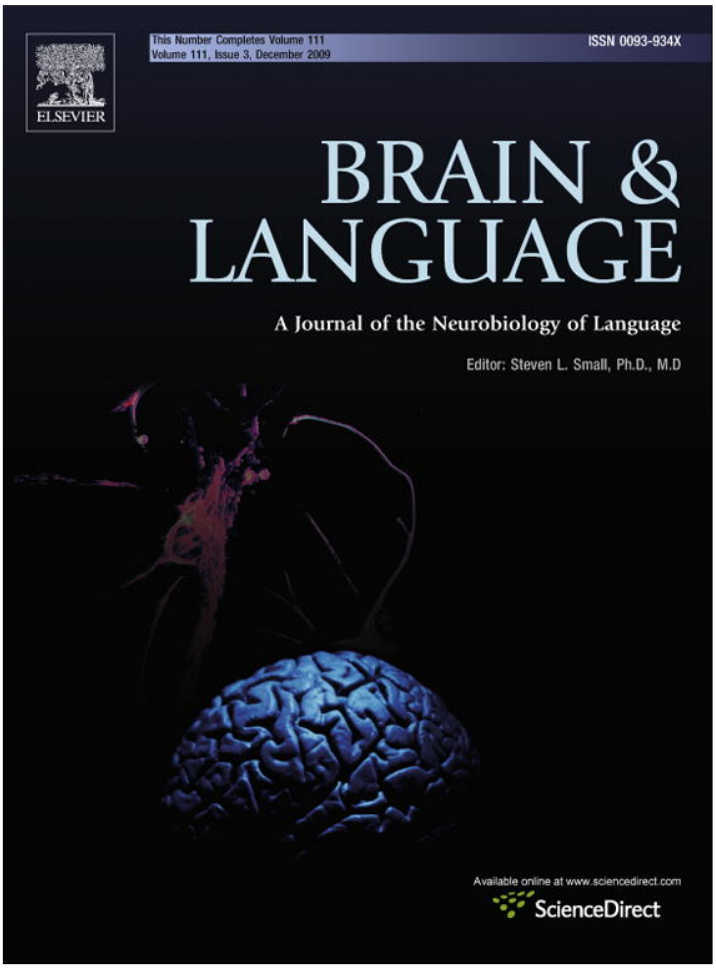
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Brain & Language 111 (2009) 161–172

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

A lexical basis for N400 context effects: Evidence from MEG

Ellen Lau [a,*], Diogo Almeida [a], Paul C. Hines [a], David Poeppel [a,b,c,d]

[a] Department of Linguistics, University of Maryland, College Park, USA
[b] Neuroscience and Cognitive Science Program, University of Maryland, College Park, USA
[c] Department of Biology, University of Maryland, College Park, USA
[d] Department of Psychology, New York University, USA

ARTICLE INFO

Article history:
Accepted 21 August 2009
Available online 7 October 2009

Keywords:
N400
MEG
Semantic priming
Semantic anomaly
Prediction

ABSTRACT

The electrophysiological response to words during the 'N400' time window (~300–500 ms post-onset) is affected by the context in which the word is presented, but whether this effect reflects the impact of context on *access* of the stored lexical information itself or, alternatively, post-access *integration* processes is still an open question with substantive theoretical consequences. One challenge for integration accounts is that contexts that seem to require different levels of integration for incoming words (i.e., sentence frames vs. prime words) have similar effects on the N400 component measured in ERP. In this study we compare the effects of these different context types directly, in a within-subject design using MEG, which provides a better opportunity for identifying topographical differences between electrophysiological components, due to the minimal spatial distortion of the MEG signal. We find a qualitatively similar contextual effect for both sentence frame and prime-word contexts, although the effect is smaller in magnitude for shorter word prime contexts. Additionally, we observe no difference in response amplitude between sentence endings that are explicitly incongruent and target words that are simply part of an unrelated pair. These results suggest that the N400 effect does not reflect semantic integration difficulty. Rather, the data are consistent with an account in which N400 reduction reflects facilitated access of lexical information.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

The role of contextual information in the access of stored linguistic representations has been a major concern of language processing research over the past several decades. Results from many behavioral studies showing contextual effects on phonemic and/or lexical tasks (Duffy, Henderson, & Morris, 1989; Elman & McClelland, 1988; Fischler & Bloom, 1979; Ganong, 1980; Ladefoged & Broadbent, 1957; Stanovich & West, 1981; Warren, 1970) can be taken as evidence that top-down information based on the context influences the access process. However, such results can also be interpreted as post-access effects, in which the contextual information impacts only integration, decision, or response processes (e.g., Norris, McQueen, & Cutler, 2000). Certain measures have been argued to be less affected by response processes than others (i.e., naming vs. lexical decision), but it has been difficult to identify a measure that solely reflects effects on the access mechanism.

Electrophysiological measures, which do not require a behavioral response, provide a means of studying the role of contextual information on access without the interference of overt decision or response processes. Indeed, the ERP response to words known as the N400 component has been shown to be consistently modulated by both lexical and contextual variables, and would thus seem to provide an excellent tool for examining the time-course of top-down information during access of stored lexical information. However, this electrophysiological measure still retains the same problems of interpretation as in the behavioral literature: while on one view, the N400 reflects ease of access due to priming or pre-activation (Federmeier, 2007; Kutas & Federmeier, 2000), on another the N400 reflects the relative difficulty of integration (Brown & Hagoort, 1993; Hagoort, 2008; Osterhout & Holcomb, 1992). Without significant consensus on the functional interpretation of the N400, the component cannot be used unambiguously as a means of answering questions about mechanisms of access or integration.

In this paper, we address the functional interpretation of the N400 by comparing the two paradigms that are typically used to elicit N400 contextual effects, semantic priming in word pairs and semantic fit in sentential contexts. Potential effects of context on access in these two paradigms should be similar in mechanism, while effects of context on integration should be more extensive in sentence contexts, where the critical word must be integrated into

* Corresponding author. Address: Department of Linguistics, 1401 Marie Mount Hall, College Park, MD 20742, USA.
E-mail address: ellenlau@umd.edu (E. Lau).

0093-934X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2009.08.007

Author's personal copy

a larger number of different structured representations (syntactic, thematic, discourse) than in word pairs. For the same reason, integration should also be more difficult for incongruent completions in sentences (where the critical word must be integrated into multiple complex structures) than in word pairs (where the critical word can only be entered into a relatively nonstructured relationship with the prime word). Using MEG, we compare the response profile and topographic distribution of the N400 contextual effect in sentences and word pairs for participants who completed both tasks in the same session, to assess the similarity of the two responses. We also compare the strength of the response to incongruent completions in both paradigms. The results support the hypothesis that N400 amplitude reductions reflect ease of access and not integration.

1.1. Contextual effects on the N400 component

The N400 component is an EEG evoked response to the presentation of words and other meaningful stimuli that appears as a negative-going deflection in the ERP waveform. The N400 usually onsets between 200 and 300 ms and reaches its peak amplitude at around 400 ms post-stimulus onset. An ERP response with this profile has been observed to words, word-like nonwords (Bentin, McCarthy, & Wood, 1985), faces (Barrett & Rugg, 1989), pictures (Barrett & Rugg, 1990; Willems, Ozyurek, & Hagoort, 2008), and environmental sounds (Orgs, Lange, Dombrowski, & Heil, 2008; van Petten & Rheinfelder, 1995), leading to the generalization that the component is associated with processing of "almost any type of meaningful, or potentially meaningful, stimulus" (Kutas & Federmeier, 2000). However, the spatial distribution of both the N400 component and contextual manipulations of the response have been shown to differ slightly across different classes of stimuli (e.g., Barrett & Rugg, 1990; Federmeier & Kutas, 2001; Ganis, Kutas, & Sereno, 1996; Holcomb & McPherson, 1994; Kounios & Holcomb, 1994; Van Petten & Rheinfelder, 1995), which suggests that the N400 component itself is composed of a number of sub-responses, some of which are common to different types of meaningful stimuli and some which are not. This latter point has not been sufficiently appreciated in the literature, where there have been numerous attempts to identify single causes (and often single neural sources) of the N400 effect.

Our study is not aimed at characterizing all of the subprocesses that give rise to the N400 evoked response; we focus on understanding the mechanisms underlying the contextual modulation in amplitude of the N400 response to words which has driven much of ERP language research. Several standard paradigms have been used to show that the N400 response to words is strongly dependent on contextual information. The first and most famous is semantic anomaly: when a sentence is completed with a highly predictable word (*I like my coffee with cream and sugar*), the amplitude of the N400 is much smaller than when the same sentence is completed with a semantically incongruous word (*I like my coffee with cream and socks*; Kutas & Hillyard, 1980). A second frequently used paradigm is semantic priming: when a word target is preceded by a semantically associated word (*salt–pepper*), the amplitude of the N400 is smaller than when a word target is preceded by an unrelated word (*car–pepper*) (Bentin et al., 1985; Rugg, 1985). These effects parallel reaction time data for similar behavioral paradigms in which naming or lexical decision of words in supportive semantic contexts are processed faster than words in unsupportive context (Fischler & Bloom, 1979; Schuberth & Eimas, 1977; West & Stanovich, 1978). Throughout this paper we will refer to this contextual modulation of N400 amplitude as the *N400 effect*, to distinguish it from the component itself.

What is the functional interpretation of this contextual effect? Many years of research have allowed us to rule out certain possibilities. N400 amplitude seems not to reflect degree of semantic anomaly or implausibility per se, as less expected sentence endings show larger N400 amplitudes than expected ones even when they are both plausible (e.g., *I like my coffee with cream and Splenda*; Kutas & Hillyard, 1984). N400 amplitude increases also seem not to reflect a mismatch response to a violation of the predicted ending, as the amplitude of the N400 response to unexpected endings is not modulated by the strength of the expectation induced by the context (e.g., strong bias context – *The children went outside to look (preferred:play)* vs. weak bias context – *Joy was too frightened to look (preferred:speak)*; Kutas & Hillyard, 1984; Federmeier, 2007). Finally, although simple lexical association (*cat–dog*) does seem to contribute to the N400 effect (Federmeier, Van Petten, Schwartz, & Kutas, 2003; Van Petten, 1993; Van Petten, Weckerly, McIsaac, & Kutas, 1997), it does not appear to be the main factor behind N400 effects in sentences. Effects of simple lexical association are severely reduced or eliminated when lexical association and congruency with the larger sentence or discourse context are pitted against each other (Camblin, Gordon, & Swaab, 2007; Coulson, Federmeier, Van Petten, & Kutas, 2005; Van Petten, Coulson, Rubin, Plante, & Parks, 1999), and strong effects of implausibility on sentence endings are observed even when local lexical association is controlled (Camblin et al., 2007; St. George, Mannes, & Hoffman, 1994; van Berkum, Hagoort, & Brown, 1999; Van Petten, 1993; Van Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991).

Despite these advances in our understanding of what the N400 effect does *not* reflect, several accounts are still compatible with the existing data. On one view, N400 amplitude is an index of processes of semantic integration of the current word with the semantic and discourse context built up on the basis of previous words (Brown & Hagoort, 1993). On this view, increased N400 amplitudes reflect increased integration difficulty of the critical word with either a prior sentence context or with the prime word. Alternatively, N400 amplitude indexes processes associated with basic lexical activation and retrieval (Kutas & Federmeier, 2000). On this view, reduced N400 amplitudes reflect facilitated access of lexical information when the word or sentence context can pre-activate (aspects of) the representation of the critical word. Hybrid hypotheses argue that the N400 actually reflects a summation of several narrower component processes (Pylkkänen & Marantz, 2003; van den Brink, Brown, & Hagoort, 2001). These positions parallel the earlier debate over whether the locus of context effects on reaction times is pre- or post-lexical in nature.

Evidence on this debate has been inconclusive. Supporters of the integration view have argued that the N400 effects due to discourse context suggest that the effects reflect the difficulty of integrating the incongruous word into the discourse, but supporters of the access view could account for this by arguing that the discourse predicts or facilitates processing of the congruous ending. Early findings that masked priming paradigms did not show N400 amplitude effects of semantic relatedness (Brown & Hagoort, 1993) were also taken as evidence that the N400 effect reflected integration with context (and so would not be observed when the context was not consciously perceived). However, later work has suggested that masked priming can affect N400 amplitude under appropriate design parameters (Grossi, 2006; Kiefer, 2002). It has also been argued that, at 400 ms post-stimulus onset, the N400 component is too late to reflect the access process (Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006; Sereno, Rayner, & Posner, 1998); however, although the context effect peaks at 400 ms, it usually onsets much earlier (~200 ms for visual stimuli), and on some theories of lexical recognition the access process is extended over time rather than occurring at a single point.

Supporters of the lexical view have argued that effects of basic lexical parameters on the N400 like frequency and concreteness (Holcomb & McPherson, 1994; Rugg, 1990; Smith & Halgren,

Author's personal copy

1987), and effects of predictability over and above plausibility (Federmeier & Kutas, 1999a; Fischler, Bloom, Childers, Roucos, & Perry, 1983; Kutas & Hillyard, 1984) are more consistent with a lexical basis for N400 effects. However, any factors that make access easier could correspondingly be argued to make semantic integration easier (van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Finally, recent MEG studies using distributed source models provide some support for a multiple-component view of the context effect (Halgren et al., 2002; Maess, Herrmann, Hahne, Nakamura, & Friederici, 2006; Pylkkänen & McElree, 2007), but disparities across studies have made it difficult to assess whether the multiple sources reported contribute to the N400 effect and not other parts of the response, such as the post-N400 positivity (Van Petten & Luka, 2006).

In a recent review, we have argued that data on the *localization* of contextual effects on word processing in N400 paradigms support the view that N400 amplitude reflects access processes (Lau, Phillips, & Poeppel, 2008). In this study we take a different approach to the problem, examining the predictions of existing models with respect to how they interpret the N400 effects observed in sentential and single-word contexts and comparing these responses directly using MEG. Our goal is not to localize the contextual effect in these two paradigms, but rather to compare the similarity of the effect across these very different context types, and to further assess the degree to which the effect reflects facilitation or inhibition processes.

Sentential and word-pair tasks differ significantly in the degree of integration required. The final word in a sentence must be integrated with the syntactic, propositional semantic, and discourse structures which have been constructed on the basis of the previous input. Although the experimental task may not require this integration, it is presumably an automatic part of sentence reading. However, while the second word in a word pair may be integrated in some way with the first word, this integration is of a much different nature, as there is no kind of syntactic structure indicating thematic or other kinds of relationships between the two words. Therefore, based on the integration view, we would expect to see qualitative differences in the nature of the N400 effect for sentential contexts and word-pair contexts. On the other hand, access of lexical information proceeds when a word is presented regardless of context, so according to the access view we should not expect to see any qualitative differences. On the multi-component view of the N400, different effects would be expected at different parts of the time-course.

A number of studies have compared contextual effects across the two paradigms using ERPs. Kutas (1993) compared the N400 effect for sentential contexts (expected plausible endings vs. unexpected plausible endings) and the N400 effect for single-word contexts (semantically associated vs. unassociated). While the size of the effect was larger for sentential contexts, replicating previous findings, she found that the latency and scalp distribution of the effect was indistinguishable for the sentential and single-word contexts. Nobre and McCarthy (1994) reported subtle differences between the paradigms with a larger electrode array, but as they did not use the same participants across the different context manipulations, this conclusion is somewhat less reliable. These studies were also subject to the concern that the sentence contexts may have confounded both lexical priming and sentential integration effects, even though care was taken to avoid including lexical associates in the sentence materials. Van Petten (1993) mitigated these concerns in a seminal study in which she isolated lexical and sentential context effects by contrasting the effect of lexical association in congruent sentences and in syntactically well-formed nonsense sentences. She found that lexical and sentential context effects thus isolated had a similar scalp distribution and indistinguishable onset latency, although the sentential context effect lasted longer, into the 500–700 ms window.

Although the ERP studies have thus largely supported the hypothesis that lexical and sentential context effects are due to the same underlying functional and neurobiological mechanism, one might argue that the limited spatial resolution of ERP has caused researchers to miss differences in the neural generators which give rise to these effects. The current study is designed to address this concern. Previous studies have established an MEG correlate to the N400 that shows the same time-course and response properties as the N400 observed in EEG (Halgren et al., 2002; Helenius, Salmelin, Service, & Connolly, 1998; Uusvuori, Parviainen, Inkinen, & Salmelin, 2008). In this study, we compare the N400 effect across these two context types using MEG. MEG measurements are subject to less spatial distortion than ERPs, and thus can provide a better test of whether there exist qualitative differences in the distribution of the effect across these context types.

The access and integration interpretations of the N400 effect also make different predictions about the relative amplitude of the component across the conditions of the sentence and word pair paradigms. An account under which N400 amplitude reflects integration difficulty views the N400 effect as being driven by an *increase* in amplitude for anomalous sentence endings, which are clearly difficult to integrate, while an account under which N400 amplitude reflects access views the N400 effect as driven by a *decrease* in amplitude for predictable sentence endings, where access would be facilitated by contextual support. Therefore, early findings that predictable endings show smaller N400s than congruent but less predictable endings (Kutas & Hillyard, 1984) were taken as support for a facilitated access account, as were findings that N400s to words in congruent sentences are large at the beginning of the sentence and become smaller as the sentence progresses and the next word becomes more predictable (Van Petten, 1993; Van Petten & Kutas, 1990, 1991). These studies further showed that in semantically random sentences, N400 amplitude does not change with word position, even though processing the first word, when the sentence is not yet anomalous, should presumably elicit less integration difficulty than the subsequent words (Van Petten, 1993; Van Petten & Kutas, 1991).

Recently it has been suggested that the integration view can also account for apparent facilitation effects, on the assumption that integration of an item is easier when it can be predicted in advance (Hagoort, 2008; van Berkum et al., 2005). However, our inclusion of both word pair and sentence stimuli in the same session allows an additional test of the directionality of the context effect, through comparison of the unrelated word pair condition with the anomalous sentence ending condition. Even if the N400 effect for sentence completions is partly driven by facilitation of integration in the congruent condition, integration should be more difficult for the anomalous sentence ending, where the final word must be integrated into a semantic and discourse model which clearly violates world knowledge, than in the unrelated word condition, where there is no need to connect the prime and target in a structured way. Therefore, if the amplitude of the N400 response is shown to be bigger for the anomalous sentence ending than the unrelated word target, it would provide novel support for the integration account.

2. Methods

2.1. Participants

Eighteen native English speakers (14 women) were included in the analysis (mean age: 20.9, age range: 18–29). All were right-handed (Oldfield, 1971), had normal or corrected to normal vision, and reported no history of hearing problems, language disorders or mental illness. All participants gave written informed consent and

Author's personal copy

were paid to take part in the study, which was approved by the University of Maryland Institutional Review Board.

2.2. Design and materials

The experiment was comprised of two separate tasks, a sentence task and a word pair task. Each task included two conditions, *contextually supported* and *contextually unsupported*. In the sentence task, this contrast was achieved by using high cloze probability sentences that ended in either expected or semantically anomalous endings. In the word task, this contrast was achieved by using semantically related or semantically unrelated word pairs.

For the sentence task, 160 sentence frames (4–9 words in length) were chosen for which published norms had demonstrated a cloze probability of greater than 70% for the most likely completion word (Bloom & Fischler, 1980; Kutas, 1993; Lahar, Tun, & Wingfield, 2004). Only sentence frames for which the most likely completion was a noun were included. To form the semantically anomalous versions of these 160 sentences, the same 160 sentence-final words were rearranged among the sentences and the resulting sentences were checked for semantic incongruity. The sentence-final target words had an average log frequency of 1.64 and an average length of 5.1 letters. On each of two lists, 80 of the sentence frames appeared with the highly expected ending and 80 of the sentence frames appeared with the semantically anomalous ending. No sentence frame or ending appeared more than once on a given list. The two lists were balanced for surface frequency of the final word.

For the word task, 160 semantically associated word pairs were chosen from existing databases and previous studies that showed semantic priming effects (Holcomb & Neville, 1990; Kutas, 1993; Mari-Beffa, Valdes, Cullen, Catena, & Houghton, 2005; Nelson, McEvoy, & Schreiber, 2004; Thompson-Schill, Kurtz, & Gabrieli, 1998). Both members of each word pair (the prime and the target) were nouns. To form the unrelated word pairs, the primes and targets were rearranged and the resulting pairs were checked for semantic unrelatedness. The target words had an average log frequency of 1.66 and an average length of 4.7 letters. On each of two lists, 80 of the primes were followed by a related target and 80 were followed by an unrelated target. No prime or target appeared more than once on a given list. The two lists were balanced for frequency of the target word.

Following Kutas (1993), we chose an end-of-trial memory probe for the experimental task in order to match the sentence and word parts of the experiment as closely as possible. In the sentence task, following each sentence participants were presented with a probe word and asked whether the word had appeared in the previous sentence. In the word task, following each word pair participants were presented with a probe letter and asked whether the letter had appeared in the previous words. Word probes were taken from various positions of the sentence, but were always content words. Letter probes were taken from various positions in both the prime and target words. On each list, the number of yes/no probe responses was balanced within and across conditions.

2.3. Procedure

Materials for both tasks were visually presented using DMDX software (Forster & Forster, 2003). Sentences were presented one word at a time in the center of the screen using RSVP. Presentation parameters were matched for the sentence and word portions of the experiment as tightly as possible. Except for the first word of sentences and proper names, words were presented only in lower-case. In both tasks words were on the screen for a duration of 293 ms, with 293 ms between words, for a total of 596 ms SOA (stimulus-onset asynchrony). Following the offset of the final word of the sentence or the target word, a 692 ms blank screen was presented before the probe appeared, allowing a 985 ms epoch from the onset of the critical word before the probe was presented in both tasks. In order to make the probes distinct from the targets, probes were presented in capital letters, followed by a question mark. The probe remained on the screen until a response was made. All words in the experiment were presented in 12-point yellow Courier New font.

In order to maximize attentiveness across the session, the longer sentence task was presented first, and the faster-paced word pair task was presented second. For the sentence task, participants were instructed that after the sentence was complete, they would be presented with a word and asked to make a button-press response indicating whether the word was present in the previous sentence. For the word task, participants were instructed that after each word pair was presented, they would be presented with a letter and asked to make a button-press response indicating whether the letter was present in either of the two words. Participants were allowed up to 3.5 s to make their response. Both parts of the experiment were preceded by a short practice session, and both parts of the experiment were divided into four blocks, with short breaks in between. Including set-up time, the experimental session took about 1.5 h.

2.4. Recordings and analysis

Subjects lay supine in a dimly lit magnetically shielded room (Yokogawa Industries, Tokyo, Japan) and were screened for MEG artifacts due to dental work or metal implants. A localizer scan was performed in order to verify the presence of identifiable MEG responses to 1 kHz and 250 Hz pure tones (M100) and determine adequate head positioning inside the machine.

MEG recordings were conducted using a 160-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan). Data were sampled at 500 Hz (60 Hz online notch filter, DC – 200 Hz recording bandwidth). A time-shift PCA filter (de Cheveigné & Simon, 2007) was used to remove external sources of noise artifacts. Epochs with artifacts exceeding 2 pT in amplitude were removed before averaging. Incorrect behavioral responses and trials in which subjects failed to respond were also excluded from both behavioral and MEG data analysis. Based on these criteria, 9% of trials were excluded. For the analyses presented below, data were averaged for each condition in each participant and baseline corrected using a 100 ms prestimulus interval. For the figures, a low-pass filter of 30 Hz was used for smoothing (note that this is 'visually generous,' including higher frequencies than often shown – typically, much lower frequencies are used for visualization of such late effects).

Two participants were excluded from the analysis based on low accuracy (<80%) in one or both sessions of the experiment, and one participant was excluded from the analysis based on high levels of signal artifact, leaving 18 participants in the analysis. Overall, there was a tendency for accuracy on the word pair task to be lower than in the sentence task (95% vs. 99%). Data from the 18 remaining participants (nine from each counterbalanced stimulus list) were entered into the MEG analysis.

The response topography generally observed to written words in the time window associated with the N400 (300–500 ms) has a similar distribution to the response topography previously associated with the M350 component, which has been posited to reflect lexical access processes (Pylkkänen & Marantz, 2003; Pylkkänen, Stringfellow, & Marantz, 2002). However, we found that not all conditions in the current experiment displayed this topography. While most participants displayed a M350-like pattern in the 300–500 ms window for the word-pair conditions and the sentence-incongruous condition, most displayed a qualitatively differ-

Author's personal copy

ent topography to the sentence-congruous condition. Thus, an analysis comparing peak latency and amplitude across all four conditions would be inappropriate.

As an alternative measure, we created statistically thresholded difference maps for the sentence and word conditions. For each participant, we created two difference maps, one for sentence-incongruous–sentence-congruous, and one for word-unrelated–word-related. We grand-averaged these individual participant maps to create a composite difference map for each task. However, because head position can vary in MEG, participants may contribute unequally to the differential activity observed at individual sensors. Therefore, we created a statistically thresholded difference map for each task by displaying activity only for sensors at which the difference between conditions was significantly different from zero across participants ($p < .01$). To correct for multiple comparisons across the large number of recording sites, we clustered the significant sensors based on spatial proximity and polarity and then conducted a randomization test of the summary test statistic for clusters of sensors that crossed this initial threshold (Maris & Oostenveld, 2007). For each thresholded difference map, we assigned each significant sensor to a cluster. Sensors within 4 cm of each other that demonstrated the same polarity were assigned to the same cluster. A summary test statistic was calculated for each cluster by summing the t-values of all the sensors in that cluster (t-sum). This test statistic (t-sum) was then recomputed in 4000 simulations in which the condition labels for each participant were randomly reassigned for each subject. This generated a random sample of the empirical distribution of t-sum under the null hypothesis of no treatment effect. If the t-sum values from the clusters obtained in our experiment were more extreme than 95% of the t-sum values calculated by the randomization simulations, the cluster was considered statistically significant.

Finally, to test the simple directionality of contextual effects (facilitatory vs. inhibitory) across different conditions, we computed the grand-average RMS over all sensors in each hemisphere (75 left-hemisphere sensors and 75 right-hemisphere sensors; 6 midline sensors were excluded for this analysis) for each condition across the M400 time-window. This provided a relatively insensitive but conservative measure of the most robust experimental effects. $2 \times 2 \times 2$ ANOVAs (hemisphere $\times$ task $\times$ contextual support) were conducted on the average waveform for each condition over 200 ms windows: 300–500 ms (N/M400), 500–700 ms, and 700–900 ms, followed by planned comparisons. We also examined contextual effects on the M170 component by computing the grand-average RMS over posterior sensor quadrants only, in the 100–300 ms time-window.

3. Results

The grand-average MEG waveform and average topography over the 300–500 ms window for the context effect (unsupportive–supportive) for each context type prior to statistical thresholding is presented at the bottom right of Fig. 1. The context effect was much larger in amplitude for sentences than word pairs, consistent

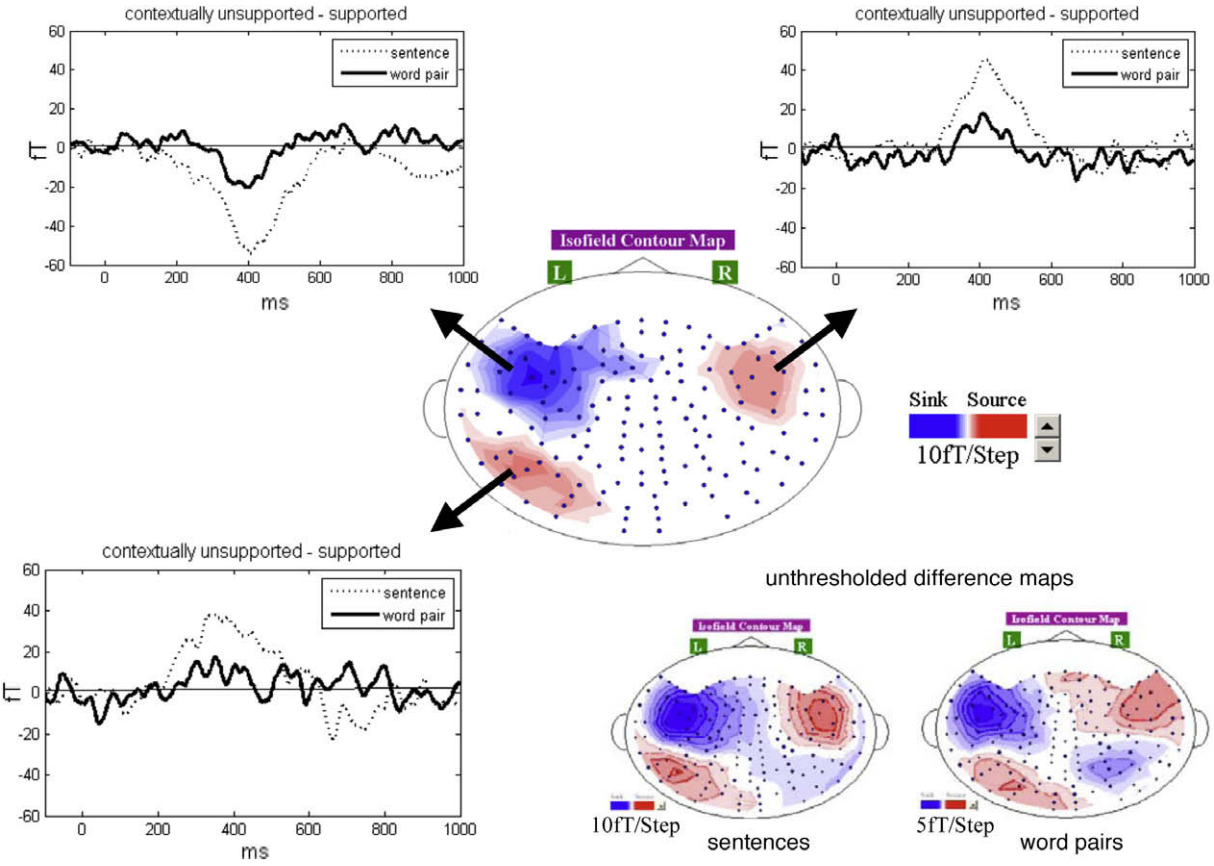


Fig. 1. Statistically thresholded grand-average whole-head topography for the sentence ending contrast (contextually unsupported–contextually supported). The central image shows activity only for sensors for which the difference between sentence conditions was significant across participants between 300 and 500 ms ($p < .01$). The waveforms show the average difference waves (contextually unsupported–supported) for sentences and word-pair conditions across the significant sensors in the cluster indicated by the arrow; all three clusters were larger than expected by chance (see text). In the lower right, the unthresholded grand-average whole-head topography at 400 ms for the contextually unsupported–contextually supported contrast is presented unthresholded to show general similarities in the topography of the contextual effect for the two context types; because the magnitude of the effect was smaller in the word-pair conditions, the scale has been decreased for this contrast relative to the sentence contrast.

Author's personal copy

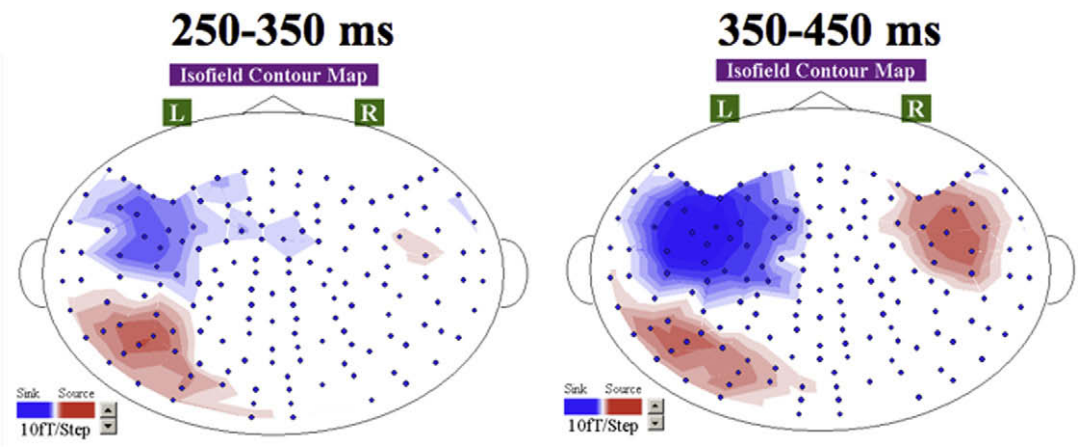


Fig. 2. Statistically thresholded grand-average whole-head topography for the sentence ending contrast (contextually unsupported–contextually supported) averaged across two time-windows chosen by visual inspection, showing only those sensors for which the difference between conditions was significant across participants in that time-window ($p < .01$). In the first time-window (250–350 ms), the two left-hemisphere clusters were larger than expected by chance, while in the second time-window (350–450 ms), an additional right-hemisphere cluster was also larger than expected by chance (see text).

with previous literature (Kutas, 1993), but the timing and topographical distribution appeared quite similar.

3.1. Cluster analysis

To compare the contextual effect across context types, it was necessary to select a subset of sensors of interest. As the contextual effect is known to be strongest for sentence contexts, we used the response observed for the sentence conditions as the basis for sensor selection. We created a statistically thresholded topographical map for the unsupportive–supportive sentence contrast (Fig. 1). This map displays sensors that showed a reliable difference ($p < .01$) between the two sentence contexts across subjects in the 300–500 ms window. This procedure identified three clusters of sensors distinguished by polarity and hemisphere: a left anterior cluster that showed a negative difference, a left posterior cluster which showed a positive difference, and a right anterior cluster which showed a positive difference. The randomization clustering analysis showed that the three clusters of activity observed in the sentence context map had a less than 5% probability of having arisen by chance (sums of t-values over significant sensors: left anterior sink = −116.1; right anterior source = 54.0; left posterior source = 87.0; 2.5% and 97.5% quantiles = $-15.1 > t\text{-sum} > 16.0$), but that only the left anterior cluster was reliable for the word context effect (sum of t-values = −26.5; 2.5% and 97.5% quantiles = $-14.2 > t\text{-sum} > 13.6$). This could indicate a qualitative difference between context types, but it could also reflect the difference in magnitude between the context types; the word effect may have been too weak to survive the analysis in all but the largest cluster.

Fig. 1 supports this latter interpretation. Here we present the difference waves for the contextual effect (contextually unsupported–supported) for sentences and word pairs. As the sensors here were defined as those that showed a significant difference between sentence conditions, the fact that the sentence conditions show strong differences in the N400 time-window is unsurprising. The question of interest is whether the word-pair conditions also showed a context effect across the same sensors. Indeed, in the word pair comparison we observe a difference between contextual conditions in the 300–450 ms window in all three clusters, significant in the left anterior cluster ($t(17) = 3.66$, $p < .01$), and marginally significant in the right anterior cluster ($t(17) = 2.05$, $p = .057$) and left posterior cluster ($t(17) = 2.04$, $p = .057$). No sensors showed a significant contextual effect for word pairs but not for sentences.

A further similarity between the sentence and word pair context effects can be observed in the timing of the effects across hemispheres. Visual observation of the grand-average difference map for the sentence condition revealed that the contextual effects observed bilaterally in anterior sensors onset earlier over the left hemisphere than the right hemisphere (Fig. 2). This was confirmed by statistically thresholding the sensors across subjects ($p < .01$); in an early time-window (250–350 ms) only the two left-hemisphere clusters were significant (sums of t-values over significant sensors: left anterior sink = −54.6; right anterior source = 6.6; left posterior source = 83.6; 2.5% and 97.5% quantiles = $-16.8 > t\text{-sum} > 18.1$), while in a later time-window (350–450 ms) the right anterior cluster was also significant (sums of t-values over significant sensors: left anterior sink = −54.6; right anterior source = 6.6; left posterior source = 83.6; 2.5% and 97.5% quantiles = $-16.9 > t\text{-sum} > 17.7$).

Crucially, the same right-hemisphere delay was observed in the word conditions. Fig. 3 plots the average context difference wave (contextually unsupported–supported) across the left and right anterior clusters depicted in Fig. 1, with the polarity of the left cluster waveform reversed to facilitate visual comparison. We tested consecutive 50-ms windows to determine when significant effects of context began ($p < .05$). For both the sentence and the word-pair conditions, the contextual effect began earlier in the left anterior sensors than in the right, although effects over both hemispheres became significant later for words than for sentences (200–250 ms (left) vs. 300–350 ms (right) for sentences; 300–350 ms (left) vs. 350–400 (right) ms for words). Importantly, however, both stimulus types showed the same hemispheric asymmetry in latency: left-hemisphere effects of context earlier than right-hemisphere effects.

Figs. 1–3 depict the difference waves and topographical maps for the contextually unsupported–supported contrast. Fig. 4 presents the grand-average MEG waveforms by sensor quadrant. This figure indicates that the response to contextually unsupported words was of similar magnitude in word pairs and sentences in left-hemisphere clusters, and that the contextually supported words showed a shift towards baseline that was greater in sentences than in word pairs. However, even though the analyses above suggest that the effect of the contextual manipulation is similar in word pairs and sentences, the condition waveforms suggest that the base response to the unsupported condition in the right-hemisphere cluster is greater when the target word is

Author's personal copy

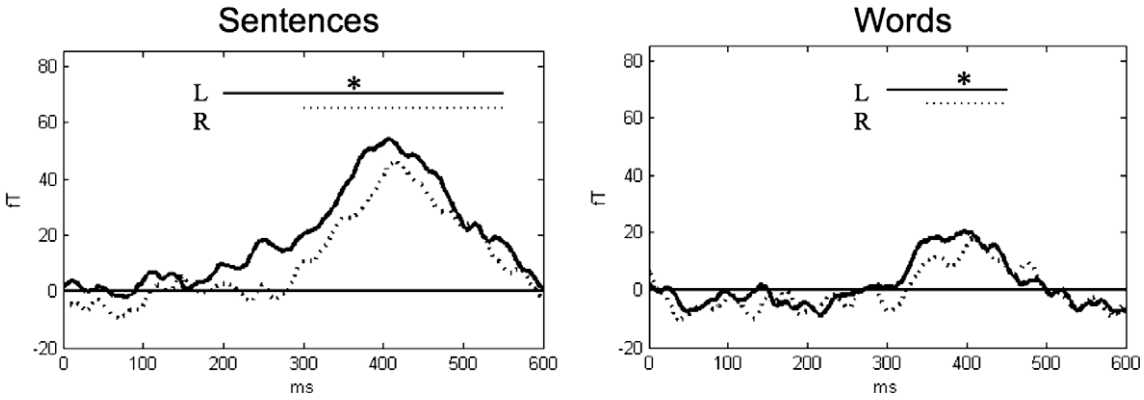


Fig. 3. Difference waves (contextually unsupported–supported) averaged across the left and right anterior clusters of sensors depicted in Fig. 1 for sentences and word pairs. The polarity of the difference wave in the left cluster has been reversed to facilitate comparison between the timing of the responses across hemispheres.

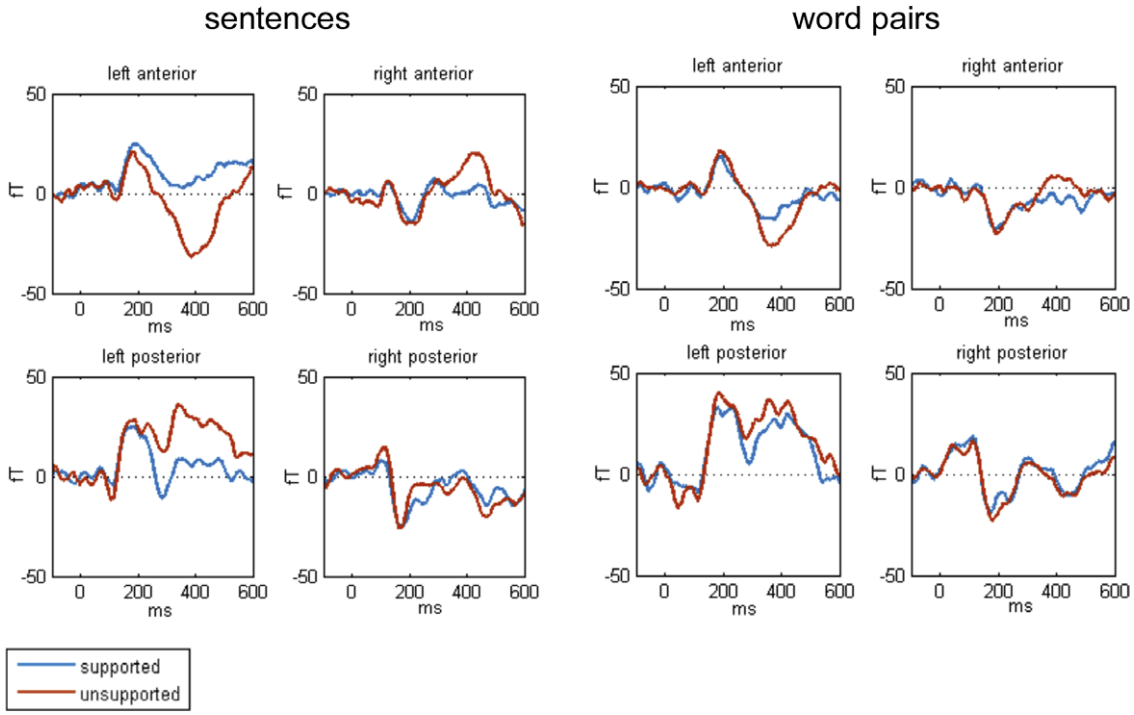


Fig. 4. Grand-average MEG waveforms for contextually supported and unsupported targets in both sentence (left) and word pair (right) contexts, presented by sensor quadrant (anterior/posterior and left/right).

embedded in a sentence context. We review possible interpretations of this pattern in the Discussion.

Overall, consistent with previous ERP studies, we find that sentence and word pairs show similar effects of contextual support in the N400 time window. While the contextual effect for word pairs was much smaller in magnitude, this is plausibly due to differences in the strength of the lexical prediction made possible by sentence contexts and single prime words. The contextual effect for word pairs was significant across a shorter time window than the context effect for sentences, but this could be due to the smaller magnitude of the effect.

3.2. Late effects of context

We also examined activity in a later time window (600–900 ms) within which a post-N400 positivity is often visible in ERP recordings (e.g., Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Friederici & Frisch, 2000; Kuperberg, 2007; Kutas & Hillyard, 1980; Van Petten & Luka, 2006 for review), but which has rarely been reported in MEG studies. No significant clusters were found using a threshold of $p < .01$. At a more conservative threshold ($p < .05$) we found two clusters of sensors in the right hemisphere showing a difference between contextually supported and unsupported sentence endings (Fig. 5); these clusters were marginally reliable in the nonparametric test (sums of t-values over significant sensors: right anterior sink = −42.7; right posterior source = 32.2; 2.5% and 97.5% quantiles = −34.3 > t-sum > 34.3). No clusters showed a difference between contextually supported and unsupported word pairs.

3.3. RMS analysis by hemisphere

It is difficult to generalize MEG analyses of subsets of sensors across studies because the position of the head with respect to the sensors is different across participants and MEG systems. Averaging MEG signal strength across all sensors in each hemisphere is one means of avoiding sensor selection and thus making results

Author's personal copy

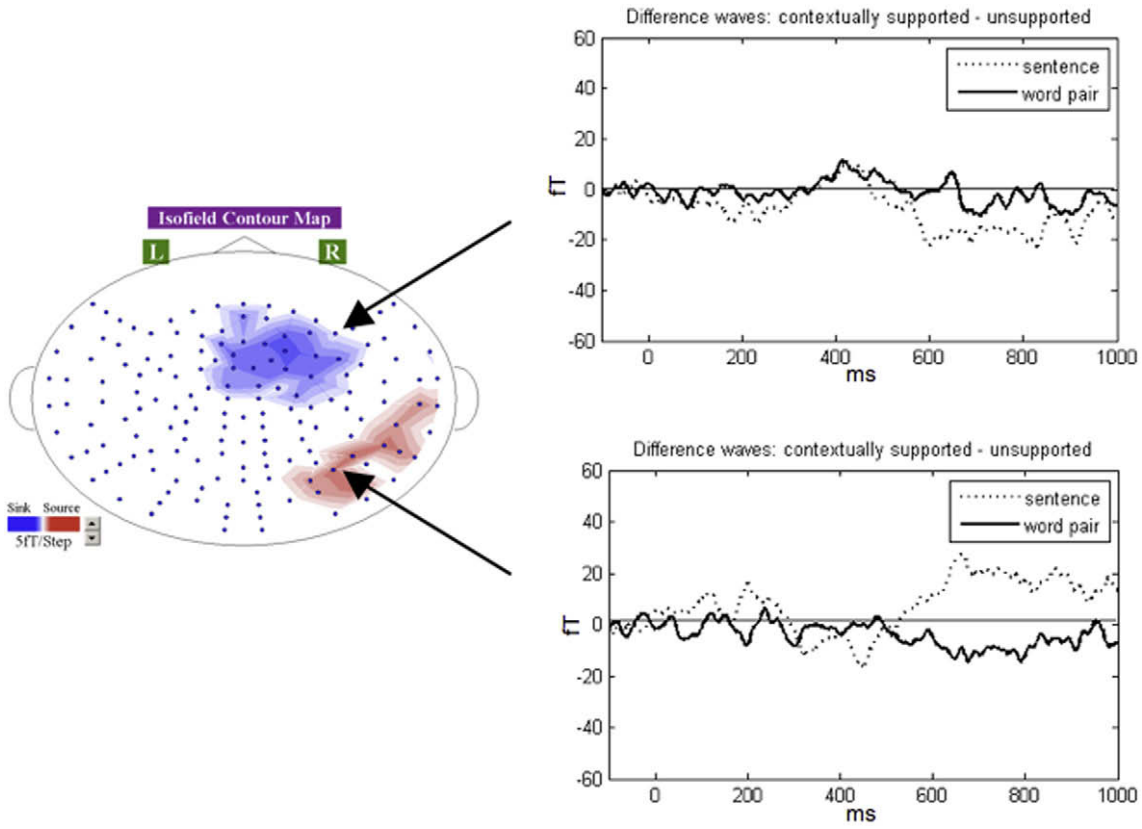


Fig. 5. Statistically thresholded grand-average whole-head topography for the difference between contextually unsupported–contextually supported sentence endings. This image only displays activity over those sensors for which this difference was significant across participants between 600 and 900 ms ($p < .05$). Difference waves for the contextual effect (contextually unsupported–contextually supported) are presented for sentences and word pairs across the sensors in the clusters indicated; these clusters were only marginally reliable, however (see text).

potentially more generalizable, even though it means: (i) allowing for greater anatomic variability and (ii) conceding that many sensors which do not contribute to the effect of interest will be included, making this a less sensitive measure. Fig. 6 presents the grand-average RMS across all sensors over each hemisphere for all four conditions. We observed three patterns of interest, described in more detail below. First, regardless of context, sentences showed a larger M170 response than words in right-hemisphere sensors. Second, in left-hemisphere sensors, the response to incongruous sentence endings was relatively similar in amplitude to both related and unrelated word targets during the N400 time window, while the congruous sentence ending was strongly reduced. Third, during the same time window in right-hemisphere sensors, the response to incongruous sentence endings showed increased amplitude relative to the other three conditions.

Visual inspection indicated an amplitude difference between conditions peaking at around 150 ms in the right hemisphere. This is the approximate time-window of the M170 component in MEG, which is an evoked response to visual stimuli associated with higher-level visual processing. The M170 is thought to be generated in occipital cortex and is typically observed in posterior sensors (Tarkiainen, Cornelissen, & Salmelin, 2002; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999). We therefore restricted our statistical analysis to the right and left posterior quadrants of sensors. Across these 57 sensors in the M170 window (100–300 ms) we found a main effect of hemisphere ($F(1, 17) = 7.26$; $p < .05$) and a significant interaction between task and hemisphere ($F(1, 17) = 5.41$; $p < .03$). These effects seemed to be driven by a hemispheric asymmetry in M170 amplitude for the word conditions (increased amplitude in the left hemisphere) but not the sentence conditions. Such a leftward asymmetry in the word conditions may have been due to the 'local' attention to the letters of the word required to perform the probe task (e.g., Robertson & Lamb, 1991). No other early effects were observed.

In the N400 window (300–500 ms) there were main effects of hemisphere ($F(1, 17) = 27.17$; $p < .01$), and contextual support ($F(1, 17) = 11.93$; $p < .01$), such that signal strength was greater in the left hemisphere than the right and in the unsupportive context than in the supportive context. There were marginally significant interactions between contextual support and hemisphere ($F(1, 17) = 4.05$; $p < .06$) and between context type and hemisphere ($F(1, 17) = 3.88$; $p < .07$), and a significant interaction between contextual support and context type ($F(1, 17) = 16.03$; $p < .01$). These interactions seemed to be driven by the strong reduction in activity observed in the left hemisphere for the congruent sentence ending relative to the other three conditions and the strong increase in activity observed for the incongruent sentence ending in the right hemisphere (Fig. 6B). Paired comparisons between the sentence endings and the unprimed word target confirmed this visual impression. In the left hemisphere there was a significant difference between the unprimed word target and the congruent sentence ending ($F(1, 17) = 14.0$; $p < .01$) but not the incongruent sentence ending ($F(1, 17) = 1.07$; $p > .1$). In the right hemisphere there was no difference between the unprimed word and the congruent ending ($F(1, 17) = .02$; $p > .1$), although there was a marginally significant difference between the unprimed word and the incongruent ending ($F(1, 17) = 4.11$; $p < .06$), which we return to in the discussion.

No significant main effects or interactions were observed for the late (600–900 ms) window in which post-N400 positivities are sometimes found ($ps > .1$).

Author's personal copy

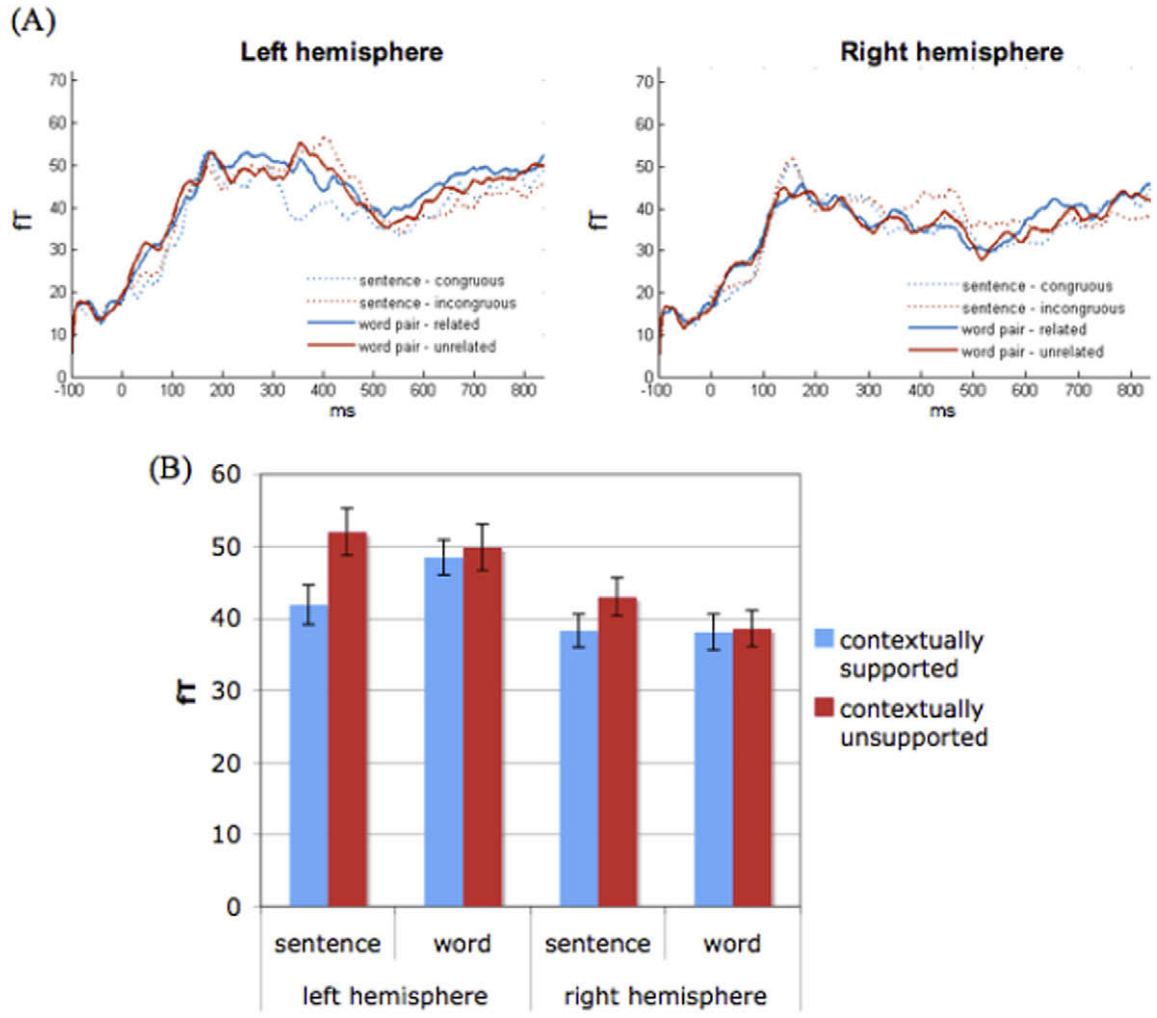


Fig. 6. (A) Grand-average RMS across all sensors in each hemisphere (75 sensors in left; 75 sensors in right) for all four conditions. (B) Grand-average RMS amplitude across the 300–500 ms window in each hemisphere.

Finally, visual inspection suggested a main effect of task in an early time-window. We therefore conducted an exploratory analysis of the 0–100 ms window. We found main effects of hemisphere ($F(1, 17) = 5.59$; $p < .05$) and context type ($F(1, 17) = 9.32$, $p < .01$), such that activity was greater in left-hemisphere sensors in this time window across all conditions, and activity was greater in this time-window for words presented in single-word contexts than sentence contexts across conditions and hemispheres. A number of global differences between the two tasks could be responsible for this early main effect. Many more words were presented per trial in the sentence task than in the word task; this increased density of stimuli per trial may have reduced the baseline response to each individual stimulus, perhaps partially as a consequence of the increased temporal predictability of the stimulus presentation. Attentiveness to the stimuli may also have differed between the two tasks, as the behavioral results suggested the letter probe task was more difficult than the word probe task.

4. General discussion

We used MEG to compare the electrophysiological effects of contextual support in more structured sentence contexts with less structured word pairs in the time window associated with the N400 effect. The results show a significant quantitative difference in effects of contextual support—much bigger effects in sentence contexts than word pairs—but no evidence of qualitative differences. Both sentence and word contexts showed the same distinctive topographical MEG profile (left anterior negativity–left posterior positivity–right anterior positivity) across most of the N400 time window, and both showed timing differences, onsetting earlier in left-hemisphere sensors than in right-hemisphere sensors. A comparison of the amplitudes of the individual word and sentence conditions also revealed that the N400 response for incongruous sentence endings did not differ from the response to unprimed words over the left hemisphere; rather, it was the congruous or predicted sentence endings that showed a significant difference from unprimed words, in the form of a reduction.

4.1. Contextual effects in sentences vs. word pairs

The finding that the N400 contextual effect in MEG is qualitatively similar for sentence contexts and word contexts is consistent with earlier work which showed that the timing of the N400 effect in EEG and its topographical distribution over the scalp were indistinguishable for the two context types (Kutas, 1993; Van Petten, 1993). MEG signals are not subject to the same field distortions that affect EEG, and thus, using MEG provides an opportunity of spatially separating components that could appear to be the same in EEG. Although future studies could conduct more detailed spatial analyses on this type of MEG data to test for even subtler dis-

Author's personal copy

tinctions in the response, the fact that the MEG profile is so similar for N400 effects across the two context types provides compelling evidence that the effects observed across these quite different paradigms do indeed share a common locus.

Many of the operations set in motion by the onset of a word in a sentence context are quite different from those elicited by the presentation of a word in a pair or list. When a word is encountered in a sentence, it must be integrated into a number of structured representations: a syntactic representation, a propositional semantic representation, a mental model of the entities and events described in the sentence or the text that the sentence is a part of. When a word is encountered in a pair or list, any integrative operations elicited will be much more variable and more dependent on the pragmatic context in which the words are embedded: in a game or experiment context, the reader may attempt to access simple stored relationships between the words presented; in a memory task, the reader may try to build new relationships between the words in order to better encode them. The operations that seem more likely to be shared across structured and unstructured language contexts are those underlying access of stored representations. In whatever context a word is encountered, the reader must go through the process of matching the input pattern to a stored representation of the form-meaning pair it is associated with. Our finding that the contextual effect on the N400 is similar in both structured and unstructured language contexts thus provides support for the view that the N400 effect reflects facilitated access of lexical information.

In this paper we focused on comparing the electrophysiological profile of contextual effects across the two context types rather than attempting to localize the source of those effects. However we did observe a clear contrast in the topography of the effect across the two hemispheres. Under the facilitated access view of the N400 effect, the difference map obtained by subtracting predicted words from unpredicted words is essentially a map of the activity resulting from normal word processing (because it is these processes which are thought to be less effortful when the word can be predicted); therefore, the observed hemispheric asymmetry could reflect differences in the left and right hemisphere structures used for word processing (e.g., Beeman & Chiarello, 1998; Federmeier, 2007; Federmeier & Kutas, 1999b). Interestingly, while previous studies in fMRI have identified many left hemisphere temporal and frontal areas to be involved in language processing, one of the few right hemisphere structures to be consistently implicated is the right anterior temporal cortex (Davis & Johnsrude, 2003; Giraud et al., 2004; Lindenberg & Scheef, 2007). However, another possibility is that activity across right and left anterior sensors in the later part of the N400 time window reflects the two poles of an anterior medial source (Pylkkänen & McElree, 2007). Future work with MEG source localization is needed to dissociate these two possibilities.

4.2. Directionality of the N400 effect

If the amplitude of the N400 reflected integration difficulty, one would expect that N400 amplitude should be greater when a word must be integrated with an obviously incongruous highly-structured prior context (*I like my coffee with cream and socks*) than when a word must be integrated with a nonstructured, and thus more neutral, prime-word context (*priest–socks*). In the first case, the structure highly constrains the possible interpretations of the sentence, such that any licit interpretation is at least unusual, while in the second case, the absence of structure means that: (1) a tight integration of the two words is not required and (2) a number of relations are possible between the two words, increasing the likelihood that a congruous integration of the two concepts can be easily achieved (e.g., *priests wear socks*, *the socks belonging to the priest*). However, the results here show no difference between incongruous sentence endings and unprimed words in left-hemisphere sensors, where the N400 effect was earliest and more spatially widespread. The congruous endings, on the other hand, showed a significant reduction in amplitude relative to unprimed words, consistent with lexical facilitation accounts in which the lexical entry for the word forming the sentence ending can be pre-activated by the highly predictive sentence frames used in most N400 studies. This is consistent with previous work suggesting that the N400 effect is driven by a reduction in activity relative to a neutral baseline (Kutas & Federmeier, 2000; Van Petten, 1993; Van Petten & Kutas, 1990, 1991).

Previous authors have suggested that predictability effects not due to congruity can still be explained by an integration account of the N400, if prediction of an upcoming word can facilitate integration processes (Hagoort, 2008; van Berkum et al., 2005). In other words, prediction can take the form of not only *pre-activation* of a stored lexical entry, but also *pre-integration* of the predicted lexical item with the current context. Therefore, if the N400 reflects integration difficulty, it will be reduced in predictive cases where little integration work is left to be done by the time the bottom-up input is encountered. While this can explain the reduction observed in the congruent sentence endings relative to a neutral baseline, the lack of an increase in the incongruent sentence endings still seems to be unaccounted for. If the N400 does not reflect the difference between a case where there are few pragmatic constraints on integration and a case where there are strong pragmatic constraints that make a felicitous derived representation hard to achieve, it is hard to see how its amplitude could be said to reflect integration difficulty. At the least, it requires a re-defining and sharpening of what is meant by 'integration' (van Berkum, 2009).

In right-hemisphere sensors, we did observe evidence of the pattern predicted by the integration account: a marginally significant difference between incongruous sentence endings and unprimed words, and no significant difference between congruous endings and unprimed words. One potential explanation of this asymmetry is that left-hemisphere sensors reflect predictability in the N400 time window, while right-hemisphere sensors reflect integration difficulty; Federmeier and colleagues have previously suggested that the left hemisphere may be preferentially dedicated to prediction in comprehension (e.g.,Federmeier & Kutas, 1999b; Federmeier, 2007). However, the size and timing of the context effect was so similar across corresponding left and right anterior sensor clusters (Fig. 4) that this seems somewhat unlikely; if sensors over different hemispheres reflected different processes, this similarity would have to be seen as largely coincidental.

An alternative is that it is not differences in the size of the context effect that drive the hemispheric asymmetry, but differences in the base level of activity for sentences and words. For sentences, sensors in both hemispheres show a strong peak of activity in incongruent endings and activity close to baseline for predicted endings. For words in word pairs, on the other hand, both unrelated and related targets show a broad peak of activity in left-hemisphere sensors, but in right-hemisphere sensors, activity is close to baseline for both conditions. An explanation for the asymmetry consistent with this pattern is that sources reflected in right-hemisphere sensors are recruited to a greater degree during normal sentence processing than in processing of isolated words. If this were true, activity in right-hemisphere sensors would *always* be higher in amplitude for words in sentences than words in pairs, unless contextual support facilitated processing across the board. Future work will be needed to determine which of these accounts best explains the asymmetry observed.

We also observed a number of right-hemisphere sensors that showed a significant effect of context for sentence endings in the

Author's personal copy

later part of the processing time-course, between 600 and 900 ms. The timing of this effect is consistent with a late positivity often observed following N400 contextual effects in ERP which has sometimes been called the *post-N400 positivity* (Van Petten & Luka, 2006). As this effect was observed for the sentence contexts, which differed on predictability and semantic congruity, but not for the word pairs, for which semantic congruity is less well-defined, this later effect may reflect difficulty in compositional semantic integration, reanalysis, or some other response specific to semantic incongruity. However, it could also reflect a sentence-specific mechanism engaged during normal processing that is simply facilitated in predictive contexts. More studies must be done to tease these possibilities apart.

5. Conclusion

Using MEG, we have shown that: (1) the effect of contextual support between 300 and 500 ms is qualitatively similar in timing and topography for words presented in structured sentences and words presented in unstructured word pairs and (2) in left-hemisphere sensors, at least, the effect is driven by a reduction in amplitude in the predicted sentence ending relative to an unprimed target word, rather than an increase in amplitude in the incongruous ending. These results support a view in which the N400 effect observed in ERP reflects facilitated access of stored information rather than relative difficulty of semantic integration. Together with converging results from other directions (e.g., DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999a, 1999b), these data renew the hope that the N400 effect may be fit for use as a tool for indexing 'true' top-down effects on access.

Acknowledgments

We thank Jeff Walker for invaluable technical assistance. Preparation of this manuscript was supported by an NSF Graduate Research Fellowship to EFL, R01DC05660-06 to DP, and NSF DGE-0801465 to the University of Maryland.

References

Barrett, S. E., & Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia, 27*(7), 913–922.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition, 14*(2), 201–212.

Beeman, M. J., & Chiarello, C. (1998). Complementary right and left hemisphere language comprehension. *Current Directions in Psychological Science, 7*, 2–8.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60*(4), 343–355.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition, 8*(6), 631–642.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience, 5*(1), 34–44.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*(1), 103–128.

Coulson, S., Federmeier, K. D., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 129–147.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *Journal of Neuroscience, 23*(8), 3423–3431.

de Cheveigné, A., & Simon, J. Z. (2007). Denoising based on time-shift PCA. *Journal of Neuroscience Methods, 165*(2), 297–305.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121.

Duffy, S. A., Henderson, J. M., & Morris, R. K. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(5), 791–801.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language, 27*(2), 143–165.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491–505.

Federmeier, K. D., & Kutas, M. (1999a). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41*, 469–495.

Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research, 8*, 373–392.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 202–224.

Federmeier, K. D., Van Petten, C., Schwartz, T. J., & Kutas, M. (2003). Sounds, words, sentences: Age-related changes across levels of language processing. *Psychology and Aging, 18*(4), 858–872.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research, 1146*, 75–84.

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on world recognition. *Journal of Verbal Learning and Verbal Behavior, 18*(1), 1–20.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology, 20*(4), 400–409.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116–124.

Friederici, A. D., & Frisch, S. (2000). Verb argument structure processing: The role of verb-specific and argument-specific information. *Journal of Memory and Language, 43*(3), 476–507.

Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8*(2), 89–106.

Ganong, W. F. 3rd, (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 110–125.

Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., et al. (2004). Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cerebral Cortex, 14*(3), 247–255.

Grossi, G. (2006). Relatedness proportion effects on masked associative priming: An ERP study. *Psychophysiology, 43*(1), 21–30.

Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363*(1493), 1055–1069.

Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage, 17*(3), 1101–1116.

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word-recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*(4), 1383–1400.

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*(Pt 6), 1133–1142.

Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition, 24*(2), 259–276.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes, 5*(4), 281–312.

Kiefer, M. (2002). The N400 is modulated by unconsciously perceived masked words: Further evidence for an automatic spreading activation account of N400 priming effects. *Cognitive Brain Research, 13*(1), 27–39.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 804–823.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23–49.

Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes, 8*(4), 533–572.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203–205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161–163.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the Acoustical Society of America, 29*, 98.

Lahar, C. J., Tun, P. A., & Wingfield, A. (2004). Sentence-final word completion norms for young, middle-aged, and older adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 59*(1), 7–10.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De) constructing the N400. *Nature Reviews Neuroscience, 9*(12), 920–933.

Lindenberg, R., & Scheef, L. (2007). Supramodal language comprehension: Role of the left temporal lobe for listening and reading. *Neuropsychologia, 45*(10), 2407–2415.

Author's personal copy

Maess, B., Herrmann, C. S., Hahne, A., Nakamura, A., & Friederici, A. D. (2006). Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Research, 1096*(1), 163–172.

Mari-Beffa, P., Valdes, B., Cullen, D., Catena, A., & Houghton, G. (2005). ERP analyses of task effects on semantic processing from words. *Cognitive Brain Research, 23*, 293–305.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The university of south Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402–407.

Nobre, A. C., & McCarthy, G. (1994). Language-related ERPS–scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience, 6*(3), 233–255.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*(3), 299–325.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2008). N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neuroscience Letters, 436*(2), 133–137.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785–806.

Pylkkänen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Science, 7*(5), 187–189.

Pylkkänen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience, 19*(11), 1905–1921.

Pylkkänen, L., Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language, 81*, 666–678.

Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology, 23*(2), 299–330.

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22*(6), 642–647.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition, 18*(4), 367–379.

Schuberth, R. E., & Eimas, P. D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance, 3*(1), 27–36.

Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *Neuroreport, 9*(10), 2195–2200.

Smith, M. E., & Halgren, E. (1987). Event-related potentials during lexical decision: Effects of repetition, word frequency, pronounceability, and concreteness. *Electroencephalography Clinical Neurophysiology Supplement, 40*, 417–421.

St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience, 6*, 70–83.

Stanovich, K. E., & West, R. F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7*(3), 658–672.

Tarkiainen, A., Cornelissen, P. L., & Salmelin, R. (2002). Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain, 125*(5), 1125.

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*(11), 2119.

Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language, 38*, 440–458.

Uusvuori, J., Parviainen, T., Inkinen, M., & Salmelin, R. (2008). Spatiotemporal interaction between sound form and meaning during spoken word perception. *Cerebral Cortex, 18*(2), 456–466.

van Berkum, J. J. A. (2009). The neuropragmatics of 'simple' utterance comprehension: An ERP review. In: U. Sauerland & K. Yatsushiro (Eds.), *Semantic and pragmatics:From experiment to theory*.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPS and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 443–467.

van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*(6), 657–671.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13*(7), 967–985.

Van Petten (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes, 8*, 485–531.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 394–417.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18*(4), 380–393.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory & Cognition, 19*(1), 95–112.

Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience, 3*, 131–150.

Van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language, 97*(3), 279–293.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33*(4), 485–508.

Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science, 8*(3), 238–242.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*(3917), 392–393.

West, R. F., & Stanovich, K. E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development, 49*, 717–727.

Willems, R. M., Ozyurek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience, 20*(7), 1235–1249.