Distinct effects of semantic plausibility and semantic composition in MEG

Liina Pylkkänen (1, 2), Rodolfo Llinás (3), Brian McElree(1)

(1) Department of Psychology, New York University

(2) Department of Linguistics, New York University

(3) Department of Physiology and Neuroscience, New York University School of Medicine

Please address correspondence to

Liina Pylkkänen
Department of Linguistics
New York University
719 Broadway, 4th Floor
New York, NY 10003
(212) 992-8764
liina.pylkkanen@nyu.edu

Keywords: semantic composition, plausibility, coercion, N400, MEG

Abstract

To date, research on the neurobiology of semantic processing has focused almost entirely on the processing of semantic anomaly, inspired by Kutas & Hillyard's (1980) discovery of an electrophysiological anomaly detector, the N400 ERP. As a consequence, little progress has been made in understanding how the brain computes the meanings of well-formed plausible expressions. In this study we contrasted semantically implausible and semantically plausible but difficult sentences with control sentences in MEG, in order to elucidate the neural bases of "normal" semantic processing outside the classic semantic violation paradigm. Plausibility, but not semantic composition effort, affected the amplitudes of a left temporal source at 300-400ms (M350), consistent with previous MEG results. In contrast, semantic composition effort, but not plausibility, affected the amplitudes of an anterior midline field (AMF) at 350-500ms. The N400 is commonly hypothesized to reflect semantic "integration." But these results show that when semantic composition/integration is manipulated in well-formed plausible expressions, effects occur in the frontal lobe and not in the generator of the N400 effect.

1. Introduction

During the last twenty years, the newly emerging field of cognitive neuroscience of language has made significant progress in articulating theories about the functional neuroanatomy of language (Hickok & Poeppel, 2004). However, what is currently almost entirely missing from this work is research on the neural bases of semantic processing that connects with linguistic theories of semantic representation. Owing to Kutas & Hillyard's (1980) discovery of the N400 anomaly detector, there is now a rich literature on the electrophysiological effects of semantic plausibility. But plausibility depends on real-world knowledge rather than linguistic knowledge about semantic representations. Thus, although the N400 is commonly hypothesized to reflect "semantic integration" (Brown & Hagoort, 1999; Friederici, 1997; Osterhout & Holcomb, 1992; Rugg, 1990; Van Berkum, Hagoort & Brown, 1999), there is no particular evidence suggesting that it is associated with the operation of building complex *linguistic* semantic representations.

In this study we aimed to elucidate the neural bases of linguistic semantic composition by manipulating the *compositionality* of well-formed plausible expressions. Given the systematicity and productivity of natural language, sentence meanings must be largely compositional, i.e., determined by the meanings of their constituents. For linguists, compositionality is both an empirical observation and a working hypothesis: much of language is compositional and maintaining compositionality imposes constraints on possible theories of meaning. While striving for compositionality is the right methodology in linguistics, it is uncontroversial that some meanings are not compositional, at least not in any obvious way. Interestingly, there is now a body of psycholinguistic results suggesting that noncompositional meanings are taxing for the language processor (McElree, Traxler, Pickering, Seely & Jackendoff, 2001; Piñango, Zurif & Jackendoff, 1999; Todorova, Straub, Badecker & Frank, 2000; Traxler, Pickering & McElree, 2002). These results offer a promising basis for investigations of the neural bases of semantic processing. Understanding the neural sources of this type of compositionality cost would not only serve as a starting point for characterizing the neurobiology of semantic structure building, but it would also provide a valuable dependent measure for addressing foundational questions of to what extent the syntactic and semantic systems of natural language are separable and to what extent semantic and pragmatic systems are separable. In this study we investigated the neural sources of semantic composition with magnetoencephalography (MEG).

1.1 Complement Coercion

To manipulate the compositionality of plausible, well-formed expressions, we employed *complement coercion* (e.g., Pustejovsky, 1995). In these expressions the selectional restrictions of an event-selecting verb appear to be violated. For example, while the aspectual verb *begin* would seem to semantically require an event complement, it is perfectly grammatical, and plausible, with a complement denoting an entity: *the author began the book*. Typically, readers interpret this sentence as *the author began writing the book* or, less frequently, as *the author began reading the book* (McElree, Traxler, Pickering, Seely, & Jackendoff, 2001; Lapata & Lascarides, 2003; Lapata, Keller, & Scheepers, 2003; Traxler, Pickering, & McElree, 2002). This suggests that readers "coerce" the book into the required semantic type by interpreting it as part of an unstated eventive complement (e.g., "writing the book"). This hypothesis is supported by evidence from self-paced reading (McElree et al., 2001; Pickering, McElree, & Traxler, in press; Traxler et al, 2002), eye tracking during reading (Pickering et al., in press; Traxler et al., 2002) and the speed-accuracy trade-off procedure (McElree, Pylkkänen, Pickering, & Traxler, submitted): sentences involving complement coercion systematically elicit a processing delay.

The coercion delay has been argued to reflect increased effort in semantic composition. Alternatively, it could reflect effort associated with the process of retrieving an appropriate activity to relate to the "mismatching" entity-complement. However, the coercion cost remains even when the activity is explicitly given in the context (Traxler et al., submitted). For example, a context sentence stating that a student was reading in the library does not eliminate the cost associated with processing a subsequent expression *Before the student began the article…* Further, the delay does not reflect a general difficulty in combining verbs such as *begin* with NP complements since a coercion effect is elicited 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) vs. *begin the fight* (event NP) (Traxler et al., 2002). Finally, a delay is elicited even when the aspectual properties of the coercion and control stimuli are entirely controlled for (*the author began the book* vs. *the author began writing the book*), suggesting that the coercion cost is not due to the telicity asymmetry between sentences such as *the author began the book* and *the author wrote the book* (Pickering et al., in press). Thus, abundant behavioral evidence point to the conclusion that complement coercion is precisely the type of manipulation that directly taps onto the linguistic operation of semantic composition. The goal of this study was to identify the neural sources of the coercion effect.

1.2. Neuromagnetic predictions

To identify the neural sources of the coercion cost and to contrast these sources to standard N400 plausibility effects, we examined the 3 sentence types illustrated in (1)-(3).

(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.

Sentences like (1), which involve complement coercion, were contrasted to compositionally simpler control sentences like (3), where the complement of the verb can be interpreted without coercing it into a different semantic type. We also directly compared sentences like (1) to anomalous sentences like (2), where the complement violated the animacy restrictions of the verb.

The ill-formedness of anomalous expressions such as (2) depends on the real-world knowledge that articles are inanimate entities and thus do not have mental states (such as astonishment). Although the well-formedness of coerced sentences such as (1) depends on real-world knowledge—that articles are the types of things that can written and read— it also crucially depends on the linguistic knowledge that entity-denoting NPs can be coerced into suitable objects for event-selecting verbs. Both (1) and (2) plausibly involve integration difficulty, although for different reasons. In the anomalous case, *article* may be difficult to integrate into the preceding context because doing so requires imagining an extremely implausible state of affairs (i.e. an article that in fact does have mental states). In contrast, the coerced case does not require imagining an implausible state of affairs. However, it does require a more taxing form of semantic integration than (3); specifically, the behavioral evidence is consistent with notion that interpreting expression like (1) require an extra compositional process to coerce the NP into a suitable semantic form (McElree et al, 2001, submitted; Pickering et al, in press; Traxler, Pickering & McElree, 2002).

In electrophysiology, perhaps the most robust neurolinguistic result to date is that anomalous sentences such as (2) elicit increased amplitudes of the N400 ERP. Consequently, the N400 has been hypothesized to reflect semantic integration (Brown & Hagoort, 1999; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Osterhout & Holcomb, 1992; Rugg, 1990; Friederici, 1997). In this study we tested whether N400 source amplitudes are modulated by coercion. If the N400 generator is indeed the principle locus of semantic processing, then N400 effects should be obtained even when semantic composition is manipulated outside the anomaly paradigm. Alternatively, if the N400 generator does not perform composition of linguistic, semantic representations, we would expect the neural generator(s) of the coercion effect to be distinct from the N400 source. Finally, if coercion and anomaly elicit clearly different effects, this would also show that coerced sentences are not simply slightly "anomalous," but rather engage qualitatively different processing from anomaly detection.

While semantic composition and semantic plausibility have not been previously contrasted in neurolinguistic research, it has recently been argued that semantic and world knowledge integration engage the N400 in exactly the same way. In order to investigate the time course of semantic and world-knowledge integration, Hagoort et al. (2004) contrasted sentences that were either false for Dutch speakers, given their culture specific world knowledge, or in general anomalous, in ways similar to (2). Anomaly was taken to depend on semantic knowledge and truth/falsity on world knowledge. Although results patterned differently for falsity and anomaly in the spectral domain, the N400 effects elicited by false and anomalous sentences were identical, leading the authors to suggest that there is no principled distinction between semantic and world-knowledge integration. However, as we elaborate in our General discussion, from a linguistic point of view, this experiment involved two manipulations of world-knowledge, one general and the other culture specific, and it did not manipulate semantic processing directly. Thus, we believe the question whether semantic and world-knowledge processing engage distinct neural sources remains unsettled.

1.3. Magnetoencephalography (MEG)

Brain activity was monitored with MEG. MEG differs from EEG in that it measures the magnetic fields, instead of electric potentials, generated by postsynaptic currents in nerve cells. Unlike electric potentials, magnetic fields are not distorted by the skull. Therefore, localization of the currents underlying the activity measured outside the head is more straightforward in MEG than in EEG. (For a review of magnetoencephalography, see Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993.) To take advantage of this, we modeled the complex neural activity elicited by the complement noun (e.g., *article* in 1-3) with multidipole analysis. According to previous MEG results, the classic N400 effect should localize to left superior temporal areas (Helenius, Salmelin, Service & Connolly, 1998; Halgren et al., 2002; Simos et al. 2002).

Experiment 1 investigated the neural sources of the coercion effect. Experiment 2 involved a different kind of semantic manipulation, targeted at narrowing down the interpretation of the results of Experiment 1.

2. Experiment 1: Coercion

2.1. Participants

17 native English speakers participated in the study. All were graduate or undergraduate students at NYU. One subjects' data were excluded from the analysis due to excessive noise artifacts.

2.2. Materials

We tested 70 triplets like (1) – (3) (see Appendix 1), which we combined with 135 filler sentences with different syntactic structures. To further diversify the materials, we embedded 35 of the triples as complements of clause-selecting verbs, as in (4) – (6).

(4) Coerced: The client knew that the seamstress began the **dress** after her short vacation.

(5) Anomalous: The client knew that the seamstress captivated the **dress** after her short vacation.

(6) Control: The client knew that the seamstress sewed the **dress** after her short vacation

Anomalous variants of the coerced sentences were generated by replacing the coercing verbs with verbs like *astonish, amuse*, etc., that 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 like *article*, *dress*, etc.

It was important that our control sentences (3 and 6) employed a verb that expressed the eventive interpretation readers most often ascribed to the coerced sentences (1 and 4). An additional 24 NYU participants provided 1-2 word fill-in-the-blank responses indicating how they would interpret a coerced sentence like "The journalist began the ______ 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 (e.g., wrote). This verb occurred as a fill-in-the-blank response on average 74% of the time (s.d. = 14.6%). We had another group of 20 NYU students rate 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. Additionally, we have collected processing time measures on a subset of these triples. Forty-five of the 70 triples were used as materials in a speed-accuracy tradeoff study (McElree, Pylkkänen, Pickering, & Traxler, submitted), 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 like (1) and (4) engendered more taxing compositional operations than control expressions like (3) and (6).

2.3. Procedure

During the experiment the participants lay in a dimly lit magnetically shielded room and viewed the experimental stimuli via fiberoptic goggles (Avotec, 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 (300ms on, 300ms 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 100 1 kHz tones. While we did not have MRI's for our subjects, source locations of the auditory M100 allowed us to visualize activity with respect to a functional landmark representing primary auditory cortex. The entire recording session lasted approximately one hour.

2.4. MEG data analysis

MEG data were cleaned of artifacts and then averaged according to stimulus category by using an epoch length of 800 ms, preceded by a prestimulus interval of 300 ms. Prior to source modeling, MEG averages were high-pass filtered at 1Hz and low-pass filtered at 40Hz.

A multiple-source model, BESA (Brain Electric Source Analysis, 5.0), was applied to all MEG activity elicited at 0-600ms. First, separate source models were created for each condition within a subject and compared to each other. All sensors were used in localization. Sources were fit at the peaks of prominent response components using both minimum norm estimates and magnetic field patterns to constrain the number of dipoles. Only dipoles whose location and orientation were consistent with the magnetic field patterns were accepted. These component-specific dipole models were then introduced into a single multi-dipole model, characterizing most of the activity elicited by the stimulus category. All conditions revealed a response pattern familiar from previous MEG language studies in the visual modality (Embick, et al., 2001; Fiorentino & Poeppel, 2004; Helenius, Salmelin, Service, & Connolly, 1998, 1999; Pylkkänen, et al., 2001, 2002, 2004, submitted; Stockall, Stringfellow & Marantz, 2004; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999). Early visual responses at ~100 ms and ~150-200 ms were followed by two major response components in the temporal regions, one at around 250 ms (M250) and the second around 350 ms (M350). M350 field patterns were the clearest for the implausible sentences. In addition, all subjects showed an anterior midline field (AMF) at 350-500ms, which was particularly prominent in the coercion condition. This component localized consistently to anterior, inferior, midline areas. Anterior midline areas have been found to be activated by linguistic stimuli in previous MEG studies as well (Halgren et al., 2002; Marinkovic et al., 2003).

For statistical analysis, we created a single multidipole model for each subject, which was then kept constant across conditions. Given the extreme similarity of the M100-M170-M250 complex across conditions, individual grandaveraged data were used to model these early sources. Since M350 field patterns were generally clearest for the anomalous stimuli, M350 sources localized on the basis of the implausible sentences were used in the models. Similarly, since AMF fields were the clearest for coerced sentences, we used AMF sources localized on the basis of the coerced sentences in the models.

2.5. Results

2.5.1. Sensicality judgment data

The end of sentence sensicality judgment data are summarized in Table 1. Response speed and accuracy were both reliably affected by the stimulus manipulation, anomalous sentences eliciting faster and more accurate responses than both coerced and control sentences ($F(2,15)= 82.243$, $p < 0.001$ and $F(2,15) = 3.41$, $p < 0.05$ respectively). This is unsurprising, given that subjects were able to determine the nonsensicality of the anomalous stimuli already in mid sentence. Most importantly, coerced and control sentences did not differ in sensicality judgment accuracy, suggesting that the subjects found them equally plausible.

INSERT TABLE 1

INSERT FIGURE 1

2.5.2. Multiple source models

Figure 1 depicts all source localizations broken down by time window. Source models were kept constant across experimental conditions. Goodness of fit did not vary as a function of stimulus category: the multiple source models explained approximately 80% of all activity between 0 and 650 ms in all conditions (coercion: 82%, SD = 5%; anomalous: 81%, SD = 5.2%; control: 81%, SD = 4.6%).

Activity at 0-300 ms. Consistent with previous MEG findings (Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999), our visually presented stimuli elicited a clear visual M100 response at 90-150 ms, followed by bilateral M170 activity at 150-230ms. The visual M100 is associated with a right-lateralized outgoing field and a left-lateralized re-entering field over the occipital sensors. Visual M100 activity was clearly 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., submitted), localization of temporal activity at 200-300ms (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 200-300ms that was distinct from the M170 and from later M350 activity. This mid-latency activity localized to temporal areas, bilaterally in 6 subjects, in the left hemisphere only in 3 subjects, and in the right hemisphere only in 6 subjects.

M350 and "M350-R." At 300-400ms, activity elicited by visual words typically localizes in temporal areas, close to primary auditory cortex, either bilaterally, or in the left hemisphere only (Halgren et al., 2002; Helenius et al. 1998, 1999; Pylkkänen, et al., 2002, 2004, submitted). 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 (Embick, et al., 2001; Fiorentino & Poeppel, 2004; Pylkkänen, et al., 2001, 2002, 2004, submitted; Pylkkänen & Marantz, 2003; Stockall, Stringfellow & Marantz, 2004).

A clear left hemisphere M350 field pattern was identified in 13 of our 16 subjects and this activity localized consistently in left superior and middle temporal areas. For 9 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 3 subjects' data, activity at 300-400ms 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 post the M350.

The anterior midline field (AMF) at 350-500ms. All subjects showed activity at 350-500ms 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.

INSERT FIGURE 2

2.5.3. Frontal effect of coercion, temporal effects of plausibility

As shown in Figure 2, the stimulus manipulation had a reliable effect on AMF source amplitudes ($F(2,15) = 4.5$; $p < 0.05$), coerced sentences eliciting larger AMF amplitudes than control sentences ($p < 0.05$, Scheffe). Anomalous sentences did not reliably differ from control sentences in AMF amplitude ($p = .6$, Scheffe). Thus, the AMF amplitude effect was specific to coercion.

INSERT FIGURE 3

As shown in Figure 3, the left hemisphere M350 was also sensitive to the stimulus manipulation ($F(2,15) = 4.5$; $p < 0.05$) and showed the classic N400 anomaly effect. Anomalous sentences elicited larger M350 amplitudes than control stimuli ($p < 0.05$, Scheffe), while coercion did not modulate the M350. Anomalous stimuli also elicited longer M350 latencies than controls ($F(2,12) = 5.49$; $p < 0.05$), likely a secondary effect of the M350 amplitude effect. These results conform to previous MEG findings by Helenius et al. (1998), Halgren et al. (2002) and Simos et al. (1997), who localized the classic N400 effect in 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 be the same source.

Right hemisphere activity in the M350 time window, the “M350-R,” also tended towards larger amplitudes for anomalous stimuli, but this effect did not reach significance. An effect was, however, obtained in M350-R latency ($F(2,9) = 3.6$; $p < 0.05$), anomalous stimuli eliciting longer latencies than controls ($p < 0.05$, 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, Llinás & Murphy, submitted).

Since 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 “inhibition.” Indeed, AMF amplitudes were delayed for anomalous sentences ($F(2,15) = 4.9$; $p < 0.05$).

In addition to the AMF and midlatency temporal effects, left-hemisphere M250 source amplitudes were affected by the stimulus manipulation ($F(2,8) = 3.8$; $p < 0.05$), 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 subject 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 activity 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 plausibility. All latency and amplitude data are summarized in Table 2.

INSERT TABLE 2

2.6. Discussion

We used complement coercion to test whether the N400 source, commonly hypothesized to perform semantic “integration,” is sensitive to semantic composition effort in plausible grammatical expressions. Our results show that this is not the case: semantic composition modulates source amplitudes of the AMF, not the N400 generator. Thus we have identified a potential neural index of the linguistic operation of semantic structure building that is distinct from the neural correlates of processing various types of anomalies.

One alternative explanation of our results is that the AMF amplitude effect is associated with our experimental task, rather than composition effort. The subjects' task was to perform a sensicality judgment at the end of the sentence. Off-line sensicality judgments were not different for coerced and control sentences in speed or accuracy, but it is possible that evaluation of sensicality was more difficult for coerced than for non-coerced complements, although it would have to be at a point prior to the end of the sentence to engender our results.

Could the AMF effect be due to differential effort in judging the sensicality of coerced and non-coerced complements? The semantic composition hypothesis of the AMF effect predicts that (a) the coercion effect should remain even if the task is changed and (b) that AMF amplitude should not be predictive of *on-line* sensicality judgment performance. Testing prediction (a) will require further experiments. Prediction (b), however, is born out in the results of Pylkkänen, Llinás & Murphy (submitted), where subjects performed sensicality judgments on noun phrases consisting of a modifier and a noun. Main dependent measures were the M350 and sensicality judgments that were performed immediately after the presentation of the phrase. The pattern of effects on the behavioral data did not directly reflect the pattern of effects on the M350, suggesting that the stimulus manipulation affected early lexical processing and late decision related processing distinctly. If the AMF was associated with semantic composition, it should have been elicited in this experiment since the stimuli were two-word phrases and therefore required semantic composition. Consistent with this prediction, the AMF was clearly present in most subjects' data. However, contrary to the predictions of the hypothesis that the AMF is associated with sensicality judgments, the AMF showed no sensitivity to the stimulus manipulation, even though sensicality judgments did. Thus it is unlikely that that the AMF reflects sensicality judgment effort.

A second alternative explanation of the AMF effect is that it reflects effort in retrieving a suitable activity to associate with the entity complement instead of reflecting semantic composition effort in a general sense. Previous results suggest that the effect of coercion on reading times does not reflect retrieval effort (Traxler et al., submitted), but these findings do not directly address the possibility that the AMF effect is linked to retrieval. In order to test this, we manipulated semantic composition effort with different stimuli, where activity retrieval did not distinguish the stimulus categories.

3. Experiment 2: Secondary predicate interpretation

In order to test whether the AMF amplitude effect generalizes beyond constructions requiring retrieval of an unstated activity, we varied the interpretation of verb phrase (VP) final adjectives, or "secondary predicates." Secondary predicates can be interpreted either as describing a result of the activity described by the main predicate, as shown in (7), or a state in which one of the event participants is during the event described by the verb, (8). The latter type of interpretation is generally called a "depictive."

(7) Resultative: The artist knocked the picture frames **crooked**.

(8) Depictive: The artist returned the picture frames **crooked**.

Understanding (8) involves understanding that the picture frames were crooked during the returning event. Thus, interpretation of depictives requires relating the post-verbal adjective both to the event described by the verb and to one of the event participants (in this case the object). In contrast, resultatives only involve construing the post-verbal adjective as a resultant state of the event described by the verb. Thus, plausibly, depictives are semantically more complex than resultatives (cf., Frazier & Clifton, 1996). Consequently, if the AMF performs semantic composition, it should be more active for depictives than for resultatives.

3.1. Self-paced reading data

If depictives are indeed semantically more complex than resultatives, then they should take longer to process than resultatives. We first carried out a behavioral reading time pre-test to verify that this is the case. 68 NYU undergraduates participated in the study for course credit or payment. Eighteen sentence pairs of the kind shown in (7-8) served as stimuli. The choice of main verb always forced the secondary predicate to be interpreted either as a depictive or as a resultative. The secondary predicate was always followed by a prepositional phrase, in order to avoid sentence wrap-up effects at the critical adjective (e.g. *The artist knocked/returned the picture frames crooked despite the warnings*). The materials were divided into two lists such that no subject read saw both the resultative and depictive versions of a sentence. Stimulus presentation was self-paced, with a moving window procedure. Each trial began with a series of dashes on the computer screen in place of the letters in the words. The first press of the space bar replaced the first set of dashes with the first word in the sentence. With subsequent space-bar presses, the next set of dashes was replaced by the next word, and the preceding word was replaced by dashes. A yes-or-no question followed each sentence and participants did not receive feedback on their answers.

INSERT FIGURE 4

Figure 4 plots the reading time data. Reading times were marginally slower for depictives than for resultatives at the critical adjective ($F(1,67) = 2.5$; $p = 0.11$) and reliably so at the first word of the spill-over region ($F(1,67) = 3.7$, $p = 0.05$), although this result reached significance only in the by-subjects analysis, and not in the items analysis (all p's > 0.05 for the latter). The overall results do, however, suggest that depictives are harder to process than resultatives.

3.2. AMF data

The materials of the reading time study formed one set of the filler items of Experiment 1 (see Appendix 2). Our analysis of the secondary predicate data was focused on testing whether secondary predicate interpretation affected AMF amplitudes in a way that paralleled the reading time data. For this purpose, the AMF dipoles of Experiment 1 were imported into the secondary predicate data.

INSERT FIGURE 5

As shown in Figure 5, secondary predicate interpretation clearly modulated AMF amplitude, depictives eliciting larger source amplitudes than resultatives ($F(1,15) = 5.8$; $p < 0.05$). Thus, the AMF effect found in Experiment 1 cannot be due to activity retrieval, since such a hypothesis could not explain why depictives should engage the AMF source more than resultatives. In contrast, if the AMF reflects semantic composition, both data sets are explained.

4. General discussion

We manipulated semantic composition in two ways and identified a neural source in anterior inferior areas that was more active when semantic composition was effortful. Crucially, this source was insensitive to a direct manipulation of plausibility (Experiment 1). Rather, plausibility modulated activity in superior and middle temporal areas, which accords with previous MEG results on the neural sources of the classic N400 anomaly effect. The same source, the so-called M350, has also been shown to vary in latency and/or amplitude with a number of stimulus factors affecting lexical access. Collectively, then, these results suggest that the M350 indexes lexical access, which is modulated by the predictability of a stimulus, and therefore shows an "N400 effect."

How do our AMF and superior temporal M350 results fit into the larger N400 literature? Because the N400 is sensitive to plausibility/anomaly, it is commonly taken to reflect *semantic integration* (Brown & Hagoort, 1999; Friederici, 1997; Hagoort et al., 2004; Osterhout & Holcomb, 1992; Rugg, 1990;), or "some aspect(s) of the processes that integrate the meaning of a particular word into a higher-order semantic interpretation" (Van Berkum, Hagoort & Brown, 1999, p. 658). However, N400 research has generally been conducted on rather coarse-grain contrasts that are not necessarily motivated by precise linguistic theories of semantic interpretation. Hence, the relationship between semantic integration and semantic *composition* in the linguistic sense of the term (i.e. the building of a semantic representation for a sentence) has not been made explicit. Thus, the fact that we did not find the N400 generator to be sensitive to semantic composition effort does not necessarily falsify the integration hypothesis, provided that "integration" does not refer to the composition of syntactic/semantic representations.

Hagoort et al. (2004) eschew any principled distinction between semantic and pragmatic (world-knowledge) integration based on findings of comparable N400 effects in pragmatically false sentences, *The Dutch trains are white…,* and what they take to be semantically anomalous sentences *The Dutch trains are sour….* Because Dutch trains are yellow not white, a fact well understood by the Dutch participants in their experiments, the first type of anomaly clearly represents a violation of pragmatic knowledge. However, it is by no means clear that the second anomaly should be regarded as a violation of semantic rather than pragmatic knowledge. Hagoort et al. suggest that advocates of separate semantic and pragmatic processing systems would indeed assume that the second type of anomaly constitutes a violation of semantic integration:

> "The core meaning of sour is related to taste and food. Under standard interpretation conditions, a predicate requires an argument whose semantic features match that of its predicate…this is clearly not the case, because semantic features related to taste and food do not apply to trains. One could thus argue that for semantic internal reasons, the third sentence is false or incoherent" (p. 438).

However, it is unclear whether modular theories of semantic composition would include argument-predicate feature checks based on properties like "taste" and "food"; indeed, to do so appears to preclude constructing various "figurative" interpretations that routinely require comprehenders to gloss over these types of feature mismatches. Even if compositional processes "detect" feature mismatches in constructions like *The Dutch trains are sour*, one has to be concerned that comprehenders might access pragmatic knowledge in an attempt to construct a pragmatically plausible construal of the anomalous expression, and that the N400 component reflects the failure to derive a pragmatically acceptable interpretation. From this perspective, it is not at all surprising that anomalous expressions engender ERP patterns similar to more obvious violations of pragmatic knowledge.

Confounding matters further, we know that, independent of the present findings, the integration hypothesis is challenged by a body of evidence suggesting that, similarly to the M350, the N400 reflects lexical and not combinatory processing (Kutas & van Petten, 1994). First, the N400 is elicited by all words even when they are not presented in a sentential context. Second, the N400 is sensitive to lexical factors such as frequency (Smith and Halgren, 1987; Van

Petten & Kutas, 1990, 1991; Van Petten & Rheinfelder, 1995) and repetition (Besson et al., 1992; Rugg, 1987). Third, the N400 shows not only semantic priming effects (Bentin et al., 1985; Nobre & McCarthy, 1994), but also *phonological* priming (Praamstra et al., 1994; Radeau et al., 1998, Rugg, 1984). Finally, N400 priming effects can be obtained even when the prime is made invisible to conscious perception, suggesting that the N400 generator cannot be performing post-lexical integration to a preceding context (Deacon et al., 2000; Kiefer & Spitzer, 2000).

In sum, the properties of the N400 ERP and the properties of the M350 conform to a great extent, and the M350 source is a likely generator of a significant portion of the activity generating the N400 ERP (Pylkkänen & Marantz, 2003). Further, when single dipole modeling is used, the N400 generator localizes precisely in the same areas as the M350 source (Halgren et al., 2002; Helenius et al., 1998; Simos et al. 2002.). However, distributed source modeling and intracranial recordings have shown that the neural generators of the N400 are not limited to left temporal areas, but also include widespread activation in perisylvian, orbital, frontopolar and dorsolateral prefrontal cortices (Halgren *et al.,* 1994a,b; Halgren et al., 2002). Thus, some of the reported sources of the N400 in fact appear very close to the source of the AMF, which we found to be sensitive to semantic composition but *not* to plausibility. For example, in addition to their pragmatic vs. semantic N400 ERP results, Hagoort et al. (2004) reported fMRI results localizing the source of their N400 effect to left inferior prefrontal cortex (BA 45 and 47), i.e., somewhere between our AMF and N400 sources. This obviously complicates the comparison of the two studies. In sum, the literature is currently not entirely consistent with respect to the localization of the N400, results from different techniques and analysis methods pointing to somewhat different conclusions. Thus, further experimentation is clearly needed to elucidate the precise functional dynamics of temporal and frontal regions in semantic processing.

5. Conclusion

The goal of this study was to elucidate the neural bases of semantic composition. In order to achieve this, we manipulated the compositionality of well-formed plausible expressions. As a result, we identified an anterior midline field (AMF) whose neural generator was more active when semantic composition was difficult. Importantly, anomaly did not affect AMF amplitudes. This shows that, instead of being associated with anomaly/violation detection, the AMF is likely to perform computations that are part of the building of "regular" well-formed semantic structures. As far as we are aware, such neural activity has not been previously reported.

In addition to informing models of the neuroanatomy of language, the AMF provides a new dependent measure for addressing foundational issues in theoretical semantics. The precise way in which natural language is compositional is perhaps the most important, unresolved, foundational question in the investigation of linguistic meaning. Noncompositionality of meaning has been dealt with in a variety of ways, including syntactic movement (Heim & Kratzer, 1998), purely semantic "type-shifting" rules (e.g., Barker, 2002; Hendriks, 1988; Jacobson, 1999; Partee & Rooth, 1983), as well by divorcing syntactic and semantic composition entirely so that the grammar no longer architecturally forces compositionality at any level (Jackendoff, 2002).

In this study we have identified neural activity that is sensitive to compositionality. This constitutes a promising start for a neurolinguistic investigation of natural language semantics. Clearly though, these results do not yet tell us what precise computations the AMF source is performing in order to resolve, for example, the noncompositionality of coerced sentences. The mechanisms could be purely semantic, consistent with type-shifting theories, but they could also be syntactic. For example, the AMF effect could reflect the insertion of a syntactic verbal projection to mediate between event-selecting verbs, such as *begin*, and entity-denoting complements. Finally, the mechanisms could be entirely non-linguistic, demanding the recruitment of extra-linguistic cognitive skills. Although our current results do not yet determine what computations *other* than coercion resolution the AMF performs, the identification of the AMF will provide researchers with an entirely novel means of investigating how the language processor deals with noncompositionality, whether it is syntactically, semantically or non-linguistically.

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.

Besson, M., Kutas. M., & Van Petten, C. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience,* 4(2)*,* 132-149.

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

Deacon, D. et al. (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.

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,* Mahwah, NJ: Erlbaum. 1342.

Frazier, L., & Clifton, C., Jr. (1996) *Construal*. Cambridge,MA: MIT Press.

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

Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M. (2004) Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438-41.

Halgren, E., Dhond, R.P., Christensen, N., VanPetten, C., Marinkovic, K., Lewine, J. D., Dale, A. M. (2002). N400-like MEG responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage*, 17; 1101-16.

Halgren, E., Baudena, P., Heit, G., Clarke, J. M., and Marinkovic, K. (1994a). Spatio-temporal stages in face and word processing. 1. Depth-recorded potentials in the human occipital, temporal and parietal lobes. *Journal of Psychophysiology,* 88, 1–50.

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

Hendriks, H. (1988). Type Change in Semantics: the Scope of Quantification and Coordination. In Klein, E. & Van Benthem, J. (Eds.) *Categories, Polymorphism and Unification*, ITLI, Amsterdam, 96–119.

Heim, I. and A. Kratzer. (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.

Hickok, G. & Poeppel, D. (2004). Towards a new functional anatomy of language. *Cognition,* 92, 1–12.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J. Knuutila, J., & Lounasmaa, O. L. (1993). Magnetoencephalography: Theory, instrumentation, and applications to noninvasive studies of the working human brain. *Review of Modern Physics,* 65, 413–495.

Jacobson, P. (1999). Towards a Variable-Free Semantics, *Linguistics and Philosophy* 22, 117-184.

Lapata, M., Keller, F., & Scheepers, C. (2003). Inter-sentential context effects on the interpretation of logical metonymy. *Cognitive Science*, 27, 649-668.

Lapata, M., & Lascarides, A. (2003). A probabilistic account of logical metonymy, *Computational Linguistics*, 29, 263-317.

Lascarides, A., & Copestake, A. (1998). Pragmatics and word meaning. *Journal of Linguistics*, 34, 387-414.

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

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

Kutas, M. and Van Petten, C. (1994).Psycholinguistics Electrified: Event-related potential investigations. In: M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*, Academic Press, pp. 83-143.

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*, 8;38(3):487-97.

McElree, B., Pylkkänen, L., Pickering, M. J. & Traxler, M.J. (submitted). Time-course measures of enrich composition.

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.

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

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

Partee, B. H. & Rooth, M. (1983). Generalized conjunction and type ambiguity. In Bäuerle, R., Schwarze, C. & von Stechow, A. (Eds.). *Meaning, Use, and Interpretation of Language*. Berlin: Walter de Gruyter. 361-83.

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.

Pickering, M. J., McElree, B., & Traxler, M. J. (in press). The difficulty of coercion: A response to de Almedia. *Brain and Language*.

Piñango, M., Zurif, E., & Jackendoff, R. (1999). Real-time processing implications of enriched composition at the syntax-semantics interface. *Journal of Psycholinguistic Research, 28*, 395-414.

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 CognitiveNeuroscience*, 6, 204–219.

Pustejovsky, J. (1995*). The Generative Lexicon*. Cambridge, MA: 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., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*, 7, 187-189.

Pylkkänen, L., Llinás, R. & Murphy, G. (submitted). Representation of polysemy: MEG evidence.

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 SL. (1998). Semantic, repetition and rime priming between spokenwords: behavioral and electrophysiological evidence. *Biological Psychology*, 8(2),183-204.

Rugg, M. D. (1984). Event-related potentials in phonological matching tasks. *Brain and Language*,23, 225-40.

Rugg, M. D. (1987). Dissociation of semantic priming, word and non-word repetition effects by event-related potentials, *Quarterly Journal of Experimental Psychology* 39A, l23–l47.

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

Simos, P. G., Basile, L. F., and 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., and 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.

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* (Pt 11), 2119-32.

Todorova, M., Straub, K., Badecker, W., & Frank, R. (2000). Aspectual coercion and the on-line computation of sentential aspect. In the *Proceedings of the twenty-second annual conference of the Cognitive Science Society*, Philadelphia, PA.

Traxler, M.J., McElree, B., Williams, R.S., & Pickering, M.J. (submitted). Context Effects in Coercion: Evidence from Eye-Movements.

Traxler, M. J., Pickering, M. J., & McElree, B. D. (2002). Coercion in sentence processing:

evidence from eyemovements and self-paced reading. *Journal of Memory and Language,* 47, 530–547.

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

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

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

AUTHOR NOTES

This research was supported by National Science Foundation grant BCS-0236732 (awarded to BM).

APPENDIX 1
Stimuli of Experiment 1

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 kindergartener 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

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/frighten/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 alter-boy tried/scared/drank the wine before the evening service

APPENDIX 2
Stimuli of Experiment 2

(depictive/resultative)

1 the angry caterer served/made the cookies black to spite the hosts

2 the young cat left/sucked the carpet clean after misbehaving all day

3 the artist returned/knocked the picture frames crooked despite the extensive warnings

4 the senior technician found/cut the circuit dead upon arrival to work

5 the hair-dresser cut/blotted the woman's hair dry during the first appointment

6 the arrogant guest passed/licked the bowl empty without anyone noticing it

7 the trucker recalled receiving/squashing the tire flat after making his complaint

8 the customer returned/poured the soup bowl full before leaving the restaurant

9 the unauthorized doctor inspected/rendered the wound infected in his suspicious practice

10 the police officer saw/set the suspect loose right outside the station

11 the new baker sold/sprinkled the cakes moist during his morning shift

12 the unlucky vacationer carried/broke his suitcase open in the departure lobby

13 the elderly photographer left/rolled the blinds shut to avoid the daylight

14 the jovial landlord delivered/sanded the boards smooth for the new tenants

15 the skilled caterer served/grated the cheese thin to please the guests

16 the ballet students held/squeezed their abs tight during the last pose

17 the incompetent renovators displayed/installed the floor uneven for the disapproving customers

18 the volunteer fireman wore/sprayed his jacket wet despite the nasty breeze

Table 1. Offline sensicality judgment data

	Coerced	Implausible	Control
RT	791	544	799
Accuracy	86%	93%	86%

Table 2. Mean source latencies and amplitudes in all conditions. (Statistically reliable effects are shaded.)

	LATENCY (msec)			AMPLITUDE (nAm)		
	Coerced	Anomalous	Control	Coerced	Anomalous	Control
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

FIGURE LEGENDS

Figure 1: All source localization broken down by time window. On the right, auditory M100 localizations are plotted, to yield a functional landmark representing primary auditory cortex.

Figure 2: Effect of the stimulus manipulation on source amplitudes of the anterior midline field (AMF) (left: single subject; right: group).

Figure 3: Effect of the stimulus manipulation on M350 source amplitudes (left: single subject; right: group).

Figure 4: Effect of secondary predicate interpretation on self-paced reading times.

Figure 5. Effect of secondary predicate interpretation on AMF amplitudes.

Figure 1.

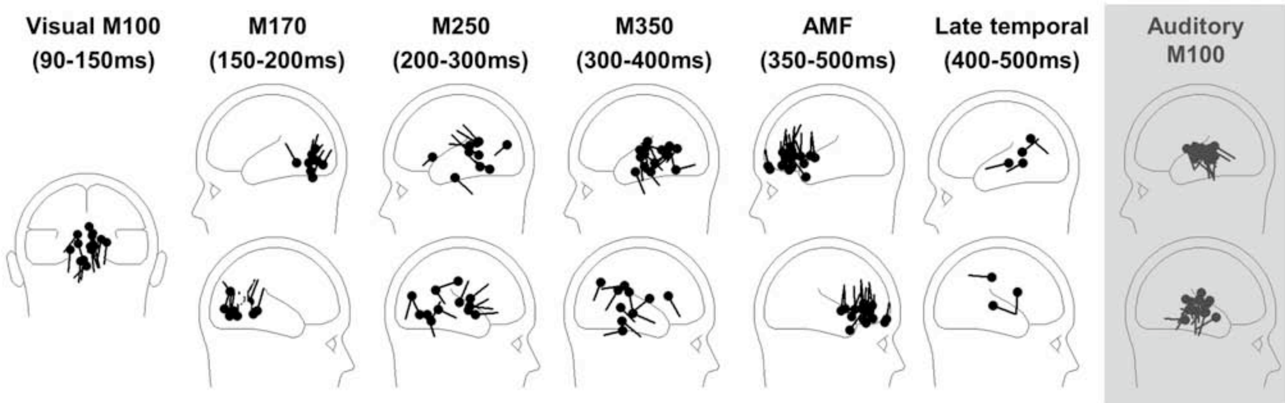
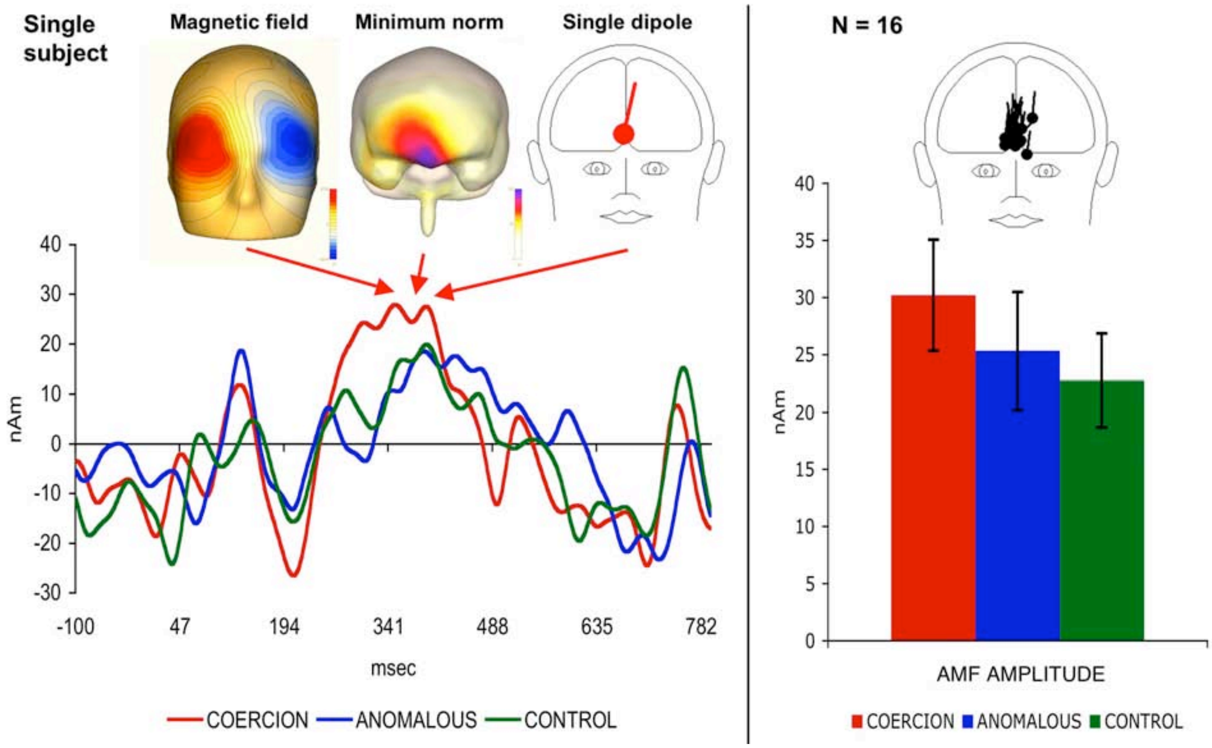
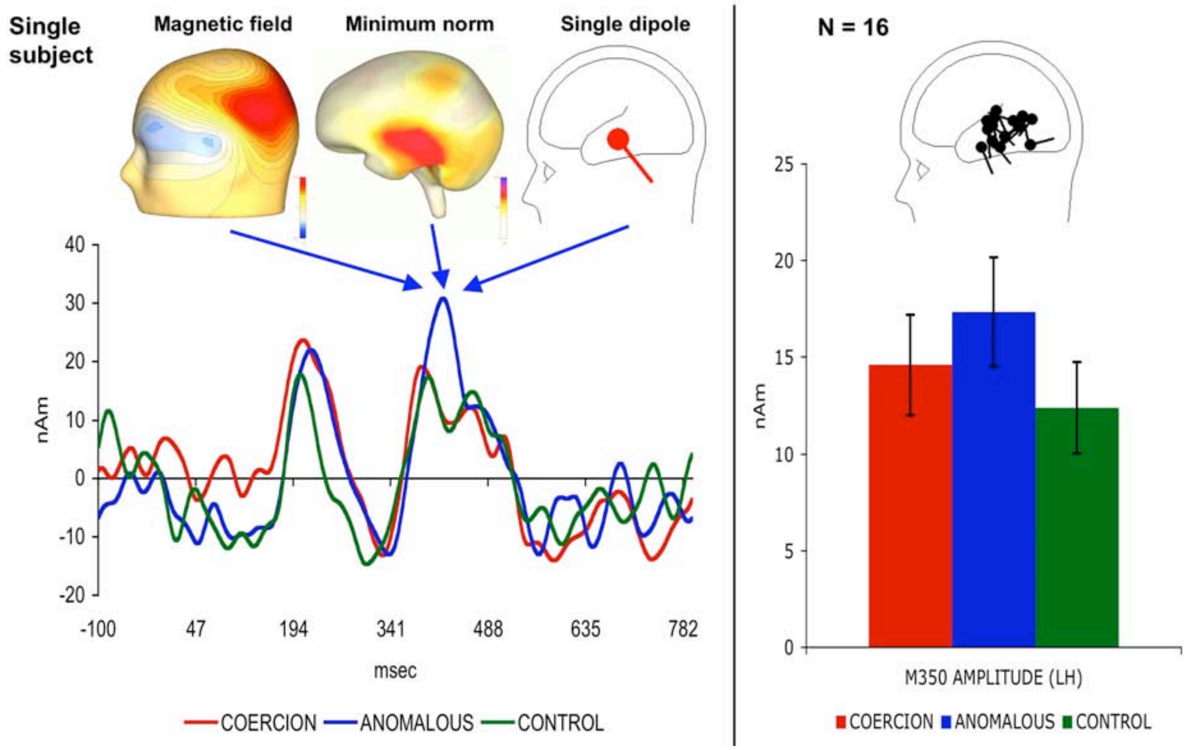

Visual M100
(90-150ms)
M170
(150-200ms)
M250
(200-300ms)
M350
(300-400ms)
AMF
(350-500ms)
Late temporal
(400-500ms)
Auditory
M100

Figure 2.



Single subject
Magnetic field
Minimum norm
Single dipole
N = 16
40
30
20
10
0
-10
-20
-30
nAm
-100
47
194
341
488
635
782
msec
COERCION
ANOMALOUS
CONTROL
40
35
30
25
20
15
10
5
0
nAm
AMF AMPLITUDE
COERCION
ANOMALOUS
CONTROL

Figure 3.



Single subject
Magnetic field
Minimum norm
Single dipole
40
30
20
10
0
-10
-20
nAm
-100
47
194
341
488
635
782
msec
COERCION
ANOMALOUS
CONTROL
N = 16
25
20
15
10
5
0
nAm
M350 AMPLITUDE (LH)
COERCION
ANOMALOUS
CONTROL

Figure 4.

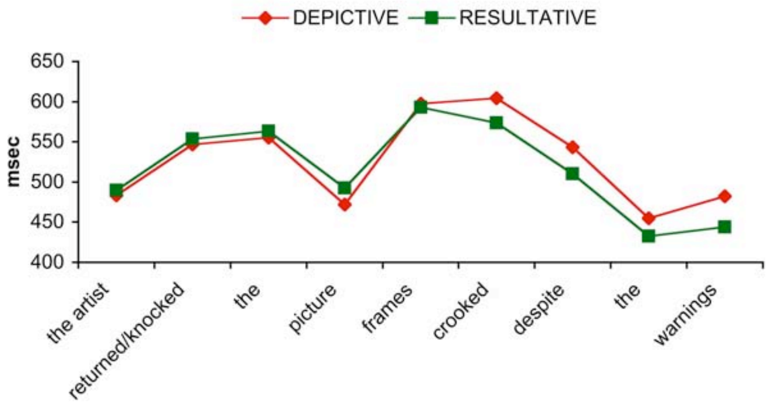
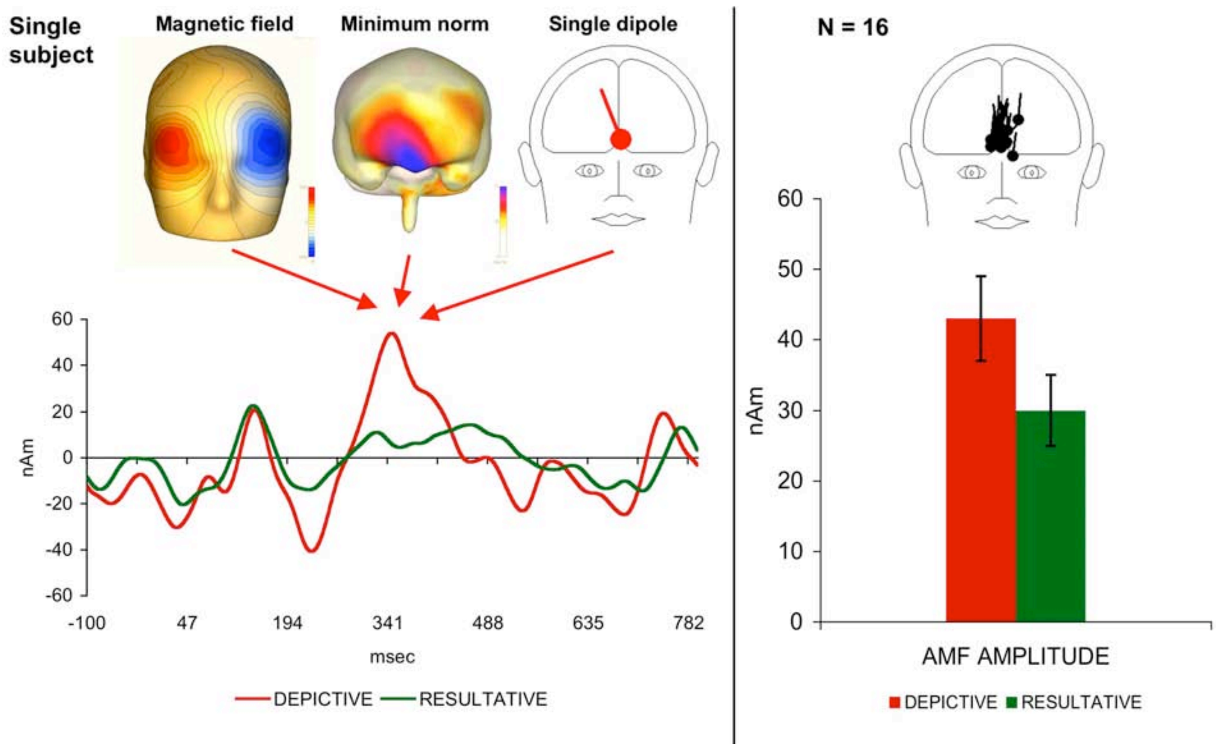

DEPICTIVE
RESULTATIVE
msec
650
600
550
500
450
400
the artist
returned/knocked
the
picture
frames
crooked
despite
the
warnings

Figure 5.


Single subject
Magnetic field
Minimum norm
Single dipole
60
40
20
0
-20
-40
-60
nAm
-100
47
194
341
488
635
782
msec
DEPICTIVE
RESULTATIVE
N = 16
60
50
40
30
20
10
0
nAm
AMF AMPLITUDE
DEPICTIVE
RESULTATIVE