An MEG Study of Silent Meaning

Liina Pylkkänen and Brian McElree

Abstract

■ Although research on the neural bases of language has made significant progress on how the brain accesses the meanings of words, our understanding of sentence-level semantic composition remains limited. We studied the magnetoencephalography (MEG) responses elicited by expressions whose meanings involved an element not expressed in the syntax, which enabled us to investigate the brain correlates of semantic composition without confounds from syntactic composition. Sentences such as *the author began the book*, which asserts that an activity was begun although no activity is mentioned in the syntax, were contrasted with control sentences such as *the author wrote the book*, which involved no implicit meaning. These conditions were further compared with a semantically anomalous condition (*the author disgusted the book*). MEG responses to the object noun showed that silent meaning and anomaly are associated with distinct effects, silent meaning, but not anomaly, eliciting increased amplitudes in the anterior midline field (AMF) at 350–450 msec. The AMF was generated in ventromedial prefrontal areas, usually implicated for social cognition and theory of mind. Our results raise the possibility that silent meaning interpretation may share mechanisms with these neighboring domains of cognition. ■

INTRODUCTION

The expressive power of human language is practically limitless. In order to explain our ability to effortlessly understand and produce sentences never encountered before, models of language processing must include productive operations that assemble complex meanings from simpler ones. A central goal for the cognitive neuroscience of language is to characterize the neural bases of these operations.

Given the systematicity and productivity of natural language, the meanings of sentences must relate in some fairly transparent way to the meanings of their parts and the syntactic way they are combined. This *principle of compositionality* dates back at least to Frege (1892), and some version of it is generally taken as necessary for any theory of sentence meaning (Fodor & Lepore, 2002; Partee, 1984). However, the degree to which compositionality holds across all possible expressions in a language is a matter of considerable debate. The most strongly compositional theories maintain that each operation of semantic composition corresponds to an operation of syntactic composition (Fodor & Lepore, 2002; Heim & Kratzer, 1998; Montague, 1970). Theories that assume weaker versions of compositionality involve some inventory of semantic rules that do not correspond to operations in the syntax (e.g., Barker, 2002; Jacobson, 1999; Hendriks, 1988; Partee & Rooth, 1983).

Maximally strong compositionality is a good working hypothesis for the syntax–semantics mapping as it offers a straightforward explanation for our ability to understand novel utterances. However, if meanings are strictly compositional, the task of understanding real-time semantic processing is extremely challenging: If syntactic and semantic composition are indeed entirely homomorphic to each other, we cannot experimentally manipulate semantic composition without also manipulating syntactic composition. Given this, it would appear that the only way to study semantic processing without confounding it with syntax would be to focus on those cases where the syntax–semantics homomorphism breaks down, if such cases can be identified.

Demonstrating that an expression has a noncompositional interpretation is nontrivial, given that most syntactic theories include elements that are syntactic but, nevertheless, unpronounced. In this research, we studied the processing of sentences such as *the author began the book*, for which there exists a sizeable theoretical and experimental literature suggesting that these expressions involve exactly the type of syntax–semantics mismatch that should allow isolating processes that are purely semantic (for a review, see Pylkkänen & McElree, 2006). Our study builds on this literature with the aim of characterizing the neural bases of these operations in both time and space.

Manipulating compositionality diverges from most previous investigations of sentence-level semantic processing. The leading paradigm for studying the neural bases of semantic comprehension has been the violation paradigm, where a word fails to meet the semantic expectations set up by a sentence frame, as in *he spread the warm bread with socks*. This type of manipulation systematically elicits

New York University

an increased amplitude in the N400 event-related potential (ERP; Kutas & Hillyard, 1980), which is perhaps the most well established and widely replicated result in the electrophysiology of language processing.

Because the N400 is sensitive to how well a word fits into the preceding context, it has commonly been hypothesized to reflect ''semantic integration'' (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Brown & Hagoort, 1999; Van Berkum, Hagoort, & Brown, 1999; Friederici, 1997; Osterhout & Holcomb, 1992; Rugg, 1990). However, this interpretation is challenged by results showing that the N400 is affected by a wide variety of lexical-level factors that are unlikely to modulate sentence-level semantic composition, such as lexical frequency (Van Petten & Rheinfelder, 1995; Van Petten & Kutas, 1990, 1991; Smith & Halgren, 1987), phonological relatedness (Radeau, Besson, Fonteneau, & Castro, 1998; Praamstra, Meyer, & Levelt, 1994; Praamstra & Stegeman, 1993; Rugg, 1984), and semantic relatedness (Nobre & McCarthy, 1994; Bentin, McCarthy, & Wood, 1985), even when the semantically related prime is masked (Deacon, Hewitt, Yang, & Nagata, 2000; Kiefer & Spitzer, 2000). These effects suggest that the N400 reflects aspects of lexico-semantic activation, not integration. Accordingly, it has been hypothesized that N400 amplitude is determined not by the amount of effort needed to perform semantic integration but rather by the degree to which the target word is preactivated by the context (DeLong, Urbach, & Kutas, 2005; Kutas & Federmeier, 2000; Kutas & Hillyard, 1984). However, what is most important for the purposes of the present research is that a sentential context is not a requirement for N400 effects. This, by itself, precludes the possibility that the N400 indexes semantic composition, viz., the computations by which complex meanings are built from simpler ones by operations such as modification and saturation of predicate–argument relations (see Pylkkänen & McElree, 2006). Apart from the N400, no other neural correlates of semantic composition/integration have been proposed. Thus, there is a gap in our understanding of the neural bases of sentence-level semantic processing.

Manipulating Semantic Composition: Complement Coercion

In order to vary the effort of semantic composition while keeping syntax completely constant, we employed verbs such as *begin* which semantically select for event-denoting objects: *the author began [writing the book]*. With these verbs, explicit mention of the object activity (writing in this case) is optional, as exemplified by the fully grammatical and plausible *the author began the book*. However, although the activity begun by the agent is not mentioned in this sentence, the sentence necessarily asserts that *some* activity involving the book was begun. Consequently, there must be a process not reflected in the pronunciation of the sentence that serves to convert the entity-denoting object NP, *the book*, into an event description. This process has been dubbed *complement coercion*, reflecting the hypothesis that the complement of the verb is ''coerced'' from an entity to an event (e.g., Pustejovsky, 1995). Importantly, complement coercion does not appear to involve a *syntactic* process of converting an NP into a silent VP as that analysis makes the wrong distributional predictions about movement possibilities and adverbial modification (Pylkkänen & McElree, 2006). Thus, complement coercion is a good case for a counterexample to strong compositionality because it seems truly invisible to syntax.

A series of psycholinguistic studies, using reading time and related measures, have shown that although expressions involving complement coercion are fully grammatical and easy to understand, they are associated with a clearly measurable processing cost (McElree, Frisson, & Pickering, 2006; McElree, Pylkkänen, Pickering, & Traxler, 2006; Pickering, McElree, & Traxler, 2005; Traxler, McElree, Williams, & Pickering, 2005; Traxler, Pickering, & McElree, 2002; McElree, Traxler, Pickering, Seely, & Jackendoff, 2001; for a review, see Pylkkänen & McElree, 2006). Importantly, these studies have systematically tested a large number of hypotheses pertaining to the source of the coercion cost, and collectively offer a detailed understanding of the processes driving the effect. The combined results indicate that what is costly in complement coercion is the operation of adjusting the event structure of the complement, rather than a host of other plausible sources. For example, the cost does not reflect a general difficulty in combining verbs such as *begin* with NP complements because the effect occurs also when the verb and the syntactic category of its complement are kept constant, and only the denotation of the complement is varied: *begin the puzzle* (entity NP) versus *begin the fight* (event NP) (Traxler et al., 2002). Further, the effect is found even when the aspectual properties of the coercion and control stimuli are entirely controlled for (*the author began the article* vs. *the author began writing the article*), suggesting that the coercion cost is not due to the telicity asymmetry between sentences such as *the author began the article* and *the author wrote the article* (Pickering et al., 2005).

Another possible explanation of the coercion cost is that it is due to the inherent ambiguity of expressions such as *begin the book*, where the nature of the initiated activity is left unspecified. However, complement coercion is costly even when the context disambiguates the nature of the implicit activity (Traxler et al., 2005), suggesting that ambiguity does not drive the effect. Further, when the degree of ambiguity of coerced expressions is manipulated directly, reading times are unaffected. This was shown by Frisson and McElree (in press), who contrasted expressions such as *the student finished the essay*, involving a single dominant interpretation for the implicit activity (i.e., ''writing''), with sentences such as *the director started the script*, where the implicit event could be interpreted as a number of activities, including shooting, reading, writing, editing, and reviewing. Although both types of expressions elicited a coercion cost, the size of the effect was unaffected by the degree of ambiguity. The effect size was also not correlated with either dominance or the number of alternative interpretations. These data strongly suggest that the difficulty of complement coercion is not due to the effort involved in retrieving a suitable activity, but rather to the process of constructing an event predicate from the entity-denoting NP. In other words, upon encountering *the book* as the complement of *begin*, comprehenders must posit an abstract activity, relate *the book* to that activity, and then construe the activity as the complement of the event-selecting verb.[1]

Given that the behavioral effect of complement coercion is well understood and that the operations involved are unlikely to be syntactic in nature, this construction offers an ideal starting point for a neurolinguistic investigation of semantic composition. Here we employed magnetoencephalography (MEG) to characterize the neural correlates of complement coercion. We contrasted expressions such as *the author began the article* with simple transitive sentences such as *the author wrote the article*, involving verbs that directly selected for an entity complement. In order to match the coerced and simple sentences maximally in meaning, the verbs in the simple condition were chosen to reflect the interpretation comprehenders most commonly assign to the implicit activity in the coerced versions (as assessed by a fill-in-the-blank test; see Materials). This resulted in a difference in predictability between the simple and the coerced conditions, with the complements of the simple transitives being more predictable than the complements of coerced verbs (see cloze norms in Materials). Predictability (as measured by cloze probability) cannot account for the coercion effect, as the effect is found even when predictability is matched (see Traxler et al., 2002). Here, we assessed the effect of predictability separately by including an entirely unpredictable control condition, involving anomaly. We expected that sources sensitive to predictability should show the largest effect for the anomalous complements and the smallest for the control stimuli, with the coerced condition patterning in the middle. Sources sensitive to coercion, on the other hand, should show the largest effect for the coerced complements, and no effect for the anomalous stimuli.

The anomalous condition also allowed us to relate our investigation to existing sentence-level studies of semantic processing, given that anomaly notoriously elicits an N400. Previous MEG studies investigating the neural generators of the N400m have found effects predominantly in the left superior temporal areas (Halgren et al., 2002; Helenius, Salmelin, Service, & Connolly, 1998; Simos, Basile & Papanicolaou, 1997), although distributed source localization methods and intracranial recordings have implicated a much larger network of regions (e.g., Halgren et al., 1994, 2002; Guillem, Rougier, & Claverie, 1999).

Based on the behavioral literature reviewed above, we expected coercion to affect the processing of the head noun, for example, *article* in sentences such as *the author began the article*. However, given that our stimuli did not include any items where the direct object was an event-denoting noun (as in *begin the fight*), it was conceivable that subjects might have predicted the necessity of coercion before encountering the noun, possibly at the determiner *the*. To determine whether this was the case, we also performed a full analysis of the MEG signals elicited by the determiner.

As regards the suitability of complement coercion for a neurolinguistic study of semantic composition, a caveat should be kept in mind. It is possible that the computation of implicit meaning is qualitatively different from the computation of meaning that has a transparent phonological expression. Thus, a neural correlate of complement coercion might not reflect activity that is related to semantic composition in general. However, as noted, the investigation of transparent meaning composition faces the methodological challenge of separating syntactic operations from semantic operations. Consequently, we take noncompositional interpretation to be a more promising starting point for a neural investigation. If a neural correlate of implicit meaning computation is identified, it then becomes possible to investigate to what extent the identified area is also active in transparent interpretation.

METHODS

Participants

Seventeen native English speakers participated in the study (10 women). All were graduate or undergraduate students at New York University (NYU, ages 20–32 years). The data from one subject were excluded from the analysis because of excessive noise artifacts.

Materials

We tested 70 triplets like (1)–(3), which we combined with 135 filler sentences with different syntactic structures. To further diversify the materials, we embedded 35 of the triplets as complements of clause-selecting verbs (see Appendix).

(1) Coerced: The journalist began the **article** after his coffee break.
(2) Anomalous: The journalist astonished the **article** after his coffee break.
(3) Control: The journalist wrote the **article** after his coffee break.

Anomalous variants of the coerced sentences were generated by replacing the coercing verbs with verbs such as *astonish*, *amuse*, and so forth, which require direct objects that denote an experiencer of the psychological state expressed by the verb, a requirement which was not fulfilled by an inanimate entity such as *article*, *dress*, and so forth. Each subject saw all three versions of each stimulus in a pseudorandom order so that the effect of repetition at the sentence was equal in all conditions.

It was important that our control sentences (3) employed a verb that expressed the eventive interpretation readers most often ascribed to the coerced sentences (1). Following other studies on complement coercion (e.g., McElree, Frisson, et al., 2006; McElree, Pylkkänen, et al., 2006; Traxler et al., 2002, 2005; McElree et al., 2001), an additional 24 NYU participants provided one- to two-word fill-in-the-blank responses indicating how they would interpret a coerced sentence such as ''The journalist began ______ the article.'' From a large pool of candidate sentences, we selected 70 in which there was a dominant response that could serve as a control verb. For example, *wrote* occurred as a response to the example sentence above 65% of the time, *read* occurred 26% of the time, with the remainder consisting of more idiosyncratic responses such as *penned* or *reviewed*. On average, the chosen control verb occurred as a fill-in-the-blank response 74% of the time (SD = 14.6%).

As discussed in the Introduction, choosing the most natural verbs for the control stimuli was likely to result in a difference in predictability between the coerced and control conditions. This was confirmed by a cloze probability norm in which 32 NYU students wrote sentence completions for the coerced and control versions of the stimuli. The participants were presented with the experimental sentences up to and including the determiner following the verb (e.g., The author began the ____, The author wrote the ____). The materials were divided into two lists in such a way that no participant saw both the coerced and control versions of the same stimulus. Participants' responses were compared with the actual experimental sentences to assess the predictability of the target words. Cloze proportions (proportion of responses completed with the target words) for the coerced and control conditions were 0.08 and 0.22, respectively, a reliable difference as revealed by a one-way analysis of variance with condition as a within-items factor [$F(1, 69) = 27.8$, $p < .0001$]. However, as explained above, our approach was not to attempt to match the coerced and simple conditions in predictability, which would have been very difficult to do, but rather to assess the effect of predictability with a separate control, the anomalous condition.

Another group of 20 NYU students rated on a 7-point scale the plausibility (7 = highly plausible) of the coerced and control sentences to ensure both were highly plausible. Raters judged how likely they believed the events described by the sentence were (e.g., Pickering & Traxler, 1998). The coerced and control sentences were split into two sets (so that no rater saw both the coerced and control pairs), mixed with various implausible sentences, randomized, and presented to two groups of 10 raters. Mean plausibility ratings were 6.43 for the coerced sentences and 6.70 for the control sentences. The plausibility ratings were significantly higher for the control sentences, $t(68) = 4.1$, $p = .0001$, but importantly for our purposes, both sentence forms were rated as highly plausible.

Our selection procedure resulted in verbs in the coerced condition being used more often than the verbs in the anomalous and control conditions (see Appendix). Although not ideal, we note that greater repetition of a verb should facilitate its processing. Thus, if differences in exposure to the respective verbs in an experimental session have any effect on processing, we believe it would be to reduce the cost of coercion, lessening our chances of detecting neural correlates for the coercion cost.[2]

Additionally, we have collected processing time measures on a subset of these triplets. Forty-five of the 70 triplets were used as materials in a speed–accuracy tradeoff study (McElree, Pylkkänen, et al., 2006), which served to contrast the full time course of processing coerced and control sentences. Time-course measures demonstrated that the coerced expressions were interpreted less accurately and more slowly than the minimally contrasting control expressions. These differences are consistent with the claim that expressions such as (1) engender more taxing compositional operations than control expressions such as (3).

Procedure

During the experiment, the participants lay in a dimly lit magnetically shielded room and viewed the experimental stimuli via fiber-optic goggles (Avotec, Stuart, FL). Each trial started with a fixation point in the middle of the screen. Subjects initiated each trial themselves by pressing a button. The sentences were presented in nonproportional Courier font (font size = 90), word by word (300 msec on, 300 msec off). At the end of the sentence, a question mark was presented. At the question mark, subjects were instructed to judge whether the sentence made sense or not.

Neuromagnetic fields were recorded with a whole-head, 148-channel neuromagnetometer array (4-D Neuroimaging, Magnes WH 2500) at a sampling rate of 678 Hz in a band between 0.1 and 200 Hz. At the end of the recording session, an auditory baseline test was conducted, during which the participants listened to one hundred 1-kHz tones. Although we did not have MRIs for our subjects, source locations of the auditory M100 allowed us to visualize activity with respect to a functional landmark representing the primary auditory cortex. The entire recording session lasted approximately 1 hour.

MEG Data Analysis

MEG data were cleaned of artifacts and then averaged according to stimulus category. On average, 5% of each subject's data were excluded due to artifacts (SD = 0.8%). Prior to source modeling, MEG averages were high-pass filtered at 1 Hz and low-pass filtered at 40 Hz.

Two multiple source models (BESA, Brain Electric Source Analysis, 5.0) were created for each subject, one for the activity elicited by the determiner *the* and another for the activity associated with the noun. Although there is now a sizeable literature on the MEG responses elicited by visual open class words, there is no such body of work addressing the processing of function words. In order to be able to relate the source models obtained for the determiner to those of the noun, we first describe the modeling of the MEG signals associated with the noun.

Source Modeling of the Noun Activity

Because our goal was to identify activity associated with coercion, we aimed to characterize all activity associated with the processing of the critical noun. Sources were fit at the peaks of prominent response components using both minimum norm estimates (MNEs) and magnetic field patterns to constrain the number of dipoles. All sensors were used in localization. Only dipoles whose location and orientation were consistent with the magnetic field patterns were accepted.

Visual inspection of the MEG signals associated with the noun at 0–350 msec revealed a similar pattern of magnetic fields across conditions and subjects, familiar from previous MEG language studies in the visual modality (Pylkkänen, Llinas, & Murphy, 2006; Fiorentino & Poeppel, 2004; Pylkkänen, Feintuch, Hopkins, & Marantz, 2004; Stockall, Stringfellow, & Marantz, 2004; Pylkkänen, Stringfellow, & Marantz, 2002; Embick, Hackl, Schaeffer, Kelepir, & Marantz, 2001; Pylkkänen, Stringfellow, Flagg, & Marantz, 2001; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999; Helenius et al., 1998). Early visual responses at ~100 msec and ~150–200 msec were followed by two major response components in the temporal regions, one at around 250 msec (M250) and the second around 350 msec (M350). Because the early visual activity did not appear to vary by condition, these sources were modeled from each individual's grand-averaged data across conditions. M350 field patterns, on the other hand, were the clearest for the anomalous sentences, and therefore, the anomalous data were used in M350 source modeling. In addition to these early- and mid-latency response components, all subjects showed an anterior midline field (AMF) at 350–500 msec, which was particularly prominent in the coercion condition, and was therefore localized on the basis of the coercion condition. This component localized consistently to anterior, inferior, midline areas. Although not part of the usual complex of language-related MEG responses, anterior midline areas have been found to be activated by linguistic stimuli in a few previous MEG studies as well (Pylkkänen et al., 2006; Marinkovic et al., 2003; Halgren et al., 2002). After the AMF, there was little consistency across subjects in the late activity. After modeling the neural generators of all the major response components, the component-specific dipole models were introduced into a single multidipole model, which was kept constant across conditions.

Source Modeling of the Determiner Activity

Visual inspection of the activity elicited by the determiner revealed no consistent condition-specific field patterns, and therefore, the grand-average of each individual's data across conditions was used in source modeling. Early visual responses elicited by the determiner were very similar to those associated with the noun. However, activity at 200–300 msec revealed a source cluster in left anterior temporal regions that was not present in the noun data. After 300 msec, magnetic field patterns varied considerably across subjects, making it difficult to form meaningful groupings of sources for statistical analysis (for details, see Results/Multiple Source Models).

RESULTS

Sensicality Judgment Data

The end of sentence sensicality judgment data are summarized in Table 1. Anomalous sentences elicited faster and more accurate responses than both coerced and control sentences [$F(2, 15) = 82.243$, $p < .001$ and $F(2, 15) = 3.41$, $p < .05$, respectively]. However, the higher accuracy rates for anomalous sentences are uninformative, as these sentences required an ''unacceptable'' response, whereas coerced and control sentences both required an ''acceptable'' response. Additionally, the speed advantage is unsurprising, as subjects were able to determine the nonsensicality of the anomalous stimuli mid sentence. What is most important for our purposes is that the coerced and control sentences did not differ in sensicality judgment accuracy, suggesting that the subjects found them equally plausible. This accords with the ratings of plausibility collected from a separate group of participants.

Table 1. Off-line Sensicality Judgment Data

	Coerced	*Implausible*	*Control*
RT (msec)	791	544	799
Accuracy (%)	86	93	86

Multiple Source Models

Figure 1 depicts all source localizations both for the determiner and the noun broken down by time window and hemisphere. Source models were kept constant across experimental conditions. Mean goodness of fit (GOF) was around 80% both for the determiner and the noun solutions, and it did not vary as a function of stimulus category for either. The multiple source models for the determiner explained, on average, 82% (*SD* = 5%) of the activity elicited in the coercion condition at 0–400 msec poststimulus onset and 82% (*SD* = 6%) and 81% (*SD* = 6%) of the activity in the anomalous and control conditions, respectively. Similar GOFs were obtained for the noun models at a slightly longer interval, lasting until 600 msec poststimulus onset (coercion: 82%, *SD* = 5%; anomalous: 81%, *SD* = 5.2%; control: 81%, *SD* = 4.6%). As above, we first report the results of multidipole modeling for the noun, and then the modeling for the determiner.

Activity Elicited by the Noun

Consistent with previous MEG findings (Tarkiainen et al., 1999), our visually presented nouns elicited a clear visual M100 response at 90–150 msec, followed by bilateral M170 activity at 150–230 msec. The visual M100 is associated with a right-lateral outgoing field and a left-lateral re-entering field over the occipital sensors. Visual M100 activity was elicited in all subjects and was best explained by a single midline dipole in occipital areas. In 14 subjects' data, the visual M100 was followed by the M170, which exhibits a polarity reversal as compared to the M100. The M170 was best explained by a bilateral two-dipole solution in 11 subjects and by a single right-lateralized dipole in 3 subjects.

After early visual responses, activity moved to the temporal lobes. Consistent with previous findings (Pylkkänen et al., 2006), localization of temporal activity at 200–300 msec (M250) is less consistent across subjects than localization of activity before or after this time window. All subjects but one showed a clear peak at 200–300 msec that was distinct from the M170 and from later M350 activity. This mid-latency activity localized to temporal areas, bilaterally in six subjects, in the left hemisphere only in three subjects, and in the right hemisphere only in six subjects.

At 300–400 msec, activity associated with visual words typically localizes in temporal areas, close to the pri-

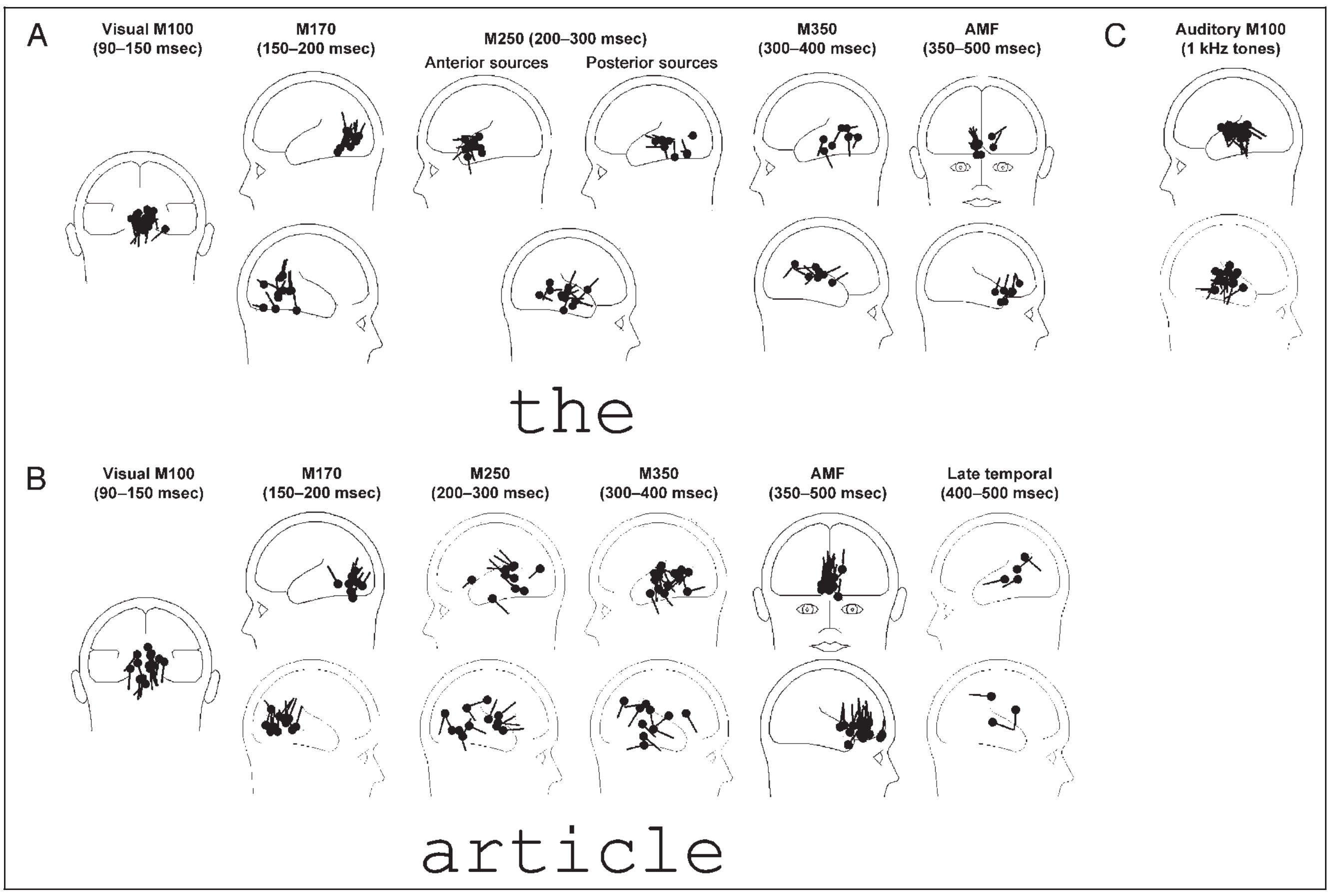

Figure 1. Results of multisource modeling for all 16 subjects, broken down into the groups that were entered into statistical analysis. Dipoles are plotted inside one average-sized participant's spherical head model. Because MRIs were not available for our subjects, we collected auditory M100 data at the end of the experiment to serve as a functional landmark representing the primary auditory cortex (see Methods).

mary auditory cortex, either bilaterally or in the left hemisphere only (Pylkkänen et al., 2002, 2004, 2006; Halgren et al., 2002; Helenius et al., 1998). The left hemisphere source, the M350, varies in latency and/or amplitude with stimulus factors affecting lexical access and has consequently been interpreted as a neural index of lexical activation (Pylkkänen et al., 2001, 2002, 2004, 2006; Beretta, Fiorentino, & Poeppel, 2005; Fiorentino & Poeppel, 2004; Stockall et al., 2004; Pylkkänen & Marantz, 2003; Embick et al., 2001). A clear left hemisphere M350 field pattern was identified in 13 of our 16 subjects, and this activity localized in left superior and middle temporal areas. For nine of these subjects, the M350 localized in a physiologically plausible way only if a second dipole was added to the model, and this second dipole, the ''M350-R,'' always localized in right temporal areas. In three subjects' data, activity at 300–400 msec localized midline. These sources could not be grouped together with the M350 or the M350-R and were therefore excluded from statistical analysis. Five subjects showed further temporal activity after the M350.

At 350–500 msec, all subjects showed activity that localized in anterior inferior areas, close to midline. This activity, dubbed the ''anterior midline field'' (AMF), was particularly clear for the coerced stimuli and was associated with a right hemisphere outgoing field and a left hemisphere re-entering field. The AMF peaked, on average, 24 msec after the left hemisphere M350, but in general, the AMF and the bilateral temporal area tended to be concurrently active, as illustrated for single-subject data in Figure 2.

Because frontal activity is notoriously sensitive to eye artifacts, we carried out an additional analysis to verify that the AMF field was not generated by some very subtle eye movement that may have survived our artifact rejection procedure. All eye artifacts are optimally modeled by a two-dipole solution with one dipole in each eye (Berg & Scherg, 1991). If the AMF was generated in the eyes instead of in ventromedial prefrontal areas of the brain, the AMF field pattern should be better explained by two dipoles placed in the eyes as opposed to a single dipole slightly above the eyes. We tested this by placing dipoles in each subjects' eyes and compared the GOF of this solution to the actual AMF dipole of that subject. For simplicity, the comparison was carried out between the two-dipole eye-solution and the one-dipole midline solution without incorporating these dipoles into the multidipole solution of the subject. The eye dipoles were fixed to the eyes in location but were allowed to orient themselves optimally, which gave them the best possible chance to explain the data. Regardless of this, the GOF of the eye dipoles was worse than the GOF of the midline dipole for all of our 16 subjects [$t(15) = -5.68$, $p < .001$], showing that the AMF is unlikely to have an ocular source.

Activity Elicited by the Determiner

The MEG signals associated with the determiner showed a clear visual M100 response in 15 out of our 16 subjects, and this response was optimally modeled by a single midline dipole in occipital areas in all cases. The M170 field pattern was observed in all subjects' data and it was best explained by a bilateral two-dipole solution in eight subjects, a single right-lateral dipole in three subjects, a single left-lateral dipole in two subjects, and a midline dipole in three subjects. The midline dipoles were not entered into statistical analysis as they were likely to represent the combined activity of the left and right hemisphere generators of the M170, which are known to be functionally distinct (Tarkiainen, Cornelissen, & Salmelin, 2002; Tarkiainen et al., 1999).

After the early visual responses, activity moved to the temporal lobes bilaterally. In the left hemisphere, 10 subjects showed a temporal source anterior to midline at 200–300 msec. In statistical analyses, these sources were grouped together as the ''M250 anterior.'' This activity may be related to the left anterior N280 elicited for closed-class words in ERP studies (Brown, Hagoort, & ter Keurs, 1999; Osterhout, Bersick, & McKinnon, 1997;

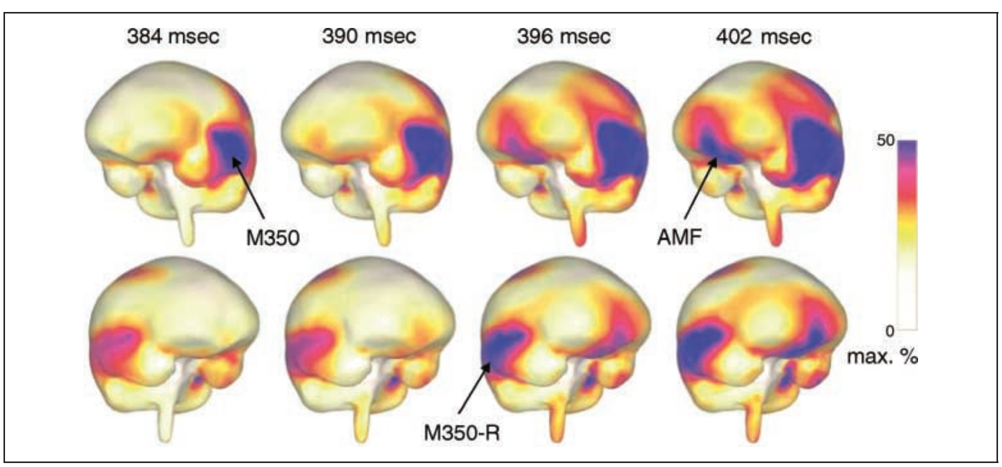

Figure 2. Minimum norm estimates of single-subject data illustrating the progression of activation from the temporal lobes bilaterally to the AMF generator in ventromedial prefrontal areas.

King & Kutas, 1995; Neville, Mills, & Lawson, 1992). Temporal sources posterior to midline (n = 9) were labeled ''M250 posterior'' and were treated separately. Right hemisphere temporal sources at 200–300 msec did not form distinct clusters in the same way and were therefore treated as a single dependent measure.

After the M250 time window, there was substantial variability in the observed magnetic fields. Nine subjects did, however, exhibit anterior midline activity resembling the AMF at ~350 msec, and this activity was included in our analyses. Further, seven subjects showed an M350-like field in the left hemisphere and seven in the right; these two source clusters were also included as dependent measures. The combination of the hypotheses that the left hemisphere M350 generator performs the access of open-class lexical items (Pylkkänen & Marantz, 2003) and that function words are stored separately from open-class words (e.g., Bradley & Garrett, 1983; Bradley, Garrett, & Zurif, 1980) would predict that the M350 should not be observed for function words. Because fewer than half of our subjects showed field patterns at 300–400 msec that were like those observed for content words, our data are indeed broadly consistent with a neural separation of content and function words.

Frontal Effect of Coercion, Temporal Effects of Anomaly

Table 2 summarizes mean peak latencies and amplitudes for all sources associated with processing of the determiner and the noun. Our stimulus manipulation had no reliable effects on the activity associated with the determiner, consistent with the hypothesis that coercion does not occur until the noun denotation is available. Activity elicited by the noun was, however, affected both by coercion and by anomaly. As shown in Figure 3, the stimulus manipulation had a reliable effect on the AMF source amplitudes of the noun [$F(2, 15) = 4.5, p < .05$], with coerced sentences eliciting larger AMF amplitudes than control sentences ($p < .05$, Scheffe). Anomalous sentences did not reliably differ from control sentences in AMF amplitude ($p = .6$, Scheffe). Thus, the AMF effect was specific to coercion.

Table 2. Mean Source Latencies and Amplitudes in All Conditions

	Latency (msec)			*Amplitude (nA)*		
	Coerced	*Anomalous*	*Control*	*Coerced*	*Anomalous*	*Control*
A. ''the''						
Visual M100	123	124	120	29	29	28
M170-L	165	172	177	7	8	8
M170-R	222	195	219	17	14	16
M250-L ant.	243	244	244	20	19	20
M250-L post.	254	251	250	11	11	13
M250-R	247	249	247	14	14	13
M350-L	353	367	349	14	13	16
M350-R	349	333	348	13	9	10
AMF	351	357	345	19	20	20
B. ''article''						
Visual M100	137	134	136	42	43	42
M170-L	193	193	195	16	17	17
M170-R	192	196	202	25	23	21
M250-L	263	276	270	20	16	19
M250-R	264	262	262	12	11	12
M350-L	368	382	363	14	17	12
M350-R	370	379	353	16	19	15
AMF	394	408	384	30	25	22

Statistically reliable effects are shaded.

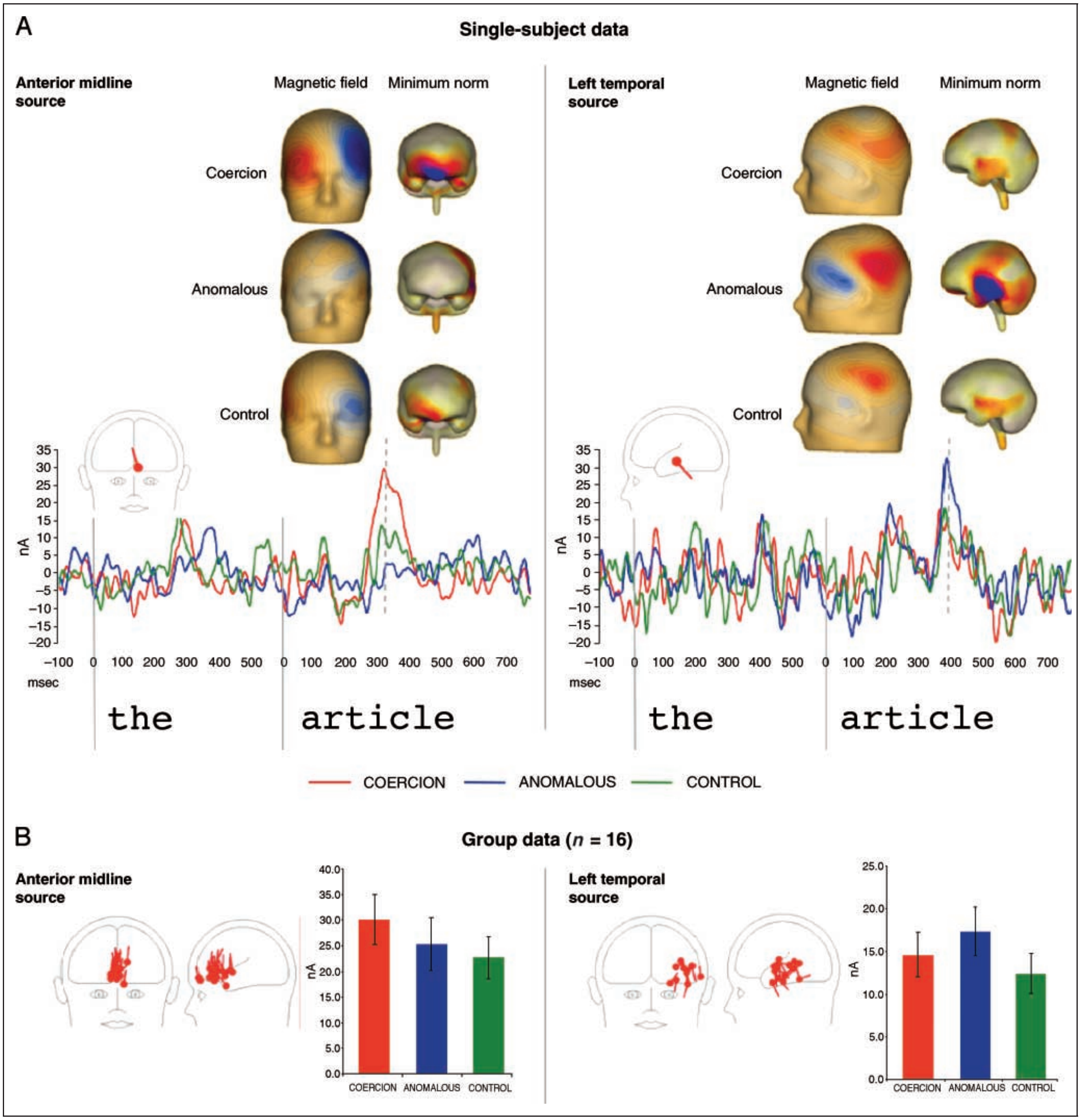

Figure 3. Effect of the stimulus manipulation on the source amplitudes of the anterior midline field (AMF) and the left hemisphere M350 in single-subject (A) and group data (B). Magnetic fields and minimum norm estimates are plotted for the AMF and M350 peaks indicated by the dotted line.

Although the AMF showed no effect of anomaly, the M350 did (Figure 3). The effect of the stimulus manipulation on left hemisphere M350 amplitudes was reliable [$F(2, 15) = 4.5, p < .05$], with anomalous sentences eliciting larger M350 amplitudes than control stimuli ($p < .05$, Scheffe). Anomalous stimuli also elicited longer M350 latencies than controls [$F(2, 12) = 5.49, p < .05$], likely a secondary effect of the M350 amplitude effect. These results conform to previous MEG findings by Halgren et al. (2002), Helenius et al. (1998), and Simos et al. (1997), which localized the classic N400 effect mainly to superior temporal areas. The N400m identified in these studies and the M350 have similar magnetic fields and localizations. Thus, the N400m and the M350 are likely to have the same primary generator.

Right hemisphere activity in the M350 time window, the "M350-R," also tended toward larger amplitudes for anomalous stimuli, but this effect did not reach significance. An effect was, however, obtained in M350-R latency [$F(2, 9) = 3.6$, $p < .05$], with anomalous stimuli eliciting longer latencies than controls ($p < .05$, Scheffe). Coerced and control stimuli did not reliably differ with respect to the M350-R. The functional role of the M350-R is unclear, although it has been previously reported to be sensitive to lexical–semantic factors (Pylkkänen et al., 2006).

Because M350 amplitudes and peak latencies were increased for anomalous stimuli as compared to controls, one might expect subsequent activity to be delayed as well, as a secondary effect of the M350 delay. Indeed, AMF amplitudes were delayed for anomalous sentences [$F(2, 15) = 4.9$, $p < .05$].

In addition to the AMF and mid-latency temporal effects, left hemisphere M250 source amplitudes were affected by the stimulus manipulation [$F(2, 8) = 3.8$, $p < .05$], anomalous stimuli eliciting smaller amplitudes than coerced and control stimuli. Importantly, coerced stimuli did not reliably differ from controls. Thus, this unpredicted effect does not challenge the conclusion that the AMF is the first coercion-sensitive neural source. Further, the M250 localizations do not exhibit sufficient consistency across subjects to warrant any strong conclusions (see Figure 1).

No effects of the stimulus manipulation were obtained in the latencies or amplitudes of right hemisphere M250 activity, or in the visual M100 or the M170. Post-M350 temporal activities were not entered into statistical analysis, due to a very small number of data points.

In sum, we found that the AMF is specifically sensitive to coercion and the left hemisphere M350 to anomaly.

In order to further evaluate the robustness of the effect of coercion on AMF amplitudes, we tested whether the AMF effect would be replicated with a maximally simple single-dipole spatial filter that is kept constant not only across conditions but also across subjects. We planted in each subject's data a single dipole representing the mean location and orientation of all the AMF dipoles across our 16 subjects (Figure 4). The source waves of these dipoles were entered into a point-by-point t test, which showed reliably larger amplitudes for the coercion condition starting at 379 msec and lasting until 400 msec ($p < .05$). Thus, although the AMF source was often concurrently active with the M350/M350-R, which did *not* show sensitivity to coercion, the coercion effect was replicated even though no other sources apart from the AMF dipole were included in the model.

Minimum Norm Estimates

Given that our coercion effect localized in an area outside the fronto-temporal network traditionally thought to be involved in language processing, we sought to further assess the validity of our AMF localization with minimum norm estimates (MNE). MNEs provide an estimation of current density across a large number of sources evenly distributed across the brain surface (Hämäläinen & Ilmoniemi, 1984). Compared to discrete source models, distributed source models such as the MNE have the advantage that they require less user intervention (e.g., specification of fit intervals or number of sources). Rather, side assumptions are implicitly incorporated in the method (here, e.g., a minimum L2-norm of the final current distribution is assumed). Further, unlike discrete source models, which are most accurate for focal sources, MNEs are also suitable for the representation of distributed sources (Uutela, Hämäläinen, & Somersalo, 1999).

The minimum norm images were calculated in BESA 5.1. Each MNE was based on the activity of 1426 regional sources evenly distributed in two shells 10% and 30% below a smoothed standard brain surface. Regional sources in MEG can be regarded as sources with two single dipoles at the same location but with orthogonal orientations. The total activity of each regional source was computed as the root mean square of the source activities of its two components. The minimum norm images were depth weighted as well as spatio-temporally weighted, using a signal subspace correlation measure introduced by Mosher and Leahy (1998). The minimum total power of the current distribution (the minimum L2

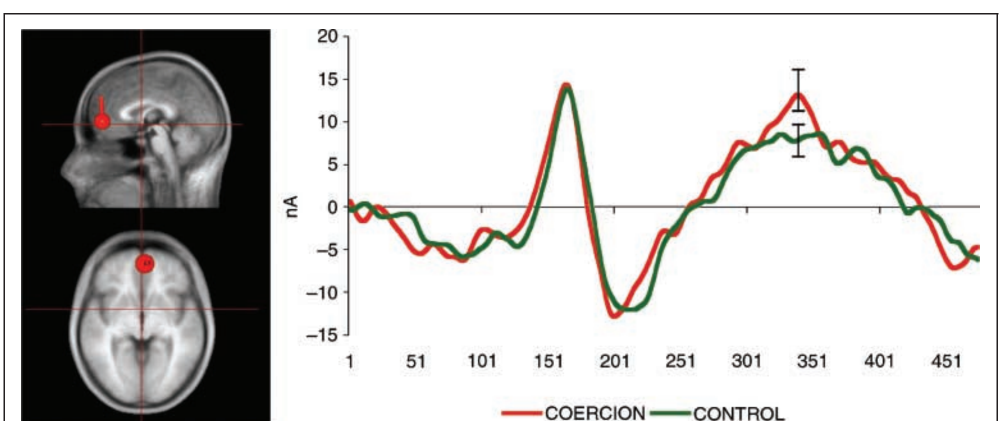

Figure 4. Mean AMF dipole plotted inside a standard MRI (left) and the grand-averaged waveform of this source across subjects (right). The source models included only the mean AMF dipole (x-loc, y-loc, z-loc = 0.12, 48.9, 36.3; x-ori, y-ori, z-ori = −0.13, 0.12, 0.98; the x-axis runs from left to right, the y-axis from posterior to anterior, and the z-axis from inferior to superior).

norm) was then compared sample by sample between the coercion and control conditions, as well as the anomalous and control conditions, for completeness. Each observation (i.e., a source amplitude at a time point) was included in a point-by-point *t* test, but a difference between conditions was considered reliable only when it remained significant ($p < .01$) for a period of time and was not limited to a single point in space. Specifically, a difference at a source was considered reliable only if it remained significant for at least 20 msec (15 samples) and if 15 of its closest neighbors also showed a significant effect. Thus, each significant effect represented the center of a spatio-temporal neighborhood showing a reliable difference between conditions.

Figure 5 plots all sources that met our significance criteria. For the coercion versus control comparison, our dipole modeling results were replicated straightforwardly: At around 400 msec source amplitude was reliably larger for the coerced stimuli in a cluster of sources in the anterior inferior midline regions of the frontal lobe (Figure 5A), consistent with our AMF localization. This anterior source cluster showed no effect of anomaly. Instead, anomalous stimuli elicited increased amplitudes in three other regions. At ~400 msec, anomalous targets were associated with larger amplitudes in medial structures of the left anterior temporal lobe. Slightly later, around 450 msec, increased amplitudes were observed for anomaly around the inferior part of the superior frontal gyrus. Finally, at around 470 msec, anomalous stimuli elicited increased activation in left inferior frontal regions. Overall, these source modeling results are broadly compatible with the MEG results obtained by Halgren et al. (2002), who also found semantic anomaly to affect a distributed cortical network. However, our distributed source solutions revealed no effect of anomaly in left superior or middle temporal region, contrary to our dipole modeling results as well as the distributed source results of Halgren et al. A likely reason for the discrepancy between our multidipole and distributed source analyses is the relatively large between-subjects variance that was obtained in the dipole localization of the M350/N400m field: Although dipole orientation stayed relatively constant across subjects, dipole location ranged from anterior parts of the

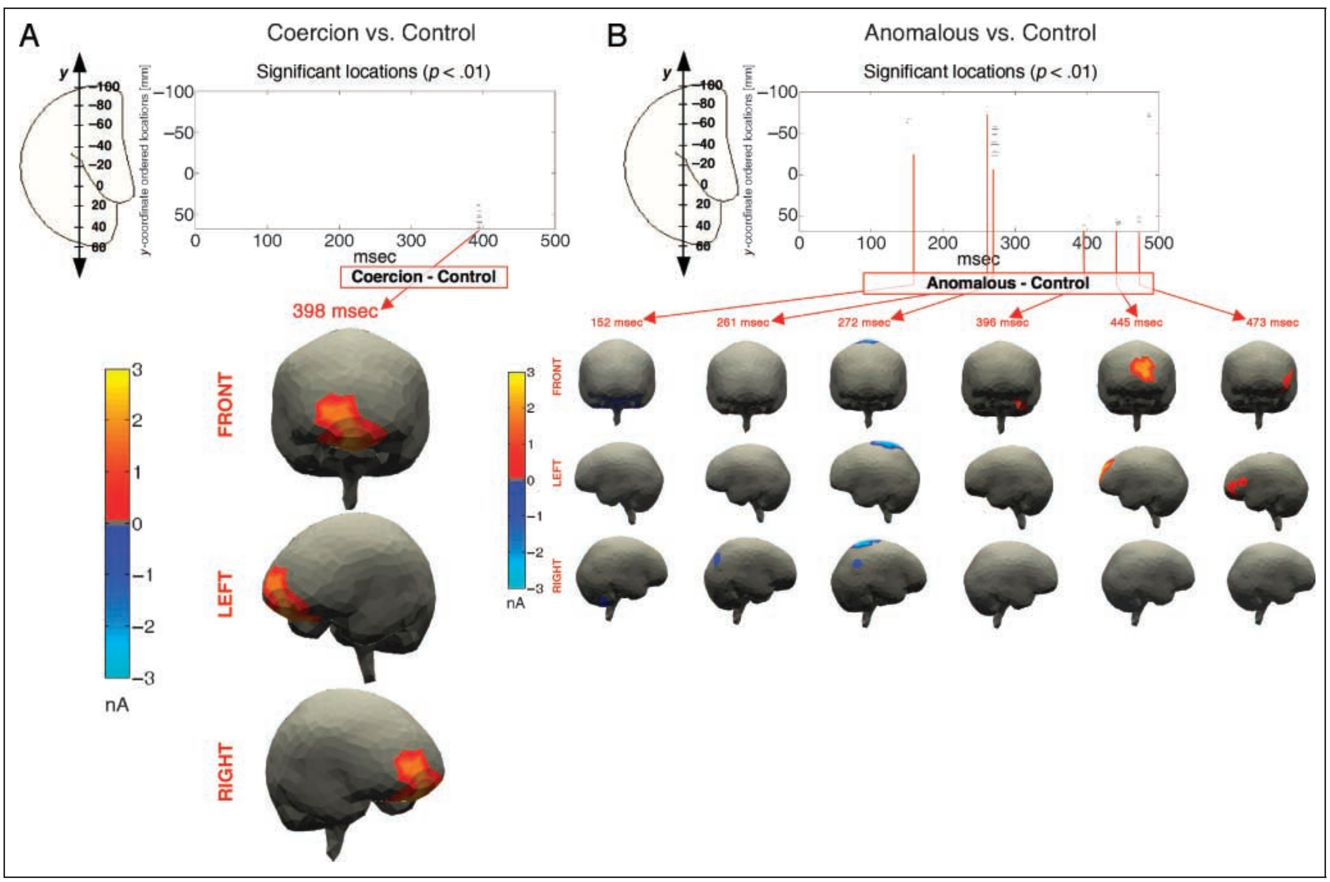

Figure 5. Minimum norm estimates of the activity elicited by the object noun (''book'') for the Coercion vs. Control (A) and the Anomaly vs. Control (B) contrasts. In the top panel 2-D plots, the *x*-axis represents time and the *y*-axis plots all sources on the smoothed brain surface, ordered according to the *y*-coordinate (running from the front of the brain to the back). In the 2-D plot, the centers of spatiotemporal neighborhoods where the conditions differed reliably (see text) are plotted in black. The bottom panel displays the spatial distributions of the reliable differences on the smoothed brain surface. Front, left, and right views are shown for each difference. As in the dipole analysis, coercion, but not anomaly, elicited increased activity in a cluster of sources in inferior anterior midline regions.

left middle temporal gyrus to posterior parts of the left superior temporal gyrus. Thus, there may have been too much variance in M350 source location for the amplitude effect obtained in the dipole modeling to survive the across-subjects grand-averaging of the distributed source analysis.

In addition to the amplitude increases at frontal sites at 400–500 msec, anomalous and control stimuli diverged slightly but reliably at ∼150, ∼260, and ∼270 msec at cerebellar, left posterior parietal, and midline fronto-parietal sources, respectively. The directionality of these early effects was, however, opposite from those obtained in the later time widow, that is, they exhibited larger amplitudes for the control than for the anomalous condition. These effects cannot be interpreted as reflecting increased processing effort due to anomaly, and thus, we refrain from speculating about their role. The important conclusion from the distributed source analysis is that the MNEs confirmed both the localization of the AMF as well as its specificity to coercion.

DISCUSSION

Coercion, Compositionality, and the AMF

We investigated the neural bases of sentence-level semantic interpretation by manipulating complement coercion, a composition process that has been argued to be purely semantic (Pylkkänen & McElree, 2006). Coercion did not modulate activity in any ''traditional'' language area, but rather elicited increased amplitudes in the AMF generated by a midline source in the ventromedial prefrontal cortex (vmPFC).

Importantly, the AMF showed no sensitivity to anomaly, which rules out several interpretations of the effect in terms of factors other than coercion. These include plausibility, predictability, and semantic relatedness, as the anomalous complements were less plausible, less predictable, and less semantically related to their contexts than the coerced complements. As a measure of semantic relatedness, we used latent semantic analysis (LSA), which showed that the co-occurrence of the coercion verbs and the critical nouns was significantly higher (cosine = 0.13) than the co-occurrence of the anomalous verbs and the noun [cosine = 0.07; $t(138) = -3.48$, $p < .001$]. Whether the AMF effect could be due to the ambiguity of coerced expressions was not directly addressed in our study, but given that the behavioral effect of coercion is not ambiguity-related (see Introduction), it is unlikely that the corresponding neural effect would be either. Nevertheless, whether the AMF is, in fact, sensitive to ambiguity is an important question for future studies.

The generator of the AMF, the vmPFC, is not part of extant models of the neurobiology of language. Thus, our findings suggest that syntactically unexpressed meaning is not computed within the network of what are generally considered uncontroversial ''language areas,'' such as the left superior temporal gyrus, middle temporal gyrus, the left temporal pole, and left inferior frontal areas (for a review, see e.g., Stowe, Haverkort, & Zwarts, 2005). Interestingly, however, the vmPFC has been shown to be activated by nonliteral interpretation, such as sarcasm and irony (Shamay-Tsoory et al., 2005). Although nonliteral meaning is, in many ways, different from the type of semantic mismatch phenomena that complement coercion exemplifies, the comprehension of both coercion and sarcasm requires generating meaning that is not part of the overt linguistic input. Outside language, vmPFC damage consistently results in impaired social cognition and, in particular, theory of mind (Rowe, Bullock, Polkey, & Morris, 2001). Theory of mind is necessary for understanding irony (Sullivan, Winner, & Hopfied, 1995); whether it is necessary for comprehension of noncompositional meaning, in general, is an interesting open question. Overall, our results raise the intriguing possibility that the interpretation of noncompositional meaning falls into the interface of language comprehension and social cognition. If true, this would have important consequences for theories of the syntax–semantics interface: If noncompositional interpretation is computed by essentially nonlinguistic systems, language itself might involve a core computational system that is, in fact, strongly compositional.

Although our study was the first to monitor brain activity during noncompositional interpretation, the comprehension of complement coercion has also been investigated in one deficit/lesion study (Piñango & Zurif, 2001). In this study, two Wernicke's and three Broca's aphasics listened to coerced (*The boy began the book*) and transparent sentences (*The boy began reading the book*) in a picture-matching task. The picture depicted either the correct scenario of a boy reading a book or an incorrect scenario of a boy buying a book. Both patient groups performed numerically worse on coerced than transparent sentences, although this effect only reached significance for the two Wernicke's aphasics. Thus, these data are more suggestive of a main effect of coercion than the interaction that would be necessary for drawing conclusions about the role of the damaged areas in coercion interpretation. Piñango and Zurif (2001) also investigated the comprehension of so-called aspectual coercion, that is, expressions such as *the horse jumped for an hour yesterday*, where the verb denotes a punctual event but a repetitive interpretation is required. Here, Wernicke's aphasics were severely impaired, but Broca's aphasics performed well above chance, suggesting that aspectual coercion might depend on the intactness of left posterior areas. If these areas, in fact, perform aspectual coercion, rather than provide input for it, the combination of our results and Piñango and Zurif's could be taken as suggesting different neural mechanisms for aspectual and complement coercion. This would be consistent with the fact that the two constructions are, in fact, representationally very different (Pylkkänen & McElree, 2006).

Distributed Fronto-Temporal Effects of Anomaly

In order to assess whether any observed effect of coercion could be explained in terms of decreased predictability, our design included an anomalous control condition, with an entirely unpredictable object NP. Rather than modulating the AMF, anomaly was associated with a complex set of effects including an amplitude increase in the M350 dipole as well as enhanced activity in the left and right inferior frontal areas and in superior midline regions of the frontal lobe, as revealed by our distributed source analysis. The M350 effect was straightforwardly accountable in terms of the predictability of the stimuli: The control stimuli elicited the smallest M350 amplitudes and the anomalous stimuli the largest, with the coercion condition patterning in the middle.

Our anomalous stimuli all involved verbs that select for animate experiencer objects but were combined with inanimate objects. Several recent ERP studies have also manipulated animacy, reporting the surprising finding that animacy violations elicit a P600 and no N400 effect (Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2006; Kim & Osterhout, 2005; Hoeks, Stowe, & Doedens, 2004; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). This is in apparent conflict with our results, where animacy violations elicited an effect similar to previously reported N400m effects. Importantly, however, in our study, the violation was revealed at the argument (*the author astonished the article*), whereas in the ERP studies, the violation has systematically occurred at the predicate, as in *the eggs would eat* (Kuperberg, Caplan, et al., 2006; Kuperberg, Kreher, et al., 2006; Kuperberg et al., 2003) or *the meal was devouring* (Kim & Osterhout, 2005). Given the differences between these results, further studies are needed to characterize how the syntactic position of animacy violations affects processing.

Conclusion

In this study, we manipulated compositionality in order to vary semantic composition effort within syntactically matched well-formed expressions. Our results show that noncompositional interpretation elicits increased activity in regions of the vmPFC. Our anomalous control stimuli elicited no effect in this area, which rules out predictability-based explanations of the coercion effect and shows that effortful semantic composition of well-formed expressions is qualitatively different from anomaly detection. The vmPFC has commonly been implicated for social cognition and theory of mind, raising the intriguing possibility that the mechanisms responsible for noncompositional interpretation may not be ''linguistic'' per se.

APPENDIX

Nonembedded Stimuli (Coerced/Anomalous/Control):

1 the pilot mastered/amazed/flew the airplane after the intense class
2 the journalist began/astonished/wrote the article before his coffee break
3 the baby tried/panicked/ate the banana before the short nap
4 the gymnast attempted/alarmed/walked the beam during the morning competition
5 the professor started/disgusted/read the book before his psychological incident
6 the baker finished/panicked/iced the cake just before the wedding
7 the pastor endured/terrified/wore the collar through his long sermon
8 the schoolboy completed/panicked/read the comic while on the bus
9 the assistant completed/offended/made the copies before the afternoon meeting
10 the cripple mastered/angered/used the crutches before the wheelchair arrived
11 the toddlers enjoyed/comforted/ate the cupcakes before their afternoon nap
12 the celebrity enjoyed/amazed/read the email despite the foul language
13 the student finished/provoked/wrote the essay after the review session
14 the nutritionist tried/excited/took the ginseng after the vendor left
15 the Northerner tried/offended/ate the grits after the long walk
16 the photographer completed/irritated/took the headshot while the actor posed
17 the bridesmaid endured/scared/wore the heels despite the foot pain
18 the vendor completed/horrified/wrote the invoice after the client called
19 the secretary began/delighted/wrote the memo after the long meeting
20 the newborn enjoyed/displeased/drank the milk after the afternoon rest
21 the hiker attempted/amazed/climbed the mountain before the trip ended
22 the customer enjoyed/frightened/ate the nachos during the short outing
23 the editor finished/infuriated/read the newspapers during his long flight
24 the kindergartner started/astonished/drew the picture before the teacher left
25 the brothers enjoyed/astonished/ate the pizza during the shopping break
26 the artist began/excited/painted the portrait as the client arrived
27 the grandmother endured/climbed/finished the stairs before buying the lift

APPENDIX (*continued*)

28 the cook finished/cooked/endured the steak before the water boiled
29 the couple enjoyed/angered/ate the sundae after the morning service
30 the knight mastered/terrified/used the sword before the monthly competition
31 the carpenter began/comforted/built the table during the morning break
32 the marine enjoyed/excited/watched the television instead of physical exercise
33 the camper started/captivated/pitched the tent as the storm broke
34 the programmer finished/scared/made the Website after his friends arrived
35 the barfly tried/shocked/drank the whiskey before his beer arrived

Embedded Stimuli (Coerced/Anomalous/Control):

36 the architect knew that the contractor started/astounded/built the apartment after the inspector left
37 the housewife hated that the husband enjoyed/captivated/drank the beer after the church service
38 everyone here knew that the tailor finished/horrified/sewed the blouse as the customer requested
39 the maid saw that the dog enjoyed/amused/chewed the bone during the rain storm
40 the therapist knew that the youth endured/shocked/wore the braces for a long time
41 the waitress knew that the diner enjoyed/annoyed/ate the casserole while talking to friends
42 the writer reported that the quarterback enjoyed/enraged/drank the champagne after the stunning victory
43 the host understood that the novelist began/annoyed/wrote the chapter before his morning walk
44 the crowd recognized that the performer endured/alarmed/wore the costume until the show finished
45 the naturalist reasoned that the beaver completed/infuriated/built the dam after the summer rains
46 the client knew that the seamstress began/captivated/sewed the dress after her short vacation
47 the anthropologist reasoned that the tribesman mastered/shocked/played the drums before the sacred ritual
48 the assistant noted that the chef started/comforted/cooked the entree before the soup boiled
49 almost everyone thought that the caterer finished/irritated/prepared the food before the guest arrived
50 the father saw that his son enjoyed/disgusted/ate the hamburger while his mother talked
51 the police claimed that the girl completed/frightened/wrote the journal during the long trip
52 the director saw that the dancer endured/shocked/wore the leotard during the long performance
53 the warden accepted that the prisoner began/disgusted/wrote the letter after the evening meal
54 everyone was thankful that the man completed/offended/read the manual before the product shipped
55 the secretary saw that the publisher started/annoyed/read the manuscript after the author called
56 the teacher saw that the teenager began/enraged/read the novel before the evening break
57 the father saw that the grandpa tried/amused/smoked the pipe as his son watched
58 everyone was shocked that the passenger completed/pleased/read the poem while in the terminal
59 the janitor observed that the pharmacist finished/repelled/filled the prescription while the customer waited
60 the housewife knew that the guests tried/displeased/ate the salmon after the music started
61 the clerk saw that the gentleman finished/pleased/ate the sandwich while at the counter
62 the nanny said that the toddler mastered/alarmed/used the seesaw before his second birthday
63 the homeowner remembered that the handyman completed/intrigued/fixed the sink before the bathtub arrived
64 the investigator believed that the infant attempted/amused/climbed the stairs while the nanny left
65 the waiter noted that the patron tried/excited/ate the steak before the fries arrived
66 the nurse noted that the pediatrician tried/shocked/used the stethoscope as his patient shuddered
67 the proprietor liked that the lady enjoyed/fascinated/drank the tea since it was hot
68 the ringmaster knew that the acrobat mastered/amused/walked the tightrope before the circus toured
69 the client understood that the mechanic finished/intrigued/fixed the truck before the boss returned
70 the priest saw that the altar-boy tried/scared/drank the wine before the evening service

Acknowledgments

This research was supported by the National Science Foundation grants BCS-0236732 (to B. M.) and BCS-0545186 (to L. P.) as well as the NIH/NINDS grant NS13742 (to Rodolfo Llinás). We thank Rodolfo Llinás for comments on the manuscript, Andrew Smart and Olla Somolyak for substantial help in the data analysis, and Karsten Hoechstetter for valuable feedback on the distributed source analysis.

Reprint requests should be sent to Liina Pylkkänen, Department of Psychology, New York University, 6 Washington Place, Room 870, New York, NY 10003, or via e-mail: liina.pylkkanen@nyu.edu.

Notes

1. A reviewer suggested that complement coercion might involve a type of garden-path, in which an initial, anomalous interpretation is corrected later in processing. Garden-path effects

are typically seen on the earliest measures of the processing of the constituent that triggers the reanalysis (e.g., first-fixation times in eye movement during reading studies). However, the effects of complement coercion typically do not emerge on the noun but rather in later regions and on measures considered to reflect integrative processing in eye movement data (Pickering et al., 2005; Traxler et al., 2005). These delayed and sustained effects are not consistent with a simple garden-path explanation.

2. The amount of repetition of semantic content in our materials obviously depends on the degree to which readers derive specific interpretations of the implicit activity. For example, *finish the book* and *finish the steak* might receive specific interpretations, such as ''finish [reading] the book'' and ''finish [eating] the steak.'' If readers adopt a fully specified interpretation, verb repetition should affect the processing of the target noun minimally, given that the implicit activity that the target noun is related to does not repeat across trials. Alternatively, however, the interpretations of these expressions might be somewhat underspecified, with both being interpreted as ''finish [SOME ACTIVITY INVOLVING] the book/steak.'' If readers underspecify, repetition should have a facilitatory effect, biasing our materials against a coercion cost. Either way, repetition should not increase the likelihood of a coercion cost.

REFERENCES

Barker, C. (2002). Continuations and the nature of quantification. *Natural Language Semantics, 10,* 211–242.

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

Berg, P., & Scherg, M. (1991). Dipole models of eye movements and blinks. *Electroencephalography and Clinical Neurophysiology, 79,* 36–44.

Beretta, A., Fiorentino, R., & Poeppel, D. (2005). The effects of homonymy and polysemy on lexical access: An MEG study. *Cognitive Brain Research, 24,* 57–65.

Bradley, D. C., & Garrett, M. F. (1983). Hemispheric differences in the recognition of closed- and open-class words. *Neuropsychologia, 21,* 155–159.

Bradley, D. C., Garrett, M. F., & Zurif, E. B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes* (pp. 269–286). Cambridge: MIT Press.

Brown, C. M., & Hagoort, P. (1999). On the electrophysiology of language comprehension: Implications for the human language system. In M. Crocker, M. Pickering, & C. Clifton, Jr. (Eds.), *Architectures and mechanisms for language processing* (pp. 213–237). Cambridge: Cambridge University Press.

Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience, 11,* 261–281.

Deacon, D., Hewitt, S., Yang, C.-M., & Nagata, M. (2000). Event-related potential indices of semantic priming using masked and unmasked words: Evidence that the N400 does not reflect a post-lexical process. *Cognitive Brain Research, 9,* 137–146.

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

Embick, D., Hackl, M., Schaeffer, J., Kelepir, M., & Marantz, A. (2001). A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive Brain Research, 10,* 345–348.

Fiorentino, R., & Poeppel, D. (2004). Decomposition of compound words: An MEG measure of early access to constituents. In R. Alterman & D. Kirsch (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (p. 1342). Mahwah, NJ: Erlbaum.

Fodor, J. A., & Lepore, E. (2002). *The compositionality papers*. Oxford: Clarendon Press.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosopic und Philosophische Kritif* 100. 25–50; trans. by M. Black as ''On sense and reference''. In P. Geach & M. Black (Eds.), *Translations from the philosophic writings of Gottlob Frege* (1960, pp. 56–78). Oxford: Basil Blackwell.

Friederici, A. D. (1997). Neurophysiological aspects of language processing. *Journal of Clinical Neuroscience, 4,* 64–72.

Frisson, S., & McElree, B. (in press). Complement coercion is not ambiguity resolution: Evidence from eye movement. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Guillem, F., Rougier, A., & Claverie, B. (1999). Short- and long-delay intracranial ERP repetition effects dissociate memory systems in the human brain. *Journal of Cognitive Neuroscience, 11,* 437–458.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304,* 438–441.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., Marinkovic, K., & Chauvel, P. (1994). Spatio-temporal stages in face and word processing. 2. Depth-recorded potentials in the human frontal and Rolandic cortices. *Journal of Physiology (Paris), 88,* 51–80.

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

Hämäläinen, M. S., & Ilmoniemi, R. J. (1984). Interpreting measured magnetic fields of the brain: Estimates of current distributions. Technical Report TKK-F-A559, Helsinki University of Technology.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell Publishers.

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.

Hendriks, H. (1988). Type change in semantics: The scope of quantification and coordination. In E. Klein & J. Van Benthem (Eds.), *Categories, polymorphism and unification* (pp. 96–119). Amsterdam: ITLI.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19,* 59–73.

Jacobson, P. (1999). Towards a variable-free semantics. *Linguistics and Philosophy, 22,* 117–184.

Kiefer, M., & Spitzer, M. (2000). Time course of conscious and unconscious semantic brain activation. *NeuroReport, 11,* 2401–2407.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52,* 205–225.

King, J. W., & Kutas, M. (1995). A brain potential whose latency indexes the length and frequency of words. *CRL Newsletter, 10,* 1–9.

Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. (2006). Neural correlates of processing syntactic, semantic and thematic relationships in sentences. *Language and Cognitive Processes, 21,* 489–530.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2006). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language, 100,* 223–237.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research, 17,* 117–129.

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

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

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

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron, 38,* 487–497.

McElree, B., Frisson, S., & Pickering, M. J. (2006). Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. *Cognitive Science, 30,* 115–124.

McElree, B., Pylkkänen, L., Pickering, M. J., & Traxler, M. J. (2006). The time course of enriched composition. *Psychonomic Bulletin and Review, 13,* 53–59.

McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched semantic composition. *Cognition, 78,* B15–B25.

Montague, R. (1970). Universal grammar. *Theoria, 36,* 373–398.

Mosher, J. C., & Leahy, R. M. (1998). Recursive MUSIC: A framework for EEG and MEG source localization. *IEEE Transactions on Biomedical Engineering, 45, 11,* 1342–1354.

Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex, 2,* 244–258.

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

Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity. *Biological Psychology, 46,* 143–168.

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

Partee, B. (1984). Compositionality. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics* (pp. 281–311). Dordrecht: Foris.

Partee, B. H., & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use, and interpretation of language* (pp. 361–383). Berlin: Walter de Gruyter.

Pickering, M. J., McElree, B., & Traxler, M. (2005). The difficulty of coercion: A response to de Almeida. *Brain and Language, 93,* 1–9.

Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 940–961.

Piñango, M., & Zurif, E. (2001). Semantic operations in aphasic comprehension: Implications for the cortical organization of language. *Brain and Language, 79,* 297–308.

Praamstra, P., Meyer, A. S., & Levelt, W. J. M. (1994). Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience, 6,* 204–219.

Praamstra, P., & Stegeman, D. F. (1993). Phonological effects on the auditory N400 event-related brain potential. *Cognitive Brain Research, 1,* 73–86.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge: MIT Press.

Pylkkänen, L., Feintuch, S., Hopkins, E., & Marantz, A. (2004). Neural correlates of the effects of morphological family frequency and family size: An MEG study. *Cognition, 91,* B35–B45.

Pylkkänen, L., Llinas, R., & Murphy, G. (2006). Representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience, 18,* 1–13.

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

Pylkkänen, L., & McElree, B. (2006). The syntax–semantics interface: On-line composition of sentence meaning. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 537–577). New York: Elsevier.

Pylkkänen, L., Stringfellow, A., Flagg, E., & Marantz, A. (2001). A neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. *Proceedings of Biomag 2000, 12th International Conference on Biomagnetism* (pp. 363–367). Espoo, Finland: Helsinki University of Technology.

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.

Radeau, M., Besson, M., Fonteneau, E., & Castro, S. L. (1998). Semantic, repetition and rime priming between spokenwords: Behavioral and electrophysiological evidence. *Biological Psychology, 8,* 183–204.

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). ''Theory of mind'' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain, 124,* 600–616.

Rugg, M. D. (1984). Event-related potentials in phonological matching tasks. *Brain and Language, 23,* 225–240.

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

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology, 19,* 288–300.

Simos, P. G., Basile, L. F., & Papanicolaou, A. C. (1997). Source localization of the N400 response in a sentence-reading paradigm using evoked magnetic fields and magnetic resonance imaging. *Brain Research, 762,* 29–39.

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

Stockall, L., Stringfellow, A., & Marantz, A. (2004). The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. *Brain and Language, 90,* 88–94.

Stowe, L. A., Haverkort, M., & Zwarts, F. (2005). Rethinking the neurological basis of language. *Lingua, 115,* 997–1042.

Sullivan, K., Winner, E., & Hopfied, N. (1995). How children tell a lie from a joke: The role of second-order mental state attributions. *British Journal of Developmental Psychology, 13,* 191–204.

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,* 1125–1136.

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain, 22,* 2119–2132.

Traxler, M. J., McElree, B., Williams, R. S., & Pickering, M. J. (2005). Context effects in coercion: Evidence from eye-movements. *Journal of Memory and Language, 53,* 1–25.

Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: Evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47,* 530–547.

Uutela, K., Hamalainen, M., & Somersalo, E. (1999). Visualization of magnetoencephalographic data using minimum current estimates. *Neuroimage, 10,* 173–180.

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,* 657–671.

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

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

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