31–79. Paris: Hermes.

Marantz, Alec. 1993. "Implications of Asymmetries in Double Object Constructions." In *Theoretical Aspects of Bantu Grammar 1,* edited by Sam A. Mchombo, 113–151. Stanford University: CSLI Publications.

———. 2005. "Generative linguistics within the cognitive neuroscience of language." *The Linguistic Review* 22 (2): 429–445.

———. 2013. "Verbal argument structure: Events and participants." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:152–168.

Margulies, Daniel S., Satrajit S. Ghosh, Alexandros Goulas, Marcel Falkiewicz, Julia M. Huntenburg, Georg Langs, Gleb Bezgin, et al. 2016. "Situating the default-mode network along a principal gradient of macroscale cortical organization." *Proceedings of the National Academy of Sciences* 113 (44): 12574–12579.

Marr, David. 1983. *Vision.* New York: W.H. Freeman & Co.

Marshuetz, Christy. 2005. “Order information in working memory: an integrative review of evidence from brain and behavior.” *Psychological Bulletin* 131 (3): 323–339.

Martin, Samuel Elmo. 2004. *A Reference Grammar of Japanese.* Honolulu, HI: University of Hawaii Press.

Martín-Loeches, Manuel, Roland Nigbur, Pilar Casado, Annette Hohlfeld, and Werner Sommer. 2006. “Semantics prevalence over syntax during sentence processing: a brain potential study of noun-adjective agreement in Spanish.” *Brain Research* 1093 (1): 178–189.

Maslova, Elena, and Giuliano Bernini. 2006. “Sentence topics in the languages of Europe and beyond.” In *Pragmatic Organization of Discourse in the Languages of Europe,* edited by Giuliano Bernini and Marcia L. Schwartz. Berlin: De Gruyter.

Matchin, William, Christopher Hammerly, and Ellen Lau. 2017. “The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI.” *Cortex* 88 (Supplement C): 106–123.

Matisoff, James A. 1973. *The Grammar of Lahu.* Berkeley: University of California Press.

Matzke, Mike, Heinke Mai, Wido Nager, Jascha Rüsseler, and Thomas F. Münte. 2002. “The costs of freedom: an ERP – study of non-canonical sentences.” *Clinical Neurophysiology* 113 (6): 844–852.

Mazoyer, Bernard, Nathalie Tzourio, Victor Frak, A. Syrota, N. Murayama, O. Levrier, G. Salamon, Stanislas Dehaene, L. Cohen, and Jacques Mehler. 1993. “The cortical representation of speech.” *Journal of Cognitive Neuroscience* 5 (4): 467–479.

McBride-Chang, Catherine, and Connie Suk-Han Ho. 2000. “Developmental issues in Chinese children’s character acquisition.” *Journal of Educational Psychology* 92 (1): 50–55.

McBride-Chang, Catherine, and Robert V. Kail. 2002. “Cross-cultural similarities in the predictors of reading acquisition.” *Child Development* 73 (5): 1392–1407.

McCawley, James D. 1989. “Notes on Li and Thompson, Mandarin Chinese: a functional reference grammar.” *Journal of the Chinese Language Teachers Association* 24 (1): 19–42.

Mecklinger, Axel, Herbert Schriefers, Karsten Steinhauer, and Angela D. Friederici. 1995. “Processing relative clauses varying on syntactic and semantic dimensions: an analysis with event-related potentials.” *Memory & Cognition* 23 (4): 477–494.

Mehler, Jacques. 1963. “Some effects of grammatical transformations on the recall of english sentences.” *Journal of Verbal Learning and Verbal Behavior* 2 (4): 346–351.

Mei, Leilei, Gui Xue, Zhong-Lin Lu, Chuansheng Chen, Miao Wei, Qinghua He, and Qi Dong. 2015. “Long-term experience with Chinese language shapes the fusiform asymmetry of English reading.” *NeuroImage* 110:3–10.

Mei, Leilei, Gui Xue, Zhong-Lin Lu, Qinghua He, Mingxia Zhang, Feng Xue, Chuansheng Chen, and Qi Dong. 2013. “Orthographic transparency modulates the functional asymmetry in the fusiform cortex: An artificial language training study.” *Brain and Language,* The Neural Basis of Reading, 125 (2): 165–172.

Meisel, Jürgen M. 1994. "Getting FAT: Finiteness, Agreement and Tense in Early Grammars." In *Bilingual First Language Acquisition: French and German grammatical development,* edited by Jürgen M. Meisel, 7:89–129. Amsterdam: John Benjamins.

Menenti, Laura, Sarah M. E. Gierhan, Katrien Segaert, and Peter Hagoort. 2011. "Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI." *Psychological Science* 22 (9): 1173–1182.

Menenti, Laura, Karl Magnus Petersson, René Scheeringa, and Peter Hagoort. 2008. "When Elephants Fly: Differential Sensitivity of Right and Left Inferior Frontal Gyri to Discourse and World Knowledge." *Journal of Cognitive Neuroscience* 21 (12): 2358–2368.

Meulen, Ineke van der, Roelien Bastiaanse, and Johan Rooryck. 2002. "Wh-movement in French agrammatism." *Brain and Language* 83:184–185.

———. 2005. "Wh-questions in agrammatism: a movement deficit?" *Stem-, Spraak- en Taalpatholgiee* 13 (1): 24–36.

Miller, George A., George A. Heise, and William Lichten. 1951. "The intelligibility of speech as a function of the context of the test materials." *Journal of Experimental Psychology* 41 (5): 329–335.

Miller, George A., and Stephen Isard. 1963. "Some perceptual consequences of linguistic rules." *Journal of Verbal Learning and Verbal Behavior* 2 (3): 217–228.

Millotte, Séverine, Elodie Cauvet, Perrine Brusini, and Anne Christophe. 2013. "Discovering word forms and word meanings: The role of phrasal prosody and function words." In *The Cambridge Handbook of Biolinguistics,* edited by Cedric Boeckx and Kleanthes K.Editors Grohmann, 86–93. Cambridge University Press.

Mitchell, Rachel L. C., and Tim J. Crow. 2005. "Right hemisphere language functions and schizophrenia: the forgotten hemisphere?" *Brain: A Journal of Neurology* 128 (Pt 5): 963–978.

Miyamoto, Edison T., and Michiko Nakamura. 2003. "Subject/object asymmetries in the processing of relative clauses in Japanese." In *WCCFL 22: Proceedings of the 22nd West Coast Conference on Formal Linguistics,* edited by Gina Garding and Mimu Tsujimura, 342–355. Somerville, MA: Cascadilla Press.

Mo, Ce, Mengxia Yu, Carol Seger, and Lei Mo. 2015. "Holistic neural coding of Chinese character forms in bilateral ventral visual system." *Brain and Language* 141:28–34.

Molinaro, Nicola, Francesco Vespignani, and Remo Job. 2008. "A deeper reanalysis of a superficial feature: an ERP study on agreement violations." *Brain Research* 1228:161–176.

Momma, Shota, and Colin Phillips. 2018. "The Relationship Between Parsing and Generation." *Annual Review of Linguistics* 4 (1).

Morgenstern, Aliyah, and Laurent Danon-Boileau. 2009. "Peut-on parler de prédication dans les premiers énoncés de enfant?" *Faits de langues,* no. 31: 57–65.

Moro, Andrea. 2006. *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure.* Cambridge: Cambridge University Press.

———. 2008. *The Boundaries of Babel: The Brain and the Enigma of Impossible Languages.* Current Studies in Linguistics. Cambridge, MA: MIT Press.

———. 2016. *Impossible Languages.* 1st ed. The MIT Press.

Moro, Andrea, Marco Tettamanti, Daniela Perani, Caterina Donati, Stefano F. Cappa, and Ferrucio Fazio. 2001. "Syntax and the Brain: Disentangling Grammar by Selective Anomalies." *NeuroImage* 13 (1): 110–118.

Mosegaard Hansen, Maj-Britt. 2001. "Syntax in interaction: Form and function of yes/no interrogatives in spoken standard French." *Studies in Language* 25 (3): 463–520.

Muayqil, Taim, Jodie Davies-Thompson, and Jason J. S. Barton. 2015. "Representation of visual symbols in the visual word processing network." *Neuropsychologia* 69:232–241.

Mueller, Jutta L., Shirley-Ann Rueschemeyer, Kentaro Ono, Motoaki Sugiura, Norihiro Sadato, and Akinori Nakamura. 2014. "Neural networks involved in learning lexical-semantic and syntactic information in a second language." *Frontiers in Psychology* 5:1209.

Munnich, E., B. Landau, and B. A. Dosher. 2001. "Spatial language and spatial representation: a cross-linguistic comparison." *Cognition* 81 (3): 171–207.

Münte, Thomas F., and Hans-Jochen Heinze. 1994. "ERP Negativities During Syntactic Processing of Written Words." In *Cognitive Electrophysiology,* edited by Hans-Jochen Heinze, Thomas F. Münte, and George R. Mangun, 211–238. Boston, MA: Birkhäuser.

Musso, Mariacristina, Andrea Moro, Volkmar Glauche, Michel Rijntjes, Jürgen Reichenbach, Christian Büchel, and Cornelius Weiller. 2003. "Broca's area and the language instinct." *Nature Neuroscience* 6 (7): 774–781.

Nagashima, Tatsuya. 1988. *A handbook of Japanese grammar: English-Japanese.* Tokyo: Pana-Lingua Institute.

Nakanishi, Kimiko. 2001. "Prosody and information structure in Japanese: A case study of topic marker wa." *Japanese/Korean Linguistics* 10:434–447.

———. 2008. "Prosody and Scope Interpretations of the Topic Marker 'Wa' in Japanese." In *Topic and Focus,* edited by Chungmin Lee, Matthew Gordon, and Daniel Büring, 177–193. Studies in Linguistics and Philosophy. Dordrecht: Springer.

Nastase, Samuel, Vittorio Iacovella, and Uri Hasson. 2014. "Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems." *Human Brain Mapping* 35 (4): 1111–1128.

Neville, Helen, Janet L. Nicol, Andrew Barss, Kenneth I. Forster, and Merill F. Garrett. 1991. "Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials." *Journal of Cognitive Neuroscience* 3 (2): 151–165.

Newman, Sharlene D., Toshikazu Ikuta, and Thomas Burns. 2010. "The effect of semantic relatedness on syntactic analysis: An fMRI study." *Brain and Language* 113 (2): 51–58.

Newmeyer, Frederick J., and Laurel B. Preston, eds. 2014. *Measuring Grammatical Complexity.* Oxford, New York: Oxford University Press, December 30.

Newport, Elissa L. 2011. "The Modularity Issue in Language Acquisition: A Rapprochement? Comments on Gallistel and Chomsky." *Language Learning and Development* 7 (4): 279–286.

Nicol, Janet L., Janet Dean Fodor, and David A. Swinney. 1994. "Using cross-modal lexical decision tasks to investigate sentence processing." *Journal of Experimental Psychology. Learning, Memory, and Cognition* 20 (5): 1229–1238.

Nieuwland, Mante S., and Jos J. A. Van Berkum. 2005. "Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension." *Cognitive Brain Research* 24 (3): 691–701.

Ning, Chunyan. 1993. "The Overt Syntax of Relativization and Topicalization in Chinese." PhD diss., University of California.

Nishimura, Miwa. 1989. "The topic-comment construction in Japanese-English code-switching." *World Englishes* 8 (3): 365–377.

Nolda, Andreas. 2008. "Topic integration: syntax and semantics of German 'split topicalization'." In *Explorations in Integrational Linguistics: Four essays on German, French, and Guaraní,* edited by Robin Sackmann, 191–220. Current Issues in Linguistic Theory 285. Amsterdam: John Benjamins Publishing Company.

Noppeney, Uta, J. Phillips, and Cathy J. Price. 2004. "The neural areas that control the retrieval and selection of semantics." *Neuropsychologia* 42 (9): 1269–1280.

Noppeney, Uta, and Cathy J. Price. 2004. "Retrieval of abstract semantics." *NeuroImage* 22 (1): 164–170.

Norman, Jerry. 1988. *Chinese.* New York: Cambridge University Press.

Novick, Jared M., John C. Trueswell, and Sharon L. Thompson-Schill. 2005. "Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension." *Cognitive, Affective & Behavioral Neuroscience* 5 (3): 263–281.

Ohta, Shinri, Masatoshi Koizumi, and Kuniyoshi L. Sakai. 2017. "Dissociating Effects of Scrambling and Topicalization within the Left Frontal and Temporal Language Areas: An fMRI Study in Kaqchikel Maya." *Frontiers in Psychology* 8.

Olsen, Rosanna K., Melissa M. Pangelinan, Cari Bogulski, M. Mallar Chakravarty, Gigi Luk, Cheryl L. Grady, and Ellen Bialystok. 2015. "The effect of lifelong bilingualism on regional grey and white matter volume." *Brain Research,* The Cognitive Neuroscience of Aging and Memory, 1612:128–139.

Opitz, Bertram, and Sonja A. Kotz. 2012. "Ventral premotor cortex lesions disrupt learning of sequential grammatical structures." *Cortex* 48 (6): 664–673.

Osterhout, Lee, and David A. Swinney. 1993. "On the temporal course of gap-filling during comprehension of verbal passives." *Journal of Psycholinguistic Research* 22 (2): 273–286.

Ouden, Dirk-Bart den den, Hans Hoogduin, Laurie A. Stowe, and Roelien Bastiaanse. 2008. "Neural correlates of Dutch Verb Second in speech production." *Brain and Language* 104 (2): 122–131.

Ouden, Dirk-Bart den, Steve Fix, Todd B. Parrish, and Cynthia K. Thompson. 2009. "Argument structure effects in action verb naming in static and dynamic conditions." *Journal of neurolinguistics* 22 (2): 196–215.

Ouden, Dirk-Bart den, Dorothee Saur, Wolfgang Mader, Björn Schelter, Sladjana Lukic, Eisha Wali, Jens Timmer, and Cynthia K. Thompson. 2012. "Network modulation during complex syntactic processing." *NeuroImage* 59 (1): 815–823.

Owen, Adrian M., Chantal E. Stern, Rodney B. Look, Irene Tracey, Bruce R. Rosen, and Michael Petrides. 1998. "Functional organization of spatial and nonspatial working memory processing within the human lateral frontal cortex." *Proceedings of the National Academy of Sciences* 95 (13): 7721–7726.

Owens, Robert E. 1984. *Language development: an introduction.* 8th ed. Colombus, Ohio: Charles E. Merrill.

Packard, Jerome L. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach.* Cambridge: Cambridge University Press.

Pallier, Christophe, Anne-Dominique Devauchelle, and Stanislas Dehaene. 2011. "Cortical representation of the constituent structure of sentences." *Proceedings of the National Academy of Sciences* 108 (6): 2522–2527.

Palomar-García, María-Ángeles, Elisenda Bueichekú, César Ávila, Ana Sanjuán, Kristof Strijkers, Noelia Ventura-Campos, and Albert Costa. 2015. "Do bilinguals show neural differences with monolinguals when processing their native language?" *Brain and Language* 142:36–44.

Pan 潘, Haihua 海华. 1997. "词汇映射理论在汉语句法研究中的应用 Cihui yingshe lilun zai hanyu jufa yanjiu zhong de yingyong." *现代外语,* no. 4: 4–18.

Pan 潘, Haihua 海华, and Jingquan 景泉 Han 韩. 2005. "显性非定格动词结构的句法研究 Xianxing fei bin'ge dongci jiegou de jufa yanjiu." *语言研究 Yuyan yanjiu* 25 (3): 1–13.

Pan 潘, Yunzhong 允中. 1982. *汉语语法史概要 Hanyu yufashi gaiyao [A Summary of Chinese Grammar History].* Zhengzhou: Ancient Books Publishing House.

Pan, Haihua. 2000. "Why the blocking effect?" In *Long Distance Reflexives,* BRILL, edited by Peter Cole, Gabriella Hermon, and C.-T. James Huang, 279–316. Syntax and semantics 33. London: Academic Press, October 17.

Pan, Haihua, and Jianhua Hu. 2000. "Deriving the Subject-Object Asymmetry in Topicalization." Hong Kong Polytechnic University, June 21.

———. 2001. "Chinese Topics: Are They Syntactically or Semantically Licensed?" *Manuscript.*

———. 2002. "Representing Topic-Comment Structures in Chinese." In *Proceedings of the 16th Pacific Asia Conference on Language, Information and Computation,* 382–390. Jeju, Korea.

Pan, Victor Junnan. 2006. “Wh-topicalization and wh-in-situ in Mandarin Chinese.” CRLAO, EHESS, Paris.

———. 2007. “Interrogation et quantification : le rôle et la fonction des particules et des syntagmes interrogatifs en chinois mandarin [Interrogation and quantification : the rôle and the function of particles and interrogative phrases in mandarin Chinese].” PhD diss., Université de Nantes.

———. 2014. “Wh-ex-situ in Mandarin Chinese: Mapping Between Information Structure and Split CP.” *Linguistic Analysis* 39 (3): 371–413.

Papeo, Liuba, and Angelika Lingnau. 2015. “First-person and third-person verbs in visual motion-perception regions.” *Brain and Language* 141:135–141.

Paradis, Michel. 2008. “Bilingualism and neuropsychiatric disorders.” *Journal of Neurolinguistics* 21 (3): 199–230.

Paris, Marie-Claude. 1979. “Some aspects of the syntax and semantics of the ”lián...ye/dou” construction in Mandarin.” *Cahiers de linguistique - Asie orientale* 5 (1): 47–70.

———. 1998. “Focus operators and types of predication in Mandarin.” *Cahiers de linguistique - Asie orientale* 27 (2): 139–159.

———. 1999. “Ordre des mots, topic et focus en Chinois contemporain.” In *La thématisation dans les langues, actes du colloque de Caen (9-11 octobre 1997),* edited by Claude Guimier, 201–216. Bern: Peter Lang.

Park, Soojin, and Marvin M. Chun. 2009. “Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception.” *NeuroImage* 47 (4): 1747–1756.

Paul, Waltraud. 2002. “Sentence-internal topics in Mandarin Chinese: the case of object preposing.” *Language and Linguistics* 3 (4): 695–714.

———. 2005. “Low IP and Left Periphery in Mandarin Chinese.” *Recherches linguistiques de Vincennes,* no. 33: 111–134.

———. 2010. “What the topic is (not) about: The case of (the split CP in) Mandarin Chinese.” Venice, September.

———. 2012. “Topic(s) in Chinese revisited.” Université Paris 8, Paris, November.

———. 2014. *New Perspectives on Chinese Syntax.* Berlin, Boston: De Gruyter Mouton.

Paul, Waltraud, and John Whitman. 2008. “Shi ···de focus clefts in Mandarin Chinese.” *Linguistic Review* 25:413–451.

Payne, Doris L., ed. 1992. *Pragmatics of Word Order Flexibility.* Vol. 22. Typological Studies in Language. Amsterdam: John Benjamins Publishing Company, July 1.

Pellikka, Janne, Päivi Helenius, Jyrki P. Mäkelä, and Minna Lehtonen. 2015. “Context affects L1 but not L2 during bilingual word recognition: An MEG study.” *Brain and Language* 142:8–17.

Peng, Danling, and Chunmao Wang. 1997. “Basic processing unit of Chinese character recognition: Evidence from stroke number effect and radical number effect.” *Acta Psychologica Sinica* 29 (1): 9.

Peng, Danling, H. Yang, and Y. Chen. 1994. “Consistency and phonetic independency effects in naming task of Chinese phonograms.” In *Information processing of Chinese language,* edited by Q. C. Jing and H. C. Zhang. Beijing: Beijing Normal University Publishing.

Perdue, Clive. 1991. “Les Pronoms chez Paula.” In *Interactions en langue étrangère,* edited by Colette Russier, Henriette Stoffel, and Daniel Véronique. Aix-en-Provence: Publications de l’Université de Provence.

Perfetti, Charles A., and Susan R. Goldman. 1974. “Thematization and sentence retrieval.” *Journal of Verbal Learning and Verbal Behavior* 13 (1): 70–79.

Petrone, Caterina, Hubert Truckenbrodt, Caroline Wellmann, Julia Holzgrefe-Lang, Isabell Wartenburger, and Barbara Höhle. 2017. “Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists.” *Journal of Phonetics* 61 (Supplement C): 71–92.

Petten, Cyma Van, and Marta Kutas. 1990. “Interactions between sentence context and word frequencyinevent-related brainpotentials.” *Memory & Cognition* 18 (4): 380–393.

Peyraube, Alain. 2003. “On the history of place words and localizers in Chinese: a cognitive approach.” In *Functional Structure(s), Form and Interpretation: Perspectives from East Asian Languages,* edited by Yen-Hui Audrey Li and Andrew Simpson, 3–27. London: RoutledgeCurzon.

Peyraube, Alain, and Marie-Claude Paris. 1993. “L’iconicité : un nouveau dogme de la syntaxe chinoise ?” *Faits de langues* 1 (1): 69–78.

PhD, Douglas Medin. 2004. *Stevens’ Handbook of Experimental Psychology, Memory and Cognitive Processes.* John Wiley & Sons, February 5.

Philipp, Markus, Ina Bornkessel-Schlesewsky, Walter Bisang, and Matthias Schlesewsky. 2008. “The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials.” *Brain and Language* 105 (2): 112–133.

Phillips, Colin, and Shevaun Lewis. 2013. “Derivational order in syntax: Evidence and architectural consequences.” *Studies in Linguistics* 6:11–47.

Pickering, Martin J., and Victor S. Ferreira. 2008. “Structural Priming: A Critical Review.” *Psychological bulletin* 134 (3): 427–459.

Pierrehumbert, Janet Breckenridge. 1980. “The phonology and phonetics of English intonation (unpublished).” PhD diss., Massachusetts Institute of Technology.

Pierrehumbert, Janet Breckenridge, and Julia Hirschberg. 1992. “The Meaning of Intonational Contours in the Interpretation of Discourse.” In *Intentions in Communication,* edited by Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, 271–311. Cambridge, Massachusetts: MIT Press.

Poarch, Gregory J., and Ellen Bialystok. 2015. "Bilingualism as a model for multitasking." *Developmental Review,* Special Issue: Living in the "Net"Generation: Multitasking, Learning, and Development, 35:113–124.

Poeppel, David, Karen Emmorey, Gregory Hickok, and Liina Pylkkänen. 2012. "Towards a new neurobiology of language." *The Journal of Neuroscience* 32 (41): 14125–14131.

Poldrack, Russell A. 2006. "Can cognitive processes be inferred from neuroimaging data?" *Trends in Cognitive Sciences* 10 (2): 59–63.

Polinsky, Maria, and Eric Potsdam. 2001. "Long-Distance Agreement And Topic In Tsez." *Natural Language & Linguistic Theory* 19 (3): 583–646.

Pollard, Carl. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Prévost, Sophie. 1998. "La notion de thème : flou terminologique et conceptuel." *Cahiers de praxématique* 30:1–35.

———. 2003a. "Détachement et topicalisation : des niveaux d'analyse différents." *Cahiers de praxématique,* no. 40: 97–126.

———. 2003b. "Les compléments spatiaux : du topique au focus en passant par les cadres." *Travaux de linguistique* 47 (2): 51–77.

Price, Cathy J. 2013. "Current themes in neuroimaging studies of reading." *Brain and Language,* The Neural Basis of Reading, 125 (2): 131–133.

Primus, Beatrice. 1999a. *Cases and Thematic Roles: Ergative, Accusative and Active.* Tübingen: Niemeyer.

———. 1999b. *Cases and Thematic Roles, Ergative, Accusative and Active.* Tübingen: Niemeyer.

Prince, Ellen F. 1981. "Toward a taxonomy of given-new information." In *Syntax and semantics: Vol. 14. Radical Pragmatics,* edited by Peter Cole, 223–255. New York: Academic Press.

———. 1985. "Fancy syntax and 'shared knowledge'." *Journal of Pragmatics* 9 (1): 65–81.

———. 1992. "The ZPG Letter: Subjects, Definiteness, and Information-status." In *Discourse Description: Diverse Analyses of a Fund Raising Text,* edited by William C. Mann and Sandra A. Thompson, 16:295. Pragmatics & Beyond New Series. Amsterdam: John Benjamins Publishing Company.

Pulleyblank, Edwin G. 1995. "Chap. 8 Topicalization and exposure." In *Outline of Classical Chinese Grammar,* 68–75. Vancouver: UBC Press.

Qi, Zhenghan, Michelle Han, Keri Garel, Ee San Chen, and John D. E. Gabrieli. 2015. "White-matter structure in the right hemisphere predicts Mandarin Chinese learning success." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:14–28.

Qu, Yanfeng. 1993. "Locative Inversion in Mandarin Chinese." *Canadian Journal of Linguistics/Revue canadienne de linguistique* 38 (3): 305–330.

———. 1994. “Object Noun Phrase Dislocation in Mandarin Chinese.” PhD diss., University of British Columbia.

Quine, Willard van Orman. 1987. “Indeterminacy of Translation Again.” *The Journal of Philosophy* 84 (1): 5–10.

Raettig, Tim. 2010. “The cortical infrastructure of language processing: Evidence from functional and anatomical neuroimaging.” PhD diss., MPI for Human Cognitive and Brain Sciences.

Raettig, Tim, Stefan Frisch, Angela D. Friederici, and Sonja A. Kotz. 2010. “Neural correlates of morphosyntactic and verb-argument structure processing: An EfMRI study.” *Cortex* 46 (5): 613–620.

Ramchand, Gillian, Charles Reiss, and Sara Thomas Rosen, eds. 2007. “Structured Events, Structured Discourse.” In *The Oxford Handbook of Linguistic Interfaces,* 181–268. Oxford: Oxford University Press.

Ramer, Andrya L. H. 1976. “Syntactic styles in emerging language.” *Journal of Child Language* 3 (1): 49–62.

Ramsay, Allan. 1993. “Topicalisation and Attachment Preferences.” In *AI and Cognitive Science ’91,* edited by Humphrey Sorensen, redacted by C. J. van Rijsbergen, 190–198. London: Springer.

Reinhart, Tanya. 1981. “Pragmatics and Linguistics: An Analysis of Sentence Topics in Pragmatics and Philosophy I.” *Philosophica anc Studia Philosophica Gandensia Gent* 27 (1): 53–94.

Repp, Sophie. 2009. “Topic marking pays off. Corpus and experimental evidence.” University of Stuttgart.

Riemsdijk, Henk van, ed. 1999. “Italian/Romance clitics: Structure and derivation.” In *Clitics in the Languages of Europe,* 543–579. Berlin: De Gruyter.

Ritter, Elizabeth, and Sara Thomas Rosen. 2005a. “Agreement without A-Positions: Another Look at Algonquian.” *Linguistic Inquiry* 36 (4): 648–660.

———. 2005b. “Topic or Aspect.” In *Aspectual Inquiries,* edited by Paula Kempschinsky and Roumyana Slabakova, 21–39. Studies in Natural Language and Linguistic Theory. Springer, Dordrecht.

Rizzi, Luigi. 1990. *Relativized minimality.* Vol. x. Linguistic inquiry monographs, 16. Cambridge, MA, US: The MIT Press.

———. 1997. “The Fine Structure of the Left Periphery.” In *Elements of Grammar. Handbook of generative syntax,* edited by Liliane Haegeman, 281–337. Dordrecht: Kluwer Academic.

———. 2001. “Relativized Minimality Effects.” In *The Handbook of Contemporary Syntactic Theory,* edited by Mark Baltin and Chris Collins, 89–110. Oxford: Blackwell Publishers Ltd.

Rizzi, Luigi. 2004a. "Locality and Left Periphery." In *Structures and Beyond: The Cartography of Syntactic Structures, Volume 3,* edited by Adriana Belletti, 3:223–251. Oxford Studies in Comparative Syntax 3. Oxford, New York: Oxford University Press, July 29.

———, ed. 2004b. *The Structure of CP and IP: The Cartography of Syntactic Structures: Volume 2.* Oxford studies in comparative syntax. New York: Oxford University Press.

———. 2013a. "Introduction: Core computational principles in natural language syntax." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:1–13.

———. 2013b. "Locality." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:169–186.

———. 2015. "Notes on labeling and subject positions." In *Structures, Strategies and Beyond: Studies in honour of Adriana Belletti,* edited by Elisa Di Domenico, Cornelia Hamann, and Simona Matteini, 223:17–46. Linguistik Aktuell/Linguistics Today. Amsterdam: John Benjamins Publishing Company.

Rizzi, Luigi, and Ian Roberts. 1989. "Complex inversion in French." *Probus* 1 (1): 1–30.

Roark, Brian, Asaf Bachrach, Carlos Cardenas, and Christophe Pallier. 2009. "Deriving Lexical and Syntactic Expectation-based Measures for Psycholinguistic Modeling via Incremental Top-down Parsing." In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing,* 1:324–333. EMNLP '09. Stroudsburg, PA, USA: Association for Computational Linguistics.

Robert, Stéphane. 1993. "Structure et sémantique de la focalisation." *Bulletin de la Société de Linguistique de Paris* 88 (1): 25–47.

Roberts, Craige. 1997. "Focus, the flow of information, and universal grammar." In *The Limits of Syntax,* edited by Peter W. Culicover and Louise McNally, 109–160. Syntax and semantics 29. San Diego, CA: Academic Press.

Roberts, Ian. 2017. *The Oxford Handbook of Universal Grammar.* Oxford University Press, January 12.

Rogalsky, Corianne, Diogo Almeida, Jon Sprouse, and Gregory Hickok. 2015. "Sentence processing selectivity in Broca's area: evident for structure but not syntactic movement." *Language, Cognition and Neuroscience* 0 (0): 1–13.

Rogalsky, Corianne, and Gregory Hickok. 2009. "Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex." *Cerebral Cortex (New York, N.Y.: 1991)* 19 (4): 786–796.

———. 2010. "The role of Broca's area in sentence comprehension." *Journal of Cognitive Neuroscience* 23 (7): 1664–1680.

Rogalsky, Corianne, William Matchin, and Gregory Hickok. 2008. "Broca's Area, Sentence Comprehension, and Working Memory: An fMRI Study." *Frontiers in Human Neuroscience* 2.

Rollinger, Robin. 2014. "Anton Marty." In *The Stanford Encyclopedia of Philosophy,* Spring 2014, edited by Edward N. Zalta.

Rosenstein, Ofra. 2001. “Israeli sign language – a Topic prominent language.” MA Dissertation, Haifa University.

Rösler, Frank, Thomas Pechmann, Judith Streb, Brigitte Röder, and Erwin Hennighausen. 1998. “Parsing of Sentences in a Language with Varying Word Order: Word-by-Word Variations of Processing Demands Are Revealed by Event-Related Brain Potentials.” *Journal of Memory and Language* 38 (2): 150–176.

Rouveret, Alain. 2004. “Grammaire formelle et cognition linguistique.” In *La Linguistique cognitive,* edited by Catherine Fuchs. Cogniprisme. Paris: Ophrys.

Rouveret, Alain, and Philippe Schlenker. 1998. “La grammaire générative entre comparatisme et cognition.” *Langages* 32 (129): 9–23.

Ruigendijk, Esther, and Sergey Avrutin. 2003. “The comprehension of pronouns and reflexives in agrammatic and Wernicke’s aphasia.” *Brain and Language,* Academy of Aphasia Abstract Issue 2003, 87 (1): 17–18.

Ruigendijk, Esther, Naama Friedmann, Rama Novogrodsky, and Noga Balaban. 2010. “Symmetry in comprehension and production of pronouns: A comparison of German and Hebrew.” *Lingua,* Asymmetries in Language Acquisition, 120 (8): 1991–2005.

Ruigendijk, Esther, Petra Schumacher, and Sergey Avrutin. 2005. “Slow sentence processing in agrammatic Broca’s aphasia: Evidence from Dutch reflexive-antecedent dependencies.” *Brain and Language* 95:84–85.

Ruigendijk, Esther, Nada Vasić, and Sergey Avrutin. 2006. “Reference assignment: Using language breakdown to choose between theoretical approaches.” *Brain and Language* 96 (3): 302–317.

Rygaloff, Alexis. 1973. *Grammaire élémentaire du chinois.* Paris: Presses Universitaires de France.

Saito, Mamoru. 2007. “Notes on East Asian Argument Ellipsis.” *Language Research* 43:203–227.

Sakai, Kuniyoshi L., Fumitaka Homae, and Ryuichiro Hashimoto. 2003. “Sentence processing is uniquely human.” *Neuroscience Research* 46 (3): 273–279.

Salzmann, Martin. 2011. “Towards a Typology of Locative Inversion - Bantu, Perhaps Chinese and English - But Beyond?” *Language and Linguistics Compass* 5 (4): 169–189.

Sanders, Lisa D., Elissa L. Newport, and Helen J. Neville. 2002. “Segmenting Nonsense: An Event-Related Potential Index of Perceived Onsets in Continuous Speech.” *Nature neuroscience* 5 (7): 700–703.

Santi, Andrea, and Yosef Grodzinsky. 2007a. “Taxing working memory with syntax: Bihemispheric modulations.” *Human Brain Mapping* 28 (11): 1089–1097.

———. 2007b. “Working memory and syntax interact in Broca’s area.” *NeuroImage* 37 (1): 8–17.

———. 2010. “fMRI adaptation dissociates syntactic complexity dimensions.” *NeuroImage* 51 (4): 1285–1293.

Santi, Andrea, and Yosef Grodzinsky. 2012. "Broca's area and sentence comprehension: A relationship parasitic on dependency, displacement or predictability?" *Neuropsychologia* 50 (5): 821–832.

Sapir, Edward. 1991. *Linguistique.* Translated by Jean-Elie Boltanski and Nicole Soulé-Susbielles. Folio Essais 149. Paris: Gallimard.

Sasse, Hans-Jürgen. 1987. "The thetic/categorical distinction revisited." *Linguistics* 25 (3): 511–580.

Saxe, Rebecca, and Nancy Kanwisher. 2003. "People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind"." *NeuroImage* 19 (4): 1835–1842.

Schenker, Natalie M., William D. Hopkins, Muhammad A. Spocter, Amy R. Garrison, Cheryl D. Stimpson, Joseph M. Erwin, Patrick R. Hof, and Chet C. Sherwood. 2010. "Broca's Area Homologue in Chimpanzees (Pan troglodytes): Probabilistic Mapping, Asymmetry, and Comparison to Humans." *Cerebral Cortex (New York, NY)* 20 (3): 730–742.

Schlesewsky, Matthias, and Ina Bornkessel. 2004. "On incremental interpretation: degrees of meaning accessed during sentence comprehension." *Lingua,* Variation in form versus variation in meaning, 114 (9): 1213–1234.

Schlesewsky, Matthias, Ina Bornkessel, and Stefan Frisch. 2003. "The neurophysiological basis of word order variations in German." *Brain and Language* 86 (1): 116–128.

Schlesewsky, Matthias, Ina Bornkessel, and Martin Meyer. 2002. "Why a "word order difference" is not always a "word order" difference: a reply to Weyerts, Penke, Münte, Heinze, and Clahsen." *Journal of Psycholinguistic Research* 31 (5): 437–445.

Schlobinski, Peter, and Stephan Schütze-Coburn. 1992. "On the topic of topic and topic continuity." *Linguistics* 30 (1): 89–122.

Schubotz, Ricarda I., and D. Yves von Cramon. 2004. "Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 24 (24): 5467–5474.

Schuchard, Julia, Michaela Nerantzini, and Cynthia K. Thompson. 2017. "Implicit learning and implicit treatment outcomes in individuals with aphasia." *Aphasiology* 31 (1): 25–48.

Schumacher, Eric H., Erick Lauber, Edward Awh, John Jonides, Edward E. Smith, and Robert A. Koeppe. 1996. "PET Evidence for an Amodal Verbal Working Memory System." *NeuroImage* 3 (2): 79–88.

Schumacher, Petra B. 2009. "Definiteness Marking Shows Late Effects during Discourse Processing: Evidence from ERPs." In *Anaphora Processing and Applications,* edited by Sobha Lalitha Devi, António Branco, and Ruslan Mitkov, 91–106. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, November 5.

Schumacher, Petra B., Jana Backhaus, and Manuel Dangl. 2015. "Backward- and Forward-Looking Potential of Anaphors." *Frontiers in Psychology* 6:1746.

Schumacher, Petra B., and Stefan Baumann. 2010. “Pitch accent type affects the N400 during referential processing.” *NeuroReport* 21 (9): 618–622.

Schwarz-Friese, Monika, Manfred Consten, and Mareile Knees. 2007. *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference.* John Benjamins Publishing, January 1.

Segaert, Katrien, Gerard Kempen, Karl Magnus Petersson, and Peter Hagoort. 2013. “Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: an fMRI study.” *Brain and Language* 124 (2): 174–183.

Shao 邵, Jingmin 敬敏. 1989. “语气词 “呢” 在疑问句中的作用 Yuqici ‘ne’ zai yiwenju zhong de zuoyong [The Function of the Expressive Particle ‘ne’ in Interrogatives].” *中国语文 Zhongguo Yuwen [Studies of the Chinese Language],* no. 3: 170–175.

Shapiro, Kevin A., Lauren R. Moo, and Alfonso Caramazza. 2012. “Neural Specificity for Grammatical Operations is Revealed by Content-Independent fMR Adaptation.” *Frontiers in Psychology* 3.

Shapiro, Lewis P., Betsy Gordon, N. Hack, and J. Killackey. 1993. “Verb-Argument Structure Processing in Complex Sentences in Broca s and Wernicke s Aphasia.” *Brain and Language* 45 (3): 423–447.

Shen 沈, Jiaxuan 家煊. 2006. ““王冕死了父亲” 的生成方式——兼说汉语 “糅合” 造句 [The Generative mechanism of sentences like “Wangmian died father” - Sentence generation by blending in Chinese].” *[中国语文 - Zhongguo Yuwen] Studies of the Chinese Language* 4:291–300.

Shen, Ming-yueh. 2011. “Errors in the Written Production of Chinese Learners of English: A Typological Perspective.” *International Journal of Linguistics* 3 (1): 8.

Shen, Xiao-Nan Susan. 1990. *The Prosody of Mandarin Chinese.* Publications in Linguistics 118. Berkeley: University of California Press.

Shetreet, Einat, and Naama Friedmann. 2012. “Stretched, jumped, and fell: an fMRI investigation of reflexive verbs and other intransitives.” *NeuroImage* 60 (3): 1800–1806.

———. 2014. “The processing of different syntactic structures: fMRI investigation of the linguistic distinction between wh-movement and verb movement.” *Journal of Neurolinguistics* 27 (1): 1–17.

Shetreet, Einat, Naama Friedmann, and Uri Hadar. 2009a. “An fMRI study of syntactic layers: sentential and lexical aspects of embedding.” *NeuroImage* 48 (4): 707–716.

———. 2009b. “Cortical representation of unaccusative verbs.” *Language and Brain* 8:21–29.

———. 2010a. “Cortical representation of verbs with optional complements: the theoretical contribution of fMRI.” *Human Brain Mapping* 31 (5): 770–785.

———. 2010b. “The neural correlates of linguistic distinctions: unaccusative and unergative verbs.” *Journal of Cognitive Neuroscience* 22 (10): 2306–2315.

Shetreet, Einat, Tal Linzen, and Naama Friedmann. 2016. "Against all odds: exhaustive activation in lexical access of verb complementation options." *Language, Cognition and Neuroscience* 31 (9): 1206–1214.

Shetreet, Einat, Dafna Palti, Naama Friedmann, and Uri Hadar. 2007. "Cortical representation of verb processing in sentence comprehension: number of complements, subcategorization, and thematic frames." *Cerebral Cortex* 17 (8): 1958–1969.

Shi, Dingxu. 1989. "Topic chain as a syntactic category in Chinese." *Journal of Chinese Linguistics* 17:223–262.

———. 1992. "The Nature of Topic Comment Construction and Topic Chains." PhD diss., University of Southern California.

———. 2000. "Topic and Topic-Comment Constructions in Mandarin Chinese." *Language* 76 (2): 383–408.

Shopen, Timothy. 2007. *Language Typology and Syntactic Description: Volume 2, Complex Constructions.* 2nd ed. Cambridge University Press.

Shu, Chih-hsiang. 2011. "Sentence Adverbs in the Kingdom of Agree." PhD diss., Stony Brook University.

Shu, Hua, and Richard C. Anderson. 1999. "Learning to read Chinese: The development of metalinguistic awareness." In *Reading Chinese script: A cognitive analysis,* edited by J. Wang, A. W. Inhoff, and H.-C. Chen, 1–18. Mahwah, NJ: Lawrence Erlbaum Associates.

Shu, Hua, Xi Chen, Richard C. Anderson, Ningning Wu, and Yue Xuan. 2003. "Properties of school Chinese: implications for learning to read." *Child Development* 74 (1): 27–47.

Shyu 徐, Shu-ing 淑英. 2004. "(A)symmetries between Mandarin Chinese Lian-dou and Shenzhi / 「连…都」及「甚至」的探讨." *Journal of Chinese Linguistics* 32 (1): 81–128.

Shyu, Shu-ing. 1995. "The Syntax of Focus and Topic in Mandarin Chinese." PhD diss., University of Southern California.

———. 2001. "Remarks on Object Movement in Mandarin SOV Order." *Language and Linguistics* 2 (1): 93–124.

Shyu, Shu-Ing. 2014. "Topic and Focus." In *The Handbook of Chinese Linguistics,* edited by Cheng-Teh James Huang, Yen-Hui Audrey Li, and Andrew Simpson, 100–125. Hoboken, USA: John Wiley & Sons, Inc., February 21.

Siblot, Paul. 1998. "Variations sur le thème qui 'toujours déjà' prédique." *Cahiers de praxématique* 30:37–53.

Siewierska, Anna. 1993. "On the interplay of factors in the determination of word order." In *Syntax: An Internatinal Handbook of Contemporary Research,* edited by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 826–846. Handbücher zur Sprach- und Kommunikationswissenschaft ; Handbooks of linguistics and communication science. Berlin: Walter de Gruyter.

———. 1994. "Word order and linearization." In *The Encyclopedia of Language and Linguistics,* edited by R. E. Asher and J. M. Y. Simpson, 9:4993–4999. Oxford: Pergamon Press.

———. 1995. "Doris L. Payne (ed.), Pragmatics of word order flexibility. Amsterdam: John Benjamins, 1992. Pp. 320." *Journal of Linguistics* 31 (2): 458–463.

———. 2001. "Word order." In *International Encyclopedia of the Social & Behavioral Sciences,* edited by Neil J. Smelser and Paul B. Baltes, 16552–16555. Amsterdam: Elsevier.

Siok, Wai Ting, and Paul Fletcher. 2001. "The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition." *Developmental Psychology* 37 (6): 886–899.

Slobin, Dan I., and Thomas G. Bever. 1982. "Children use canonical sentence schemas: A crosslinguistic study of word order and inflections." *Cognition* 12 (3): 229–265.

Slobin, Dan Isaac, ed. 1992. "The acquisition of Mandarin." In *The Crosslinguistic Study of Language Acquisition,* 3:373–455. Hillsdale, NJ: Lawrence Erlbaum Associates.

Smaers, Jeroen B., Aida Gómez-Robles, Ashley N. Parks, and Chet C. Sherwood. 2017. "Exceptional Evolutionary Expansion of Prefrontal Cortex in Great Apes and Humans." *Current Biology* 27 (5): 714–720.

Smith, Edward E., and John Jonides. 1999. "Storage and Executive Processes in the Frontal Lobes." *Science* 283 (5408): 1657–1661.

Sorace, Antonella, and Francesca Filiaci. 2006. "Anaphora resolution in near-native speakers of Italian." *Second Language Research* 22 (3): 339–368.

Sperber, Dan, and Deirdre Wilson. 1986. *Relevance: Communication and Cognition.* Cambridge, MA: Harvard University Press.

Spinelli, Elsa. 2005. *Psychologie du langage: L'écrit et le parlé, du signal à la signification.* Armand Colin, September 1.

Sportiche, Dominique. 2013. "Binding theory—Structure sensitivity of referential dependencies." *Lingua,* SI: Syntax and cognition: core ideas and results in syntax, 130:187–208.

Spreng, R. Nathan, W. Dale Stevens, Jon P. Chamberlain, Adrian W. Gilmore, and Daniel L. Schacter. 2010. "Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition." *NeuroImage* 53 (1): 303–317.

St George, Marie, Marta Kutas, A. Martinez, and Marty I. Sereno. 1999. "Semantic integration in reading: engagement of the right hemisphere during discourse processing." *Brain: A Journal of Neurology* 122 (7): 1317–1325.

Staub, Adrian, and Keith Rayner. 2007. "Eye movements and on-line comprehension processes." In *The Oxford Handbook of Psycholinguistics,* edited by Gareth Gaskell, 327–342. Oxford: Oxford University Press, August 2.

Stavrakaki, Stavroula. 2001. "Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children." *Brain and Language* 77 (3): 419–431.

Stavrakaki, Stavroula, and Heather van der Lely. 2010. "Production and comprehension of pronouns by Greek children with specific language impairment." *The British Journal of Developmental Psychology* 28 (Pt 1): 189–216.

Steinhauer, Karsten, Kai Alter, and Angela D. Friederici. 1999. "Brain potentials indicate immediate use of prosodic cues in natural speech processing." *Nature Neuroscience* 2 (2): 191–196.

Steinhauer, Karsten, and John F. Conolly. 2008. "Event-Related Potentials in the Study of Language." In *Handbook of the Neuroscience of Language,* edited by Brigitte Stemmer and Harry A. Whitaker, 91–104. London: Academic Press.

Stowe, Laurie A. 1986. "Parsing WH-constructions: Evidence for on-line gap location." *Language and Cognitive Processes* 1 (3): 227–245.

Stowe, Laurie A., Cees A. J. Broere, Anne M. J. Paans, Albertus A. Wijers, Gijsbertus Mulder, Wim Vaalburg, and Frans Zwarts. 1998. "Localizing components of a complex task: sentence processing and working memory." *Neuroreport* 9 (13): 2995–2999.

Streb, Judith, Erwin Hennighausen, and Frank Rösler. 2004. "Different Anaphoric Expressions Are Investigated by Event-Related Brain Potentials." *Journal of Psycholinguistic Research* 33 (3): 175–201.

Stromswold, Karin, David Caplan, Nathaniel Alpert, and Scott Rauch. 1996. "Localization of Syntactic Comprehension by Positron Emission Tomography." *Brain and Language* 52 (3): 452–473.

Su, Mengmeng, Jiuju Wang, Urs Maurer, Yuping Zhang, Jun Li, Catherine McBride, Twila Tardif, Youyi Liu, and Hua Shu. 2015. "Gene – environment interaction on neural mechanisms of orthographic processing in Chinese children." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:172–186.

Suh, Soyoung, Hyo Woon Yoon, Seungbok Lee, Jun-Young Chung, Zang-Hee Cho, and HyunWook Park. 2007. "Effects of syntactic complexity in L1 and L2; An fMRI study of Korean–English bilinguals." *Brain Research* 1136:178–189.

Sun, Chaofen. 2006. *Chinese: a linguistic introduction.* Cambridge: Cambridge University Press.

Sun, Chao-Fen, and Talmy Givón. 1985. "On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications." *Language* 61 (2): 329–351.

Sussman, Rachel Shirley, and Julie Sedivy. 2003. "The time-course of processing syntactic dependencies: Evidence from eye movements." *Language and Cognitive Processes* 18 (2): 143–163.

Swart, Henriette de, and Helen de Hoop. 1996. "Topic and Focus." *GLOT International* 7:3–7.

Swinney, D., and E. Zurif. 1995. “Syntactic Processing in Aphasia.” *Brain and Language* 50 (2): 225–239.

Sybesma, Rint. 1996. “Review of the Syntax of Focus and Topic in Mandarin Chinese by Shu-ing Shyu.” *GLOT International* 2 (4): 12–14.

Sybesma, Rint, and Boya Li. 2007. “The dissection and structural mapping of Cantonese sentence final particles.” *Lingua* 117 (10): 1739–1783.

Sze, Felix. 2008. “Topic Constructions in Hong-Kong Sign Language.” PhD diss., University of Bristol.

Tai, James H.-Y. 1973. “On the center of predication in Chinese verb-complement construction.” San Diego, CA.

———. 1984. “Word order in Chinese and the X theory.” *Journal of the Chinese Language Teachers Association* 19 (2): 23–36.

———. 1989. “Toward a cognition-based functional grammar of Chinese.” In *Functionalism and Chinese Grammar,* edited by James H.-Y. Tai and Frank F.S. Hsueh, 186–226. Chinese Teachers Association Monograph Series 1. South Orange, NJ: Chinese Language Teachers Association.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics.* Cambridge, MA: MIT.

———. 2005. “The fundamental system of spatial schemas in language.” In *From Perception to Meaning: Image Schemas in Cognitive Linguistics,* edited by Beate Hampe, 199–234. Berlin: Mouton de Gruyter.

Tan, Li Hai, and Ping Li. 2015. “Towards an integrative understanding of the neuroanatomical and genetic bases of language: The Chinese context.” *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:1–2.

Tanenhaus, Michael K., and John C. Trueswell. 1995. “Sentence comprehension.” In *Speech, Language, and Communication,* edited by Joanne L. Miller and Peter D. Eimas, 217–262. Handbook of perception and cognition (2nd ed.), Vol. 11. San Diego, CA: Academic Press.

Tang 湯, Tingchi 廷池. 1979. 國語語法研究論集. 台北: 學生書局.

Teng, Shou-Hsin. 1974. “Double Nominatives in Chinese.” *Language* 50 (3): 455–473.

Tettamanti, Marco, Rosa Manenti, Pasquale A. Della Rosa, Andrea Falini, Daniela Perani, Stefano F. Cappa, and Andrea Moro. 2008. “Negation in the brain: modulating action representations.” *NeuroImage* 43 (2): 358–367.

Tettamanti, Marco, Irene Rotondi, Daniela Perani, Giuseppe Scotti, Ferruccio Fazio, Stefano F. Cappa, and Andrea Moro. 2009. “Syntax without language: neurobiological evidence for cross-domain syntactic computations.” *Cortex* 45 (7): 825–838.

Thompson, Cynthia K. 2003. “Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis.” *Journal of Neurolinguistics,* Objects vs. actions and nouns vs. verbs, 16 (2): 151–167.

Thompson, Cynthia K., Kirrie J. Ballard, and Lewis P. Shapiro. 1998. "The role of syntactic complexity in training wh-movement structures in agrammatic aphasia: optimal order for promoting generalization." *Journal of the International Neuropsychological Society: JINS* 4 (6): 661–674.

Thompson, Cynthia K., JungWon Janet Choy, Audrey Holland, and Ronald Cole. 2010. "Sentactics®: Computer-Automated Treatment of Underlying Forms." *Aphasiology* 24 (10): 1242–1266.

Thompson, Cynthia K., Dirk-Bart den Ouden, Borna Bonakdarpour, Kyla Garibaldi, and Todd B. Parrish. 2010. "Neural plasticity and treatment-induced recovery of sentence processing in agrammatism." *Neuropsychologia* 48 (11): 3211–3227.

Thompson, Cynthia K., and Lewis P. Shapiro. 2005. "Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms." *Aphasiology* 19 (10): 1021–1036.

———. 2007. "Complexity in Treatment of Syntactic Deficits." *American journal of speech-language pathology / American Speech-Language-Hearing Association* 16 (1): 30–42.

Thompson, Cynthia K., Lewis P. Shapiro, Kirrie J. Ballard, Beverly J. Jacobs, Sandra S. Schneider, and Mary E. Tait. 1997. "Training and generalized production of wh- and NP-movement structures in agrammatic aphasia." *Journal of speech, language, and hearing research: JSLHR* 40 (2): 228–244.

Thompson, Cynthia K., Lewis P. Shapiro, Swathi Kiran, and Jana Sobecks. 2003. "The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: the complexity account of treatment efficacy (CATE)." *Journal of speech, language, and hearing research: JSLHR* 46 (3): 591–607.

Thothathiri, Malathi, Albert Kim, John C. Trueswell, and Sharon L. Thompson-Schill. 2012. "Parametric effects of syntactic-semantic conflict in Broca's area during sentence processing." *Brain and Language* 120 (3): 259–264.

Ting 丁, Jen 仁, and Yu-Chi 玉琪 Huang 黃. 2008. "Some Remarks on Parasitic Gaps in Chinese [論中文寄生空位結構]." *Concentric: Studies in Linguistics* 34 (1): 27–52.

Tomasello, Michael. 1988. "The role of joint attentional processes in early language development." *Language Sciences* 10 (1): 69–88.

Tomlin, Russell S. 1985. "Foreground-background information and the syntax of subordination." *Text* 5 (1): 85–122.

———. 1986. *Basic word order. Functional principles.* London: Croom Helm.

Townsend, David J., and Thomas G. Bever. 2001. *Sentence Comprehension: The Integration of Habits and Rules.* Cambridge, MA: MIT Press.

Treiman, Rebecca, and Andrea Zukowski. 1991. "Levels of phonological awareness. In." In *Phonological Processes in Literacy: A Tribute to Isabelle P. Liberman.* Edited by S.A. Brady and D. P. Shankweiler, 67–83. Hillsdale, NJ: Erlbaum.

Tsao, Feng-Fu. 1977. "A Functional Study of Topic in Chinese: the First Step towards Discourse Analysis." PhD diss., University of Southern California.

———. 1990. *Sentence and Clause Structure in Chinese: a functional perspective.* Taipei: Student Book Co.

Uchiyama, Yuji, Hiroshi Toyoda, Manabu Honda, Haruyo Yoshida, Takanori Kochiyama, Kazutoshi Ebe, and Norihiro Sadato. 2008. "Functional segregation of the inferior frontal gyrus for syntactic processes: A functional magnetic-resonance imaging study." *Neuroscience Research* 61 (3): 309–318.

Ullmann, Stephen. 1952. "Inversion as a Stylistic Device in the Contemporary French Novel." *The Modern Language Review* 47 (2): 165–180.

Vallduví, Enric. 1992. *The Informational Component.* New York: Garland Publishing.

Van Valin, Robert D., and Randy J. LaPolla. 1997. *Syntax. Structure, Meaning and Function.* Cambridge: Cambridge University Press.

Vandenberghe, Rik, Anna Christina Nobre, and Cathy J. Price. 2002. "The response of left temporal cortex to sentences." *Journal of Cognitive Neuroscience* 14 (4): 550–560.

Varlokosta, Spyridoula, Adriana Belletti, João Costa, Naama Friedmann, Anna Gavarró, Kleanthes K. Grohmann, Maria Teresa Guasti, et al. 2015. "A cross-linguistic study of the acquisition of clitic and pronoun production." *Language Acquisition* 23:1–26.

Vermeulen, Reiko. 2013. "On the position of topics in Japanese." *The Linguistic Review* 30 (1): 117–159.

Véronis, Jean, and Emilie Guimier. 2006. "Le traitement des nouvelles formes de communication écrite." In *Compréhension Automatique des Langues et Interaction,* edited by Gérard Sabah, 227–248. Paris: Hermès Science Publications.

Villiers, Jill G. de, and Peter A. de Villiers. 1978. *Language Acquisition.* Cambridge, MA: Harvard University Press.

Vos, Sandra H., Thomas C. Gunter, Herman H.j. Kolk, and Gijsbertus Mulder. 2001. "Working memory constraints on syntactic processing: An electrophysiological investigation." *Psychophysiology* 38 (1): 41–63.

Wagers, Matthew W., and Colin Phillips. 2009. "Multiple dependencies and the role of the grammar in real-time comprehension." *Journal of Linguistics* 45 (2): 395–433.

Wagner, Michael. 2010. "Prosody and Incremental Structure Building." University of Ulster, Belfast, October 1.

Wakefield, John C. 2011. "The English equivalents of Cantonese sentence-final particles : a contrastive analysis." Thesis, The Hong Kong Polytechnic University.

Wang 王, Li 力. 1958. 汉语史稿 *Hanyu shi gao.* Beijing: 中华书局 Zhonghua shu ju.

Wang, Bei, and Yi Xu. 2017. "Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese." *Journal of Phonetics* 39 (4): 595–611.

Wang, Guoqing, Yuxi Zhou, Yong Gao, Huan Chen, Jiguang Xia, Junquan Xu, Michael S. Y. Huen, et al. 2015. "Association of specific language impairment candidate genes CMIP and ATP2C2 with developmental dyslexia in Chinese population." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:163–171.

Wang, Li. 1956. "A definition of subject and its application in Chinese." *Yuwen Huiban* 9:169–180.

Wang, Luming, Matthias Schlesewsky, Markus Philipp, and Ina Bornkessel-Schlesewsky. 2012. "The Role of Animacy in Online Argument Interpretation in Mandarin Chinese." In *Case, Word Order and Prominence,* edited by Monique Lamers and Peter de Swart, 40:91–119. Studies in Theoretical Psycholinguistics. Dordrecht: Springer.

Wang, Luming, and Petra B. Schumacher. 2013. "New is not always costly: evidence from online processing of topic and contrast in Japanese." *Frontiers in Psychology* 4:363.

Wang, Xiaojuan, Jianfeng Yang, Hua Shu, and Jason D. Zevin. 2011. "Left fusiform BOLD responses are inversely related to word-likeness in a one-back task." *NeuroImage* 55 (3): 1346–1356.

Wang, Xiaojuan, Jianfeng Yang, Jie Yang, W. Einar Mencl, Hua Shu, and Jason David Zevin. 2015. "Language Differences in the Brain Network for Reading in Naturalistic Story Reading and Lexical Decision." *PLOS ONE* 10 (5): e0124388.

Warrington, Elizabeth K. 1975. "The selective impairment of semantic memory." *Quarterly Journal of Experimental Psychology* 27 (4): 635–657.

Wartenburger, Isabell, Hauke R Heekeren, Frank Burchert, Ria De Bleser, and Arno Villringer. 2003. "Grammaticality judgments on sentences with and without movement of phrasal constituents—an event-related fMRI study." *Journal of Neurolinguistics,* FUNCTIONAL NEUROIMAGING CONTRIBUTIONS TO NEUROLINGUISTICS, 16 (4): 301–314.

Watumull, Jeffrey, Marc D. Hauser, Ian G. Roberts, and Norbert Hornstein. 2014. "On recursion." *Frontiers in Psychology* 4.

Weber, Kirsten, and Peter Indefrey. 2009. "Syntactic priming in German-English bilinguals during sentence comprehension." *NeuroImage* 46 (4): 1164–1172.

Weckerly, Jill, and Marta Kutas. 1999. "An electrophysiological analysis of animacy effects in the processing of object relative sentences." *Psychophysiology* 36 (5): 559–570.

Wehbe, Leila, Brian Murphy, Partha Talukdar, Alona Fyshe, Aaditya Ramdas, and Tom Mitchell. 2014. "Simultaneously Uncovering the Patterns of Brain Regions Involved in Different Story Reading Subprocesses." *PLOS ONE* 9 (11): e112575.

Weil, Henri. 1844. *De l'Ordre des mots dans les langues anciennes comparées aux langues modernes.* Paris: Joubert.

Wellmann, Caroline, Julia Holzgrefe, Hubert Truckenbrodt, Isabell Wartenburger, and Barbara Höhle. 2012. "How Each Prosodic Boundary Cue Matters: Evidence from German Infants." *Frontiers in Psychology* 3.

Westerlund, Masha, Itamar Kastner, Meera Al Kaabi, and Liina Pylkkänen. 2015. "The LATL as locus of composition: MEG evidence from English and Arabic." *Brain and Language* 141:124–134.

Weyerts, Helga, Martina Penke, Thomas F. Münte, Hans-Jochen Heinze, and Harald Clahsen. 2002. "Word order in sentence processing: an experimental study of verb placement in German." *Journal of Psycholinguistic Research* 31 (3): 211–268.

Wilkinson, Ray. 1995. "Aphasia: conversation analysis of a non-fluent aphasic person." In *Case Studies In Clinical Linguistics,* edited by Mick Perkins and Sara Howard, 271–292. London: Whurr Publishers.

Wilkinson, Ray, Suzanne Beeke, and Jane Maxim. 2010. "Formulating Actions and Events With Limited Linguistic Resources: Enactment and Iconicity in Agrammatic Aphasic Talk." *Research on Language & Social Interaction* 43 (1): 57–84.

Willems, Roel M., and Peter Hagoort. 2009. "Broca's region: battles are not won by ignoring half of the facts." *Trends in Cognitive Sciences* 13 (3): 101–102.

Williams, Edwin. 1980. "Predication." *Linguistic Inquiry* 11 (1): 203–238.

Windfuhr, Kirsten L., and Margaret J. Snowling. 2001. "The Relationship between Paired Associate Learning and Phonological Skills in Normally Developing Readers." *Journal of Experimental Child Psychology* 80 (2): 160–173.

Wlodarczyk, Hélène, and André Wlodarczyk, eds. 2006. *La Focalisation dans les langues.* Sémantiques. Paris: L'Harmattan.

Wolff, Susann, Matthias Schlesewsky, Masako Hirotani, and Ina Bornkessel-Schlesewsky. 2008. "The neural mechanisms of word order processing revisited: electrophysiological evidence from Japanese." *Brain and Language* 107 (2): 133–157.

Wong, Becky, Bin Yin, and Beth O'Brien. 2016. "Neurolinguistics: Structure, Function, and Connectivity in the Bilingual Brain." *BioMed Research International* 2016.

Wong, Cara, and Jason Gallate. 2012. "The function of the anterior temporal lobe: a review of the empirical evidence." *Brain Research* 1449:94–116.

Wu 吴, Weizhang 为章. 1995. "语序重要." 中国语文, no. 6: 429–436.

Wu, Chiao-Yi, Moon-Ho Ringo Ho, and Shen-Hsing Annabel Chen. 2012. "A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing." *NeuroImage* 63 (1): 381–391.

Wu, Hsiao-hung Iris. 2013. "Experiencers as mental locations." EHESS - Paris, September.

Wydell, T. N., and B. Butterworth. 1999. "A case study of an English-Japanese bilingual with monolingual dyslexia." *Cognition* 70 (3): 273–305.

Xu 徐, Dan 丹. 1988. "浅谈这/那的不对称性 [Brief Overlook on Asymmetry of zhe and na]." 中国语文 *Zhongguo Yuwen [Studies of the Chinese Language]* 2:129–130.

Xu 徐, Feng 峰. 1998. "现代汉语置放动词配价研究 Xiandai Hanyu zhifang donci peijia yanjiu." 语言教学与研究, no. 3: 86–101.

Xu 徐, Jie 杰. 1999. “两种保留宾语句式及相关句法理论问题 Liang zhong baoliu binyu jushi ji xiangguan jufa lilun [Some theoretical issues of the two types of Chinese ‘retained-Object’constructions].” 当代语言学 *Dangdai yuyanxue [Contemporary Linguistics],* no. 1: 16–29.

———. 2005. 汉语研究的类型学视角 *[Studies in Chinese Linguistics: A typological perspective].* Beijing: BLCU Press [北京语言大学出版社].

Xu 徐, Jie 杰, and Yingzhe 英哲 Li 李. 1993. “焦点和两个非线性语法范畴 [否定][疑问].” 中国语文 *Zhongguo Yuwen [Studies of the Chinese Language],* no. 2: 81–92.

Xu 徐, Liejong 烈炯. 1996. “汉语语义研究的空白地带.” 中国语文 *Zhongguo Yuwen [Studies of the Chinese Language]* 4:255–259.

Xu 徐, Liejong 烈炯, and Danqing 丹青 Liu 刘. 1998. 话题的结构与功能 *Huati de jiegou yu gongneng.* Shanghai: Shanghai educational publishing house [上海教育出版社].

Xu, Dan. 2000. “Ordre des mots et actance dans la diachronie du chinois.” In *Cahier de linguistique de l’INALCO,* edited by Anaïd Donabédian and Dan Xu, 3:111–126. Paris: Langues’O (publications).

———. 2006. *Typological Change in Chinese Syntax.* Oxford: Oxford University Press.

———. 2008a. “Asymmetry in the Expression of Space in Chinese —The Chinese Language Meets Typology.” In *Space in Languages of China,* edited by Dan Xu, 175–198. Dordrecht: Springer.

———. 2008b. “Introduction: How Chinese Structures Space.” In *Space in Languages of China,* edited by Dan Xu, 1–14. Dordrecht: Springer.

———, ed. 2008c. *Space in Languages of China: cross-linguistic, synchronic and diachronic perspectives.* Dordrecht: Springer.

———. 2015. “Sinitic languages of Northwest China: Where did their case marking come from?” In *Languages in contact in Northwestern China,* edited by Alain Peyraube, Redouane Djamouri, and Guangshun Cao, 217–244. Collection des Cahiers de Linguistique Asie Orientale. Paris: Monograph series of the EHESS-CRLAO.

Xu, Guiping, Weifang Zhong, Hua Jin, and Lei Mo. 2015. “An ERP study on how subsequent sentence context can influence previous world knowledge constraints.” *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:96–103.

Xu, Jie. 2003. “Subject as a Constituent, Topic as a Feature and Their Respective Typologies.” *Linguistic Sciences* 2 (1): 3.

Xu, Liejiong. 1986. “Free Empty Category.” *Linguistic Inquiry* 17 (1): 75–93.

———. 1990. “Are they parasitic gaps?” In *Grammar in progress: Glow essays for Henk van Riemsdijk,* edited by Marina Nespor and Joan Mascaró, 455–461. Dordrecht: Foris Publications.

———. 1995. “Definiteness effects on Chinese word order.” *Cahiers de linguistique - Asie orientale* 24 (1): 29–48.

———. 1996. “Construction and destruction of theories by data : a case study.” In *CLS : proceedings from the meeting of the Chicago Linguistic Society,* 32:107–118. University of Chicago.

———. 2000. “The Topic-Prominence Parameter.” *ZAS Papers in Linguistics* 20:21–41.

———. 2006. “Topicalization in Asian Languages.” In *The Blackwell Companion to Syntax,* edited by rtin Everaert and Henk van Riemsdijk, 137–174. Hoboken, NJ: Blackwell Publishing.

Xu, Liejiong, and D. Terence Langendoen. 1985. “Topic Structures in Chinese.” *Language* 61 (1): 1–27.

Xu, Min, Jing Yang, Wai Ting Siok, and Li Hai Tan. 2015. “Atypical lateralization of phonological working memory in developmental dyslexia.” *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:67–77.

Yang 杨, Qiguang 启光. 2012. “汉语多项 NP 句主题句表意功能探析 Hanyu duoxiang NP ju zhutiju biaoyi gongneng tanxi [A deeper analysis of the referentiality/meaning of Chinese Topic multiple NP sentences].” 暨南大学中文系.

Yang, Chin Lung, Charles A. Perfetti, and Franz Schmalhofer. 2007. “Event-related potential indicators of text integration across sentence boundaries.” *Journal of Experimental Psychology. Learning, Memory, and Cognition* 33 (1): 55–89.

Yang, Jianfeng, Bruce D. McCandliss, Hua Shu, and Jason D. Zevin. 2009. “Simulating language-specific and language-general effects in a statistical learning model of Chinese reading.” *Journal of Memory and Language* 61 (2): 238–257.

Yang, Jianfeng, Xiaojuan Wang, Hua Shu, and Jason D. Zevin. 2011. “Brain networks associated with sublexical properties of Chinese characters.” *Brain and language* 119 (2): 68–79.

Yang, Jing, Kathleen Marie Gates, Peter Molenaar, and Ping Li. 2015. “Neural changes underlying successful second language word learning: An fMRI study.” *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:29–49.

Ye, Zheng, Yue-jia Luo, Angela D. Friederici, and Xiaolin Zhou. 2006. “Semantic and syntactic processing in Chinese sentence comprehension: evidence from event-related potentials.” *Brain Research* 1071 (1): 186–196.

Ye, Zheng, Weidong Zhan, and Xiaolin Zhou. 2007. “The semantic processing of syntactic structure in sentence comprehension: an ERP study.” *Brain Research* 1142:135–145.

Ye, Zheng, and Xiaolin Zhou. 2009. “Conflict control during sentence comprehension: fMRI evidence.” *NeuroImage* 48 (1): 280–290.

Yeatman, Jason D., Andreas M. Rauschecker, and Brian A. Wandell. 2013. “Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections.” *Brain and Language,* The Neural Basis of Reading, 125 (2): 146–155.

Yiming, Yang, and Liu Tao. 2014. “Neural Mechanisms of Trace in Chinese Topicalized Constructions.” *Social Sciences in China* 35 (1): 86–111.

Yokoyama, Satoru, Tadao Miyamoto, Jorge Riera, Jungho Kim, Yuko Akitsuki, Kazuki Iwata, Kei Yoshimoto, Kaoru Horie, Shigeru Sato, and Ryuta Kawashima. 2006. "Cortical Mechanisms Involved in the Processing of Verbs: An fMRI Study." *Journal of Cognitive Neuroscience* 18 (8): 1304–1313.

Yuan 袁, Yulin 毓林. 1996. "话题化及相关的语法过程 Huatihua ji xiangguan de yufa guocheng [Topicalization and grammatical process]." 中国语文 *Zhongguo Yuwen [Studies of the Chinese Language]* 4.

Yuan 袁, Yulin 毓林, Xiang 湘 Li 李, Hong 宏 Cao 曹, and Jian 健 Wang 王. 2009. "A Situational Semantic Analysis of Yǒu-constructions in Chinese." 世界漢語教學 *Shijie Hanyu Jiaoxue [Chinese Teaching in the word]* 3:291–307.

Zaccarella, E., and A. D. Friederici. 2017. "The neurobiological nature of syntactic hierarchies." *Neuroscience & Biobehavioral Reviews.*

Zaccarella, Emiliano, Lars Meyer, Michiru Makuuchi, and Angela D. Friederici. 2015. "Building by Syntax: The Neural Basis of Minimal Linguistic Structures." *Cerebral Cortex.*

Zhang 张, Bojiang 伯江, and Mei 梅 Fang 方. 1994. "汉语口语里的主位结构 Hanyu kouyuli de zhuwei jiegou." 北京大学学报 *Beijing daxue xuebao [Journal of Peking University],* no. 2.

———. 1996. 汉语功能语法研究 *Hanyu gongneng yufa yanjiu [Fonctional studies of Chinese grammar].* 江西教育出版社 Jiangxi Education Press. 南昌 Nanchang.

Zhang 张, Zhenlin 振林. 1982. "先秦古文字材料中的语气词 Xian Qin guwenzi cailiao zhong de yuqici." 古文字研究 *Guwenzi Yanjiu* 7:289–307.

Zhang, Caicai, Quansheng Xia, and Gang Peng. 2015. "Mandarin third tone sandhi requires more effortful phonological encoding in speech production: Evidence from an ERP study." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:149–162.

Zhang, Mingxia, Jin Li, Chuansheng Chen, Leilei Mei, Gui Xue, Zhonglin Lu, Chunhui Chen, Qinghua He, Miao Wei, and Qi Dong. 2013. "The contribution of the left mid-fusiform cortical thickness to Chinese and English reading in a large Chinese sample." *NeuroImage* 65:250–256.

Zhang, Ning. 1996. "Three Types of Object Shift in Mandarin Chinese," 413–425. University of Illinois, Urbana-Champaign.

———. 2000. "Object Shift in Mandarin Chinese." *Journal of Chinese Linguistics* 28 (2): 201–246.

Zhao, Jingjing, Xiaoyi Wang, Stephen J. Frost, Wan Sun, Shin-Yi Fang, W. Einar Mencl, Kenneth R. Pugh, Hua Shu, and Jay G. Rueckl. 2014. "Neural division of labor in reading is constrained by culture: a training study of reading Chinese characters." *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 53:90–106.

Zhou, Xiaolin, Xiaoming Jiang, Zheng Ye, Yaxu Zhang, Kaiyang Lou, and Weidong Zhan. 2010. "Semantic integration processes at different levels of syntactic hierarchy during sentence comprehension: an ERP study." *Neuropsychologia* 48 (6): 1551–1562.

Zhou, Xiaolin, and Hua Shu. 2011. "Neurocognitive processing of the Chinese language." *Brain and Language,* Neurocognitive Processing of the Chinese Language, 119 (2): 59.

Zhu 朱, Dexi 德熙. 1982. *语法讲义 Yufa jiangyi.* Shanghai: Commercial Press [商务印书馆].

Ziegler, Johannes C., and Usha Goswami. 2005. "Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory." *Psychological Bulletin* 131 (1): 3–29.

Zimmermann, Malte, and Caroline Féry. 2009. *Information Structure: Theoretical, Typological, and Experimental Perspectives.* Oxford, New York: Oxford University Press.

Zinszer, Benjamin D., Peiyao Chen, Han Wu, Hua Shu, and Ping Li. 2015. "Second language experience modulates neural specialization for first language lexical tones." *Journal of Neurolinguistics,* Language, Brain, and Gene: The Chinese Context, 33:50–66.

Zribi-Hertz, Anne. 1986. "Relations anaphoriques en Français: esquisse d'une grammaire générative raisonné de la réfléxivité et de l'ellipse structurale." PhD diss., Université de Paris VIII.

Zurif, E., D. Swinney, P. Prather, J. Solomon, and C. Bushell. 1993. "An On-Line Analysis of Syntactic Processing in Broca s and Wernicke s Aphasia." *Brain and Language* 45 (3): 448–464.

Zurif, Edgar, David Swinney, Penny Prather, and Tracy Love. 1994. "Functional localization in the brain with respect to syntactic processing." *Journal of Psycholinguistic Research* 23 (6): 487–497.

潘 Pan, 允中 Yunzhong. 1982. *汉语语法史概要 Hanyu yufa shi gaiyao [Essentials of Chinese Grammar].* 鄭州 Zhengzhou: 中州書畫社.

Murielle Fabre

The Sentence as a cognitive object

The Neural underpinnings of syntactic complexity in Chinese and French

Résumé

En associant les récentes techniques de neuro-imagerie (IRMf et Potentiels Evoqués) à la finesse des analyses syntaxiques des approches typologiques et formelles, cette recherche pluridisciplinaire se penche sur la question de la représentation des structures hiérarchiques qui caractérisent l'unité-phrase à travers les langues. La façon dont le cerveau humain représente, construit et comprend les diverses structures de phrase, est en effet une des plus importantes questions qui restent encore largement irrésolues dans l'organisation cérébrale du langage.

En nous appuyant sur la diversité des langues dans leur articulation syntaxique de la phrase, nous avons pu analyser les dimensions phonologiques, discursives et syntaxiques qui caractérisent l'unité-phrase, et ce grâce aux spécificités des articulations Topique-Commentaire en chinois mandarin ainsi qu'aux propriétés syntaxiques du français dans la formation des questions. Suite à une étude du marquage intonationel de la hiérarchie entre Topique et Commentaire, nous avons pu enregistrer les réponses cérébrales (PE) à ce type de constructions en contexte. Cela a permis d'observer l'influence de sa signature prosodique sur le traitement en temps réel de ces constructions dans leur interface avec l'information contextuelle. Nos deux études d'IRMf apportent quant à elles un éclairage sur les bases neurales de deux dimensions de la complexité syntaxique de la phrase : sa structure hiérarchique et la complexité des transformations structurelles dont elle témoigne en cas de dislocation d'un ou plusieurs de ses éléments. La première étude, sur les interrogatives en français, met en lumière les corrélats cérébraux de différents types de movements syntaxiques, la seconde, sur les différents phénomènes topicaux du chinois, révèle les différentes représentations cérébrales qui sont liées à la présence d'un Topique dans la phrase et aux liens de dépendance entre topique et commentaire en cas de dislocation.

Mots-clés neuro-syntax, movement wh, Topique-Commentaire, clitiques, IRMf, EEG-PE, Français, Chinois mandarin

Abstract

Combining fine-grained linguistic analyses—from both typological and formal approaches to syntax—with neuro-imaging techniques (fMRI and ERP), this trans-disciplinary research experimentally investigates the hierarchical nature and complexity of the linguistic representation of sentence structure and its processing strategies across languages, specifically focusing on the case of Chinese Topic-Comment articulations and French Interrogative constructions.

The question of how the brain achieves sentence structure representation, building, and understanding is often seen as one of the most important and unsolved issues of the neural organization of language. Leveraging on cross-linguistic invariance and variability in sentence hierarchical structure organization and building, we found in Chinese and French two exceptional testing grounds to isolate different syntactic complexity dimensions of the sentence-unit encoding. The on-line auditory comprehension of sentence hierarchical structure in the case of minimal intonational cues is investigated thanks to ERP recordings of Topic-Comment articulations in Chinese. The two fMRI studies isolated two different syntactic complexity dimensions, respectively reflecting the sentence's hierarchy and syntactic transformations and revealed a broad fronto-temporal division of labor. The first study, on French interrogatives, isolated the neural correlates of different syntactic movement types, while the second enabled us to distinguish word-order surface complexity factors from syntactic movement transformations. Thanks to the Chinese sentence-discourse interface properties and its different Topics types, we could observe neurally distinct activation patterns for the establishment of intra-sentential dependency-links and syntactic hierarchy complexity.

Keywords neuro-syntax, wh-movement, Topic-Comment, anaphors, clitics, fMRI, EEG-ERP, French, Mandarin Chinese