Can the meaning of multiple words be integrated unconsciously?

Simon van Gaal, Lionel Naccache, Julia D. I. Meuwese, Anouk M. van Loon, Alexandra H. Leighton, Laurent Cohen and Stanislas Dehaene

Supplementary data	"Data Supplement" http://rstb.royalsocietypublishing.org/content/suppl/2014/03/10/rstb.2013.0212.DC1.html
References	**This article cites 72 articles, 10 of which can be accessed free** http://rstb.royalsocietypublishing.org/content/369/1641/20130212.full.html#ref-list-1
Subject collections	Articles on similar topics can be found in the following collections behaviour (510 articles) cognition (332 articles) neuroscience (451 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

Research

Royal Society Publishing

Can the meaning of multiple words be integrated unconsciously?

Simon van Gaal[1,2,3,4,5], Lionel Naccache[3,6,7], Julia D. I. Meuwese[4], Anouk M. van Loon[4], Alexandra H. Leighton[4], Laurent Cohen[3,6,7] and Stanislas Dehaene[1,2,8,9]

[1]INSERM, Cognitive Neuroimaging Unit, Gif-sur-Yvette, France
[2]Commissarìat à l'Energie Atomique, Neurospin Center, Gif-sur-Yvette, France
[3]INSERM, ICM Research Center, UMRS 975, Paris, France
[4]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
[5]Donders Institute for Brain, Cognition and Behavior, Nijmegen, The Netherlands
[6]University Paris 6, Faculté de Médecine Pitié-Salpêtrière, Paris, France
[7]Department of Neurophysiology and Neurology, AP-HP, Groupe hospitalier Pitié-Salpêtrière, Paris, France
[8]Université Paris-Sud 11, Orsay, France
[9]Collège de France, Paris 75005, France

What are the limits of unconscious language processing? Can language circuits process simple grammatical constructions unconsciously and integrate the meaning of several unseen words? Using behavioural priming and electroencephalography (EEG), we studied a specific rule-based linguistic operation traditionally thought to require conscious cognitive control: the negation of valence. In a masked priming paradigm, two masked words were successively (Experiment 1) or simultaneously presented (Experiment 2), a modifier ('not'/'very') and an adjective (e.g. 'good'/'bad'), followed by a visible target noun (e.g. 'peace'/'murder'). Subjects indicated whether the target noun had a positive or negative valence. The combination of these three words could either be contextually consistent (e.g. 'very bad - murder') or inconsistent (e.g. 'not bad - murder'). EEG recordings revealed that grammatical negations could unfold partly unconsciously, as reflected in similar occipito-parietal N400 effects for conscious and unconscious three-word sequences forming inconsistent combinations. However, only conscious word sequences elicited P600 effects, later in time. Overall, these results suggest that multiple unconscious words can be rapidly integrated and that an unconscious negation can automatically 'flip the sign' of an unconscious adjective. These findings not only extend the limits of subliminal combinatorial language processes, but also highlight how consciousness modulates the grammatical integration of multiple words.

1. Introduction

Mounting evidence suggests that unconscious cognition is very powerful. Brain-imaging studies have revealed subliminal information processing in many different brain areas, from low-level perceptual regions up to 'executive' areas in the prefrontal cortex [1,2]. To illustrate, recently, we have tested whether the ability to control ourselves and to inhibit our routine actions, a marker of the mind's central executive system, can be influenced unconsciously. In a so-called 'stop signal' task, it was tested whether refraining from responding requires consciousness of the instructing stimulus. When an unconscious 'stop' signal was briefly flashed and subsequently masked by the second stimulus, participants' slowed down their response to a subsequently presented 'go signal', without knowing why [3]. The invisible stop signal triggered a wave of brain activity that could be tracked by electroencephalographic recordings (EEG) and functional magnetic resonance imaging deep into the 'executive networks' of the frontal cortex that are crucial for controlling our actions [1,4–6]. Similarly, in recent years it has been shown that frontal executive networks of our brain can unconsciously register the occurrence of response errors [7–9], conflicting response alternatives [10–12] and competing task sets [13,14].

2

Although great progress has been made in characterizing the flow of information triggered by a single unconscious visual stimulus, whether and how multiple sources of unconscious information can be integrated are strongly debated. Generally, although several high-level cognitive processes can be influenced by subliminal stimuli they are likely not executed in the same way as during their conscious versions [2,15]. Influential models as well as experimental data suggest that consciousness is required for multiple-step rule-based algorithms, for example in doing mental arithmetic or performing grammatical language operations [16–18]. In contrast to unconscious stimulus processing, which seems relatively fleeting, conscious information processing might be accompanied by recurrent and long-lasting information sharing between distant brain regions. This unique feature of consciousness might allow information to be kept active in the brain for durations exceeding stimulus presentation times, and therefore facilitates the active integration and accumulation of information across time. Here, we used behavioural priming and EEG to explore whether, and if so how, the meaning of multiple subliminal (masked) words can also be integrated.

(a) Effects of subliminal words

Previous masking studies have observed that the processing of a visible target word can be modulated by a preceding single subliminal prime word, thereby demonstrating relatively high-level *semantic* analysis of unconscious words [11,19–21]. For example, it has been observed that when a masked word (e.g. 'table'), that activates a specific context, is rapidly followed by a related conscious target word (e.g. 'chair'), subjects respond faster and make fewer errors than when the same unconscious prime word is followed by an unrelated conscious target word (e.g. 'dog') (for a review of this much-debated but now-established field of research, see [22]). In EEG, incongruent prime–target pairs elicit a larger N400 event-related potential (ERP) component than congruent word pairs [19,23–26]. The N400 is a negative ERP deflection around 200–500 ms after the onset of the critical word and is typically associated with automatic semantic and syntactic violations [26–32]. Although it has been assumed for a long time that such unconscious processes are rather automatic and inflexible, recent studies show they are not [1,33,34]. One of the key criteria for automaticity is independence from top-down (cognitive) influences, and recent studies have demonstrated that several top-down factors (i.e. temporal and spatial attention, task strategy) affect the impact that unconscious stimuli have on behaviour and brain activity [25,35].

Recently, Batterink & Neville [36] studied whether *syntactic* violations can also be processed in the absence of awareness by using a cross-modal distraction task. They presented a tone either immediately before or after the onset of a visually presented syntactic violation (e.g. 'we drank Lisa's *by* brandy the fire in the lobby'). Subjects sometimes missed these violations because the tone attracted their attention and prevented the violation from reaching awareness (a cross-modal attentional blink paradigm). Interestingly, Batterink and Neville observed that both detected and undetected syntactic violations triggered a similar N400 response. However, crucially, the P600 was only observed when syntactic violations were consciously detected. The P600 peaks later (approx. 500–1000 ms; the exact timing depends on specific task parameters) and is associated with a wide range of grammaticality violations, and is thought to reflect more 'controlled' (and possibly conscious) language

processes. Overall, these results highlight that semantic and syntactic violations elicited by a single word can be automatically and unconsciously extracted.

(b) The possibility to integrate multiple unconscious stimuli over time

Recent studies even went one step further and tested whether the meaning of multiple subliminal stimuli (e.g. words, numbers) can also be integrated. Most prominently, Sklar *et al.* [37] used continuous flash suppression (CFS) to show that incongruent sentences ('I ironed coffee') break through interocular suppression ('pop into consciousness') faster than congruent sentences ('I ironed clothes'), suggesting that multiple words can be integrated, and semantic violations can be detected unconsciously (see also [38] for graphical violations). Similar results have been obtained for solving simple arithmetic equations, such as multiplication [39], addition [40] and subtraction [37]. Although intriguing, we and others [41] argue that it remains an open question whether such findings actually reflect the active combination of multiple unconscious stimuli. In fact, these effects might be explained by a much simpler mechanism, namely by stronger mnemonic associations between some pairs of individual concepts ('ironed'/'clothes') compared with others ('ironed'/'coffee'), without the need to integrate all words into a sentence (see [41] for a similar argument). Similarly, some complex linguistic constructions or mathematical facts may become stored in memory with common practice and sufficient exposure, possibly allowing *automatic* access to these stored representations via the mere spreading of activation [42–45]. Recently, we have shown that stimulus visibility matters when several successive arrows have to be integrated across time for accurate decision-making based on the total amount of evidence [46]. Although low-visibility masked arrows (slightly above the threshold of objective awareness) could be accumulated over time in a linear fashion, high-visibility arrows seemed to be accumulated up to a much higher level, leading to important changes in strategic top-down decision-making.

(c) The present study

Here, we studied directly the possibility of integrating multiple pieces of unconscious information and the potential neural differences between conscious and unconscious integration mechanisms, using a design that carefully separates single-word and multiple-word effects. Specifically, we focused on a rule-based grammatical operation traditionally thought to require conscious cognitive control: the negation of valence [47]. In our masking paradigm, the meaning of two (un)conscious words had to be integrated to either negate or strengthen the valence of an adjective (e.g. 'not happy'/'very happy'). Importantly, this grammatical operation could not rely on stored memory representations [41,42] and the two critical words were presented unconsciously (Greenwald's 'two-word challenge' [48]). We reasoned that if it is possible to integrate negation into the meaning of sequences of multiple unconscious words this would be reflected in N400 and/or P600 effects. Based on previous findings [19,25,36,49], we also expected that qualitative differences between conscious and unconscious integration/negation processes might be reflected in the relative modulations of the N400 (present
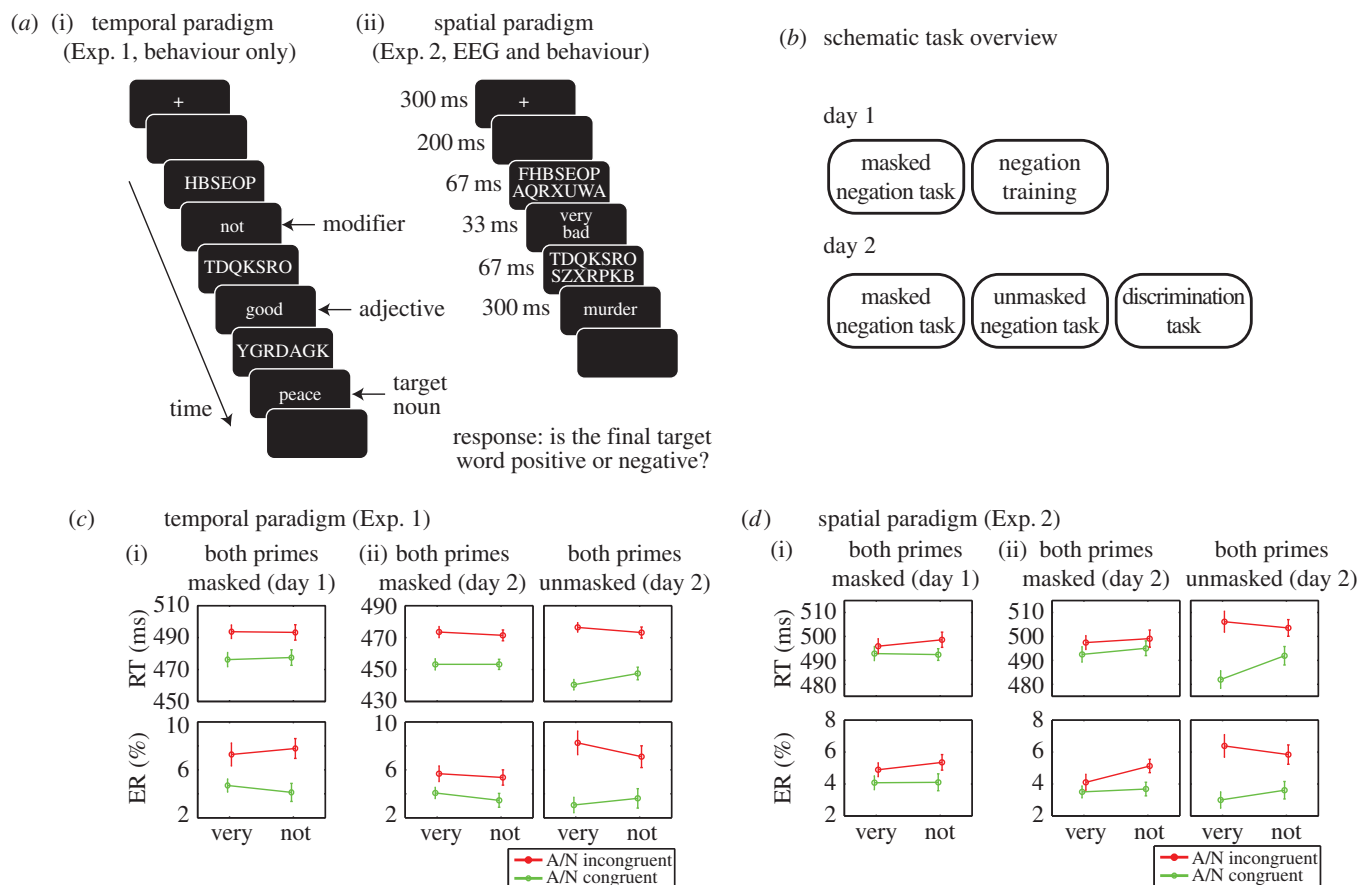
Figure 1. General set-up and behavioural results. (*a*) Overall trial structure and stimuli in the temporal (i) and spatial (ii) versions of the semantic negation task. The trials contained a modifier (not/very), an adjective (e.g. bad/good/sad/happy) and a noun target (e.g. peace/murder/war/love). (*b*) Schematic of the general structure of Experiments 1 and 2. (*c*) Response times (RT) and ER for congruent (e.g. bad – murder) and incongruent (e.g. bad – peace) adjective – noun (A/N) pairs separated for modifier identity (not/very) and visibility (masked/unmasked) are presented for Experiment 1. Results are presented for day 1 (before negation training (i)) and for day 2 (after negation training (ii)). (*d*) RT and ER results for congruent (e.g. bad – murder) and incongruent (e.g. bad – peace) adjective – target pairs, separated for modifier identity (not/very) and visibility (masked/unmasked) are presented for Experiment 2. Results are presented for day 1 (before negation training (i)) and for day 2 (after negation training (ii)). All error bars represent one standard error of the mean, after subtraction of each participant's grand mean. A, adjective; N, noun.

on both conscious and unconscious sentences) and the P600 (uniquely present on conscious sentences).

2. Material and methods

(a) Subjects

Seventy-six subjects participated in this study (18 in Experiment 1, 25 in Experiment 2, 17 in Experiment 3 and 16 in Experiment 4). Experiments 3 and 4 are reported in the electronic supplementary material. All subjects gave their written informed consent prior to participation, had normal or corrected-to-normal vision and were naive to the purpose of the experiments.

(b) Experiment 1: temporal negation paradigm

On masked trials, we presented a fixation cross (300 ms), a blank screen (200 ms), a mask (67 ms), the modifier (50 ms), a mask (67 ms), the adjective (50 ms), a mask (67 ms) and finally the target noun (300 ms). On unmasked trials, the same sequence of events was presented but all the masks were replaced by blank screens (figure 1*a*). Note that the duration of the modifier and adjective was rather long (50 ms), and therefore not fully masked from conscious awareness (see §3*a*). This was done to optimize our chances of finding automatic (although not unconscious) negation effects in behaviour, because previous studies had failed to find these effects for fully unconscious words [42]. Subjects were instructed to respond as fast as possible to

the target noun by indicating whether it had a positive or a negative valence by pressing a button with their right or left hand. In all experiments, the mapping of the response (negative/positive) was counterbalanced (left/right) across subjects. The interval between trials varied between 750 and 1250 ms (drawn from a random distribution).

We employed a 2 (modifier: 'not' or 'very') × 2 (adjective identity: positive or negative) × 2 (target identity: positive or negative) × 2 (visibility: masked/unmasked) factorial design. Five adjective–target pairs were used, each consisting of a positive and negative noun (e.g. death, peace, war, murder, smile, party; average word length of 4.9 letters, average frequency = 81.1) and a positive and negative adjective (e.g. happy, good, bad, sweet, angry, nice, scary). Two modifiers were used: 'not' and 'very' ('niet' and 'heel' in Dutch). For all adjective–noun pairs, nouns and adjectives were matched as much as possible in terms of overall frequency of appearance in daily Dutch language (80 versus 82 per 1 million, respectively, as stated in the Celex database [50]) and word length.

The masks consisted of seven randomly chosen uppercase letters, which were slightly overlapping to increase the density of the mask. The spacing between the centres of the letters was 10 pixels. Arial font was used with a font size of 20. Stimuli were presented in white against a black background at the centre of a 17 inch VGA monitor (frequency 60 Hz) using PRESENTATION software (Neurobehavioral Systems, Albany, NY, USA). Participants viewed the monitor from a distance of approximately 90 cm, so that each centimetre subtended a visual angle of 0.64°.

4

(c) Procedure

The experiment consisted of two separate sessions on separate days (see figure 1b for a schematic overview of the procedure). In the first session, participants first performed the temporal negation paradigm with masked modifiers/adjectives (described in §2b) for four blocks of 160 trials (two repetitions of each possible three-word combination per block, 'masked negation task' in figure 1b). Note that subjects had not been introduced to the words that were used as modifiers or adjectives before performing this task. Next, they performed the same task with the same words, except that the target noun was omitted and the modifier and adjective were presented consciously because all masks were replaced by blanks. Participants indicated as fast as possible whether the sequentially presented modifier–adjective combination was positive (e.g. very good, not bad) or negative (e.g. not good, very bad) (four blocks of 160 trials, 'negation training' in figure 1b). In the second session (1–7 days spaced from the first), subjects first performed four blocks of the masked (temporal) negation task with fully masked modifier–adjective pairs. Thereafter, they performed three blocks of the same task with only unmasked modifier–adjective pairs ('unmasked negation task' in figure 1b). Before each task, the subjects were briefly familiarized again with the task and performed 20 practice trials. By comparing performance on the masked negation paradigm on day 1 and day 2, we were able to test whether priming would get stronger owing to familiarity with the word-set as well as owing to the practice of the negation task. It has previously been suggested that practice in the negation process leads to changes in the underlying associative (memory-based) representation, without affecting the rule-based, more procedural components of performing negations [47].

At the end of the experiment (day 2), participants performed a forced-choice discrimination task to test the visibility of the modifiers and adjectives (80 trials for masked trials, 40 trials for unmasked trials), in separate blocks ('discrimination task' in figure 1b). The timing of the trials was the same as in the (masked/unmasked) negation task, but this time participants were instructed to focus on the modifier and the adjective and ignore the noun. At the end of the trial, subjects had to determine which of the four possible modifier–adjective combinations was presented. Each target noun was followed after 500 ms by four-choice options presented in capital letters (e.g. VERY HAPPY, NOT HAPPY, VERY SAD, NOT SAD) and presented around fixation (in a squarewise configuration). There was no speed stress on this discrimination response.

(d) Experiment 2: spatial negation paradigm— electroencephalographic experiment

In Experiment 2, a spatial configuration for the modifier and the adjective was used such that both words were presented at the same time, just above (modifier) and below (adjective) fixation (figure 1a(ii)) instead of the temporal configuration, as in the previous experiment. All procedures were similar to Experiment 1 except that subjects performed the negation training for two blocks in the second session, before we started our EEG recordings.

(e) Electroencephalographic measurements

EEGs were recorded and sampled at 256 Hz using a BioSemi ActiveTwo system (BioSemi, Amsterdam, The Netherlands). Sixty-four scalp electrodes were measured, as well as four electrodes for horizontal and vertical eye movements (each referenced to their counterpart) and two reference electrodes on the ear lobes. After acquisition, the EEG data were referenced to the average of both ears. EEG data were high-pass filtered at 0.5 Hz and then epoched from −1.5 to +2 s surrounding each trial. All trials were visually inspected and those containing artefacts not related to blinks were manually removed. Independent

components analysis was computed and components containing blink/oculomotor artefacts or other artefacts that could be clearly distinguished from brain-driven EEG signals were subtracted from the data. Baseline correction was applied by aligning time series to the average amplitude of the interval—from 400 to the 200 ms—preceding target onset (note that this is before the first stimulus in the trial sequence). For visualization purposes only, we applied a 30 Hz low-pass filter (figures 2 and 3). All pre-processing steps were done with EEGLAB. Statistical analysis was conducted using MATLAB (The Mathworks, Natick, MA, USA).

(f) Electroencephalographic analyses

To isolate the activity related to the semantic violations, congruent and incongruent target-locked trials were compared directly (congruency depended on the specific comparison performed). First, target-locked ERPs were calculated from the EEG data for all relevant comparisons for masked and unmasked trials separately. We focused our analysis on two typical ERP components with different latencies and polarities: the N400 and P600 [19,25–29,31,32,36,51–54]. The focus on the N400 and P600 components follows naturally from the previous literature on similar topics (e.g. semantics in language) [24,26–32,36,49,52]. Previous studies have shown that the N400 generally peaks at occipito-parietal electrodes, although sometimes more anteriorly, around frontal electrode sites [26–32,36,49,52]. It is difficult to anticipate the exact scalp topography of the N400 beforehand, because its topography has been observed to vary depending on the stimulus material, the exact manipulations, the modality and the specific task structure. The P600 seems to peak typically at (occipito-)parietal, or central sites, although also there is some variance. Because of the interest in two ERP effects (N400/P600), two different mechanisms ('adjective priming' versus 'negation') and differences related to stimulus visibility (masked/unmasked), the strongest hypotheses could be formed about the temporal aspects of the anticipated effects (N400: approx. 250–500 ms; P600: approx. 500–700 ms) and their polarity (N400: negative difference between incongruent and congruent trials; P600: positive difference between incongruent and congruent trials) (for reviews, see [27,51–54]). To zoom in on these specific effects, two regions of interest (ROI) were defined: a frontal ROI (FCz, FC1, FC2, FC3, FC4, Fz, F1, F2, F3, F4, AFz, AF3, AF4, FPz, see inset figure 2a) and an occipito-parietal ROI (Oz, POz, PO3, PO4, Pz, P1, P2, P3, P4, P5, P6, CPz, CP1, CP2, see inset figure 3a). Both ROIs consisted of 14 electrode channels, which increase the signal-to-noise ratio [31] (see e.g. [19,24,36] for similar electrode selection procedures).

To calculate the time-window at which the specific comparisons were different for incongruent from congruent conditions, we employed a moving average of 40 ms shifting every 10 ms (repeated-measures ANOVAs, factors depending on the comparisons, see §3c,d). As the first step, we focused on the 'adjective priming contrast' (figure 2) and applied false discovery rate (FDR) correction across the 0–700 ms time-window to the statistics of this effect. In this initial analysis, a significant interval was defined by all significant time-points between 0 and 700 ms after target presentation ($p_{fdr} < 0.05$, two-tailed, see §3c). For follow-up analyses, we focused only on those time-windows where a significant 'adjective priming effect' was found (see above and §3c,d: the N400 and P600 effects). In follow-up analyses, for the N400 we tested for a difference between 250 and 500 ms, and for the P600 for a difference between 450 and 700 ms ($p < 0.05$, two-tailed).

Although the early ERP components (less than 200 ms) were strongly affected by the differences in the stimulation procedure between masked and unmasked trials (basically by introducing masks in only the masked condition), the N400 and P600 ERP components peaked at roughly the same time for masked and unmasked
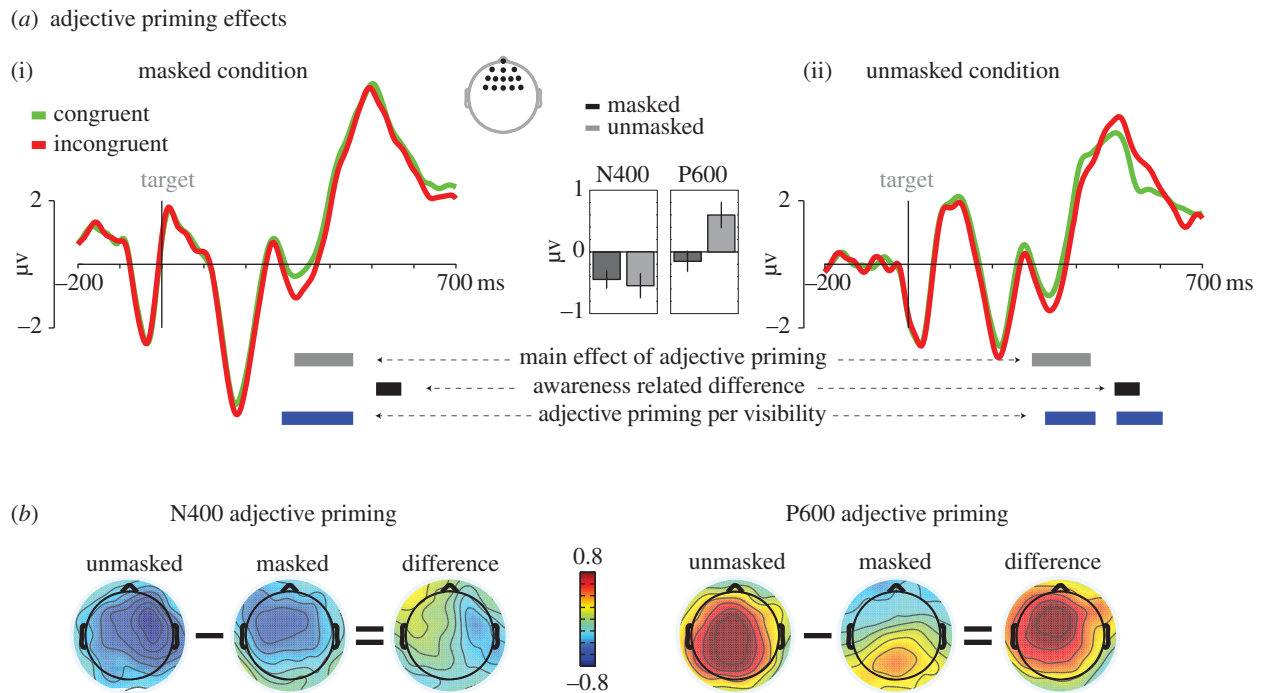
Figure 2. Adjective priming effect. (*a*) ERPs for a frontal region of interest (see inset) for congruent (green) and incongruent (red) adjective/noun (A/N) combinations, collapsed across modifier identity (not/very). Time 0 represents target noun onset. Horizontal grey bars represent time-windows at which the adjective priming effect ('main effect of adjective priming', two-way interaction: adjective × target noun) was significant across visibility conditions. Horizontal black bars represent the time-window at which the adjective priming was significantly larger for unmasked than masked trials ('awareness-related difference', three-way interaction: adjective × target noun × visibility). Horizontal blue bars represent time-windows at which the adjective priming effect was significant for each visibility condition separately (masked (i); unmasked (ii)) ('adjective priming per visibility'). The bar graph in the middle displays the adjective priming effect (incongruent − congruent) for the N400 (320−440 ms) and the P600 (490−600 ms) effects. (*b*) Head maps for the adjective priming effect for masked and unmasked trials, and their difference, for the N400 and P600 effects (same time-windows as bar graphs). Head maps represent the difference between incongruent and congruent A/N combinations. A, adjective; N, noun.

trials. This allowed us to directly compare these components between visibility conditions (using repeated-measures ANOVAs).

Three subjects were removed from all analyses in the EEG experiment: one subject because of excessive miss rates to the target (more than 40%) and two subjects because they scored significantly above-chance level in the forced-choice discrimination task administered after the main experiment (binominal test evaluated at $p < 0.05$, one-tailed). Therefore, we could not be sure that they were truly unable to perceive the masked words. This led to a total of 22 subjects in the EEG analyses.

3. Results

Our first aim was to explore whether negations can be performed unconsciously using a sequential presentation of the modifier and adjective preceding the nouns. In figure 1, conditions were grouped according to their adjective/noun (A/N) congruency as in typical masked priming paradigms (e.g. [19]) (e.g. congruent: bad/murder; incongruent: good/murder) and separated as a function of modifier identity ('very' versus 'not').

(a) Experiment 1: behaviour

Day 1: Before any negation training, adjective priming effects could already be observed on masked trials in response times (RT) and error rates (ER), as subjects were faster and more accurate on congruent than incongruent A/N pairs (two-way interaction of adjective × noun, RT: $F_{1,17} = 27.3$, $p < 0.001$; ER: $F_{1,17} = 13.8$, $p = 0.002$). This effect was not modulated by modifier identity (no three-way interaction of modifier ×

adjective × noun, hereafter termed 'negation effect', RT: $F_{1,17} = 0.4$, $p = 0.54$; ER: $F_{1,17} = 1.1$, $p = 0.31$, figure 1c). Note that we did not run the unmasked negation task on day 1. After the masked negation task, participants were familiarized with the negation process by indicating, as fast as possible, whether a sequentially presented modifier–adjective combination was either positive (e.g. 'very good', 'not sad') or negative (e.g. 'not good', 'very bad', see §2c). Subjects got considerably faster (from 723 ms in block 1 to 661 ms in block 4, $p < 0.001$) and more accurate (from 82% in block 1 to 90% in block 4, $p = 0.016$) across blocks while performing the negation training task.

Day 2: On day 2, we observed a clear adjective priming effect for masked trials (RT: $F_{1,17} = 35.2$, $p < 0.001$; ER: $F_{1,17} = 7.2$, $p = 0.016$). This effect was similar to day 1 (all ps > 0.087 for all possible interactions with the factor Session), so this form of priming was not affected by training. For unmasked trials, adjective priming was also observed (RT: $F_{1,17} = 109.3$, $p < 0.001$; ER: $F_{1,17} = 12.8$, $p = 0.002$), which was stronger than that on masked trials (RT: $F_{1,17} = 9.8$, $p = 0.006$; ER: $F_{1,17} = 6.5$, $p = 0.021$). Crucially, this adjective priming effect was only modulated by modifier identity (negation effect) in the unmasked condition (RT: $F_{1,17} = 10.7$, $p = 0.005$; ER: $F_{1,17} = 2.9$, $p = 0.105$), but not in the masked condition (RT: $F_{1,17} = 0.5$, $p = 0.48$; ER: $F_{1,17} = 0.3$, $p = 0.58$). These findings led to a trend towards a stronger negation effect for unmasked trials than masked trials (RT: $F_{1,17} = 3.0$, $p = 0.101$; ER: $F_{1,17} = 2.9$, $p = 0.108$). Note that, even for unmasked trials, there was no complete reversal of the effect depending on the nature of the modifier, i.e. 'not happy' was
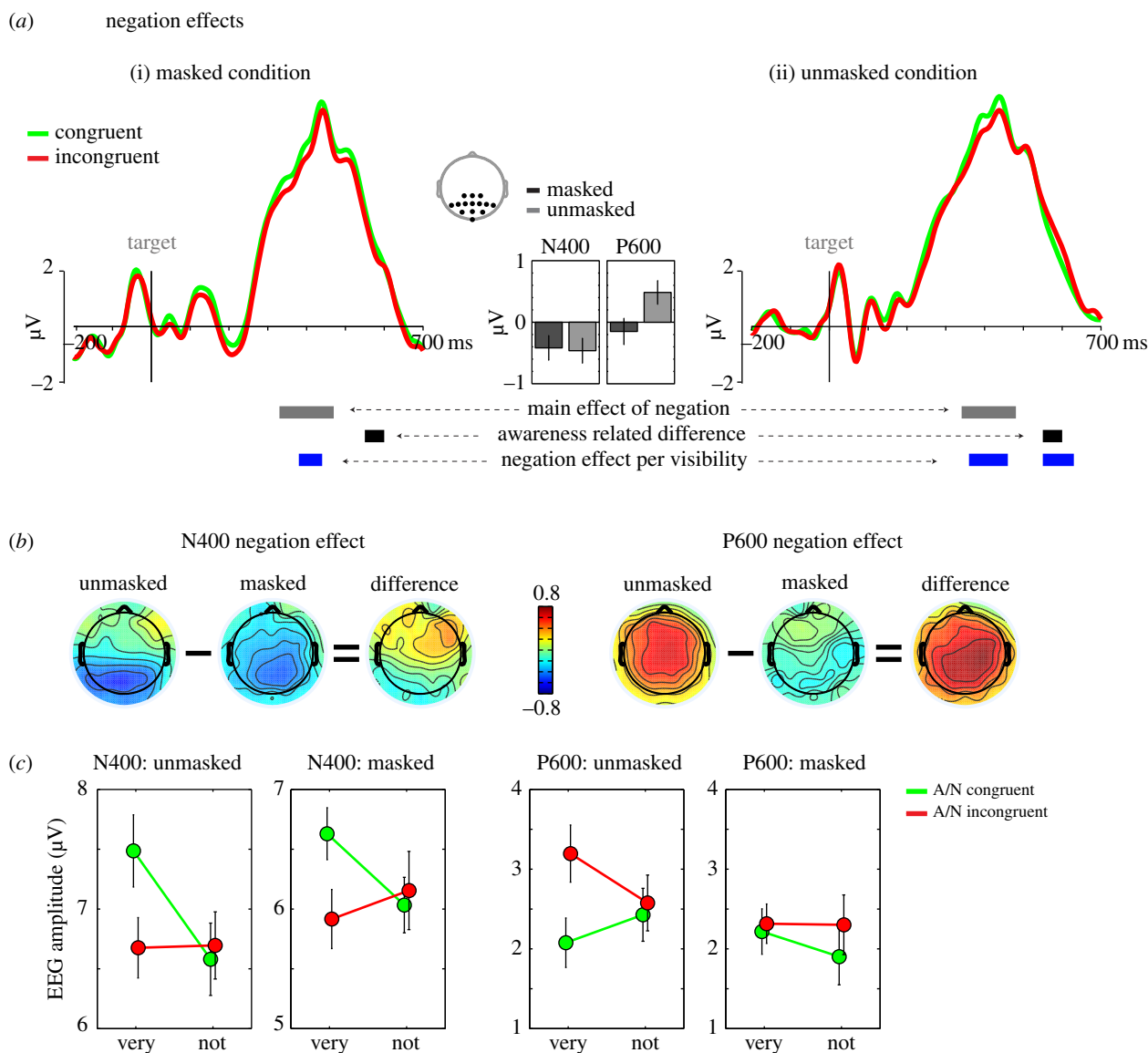
Figure 3. Negation effect. (*a*) ERPs for an occipito-parietal region of interest (see inset) for congruent (green) and incongruent (red) modifier/adjective/noun combinations. Time 0 represents target noun onset. Horizontal grey bars represent time-windows at which the negation effect ('main effect of negation', three-way interaction: modifier × adjective × target noun) was significant across visibility conditions. Horizontal black bars represent the time-window at which the negation effect was larger for unmasked than masked trials (four-way interaction: modifier × adjective × target noun × visibility, 'awareness-related difference'). Horizontal blue bars represent time-windows at which the negation effect was significant for each visibility condition separately ('negation effect per visibility'; masked (i); unmasked (ii)). (*b*) Head maps represent the negation effect for unmasked and masked trials (incongruent − congruent trials), and the difference, for the N400 (380–440 ms) and P600 (550–630 ms) effects. (*c*) Line plots showing the negation effect for the N400 and P600 (same time-windows as in (*b*)). Error bars represent one standard error of the mean, after subtraction of each participant's grand mean. A, adjective; N, noun.

not equivalent to 'sad' (no full crossover interaction, see [42] for similar behavioural results).

The forced-choice discrimination task administered at the end of the experiment revealed that masked words were perceived much more poorly than unmasked words (39.9% versus 90.8%, respectively, difference: $p < 0.001$), although subjects scored significantly above-chance level (25%) in both conditions (both ps < 0.001).

Our masking procedure did not render the words fully unconscious because of the relatively long prime duration (50 ms), which was chosen to maximize the chance of finding negation modulations in behaviour. However, still, no behavioural effects of negation were observed, replicating previous observations using the same prime duration [42]. In two follow-up behavioural experiments, we tested two outstanding questions. First, can a negation effect be observed when only the modifier is masked but the adjective is presented

consciously? Second, did the overall complexity of the paradigm or the rather long time between modifier and adjective (233 ms) prevent us from observing a negation effect? In these follow-up experiments, we reduced the prime duration to 33 ms per prime word to ensure that the primes were below the threshold of conscious awareness. These experiments revealed that, even when the adjective was presented fully consciously and only the modifier was masked, negation effects were not observed (see the electronic supplementary material). Further, the absence of unconscious negation effects is not likely owing to the overall complexity of the paradigm, because negation was also not observed in an experiment using a strongly simplified set-up (see the electronic supplementary material).

In Experiment 2, we took advantage of the high temporal resolution of EEG to study the neural processes related to conscious and unconscious negation processes. In this experiment, we made one more adjustment to the overall task structure: the

words were presented in a spatial arrangement instead of a sequential presentation (figure 1*a*(ii)). This allowed us to track the spatio-temporal neural processing of the modifier–adjective combination directly. Further, this change led to an SOA of 100 ms between the modifier–adjective pair and the target noun. As we shall see, although the spatial and temporal tasks were considerably different, the behavioural effects were remarkably similar.

(b) Experiment 2: behaviour

Day 1: Adjective priming effects were again observed on masked trials (RT: $F_{1,21} = 7.6$, $p = 0.012$; ER: $F_{1,21} = 9.3$, $p = 0.006$). Again, crucially, this effect was not modulated by modifier identity (RT: $F_{1,21} = 0.7$, $p = 0.42$; ER: $F_{1,21} = 0.3$, $p = 0.57$, figure 1*d*). Further, across the four blocks of negation training, participants got considerably faster (from 752 ms in block 1 to 696 ms in block 4, $p = 0.001$) and better (from 81% in block 1 to 89% in block 4, $p = 0.046$) at performing the negation operation.

Day 2: The adjective priming effects were again present for masked (RT: $F_{1,21} = 4.6$, $p = 0.043$; ER: $F_{1,21} = 7.2$, $p = 0.016$) and unmasked trials (RT: $F_{1,21} = 41.8$, $p < 0.001$; ER: $F_{1,21} = 12.8$, $p = 0.002$); however, these effects were stronger in the unmasked condition (RT: $F_{1,21} = 20.9$, $p < 0.001$; ER: $F_{1,21} = 6.5$, $p = 0.021$). Again, the negation effect was only present in the unmasked condition (RT: $F_{1,21} = 8.3$, $p = 0.009$; ER: $F_{1,21} = 2.9$, $p = 0.105$), and not in the masked condition (RT: $F_{1,21} = 0.1$, $p = 0.822$; ER: $F_{1,21} = 0.3$, $p = 0.581$), inducing a trend towards a stronger negation effect for unmasked than for masked trials (RT: $F_{1,21} = 3.8$, $p = 0.065$; ER: $F_{1,21} = 2.9$, $p = 0.108$). There were no significant differences between both sessions (all $p > 0.078$ for all possible interactions with the factor Session).

Discrimination performance for masked words was at chance level (25%) on a forced-choice discrimination task performed after the EEG experiment (26.4% correct, $t_{21} = 1.44$, $p = 0.16$; 75.1% correct for unmasked pairs, $t_{21} = 18.4$, $p < 0.001$, difference between masking conditions: $p < 0.001$). The commonly used frequentist statistics provide a measure of confidence in rejecting the null hypothesis, but not a measure of confidence in the null hypothesis itself. In order to quantify the confidence in the null hypothesis, we calculated Bayes factors for this experiment [55]. The performance on masked trials corresponds to a Bayes factor of 2.35 in favour of the null hypothesis, which reflects that there is moderate evidence in favour of the null hypothesis.

(c) Electroencephalographic correlates of adjective priming

Given the complicated structure of the task compared with 'typical' language studies, in conducting the ERP analyses, our first aim was to verify whether semantic violations were associated with the same electrophysiological markers as observed in previous studies, most prominently the N400 and P600 [19,21,25,27,35,36,49,52]. Therefore, target-locked ERPs from congruent and incongruent adjective/noun (A/N) pairs were compared ('adjective priming effect', figure 2*a*). For this initial analysis, we considered only the two-way congruency of the adjective and the noun, and therefore collapsed across modifier identity (e.g. 'very bad'/'murder' and 'not bad'/'murder' were both considered congruent) and observed that the adjective priming effect peaked at frontal electrode sites.

To isolate the ERP characteristics of significance (figure 2*a*, see also inset), we performed repeated-measures ANOVAs across the whole ERP time-window (0–700 ms) with three factors: visibility (masked, unmasked), adjective identity (positive, negative) and target identity (positive, negative). To solve the multiple comparison problem, we applied FDR correction across the entire 0–700 ms time-window to the statistics of these effects (see also [5,56] for similar procedures).

At the frontal ROI, only two significant effects were observed, corresponding to the N400 and P600 time periods (and no other moments in time). A main effect of adjective priming was observed for the N400 ($p_{fdr} < 0.05$ between 310 and 450 ms; corresponding to $p_{uncorrected} < 0.0072$, *ns* at the parietal ROI) and a significant adjective priming by visibility interaction for the P600 ($p_{fdr} < 0.05$ between 510 and 570 ms; corresponding to $p_{uncorrected} < 0.0015$; *ns* at the parietal ROI). See figure 2 for topographical maps and ERP waveforms for incongruent and congruent A/N pairs and the bar plots in figure 2*a* for the differences between conditions. Next, post hoc tests were performed on these two ERP effects, separately for masked and unmasked conditions. This revealed a significant N400 adjective priming effect for the masked ($p < 0.05$ between 280 and 450 ms; $F_{1,21} = 9.7$, $p = 0.005$ across the whole time-window), as well as the unmasked condition ($p < 0.05$ between 320 and 440 ms; $F_{1,21} = 7.7$, $p = 0.011$ across the whole time-window), which did not differ between both visibility conditions ($F_{1,21} = 0.2$, $p = 0.630$ for the entire 320–440 ms time-window). The pattern of results observed for the P600 was remarkably different. A significant P600 effect was observed for unmasked adjective priming ($p < 0.05$ between 490 and 600 ms; $F_{1,21} = 8.3$, $p = 0.009$ across the entire time-window), whereas no such priming was present for masked trials ($F_{1,21} = 0.8$, $p = 0.368$ for the same time-window).

At the occipito-parietal ROI, there was only a main effect of the P600, which was significant between 480 and 630 ms ($p_{fdr} < 0.05$, $p_{uncorrected} = 0.008$). A rather small P600 effect seemed to be present for masked trials, but this effect failed to reach statistical significance ($F_{1,21} = 2.6$, $p = 0.12$ for the 480–630 ms time-window, see figure 3*a*). The P600 effect was significant in the unmasked condition ($p < 0.05$ between 470 and 630 ms).

(d) Electroencephalographic correlates of negation

Next, we focused on the EEG pattern driven by the combined meaning of the modifier–adjective pair: the crucial 'negation effect'. We focused our analysis on the N400 (negative difference between 250 and 500 ms) and the P600 (positive difference between 450 and 700 ms) time-windows reported above (see also §2*f*). Target-locked ERPs from congruent and incongruent modifier–adjective–noun combinations are shown in figure 3*a*. For this analysis, modifier identity was crucial as we looked for the three-way congruity of modifier, adjective and noun (e.g. 'very bad'/'murder' = congruent; 'not bad'/'murder' = incongruent). Inspection of the differential ERP response between incongruent and congruent trials (figure 3*b*) revealed similar N400 and P600 effects, but this time at a more occipito-parietal cluster of electrodes, where the N400 and P600 often peak [26–29,51–53]. Note that these N400/P600 effects are relative ERP changes that ride on top of a large P3-like component (figure 3*a*).

Repeated-measures ANOVAs revealed a significant main effect of negation at the occipito-parietal ROI for the N400 ($p < 0.05$ between 330 and 460 ms; $F_{1,21} = 11.2$, $p = 0.003$ across this time-window; ns at the frontal ROI) and a significant negation by visibility interaction for the P600 ($p < 0.05$ between 550 and 600 ms; $F_{1,21} = 4.8$, $p = 0.039$ across this time-window; ns at the frontal ROI, see bar plots in figure 3a for differences between conditions). The N400 did not vary as a function of visibility ($F_{1,21} = 0.008$, $p = 0.929$ between 330 and 460 ms, figure 3b(i)) and was significant both for masked ($p < 0.05$ between 380 and 440 ms; $F_{1,21} = 4.4$, $p = 0.048$ across entire time-window) and unmasked trials ($p < 0.05$ between 360 and 460 ms; $F_{1,21} = 5.3$, $p = 0.032$ across entire time-window). Again, the P600 effects were rather different between visibility conditions. We observed a significant negation effect on the P600 for the unmasked condition ($p < 0.05$ between 550 and 630 ms; $F_{1,21} = 8.3$, $p = 0.009$ across the entire time-window), but not for the masked condition ($F_{1,21} = 0.8$, $p = 0.368$ for the 550–630 ms time-window). Figure 3c shows the average amplitude for congruent A/N and incongruent A/N combinations, either combined with the modifier 'very' or the modifier 'not'. These plots can be compared qualitatively with the behavioural data plots shown in figure 1d.

4. Discussion

In this work, we used behavioural priming and EEG to explore whether and how the semantic meaning of multiple subliminal words can be integrated. Specifically, we examined the possibility of integrating the meaning of two masked words and we tested how the combined meaning of these words could affect the processing of a consciously presented target word later in time. The processing of grammatical negation is thought to a rather high level and demanding linguistic operation, which has been proposed to require slow, conscious cognitive resources [42,57].

First, we replicated the single-word unconscious priming effects observed previously. Subjects responded faster and more accurately to congruent adjective–noun pairs (e.g. 'bad/murder', 'good/peace') than incongruent pairs (e.g. 'good/murder', 'bad/peace'), even when the adjective was fully masked (Experiment 2). This adjective priming effect did not depend on initial word exposure or training and was observed even when the adjective had never been seen consciously before, thus refuting any interpretation based on automatized response mappings [58,59]. In EEG, we observed a typical biphasic sequence of effects, often observed in language studies when comparing incongruent and congruent conditions, namely an N400 effect followed by a P600 effect [27–29,32]. As expected, the N400 effect was observed when comparing incongruent adjective–noun pairs with congruent adjective–noun pairs [19]. These N400 effects were remarkably similar for adjectives presented consciously (unmasked) and unconsciously (masked). The N400 thus seems to index an unconscious process of semantic integration whose amplitude and latency is largely unaffected by manipulations of awareness, whether induced by masking (present work) or by the attentional blink [23]. In sharp contrast, the P600 effect was strongly modulated by adjective awareness: adjective–noun congruency modulations were observed on the P600 in the conscious case, but not (or at least much less so) in the unconscious case.

Although the adjective priming effect was clearly present, we did not observe any grammatical negation effects in behaviour, replicating previous behavioural findings using a similar task set-up [42]. Potential negation effects remained undetectable even after substantial negation training (see also [47]). Based on these behavioural results alone, one might have concluded that multiple-word integration, and specifically the negation of valence, cannot be performed unconsciously (see Draine [42] for similar conclusions).

However, remarkably, the EEG results revealed a rather different picture and demonstrated that neuroimaging can sometimes dissect unconscious processing stages that cannot be detected by behaviour. In EEG, subliminal negation effects were observed when we examined the combinatorial effect of the modifier–adjective pair on the processing of the subsequent target noun. Most importantly, we observed an N400 effect (incongruent–congruent conditions) when taking into account the combined meaning of the modifier and the adjective, both for masked and unmasked word compositions. Interestingly, this N400 effect was again similar in strength for both visibility conditions. However, only conscious grammatical negations revealed a P600 effect. These results fit nicely with recent results obtained by Batterink & Neville [36], who showed that undetected (missed) syntactic violations triggered an early negativity, but that the P600 was uniquely modulated by the conscious registration of syntactic violations. In their case, all words of the sentence were perceived consciously, but only the crucial 'violating' word was missed because of attentional distraction. The present results extend these findings by showing that the same ERP modulations are observed when multiple-word sequences are fully masked, and therefore words have to be integrated unconsciously in order to detect semantic violations.

In the literature, several methods are used to gauge the visibility of stimulus material, ranging from subjective assessments ('seen–unseen' or a wider range of response options), confidence judgements and post-decision wagering methods, to rigorous forced-choice discrimination tasks [60–62]. The demonstration of chance-level performance in the latter is generally considered the strongest evidence for invisibility. However, whether discrimination performance exceeds chance-level performance (i.e. reaches significance) depends on several factors, including the number of trials used for the discrimination task and the number of participants (see [63,64] for a more elaborate discussion on this issue). In the present experiments, performance on the discrimination tasks confirmed that the visibility of masked and unmasked words differed strongly and also suggested that subjects could not perceive the masked words (using traditional t-tests and Bayesian statistics). However, importantly, we have demonstrated *qualitative* differences in the processing of masked versus unmasked negated word sequences. Namely, the N400 effects were highly similar, and equally strong, for masked and unmasked three-word sequences, whereas the P600 effects were strongly affected by stimulus visibility. Such qualitative differences are considered as convincing evidence for subliminal perception and clearly show that certain semantic processes are not affected by visibility (as reflected in the N400), whereas others are (as reflected in the P600) [46,65,66]. Interestingly, recently, Armstrong & Dienes [41] have obtained behavioural evidence for unconscious negation operations using subjective threshold measurements, instead of objective threshold measurements (as in the present experiments).

How can these awareness-related neural differences on the N400 and P600 effects be explained? Generally, the N400 is associated with (simple) semantic or lexical violations [29,51], whereas the P600 has been associated with a wide range of grammaticality violations as well as sentence complexity/ambiguity [67]. Further, the N400 is thought to reflect automatic processes that are specific to language, whereas the P600 might reflect more 'controlled' processes, including the 're-analysis', 'monitoring' or 'repair' of the inconsistent preceding language material [28,29,36,51,67–70]. These latter processes might be more domain-general and even non-linguistic [36,70]. We and others [36] speculate that the time-consuming re-analysis of the preceding word sequence, which relies on active working memory mechanisms, might be strongly impeded by masking and might require conscious awareness.

There is further support for this idea from neuroimaging studies of visual awareness. It has been repeatedly observed that the *depth* of unconscious processing can be surprisingly powerful, but that the *type* of neural processing of conscious and unconscious information differs fundamentally [1,2,22,71]. Unconscious information seems to be processed mainly in a feed-forward manner, flowing from lower to higher level brain regions. This feed-forward wave of information processing is probably rather fleeting, and thus the lifetime of unconscious neural activations is severely limited. On the other hand, conscious information processing seems to be uniquely marked by dynamic and recurrent interactions between (distant) brain regions [2,72]. Recurrent interactions are self-sustained and might allow brain regions to dynamically share information over extended periods of time, arbitrarily longer than the initial duration of stimulus presentation ([18], see also [46]). With regard to the present results, unconscious neural activations probably decayed relatively rapidly and therefore did not allow for an effortful re-analysis of the past multi-word sequence, which was reflected in the absence of P600 modulations on unconscious negated three-word sequences. However, unconscious multiple-word integration mechanisms, reflected in the N400 component, might still be present, although of a rather fleeting nature.

Previously, in an impressive series of behavioural experiments, Draine [42] and Greenwald & Liu [45] explored unconscious multiple-word syntactic constructions. In line with the present results, they did not observe any behavioural evidence for negations using 'not'. In their case, this might have been caused by the fact that they presented the word 'not' at the left side of fixation (before the adjective, as in normal reading), with the potential caveat that this crucial word was mainly processed in the non-language dominant right hemisphere [73]. However, we did not observe any behavioural evidence when both words (modifier and adjective) were presented centrally. As our previous research in this domain has typically shown more robust and larger neural than behavioural effects [5], we speculate that EEG was more sensitive than behaviour in elucidating the underlying integrative (subliminal) language mechanisms because of its continuous high-resolution temporal sampling. Further, it has been shown that it takes time to integrate negation into the meaning representation of sentences and it has been suggested that this might happen just shortly before response preparation [31]. Therefore, EEG might reveal subtle effects which extend beyond the time of the response and cannot be detected by behaviour alone. A benefit of the absence of behavioural differences between congruency conditions in the unconscious conditions is that the reported N400 effects cannot be trivially explained by differences in RT.

Interestingly, as Draine [42] and Greenwald & Liu [45] argued, negations come in many forms: in grammatical two-word negations (as studied here), but also in simpler lexical operations using prefixes, such as 'un', 'dis' or 'non'. Interestingly, Draine showed that masked negated adjectives with negated prefixes (e.g. 'unbad', 'ungood') slightly influenced the processing of a polarized target noun, as compared with control compound words. Interestingly, they came to a similar conclusion and suggested that 'linguistic input needs to be buffered in working memory in order for constructive processing to take place' [42]. Recently, Sklar *et al.* [37] showed the integration of multiple unconscious words with manipulations that allow for longer stimulus durations than masking, namely CFS. As argued in §1 and by others [41], these results might be explained by the fact that some initially complex linguistic constructions, especially common expressions, as used in their experiments, may become fixed representations in long-term memory with practice and sufficient exposure [42,43]. However, differences between studies might also be caused by differences in paradigms. CFS allows for longer unconscious stimulus presentations, with the potential to observe stronger behavioural effects. Future studies, combining negation with CFS might be able to reveal stronger and conclusive behavioural evidence for unconscious negation effects.

In summary, the present findings reveal that a complex set of combinatorial language computations can partly unfold unconsciously and automatically, as reflected in N400 ERP modulations. However, at the same time, these results also demonstrated crucial neural differences between conscious and unconscious combinatorial negation processes, as reflected in the later P600 component. Therefore, these results reconcile recent behavioural evidence for unconscious multiple-word integration [37] with ERP evidence for an awareness-related dissociation between the N400 and P600 components [36] and with leading theories that propose a critical role of consciousness in multiple-step algorithms [18,74,75].

References

1. van Gaal S, Lamme VAF. 2012 Unconscious high-level information processing: implications for neurobiological theories of consciousness. *Neuroscientist* **18**, 287–301. (doi:10.1177/1073858411404079)

2. Dehaene S, Changeux JP. 2011 Experimental and theoretical approaches to conscious processing. *Neuron* **70**, 200–227. (doi:10.1016/j.neuron.2011.03.018)

3. van Gaal S, Ridderinkhof KR, van den Wildenberg WPM, Lamme VAF. 2009 Dissociating consciousness from inhibitory control: evidence for unconsciously triggered inhibitory control in the stop-signal paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 1129–1139. (doi:10.1037/a0013551)

4. van Gaal S, Lamme VAF, Fahrenfort JJ, Ridderinkhof KR. 2011 Dissociable brain mechanisms underlying the conscious and unconscious control of behavior. *J. Cogn. Neurosci.* **23**, 91–105. (doi:10.1162/jocn.2010.21431)

5. van Gaal S, Ridderinkhof KR, Fahrenfort JJ, Scholte HS, Lamme VAF. 2008 Frontal cortex mediates unconsciously triggered inhibitory control. *J. Neurosci.* **28**, 8053–8062. (doi:10.1523/JNEUROSCI.1278-08.2008)

6. van Gaal S, Ridderinkhof KR, Scholte HS, Lamme VAF. 2010 Unconscious activation of the prefrontal no-go network. *J. Neurosci.* **30**, 4143–4150. (doi:10.1523/JNEUROSCI.2992-09.2010)

7. Nieuwenhuis S, Ridderinkhof KR, Blom J, Band GPH, Kok A. 2001 Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology* **38**, 752–760. (doi:10.1111/1469-8986.3850752)

8. Charles L, Van Opstal F, Marti S, Dehaene S. 2013 Distinct brain mechanisms for conscious versus subliminal error detection. *NeuroImage* **73**, 80–94. (doi:10.1016/j.neuroimage.2013.01.054)

9. Hester R, Foxe JJ, Molholm S, Shpaner M, Garavan H. 2005 Neural mechanisms involved in error processing: a comparison of errors made with and without awareness. *Neuroimage* **27**, 602–608. (doi:10.1016/j.neuroimage.2005.04.035)

10. van Gaal S, Scholte HS, Lamme VAF, Fahrenfort JJ, Ridderinkhof KR. 2011 Pre-SMA gray-matter density predicts individual differences in action selection in the face of conscious and unconscious response conflict. *J. Cogn. Neurosci.* **23**, 382–390. (doi:10.1162/jocn.2010.21444)

11. Dehaene S, Naccache L, Le Clec'H G, Koechlin E, Mueller M, Dehaene-Lambertz G, van de Moortele P-F, Le Bihan D. 1998 Imaging unconscious semantic priming. *Nature* **395**, 597–600. (doi:10.1038/26967)

12. D'Ostilio K, Garraux G. 2012 Dissociation between unconscious motor response facilitation and conflict in medial frontal areas. *Eur. J. Neurosci.* **35**, 332–340. (doi:10.1111/j.1460-9568.2011.07941.x)

13. Lau HC, Passingham RE. 2007 Unconscious activation of the cognitive control system in the human prefrontal cortex. *J. Neurosci.* **27**, 5805–5811. (doi:10.1523/JNEUROSCI.4335-06.2007)

14. Reuss H, Kiesel A, Kunde W, Hommel B. 2011 Unconscious activation of task sets. *Conscious. Cogn.* **20**, 556–567. (doi:10.1016/j.concog.2011.02.014)

15. van Gaal S, de Lange FP, Cohen MX. 2012 The role of consciousness in cognitive control and decision making. *Front. Hum. Neurosci.* **6**, 121. (doi:10.3389/fnhum.2012.00121)

16. Baumeister RF, Masicampo EJ. 2010 Conscious thought is for facilitating social and cultural interactions: how mental simulations serve the animal–culture interface. *Psychol. Rev.* **117**, 945–971. (doi:10.1037/a0019393)

17. Morewedge CK, Kahneman D. 2010 Associative processes in intuitive judgment. *Trends Cogn. Sci.* **14**, 435–440. (doi:10.1016/j.tics.2010.07.004)

18. Sackur J, Dehaene S. 2009 The cognitive architecture for chaining of two mental operations. *Cognition* **111**, 187–211. (doi:10.1016/j.cognition.2009.01.010)

19. Kiefer M, Spitzer M. 2000 Time course of conscious and unconscious semantic brain activation. *Neuroreport* **11**, 2401–2407. (doi:10.1097/00001756-200008030-00013)

20. Greenwald AG, Draine SC, Abrams RL. 1996 Three cognitive markers of unconscious semantic activation. *Science* **273**, 1699–1702. (doi:10.1126/science.273.5282.1699)

21. Deacon D, Hewitt S, Yang C-M, Nagata M. 2000 Event-related potential indices of semantic priming using masked and unmasked words: evidence that the N400 does not reflect a post-lexical process. *Cogn. Brain Res.* **9**, 137–146. (doi:10.1016/S0926-6410(99)00050-6)

22. Kouider S, Dehaene S. 2007 Levels of processing during non-conscious perception: a critical review of visual masking. *Phil. Trans. R. Soc. B* **362**, 857–875. (doi:10.1098/rstb.2007.2093)

23. Luck SJ, Vogel EK, Shapiro KL. 1996 Word meanings can be accessed but not reported during the attentional blink. *Nature* **383**, 616–618. (doi:10.1038/383616a0)

24. Kiefer M. 2002 The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. *Brain Res. Cogn. Brain Res.* **13**, 27–39. (doi:10.1016/S0926-6410(01)00085-4)

25. Kiefer M, Brendel D. 2006 Attentional modulation of unconscious 'automatic' processes: evidence from event-related potentials in a masked priming paradigm. *J. Cogn. Neurosci.* **18**, 184–198. (doi:10.1162/jocn.2006.18.2.184)

26. Holcomb PJ. 1993 Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology* **30**, 47–61. (doi:10.1111/j.1469-8986.1993.tb03204.x)

27. Lau EF, Phillips C, Poeppel D. 2008 A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* **9**, 920–933. (doi:10.1038/nrn2532)

28. Kuperberg GR. 2007 Neural mechanisms of language comprehension: challenges to syntax. *Brain Res.* **1146**, 23–49. (doi:10.1016/j.brainres.2006.12.063)

29. Friederici AD, Wartenburger I. 2010 Language and brain. *WIREs Cogn. Sci.* **1**, 150–159.

30. Kutas M, Hillyard SA. 1980 Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* **207**, 203–205. (doi:10.1126/science.7350657)

31. Ludtke J, Friedrich CK, De Filippis M, Kaup B. 2008 Event-related potential correlates of negation in a sentence-picture verification paradigm. *J. Cogn. Neurosci.* **20**, 1355–1370. (doi:10.1162/jocn.2008.20093)

32. Neville H, Nicol JL, Barss A, Forster KI, Garrett MF. 1991 Syntactically based sentence processing classes: evidence from event-related brain potentials. *J. Cogn. Neurosci.* **3**, 151–165. (doi:10.1162/jocn.1991.3.2.151)

33. Kiefer M, Martens U. 2010 Attentional sensitization of unconscious cognition: task sets modulate subsequent masked semantic priming. *J. Exp. Psychol. Gen.* **139**, 464–489. (doi:10.1037/a0019561)

34. Naccache L, Blandin E, Dehaene S. 2002 Unconscious masked priming depends on temporal attention. *Psychol. Sci.* **13**, 416–424. (doi:10.1111/1467-9280.00474)

35. Kiefer M, Adams SC, Zovko M. 2012 Attentional sensitization of unconscious visual processing: top-down influences on masked priming. *Adv. Cogn. Psychol.* **8**, 50–61.

36. Batterink L, Neville HJ. 2013 The human brain processes syntax in the absence of conscious awareness. *J. Neurosci.* **33**, 8528–8533. (doi:10.1523/JNEUROSCI.0618-13.2013)

37. Sklar AY, Levy N, Goldstein A, Mandel R, Maril A, Hassin RR. 2012 Reading and doing arithmetic nonconsciously. *Proc. Natl Acad. Sci. USA* **109**, 19 614–19 619. (doi:10.1073/pnas.1211645109)

38. Mudrik L, Breska A, Lamy D, Deouell LY. 2011 Integration without awareness: expanding the limits of unconscious processing. *Psychol. Sci.* **22**, 764–770. (doi:10.1177/0956797611408736)

39. Garcia-Orza J, Damas-Lopez J, Matas A, Rodriguez JM. 2009 '2×3' primes naming '6': evidence from masked priming. *Atten. Percept. Psychophys.* **71**, 471–480. (doi:10.3758/APP.71.3.471)

40. Ric F, Muller D. 2012 Unconscious addition: when we unconsciously initiate and follow arithmetic rules. *J. Exp. Psychol. Gen.* **141**, 222–226. (doi:10.1037/a0024608)

41. Armstrong AM, Dienes Z. 2013 Subliminal understanding of negation: unconscious control by subliminal processing of word pairs. *Conscious. Cogn.* **22**, 1022–1040. (doi:10.1016/j.concog.2013.06.010)

42. Draine SC. 1997 *Analytic limitations of unconscious language processing, in psychology*. Washington, DC: University of Washington.

43. Swinney DA, Cuttler A. 1997 The access and processing of idiomatic expressions. *J. Verbal Learn. Verbal Behav.* **18**, 523–534. (doi:10.1016/S0022-5371(79)90284-6)

44. Rusconi E, Priftis K, Rusconi ML, Umiltà C. 2006 Arithmetic priming from neglected numbers. *Cogn. Neuropsychol.* **23**, 227–239. (doi:10.1080/13594320500166381)

45. Greenwald AG, Liu TJ. 1985 Limited unconscious processing of meaning. *Bull. Psychon. Soc.* **23**, 292–313.

46. de Lange FP, van Gaal S, Lamme VAF, Dehaene S. 2011 How awareness changes the relative weights of evidence during human decision-making. *PLoS Biol.* **9**, e1001203. (doi:10.1371/journal.pbio.1001203)

47. Deutsch R, Gawronski B, Strack F. 2006 At the boundaries of automaticity: negation as reflective operation. *J. Pers. Soc. Psychol.* **91**, 385–405. (doi:10.1037/0022-3514.91.3.385)

48. Greenwald AG. 1992 New look 3: unconscious cognition reclaimed. *Am. Psychol.* **47**, 766–779. (doi:10.1037/0003-066X.47.6.766)

49. Holcomb PJ, Reder L, Misra M, Grainger J. 2005 The effects of prime visibility on ERP measures of masked priming. *Brain Res. Cogn. Brain Res.* **24**, 155–172. (doi:10.1016/j.cogbrainres.2005.01.003)

50. Baayen H, Piepenbrock R, Gulikers L. 1995 The CELEX lexical database [CD-ROM].

51. Kutas M, Federmeier KD. 2011 Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* **62**, 621–647. (doi:10.1146/annurev.psych.093008.131123)

52. Barber HA, Kutas M. 2007 Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Res. Rev.* **53**, 98–123. (doi:10.1016/j.brainresrev.2006.07.002)

53. Kutas M, Federmeier KD. 2000 Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* **4**, 463–470. (doi:10.1016/S1364-6613(00)01560-6)

54. Grainger J, Holcomb PJ. 2009 Watching the word go by: on the time-course of component processes in visual word recognition. *Lang. Linguist. Compass* **3**, 128–156. (doi:10.1111/j.1749-818X.2008.00121.x)

55. Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G. 2009 Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* **16**, 225–237. (doi:10.3758/PBR.16.2.225)

56. Fahrenfort JJ, Scholte HS, Lamme VAF. 2007 Masking disrupts reentrant processing in human visual cortex. *J. Cogn. Neurosci.* **19**, 1488–1497. (doi:10.1162/jocn.2007.19.9.1488)

57. Baumeister RF, Masicampo EJ, Vohs KD. 2012 Do conscious thoughts cause behavior? *Annu. Rev. Psychol.* **62**, 331–361. (doi:10.1146/annurev.psych.093008.131126)

58. Naccache L, Dehaene S. 2001 Unconscious semantic priming extends to novel unseen stimuli. *Cognition* **80**, 215–229. (doi:10.1016/S0010-0277(00)00139-6)

59. Damian MF. 2001 Congruity effects evoked by subliminally presented primes: automaticity rather than semantic processing. *J. Exp. Psychol. Hum. Percept. Perform.* **27**, 154–165. (doi:10.1037/0096-1523.27.1.154)

60. Overgaard M, Timmermans B, Sandberg K, Cleeremans A. 2010 Optimizing subjective measures of consciousness. *Conscious. Cogn.* **19**, 682–684 (discussion 685–686). (doi:10.1016/j.concog.2009.12.018)

61. Sandberg K, Timmermans B, Overgaard M, Cleeremans A. 2010 Measuring consciousness: is one measure better than the other? *Conscious. Cogn.* **19**, 1069–1078. (doi:10.1016/j.concog.2009.12.013)

62. Seth AK, Dienes Z, Cleeremans A, Overgaard M, Pessoa L. 2008 Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn. Sci.* **12**, 314–321. (doi:10.1016/j.tics.2008.04.008)

63. Dehaene S. 2008 Conscious and nonconscious processes: Distinct forms of evidence accumulation? In *Decision making, the human mind, and implications for institutions* (*Strüngmann forum reports*) (eds C Engel, W Singer), pp. 21–49. Cambridge, MA: MIT Press.

64. Hannula DE, Simons DJ, Cohen NJ. 2005 Imaging implicit perception: promise and pitfalls. *Nat. Rev. Neurosci.* **6**, 247–255. (doi:10.1038/nrn1630)

65. Jacoby LL. 1991 A process dissociation framework: separating automatic from intentional uses of memory. *J. Mem. Lang.* **30**, 513–541. (doi:10.1016/0749-596X(91)90025-F)

66. Merikle PM, Smilek D, Eastwood JD. 2001 Perception without awareness: perspectives from cognitive psychology. *Cognition* **79**, 115–134. (doi:10.1016/S0010-0277(00)00126-8)

67. Kolk H, Chwilla D. 2007 Late positivities in unusual situations. *Brain Lang.* **100**, 257–261. (doi:10.1016/j.bandl.2006.07.006)

68. van Herten M, Chwilla DJ, Kolk HH. 2006 When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *J. Cogn. Neurosci.* **18**, 1181–1197. (doi:10.1162/jocn.2006.18.7.1181)

69. Hagoort P. 2008 The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Phil. Trans. R. Soc. B* **363**, 1055–1069. (doi:10.1098/rstb.2007.2159)

70. Service E, Helenius P, Maury S, Salmelin R. 2007 Localization of syntactic and semantic brain responses using magnetoencephalography. *J. Cogn. Neurosci.* **19**, 1193–1205. (doi:10.1162/jocn.2007.19.7.1193)

71. Rees G. 2007 Neural correlates of the contents of visual awareness in humans. *Phil. Trans. R. Soc. B* **362**, 877–886. (doi:10.1098/rstb.2007.2094)

72. Lamme VAF. 2010 On how neuroscience will change or view on consciousness. *Cogn. Neurosci.* **1**, 204–240. (doi:10.1080/17588921003731586)

73. Nazir TA, Jacobs AM, O'Regan JK. 1998 Letter legibility and visual word recognition. *Mem. Cogn.* **26**, 810–821. (doi:10.3758/BF03211400)

74. Dehaene S, Naccache L. 2001 Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* **79**, 1–37. (doi:10.1016/S0010-0277(00)00123-2)

75. Dehaene S, Sigman M. 2012 From a single decision to a multi-step algorithm. *Curr. Opin. Neurobiol.* **22**, 937–945. (doi:10.1016/j.conb.2012.05.006)