NeuroImage 122 (2015) 33–43

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Neural correlates of merging number words

Yi-Hui Hung [a,b,c,d], Christophe Pallier [e,f,g,h], Stanislas Dehaene [e,f,g,h], Yi-Chen Lin [a,b], Acer Chang [i,j,k], Ovid J.-L. Tzeng [l,m], Denise H. Wu [b,i,*]

[a] Institute of Neuroscience, National Yang-Ming University, Taiwan
[b] Laboratories for Cognitive Neuroscience, National Yang-Ming University, Taiwan
[c] Haskins Laboratories, 300 George St, Suite 900, New Haven, CT 06511, USA
[d] Yale University, New Haven, CT 06520, USA
[e] INSERM, U992, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France
[f] CEA, DSV/I2BM, NeuroSpin Center, F-91191 Gif/Yvette, France
[g] Univ Paris-Sud, Cognitive Neuroimaging Unit, F-91191 Gif/Yvette, France
[h] Collège de France, F-75005 Paris, France
[i] Institute of Cognitive Neuroscience, National Central University, Taiwan
[j] Department of Informatics, University of Sussex, Brighton BN1 9QJ, UK
[k] Sackler Centre for Consciousness Science, University of Sussex, Brighton BN1 9QJ, UK
[l] Brain Science Research Center, National Chiao Tung University, Hsinchu, Taiwan
[m] The Institute of Linguistics, Academia Sinica, Taiwan

ARTICLE INFO

Article history:
Received 21 April 2015
Accepted 18 July 2015
Available online 28 July 2015

Keywords:
Structure building
Number word
Cross-cultural invariance
Linguistic syntax
Functional magnetic resonance imaging

ABSTRACT

Complex number words (e.g., “twenty two”) are formed by merging together several simple number words (e.g., “twenty” and “two”). In the present study, we explored the neural correlates of this operation and investigated to what extent it engages brain areas involved processing numerical quantity and linguistic syntactic structure. Participants speaking two typologically distinct languages, French and Chinese, were required to read aloud sequences of simple number words while their cerebral activity was recorded by functional magnetic resonance imaging. Each number word could either be merged with the previous ones (e.g., ‘twenty three’) or not (e.g., ‘three twenty’), thus forming four levels ranging from lists of number words to complex numerals. When a number word could be merged with the preceding ones, it was named faster than when it could not. Neuroimaging results showed that the number of merges correlated with activation in the left inferior frontal gyrus and in the left inferior parietal lobule. Consistent findings across Chinese and French participants suggest that these regions serve as the neural bases for forming complex number words in different languages.

Introduction

Number words can denote very large quantities in a precise manner. In many languages, all integers can be named by combining simple number words from a small and finite set (for exceptions, see Gordon, 2004; Pica et al., 2004). For instance, from simple number words such as “three”, “thirteen”, “thirty” and “hundred”, one can create complex number words “thirty three”, “three hundred and thirteen”, “three hundred and thirty three”, and etc. Across various languages, Hurford (1987) observed that simple number words fell in two common categories: digit names (e.g., “zero”, “eight”) and base morphemes (e.g., “hundred”, “thousand”). To construct complex number words denoting large quantities, digits and base morphemes are merged[1] together by the operations of multiplication and addition. For instance, the merging operations underlying “one hundred and thirty two” can be expressed as $(1 \times 100) + (3 \times 10) + 2$.

These properties of number word systems are widely shared across many cultures (Hurford, 1987). However, the neural mechanisms underlying such powerful numeration systems have not been directly explored. Because simple number words are merged by multiplication and addition, the comprehension of complex number words might recruit the brain regions underlying quantity processing and mathematic operations (see Dehaene, 2009, for a review). Specifically, Piazza et al.

* Corresponding author at: Science Building 5, No. 300, Jhongda Rd., Jhongli City, Taoyuan County 32001, Taiwan.
E-mail address: denisewu@cc.ncu.edu.tw (D.H. Wu).

[1] We used “merging” to refer to the operation that combines two number words into a larger one (e.g. “twenty two”). We do not imply that it is the same operation as ‘merge’ that is used in minimalist syntax and is defined by Chomsky (1999) as “indispensable operation of a recursive system, which takes two syntactic objects and forms the new object”.

http://dx.doi.org/10.1016/j.neuroimage.2015.07.045
1053-8119/

(2007) showed that the activation of the lower bank of the intraparietal sulcus (IPS), in the inferior parietal lobule (IPL), was modulated by the abstract magnitude expressed by stimuli, irrespective of the notations (e.g., dot patterns and Arabic numbers) employed in the presentation. Pesenti et al. (2000) showed that the left IPL was more engaged when participants performed magnitude comparison on numbers than when they performed orientation judgment on numbers. Moreover, addition and magnitude comparison engaged the left IPL to a similar extent, suggesting the involvement of the left IPL in mathematic operations. Furthermore, Stanescu-Cosson et al. (2000) reported that the activation of the IPL positively correlated with the problem size of arithmetic operations (e.g., small problem size: $1 + 2$ vs. large problem size: $5 + 6$). Based on the involvement of the inferior parietal regions in magnitude representations and arithmetic operations (also see Ansari et al., 2005; Dehaene et al., 1999; Eger et al., 2003; see Arsalidou and Taylor, 2011 for a meta-analysis), these areas might also support the formation of complex number words.

In addition to mechanisms in the numerical domain, other domain-general mechanisms might also be involved in the formation of complex number words. Because the operations that merge simple number words resemble those underlying phrase and sentence construction in linguistic materials, it is plausible that the construction of structures within complex number words and phrases/sentences is achieved via the same computational mechanisms. For instance, in the complex number word “two hundred”, the digit “two” modifies the base morpheme “hundred”, perhaps in a similar fashion as the adjective “big” modifies the noun “apple” in the noun phrase “big apple”. Consequently, one might expect to observe the association between the processing of complex number words and the activations in the brain regions sensitive to structure building in sentence processing.

Previous research has associated the left inferior frontal gyrus (IFG) with linguistic structure-building operations. For instance, Hagoort (2005) suggested that the IFG performs unification operations that bind lexical items both syntactically and semantically. Friederici et al. (2006) also related this region, specifically BA 44, with phrase structure building. Several left temporal regions are also suggested to support the combinatorial operations underlying language processing. In a MEG study, Bemis and Pylkkänen (2011) reported that the combination of an adjective and a noun (e.g., *red boat*) elicited neuromagnetic responses reflecting semantic composition in the left anterior temporal lobe (ATL). In a fMRI study, Pallier et al. (2011) manipulated the size of syntactic constituents, and hence the number of merges between lexical items, and found that a greater number of merges led to increased activation in a set of areas of the language network, namely, the left IFG and four areas along the superior temporal sulcus (STS) from the temporal pole to the temporo-parietal junction. Critically, the left IFG and a mid-posterior part of the STS continued to show a sensitivity to the size of constituents regardless of whether the stimuli were composed of real or pseudo content words, hence these regions seemed particularly implicated in syntactic merging. On the other hand, the brain activation in the ATL only increased when real but not pseudo content words were used, which suggested that this region is sensitive to semantic coherence.

The present study explored the neural correlates underlying the formation of complex number words, and investigated to what extent the merging operation for number words engages numerical processing in parietal regions and/or syntactic processing in the temporal and inferior frontal regions. As we had access to Chinese and French participants, we decided to test both populations which allowed us not only to test the commonality of the neural substrates underlying numerical and linguistic merging, but also to assess whether language differences could have an impact on the neural bases of such merging operations. Regarding the creation of complex number words, the French number system is similar to the Chinese one, as both rely on base ten. However, the Chinese system does not use special tens or decade number words and therefore comprises fewer number words and a more transparent syntax. For example, “32” in French is “trente deux” (i.e., “thirty two”) and in Chinese is “三十二” (i.e., “three–ten–two”). In contrast to the perfect regularity of the morphology of the Chinese number system, some number words in the French number system are not morphologically regular. For example, French number words between 11 and 19 use both “ze” and “dix” to represent “ten”. French number words 80 and 90 use the base-twenty principle, as “quatre–vingts” (i.e., “80”) literally means “four–twenty”, and “quatre–vingt dix” (i.e., “90”) literally means “four–twenty and ten”. When we designed the experiments, we only incorporated number words that are regular in both Chinese and French. Our goal was to identify the brain regions underlying the common process of number word merging between the two languages.

Following the design of Pallier et al. (2011), we parametrically manipulated the number of merges of sequentially-presented number words. Four levels of merging were used. At one extreme, all words in the list could be merged into a single complex number word (e.g., “six hundred sixty two thousand nine hundred forty seven”). At the other extreme, the stimulus was a mere list of words, as virtually none of the successive number words could be merged (e.g., “sixty hundred hundred nine six eight five thirty hundred”). Intermediate conditions permitted partially merges of variable overall constituent size (see Figs. 1A and B for French and Chinese stimuli, respectively). Lists of stimuli were created in Chinese and in French, and native speakers of these two languages were always tested in their mother tongue. Different from Pallier et al. (2011), where participants were passively exposed to the stimuli, we asked participants to read out the simple number words one by one as the stimuli were successively presented on the screen. This task ensured that participants processed each of the simple number words naturally and attentively. Furthermore, naming times could be recorded during functional magnetic resonance imaging (fMRI), thus allowing for a fine-grained characterization of whether simple naming is also affected by the possibility of merging.

Methods

Participants

Twenty native French speakers (13 males, ages: 19 - 30 years old, $M = 23$) and 20 native Chinese speakers (10 males, ages: 19 - 30 years old, $M = 22$) with comparable educational background participated in this study. All participants were right-handed, and none of them reported any history of neurological disorders or reading difficulties. The French and Chinese participants were tested in Neurospin, France, and National Yang Ming University, Taiwan, respectively. The experiment was approved by the ethical committees in France (Comité de Protection des Personnes) and in Taiwan (Taiwan Association of Institutional Review Board).

Stimuli

For the French stimuli, 120 sequences of nine simple number words were prepared. Each sequence contained three base morphemes from the set (“mille” (thousand) or “cent” (hundred)), two decades from the set (“vingt” (twenty), “trente” (thirty), “quarante” (fourty), “cinquante” (fifty), or “soixante” (sixty)), and four digits from the set (“deux” (two), “trois” (three), “quatre” (four), “cinq” (five), “six” (six), “sept” (seven), “huit” (eight) or “neuf” (nine)). The digit “un” (one) was excluded to prevent illegal combination of “un cent” (one hundred) in French. The sequences comprised nine successive words and formed four different numbers of merges. This was achieved by varying the positions of the nine simple number words according to the templates listed in Fig. 1A (also see Table S1A in the Supplementary materials for further details). At the highest level, the number sequence (e.g., (six cent soixante deux mille neuf cent quarante sept), “six hundred sixty two thousand nine hundred forty seven”) had the largest number of possible merges (i.e., 8) because each simple number word could be

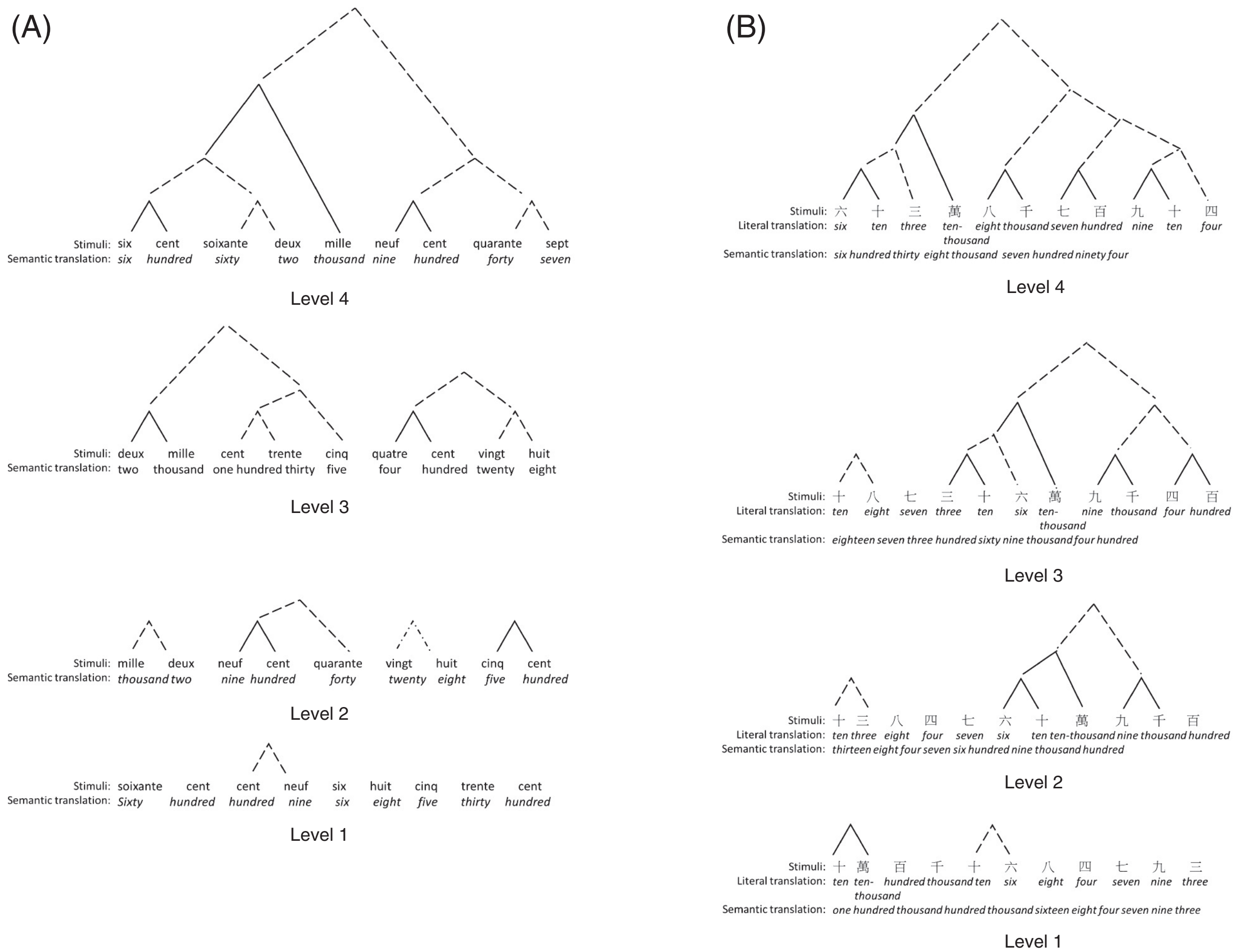

Fig. 1. Stimulus examples of the four levels with different numbers of merges for the French (A) and Chinese (B) numeric strings (solid line: multiplication, dash line: addition).

merged with the previous one. At the lowest level, the number sequence (e.g., "(soixante) (cent) (cent neuf) (six) (huit) (cinq) (trente) (cent)" meaning "sixty hundred hundred nine six eight five thirty hundred") had the smallest number of possible merges (i.e., 1) as only one of the simple number words could be merged with the previous one. In order to match the occurrence of the digits and the decades while simultaneously manipulating the numbers of merges across the four levels, 'mille' never occurred at the lowest level. Stimulus examples of the four levels with different numbers of merges were illustrated in Fig. 1A.

For the Chinese stimuli, similar to the French stimuli, sequences of 11 simple number words were prepared. Each sequence always contained three base morphemes ("萬" (ten-thousand), "千" (thousand), and "百" (hundred)), two decades ("十" (ten)), and six digits ("三" (three), "四" (four), "六" (six), "七" (seven), "八" (eight) and "九" (nine)). The digit "一" (one) was excluded to prevent ambiguous merging with the decade morpheme "ten" because the number ten can be expressed by both "ten" and "one ten" in the Chinese number system. The digit "二" (two) was excluded because of possible pronunciation changes, as "二" was pronounced as "er" in the number words 2 and 20 but as "liang" in 200. The digit "五" (five) was also excluded because it has a glide consonant in the initial which may not be reliably detected by the microphone. The Chinese sequences also comprised four levels with different numbers of merges, achieved by varying the positions of the 11 simple number words to fill the templates listed in Fig. 1B (also see Table S1B in the Supplementary material for further details). At the highest level, the number sequence (e.g., "(六十三萬八千七百九十四)", literal translation: "(six ten three ten-thousand eight thousand seven hundred nine ten four)", semantic translation: "(six hundred thirty eight thousand seven hundred ninety four)") had the largest number of possible merges (i.e., 10) as each simple number word could be merged with the previous one. At the lowest level, the number sequence (e.g., "(十六) (三) (九萬) (千) (四百) (十七) (八)", literal translation: "(ten six) (three) (nine ten-thousand) (thousand) (four hundred) (ten seven) (eight)", semantic translation: "(sixteen) (three) (ninety thousand) (thousand) (four hundred) (seventeen) (eight)") had the smallest number of possible merges (i.e., 2). Stimulus examples of the four levels with different numbers of merges were illustrated in Fig. 1B.

Procedure

Each trial started with the display of a fixation cross at the center of the screen for 800 ms for the French participants and for 2000 ms for the Chinese participants accompanied by a pure tone (50 ms). Then, the nine French number words from a sequence, each of which extended 0.64° - 1.8° horizontally and 0.29° vertically in visual angles, or 11 Chinese number characters, each of which extended 0.8° horizontally and 0.8° vertically in visual angles, were displayed successively, each stimulus being presented for 200 ms and followed by a blank screen for 200 ms (see Fig. 2). Participants were asked to read each stimulus aloud immediately when they saw it on the screen. The instruction emphasized both accuracy and speed. After the disappearance of the final stimulus of a sequence, the screen remained blank for 9.8 s for French participants and 9 s for Chinese participants before the next trial started.

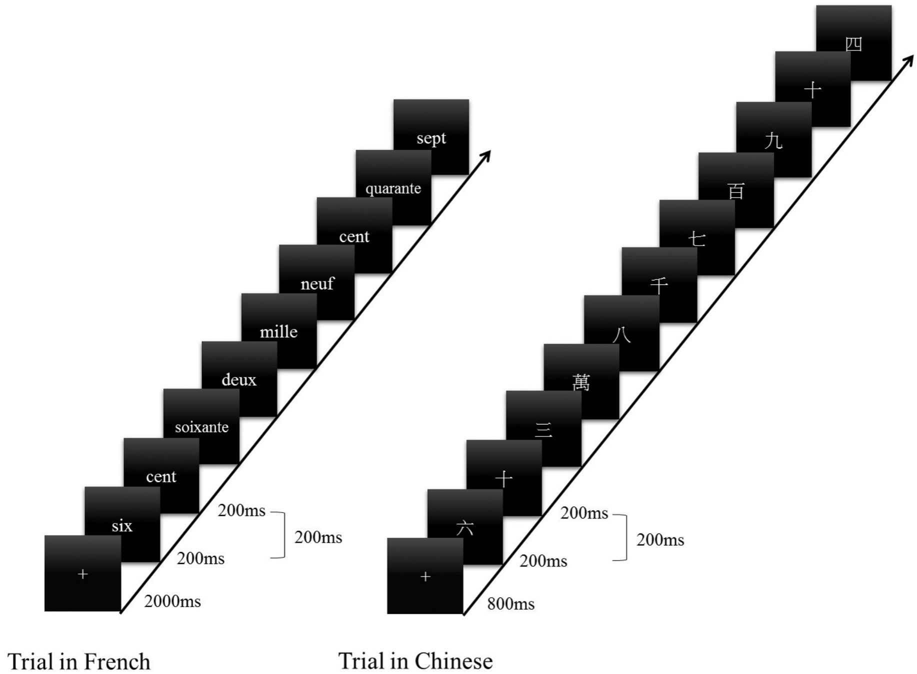

Fig. 2. In each trial, nine French number words or eleven Chinese number words were displayed sequentially for 200 ms separated from each other by a 200-ms blank screen.

In each functional run, there were five trials from each of the four merging levels, which resulted in 20 trials in total. In the end of each run, there was a baseline period during which the screen was empty for 10 s and the participants remained still. The French and Chinese participants received 16 and 10 practice trials, respectively, before entering the MRI scanner. After the acquisition of the anatomical scan, the French participants underwent six functional runs of EPI acquisitions, each of which lasted for 5 min. The Chinese participants underwent seven functional runs of EPI acquisitions, each of which lasted about 6 min. The stimuli were projected on a translucent screen with a digital-light-processing projector (Panasonic PT-D7700E/PT-D4000U; refresh rate: 60 Hz). The experiment was controlled with E-Prime 1.2 software (Psychological Software Tools, Inc., Pittsburgh, PA).

fMRI data acquisition

For both French and Chinese participants, a 3-T Siemens VISION system (Siemens Trio-Tim Syngo) was used to acquire both T1 anatomical volume images ($1 \times 1 \times 1.1$ mm voxels) and T2-weighted echoplanar images (matrix size: 64×64; spatial resolution: 3.4×3.4 mm; TR = 2000 ms, TE = 30 ms, flip angles = 77°). Each echo-planar image comprised 34 and 33 ascending axial slices of 3.4 mm in thickness without gap for the French participants and the Chinese participants, respectively.

Verbal response recording

During the fMRI scanning, the verbal responses of both groups of participants were recorded by a MR-compatible noise canceling audio recording system (FOMRI II; Optoacoustics Ltd.).

Behavioral data preprocessing

Due to a technical failure, audio recordings of 20 trials were missing for one French participant. As these missing trials only represented 17% of his data, this participant was nevertheless included in the analyses.

Any trials which contained one or more failure to respond, reading error (e.g., “seven” named as “six”) or erroneous word order (e.g., “hundred two” was named as “two hundred”) were excluded from further analysis. For all remaining trials, the onset time of each stimulus naming response was estimated by dynamic time warping (DTW) (Ellis, 2003). Specifically, the continuous recording file of each participant's naming responses across the entire experiment was segmented using Praat (Boersma and Weenink, 2013) into individual sequence naming files which spanned from the onset to the offset of all-word naming response in each trial. For each participant and each simple number word, a naming response was manually selected and extracted as a representative single-word naming file. These representative files were then assembled according to the order of the stimulus words in that particular trial to form an ideal naming file. DTW was then applied to determine the naming onset time of each simple number word within the actual naming file of each trial. The procedure warped the ideal naming file in order to minimize the difference between the two files, thus resulting in estimates of naming onset for each word.

fMRI data preprocessing

For both French and Chinese participants, the fMRI data were processed with SPM8. Anatomical images were normalized to the standard brain template defined by the Montreal Neurological 152-brains average provided by SPM8. Functional images were corrected for slice-timing differences and realigned to the mean image of all acquired images to correct for head movements. The functional images were then spatially normalized using the parameters obtained in the normalization of the anatomical images, resampled with a voxel size of $3 \times 3 \times 3$ mm, and smoothed with a 5 mm Gaussian kernel. Experimental effects at each voxel were estimated using a multi-session design matrix. A general linear model (GLM) was created, including five trial types (four regressors for the numerical sequences with correct verbal responses at different merging levels and one additional regressor for incorrect trials across different levels), each modeled by the canonical haemodynamic response function and its first-order time derivative. Regressors for 6 individual motion parameters were also included in the design matrix to capture the remaining signal variations due to head movements. The stimulus duration was set to 3.4 s for the French participants and to 4.2 s for the Chinese participants, corresponding to the display of sequences. High-pass filtering removing frequencies below 1/128 Hz was implemented. Individual contrast maps estimating the responses to each of the four merging levels were smoothed with an 8 mm Gaussian kernel and entered in a second-level group analysis based on an ANOVA model that includes the factors Group (French vs. Chinese participants) and Level (1 to 4), and Subject Unless otherwise mentioned, results are reported at a threshold corrected for multiple comparisons using the family-wise error rate correction (FWE, $p < 0.05$).

Region of interest (ROI) analysis

We performed analyses in a total of 16 regions of interest (ROIs), half located in the left hemisphere, and half in the right hemisphere (at symmetric locations, obtained by flipping the sign of the x coordinate of the left ROIs). Seven of the left hemispheric regions were shown to be sensitive to constituent structure in language in a previous study (Pallier et al., 2011). They included the left IFG pars triangularis (IFGtri), the left IFG pars orbitalis (IFGorb), the left anterior superior temporal sulcus (aSTS), the left posterior superior temporal sulcus (pSTS), the left temporal parietal junction (TPJ), the left temporal pole (TP) and the putamen. In addition, we added a region of interest in the left IFG pars opercularis (IFGop), defined as sphere of 1 cm radius centered on the coordinates of the center of mass of BA 44 described by Amunts et al. (1999).

Within each ROI, for each participant, the voxels most responsive to the task were first selected using the contrast averaging over all levels of merging (uncorrected $p = 0.01$, percentage of participants showing activation = 0.5), then, the positive linear effect of level of merging was computed in these voxels. To avoid bias (Kriegeskorte et al. 2009), this was performed using different sessions to estimate the average response and the linear trends. The spm_ss toolbox (version 11e; available at http://web.mit.edu/evelina9/www/funcloc/funcloc_toolbox.html; see Fedorenko et al., 2010) was used to perform this operation.

Results

Behavioral results

The reaction time (RT) of each word naming was defined by the interval between the word onset-time and the word naming onset-time estimated by DTW. The trials with incorrect responses (French: 10.5%; Chinese: 8.1%) or with extreme RTs (i.e., shorter than 200 ms or longer than 1200 ms, and the trials whose RTs were outside the range of 2.5 standard deviations from the mean; French: 2.8%; Chinese: 4%) were excluded from further analysis. Average naming times of each word, computed from the remaining trials, are displayed for each group as a function of the level (Fig. 3).

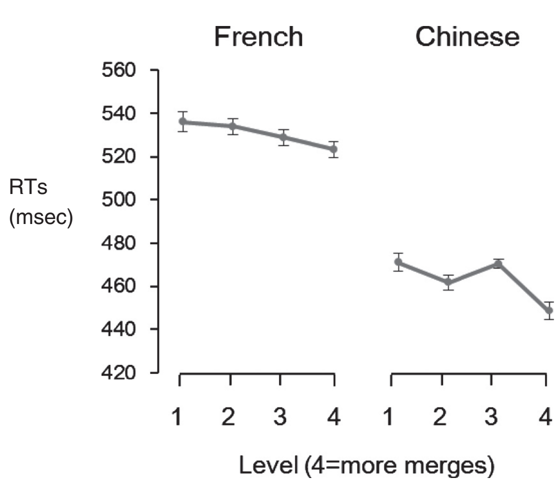

Fig. 3. Naming reaction times (RTs) decreased across levels as the number of merges increased. (Error bars indicate +-1 standard error after removing the main effect of participants.)

The RT data from French and Chinese participants were separately analyzed by a linear mixed-effect model including subject as a random factor and two fixed effects for level (1–4, as a numeric variable) and serial position (1–9 for the French group, 1–11 for the Chinese group, as a numeric variable). The analysis revealed that naming RT was affected by syntactic level, as a negative linear effect was detected for both Chinese (−5.5 ms/level, SE = 1.1, *Wald* X^2 (1) = 23.5, $p < 0.0001$) and, to a near-significant degree, in French participants (−4.6 ms/level, SE = 2, *Wald* X^2 (1) = 3.77, $p = 0.052$). These results indicated that although the naming task placed no explicit requirement on the processing of syntactic structure, both groups of participants exhibited faster naming RTs of the simple number words in the higher (i.e., more mergeable) levels than in the lower (i.e., less mergeable) levels, albeit the effect was more robust for the Chinese than the French group. The naming RTs of the simple number words were also affected by the serial position of the stimulus in both Chinese (−7.1 ms/serial position, SE = 0.4, *Wald* X^2 (1) = 333; $p < 0.00001$) and French groups (−4.2 ms/serial position, SE = 0.9, *Wald* X^2 (1) = 24; $p < 0.001$), showing that the naming responses were faster in later than earlier serial positions in the numerical sequences.

The availability of the naming RT of each word allowed us to perform a fine-grained analysis by introducing a 'mergeability' factor that specified, for each word, whether it could merge with the preceding one. We therefore reran a mixed-effect model for French and Chinese participants separately, replacing the factor of level by the factor of mergeability (1 = mergeable with the preceding word, 0 = non-mergeable). The variable of serial position was kept in the model as a covariate except that the data from the first serial position were excluded as these words were not preceded by any word. The analysis revealed that for the Chinese participants, the mergeability had a significant negative effect on naming RTs (−20 ms, SE = 3, *Wald* X^2 (1) = 49, $p < 0.001$), while the same effect was marginal for the French group (−14 ms, SE = 7, *Wald* X^2 (1) = 3.45, $p = 0.06$). That is, the naming RT of a number word was faster when the word could merge with the preceding one for both groups of participants, though the effect was more robust in the Chinese than French participants. The findings suggest that although the number words were read one by one, they were merged into a coherent speech unit whenever possible, hence facilitating their naming.

Neuroimaging results

Whole-brain analyses

The global network of areas implicated in the number–word naming task was identified by a contrast averaging over the four levels with different numbers of merges (relative to the implicit baseline). It included the left precentral cortex (extending to the left IFG), the left superior temporal region, the right middle occipital cortex, the left supplementary motor area, the bilateral thalamus, the right cerebellum, the bilateral para-hippocampus and the left hippocampal area (see Table 1).

Although the global network underlying the number–word naming task was very similar across French and Chinese participants, a direct comparison between the two groups showed that the French participants had stronger activations than the Chinese participants in small clusters in the bilateral cerebellum, the bilateral inferior occipital area, the left para-hippocampus, the left superior frontal area, and the right middle temporal area. The reversed contrast showed that the activation was stronger for Chinese than French participants in the bilateral middle occipital cortex, the bilateral fusiform gyrus, the right cuneus, the bilateral precuneus, the bilateral inferior temporal area, the left precentral area, the left inferior occipital area, and the left calcarine area (see Table 1). Because the stimuli and procedures of the number–word naming task in French and in Chinese were not exactly the same,

Table 1
Local activation maxima for the main effect of number words for the two groups in MNI space (FWEs < 0.05).

	Anatomical labels	Cluster size	T values	X	Y	Z
Average across all levels for French and Chinese	Precentral_L[a]	1945	22.03	−54	−4	49
			17.29	−45	−13	37
			16.54	−51	−10	31
	Temporal_Sup_R[a]	1579	16.59	57	−16	1
			15.46	48	−7	34
			13.52	60	2	−2
	Occipital_Mid_R[a]	3556	16.08	30	−94	4
			16.08	−12	−61	−20
			15.28	21	−61	−23
	Supp_Motor_Area_L	147	10.44	−3	5	67
	Thalamus_L	83	7.33	−12	−13	10
	Cerebellum_R	41	6.60	9	−70	13
		33	6.35	−12	−19	−11
	ParaHippocampal_L		5.61	−15	−1	−17
	Thalamus_R	28	6.32	15	−10	10
	Hippocampal_L	7	5.87	15	−22	−8
	ParaHippocampal_R	3	5.44	6	−13	−17
French vs. Chinese for all levels	Cerebellum__L	160	7.77	−36	−49	−35
			6.56	−15	−61	−26
			5.96	−48	−64	−35
	Occipital_Inf_R	19	6.65	33	−97	−8
	Cerebellum_R	68	6.58	21	−61	−29
			6.27	42	−55	−32
	ParaHippocampal_L	36	6.33	−15	2	−17
	Occipital_Inf_L	7	6.18	−27	−100	−11
	Frontal_Sup_L	6	5.52	−12	56	40
	Temporal_Mid_R	2	5.33	60	−31	−5
Chinese vs. French for all levels	Occipital_Mid_L	351	8.39	−30	−79	1
			7.23	−27	−88	16
	Fusiform_L		5.95	−30	−73	−11
	Cuneus_R	730	7.45	6	−76	25
	Occipital_Mid_R		7.03	27	−85	10
	Precuneus_R		6.78	18	−46	7
	Temporal_Inf_R	36	7.32	45	−16	−32
	Fusiform_R	19	6.94	33	−25	−26
	Precentral_L	16	6.89	−54	−4	49
	Temporal_Inf_R	4	6.20	51	2	55
	Temporal_Inf_L	1	5.52	−42	−13	−35
	Precuneus_L	2	5.47	−15	−46	4
	Occipital_Inf_L	2	5.47	−45	−70	−5
	Precuneus_L	3	5.34	−18	−58	40
	Calcarine_L	1	5.14	−15	−52	7

[a] Also significant in a conjunction analysis across the two groups of participants.

these relatively minor differences between the two groups might be due to such factors rather than to genuine differences in the computations involved in forming French and Chinese number words. A direct comparison of the effects of merging levels between the two groups (see below) will shed more light on the common and specific brain regions underlying number–word formation in French and in Chinese.

To identify the brain areas supporting the merging of number words, we used a positive linear contrast on the factor Level. The results (displayed in Fig. 4A and Table 2) showed significant increases in activation with Level in the inferior frontal gyrus (pars triangularis bilaterally; pars opercularis on the left and pars orbitalis on the right), the left insula, the left superior medial frontal gyrus, the left inferior parietal lobule and the angular gyrus.

The reverse linear contrast on the factor Level revealed regions from the left middle occipital, the right middle temporal, and the bilateral superior frontal areas in both groups of participants (see Table 2 and Fig. 4B), indicating that these regions were more activated when there were more non-mergeable number words in a sequence.

Finally, we examined the potential differences between French and Chinese, using the interactions between Group and the linear effects of Level. This analysis revealed that the positive linear effect of level was stronger in Chinese participants than in French participants in the right middle orbital frontal area (see Table 2 and Fig. 4C). No brain region significantly differed between the two groups on the negative linear effect of Level.

Region-of-interest analysis

The effect of the number of merges of successive number words was examined in regions of interest which have been previously implicated in the construction of linguistic constituents. The activation profiles are presented in Fig. 5. For the French participants, the linear effect of Level reached significance in the three left IFG regions and in the right IFGtri (the effect was significant at a corrected level (Bonferroni $c = 16$, $\alpha FWE = 0.05$) in the left IFGorb; and only at an uncorrected $p < .05$ level in the left and right IFGtri and the left IFGop). For the Chinese participants, the positive linear effect of level reached the significance level (uncorrected $p < 0.05$, but did not survive the Bonferroni correction) in the left IFGtri and the right IFGorb. No effects were detected in the temporal lobe regions.

Discussion

In the present fMRI experiment, French and Chinese participants were required to read aloud sequences of simple number words as accurately and as quickly as possible. In addition to number processing and structure building, visual processing (reading) and motor execution (speaking) were also involved in the overt naming task. Relative to baseline, activations were observed not only in the frontal and temporal regions, but also in the occipital cortex, motor cortex and the cerebellum, for both groups of participants. Although the two groups of participants exhibited a very similar neural network in reading aloud number sequences, there were small but significant differences in the direct contrast between them. Specifically, French participants showed higher activation in the bilateral cerebellum than Chinese participants. This might indicate that naming French number words, which contain one to several syllables, requires more efforts in planning and coordinating the vocal organs than naming Chinese number words, all of which are mono-syllabic. In contrast, Chinese participants showed higher activation in the occipital cortex than French participants, which might reflect that Chinese number words are visually more complex than French number words (see Tan et al., 2005, for similar findings). However, it should be noted that there were inevitable differences in the experimental designs of the number–word naming task in French and in Chinese, such as the durations of the number–word sequences in the two languages. Therefore, the differences observed in the direct comparison between French and Chinese participants should be interpreted with caution, while the contrast between the effects of merging levels within each group (i.e., the interaction) would be appropriate to reveal the differential involvement of brain regions in forming number words of the two languages.

To identify the neural correlates underlying structure building in the number domain, we systematically varied the possibility for simple number words to merge with the preceding ones and form complex number words. As the number of possible merges increased, the naming time of individual simple number words decreased, and there was an increase in the activation of the left IFG and the left IPL in both French and Chinese participants. These results suggest that the left IFG and the left IPL are involved in merging simple number words to form the basic structures of “digit × base morpheme + digit” that underlie the formation of complex number words in all languages.

The left IFG, whose activation was positively correlated with the possible merges of simple number words in the present study, has also been repeatedly identified in previous studies of syntactic processing (e.g., Bahlmann et al., 2008; Caplan et al., 2008; Musso et al., 2003; Pallier et al., 2011; Suzuki and Sakai, 2003). For example, Caplan et al. (2008) identified a widespread neural network for syntactic processing, including the left IFG, by contrasting the brain activations during grammaticality judgment on sentences with object- and subject-relative clauses. Suzuki and Sakai (2003) also demonstrated that the IFGop

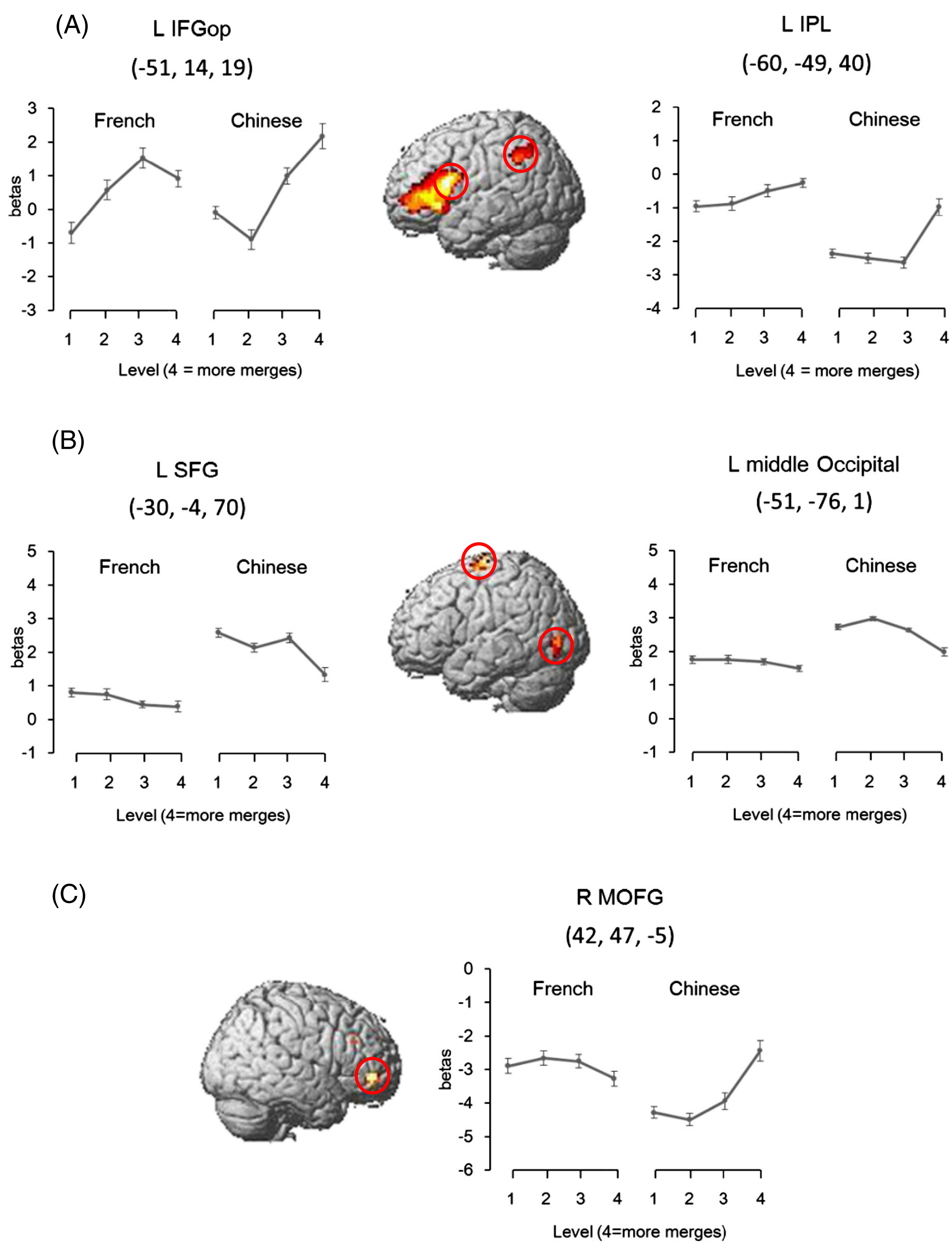

Fig. 4. (A) Brain regions where activation increased across levels as the number of merges increased (FWE corrected, $p < 0.05$). The activation at the local maxima in the left inferior parietal lobule (L IPL) and in the left inferior frontal pars opercularis (L IFGop) was plotted. (Error bars indicate +-1 standard error after removing the main effect of participants.) (B) Brain regions where activation decreased across levels as the number of merges increased (FWE corrected, $p < 0.05$). The activation at the local maxima in the left middle occipital region and in the left superior frontal gyrus (L SFG) was plotted. (Error bars indicate +-1 standard error after removing the main effect of participants.) (C) One brain area showed a difference in the effect of merging levels between the two groups (FWE corrected, $p < 0.05$). The activation at the local maximum in the right middle orbitofrontal gyrus (R MOFG) was plotted, which indicated that only Chinese participants but not French participants showed the effect. (Error bars indicate +-1 standard error after removing the main effect of participants.)

and IFGtri were particularly involved in explicit syntactic processing of transitive and intransitive verbs. Pallier et al. (2011) employed a design similar to the current study to examine the neural correlates underlying the merging of real and pseudo content words to form constituents with different sizes. The findings showed that the IFG was sensitive to the number of merges even in meaningless Jabberwocky phrases. This is

Table 2
Local activation maxima for the merging-level effect for the two groups in MNI space (FWEs < 0.05).

	Anatomical labels	Cluster size	T values	X	Y	Z
Positive level effect averaged across French and Chinese participants	Frontal_Inf_Oper_L[a]	719	7.58	−51	14	19
	Frontal_Inf_Tri_L		6.44	−45	35	4
	Insula_L		5.38	−33	20	−5
	Frontal_Inf_Tri_R	129	6.16	57	26	16
	Frontal_Sup_Medial_L	50	5.73	−3	35	43
	Parietal_Inf_L	103	5.63	−60	−49	40
	Angular_L		4.80	−45	−61	52
	Frontal_Inf_Orb_R	37	5.17	39	32	−5
Negative level effect averaged across French and Chinese participants	Occipital_Mid_L	32	5.82	−51	−79	1
	Frontal_Sup_L	45	5.31	−30	−4	70
			5.18	−24	−10	70
	Paracentral_L		4.60	−6	−16	73
	Frontal_Sup_R	28	5.04	21	2	73
	Supp_Motor_R	8	5.04	9	−10	73
	Temporal_Mid_R	4	4.76	51	−67	−2
	Frontal_Mid_L	1	4.61	−48	44	16
		1	4.56	57	−46	55
Positive level effect Chinese > French	Frontal_Mid_Orb_R	34	5.48	42	47	−5
	Frontal_Inf_Tri_R	4	4.71	36	29	28

[a] Also significant in a conjunction analysis across the two groups of participants.

also consistent with Hagoort (2005) that Broca's area is involved in unification operations inasmuch as merging two items into a larger one can be viewed as a unification process.

The contribution of the IFG to merging operations in the mathematical domain has been investigated in several previous studies, with a diversity of paradigms and, unfortunately, highly inconsistent outcomes (Bemis and Pylkkänen, 2013; Friedrich and Friederici, 2009; Makuuchi et al., 2012; Maruyama et al., 2012; Nakai and Sakai, 2014). Maruyama et al. (2012) manipulated the combinability of digits and operands in mathematical expressions by contrasting scrambled expression such as "4-+)3)(+2(1")") with nested expressions such as "((3 − 2) + 4) + 1". They only found a very small effect in the left IFG, detectable only at $p < 0.05$ in a targeted ROI analysis. The whole brain analysis, instead, revealed large effects in the converse direction (destructured > structured expressions) in visual and parietal areas, perhaps due to the high processing demands of the brief visual presentation (200 ms) and the same–different comparison task used, which was more difficult for destructured compared to structured stimuli.

Friedrich and Friederici (2009) similarly examined the neural underpinnings of rule-based hierarchical processing of mathematical expressions by presented algebraic expressions which were organized hierarchically (e.g., $(a = c + u) \wedge (v \cdot x < u + x)$) or in a list fashion (e.g., $\{a + c, x \cdot v, \varphi \wedge \psi = a, u < y\}$). Participants were asked to judge whether the expressions were legal. The results showed small activations in the left IFGtri (BA 45) and the left IFGorb (BA 47), but not the classic Broca's area (BA 44), during the hierarchical condition compared to the list condition, which was interpreted as challenging the idea of a common domain-general brain mechanism for "merge" in the IFG. In the same group, however, Makuuchi et al. (2012) using an explicit calculation task reported common involvement of the IFG in the hierarchical build-up in both the language and mathematics domains, and they speculated that the reduced involvement of the IFG in Friedrich and Friederici (2009) was due to the abstract mathematical formulae employed. Similarly, during explicit calculation, Nakai and Sakai (2014) found that the left IFG, as well as the left supramarginal gyrus, showed positive activations with the complexity of the hierarchical tree structure underlying nested computation with single digits. Unfortunately, their design is partially confounded by the number of operations involved, which increases with the amount of tree structure.

Monti et al. (2012) studied a carefully matched contrast between equivalence versus grammaticality judgments for linguistic arguments (e.g., is "Z was paid X by Y" equivalent to "It was X that Y paid Z"?), and for tightly parallel algebraic arguments (e.g., is "X minus Y is greater than Z" equivalent to "Z plus Y is smaller than X"?). They observed a double dissociation between language and mathematics, with an activation of the left IFG only during the manipulation of linguistic arguments. Conversely, there was a strong involvement of bilateral parietal regions during the processing of algebraic but not linguistic arguments. Monti and colleagues concluded that distinct neural substrates support syntactic combinatorial manipulations in these two domains. However, it should be pointed that the contrast between the equivalence and grammaticality judgments in Monti et al. is also likely to engage semantic processing, as determining the equivalence of two arguments heavily rely on understanding the meanings of both arguments. Therefore, the findings might demonstrate distinct neural computations of concepts

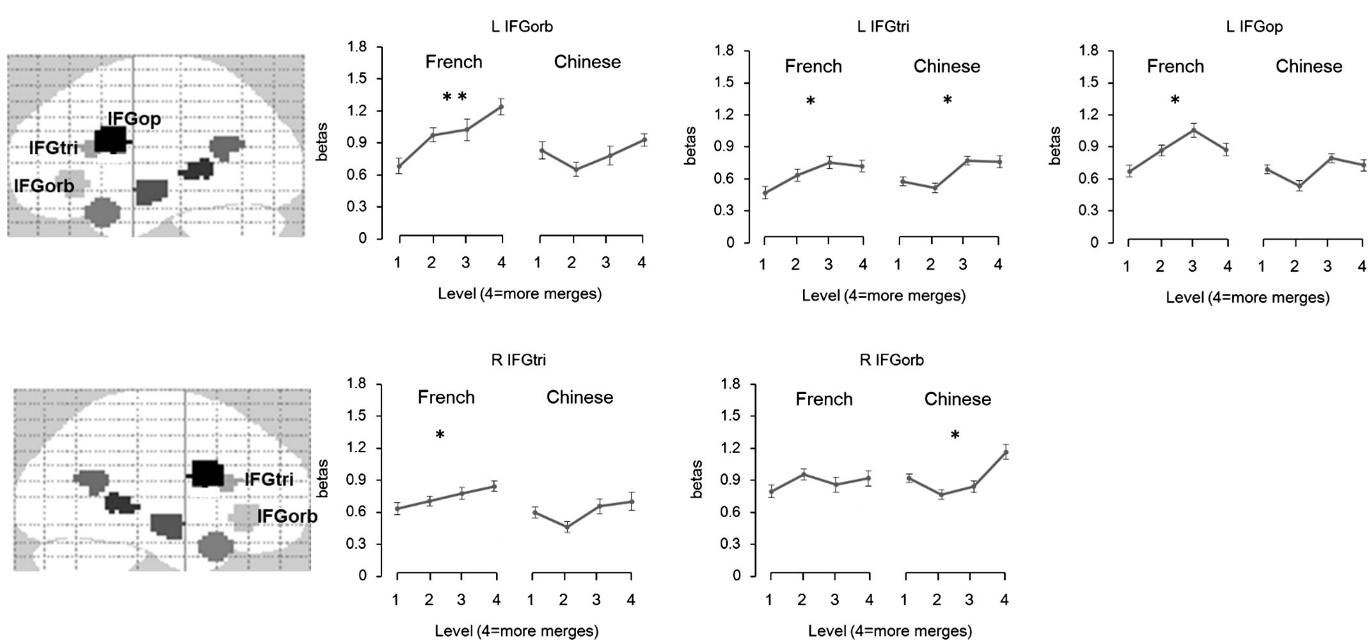

Fig. 5. The ROIs relevant to merging linguistic materials, whose activation increased across levels as the number of merges of number words increased, were shown (*uncorrected $p < 0.05$; **significant after Bonferroni correction). (Error bars indicate +-1 standard error after removing the main effect of participants.)

but not necessarily syntax in the language and mathematical domains. Similarly, in a recent MEG study, Bemis and Pylkkänen (2013) found no involvement of the frontal regions when comparing the activation associated with adding two digits (e.g., "2 + 3") and the activation associated with adding one digit and a meaningless symbol (e.g., "2 + あ."). Although the stimuli in the former condition were more mergeable than those in the latter condition, the key difference between the two conditions might lie in the meaningfulness of the stimuli. The lack of involvement of the IFG in Monti et al. and Bemis and Pylkkänen might therefore suggest that syntactic processes of structure building, rather than semantic composition, drive the activation of the IFG. This possibility is consistent with the findings from Pallier et al. (2011) and Ohta et al. (2013) in which the left IFG is sensitive to hierarchical structure building even in the Jabberwocky condition.

Taken together, the present and previous findings, using a variety of linguistic and non-linguistic materials (see also Friederici, 2011; Koechlin and Jubault, 2006), offer only weak support for the hypothesis that IFGtri and IFGorb subserve a general "merge" operation that would combine individual words, numbers or algebraic symbols according to abstract combinatorial principles across various domains (Fitch and Martins, 2014). While more systematic research is clearly needed, the existing brain-imaging literature reveals striking dissociations between mathematical and linguistic compositionality, and equally striking dissociations have been reported in aphasic patients with preserved algebraic abilities (Varley et al., 2005). To account for this complex pattern of associations and dissociations, one possibility is to assume that a large extent of the inferior frontal gyrus does play an overarching role in compositionality, but with a systematic variation in the localization of the exact cortical site depending on the posterior areas holding the information that has to be merged. There would thus be multiple parallel circuits for merging or "unification", all involving a node at or near the IFG, but with dorsal or ventral displacement depending on the domain to which the merged objects belong (Xiang et al., 2010). When numbers are merely manipulated as words to be integrated in linguistic constituent structures, as in the present work, then a clear contribution of the "Broca's area" is seen. When the task calls for meaningful manipulations of arithmetic or algebraic nested expressions, involving non-linguistic mathematical symbols (Arabic digits, algebraic symbols) encoded primarily in the intraparietal sulcus, then more dorsal parietal, precentral and prefrontal circuits would be engaged.

There is indeed a high convergence between all studies in indicating a strong role for the parietal lobe whenever mathematical objects are involved. In the present work, in addition to the IFG, the activation of the left IPL correlated positively with the number of possible merges of successive simple number words in the present study. Semantically, the formation of complex numbers requires a nested combination of multiplication and addition operations. In this respect, the activation of the IPL is consistent with many previous studies that have observed parietal lobe activations during arithmetic operations (see Arsalidou and Taylor, 2011 for a meta-analysis). It should be noted, however, that the numerical quantity of the complex number words also increased with the number of possible merges in the French stimuli: only the sequences with the largest number of merges denoted a magnitude of several hundred thousands ('cent mille'). Therefore, the increasing activation of the left IPL might have reflected the increasing magnitude of the outcome of the merging operations, rather than the merging operations per se. Unlike the French stimuli, however, in Chinese the quantity of the number words in all four levels with different numbers of merges reached the magnitude of "十萬" ("ten ten-thousand", meaning "hundred thousands"). For instance, Level 4 had "六十三萬" ("six ten three ten-thousand", meaning "six hundred thirty thousand"), Level 3 had "七十萬" ("seven ten ten-thousand", meaning "seven hundred thousand"), Level 2 had "六十萬" ("six ten ten-thousand", meaning "six hundred thousand") and Level 1 had "四十萬" ("four ten ten-thousand", meaning "four hundred thousand"). Because a positive linear trend of the activation in the left IPL across the four levels was still observed in Chinese participants, the result provides evidence that this region supported numerical merging operations, such as multiplication and addition, rather than merely reflecting the numerical quantity expressed by complex number words. Actually, Nakai and Sakai (2014) also identified this region to be sensitive to the "degree of merger" in the hierarchical tree structure in the mathematics domain, providing converging evidence with the present finding.

In contrast to several previous studies on merging processes in the linguistic domain (Bemis and Pylkkänen, 2011; Pallier et al., 2011), the present study did not detect any involvement of temporal regions in the merging of simple number words. This is so although such temporal activations are typically observed by comparing the processing of sentences with that of word lists, a contrast superficially similar to the present one (see Ferstl et al., 2008 for a meta-analysis). Such a finding could reflect the different characteristics of combinatorial processes underlying semantic composition and/or syntactic structure building in the linguistic and numerical domains. Previous studies have indicated that the anterior temporal region is sensitive to the processing of meaning and reference and may be involved in semantic composition (e.g., Bemis and Pylkkänen, 2011). The fact that activation in this region was not modulated by the number of merges of simple number words in the present study suggests that the semantic operations that it performs are not engaged by mere numbers, but perhaps involve the formation of complex object, person, or event structures. Similarly, the posterior part of the STS is sensitive to syntactic manipulations (Friederici et al., 2000; Monti et al., 2012) as well as to the formation of syntactic constituents (Pallier et al., 2011) but not to the formation of complex number words. One possibility is that it contributes to a "syntactic lexicon" that provides grammatical category and morpheme information indispensable to parsing. Because all of our conditions comprised the same type and amount of simple number words, such a region would have been activated identically in all conditions. Whether or not these speculative interpretations hold, our results are compatible with previous findings suggesting that the processing of number words occurs outside of the temporal lobe, to such an extent that patients with semantic dementia may still exhibit largely preserved processing of numbers and arithmetic (Cappelletti et al., 2001, 2012; Crutch and Warrington, 2002).

In addition to brain regions whose activations demonstrated a positive effect of the number of merges, the present study also revealed brain regions whose activations were negatively correlated with the number of merges. Specifically, activations of the left paracentral cortex and the left supplementary motor region increased as the number of merges decreased. Because the latency of naming responses of simple number words also correlated negatively with the number of merges, these regions might be involved in the planning or execution of articulation in the naming task (Price, 2010). Activations of the left middle occipital cortex also increased as the number of merges decreased. This effect might be due to the fact that simple number words in the lower levels were less easily combined to denote a large quantity than in the higher levels. Therefore, the naming responses in the lower levels might receive less help from the predictability at the lexical/semantic level and rely more on the visual analysis of the stimuli for correct verbal production, than in the higher levels, hence the higher activation in the occipital vision regions.

Our behavioral and fMRI results dissociate word merging operation and task difficulty by finding that distinct brain regions correlated with reaction time across levels in opposite directions. The time to name number words, which reflects the general task difficulty that might be driven by working memory and attention, was negatively correlated with the number of merges. This finding suggests that the activations in the left IPL and in the left IFG, which were positively correlated with mergeability but negatively correlated with reaction time, could not be due to task difficulty, working memory or attention. Conversely, the activation of the left SFG and the left middle occipital area correlated positively with reaction times but negatively with mergeability. These activations might be partly due to the putatively higher demands of working memory and/or attention when naming less mergeable stimuli (Level 1) than when naming more mergeable stimuli (Level 4). The present results are consistent with the idea that the left IFG and IPL are involved in the combinatorial processes for forming complex number words, while the left SFG and middle occipital area support bottom-up processing for naming words that cannot be predicted or integrated in phrases. Fedorenko et al. (2012) used a similar approach to dissociate the mechanisms underlying language processing and general task demands. Specifically, they observed that some voxels within BA 45 responded more to sentences than to nonwords, but did not respond to manipulations of general cognitive demands (e.g., verbal or spatial working memory). On the other hand, other voxels within BA 44 and BA 45 responded more to the nonword condition than to the sentence condition, and they also responded more to the difficult condition than to the easy condition in various general cognitive tasks (e.g., high versus low memory load). The findings from both the present study and Fedorenko et al. indicated that merging and task difficulty can be dissociated in the brain.

The behavioral and neuroimaging results from the French and Chinese participants in the present study were quite similar. The finding of parallel increases reflecting the numbers of merges for numerical stimuli in both languages are consistent with Hurford (1987) of a universal linguistic structure of numerical expressions, i.e. a universal set of constraints by which complex number words are formed by merging number words and/or morphemes. However, a small difference between the two groups was also observed: The activation of the right orbitofrontal gyrus positively correlated with the number of merges in the Chinese participants but no in the French participants. The difference might relate to the proposal that Chinese readers engage the right hemisphere to a greater degree than French readers for stimulus recognition (Tan et al., 2005), rather than specific to the operations of merging number words. However, when contrasting the main effect of number word naming of all levels between the Chinese and French participants, the right prefrontal cortex was not more involved in the Chinese than French participants (see Table 1). Further research is needed to determine whether and how the functional role of the right orbitofrontal gyrus differs in processing number words across different languages.

Conclusions

To our knowledge, the present study provides the first evidence for the neural correlates of the construction of complex number words in two typologically distinct languages. The left IFG and the left IPL are involved in combining simple number words in French and Chinese alike to form complex number words. Overall, the evidence suggests that the left IFG supports a general mechanism underlying structural building in the language domain, while the left IPL may support the numerical operations that merge digits and base morphemes to express quantity beyond the range of single digits.

Acknowledgment

We thank Marco Lai for fMRI data collection from Chinese participants and initial data processing. This work was supported by grants from the National Science Council Taiwan (NSC 99-2911-I-008-507, NSC 100-2911-I-008-505, NSC 102-2628-H-008-002-MY3, NSC 99-2410-H-008-023-MY3) and by the French Ministry of Research (ANR CONSTRUCT 2010 BLANC 1403 01).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.07.045.

References

Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H., & Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J. Comp. Neurol. 412 (2), 319–341.

Ansari, D., Garcia, N., Lucas, E., Hamon, K., & Dhital, B., 2005. Neural correlates of symbolic number processing in children and adults. Neuroreport 16 (16), 1769–1773.

Arsalidou, M., & Taylor, M.J., 2011. Is $2+2=4$? Meta-analyses of brain areas needed for numbers and calculations. NeuroImage 54 (3), 2382–2393.

Bahlmann, J., Schubotz, R.I., & Friederici, A.D., 2008. Hierarchical artificial grammar processing engages Broca's area. NeuroImage 42 (2), 525–534.

Bemis, D.K., & Pylkkänen, L., 2011. Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. J. Neurosci. 31 (8), 2801–2814.

Bemis, D.K., & Pylkkänen, L., 2013. Combination across domains: an MEG investigation into the relationship between mathematical, pictorial, and linguistic processing. Front. Psychol. 3.

Boersma, Paul, & Weenink, David, 2013. Praat: Doing Phonetics by Computer [Computer Program] Version 5.3.59, retrieved 20 November 2013 from http://www.praat.org/.

Caplan, D., Chen, E., & Waters, G., 2008. Task-dependent and task-independent neurovascular responses to syntactic processing. Cortex 44 (3), 257–275.

Cappelletti, M., Butterworth, B., & Kopelman, M., 2001. Spared numerical abilities in a case of semantic dementia. Neuropsychologia 39 (11), 1224–1239.

Cappelletti, M., Butterworth, B., & Kopelman, M., 2012. Numeracy skills in patients with degenerative disorders and focal brain lesions: a neuropsychological investigation. Neuropsychology 26 (1), 1–19.

Chomsky, N., 1999. Derivation by Phase vol. 18. MIT, Department of Linguistics, pp. 1–52.

Crutch, S.J., & Warrington, E.K., 2002. Preserved calculation skills in a case of semantic dementia. Cortex 38 (3), 389–399.

Dehaene, S., 2009. Origins of mathematical intuitions. Ann. N. Y. Acad. Sci. 1156 (1), 232–259.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S., 1999. Sources of mathematical thinking: behavioral and brain-imaging evidence. Science 284 (5416), 970–974.

Eger, E., Sterzer, P., Russ, M.O., Giraud, A.L., & Kleinschmidt, A., 2003. A supramodal number representation in human intraparietal cortex. Neuron 37 (4), 719–726.

Ellis, D., 2003. Dynamic Time Warp (DTW) in Matlab from http://www.ee.columbia.edu/~dpwe/resources/matlab/dtw/.

Fedorenko, E., Hsieh, P.J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N., 2010. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J. Neurophysiol. 104 (2), 1177–1194.

Fedorenko, E., Duncan, J., & Kanwisher, N., 2012. Language-selective and domain-general regions lie side by side within Broca's area. Curr. Biol. 22 (21), 2059–2062.

Ferstl, E.C., Neumann, J., Bogler, C., & Von Cramon, D.Y., 2008. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. Hum. Brain Mapp. 29 (5), 581–593.

Fitch, W., & Martins, M.D., 2014. Hierarchical processing in music, language, and action: Lashley revisited. Ann. N. Y. Acad. Sci. 1316 (1), 87–104.

Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91 (4), 1357–1392.

Friederici, A.D., Meyer, M., & von Cramon, D.Y., 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang. 74 (2), 289–300.

Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., & Anwander, A., 2006. The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proc. Natl. Acad. Sci. U. S. A. 103 (7), 2458–2463.

Friedrich, R., & Friederici, A.D., 2009. Mathematical logic in the human brain: syntax. PLoS ONE 4 (5), e5599.

Gordon, P., 2004. Numerical cognition without words: evidence from Amazonia. Science 306 (5695), 496–499.

Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9 (9), 416–423.

Hurford, J., 1987. Language and Number: The Emergence of a Cognitive System. Basil Blackwell, Oxford.

Koechlin, E., & Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. Neuron 50 (6), 963–974.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., & Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. Nat. Neurosci. 12 (5), 535–540.

Makuuchi, M., Bahlmann, J., & Friederici, A.D., 2012. An approach to separating the levels of hierarchical structure building in language and mathematics. Philos. Trans. R. Soc. B 367 (1598), 2033–2045.

Maruyama, M., Pallier, C., Jobert, A., Sigman, M., & Dehaene, S., 2012. The cortical representation of simple mathematical expressions. NeuroImage 61 (4), 1444–1460.

Monti, M.M., Parsons, L.M., & Osherson, D.N., 2012. Thought beyond language neural dissociation of algebra and natural language. Psychol. Sci. 23 (8), 914–922. http://dx.doi.org/10.1177/0956797612437427.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., & Weiller, C., 2003. Broca's area and the language instinct. Nat. Neurosci. 6 (7), 774–781.

Nakai, T., & Sakai, K.L., 2014. Neural mechanisms underlying the computation of hierarchical tree structures in mathematics. PLoS ONE 9 (11), e111439.

Ohta, S., Fukui, N., & Sakai, K.L., 2013. Syntactic computation in the human brain: the degree of merger as a key factor. PLoS ONE 8 (2), e56230.

Pallier, C., Devauchelle, A.D., & Dehaene, S., 2011. Cortical representation of the constituent structure of sentences. Proc. Natl. Acad. Sci. 108 (6), 2522–2527.

Pesenti, M., Thioux, M., Seron, X., & Volder, A.D., 2000. Neuroanatomical substrates of Arabic number processing, numerical comparison, and simple addition: a PET study. J. Cogn. Neurosci. 12 (3), 461–479.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S., 2007. A magnitude code common to numerosities and number symbols in human intraparietal cortex. Neuron 53 (2), 293–305.

Pica, P., Lemer, C., Izard, V., & Dehaene, S., 2004. Exact and approximate arithmetic in an Amazonian indigene group. Science 306 (5695), 499–503.

Price, C.J., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann. N. Y. Acad. Sci. 1191 (1), 62–88.

Stanescu-Cosson, R., Pinel, P., van de Moortele, P.F., Le Bihan, D., Cohen, L., & Dehaene, S., 2000. Understanding dissociations in dyscalculia: a brain imaging study of the impact of number size on the cerebral networks for exact and approximate calculation. Brain 123 (11), 2240–2255.

Suzuki, K., & Sakai, K.L., 2003. An event-related fMRI study of explicit syntactic processing of normal/anomalous sentences in contrast to implicit syntactic processing. Cereb. Cortex 13 (5), 517–526.

Tan, L.H., Laird, A.R., Li, K., & Fox, P.T., 2005. Neuroanatomical correlates of phonological processing of Chinese characters and alphabetic words: a meta-analysis. Hum. Brain Mapp. 25 (1), 83–91.

Varley, R.A., Klessinger, N.J., Romanowski, C.A., & Siegal, M., 2005. Agrammatic but numerate. Proc. Natl. Acad. Sci. U. S. A. 102 (9), 3519–3524.

Xiang, H.D., Fonteijn, H.M., Norris, D.G., & Hagoort, P., 2010. Topographical functional connectivity pattern in the perisylvian language networks. Cereb. Cortex 20 (3), 549–560.