No language unification without neural feedback: How awareness affects sentence processing

Valeria Mongelli [a,b,c,*], Erik L. Meijs [d], Simon van Gaal [b,c,d,1], Peter Hagoort [a,d,**,1]

[a] Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD, Nijmegen, the Netherlands
[b] Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WT, Amsterdam, the Netherlands
[c] Amsterdam Brain and Cognition (ABC), University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WT, Amsterdam, the Netherlands
[d] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Montessorilaan 3, 6525 HR, Nijmegen, the Netherlands

ABSTRACT

How does the human brain combine a finite number of words to form an infinite variety of sentences? According to the Memory, Unification and Control (MUC) model, sentence processing requires long-range feedback from the left inferior frontal cortex (LIFC) to left posterior temporal cortex (LPTC). Single word processing however may only require feedforward propagation of semantic information from sensory regions to LPTC. Here we tested the claim that long-range feedback is required for sentence processing by reducing visual awareness of words using a masking technique. Masking disrupts feedback processing while leaving feedforward processing relatively intact. Previous studies have shown that masked single words still elicit an N400 ERP effect, a neural signature of semantic incongruency. However, whether multiple words can be combined to form a sentence under reduced levels of awareness is controversial. To investigate this issue, we performed two experiments in which we measured electroencephalography (EEG) while 40 subjects performed a masked priming task. Words were presented either successively or simultaneously, thereby forming a short sentence that could be congruent or incongruent with a target picture. This sentence condition was compared with a typical single word condition. In the masked condition we only found an N400 effect for single words, whereas in the unmasked condition we observed an N400 effect for both unmasked sentences and single words. Our findings suggest that long-range feedback processing is required for sentence processing, but not for single word processing.

1. Introduction

Language processing involves single word and sentence processing. In both production and comprehension, single word processing is based on lexical retrieval from memory, whereas sentence processing also requires concatenation of retrieved lexical items and the combination of these items into novel meanings. Within the Memory, Unification and Control (MUC) model, the combinatorial aspect of language has been referred to as unification (Hagoort, 2017, 2013, 2005). Describing the neural network subserving unification, and what differentiates this network from related lexical retrieval mechanisms, is a major challenge in brain research (Dehaene et al., 2015; Hagoort, 2017). There is some evidence for a distribution of labor between left posterior temporal and left inferior frontal regions for lexical retrieval and unification mechanisms,

respectively (Snijders et al., 2009). According to the MUC model, lexical retrieval may be mainly dependent on feedforward propagation of semantic information from sensory regions to the left posterior temporal cortex (LPTC), whereas unification is thought to require feedback from the left inferior frontal cortex (LIFC), which includes Broca's area, and the LPTC (Hagoort, 2017, 2013; Hultén et al., 2019; Schoffelen et al., 2017; Tyler et al., 2011).

Here we tested the claim that feedback processing is required for unification. To this aim, we reduced visual awareness of words using a visual masking technique. Although still a matter of investigation, it seems that masking may disrupt both local feedback mechanisms in visual cortex (Fahrenfort et al., 2007), as well as more long-range feedback mechanisms (e.g. from frontal cortex, Del Cul et al., 2009), while leaving feedforward processing relatively intact (Fahrenfort et al., 2017, 2007;

Kovacs et al., 1995; Lamme and Roelfsema, 2000). Previous studies have shown that masked single words still trigger behavioral and neural signatures of semantic processing (see Kouider and Dehaene, 2007 for a review). For example, when a visible target (e.g. *cat*) is preceded by an unrelated masked prime (e.g. *bottle*), this results in longer reaction times and more response errors compared to prime-target pairs that are related (e.g. *dog-cat*) (Brown and Hagoort, 1993; Greenwald et al., 1996; Marcel, 1980). In electroencephalography (EEG), unrelated vs. related prime-target pairs trigger a larger N400 (Kiefer, 2002), a negative event-related potential (ERP) effect associated with violations of semantics, context and world knowledge (Kutas and Federmeier, 2010; Kutas and Hillyard, 1980; Lau et al., 2008). However, whether unification can also be achieved under reduced levels of awareness is highly controversial. There is evidence showing that simple combinatorial operations, like the negation of valence (e.g. *not good*), can be performed under masked processing (Armstrong and Dienes, 2013; van Gaal et al., 2014), although only after considerable training of the negation process (van Gaal et al., 2014). Some authors have suggested that masked combinatorial processing occurs for entire sentences (Armstrong and Dienes, 2014; Nakamura et al., 2018; Sklar et al., 2012), even though evidence is mixed (Rabagliati et al., 2018; Yang et al., 2017).

In this study, we tested the prediction that long-range feedback is necessary for sentence processing, but not (or less so) for single word processing. We addressed this issue by interfering with neural feedback by means of visual masking and investigated whether, under these conditions, sentence processing could still occur. To this aim, we designed a new experimental paradigm carefully disentangling sentence and single word processes. Using EEG, we measured brain responses to sequentially (Experiment 1) and simultaneously (Experiment 2) presented masked and unmasked short sentences (e.g. *man-pushes-woman*). Sentences were followed by a picture, which was either congruent or incongruent with the previous sentence. This sentence condition was compared with a single word condition, in which either masked/unmasked nouns (e.g. *man*) or verbs (e.g. *pushes*) were followed by congruent/incongruent pictures. In the unmasked condition, for both sentences and single words, we expected an N400 effect. In the masked condition, we also expected an N400 effect for single words, replicating previous findings (Deacon et al., 2000; Kiefer, 2002; Kiefer and Brendel, 2006; Kiefer and Spitzer, 2000). For masked sentences, there were two possible scenarios. If masked sentences triggered an N400 effect, similarly to masked single words, this would suggest that feedback processing (likely from frontal to temporal regions) is not required for unification. However, if no N400 effect was present for masked sentences, we may conclude that long-range feedback is required for unification to occur.

2. Experiment 1

2.1. Material and methods

2.1.1. Participants

44 subjects (34 females) participated in this study. 4 subjects did not complete the experiment, hence were excluded from the analyses. All subjects were right-handed, aged between 18 and 35, had normal or corrected-to-normal vision, no neurological history and were naive to the purpose of the experiment. Subjects all gave written informed consent prior to participation, according to the Declaration of Helsinki. They received 36 euros for their participation.

2.1.2. Stimuli

In the sentence condition, we employed three-word sentences composed of an agent, an action and a patient (e.g. *man pushes woman*, in Dutch *man duwt vrouw*). Importantly, sentences were obtained by combining words that are not lexically related. The nouns employed to fill the agent/patient roles were the following: *man, woman, boy, girl* (in Dutch: *man, vrouw, jongen, meisje*). The verbs employed to fill the action role were the following: *pushes, measures, drags, finds* (in Dutch: *duwt,*

meet, sleept, vindt). Sentences were presented sequentially, i.e. word by word. Each sentence was followed by a picture, either congruent or incongruent with the previous sentence (Fig. 1A). Pictures were selected from a previously tested database (Menenti et al., 2011). They depicted transitive events such as *pushing* or *dragging* and the agent/patient of this action. Each picture included one male and one female actor (either adults or children), with one of the two actors (either the male or the female) playing the agent role. The position of the agent (left or right) was randomized. The database was tested by Menenti et al. (2011) to check whether the actions depicted in the pictures were clear. Then, the verb which was most commonly used to describe each action was selected to create the word-picture pairs.
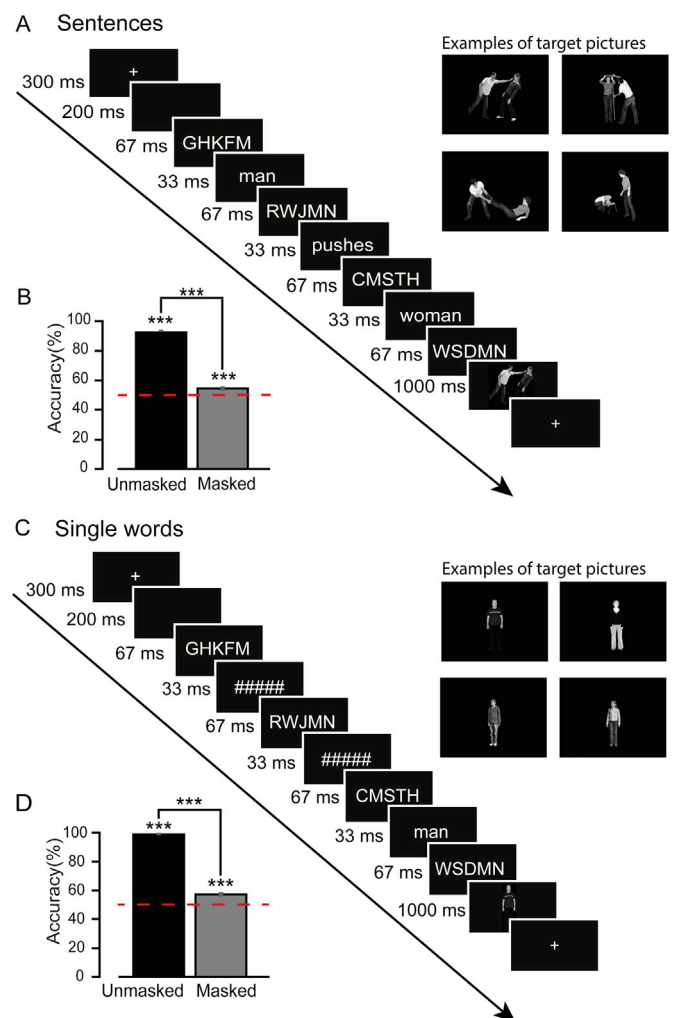
Fig. 1. Paradigm and behavioral results of Experiment 1. (A) Left panel: trial structure in the masked sentence condition. Three masked words forming a sentence were sequentially presented and followed by a target picture, which could be congruent or incongruent with the previous sentence. In the figure, a congruent trial is represented. When the fixation cross appeared after the picture, participants performed a match/no match task. Right panel: example of target pictures employed in the sentence condition. **(B)** Participants' accuracy in the sentence condition, for unmasked (black) and masked (grey) trials separately. Results are presented for Day 2 only. Error bars represent between-subject SEM. **(C)** Left panel: trial structure in the masked single word condition. One word occurring in the first, second or third prime position was presented and followed by a picture, either congruent or incongruent with the previous word. In the figure a congruent, third-position trial is represented. Right panel: example of target pictures employed in the single word condition. **(D)** Participants' accuracy in the single word condition, for unmasked (black) and masked (grey) trials separately. Results are presented for Day 2 only. Error bars represent between-subject SEM.

In the noun and verb condition, we employed respectively the same four nouns and four verbs that were used in the sentence condition. Importantly, both nouns and verbs occurred in the first, second or third prime position. The other positions were filled by hashtags, e.g.: 1. *man #### ####*; 2. *#### man ####*; 3. *#### #### man*. In the noun condition, words were followed by a picture depicting a man, a woman, a boy or a girl. In the verb condition, pictures were the same as employed in the sentence condition. As in the sentence condition, pictures were either congruent or incongruent with the previous noun/verb. In all conditions, pictures were in greyscale and had a size of 400 × 300 pixels.

The masks consisted of seven randomly chosen uppercase letters, which were slightly overlapping to increase the density of the mask. The space between the center of each letter was 10 pixels. Words and masks were presented in Arial lowercase font with a size of 20. Stimuli were presented in white against a black background. Participants were seated at a distance of approximately 110 cm from the computer screen.

2.1.3. Experimental design

The experiment was programmed using Presentation software (Neurobehavioral Systems, Albany, NY, USA). We employed a 2 (primes semantic complexity: sentence or single word) x 2 (target congruency: congruent or incongruent) x 2 (masking strength: masked or unmasked) factorial design. In the sentence condition, incongruency was equally divided across three subcases: (1) incongruent actors/congruent action (e.g. sentence: *man pushes woman*, followed by a picture of a woman pushing a man); (2) congruent actors/incongruent action (e.g. *man pushes woman*, followed by a picture of a man measuring a woman); (3) incongruent actors/incongruent action (e.g. *man pushes woman*, followed by a picture of a woman measuring a man). Each sentence block included 112 trials and lasted approximately 5 min, whereas each noun/verb block included 96 trials and lasted approximately 4 min. In total, there were 448 trials per subject in the sentence condition and 384 trials in each single word condition. In both sentence and single word conditions, each block included 50% masked and 50% unmasked trials, randomly mixed within blocks.

In masked trials, we presented a fixation cross (300 ms), a blank screen (200 ms), a first mask (67 ms), a first prime (33 ms), a second mask (67 ms), a second prime (33 ms), a third mask (67 ms), a third prime (33 ms), a fourth mask (67 ms) and a target picture (1000 ms). After the picture, a fixation cross appeared on the screen until an answer was given (see Fig. 1A and C for examples of the trial structure). After the answer, the fixation cross turned green and a new trial started. In unmasked trials, we presented a fixation cross (300 ms), a blank screen (200 ms), a first blank (67 ms), a first prime (67 ms), a second blank (33 ms), a second prime (67 ms), a third blank (33 ms), a third prime (67 ms), a fourth blank (33 ms) and a target picture (1000 ms), followed by a fixation cross. Unmasked primes had a longer duration than the masked primes in order to increase the visibility of the middle word, which otherwise would have been partially masked by the first and third words. In order to keep the duration of a trial constant (i.e. 1867 ms) across masking strength conditions, the duration of blanks in the unmasked condition was adjusted accordingly.

Participants were instructed to respond to the target picture only after the fixation cross appeared on screen. They had to indicate whether the picture matched or did not match with the previous sentence/word by pressing either a left or a right key on the keyboard with their left or right hand, respectively. Crucially, in the sentence condition, participants needed to combine the three words into a sentence in order to perform the task correctly.

In order to make the task meaningful in the masked condition, participants were told that each block included "easy" and "difficult" trials. In the difficult trials, words were "hidden", which made them very hard to be read. However, they had to try their best to respond correctly. Therefore, participants were aware that words were present in the masked condition, but they were not able to perceive them clearly. There was no speed stress on the discrimination response. Each block included

50% congruent and 50% incongruent trials, randomly mixed within blocks. The interval between trials varied between 750 and 1250 ms, drawn from a random distribution. From now on, we will refer to these word-picture trials as "picture trials".

In order to assess prime visibility, participants performed a forced-choice discrimination task in approximately 15% of the masked trials (i.e. 16 for the sentence blocks and 12 for the noun/verb blocks). Here, instead of a picture, a four-option screen occurred after masked primes. Participants were asked to decide which sentence or word they had just seen, and pick the correct option among the four presented. The four choices were presented in lowercase letters and around fixation, in a squarewise configuration. The correct choice was presented equally across the four possible positions around fixation (up-left, up-right, low-left, low-right). In the sentence condition, incorrect choices were equally divided across three incongruency subcases, as in the picture trials (incongruent actors/congruent action; congruent actors/incongruent action; incongruent actors/incongruent action). From now on, we will refer to these trials as "discrimination trials". Note that, in the main task, participants also needed to discriminate words in order to perform the match/no match task correctly. Therefore, the main task can also be considered as a discrimination task. The forced-choice discrimination task was added as a more traditional and supplementary visibility check. At the end of each block, performance feedback was provided on participants' accuracy, for both masked and unmasked trials.

2.1.4. Procedure

The experiment consisted of two separate sessions on separate days. In the first session, participants performed a short version of the task for 6 blocks (2 sentence, 2 noun and 2 verb blocks), with no EEG recording, in order to familiarize themselves with the task. They were introduced to both task and stimuli before starting the measurement. During the second session (1–7 days after the first), participants performed the main task (12 blocks: 4 sentence, 4 nouns and 4 verb blocks) and EEG was recorded. Before starting the EEG recording, participants were briefly reminded both task and stimuli and performed one practice block, randomly selected across block types.

2.1.5. Electroencephalographic measurements

The electroencephalogram (EEG) was recorded continuously from 61 active Ag/AgCL electrodes using an actiCap system (Brain Products GmbH). 59 of these electrodes were mounted on a cap with equidistant electrode montage. Two separate electrodes were placed on the left and right mastoid, respectively. Blinks were monitored through a separate electrode placed below the left eye and the closest frontal electrodes to the left eye. Horizontal eye movements were monitored through two electrodes placed on the lateral canthi next to the left and right eye. The ground electrode was placed on the forehead. EEG and EOG recordings were amplified through BrainAmp DC amplifiers. The EEG signal was recorded with a sampling rate of 500 Hz and filtered online with a band-pass filter between 0.016 and 100 Hz, then filtered offline with a high-pass filter of 0.1 Hz and a low-pass filter of 40 Hz. All electrode sites were re-referenced offline to the average of left and right mastoids and epoched from −0.867 (i.e. the beginning of a trial) to +1 s surrounding each trial. All ERPs were time-locked to target, i.e. picture, presentation. Next, trials were visually inspected and those containing artefacts not related to blinks were manually removed. Note that visual inspection was blind to the conditions of the experiment. Deleted channels were reconstructed based on a nearest neighbor approach. Independent component analysis (ICA) was computed and components containing blink/oculomotor artefacts were removed from the data. The pre-target interval between 567 and 367 ms before target (i.e. the duration of the blank before the first mask) was employed as a baseline.

2.1.6. Statistical analyses

Behavioral analyses (repeated measures ANOVAs and paired t-tests) were performed with in-house MATLAB scripts. For our ERP analyses, all

preprocessing steps were conducted with Fieldtrip (Oostenveld et al., 2011). ERP statistical analyses were performed using a combination of Fieldtrip and in-house MATLAB (The Mathworks, Natick, MA, USA) scripts. We focused on the N400 effect, defined as a difference in amplitude between two N400 components, i.e. negative event-related potentials (ERPs) evoked by meaningful stimuli (Baggio and Hagoort, 2011). Since its discovery, the N400 effect has been consistently observed in response to semantic violations (Kutas and Hillyard, 1980), although it is also known to be evoked by violations of context, expectancies and world knowledge (Baggio and Hagoort, 2011; Kutas and Federmeier, 2010; Lau et al., 2008). Previous studies showed that N400 usually peaks at centroparietal electrodes, although sometimes more anteriorly (Kutas and Federmeier, 2010; Lau et al., 2008). The typical N400 time window is 250–500 ms, although this may vary as well (Kutas and Federmeier, 2010; Lau et al., 2008). Thus, our strongest hypothesis was formulated around the polarity of the N400 effect, in that we expected a negative difference between incongruent and congruent trials.

As a first step, we performed cluster-based permutation tests (Maris and Oostenveld, 2007) on the time window 0–1 s from target (i.e. picture) onset to isolate significant ERP effects related to semantic incongruency (contrast incongruent vs. congruent, all conditions collapsed). A minimum of two neighboring electrodes had to pass the threshold of 0.05 to form a cluster. Cluster-based procedure was repeated 10000 times. This Monte-Carlo method generated a nonparametric estimate of the p-value representing the statistical significance of the originally identified cluster. This analysis isolated significant effects as clusters in time and space (in our case, the N400 effect). To visualize the evolution of significant clusters, we divided the time period (0–1 s) in 10 equal bins and plotted for each bin (i.e. 0–100, 100–200, etc.) all electrodes that were significant at least 50% of the time (as an example, cf. Fig. 2A). Subsequently, for follow-up analyses we isolated the exact time window and channels where the N400 effect was significant. To do so, the onset and offset of a cluster were defined as the time period around the maximum difference where the difference did not drop below 50% of this
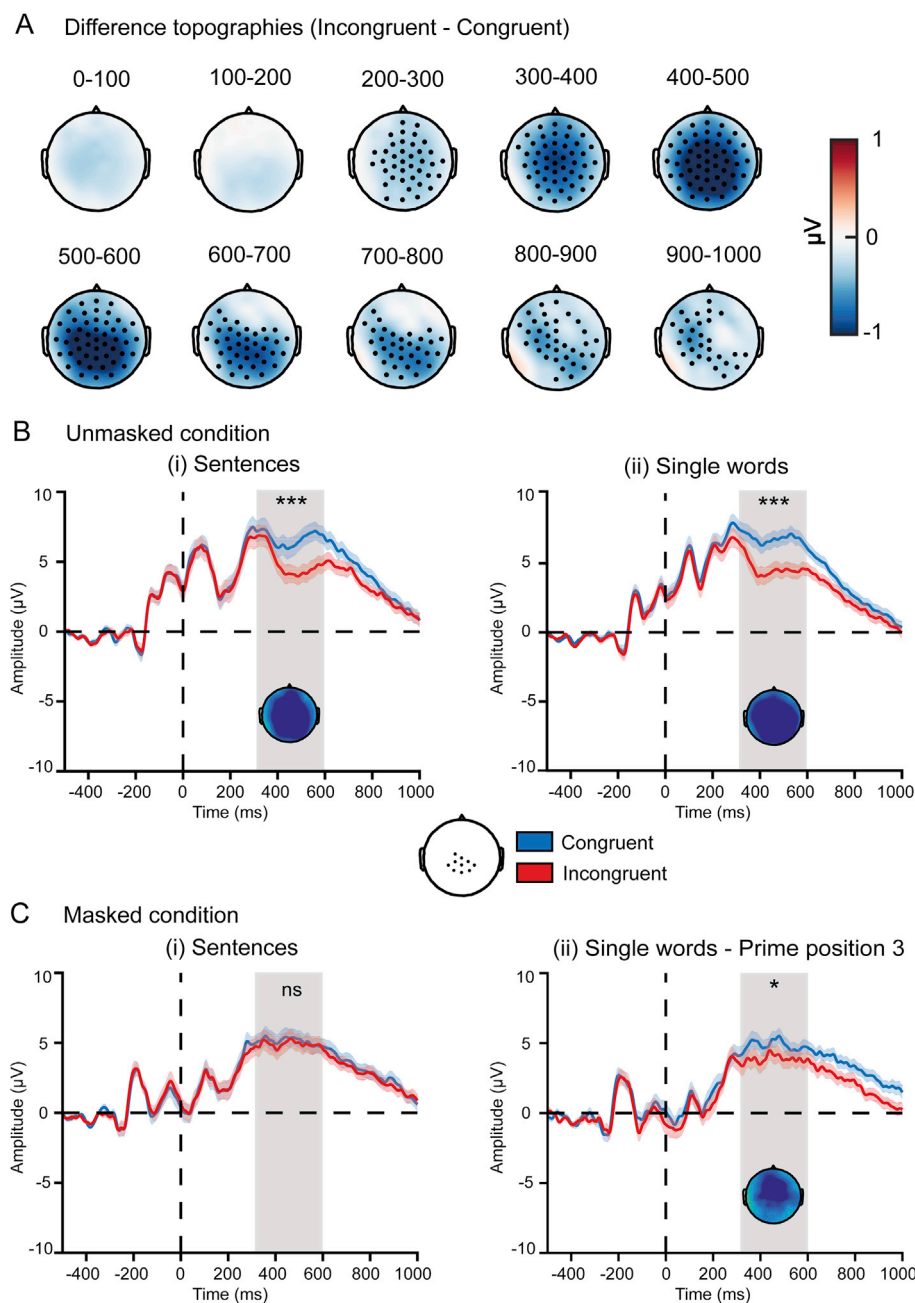
Fig. 2. N400 ERP effect of Experiment 1. (A) Topographic maps of the difference between incongruent and congruent trials over time (0 = target, i.e. picture onset). Cluster-based permutation tests were used to isolate the significant effects, while correcting for multiple comparisons across time and (electrode) space. On each head map, channels with a significant effect for at least 50% of its time window are highlighted. All headmaps are scaled from −1 to 1 μV. **(B)** ERPs for a central region of interest (see inset) for congruent (blue) and incongruent (red) unmasked trials, in the sentence (i) and single word (ii) condition. **(C)** ERPs for congruent (blue) and incongruent (red) masked trials, for sentences (i) and single words occurring in the third prime position (ii). In all figures shaded blue and red areas around the wave form indicate between-subject SEM. Time 0 represents the onset of the target picture. Shaded grey areas highlight the time window of interest for the N400 effect (336–598 ms). All headmaps are scaled from −1 to 1 μV.

maximum and where at least one channel showed a significant effect. We then selected the 10 channels within the cluster that showed the largest effect in this time window (for a similar procedure, cf. Meijs et al., 2018). The resulting region of interest (ROI) and time window were used for subsequent analyses including all the relevant factors. Note that, as a "sanity check", all analyses were also performed on a central ROI, defined based on previous literature on the N400 (Kutas and Federmeier, 2010). These supplementary analyses led to the same conclusions.

2.2. Results

Our goal was to investigate whether sentence processing requires neural feedback between distant brain areas. We addressed this issue by interfering with feedback processing by means of visual masking, and comparing masked and unmasked processing of sentences and single words.

2.2.1. Behavior

2.2.1.1. Day 1 (behavioral training). In these analyses, 36 out of the 40 participants were included, because the first four participants passed a slightly different version of the task in the training session. Overall, we observed that masking strength affected participants' accuracy (main effect of masking strength: $F_{1,35} > 1000$, $p < 0.001$), and that this effect was modulated by the semantic complexity of the primes (sentences vs. single words) (interaction masking strength x semantic complexity: $F_{1,35} = 4.24$, $p = 0.047$). Follow-up analyses revealed that performance was higher in the unmasked than in the masked condition (sentences: $t_{35} = 23.35$, $p < 0.001$; single words: $t_{35} = 45.37$, $p < 0.001$). In the unmasked condition, performances were very high and strongly above chance (sentences: $t_{35} = 31.70$, $p < 0.001$, 90% correct; single words: $t_{35} = 120.43$, $p < 0.001$, 98% correct). Although performance was poor in the masked condition, it was also above chance level (sentences: $t_{35} = 3.66$, $p < 0.001$, 54% correct; single words: $t_{35} = 10.21$, $p < 0.001$, 58% correct).

In forced-choice discrimination trials, masked words were also poorly discriminated (sentences: 29% correct; single words: 35% correct), although subjects scored significantly above chance level (sentences: $t_{35} = 3.12$, $p = 0.003$; single words: $t_{35} = 6.9$, $p < 0.001$). Note that chance level was 25% in this case, given that the task required a choice among four response options.

2.2.1.2. Day 2 (EEG recording). Participants performed similarly on day 1 and day 2 on the main task (all p-values > 0.063). Further, as on day 1, we observed that masking strength affected participants' accuracy (main effect of masking strength: $F_{1,39} > 1000$, $p < 0.001$) and performance was modulated by semantic complexity (interaction masking strength x semantic complexity: $F_{1,39} = 10.1$, $p = 0.003$). Again, performance was higher in the unmasked than in the masked condition (sentences: $t_{39} = 37.77$, $p < 0.001$; single words: $t_{39} = 54.31$, $p < 0.001$. Fig. 1B and D). In the unmasked condition, subjects scored very high and strongly above chance (sentences: $t_{35} = 38.44$, $p < 0.001$, 92% correct; single words: $t_{35} = 195.33$, $p < 0.001$, 98% correct. Fig. 1B and D). Although performance was poor in the masked condition, it was above chance level (sentences: $t_{39} = 5.63$, $p < 0.001$, 54% correct; single words: $t_{39} = 9.06$, $p < 0.001$, 57% correct. Fig. 1B and D). Moreover, we found that in single word trials, prime position affected participants' accuracy (main effect of prime position: $F_{2,78} = 6.25$, $p = 0.003$) and was modulated by masking strength (interaction prime position x masking strength: $F_{2,78} = 5.42$, $p = 0.006$). In the unmasked condition performance was equal in all prime positions (~98% correct in all positions, $F_{2,78} = 0.22$, $p = 0.790$), whereas in the masked condition it was lower when the prime was presented later in the trial (prime position 1: 60% correct, prime position 2: 56% correct, prime position 3: 54% correct, $F_{2,78} = 5.93$, $p = 0.004$), despite being above chance in all cases (all p's < 0.001).

In forced-choice discrimination trials, there was no significant difference between day 1 and day 2 ($p = 0.853$). As in the main task, masked words were poorly perceived (sentences: 29% correct; single words: 34% correct), although again subjects scored significantly above-chance level (sentences: $t_{39} = 4.63$, $p < 0.001$; single words: $t_{39} = 6.59$, $p < 0.001$).

Overall, the behavioral results revealed that participants perceived masked and unmasked trials very differently, as revealed by the fact that accuracy was much lower in the masked than in the unmasked condition (although accuracy was above chance in the main task as well as in the discrimination task for masked trials). This suggested that participants may have been partially aware of masked primes, most likely due to the nature of the task (they were told that in some trials words were present, but very difficult to be seen – see the Methods section for further details). Therefore, masking led to a quantitative difference in visibility ("matter of degree") rather than a qualitative difference.

2.2.2. ERPs: N400 modulations by prime visibility and semantic complexity

Next, we investigated whether neural signatures of semantic processing, reflected in the N400 effect, could be found for masked/unmasked sentences and single words. To this end, we contrasted congruent and incongruent trials using cluster-based permutation testing, correcting for multiple comparisons across both time (0–1 s) and (electrode) space (see Fig. 2A and Methods) (Maris and Oostenveld, 2007). When isolating significant effects as clusters in time and space (see Methods), we found a significant difference over centro-parietal electrodes from 336 to 598 ms after target onset and peaking at 490 ms (see Fig. 2A and – inset), consistently with a typical N400 time window and topography (p-value of the cluster between 336 and 598 ms: $p < 0.001$).

Within this cluster, we performed repeated measures ANOVAs on the N400 time window with three factors: masking strength (masked/unmasked), prime/target congruency (congruent/incongruent), and semantic complexity (sentence/single word) to test how, and if, the N400 effect was modulated by masking strength and semantic complexity. This ANOVA revealed that the N400 effect (main effect of congruency: $F_{1,39} = 104.06$, $p < 0.001$) was larger for unmasked than masked trials (interaction masking strength x congruency: $F_{1,39} = 66.80$, $p < 0.001$). Additionally, we found that the overall ERP amplitude was larger for unmasked than masked trials (main effect of masking strength: $F_{1,39} = 13.76$, $p < 0.001$). No other main effects or interactions were significant (all p-values > 0.234). Overall, these results show that the N400 effect was modulated by prime visibility. Therefore, in further planned post-hoc analyses we tested the N400 effect for masked and unmasked trials separately.

In the unmasked condition, a clear N400 effect was observed (main effect of congruency: $F_{1,39} = 94.42$, $p < 0.001$), which was not modulated by semantic complexity (interaction of congruency x semantic complexity: $F_{1,39} = 1.77$, $p = 0.190$). In both unmasked semantic complexity conditions, an N400 effect was observed (sentences: $t_{39} = 6.73$, $p < 0.001$; single words: $t_{39} = 9.49$, $p < 0.001$, see Fig. 2B).

In contrast, in the masked condition we did not find a significant N400 effect regardless of semantic complexity (main effect of congruency: $F_{1,39} = 2.21$, $p = 0.145$; interaction of congruency x semantic complexity: $F_{1,39} = 0.002$, $p = 0.961$). Given our hypotheses, we tested the congruency effect for sentences and single words separately, but no N400 effect was present in either condition (sentences: $t_{39=}0.78$, $p = 0.439$; single words: $t_{39} = 1.21$, $p = 0.232$).

Still, in our paradigm, single words could occur in different prime positions. Previous studies have shown that the Stimulus Onset Asynchrony (SOA) between prime and target affects the strength of semantic priming (Greenwald et al., 1996; Kiefer, 2002; Kiefer and Brendel, 2006; Nakamura et al., 2018). Therefore, we investigated whether prime position affected N400 amplitude. To do so, we split the trials according to the prime position, i.e. first, second and third, and used this as additional factor in an ANOVA. We found that prime position affected the N400 effect for single words in the masked condition (interaction congruency x

position: $F_{2,78} = 3.40$, $p = 0.038$), but not in the unmasked condition (interaction congruency x position: $F_{2,78} = 0.51$, $p = 0.603$). Interestingly, post-hoc t-tests revealed that an N400 effect for masked single words was only present when the prime was presented in the third position, i.e. immediately before target presentation ($t_{39} = 2.48$, $p = 0.017$, see Fig. 2C). In the other prime positions no significant N400 effect was observed (second prime position: $t_{39} = 1.10$, $p = 0.274$; first prime position: $t_{39} = 1.13$, $p = 0.261$).

In summary, our results for the unmasked condition revealed an N400 effect for both sentences and single words, showing that our paradigm worked as predicted. In the masked condition however, we found an N400 effect only for single words occurring in the third prime position, i.e. just before picture presentation. This suggests that processing of masked words is dependent on SOA and that, more generally, masked visual information is fleeting. It is unlikely that this effect was driven by prime visibility, because our behavioral results showed that performance (visibility) for third-position primes was actually lower than in the other two prime positions. We did not find any N400 effect for masked sentences, which may mean that sentence processing cannot be performed when disrupting feedback processing with masking. However, in our paradigm, sentences were presented sequentially, i.e. word by word. Given that, in the single word condition, we only found an effect for words presented at a short SOA, this null result could stem from the sequential presentation of the primes. If in the sentence condition the information conveyed by the first two words decayed very quickly, combining words into a sentence would not be possible. To further test this hypothesis, we performed an additional experiment in which words forming a sentence were presented simultaneously and at a short SOA (for a similar logic see van Gaal et al., 2014). We reasoned that, if we replicate the observation that there was only an N400 effect for masked single words in this second experiment, our results would suggest a true difference between sentence and single word processing, not biased by differences in the timing of stimulus presentation and the temporal profile and fleeting nature of masked word processing in the brain.

3. Experiment 2

3.1. Material and methods

3.1.1. Participants

40 subjects (25 females) participated in this study. All subjects were right-handed, aged between 18 and 35, had normal or corrected-to-normal vision, no neurological history and were naive to the purpose of the experiment. Subjects gave written informed consent prior to participation, according to the Declaration of Helsinki. They received 56 euros for their participation.

3.1.2. Stimuli

In the sentence condition, we employed two-word combinations composed by either agent/action pairs (e.g. *man pushes*, in Dutch *man duwt*) or patient/action pairs (e.g. *man is pushed*, in Dutch *man geduwd*). We will refer to these combinations as sentences. As in Experiment 1, sentences were obtained by combining words that are not lexically related. We chose to employ two-word combinations, instead of three-word combinations as in Experiment 1, for two reasons. First, we aimed at investigating whether an effect could be found for masked sentences when semantic integration was made easier. Second, we wanted to make our results more comparable to other studies finding an N400 effect for masked processing of simple two-word combinations (e.g. van Gaal et al., 2014). As in Experiment 1, the nouns employed to fill the agent/patient roles were the following: *man, woman, boy, girl* (in Dutch: *man, vrouw, jongen, meisje*). The verbs employed to fill the action role were the following: (i) active verbs: *pushes, measures, drags, finds* (in Dutch: *duwt, meet, sleept, vindt*); (ii) passive verbs: *is pushed, is measured, is dragged, is found* (in Dutch: *geduwd, gemeten, gesleept, gevonden*). The two words were presented one above the other.

In the noun and verb conditions, we employed respectively the same four nouns and four active/passive verbs that were used in the sentence condition. Pictures were extracted from the same database employed for Experiment 1. As in Experiment 1, pictures were either congruent or incongruent with the previous word(s). In all conditions, pictures had a size of 400×300 pixels. Masks were created in the same way as in Experiment 1. However, since in the case of passive verbs masked words were slightly longer, we increased the number of letters composing each mask (ten letters instead of seven). The parameters of stimuli presentation (e.g. participants' distance from the monitor) were the same as in Experiment 1.

3.1.3. Experimental design

The experiment was programmed using Presentation software (Neurobehavioral Systems, Albany, NY, USA). We employed a 2 (primes semantic complexity: sentence or single word) x 2 (target congruency: congruent or incongruent) x 2 (masking strength: masked or unmasked) factorial design. In the sentence condition, we built incongruency by reversing the actors' roles in the picture (e.g. if the prime was *man pushes*, then the picture depicted a woman pushing a man; if the prime was *man is pushed*, then the picture depicted a man pushing a woman). In both sentence and single word blocks, each block included 96 trials and lasted approximately 5 min. Each block included 50% masked and 50% unmasked trials, randomly mixed within blocks.

In masked trials, we presented a fixation cross (300 ms), a blank screen (200 ms), a first mask (67 ms), a prime (33 ms), a second mask (67 ms), a target picture (1000 ms) and a fixation cross (200 ms). After the fixation cross, an answer screen with the options *match/no match* (in Dutch: *wel match/geen match*) appeared (see Fig. 3A and C for examples of the trial structure). The position of the match answer (left or right) was randomized. In unmasked trials, we presented a fixation cross (300 ms), a blank screen (200 ms), a first blank (67 ms), a prime (67 ms), a second blank (33 ms), a target picture (1000 ms) and a fixation cross (200 ms). The unmasked prime had a longer duration than the masked prime in order to increase its visibility. To keep the duration of a trial constant (i.e. 1867 ms) across visibility conditions, the duration of blanks in the unmasked condition was adjusted accordingly. Due to a programming error, in the noun/verb main EEG session the duration of the second mask in the masked trials was 33 ms, whereas the duration of the second blank in the unmasked trials was 67 ms. However, this did not influence our results, as performances did not differ between the training session and the main EEG session (see the Results section for further details).

Contrary to Experiment 1, in which pictures in the noun condition only included one character, we employed the same two-character pictures as in the sentence and verb condition, in order to minimize task differences between the sentence and the single word condition. Since the noun primes referred to a single person (*man, woman, boy* or *girl*), and pictures depicted either a man and a woman or a boy and a girl, we employed a color-coded version of the pictures, instead of the greyscale version used in Experiment 1. The agent was depicted in green and the patient was depicted in red. In the noun condition, participants had to indicate if the green character matched with the prime by pressing a key on the keyboard with their left or right hand (depending on the location of the match/no match answer). In the sentence and verb conditions, participants had to indicate whether the picture matched or did not match the previous sentence/verb. They were instructed to respond to the target picture only after the answer screen appeared. As in Experiment 1, in the sentence condition participants needed to combine the two words into a sentence in order to perform the task correctly.

As in Experiment 1, in order to explain the task for the masked condition, participants were told that each block included "easy" and "difficult" trials. There was no speed stress on the discrimination response. Each block included 50% congruent and 50% incongruent trials randomly mixed within blocks. The interval between trials varied between 750 and 1250 ms, drawn from a random distribution. We will refer to these trials as "picture trials". Prime visibility was assessed with a

A Sentences

B

C Single words

D
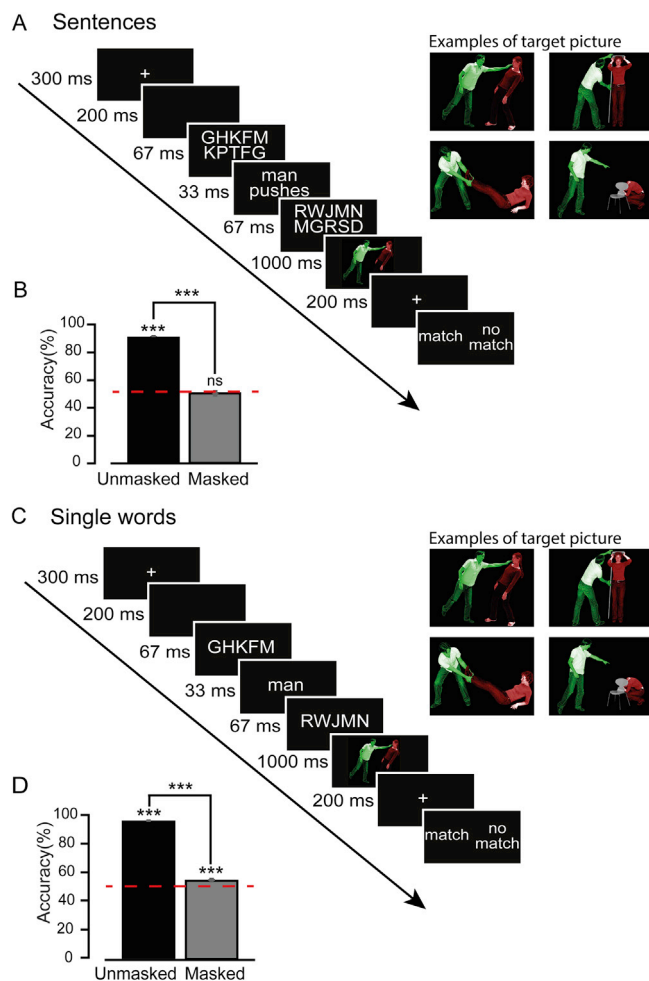
Fig. 3. Paradigm and behavioral results of Experiment 2. (A) Left panel: trial structure in the masked sentence condition. Two masked words forming a sentence were simultaneously presented and followed by a target picture, which could be congruent or incongruent with the previous sentence. In the figure, a congruent trial is represented. Participants performed a match/no match task. Right panel: example of target pictures employed in the sentence condition. **(B)** Participants' accuracy in the sentence condition, for unmasked (black) and masked (grey) trials separately. Results are presented for Day 2 only. Error bars represent between-subject SEM. **(C)** Left panel: trial structure in the masked single word condition. One word was presented and followed by a picture, either congruent or incongruent with the previous word. In the figure a congruent trial is represented. Right panel: example of target pictures employed in the single word condition. Contrary to Experiment 1, target pictures are the same for sentence and single word conditions. **(D)** Participants' accuracy in the single word condition, for unmasked (black) and masked (grey) trials separately. Results are presented for Day 2 only. Error bars represent between-subject SEM.

forced-choice discrimination task (16 trials out of 96), similarly to Experiment 1. We will refer to these trials as "discrimination trials". As in Experiment 1, this forced-choice discrimination task was added as a supplementary visibility check, given that the main task could also be considered a discrimination task. At the end of each block, feedback was provided on participants' accuracy, for both masked and unmasked trials.

3.1.4. Procedure

In Experiment 2, we increased the number of trials per condition compared to Experiment 1. A single EEG session would have been too long, hence we split the EEG recording in two sessions. Therefore, in total the experiment consisted of three separate sessions on separate days. In the first session, participants performed a short version of the task for 6 blocks (2 sentence, 2 noun and 2 verb blocks), with no EEG recording, in

order to familiarize themselves with the task. They were introduced to both task and stimuli before starting the measurement. During the second session, participants performed 12 sentence blocks while EEG was recorded. During the third session, participants performed 6 noun blocks and 6 verb blocks while EEG was recorded. The second and third sessions were both completed within 1–7 days after the first and their order was counterbalanced across subjects. In the EEG recording sessions, there were 1152 trials in the sentence condition and 1152 trials in the single word condition (576 trials in the noun condition and 576 trials in the verb condition). Before starting the EEG recordings, participants were briefly reminded of task and stimuli and performed one practice block (in the noun/verb session, this was randomly selected across block types).

3.1.5. Electroencephalographic measurements

Electroencephalographic measurements and parameters were the same as in Experiment 1 except for data epoching, which was performed from −2 to 2.5 s surrounding each target. All ERPs were time-locked to target, i.e. picture, presentation. The pre-target interval between 367 and 167 ms before target (i.e. the duration of the blank before the first mask) was employed as a baseline.

3.1.6. Statistical analyses

As for Experiment 1, we focused our analysis on the N400 effect and followed the same analysis procedure (see Methods section of Experiment 1 for further details). As in Experiment 1, in addition to the cluster-based approach, we performed all analyses on a pre-defined, central ROI, as a "sanity check". As in Experiment 1, these analyses led to the same conclusions.

3.2. Results

In Experiment 2, we compared EEG responses to simultaneously presented words forming a sentence with responses induced by single words (nouns and verbs). We reasoned that if we replicated findings of Experiment 1, i.e. if we found an N400 effect for masked single words and not for masked sentences, we could conclude that sentence processing cannot be performed under reduced awareness. On the contrary, finding an N400 effect for both masked sentences and single words would mean that the null result in Experiment 1 could be explained by primes sequential presentation in the sentence condition.

3.2.1. Behavior

3.2.1.1. Day 1 (behavioral training).
Overall, we observed that masking strength affected participants' accuracy (main effect of masking strength: $F_{1,39} > 1000$, $p < 0.001$), and that this effect was influenced by semantic complexity (sentences vs. single words) (interaction masking strength x semantic complexity: $F_{1,39} = 12.62$, $p = 0.001$). In follow-up analyses, we found that performance was higher in the unmasked than the masked condition (sentences: $t_{39} = 23.04$, $p < 0.001$; single words: $t_{39} = 46.46$, $p < 0.001$). For both sentences and single words, performance was poor in the masked condition, although it was at chance level for sentences ($t_{39} = 1.42$, $p = 0.160$, 50% correct) and above chance for single words ($t_{39} = 5.64$, $p < 0.001$, 53% correct).

In forced-choice discrimination trials, in which chance level was at 25%, masked words were also poorly perceived (sentences: 28% correct; single words: 32% correct), although subjects performed significantly above chance level for both sentences and single words (sentences: $t_{39} = 2.67$, $p = 0.011$; single words: $t_{39} = 6.97$, $p < 0.001$).

3.2.1.2. Day 2 (EEG recording).
Given that the second and third EEG sessions were counterbalanced across subjects, they were both considered as day 2. First, we compared participants' accuracy between the training session and the main session. In unmasked trials, we found a training effect for sentences ($t_{39} = 3.62$, $p < 0.001$), but not for single

words ($t_{39} = 1.49$, $p = 0.140$). In masked trials, as well in the forced-choice discrimination trials, participants performed similarly on day 1 and day 2 (all p-values>0.070). Next, we observed that overall masking strength affected participants' accuracy (main effect of masking strength: $F_{1,39}>1000$, $p < 0.001$), but was not modulated by semantic complexity (interaction masking strength x semantic complexity: $F_{1,39} = 0.06$, $p = 0.800$). For both sentences and single words, the masked and unmasked conditions differed significantly, accuracy being significantly lower in the masked condition than in the unmasked condition (sentences: $t_{39} = 34.64$, $p < 0.001$; single words: $t_{39} = 47.23$, $p < 0.001$. Fig. 3B and D). In the unmasked condition, subjects scored very high and strongly above chance (sentences: $t_{35} = 36.46$, $p < 0.001$, 90% correct; single words: $t_{35} = 66.58$, $p < 0.001$, 95% correct. Fig. 3B and D). Performance was poor in the masked condition for both sentences and single words, but it was at chance level for sentences ($t_{39} = 0.41$, $p = 0.68$, 50% correct) and above chance for single words ($t_{39} = 7.18$, $p < 0.001$, 54% correct. Fig. 3B and D).

In forced-choice discrimination trials, masked words were again poorly perceived (sentences: 27% correct; single words: 31% correct), although subjects scored significantly above chance level for both sentences and single words (sentences: $t_{39} = 2.24$, $p = 0.031$; single words: $t_{39} = 6.07$, $p < 0.001$).

In summary, behavioral results in Experiment 2 revealed that performance was significantly worse in the masked than in the unmasked condition. Still, one may ask why in masked sentence trials performance was above chance in Experiment 1 and at chance in Experiment 2. This may be explained by the fact that sentence words were presented simultaneously and, as a consequence, participants had less time to perceive and integrate them. Overall, behavioral results in Experiment 2 confirmed our main findings in Experiment 1.

3.2.2. ERPs: N400 modulations by prime visibility and semantic complexity

Following up on Experiment 1, we investigated whether the N400 effect was modulated for masked/unmasked simultaneously presented
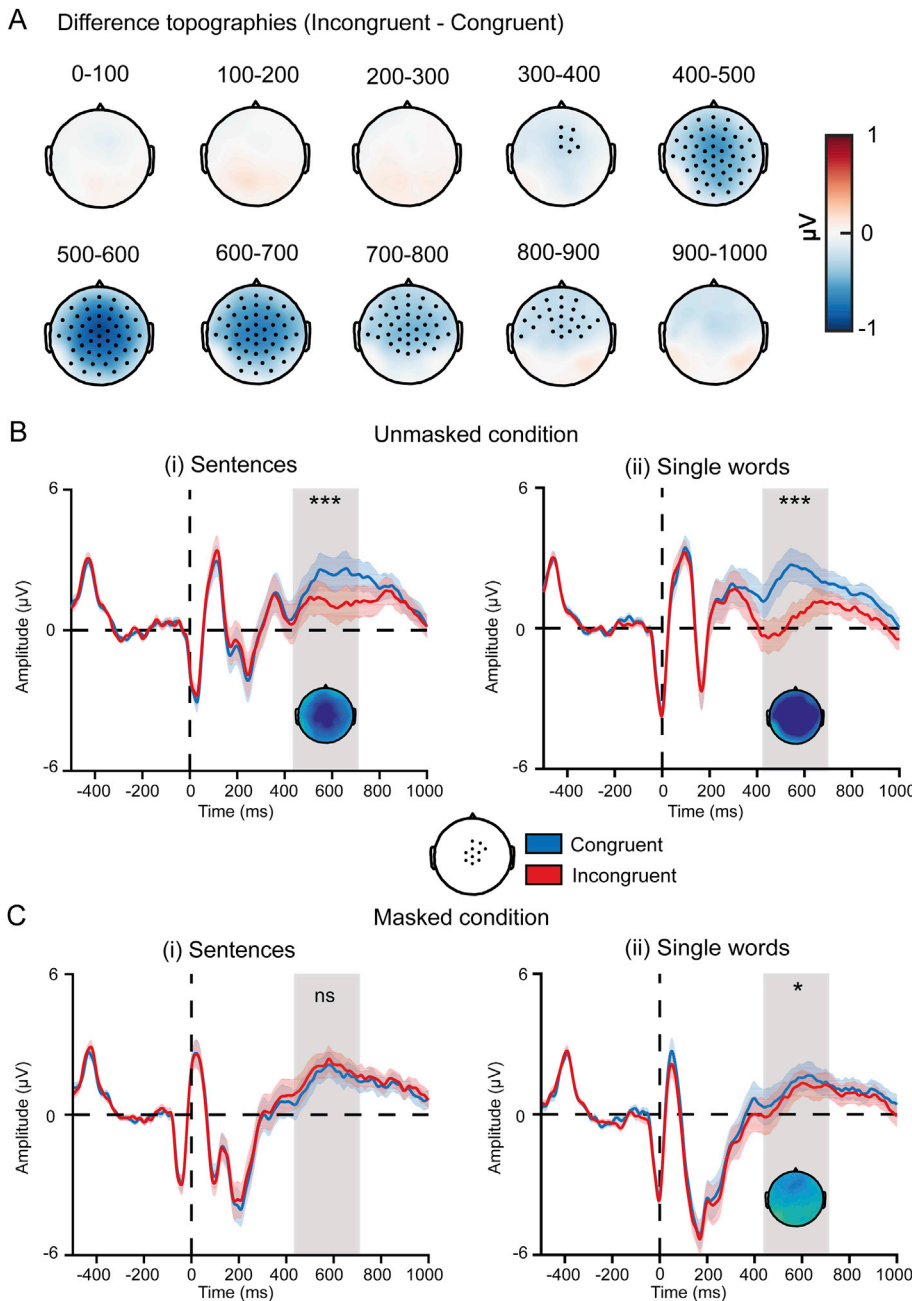
Fig. 4. N400 ERP effect of Experiment 2. (A) Topographic maps of the difference between incongruent and congruent trials over time (0 = target, i.e. picture onset). Cluster-based permutation tests were used to isolate the significant events, while correcting for multiple comparisons across time and (electrode) space. On each head map, channels with a significant effect for at least 50% of its time window are highlighted. All headmaps are scaled from −1 to 1 μV. **(B)** ERPs for a central region of interest (see inset) for congruent (blue) and incongruent (red) unmasked trials, in the sentence (i) and single word (ii) condition. **(C)** ERPs for congruent (blue) and incongruent (red) masked trials, for sentences (i) and single words (ii). In all figures shaded blue and red areas around the wave form indicate between-subject SEM. Time 0 represents the onset of the target picture. Shaded grey areas highlight the time window of interest for the N400 effect (438–708 ms). All headmaps are scaled from −1 to 1 μV.

sentences and single words. To this end, similarly to Experiment 1, we contrasted incongruent and congruent trials using cluster-based permutation testing, correcting for multiple comparisons across both time (0–1 s) and (electrode) space (see Fig. 4A and Methods section of Experiment 1) (Maris and Oostenveld, 2007). When isolating significant effects as clusters in time and space (see Methods), we found a significant difference over centro-frontal electrodes from 438 to 708 ms after target onset and peaking at 544 ms (see Fig. 4A and Fig. 4A-inset; p-value of the cluster between 438 and 708 ms: p < 0.001).

Next, we investigated how the N400 effect was modulated by masking strength and semantic complexity. To this aim, we performed repeated measures ANOVAs on the N400 time window with three factors: masking strength (masked/unmasked), prime/target congruency (congruent/incongruent) and semantic complexity (sentence/single word). We found that the N400 effect (main effect of congruency: $F_{1,39} = 64.76$, p < 0.001) was larger for unmasked than for masked trials (interaction masking strength x congruency: $F_{1,39} = 48.46$, p < 0.001). This N400 effect was modulated by semantic complexity (interaction congruency x semantic complexity: $F_{1,39} = 14.10$, p < 0.001), as the N400 effect was larger for single words than for sentences. Moreover, there was a marginally significant effect of semantic complexity, suggesting that the overall ERP amplitude may have been slightly higher for sentences than single words (main effect of semantic complexity: $F_{1,39} = 3.69$, p = 0.062). No other main effects or interactions were significant (all p-values>0.232). Next, in planned post-hoc analyses, we tested the N400 effect for masked and unmasked trials separately, investigating how it was modulated by semantic complexity.

In the unmasked condition, we observed a clear N400 effect (main effect of congruency: $F_{1,39} = 70.48$, p < 0.001) that was modulated by semantic complexity, being larger for single words than for sentences (interaction of congruency x semantic complexity: $F_{1,39} = 11.45$, p = 0.001). In both semantic complexity conditions, unmasked incongruent trials triggered a larger N400 effect compared to congruent trials (sentences: $t_{39} = 4.81$, p < 0.001; single words: $t_{39} = 9.46$, p < 0.001. Fig. 4B).

Although we did not find a significant N400 effect in the masked condition (main effect of congruency: $F_{1,39} = 0.06$, p = 0.798), the N400 effect was modulated by semantic complexity (interaction of congruency x semantic complexity: $F_{1,39} = 5.88$, p = 0.020). This effect was driven by the fact that we observed a significant N400 effect in the single word condition ($t_{39} = 2.26$, p = 0.028), but not in the sentence condition ($t_{39} = -1.71$, p = 0.093; if anything the effect was observed in the opposite direction, Fig. 4C).

In summary, in Experiment 2 we replicated the main findings of Experiment 1. In the unmasked condition, we found robust neural signatures of semantic processing, as reflected by N400 effects for both sentences and single words. On the contrary, in the masked condition the N400 effect was only present for single words.

4. Discussion

In this study, we investigated the neural mechanisms underpinning sentence and single word processing, focusing on the difference between unification and lexical retrieval processes under varying levels of awareness. According to the Memory, Unification and Control (MUC) model, feedback from the left inferior frontal cortex (LIFC) to the left posterior temporal cortex (LPTC) is crucial for unification, i.e. the process in which words are flexibly combined into a sentence, whereas lexical retrieval (single word processing) relies on a feedforward information flow from sensory regions to the LPTC (Hagoort, 2017, 2013, 2005). Here we tested the general claim that long-range feedback is required for sentence processing, but not for single word processing. To this aim, we reduced stimulus awareness with visual masking, which is thought to disrupt long-range connectivity within the brain (Del Cul et al., 2009, 2007). We investigated whether, when disrupting long-range feedback with masking, unification could still occur. To do so, we designed a novel

experimental paradigm in which we directly compared unification and lexical retrieval mechanisms with EEG.

Behaviorally, performance on the word(s)-picture matching task was higher in the unmasked than in the masked condition, suggesting that masking strongly reduced prime visibility. In EEG, incongruent word(s)-picture pairs triggered an N400 effect in both unmasked sentence and single word trials. This is in line with previous work, showing that manipulating semantic congruency within word-picture pairs elicits N400 effects (Coco et al., 2017; Federmeier and Kutas, 2001; Willems et al., 2008). On the contrary, in the masked condition, an N400 effect was only observed for single words and not for sentences. In Experiment 1, this N400 effect was limited to single words occurring in the third prime position (i.e. just before picture presentation), which supports the claim that masked word processing is fleeting, i.e. strongly dependent on the Stimulus Onset Asynchrony (SOA) between prime and target (Greenwald et al., 1996; Kiefer, 2002; Kiefer and Brendel, 2006; Nakamura et al., 2018). Overall, our findings suggest that feedback processing may be required for unification (sentence processing) but not for lexical retrieval (single word processing), which may instead rely on feedforward spreading of activation.

4.1. The role of awareness in sentence processing

Our results are consistent with previous work, in which no evidence for sentence processing under reduced awareness was found (Rabagliati et al., 2018; Yang et al., 2017). However, our findings contrast with other work, which suggests that sentence processing under reduced awareness may be possible, to some extent and under certain conditions (Armstrong and Dienes, 2014; Axelrod et al., 2015; Batterink and Neville, 2013; Nakamura et al., 2018; Sklar et al., 2012). Nevertheless, these previous findings are either controversial or may be explained by other factors than unification processes.

Axelrod et al. (2015) used continuous flash suppression (CFS) to reduce sentence awareness in fMRI, and found selective responses in the left frontal cortex to masked sentences compared to unpronounceable nonwords. However, this effect may not be specifically related to unification as sentences and chains of nonwords differ substantially on the semantic level, nonwords not involving any word-meaning mapping.

Nakamura et al. (2018) measured the N400 effect triggered by congruent (e.g. *dog-runs*) and incongruent (e.g. *dog-opens*) masked subject/verb pairs within a sentence, while varying the distance between subject and verb across sentences. They found that incongruent pairs triggered an N400 effect only when subject and verb were separated by 0 or 1 word. Although intriguing, these results may reflect simple differences in lexical associations between congruent (*dog-runs*) and incongruent (*dog-opens*) pairs, instead of true unification processes.

A similar criticism can be formulated against Sklar et al. (2012), who used CFS to show that masked sentences containing semantic anomalies (e.g. *"I ironed coffee"*) break through interocular suppression quicker than masked control sentences (e.g. *"I drank coffee"*), which they have interpreted as evidence for unconscious integration of different words into a sentence. As in the previous study, their results could be due to lexical differences between related (*drink-coffee*) and unrelated (*iron-coffee*) items, instead of unification mechanisms. Moreover, recently Sklar et al.'s findings failed to replicate (Rabagliati et al., 2018), and have been criticized based on methodological grounds as well (Shanks, 2017).

Batterink and Neville (2013) used a cross-modal attentional blink paradigm to investigate undetected syntactic violations within sentences. Using EEG, they presented a tone (first target) 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*) (second target). Sometimes participants missed these syntactic violations, because the tone attracted their attention and prevented the violation from reaching awareness (attention blinked). Interestingly, they found that both detected and undetected syntactic violations triggered an early left negative response (ELAN), whereas only detected violations triggered a late positivity (P600).

Nevertheless, it has been shown that the attentional blink impairs conscious report behaviorally, but does not disrupt perceptual integration mechanisms and may leave feedback processing intact (Dehaene et al., 2006; Fahrenfort et al., 2017; Luck et al., 1996). Therefore, this leaves open whether sentence processing can occur in the absence of feedback. Overall, it is not clear whether previous work on sentence processing under reduced awareness properly disentangled unification and lexical retrieval effects. To the best of our knowledge, this is the first work that may control for this confound, as (i) in our study, we used sentences obtained by combining words that are not lexically related, and (ii) we directly compared masked sentences and single words within one experimental paradigm.

The claim that sentence processing is impaired under reduced awareness is also supported by studies performed while subjects were asleep and in patients with disorders of consciousness (DOC). Here, N400 effects are described as weaker, partial and delayed compared to wakefulness (Bastuji et al., 2002; Brualla et al., 1998; Daltrozzo et al., 2012; Davis et al., 2007; Ibáñez et al., 2006; Rohaut et al., 2015; Steppacher et al., 2016; Strauss and Dehaene, 2019). It has been argued that during sleep, there is decreased long-distance connectivity between brain regions (Boly et al., 2012; Massimini et al., 2005; Spoormaker et al., 2012; Tagliazucchi et al., 2013), and that the sleeping brain may still process semantic information, but in a purely feedforward manner (Strauss and Dehaene, 2019). In the vegetative state (no conscious awareness), it has been found that long-distance connectivity between frontal and temporal regions is disrupted (Boly et al., 2011). These findings are compatible with our hypothesis that long-distance feedback processing may play an important role in some sentence-level processes.

4.2. Visual masking as a tool to investigate feedback processing

One important assumption of this work is that visual masking selectively disrupts feedback processing between distant brain areas, while preserving feedforward processing. In many influential theories of consciousness, unconscious information seems to be processed mainly in a feedforward manner (Lamme and Roelfsema, 2000), while conscious processing is thought to be uniquely marked by dynamic recurrent interactions between distant brain regions (Dehaene and Changeux, 2011; Lamme, 2006). A highly-distributed fronto-temporo-parietal activation was found to be a correlate of conscious reportability (Del Cul et al., 2007), and long-distance feedback processing from frontal to temporal cortices was observed to be diminished in patients with DOC (Boly et al., 2011). Further, previous research on visual perception showed that masking interferes with visual awareness by disrupting feedback processing from higher to lower visual areas, whereas feedforward processing is substantially preserved (Fahrenfort et al., 2017, 2007; Kovacs et al., 1995; Lamme et al., 2002). Other studies showed that masking also disrupts long-distance feedback processing from frontal areas to higher visual areas (Del Cul et al., 2009).

Although these studies do not provide direct proof that masking impairs feedback processing between frontal and temporal areas, as predicted by the MUC model (Hagoort, 2017; Hultén et al., 2019), they strongly suggest that masking impairs long-distance feedback processing between many distant brain regions (potentially including frontal and temporal areas), while preserving feedforward processing. Therefore, an impairment of feedback between frontal and temporal areas due to masking seems a plausible consequence, although this should be confirmed by future studies.

Within this theoretical framework, when information is consciously processed in a global, recurrent network, it can also be maintained *ad libitum* in working memory (Baars and Franklin, 2003; Dehaene et al., 2006). Therefore, when feedback processing is impaired, working memory may also be impaired. Interestingly, previous studies have also shown that masking impairs working memory without directly affecting awareness of stimulus material (Blalock, 2013; Ricker and Sandry, 2018). The extent to which working memory and consciousness can be

dissociated is debated (Soto and Silvanto, 2014). Future studies should address the question whether, and if so to what extent, working memory and awareness differentially affect unification processes. At the same time, the MUC model implements working memory, in the service of multi-word integration processes (Hagoort, 2017).

4.3. Residual levels of awareness in visual masking

Our behavioral results revealed that performance on masked single words, despite being poor, was above chance level, and thus participants may have been partially aware of the masked primes (possibly due to the fact that task instructions explicitly stated that in the masked conditions words were present). This is compatible with previous findings, showing that the N400 effect is absent when single words are presented in a fully unconscious fashion (Kang et al., 2011). Therefore, we do not claim that in our paradigm masked primes were unconscious. Still, visibility of masked and unmasked trials differed strongly, as revealed by accuracy in both the main and the discrimination task. Moreover, importantly, our results suggest that masked processing of sentences and single words may differ on a qualitative point of view. Namely, the N400 effect was present for masked single words, but not for masked sentences. If confirmed by future work, these qualitative differences would suggest that reducing stimulus visibility has a stronger effect on unification processes (fully abolishing it) compared to lexical retrieval processes (leaving it partly intact).

4.4. Differences between unification, multiple word processing and single word processing

One may argue that the differences between masked sentence and single word processing in our results are not truly qualitative, but only reflect higher difficulty of sentence processing compared to single word processing. Indeed, in Experiment 2 the N400 effects in the single word condition were stronger compared to the sentence condition (for both masking conditions). In our experimental paradigm, we aimed to minimize task differences between sentences and single words, for example by employing the same two-actor pictures in all conditions in Experiment 2. However, one intrinsic limitation of our study (but potentially of all studies investigating sentence and single word processing) is that sentence and single word processing are hardly comparable, because task difficulty at the sentence level may be harder with respect to the single word level. In fact, the same holds for the factor masking: since the N400 effect is much larger for masked and unmasked stimuli, there is also more room for experimental modulations.

Our findings speak to a broad literature aiming to link ERP effects as the N400 and the P600 to lexical and/or unification processes. Whether the N400 effect reflects lexical or unification mechanisms is still a matter of debate (Baggio and Hagoort, 2011; Kutas and Federmeier, 2010; Lau et al., 2008). Recently, it has been proposed that the N400 effect indexes lexical retrieval, while the P600 (a centro-parietal distributed positivity often associated with syntactic violations - Hagoort et al., 1993), reflects unification (Brouwer et al., 2017). In our results, we did not find any P600 effect in the sentence condition (not even in further uncorrected exploratory analyses, data not reported). Therefore, at least in our experimental paradigm, the N400 effect seems to reflect both unification and lexical retrieval processes.

It has been found that combinatorial operations simpler than sentence processing may occur under reduced levels of awareness (Armstrong and Dienes, 2013; Scott et al., 2018; van Gaal et al., 2014). van Gaal et al. (2014) investigated the negation of valence by presenting masked negation-adjective pairs (e.g. *not-bad*), which formed incongruent combinations with a visible target (e.g. *murder*). They compared this condition with a congruent condition (e.g. *not-good murder*) and found an N400 effect. The paradigm employed by van Gaal et al. is comparable to our paradigm, in terms of both masking technique and timing of stimulus presentation, although the target was a picture in our study and a word in

van Gaal et al.'s study. One may thus wonder why, in contrast to this study, they did find an effect of masked multiple words processing. There may be two reasons for that, one methodological and one theoretical.

First, in the study by van Gaal et al. (2014) participants underwent a negation training in an experimental session before the main session (that took place on a different day), in which subjects were actively trained to either categorize the modifier-adjective combination as positive (e.g. *not-bad, very-good*), or negative (e.g. *not-good, very-bad*). This training likely increased the automaticity of the negation process, and established a lexical association between the negation and the adjective. Therefore, the likelihood of observing N400 effects in the following EEG session was also increased.

A second, theoretical, explanation comes from recent psycholinguistic theories, in which word retrieval and the application of simple syntactic rules are thought to share some neural resources (Jackendoff, 2007). Accordingly, chunks of syntactic structures (like negation-adjective or negation-noun combinations) are assembled in memory retrieval regions (LPFC), exploiting local processing within these areas, and stored there along with words. Later, these pre-assembled chunks are retrieved for unification operations and combined with other chunks in order to form larger structures, i.e. sentences (Hagoort, 2017). In this framework, the negation-adjective combinations investigated by van Gaal et al. (2014) differ from the sentence structures we focused on, which are higher in the syntactic hierarchy and may exploit different neural resources. The same is true for the findings of Armstrong and Dienes (2013), who investigated subliminal processing of negation using subjective threshold measurements.

Given the findings by van Gaal et al. (2014), it is unlikely that results of the masked sentence condition in Experiment 2 are fully explained by difficulties allocating attention to two words presented simultaneously, or by easier visibility of one prime word compared to multiple prime words. Taken together, our results and van Gaal et al.'s results suggest that masking may specifically disrupt sentence processing (unification), but not multiple word processing, although this is of course only indirect proof. The way awareness differentially modulates structures on different levels of the semantic and syntactic hierarchy should be further explored in future work.

5. Conclusions

Our findings suggest that disrupting long-range feedback by means of visual masking may impair unification (sentence processing), but not lexical retrieval (single word processing). Since masking reduces visual awareness, this implies that single words may be processed under reduced awareness, whereas full awareness is required for sentence processing. If confirmed by future studies, our results may show that long-range feedback processing, potentially implying frontal and temporal areas, may be an important prerequisite of language unification, supporting a core prediction of the MUC model (Hagoort, 2017; Hultén et al., 2019; Snijders et al., 2009). This work provides new insights to ongoing debates about the specific roles of different brain mechanisms in a distributed language network (Goucha and Friederici, 2015; Hagoort, 2017; Hickok and Poeppel, 2007; Thompson-Schill et al., 2005).

Acknowledgements

Conflicts of interest

The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116063.

References

Armstrong, A.M., Dienes, Z., 2013. Subliminal understanding of negation: unconscious control by subliminal processing of word pairs. Conscious. Cognit. 22, 1022–1040. https://doi.org/10.1016/j.concog.2013.06.010.

Armstrong, A.-M., Dienes, Z., 2014. Subliminal understanding of active versus passive sentences. Psychol. Conscious. Theor. Res. Pract. 1, 32–50. https://doi.org/10.1037/css0000002.

Axelrod, V., Bar, M., Rees, G., Yovel, G., 2015. Neural correlates of subliminal language processing. Cerebr. Cortex 25, 2160–2169. https://doi.org/10.1093/cercor/bhu022.

Baars, B.J., Franklin, S., 2003. How conscious experience and working memory interact. Trends Cogn. Sci. 7, 166–172. https://doi.org/10.1016/S1364-6613(03)00056-1.

Baggio, G., Hagoort, P., 2011. The balance between memory and unification in semantics: a dynamic account of the N400. Lang. Cogn. Process. 26, 1338–1367. https://doi.org/10.1080/01690965.2010.542671.

Bastuji, H., Perrin, F., Garcia-Larrea, L., 2002. Semantic analysis of auditory input during sleep: studies with event related potentials. Int. J. Psychophysiol. 46, 243–255. https://doi.org/10.1016/S0167-8760(02)00116-2.

Batterink, L., Neville, H.J., 2013. The human brain processes syntax in the absence of conscious awareness. J. Neurosci. 33, 8528–8533. https://doi.org/10.1523/jneurosci.0618-13.2013.

Blalock, L.D., 2013. Mask similarity impacts short-term consolidation in visual working memory. Psychon. Bull. Rev. 20, 1290–1295. https://doi.org/10.3758/s13423-013-0461-9.

Boly, M., Garrido, M.I., Gosseries, O., Bruno, M., Boveroux, P., Schnakers, C., Massimini, M., Litvak, V., Laureys, S., Friston, K., 2011. Preserved feedforward but impaired top-down processes in the vegetative state. Science 332, 858–863.

Boly, M., Perlbarg, V., Marrelec, G., Schabus, M., Laureys, S., Doyon, J., Pelegrini-Issac, M., Maquet, P., Benali, H., 2012. Hierarchical clustering of brain activity during human nonrapid eye movement sleep. Proc. Natl. Acad. Sci. 109, 5856–5861. https://doi.org/10.1073/pnas.1111133109.

Brouwer, H., Crocker, M.W., Venhuizen, N.J., Hoeks, J.C.J., 2017. A neurocomputational model of the N400 and the P600 in language processing. Cogn. Sci. 41, 1318–1352. https://doi.org/10.1111/cogs.12461.

Brown, C., Hagoort, P., 1993. The processing nature of the N400: evidence from masked priming. J. Cogn. Neurosci. 5, 34–44. https://doi.org/10.1162/jocn.1993.5.1.34.

Brualla, J., Romero, M.F., Serrano, M., Valdizan, J.R., 1998. Auditory event-related potentias to semantic priming during sleep. Electroencephalogr. Clin. Neurophysiol. 108, 283–290.

Coco, M.I., Araujo, S., Petersson, K.M., 2017. Disentangling stimulus plausibility and contextual congruency: electro-physiological evidence for differential cognitive dynamics. Neuropsychologia 96, 150–163. https://doi.org/10.1016/j.neuropsychologia.2016.12.008.

Daltrozzo, J., Claude, L., Tillmann, B., Bastuji, H., Perrin, F., 2012. Working memory is partially preserved during sleep. PLoS One 7, 1–12. https://doi.org/10.1371/journal.pone.0050997.

Davis, M.H., Coleman, M.R., Absalom, A.R., Rodd, J.M., Johnsrude, I.S., Matta, B.F., Owen, A.M., Menon, D.K., 2007. Dissociating speech perception and comprehension at reduced levels of awareness. Proc. Natl. Acad. Sci. 104, 16032–16037. https://doi.org/10.1073/pnas.0701309104.

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. https://doi.org/10.1016/S0926-6410(99)00050-6.

Dehaene, S., Changeux, J.-P., 2011. Experimental and theoretical approaches to conscious processing. Neuron 70, 200–227. https://doi.org/10.1016/j.neuron.2011.03.018.

Dehaene, S., Changeux, J.P., Naccache, L., Sackur, J., Sergent, C., 2006. Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends Cogn. Sci. 10, 204–211. https://doi.org/10.1016/j.tics.2006.03.007.

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., Pallier, C., 2015. The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. Neuron 88, 2–19. https://doi.org/10.1016/j.neuron.2015.09.019.

Del Cul, A., Baillet, S., Dehaene, S., 2007. Brain dynamics underlying the nonlinear threshold for access to consciousness. PLoS Biol. 5, 2408–2423. https://doi.org/10.1371/journal.pbio.0050260.

Del Cul, A., Dehaene, S., Reyes, P., Bravo, E., Slachevsky, A., 2009. Causal role of prefrontal cortex in the threshold for access to consciousness. Brain 132, 2531–2540. https://doi.org/10.1093/brain/awp111.

Fahrenfort, J.J., Scholte, H.S., Lamme, V.A.F., 2007. Masking disrupts reentrant processing in human visual cortex. J. Cogn. Neurosci. 19, 1488–1497. https://doi.org/10.1162/jocn.2007.19.9.1488.

Fahrenfort, J.J., van Leeuwen, J., Olivers, C.N.L., Hogendoorn, H., 2017. Perceptual integration without conscious access. Proc. Natl. Acad. Sci. 114, 3744–3749. https://doi.org/10.1073/pnas.1617268114.

Federmeier, K.D., Kutas, M., 2001. Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. J. Exp. Psychol. Learn. Mem. Cogn. 27, 202–224.

Goucha, T., Friederici, A.D., 2015. The language skeleton after dissecting meaning: a functional segregation within Broca's Area. Neuroimage 114, 294–302. https://doi.org/10.1016/j.neuroimage.2015.04.011.

Greenwald, A.G., Draine, S.C., Abrams, R.L., 1996. Three cognitive markers of unconscious semantic activation. Science 273, 1699–1702.

Hagoort, P., 2017. The core and beyond in the language-ready brain. Neurosci. Biobehav. Rev. 37, 347–362. https://doi.org/10.1016/j.neubiorev.2017.01.048.

Hagoort, P., 2013. MUC (memory, unification, control) and beyond. Front. Psychol. 4, 1–13.

Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423. https://doi.org/10.1016/j.tics.2005.07.004.

Hagoort, P., Brown, C., Groothusen, J., 1993. The syntactic positive shift (sps) as an erp measure of syntactic processing. Lang. Cogn. Process. 8, 439–483. https://doi.org/10.1080/01690969308407585.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402. https://doi.org/10.1038/nrn2113.

Hultén, A., Schoffelen, J.-M., Uddén, J., Lam, N.H.L., Hagoort, P., 2019. How the brain makes sense beyond the processing of single words – an MEG study. Neuroimage 186, 586–594. https://doi.org/10.1016/j.neuroimage.2018.11.035.

Ibáñez, A., López, V., Cornejo, C., 2006. ERPs and contextual semantic discrimination : degrees of congruence in wakefulness and sleep. Brain Lang. 98, 264–275. https://doi.org/10.1016/j.bandl.2006.05.005.

Jackendoff, R., 2007. A Parallel Architecture perspective on language processing. Brain Res. 1146, 2–22. https://doi.org/10.1016/j.brainres.2006.08.111.

Kang, M.-S., Blake, R., Woodman, G.F., 2011. Semantic analysis does not occur in the absence of awareness induced by interocular suppression. J. Neurosci. 31, 13535–13545. https://doi.org/10.1523/jneurosci.1691-11.2011.

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

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. https://doi.org/10.1162/jocn.2006.18.2.184.

Kiefer, M., Spitzer, M., 2000. Time course of conscious and unconscious semantic brain activation. Neuroreport 11, 2401–2407.

Kouider, S., Dehaene, S., 2007. Levels of processing during non-conscious perception: a critical review of visual masking. In: Philosophical Transactions of the Royal Society B: Biological Sciences, pp. 857–875. https://doi.org/10.1098/rstb.2007.2093.

Kovacs, G., Vogels, R., Orban, G.A., Sprague, J.M., 1995. Cortical correlate of pattern backward masking (inferior temporal cortex/temporal integration/shape recognition). Neurobiology 92, 5587–5591.

Kutas, M., Federmeier, K.D., 2010. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.

Kutas, M., Hillyard, S.A., 1980. Event related potentials in response to semantically inappropriate and surprisingly large words. Biol. Psychol. 11, 99–116. https://doi.org/10.1016/0301-0511(80)90046-0.

Lamme, V.A.F., 2006. Towards a true neural stance on consciousness. Trends Cogn. Sci. 10, 494–501. https://doi.org/10.1016/j.tics.2006.09.001.

Lamme, V.A.F., Roelfsema, P.R., 2000. The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. 23, 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X.

Lamme, V.A.F., Zipser, K., Spekreijse, H., 2002. Masking interrupts figure-ground signals in V1. J. Cogn. Neurosci. 14, 1044–1053. https://doi.org/10.1162/089892902320474042.

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

Luck, S.J., Vogel, E.K., Shapiro, K.L., 1996. Word meanings can be accessed but not reported during the attentional blink. Nature 383, 616–618. https://doi.org/10.1038/383616a0.

Marcel, A., 1980. Conscious and preconscious recognition of polysemous words: locating the selective effects of prior verbal context. In: Nickerson, R.S. (Ed.), Attention and Performance, VIII. Erlbaum, Hillsdale, NJ.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S.K., Singh, H., Tononi, G., 2005. Breakdown of cortical effective connectivity during sleep – supporting material. Science 309, 2228–2233. https://doi.org/10.1126/science.1117256.

Meijs, E.L., Slagter, H.A., de Lange, F.P., van Gaal, S., 2018. Dynamic interactions between top-down expectations and conscious awareness. J. Neurosci. 38, 2318–2327. https://doi.org/10.1523/JNEUROSCI.1952-17.2017.

Menenti, L., Gierhan, S.M.E., Segaert, K., Hagoort, P., 2011. Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. Psychol. Sci. 22, 1173–1182. https://doi.org/10.1177/0956797611418347.

Nakamura, K., Makuuchi, M., Oga, T., Mizuochi-Endo, T., Iwabuchi, T., Nakajima, Y., Dehaene, S., 2018. Neural capacity limits during unconscious semantic processing. Eur. J. Neurosci. 47, 929–937. https://doi.org/10.1111/ejn.13890.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869. https://doi.org/10.1155/2011/156869.

Rabagliati, H., Robertson, A., Carmel, D., 2018. The importance of awareness for understanding language. J. Exp. Psychol. Gen. 147, 190–208.

Ricker, T.J., Sandry, J., 2018. The relationship between masking and short-term consolidation during recall from visual working memory. Ann. N. Y. Acad. Sci. 1424, 91–101. https://doi.org/10.1111/nyas.13641.

Rohaut, B., Faugeras, F., Chausson, N., King, J.R., Karoui, I. El, Cohen, L., Naccache, L., 2015. Probing ERP correlates of verbal semantic processing in patients with impaired consciousness. Neuropsychologia 66, 279–292. https://doi.org/10.1016/j.neuropsychologia.2014.10.014.

Schoffelen, J.-M., Hultén, A., Lam, N., Marquand, A.F., Uddén, J., Hagoort, P., 2017. Frequency-specific directed interactions in the human brain network for language. Proc. Natl. Acad. Sci. 114, 8083–8088. https://doi.org/10.1073/pnas.1703155114.

Scott, R.B., Samaha, J., Chrisley, R., Dienes, Z., 2018. Prevailing theories of consciousness are challenged by novel cross-modal associations acquired between subliminal stimuli. Cognition 175, 169–185. https://doi.org/10.1016/j.cognition.2018.02.008.

Shanks, D.R., 2017. Regressive research: the pitfalls of post hoc data selection in the study of unconscious mental processes. Psychon. Bull. Rev. 24, 752–775. https://doi.org/10.3758/s13423-016-1170-y.

Sklar, A.Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., Hassin, R.R., 2012. Reading and doing arithmetic nonconsciously. Proc. Natl. Acad. Sci. 109, 19614–19619. https://doi.org/10.1073/pnas.1211645109.

Snijders, T.M., Vosse, T., Kempen, G., Van Berkum, J.J.A., Petersson, K.M., Hagoort, P., 2009. Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. Cerebr. Cortex 19, 1493–1503. https://doi.org/10.1093/cercor/bhn187.

Soto, D., Silvanto, J., 2014. Reappraising the relationship between working memory and conscious awareness. Trends Cogn. Sci. 18, 520–525. https://doi.org/10.1016/j.tics.2014.06.005.

Spoormaker, V.I., Gleiser, P.M., Czisch, M., 2012. Frontoparietal connectivity and hierarchical structure of the brain's functional network during sleep. Front. Neurol. 3, 1–10. https://doi.org/10.3389/fneur.2012.00080.

Steppacher, I., Kaps, M., Kissler, J., 2016. Against the odds: a case study of recovery from coma after devastating prognosis. Ann. Clin. Transl. Neurol. 3, 61–65. https://doi.org/10.1002/acn3.269.

Strauss, M., Dehaene, S., 2019. Detection of arithmetic violations during sleep. Sleep 42, 1–12. https://doi.org/10.1093/sleep/zsy232.

Tagliazucchi, E., von Wegner, F., Morzelewski, A., Brodbeck, V., Jahnke, K., Laufs, H., 2013. Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep. Proc. Natl. Acad. Sci. 110, 15419–15424. https://doi.org/10.1073/pnas.1312848110.

Thompson-Schill, S.L., Bedny, M., Goldberg, R.F., 2005. The frontal lobes and the regulation of mental activity. Curr. Opin. Neurobiol. 15, 219–224. https://doi.org/10.1016/j.conb.2005.03.006.

Tyler, L.K., Marslen-Wilson, W.D., Randall, B., Wright, P., Devereux, B.J., Zhuang, J., Papoutsi, M., Stamatakis, E.A., 2011. Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. Brain 134, 415–431. https://doi.org/10.1093/brain/awq369.

van Gaal, S., Naccache, L., Meuwese, J.D.I., van Loon, A.M., Leighton, A.H., Cohen, L., Dehaene, S., 2014. Can the meaning of multiple words be integrated unconsciously? Philos. Trans. R. Soc. Biol. Sci. 369, 20130212. https://doi.org/10.1098/rstb.2013.0212.

Willems, R.M., Özyürek, A., Hagoort, P., 2008. Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. J. Cogn. Neurosci. 20, 1235–1249. https://doi.org/10.1162/jocn.2008.20085.

Yang, Y.H., Tien, Y.H., Yang, P.L., Yeh, S.L., 2017. Role of consciousness in temporal integration of semantic information. Cognit. Affect Behav. Neurosci. 17, 954–972. https://doi.org/10.3758/s13415-017-0525-9.