

How the brain makes sense beyond the processing of single words – An MEG study

Annika Hultén^{a,b,*}, Jan-Mathijs Schoffelen^{a,b}, Julia Uddén^{a,b}, Nietzsche H.L. Lam^{a,b}, Peter Hagoort^{a,b,**}

^a Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

^b Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Radboud University, Nijmegen, Netherlands



ARTICLE INFO

Keywords:
MEG
Word processing
Sentence progression
Unification
Fronto-temporal interaction

ABSTRACT

Human language processing involves combinatorial operations that make human communication stand out in the animal kingdom. These operations rely on a dynamic interplay between the inferior frontal and the posterior temporal cortices. Using source reconstructed magnetoencephalography, we tracked language processing in the brain, in order to investigate how individual words are interpreted when part of sentence context. The large sample size in this study ($n = 68$) allowed us to assess how event-related activity is associated across distinct cortical areas, by means of inter-areal co-modulation within an individual. We showed that, within 500 ms of seeing a word, the word's lexical information has been retrieved and unified with the sentence context. This does not happen in a strictly feed-forward manner, but by means of co-modulation between the left posterior temporal cortex (LPTC) and left inferior frontal cortex (LIFC), for each individual word. The co-modulation of LIFC and LPTC occurs around 400 ms after the onset of each word, across the progression of a sentence. Moreover, these core language areas are supported early on by the attentional network. The results provide a detailed description of the temporal orchestration related to single word processing in the context of ongoing language.

1. Introduction

Understanding language is more than the concatenation of individual word meanings that are retrieved from memory. It requires an interplay between the context and the incoming information provided by the word that is currently processed, to produce an interpretation of the whole utterance. These combinatorial operations, here referred to as unification, are a hallmark of human language processing. Within the framework of the Memory, Unification and Control (MUC) model for sentence processing (Hagoort, 2003, 2005, 2013), unification refers to the operations that extract sentence-level meaning from the combined lexical input. In the present study, we focused on unification processes from the single word perspective. To this aim, we exploited a time sensitive measurement technique for brain activity (magnetoencephalography, MEG) to investigate how the incremental context provided by the unfolding sentence influences the processing of the incoming words.

Reading a single word gives rise to a forward driven sweep of

activation, which generally proceeds from early pre-lexical processing in the visual and occipito-temporal cortex to meaning related activity in the temporal cortex (Salmelin, 2007). The later (>200 ms) event-related responses are, however, also mediated by backward connections (Garrido et al., 2007). In electroencephalography (EEG), the amplitude of the event-related N400 potential is known to be highly sensitive to expectations based on contextual semantic cues (Baggio and Hagoort, 2011; Kutas and Federmeier, 2011; Lau et al., 2008; Van Petten and Kutas, 1990). Using MEG, studies typically report that the activation characteristic for the N400 component (in MEG designated as the N400m) originates from areas in the temporal cortex, in particular from the middle and/or posterior superior temporal gyrus, but occasionally frontal cortex activation has also been reported (Dale et al., 2000; Halgren et al., 2002; Helenius et al., 1998; Van Petten and Luka, 2006). In order to highlight brain activity that is related to top-down influence from context, we manipulated the context in which the words were presented, by having the participants read well-formed sentences and unstructured

* Corresponding author. Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

** Corresponding author. Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

E-mail addresses: annika.hulten@aalto.fi (A. Hultén), peter.hagoort@donders.ru.nl (P. Hagoort).

¹ Current affiliation: The Department of Neuroscience and Biomedical Engineering, Aalto University, Helsinki, Finland.

word lists. Importantly, in order to capture how the context affects individual word processing in general, the analysis was not limited to one word (such as the last word in the sentence or the list), but the effect of the context was assessed over all the presented words.

To date, the sentences and word list paradigm with written words has mainly been used in hemodynamic studies (Fedorenko et al., 2011; Friederici et al., 2000; Humphries et al., 2005, 2006; Jobard et al., 2007; Snijders et al., 2009; Stowe et al., 1998; Vandenberghe et al., 2002; Xu et al., 2005). Typical findings in these studies highlight larger signals during sentences than during words list in the left inferior frontal gyrus, the (posterior) middle temporal gyrus and the anterior temporal gyri. These regions have therefore to a greater or lesser extent all been attributed a role in processing the syntactical sentence structure.

Electrophysiological studies are much scarcer. To our knowledge only one previous study has used the sentence and word list paradigm to study event-related responses in MEG, and this study had a small number of only 9 participants (Brennan and Pykkänen, 2012). The previous MEG findings report increased activation to sentences as compared to word lists in the posterior temporal cortex (~200–400 ms), in the bilateral anterior temporal regions (~300–350 ms), the left ventromedial pre-frontal cortex (250–350 ms) and the left pars opercularis of the inferior frontal gyrus (~200–500 ms). Given the limited sample size of the published data, investigations of the temporal dynamics require replication with a larger sample size. The main analyses in the current study included data from sentence reading in a subset of 68 participants, based on a strict criterion of performance in the catch trials. However, note that the same pattern of results was obtained when including the whole set of 102 participants. This large sample size also gave us the opportunity to investigate the co-modulation of different areas across subjects in event-related measures.

Sentence processing involves a dynamic modulation across the input string of words. The progression of a sentence, where new words become available in an incremental way, means that both the unification load and the information build-up changes over the course of the sentence. For each incoming word, the lexical representation is activated and, as the sentence advances, the amount of lexical information builds up. However, predicting the upcoming word becomes usually easier towards the end of a sentence, with the consequence that some processing demands decrease over the course of the sentence. At the same time, as the information content increases towards the end of the sentence, the maintenance load of the context gets higher. Previous EEG studies report a decrease in the N400 amplitude as a function of the ordinal word position (Van Petten and Kutas, 1990, 1991), but do not describe how sentence processing is affected before or after the peak activation. In addition, using source localization we will be able to decompose how different regions contribute to the effect and explore the dynamical modulation between the nodes of the language network.

Previously reported effects in left posterior temporal regions overlap in time with the left frontal activation (Brennan and Pykkänen, 2012). In order to more directly investigate how these and other active areas might co-modulate with each other, we tested for correlation between sentence context effect and sentence progression effect in all active regions across individuals. If the amplitude of the response in each region is modulated by the incremental sentence context in a similar manner within an individual, the two areas cannot be functionality independent. Instead, this would suggest that the two areas are working in unison, both contributing to the processing needed for successful unification. The correlation of amplitude differences across a pair of areas provides a measure of interaction (either direct or indirect), and can thus inform us of information exchange between areas.

Correlated amplitude differences would complement a previously published functional connectivity analysis of the same data set, where we showed that posterior temporal regions and frontal regions display bidirectional Granger Causal interactions in direction-specific frequency bands (Schiffelen et al., 2017). Although Granger Causal interactions are indicative for an established causal interaction between brain areas, they

do not provide insight into whether the local cortical computations are in effect affected by these interactions. Local cortical activation is readily quantified as event-related activity, in our study time-locked to the onset of a new word. This event-related activity reflects the local response, of how an incoming word is processed given the preceding context, which will be informative in establishing the functionality of the nodes in the language network.

The MUC poses that the network dynamics of the left inferior frontal gyrus (LIFG) and the temporal cortex are expressed in both reverberating rhythmic activity and time-locked events related to incoming information (Baggio and Hagoort, 2011). The temporal cortex is thought to capture the reverberating activity within the temporal and frontal cortices. The temporal cortex will accumulate both bottom-up driven input from inferior temporal regions, local processing of lexical information, and top-down feedback from the LIFG related to unification and on-line maintenance of the context (Baggio and Hagoort, 2011). The LIFG, in turn, receives input from the temporal cortices, but its activity may at this stage be less time-locked to the onset of individual words (Baggio and Hagoort, 2011). The time-locked measures of MEG may therefore be sensitive to capturing how context-driven expectations (semantic or syntactic) affect the processing of individual words, especially in the temporal cortex.

In addition to the dynamics between LIFG and the temporal cortex suggested by the MUC model, the anterior temporal lobe (ATL) has also been suggested to be a key region for the combinatorial operations involved in sentence comprehension (e.g. Brennan and Pykkänen, 2012; Brennan and Pykkänen, 2017; Humphries et al., 2006; Rogalsky and Hickok, 2009; Vandenberghe et al., 2002). The exact role of the ATL remains unclear, as studies suggest that it may be linked to both syntactic structure building (Humphries et al., 2006), semantic composition (Stowe et al., 2005; Vandenberghe et al., 2002) or attention to syntactic features (Rogalsky and Hickok, 2009). Previous MEG studies using the event-related measures have linked activity in the ATL to basic linguistic combinatorial processing needed in sentence comprehension (Brennan and Pykkänen, 2012; Brennan and Pykkänen, 2017).

Here, we investigated the temporal activation dynamics of individual words in a sentence context. First, we contrasted the brain response to individual words in visually presented sentences to the same words in scrambled sentences (i.e., word lists), which we call the *sentence context* effect. Second, we quantified the effect of sentence progression on the processing of single words. We call this the *sentence progression* effect. By evaluating if the time-locked response to each incoming word leads to co-modulation of the main active areas, we can assess whether these areas are working in unison or independently. The overarching goal of these analyses is to shed further light on the functional roles of the cortical regions involved in sentence processing.

2. Materials and methods

We collected MEG-data from 102 participants, all of which were native speakers of Dutch with ages ranging from 18 to 33 years (mean age 22 years). Half of the participants were males and all were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological, developmental or language deficits. The final set of participants was 68 individuals, based on the behavioural task compliance.

The present data is a subset of a larger study – MOUS (Mother of all Unification Study), with a total of 204 (102 in the visual domain) participants and has previously been explored in the frequency domain (Lam et al., 2016, 2018; Schiffelen et al., 2017). All participants were native speakers of Dutch with ages ranging from 18 to 33 years (mean age 22 years). Half of the participants were males and all were right-handed, had normal or corrected-to-normal vision, and reported no history of neurological, developmental or language deficits. 15 participants were excluded and replaced by new volunteers, due to either technical problems in either the fMRI or MEG part of the study (4), magnetic residue on the participant (2), failure to meet inclusion criteria (4), low

behavioural accuracy (2, see below for details) or because the participant chose to discontinue the study (3).

The study was approved by the local ethics committee and followed the guidelines of the Helsinki declaration. Fig. 1 presents an overview of the experimental task and examples of the stimulus set.

2.1. Stimuli and experimental design

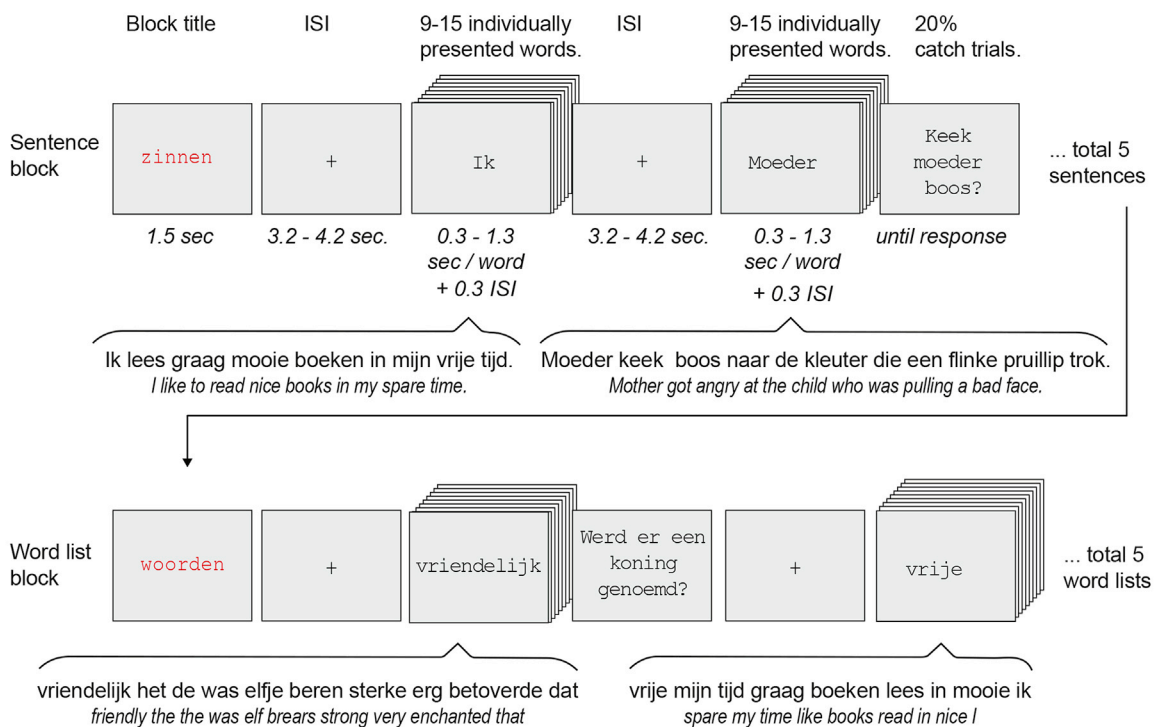
The stimuli consisted of 240 sentences and their word list counterparts. They varied between 9 and 15 words in length. The sentences consisted of a mixture of different sentence structures and contained a relative clause in half of the sentences. The individual words were on average 5 letters long (SD = 2.6, range 1–14; short and long words were exceptional with only one word of one letter and one word of 14 letters in a total of 4142 words). The word lists were created by shuffling the order

of the words in each sentence, so that fewer than three words could be interpreted as a coherent phrase. Using the same words in both the sentence and the word list condition allowed us to control for the lexical processing level. Fig. 1B presents examples of the stimulus set.

Each participant saw half of the stimuli (120) in the sentence form and half (120) in the word list form. Across participants, each stimulus word was presented the same number of times in the sentence and in the word list condition. The words in the sentences and word lists were presented sequentially, one word a time in alternating blocks of five consecutive sentences or word lists. At the beginning of each block, subjects were reminded of the upcoming block type by a word presented in red, describing the upcoming block (*zinnen* (sentences) or *woorden* (words)). Subjects were instructed to read the sentences and word lists carefully, and to not try to re-order the items in the word lists in any way.

In 20% of the trials a yes/no question, relating to the previous

A. Experimental design



B. Examples of the stimulus set

Sentences	Word lists
Moeder keek boos naar de kleuter die een flinke pruillip trok. <i>Mother got angry at the child who was pulling a bad face.</i>	flinke naar een keek trok die de boos moeder pruillip kleuter <i>really at a got pulling who the angry mother face child</i>
Het elfje dat de sterke beren betoverde was erg vriendelijk. <i>The elf that had enchanted the strong bears was very friendly.</i>	vriendelijk het de was elfje beren sterke erg betoverde dat <i>friendly the the was elf brears strong very enchanted that</i>
Ik lees graag mooie boeken in mijn vrije tijd. <i>I like to read nice books in my spare time.</i>	vrije mijn tijd graag boeken lees in mooie ik <i>spare my time like books read in nice I</i>
Questions	
Keek moeder boos? <i>Did mother get angry?</i>	Werd er een koning genoemd? <i>Was a king mentioned?</i>

Fig. 1. Experimental design. (A) Task design and timing of the stimulus presentation. Sentences/Word lists were presented one word at a time, separated by a blank screen of 300 ms (ISI) in interleaving blocks of five Sentences or Word lists. Each block started with a condition printed in red to alert the participant to the type of the upcoming stimuli. Ten percent of the trials were followed by a catch trial in which a question about the previous sentence or word list was presented. (B) Examples of the stimuli (translations are not literal).

sentence/word list was presented to check for compliance. For example, *Did grandma give a cookie to the girl? or Was the word 'grandma' mentioned?* Subjects answered the question by pressing a button for 'Yes'/'No' with their left index and left middle finger, respectively. Only participants that responded correctly to at least 80% of these questions were included in the analyses.

The visual presentation rate of the stimuli was determined in relation to the duration of the audio recording of spoken versions of the sentences and the word lists (audiotur), taking into account both the number of letters (sumnletters) and words (nwords) in the whole sentence and the number of letters within each word (nletters). The duration of a single word (in ms) was determined as: $(nletters/sumnletters) * (audiotur + 2000 - 150 * nwords)$. No word was presented for a shorter time than 300 ms. The mean duration for each word was 434 ms (range 300–1344 ms). Between the sentences/word a fixation cross was presented with a jittered duration between 3200 and 4200 ms. The stimuli were presented in a black mono-spaced font on a grey background at the centre of a screen in front of the participant, within a visual angle of 4° using the Presentation software (Version 16.0, Neurobehavioral Systems, Inc).

2.2. Data acquisition and analysis

The MEG data were collected with a 275 axial gradiometer system (CTF), with a sampling frequency of 1200 Hz. The head position of each participant was measured continuously during the measurement via the position of three coils attached to the participant's head (Stolk et al., 2013). If the movement exceeded 5 mm the task was paused (but never in the middle of a sentence or word list) and the participant was guided into the original head positions before continuing. Eye blinks and movements were monitored by two horizontal and two vertical EOG electrodes.

Anatomical MR images needed for the source localization of the MEG data were collected in a Siemens Trio 3T scanner using a high-resolution T1-weighted magnetization-prepared rapid gradient-echo pulse sequence. The anatomical image was manually co-registered to the MEG sensors using a digitized head shape obtained with a Polhemus device.

The data analysis was performed using the tools in the FieldTrip package (Oostenveld et al., 2011) unless otherwise specified. All data segments contaminated by physiological artifacts (eye blinks and movements, muscle contractions) or SQUID jumps were identified by a semi-automatic artifact identification procedure (Schoffelen et al., 2005) and removed from further analysis. The data was then band-pass filtered between 0.5 and 40 Hz, and downsampled to a sampling frequency of 300 Hz. Event-related epochs were extracted from 200 ms before word onset until the onset of the next word and averaged across all words in a sentence or word list. The 200-msec interval before the onset of each word was used as baseline. This choice of baseline ensures that any effect between sentences and word lists is not due to slow signal fluctuations across the sentence or word list, but reflects the effect of a meaningful sentence context on the neural response to the incoming words. In order to analyse the effect of sentence progression we also computed averages that were computed across words in the same ordinal position across sentences or word lists. In all statistical analyses, the time window was constrained to 0–600 ms after stimulus onset.

We proceeded to source level descriptions of the data using Minimum Norm Estimation (Hämäläinen and Ilmoniemi, 1994), as the sensor level MEG data represent a complex spatial summation of underlying neural activation. First, the anatomical MR image was used to create a volume conduction model based on a single shell description (Nolte, 2003) of the inner surface of the skull and to create a cortically constrained source model. We used *Freesurfer* (<http://surfer.nmr.mgh.harvard.edu/>) for the extraction of the cortical sheet. Next, the *Caret* software (Van Essen et al., 2001) was used to register the individual surfaces to the *fsaverage* template surface, and subsequently downsampled to result in source model with 8196 dipole locations. This procedure resulted in source models where the individual dipole locations could be directly averaged or

compared across subjects, because they were expressed in surface-registered space. For the downsampling of the high-resolution cortical sheets to 8196 dipole locations we used the MNE-suite software package (Gramfort et al., 2014). No orientation constraints were applied to the source dipoles. Minimum norm estimates of source activity were obtained using *FieldTrip*, and we estimated the current density, using depth-weighting and regularization with an estimate of the covariance of the noise (Dale et al., 2000). This noise covariance was estimated from the single trial epochs, using only the 200 ms period preceding the onset of the first word in a sentence/word list. The same co-variance estimate was used for all conditions. The resulting maps of noise-normalized current density estimates, that is, so called dynamic statistical parametric maps (dspm; Dale et al., 2000) were computed for each condition of interest.

2.3. Statistical analyses

In the present study, sentence processing was examined both in relation to *sentence context*, i.e. by contrasting the average response to words in a sentence context to the average response to words in a word list, and in relation to *sentence progression*. The effect of sentence progression (build-up and integration of information) was quantified at the source level by modelling the change in the neural responses as a function of ordinal word position. To this end, we fitted, at each time point, a simple linear function of the ordinal word position specific neural response in a sentence, with ordinal word position as the predictor. The same was repeated for the word list condition. The resulting two sets of regression coefficients, which capture the general tendency of signal change from one word to the other across the sentence/word list, were used in further statistical analyses both on the whole head level and in a ROI analysis (see below). The response to the first word was excluded from the analysis (sentence/word list onset transients may affect the signal) as well as words from the 11th position onward (to exclude biased estimates due to low number of words at these ordinal positions).

Whole head statistical analyses of both sentence context and the sentence progression effects were performed using non-parametric permutation tests (with 5000 permutations) together with clustering, to address the multiple comparisons problem (Maris and Oostenveld, 2007). We used the paired samples T-statistic as the test-statistic. As this study contained a higher than usual number of participants we chose a stricter than normal a priori cluster threshold of $T = \pm 3$ (corresponding with a one-sided $p < 0.01$) as an a priori cut-off value to grow spatio-temporal clusters.

In addition to a whole head analysis, we looked at inter-areal dynamics of the active areas through a ROI analysis. The ROIs were defined in a data-driven manner, based on the location of the grand average (across conditions and participants) peak activity in four distinct time windows: 50–150, 150–250, 250–350, and 350–450 ms.

Vertices belonging to an ROI determined earlier in time were excluded when searching for maximum amplitude in the subsequent time windows, which prevented us from discovering the same area twice. In order to establish the size of the ROIs we first determined full width at half maximum (FWHM) of the peak amplitude of the grand average. The region was allowed to grow in all directions around the peak coordinates, until the half-maximum was reached. In the temporal and frontal cortex, this resulted in two patches of cortex including 127 and 169 vertices respectively. In the occipital cortex, the FWHM was disproportionately larger (271 vertices) than in the core language regions (likely due to the higher signal-to-noise ratio in sensory areas). In order to better meet the assumption of variance homogeneity within each ROI, we chose to fix the size all the ROIs to 127 vertices, by repeating the region growing around peak activity until 127 vertices were reached. The data-driven search resulted in four ROIs. The region growing around the maximum resulted in two spatially overlapping ROIs in the left occipital cortices. These two ROIs shared a majority of the vertices (102 vertices) and the activation was largely overlapping in time (occipital peak latency 148 ms; occipito-

temporal peak latency 151 ms). The two occipital ROIs can therefore not be considered as distinct or independent signal sources, and we chose to exclude the second one. The three remaining ROIs were located in the left occipital (LO), the left posterior temporal cortex (LPTC) and the left inferior frontal cortex (LIFC). The full time course of activation for each ROI was extracted by averaging the activation across the vertices in the ROI.

For the ROI-based activity estimates, we statistically evaluated the differences across conditions in two ways. First, using a time unconstrained technique, we employed a non-parametric permutation test, with clustering across time to address the multiple comparison problem across multiple time points, followed by a Bonferroni correction for the number of ROIs tested. Second, we identified temporal peaks of mean activity (independent of condition) within each of the ROIs, and performed paired t-tests on the amplitudes (averaged across a time window surrounding each peak based on the peak width at 70% of the maximum), correcting for multiple tests (8: 2 contrasts times 4 peaks), using Bonferroni correction. The LPTC ROI displayed two distinct peaks. The peak latencies and peak width in each of the ROI were the following; LO: 148 ms (width: 111–178 ms), LPTC1: 178 ms (width: 108–301 m), LPTC2: 401 ms (width: 315–461 ms) and LIFC: 398 ms (width: 308–481 ms).

In order to evaluate whether the peak activity in the ROIs were in any way related to each other we looked at whether the ROI peak activity was modulated the same way within an individual. In the sentence context analysis, we tested for correlation between each of the four peaks (LO, LPTC1, LPTC2 and LIFC) across participants, using the amplitude difference of the words in sentences with respect to the same words in the word lists (sentence – word list). The sentence progression effect was subjected to a similar correlation analysis of the intra-individual effects between sentence and word list by testing for correlation between the difference in regression coefficients between sentences and word lists. The results of both correlational analyses were corrected for multiple comparisons with the total number of tests (12), using Bonferroni correction.

Given the degree of variation in the behavioural performance, we also tested for correlation between the behavioural performance and the amplitude differences. In this analysis, we used the relative amplitude in the sentence context and the sentence progression effect and the behavioural performance, and (as before) applied Bonferroni correction for multiple comparisons.

All correlation analyses were done on the amplitude difference of sentences versus word lists, in order to exclude variation related purely to individual variation in power. As the LPTC2 and LIFC peaks were overlapping in time, a control analysis was performed to ensure that the observed effect was distinctive for the interplay between these two regions, and not due general stimulus-induced modulations across the cortex. In this control analysis, we tested for correlation between the activity in the LPTC2 and the LIFC with the activity in the LO in the same time window. That is, the LPTC2-LO correlation analysis was computed for the activation between 315 and 461 ms and for the LIFC-LO between 308 and 481 ms.

3. Results

In total 68 participants fulfilled the behavioural requirement of 80% correct, set to guarantee that the participants were actively reading the stimuli. The mean percentage correct for the 68 included participants was 84.8% (SD 6.7%), ranging from 80.4 to 93.5%. The accuracy was slightly better for sentence questions [mean: 87.0%, SD = 6.3%] than for word list questions [mean: 82.7%, SD = 6.9%; $t(67) = 3.1$, $p < 0.01$]. As this could imply that the task difficulty might be different for sentences and word lists, we investigated this in some more detail. The sentences and word lists related questions were each split into two sets based on difficulty. Half of the sentences contained relative clauses which are deemed to be more difficult to process than other types of sentence

structures. For the word lists, long word lists (12 or more words) were considered more difficult than short word lists (less than 12 words). In a two-by-two ANOVA with difficulty (easy/difficult) and the condition (sentence/word list) as factors, we found a significant main effect of condition [$F(1,66) = 9.1$, $p < 0.01$] but not for difficulty [$F(1,66) = 3.2$, n. s.] or, importantly, an interaction between difficulty and the condition [$F(1,66) = 1.1$, n. s.]. Thus, there the main comparison between sentences and word-lists cannot be accounted for by difficulty alone.

In the brain, both sentences and word lists elicited a similar forward sweep of activation time-locked to the onset of an incoming word (see Fig. 2A and B). Activity traversed from the bilateral occipital cortex around 100 ms, via left and right occipito-temporal regions at 130–230 ms, to the left posterior and middle temporal cortex (300–500 ms) as well as to the left inferior frontal gyrus (400–500 ms). In addition, activity was observed in the middle part of bilateral motor/frontal cortex at 200–440 ms (the right hemisphere activity subsided at 400 ms).

The comparison between words in sentences and in word lists indicated that the brain responses to these conditions were significantly different ($p < 0.01$). We observed stronger activation for words in a sentence compared to a word list context, starting from 150 ms onwards; an overview of the sentence context effects is presented in Fig. 2C, and the main findings are described below.

The first difference between sentences and word lists was observed in left occipito-temporal cortex, peaking around 150 ms after word onset. Between approximately 170–200 ms differences were further observed in the left inferior parietal lobule, followed by effects in the left superior frontal cortex between 190 and 400 ms. The inferior frontal cortex and the left anterior superior temporal cortex showed significant differences between approximately 300–450 ms. Clear effects were also seen in the left middle temporal cortex, first between 150 and 230 ms and then again between 314 and 400 ms now also including parts of the inferior post-central cortices. In addition, a sustained effect between 250 and 400 ms was observed in the right inferior frontal cortex and superior anterior temporal cortex though the latter ended a little earlier.

The analysis of sentence progression revealed one distinct cluster, in which the word-position-dependent brain response differed significantly between sentences and word lists. The cluster extended from the middle frontal to the precentral cortex between 208 and 300 ms (see Fig. 3). Visualization of the waveforms in this area to the individual words as each ordinal position showed that the effect stemmed from the ascending slope of the response (Fig. 3B) and that the slope of the amplitude increase was steeper to sentences than to word lists.

The ROI analysis supported the outcome of the whole head analysis (see Fig. 4). In the sentence context analysis, the unconstrained cluster based permutation test on the time course of activity within each ROI showed that sentences and word lists differed significantly in all three ROIs. The significant differences were supported by temporal clusters exceeding the uncorrected significance level with ROI-specific latencies. These time ranges were between 0.06 – 115 ms and 141–195 ms for the LO ROI, between 155 – 275 ms and 300–421 ms for the LPTC ROI, and between 235 and 475 ms for the LIFC ROI. The same effect was also evident when contrasting the a priori defined peak amplitudes, based on the window containing 70% of the maximum activity of each peak [LO 111–178 ms: $t(67) = 3.0$, $p < 0.05$; LPTC1 108–301 ms: $t(67) = 2.9$, $p < 0.05$; LPTC2 315–461 ms: $t(67) = 3.4$, $p < 0.01$; LIFC 308–481 ms: $t(67) = 4.9$, $p < 0.001$]. There were no significant effects of sentence progression in the ROI analysis.

In order to test for the co-modulation between the ROIs, we computed correlations of relative amplitude modulation of sentences and word lists in each of the ROIs (i.e. the sentence context effect). The correlations between LPTC1 and LPTC2 [$r = 0.3$, $p = 0.02$] as well as between LIFC and LPTC2 [$r = 0.3$, $p = 0.01$ uncorrected] were not significant after Bonferroni correction.

The intra-individual effects between the ROI's were further tested by correlating the relative sentence progression coefficients. That is, relative

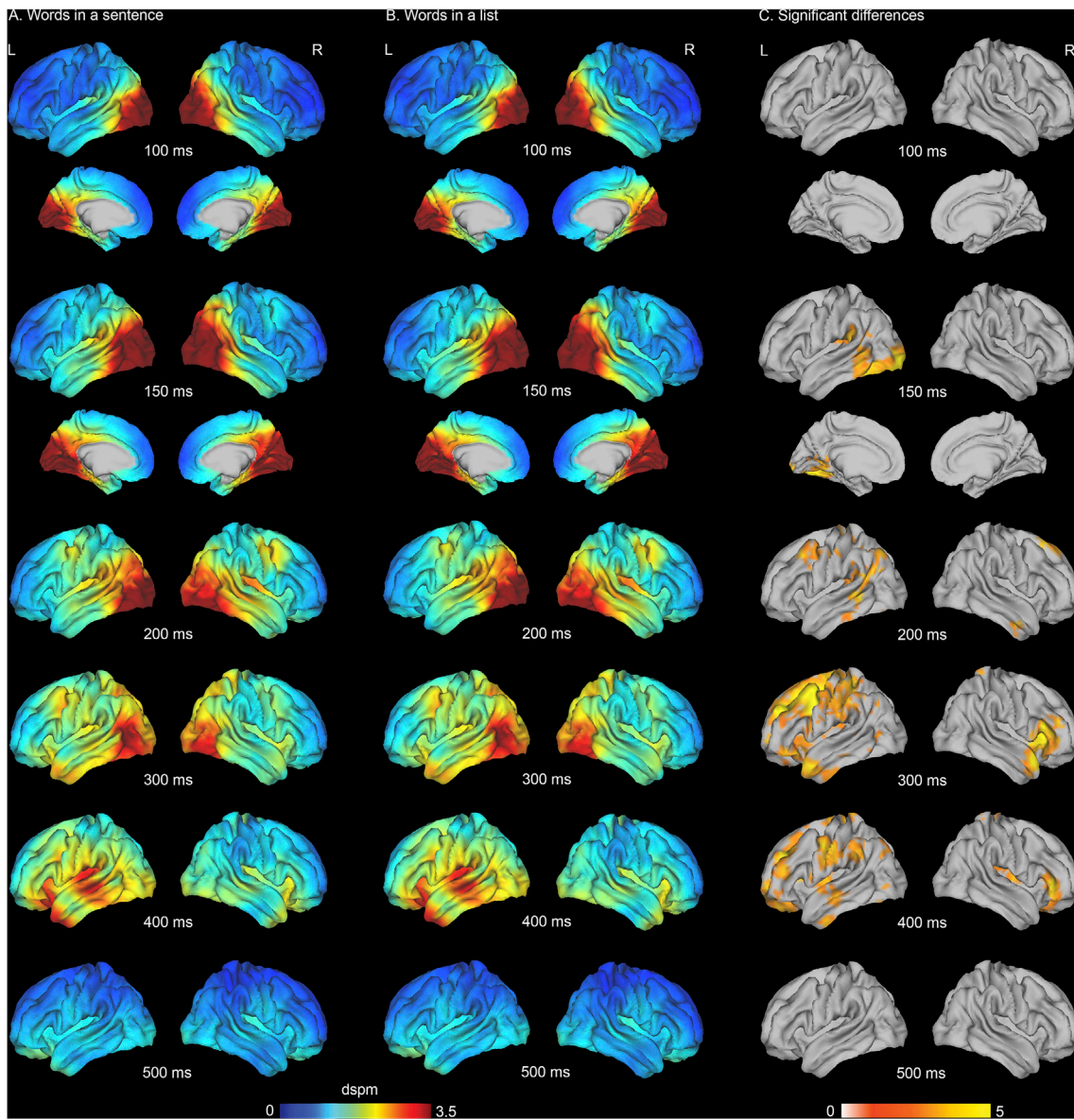
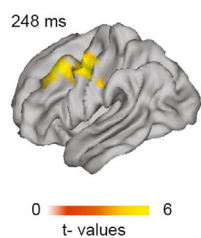


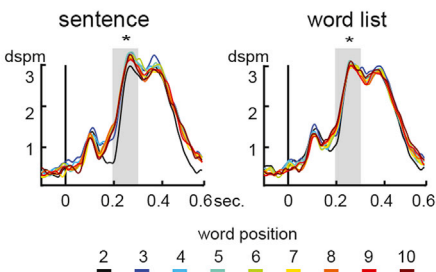
Fig. 2. Whole head sentence context results. (A) Grand average across participants of the forward sweep of activation seen for words in a sentence and (B) word list. (C) T-statistics of the difference in the spatiotemporal cluster test, showing all regions that significantly distinguish the activity level between words in a sentence context and words in word list. The data are presented in still frames at regular intervals, though the analysis was performed on the continuous data points. Medial views of the brain is presented for the 100 ms and 150 ms time frames, below the lateral view. The medial view was non-informative in the later time frames. The cortical activity is represented as noise normalized current density estimates (dspm).

Sentence progression effect

A. Location



B. Time course of activation



C. Mean amplitude at each word position

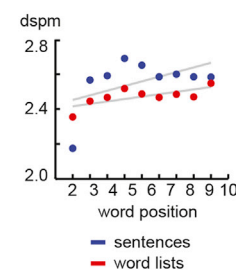
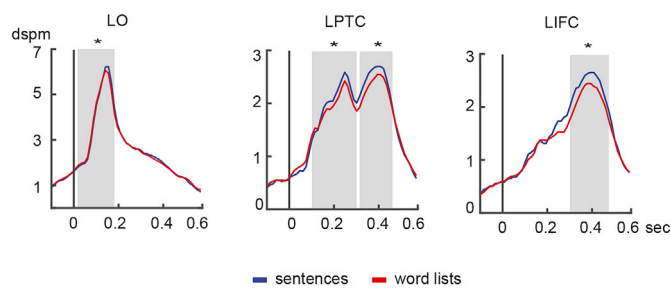


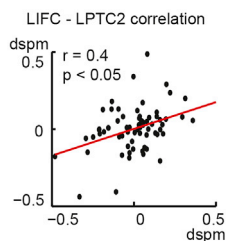
Fig. 3. Whole head sentence progression results. The cortical activity reflects the average amount of signal change from one word to the other, expressed in noise normalized current density estimates (dspm). (A) Location and (B) time course of the cluster in which the progression effect was significantly different for the sentences and word lists. The significant time window (208–300 ms) is marked in grey. (C) Visualization of the sentence progression effect: the mean of amplitude across the significant time window is plotted separately for each word.

ROI results

A. Sentence context



B. Sentence progression



C. ROI locations

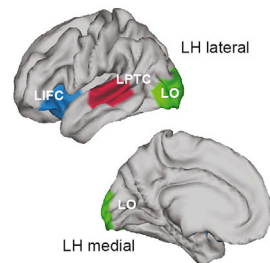


Fig. 4. ROI results. (A) Results from the Sentence context analysis. The cortical activity is represented as noise normalized current density estimates (dspm) and the grey bars mark the time windows defined as 70% of the peak amplitude; stars indicate significant differences between sentences and word lists in these windows. Words in sentences elicited higher amplitudes than words in word lists in all the ROIs, albeit the effect was minimal in the LO. (B) Sentence progression results. A positive correlation was found in the relative sentence progression (sentences vs. word lists) between the peak activity in temporal cortex (LPTC2) and frontal cortex (LIFC). The larger an individual's progression effect between 315 and 461 ms was in LPTC2, the larger the effect between 308 and 481 ms in the LIFC. (C) Location of the data-driven ROIs on the cortical surface. BLUE: LIFC (Left Inferior Frontal Cortex), RED: LPTC (Left Middle Temporal Cortex), Green: LO (Left Occipital cortex).

(sentence vs. word lists) amount of signal change from one word to the next. A significant co-modulation was found for the LIFC and LPTC2 [$r = 0.4$, $p < 0.05$] (for a scatter plot see Fig. 4B; a report of all correlation analyses accompanied by scatter plots can be found in Supplementary Fig. 1). The effect was turned out to be distinctive for the dynamics of LIFC and LPTC2, as no significant effects were discovered in the control analysis with the LO at the same time window (see Supplementary Fig. 1).

4. Discussion

In the present study, we examined how an incremental sentence context influences individual word processing. We evaluated the general effect of a sentence context by contrasting all words in a sentence to the same words presented in a scrambled order as word lists. In addition, we evaluated the effect of sentence progression by looking at millisecond temporal modulations of the amplitude of the event-related response to individual words as a function of the ordinal word position. These analyses were complemented by a correlation analysis, investigating whether effect sizes in LIFC and LPTC were co-modulated within an individual.

One of the key strengths of MEG is its excellent temporal resolution in combination with good spatial resolution. We observed a sweep of activation starting at about 100 ms in occipital cortex, which spread forward to temporal, parietal and frontal cortex. This general pattern was similar to that described by Brennan and Pykkänen (2012). Activation was not exclusively left-lateralized, but clearly more dominant in the left than the

right hemisphere. The new finding is that the initial activation in frontal and temporal cortices seemed driven by a feedforward sweep of activation, whereas the second stage was characterized by correlated activity between the LPTC2 and LIFC with respect to the sentence progression dynamics. In less than 500 ms the word processing cycle was completed, which suggests that lexical information has been successfully unified with the higher-order representation of the context. In the following sections we discuss results in more detail.

In the whole head analysis, one of the earliest effects between sentences and word lists was observed around 150 ms after word onset, in the left occipito-temporal cortex. In the ROI-analysis we observed a similar effect in the left occipital cortex ROI (LO), between 140 and 181 ms. These regions have been linked to pre-lexical processing of letter-strings and low-level feature processing respectively (Salmelin, 2007; Tarkiainen et al., 1999). Here, these processes seem to be affected by increased predictability of the upcoming word in sentences.

In addition to expected activations in language-relevant areas, the whole head analysis also revealed a prominent, temporally sustained effect in the left superior frontal cortex between 190 and 400 ms. This area was also sensitive to sentence progression between 208 and 300 ms. Curiously, in this area the sentence progression effect was in the form of a systematic increase of activity as the sentence progressed. This is contrary to the other effects in the current study, where the sentences compared to the word lists were accompanied by decreased or delayed activation. In terms of location, the effect seemed to encompass both motor and premotor areas as well as regions in the dorsolateral prefrontal cortex. While the sentences were in no way biased to actions, the motor cortex activation does overlap with the areas suggested to process action semantics (e.g., Pulvermüller, 2013). Interestingly, we found a similar effect in a previous study where we focused on the oscillatory response, using the present data (Lam et al., 2016), namely that the theta power increased for the words late in the sentence. This effect was strongest in the bilateral frontal and right parietal regions, and was interpreted as relating to the increase in the memory load, which increases as information builds up. A post-hoc interpretation of the present data and the oscillatory results is to link them to involvement of the attentional control network (Corbetta and Shulman, 2002). In order to allocate resources efficiently, humans tend to attend to some of the linguistic input more strongly. In sentences, the new and most relevant information is often marked syntactically or, in the case of spoken language, with prosody. These types of information structure markers have been shown to activate the attentional network in service of extended processing of the linguistic input (Burholt Kristensen et al., 2013). Although information structure was not explicitly manipulated in this study, it is possible that the build-up of a higher order representation spanning the entire sentence has increased attentional demands compared to a list of single words. It is worth noting, that the observed amplitude modulations in the sentence progression analysis might not be strictly linear (see Fig. 3C) though linear regression was used for quantification. The linear model used in the sentence progression analysis is admittedly a simplification of the true complexity of sentence structure, yet building more sophisticated models that include for instance word-specific syntactic information would require fitting the model to single-trial responses, which is challenging in MEG, given the poor signal-to-noise ratio at the single trial level.

All activity in the N400 window in the temporal and inferior frontal cortices is of particular interest in the present study. In the whole head analysis, we found significantly stronger activity for sentences than word lists in the temporal cortex response between 150–230 ms and 314–400 ms, and in the inferior frontal cortices between 300 and 450 ms. The temporal profiles in LIFC and LPTC provide an interesting new insight into the temporal dynamics of single word processing in a sentence context. In the temporal region, we found two clear peaks in the MEG signal. The fact that the LPTC was activated first and only then followed by activity in the LIFC and that this activity was uncorrelated implies a bottom-up driven sweep of stimulus related activity for each

incoming word. That said, nothing in the present results points towards a magic moment in time at which the processing of word information is completed in the LPTC and then fed forward to LIFC (Balota and Yap, 2006). Had this been the case, we should not have observed a difference in first temporal peak in the sentence context between words in a sentence and in word lists.

The whole head analysis showed that the ATL in terms of activity and the functionality, followed that of the LIFC. That is, in the ATL sentences evoked significantly higher amplitudes between 300 and 450 ms. The ATL was not discovered by the data driven ROI search, which is likely due to the fact that the region was never more active than the LIFC or the LPTC nor did it have a different time course of activation than these regions. This does not of course exclude the possibility that the region might independently be involved in combinatorial operations (Brennan and Pykkänen, 2012; Brennan and Pykkänen, 2017), alongside the LIFC and posterior temporal cortex. However, one should also bear in mind that estimated source activity obtained from MEG signals are not equally reliable across the cortex; simulation studies (Hillebrand and Barnes, 2002; Stenroos et al., 2014) suggest that source localizations or ERFs are harder and less reliable for ATL than for most other cortical areas. Moreover, it cannot be excluded that the activity that we see in ATL is a spillover from LIFC activity, which is less affected by limitations in source localization accuracy. Therefore, at this stage we refrain from making claims about the functional role of the ATL in the context of our experimental manipulations.

In the present study, sentence progression was estimated via the regression coefficients equivalent to the average amount of signal change across the sentence as compared to the word list condition. Over the course of the sentence (the sentence progression analysis) the first temporal peak (LPTC1) did not correlate with the LIFC whereas the second peak (LPTC2) did. This implies that the activity during the initial sweep reflects largely independent processes, but that the temporal and frontal cortices start interacting later on in time. The LPTC2 and LIFC can only be co-modulated if activity is reverberating between the two regions. The co-modulation seemed to be specific to the LPTC2 and LIFC, as their amplitudes did not correlate with the activity in the visual cortex. We can therefore rule out a general co-modulation of power across the cortex. The LPTC and LIFC have well-documented anatomical (e.g. Catani et al., 2005; Parker et al., 2005) and functional connections (Xiang et al., 2009; Schoffelen et al., 2017) which implies that the two regions are in direct contact, but it should be acknowledged that other regions may be part of the same network. Nevertheless, the correlated amplitude modulation between the LPTC2 and LIFC in the sentence progression contrast is indicative of increased interaction between frontal and temporal areas. Presumably, the two regions are functionally coupled during N400 time window (~300–450 ms) in order to unify a word's semantic and syntactic specifications with the higher order semantic and syntactic structures provided by the context.

Here, we analysed a subset of the data that has been used in previous reports (Lam et al., 2016; Schoffelen et al., 2017). In the present report, we used a more stringent criterion for task compliance, which is why we present results from a somewhat smaller sample compared to Lam et al. (2016); Schoffelen et al. (2017). However, the main findings in the present study are unchanged also when the full sample of 102 participants was included for the analysis. The event-related results presented here complement our previous reports, which evaluated brain activity in the oscillatory domain. In previous work, we observed stronger desynchronization of alpha activity for word lists compared to sentences in widespread temporal and frontal areas (Lam et al., 2016). These areas overlap with the areas that also display a difference in the event-related response. Also, using Granger Causality analysis, we observed bidirectional interactions between frontal and temporal areas, where the temporal to frontal interaction was supported by alpha oscillatory activity whereas the reverse frontal to temporal interaction was supported by beta oscillatory activity (Schoffelen et al., 2017). In both studies of the oscillatory activity the event-related response was removed, whereas the

averaging over trials in the present study will cancel out the phase and thereby any frequency specific effects. These studies therefore suggest that the LPTC is functionally connected to the LIFC in two distinct ways. According to the “Communication through Coherence” theory (Fries, 2005, 2015), oscillatory activity can regulate the responsiveness of a neuronal group, facilitating information transfer between synchronised brain areas. In the language network, this might be implemented in form of frequency specific directionality between LPTC and LIFC opening up a communication window for word specific time-locked information to smoothly pass between the two areas. However, further research is needed to fully understand the different aspects of brain activity that is captured by oscillatory and event-related measures.

5. Conclusion

This study highlights the temporal dynamics of the processing of words in context. In less than 500 ms lexical information is retrieved and unified with the context. This does not happen in a strictly feed-forward manner, but by means of dynamic interplay of the activity in the left temporal and frontal cortices, for each individual word. We found no evidence for a specific moment in time (Balota and Yap, 2006) where word retrieval ends, and then integration of the word with the contexts begins. Instead, we show that retrieval and unification go hand in hand, and that communication between the LIFG and the left posterior temporal cortex occurs at the later part of the word processing window. The processing in these areas seems to be supported early on by areas in the attentional network, which expands the view on what traditionally has been considered language-processing areas. This is the largest MEG study on sentence processing to date, and the dynamic view on language processing that it provides is an important extension of the insights gained by previous fMRI studies on the functional organization of the language network in the brain.

Acknowledgements

We are grateful to research assistants Laura Arendsen, Manuela Schuetze, Tineke de Haan, and Charlotte Poullisse for assisting with participant recruitment, stimuli construction and the manual parts of data preprocessing. The researchers were supported by the International Max Planck Research School for Language Science (for N.H.L.L.), the Spinoza Prize Award, as well as the Language in Interaction Gravitation grant from the Netherlands Organisation for Scientific Research (for P.H.), and the Finnish Academy (grant #287474) to A.H. The data is part of the MOUS (Mother of all Unification Studies) project, which will be made available to the wider research community in the near future.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2018.11.035>.

References

- Baggio, G., Hagoort, P., 2011. The balance between memory and unification in semantics: a dynamic account of the N400. *Lang. Cognit. Process.* 26, 1338–1367.
- Balota, D.A., Yap, M.J., 2006. Attentional control and the flexible lexical processor: explorations of the magic moment of word recognition. In: Andrews, S. (Ed.), *From Inkmarks to Ideas: Current Issues in Lexical Processing*. Psychology Press, Hove, UK, pp. 229–258.
- Brennan, J., Pykkänen, L., 2012. The time-course and spatial distribution of brain activity associated with sentence processing. *Neuroimage* 60, 1139–1148.
- Brennan, J., Pykkänen, L., 2017. MEG evidence for incremental sentence composition in the anterior temporal lobe. *Cognit. Sci.* 41 (Suppl. 6), 1515–1531.
- Burholt Kristensen, L., Wang, L., Petersson, K.M., Hagoort, P., 2013. The interface between language and attention: prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebr. Cortex* 23, 1836–1848.
- Catani, M., Jones, D.K., ffytche, D.H., 2005. Perisylvian language networks of the human brain. *Ann. Neurol.* 57, 8–16.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

- Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26, 55–67.
- Fedorenko, E., Behr, M.K., Kanwisher, N., 2011. Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl. Acad. Sci. U. S. A.*, 201112937.
- Friederici, A.D., Meyer, M., von Cramon, D.Y., 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* 74, 289–300.
- Fries, P., 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cognit. Sci.* 9, 474–480.
- Fries, P., 2015. Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235.
- Garrido, M.I., Kilner, J.M., Kiebel, S.J., Friston, K.J., 2007. Evoked brain responses are generated by feedback loops. *Proc. Natl. Acad. Sci. U. S. A.* 104, 20961–20966.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M.S., 2014. MNE software for processing MEG and EEG data. *Neuroimage* 86, 446–460.
- Hagoort, P., 2003. How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage* 20 (Suppl. 1), S18–S29.
- Hagoort, P., 2005. On Broca, brain, and binding: a new framework. *Trends Cognit. Sci.* 9, 416–423.
- Hagoort, P., 2013. MUC (memory, unification, control) and beyond. *Front. Psychol.* 4, 416.
- Halgren, E., Dhond, R.P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J.D., Dale, A.M., 2002. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage* 17, 1101–1116.
- Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.
- Helenius, P., Salmelin, R., Service, E., Connolly, J.F., 1998. Distinct time courses of word and context comprehension in the left temporal cortex. *Brain* 121, 1133–1142.
- Hillebrand, A., Barnes, G.R., 2002. A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *Neuroimage* 16, 638–650.
- Humphries, C., Love, T., Swinney, D., Hickok, G., 2005. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum. Brain Mapp.* 26, 128–138.
- Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J. Cognit. Neurosci.* 18, 665–679.
- Jobard, G., Vigneau, M., Mazoyer, B., Tzourio-Mazoyer, N., 2007. Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage* 34, 784–800.
- Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647.
- Lam, N.H.L., Schoffelen, J.M., Udden, J., Hultén, A., Hagoort, P., 2016. Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *Neuroimage* 142, 43–54.
- Lam, N.H., Hultén, A., Hagoort, P., Schoffelen, J.M., 2018. Robust neuronal oscillatory entrainment to speech displays individual variation in lateralisation. *Lang Cogn Neurosci* 1–12.
- Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de)constructing the N400. *Nat. Rev. Neurosci.* 9, 920–933.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190.
- Nolte, G., 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* 48, 3637–3652.
- 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, 1.
- Parker, G.J., Luzzi, S., Alexander, D.C., Wheeler-Kingshott, C.A., Ciccarelli, O., Lambon Ralph, M.A., 2005. Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage* 24, 656–666.
- Pulvermüller, F., 2013. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. *Trends Cognit. Sci.* 17, 458–470.
- Rogalsky, C., Hickok, G., 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebr. Cortex* 19, 786–796.
- Salmelin, R., 2007. Clinical neurophysiology of language: the MEG approach. *Clin. Neurophysiol.* 118, 237–254.
- Schoffelen, J.M., Hultén, A., Lam, N., Marquand, A.F., Udden, J., Hagoort, P., 2017. Frequency-specific directed interactions in the human brain network for language. *Proc. Natl. Acad. Sci. U. S. A.* 114, 8083–8088.
- Schoffelen, J.M., Oostenveld, R., Fries, P., 2005. Neuronal coherence as a mechanism of effective corticospinal interaction. *Science* 308, 111–113.
- Snijders, T.M., Vosse, T., Kempen, G., Van Berkum, J.J., 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.
- Stenroos, M., Hunold, A., Hauelsen, J., 2014. Comparison of three-shell and simplified volume conductor models in magnetoencephalography. *Neuroimage* 94, 337–348.
- Stolk, A., Todorovic, A., Schoffelen, J.M., Oostenveld, R., 2013. Online and offline tools for head movement compensation in MEG. *Neuroimage* 68, 39–48.
- Stowe, L.A., Broere, C.A., Paans, A.M., Wijers, A.A., Mulder, G., Vaalburg, W., Zwarts, F., 1998. Localizing components of a complex task: sentence processing and working memory. *Neuroreport* 9, 2995–2999.
- Stowe, L.A., Haverkort, M., Zwarts, F., 2005. Rethinking the neurological basis of language. *Lingua* 115, 997–1042.
- Tarkiainen, A., Helenius, P., Hansen, P.C., Cornelissen, P.L., Salmelin, R., 1999. Dynamics of letter string perception in the human occipitotemporal cortex. *Brain* 122, 2119–2132.
- Van Essen, D.C., Drury, H.A., Dickson, J., Harwell, J., Hanlon, D., Anderson, C.H., 2001. An integrated software suite for surface-based analyses of cerebral cortex. *J. Am. Med. Inf. Assoc.* 8, 443–459.
- Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. *Mem. Cognit.* 18, 380–393.
- Van Petten, C., Kutas, M., 1991. Influences of semantic and syntactic context on open- and closed-class words. *Mem. Cognit.* 19, 95–112.
- Van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain Lang.* 97, 279–293.
- Vandenberghe, R., Nobre, A.C., Price, C.J., 2002. The response of left temporal cortex to sentences. *J. Cognit. Neurosci.* 14, 550–560.
- Xiang, H.D., Fonteijn, H.M., Norris, D.G., Hagoort, P., 2009. Topographical functional connectivity pattern in the perisylvian language networks. *Cerebr. Cortex* 20, 549–560.
- Xu, J., Kemeny, S., Park, G., Frattali, C., Braun, A., 2005. Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage* 25, 1002–1015.