# Electrophysiological Correlates of Complement Coercion

**Gina R. Kuperberg[1,2], Arim Choi[1], Neil Cohn[1], Martin Paczynski[1], and Ray Jackendoff[1]**

## Abstract

■ This study examined the electrophysiological correlates of complement coercion. ERPs were measured as participants read and made acceptability judgments about plausible coerced sentences, plausible noncoerced sentences, and highly implausible animacy-violated sentences ("The journalist began/wrote/astonished the article before his coffee break"). Relative to noncoerced complement nouns, the coerced nouns evoked an N400 effect. This effect was not modulated by the number of possible activities implied by the coerced nouns (e.g., began reading the article; began writing the article) and did not differ either in magnitude or scalp distribution from the N400 effect evoked by the animacy-violated complement nouns. We suggest that the N400 modulation to both coerced and animacy-violated complement nouns reflected different types of mismatches between the semantic restrictions of the verb and the semantic properties of the incoming complement noun. This is consistent with models holding that a verb's semantic argument structure is represented and stored at a distinct level from its syntactic argument structure. Unlike the coerced complement noun, the animacy-violated nouns also evoked a robust P600 effect, which may have been triggered by the judgments of the highly implausible (syntactically determined) meanings of the animacy-violated propositions. No additional ERP effects were seen in the coerced sentences until the sentence-final word that, relative to sentence-final words in the noncoerced sentences, evoked a sustained anteriorly distributed positivity. We suggest that this effect reflected delayed attempts to retrieve the specific event(s) implied by coerced complement nouns. ■



## INTRODUCTION

Although it is widely acknowledged that sentences are built compositionally, there is debate over whether their meaning is determined entirely through combining individual lexical items using syntactic rule systems (strong compositionality, e.g., Montague, 1970) or whether it is possible to construct new meaning that is invisible to syntactic structure. One piece of evidence for the latter hypothesis comes from a phenomenon known as complement coercion, exemplified by the sentence, "The man began the book" (Jackendoff, 1997; Pustejovsky, 1995). Under strong compositionality, verbs like "begin," which semantically select for an activity, should be unable to take arguments denoting entities such as "book." Nonetheless, we interpret such sentences as plausible. One account of this phenomenon is that eventive verbs, like "begin," "finish," and "enjoy," when paired with a complement NP entity ("book"), "type shift" that complement into an event to meet the demands of the argument structure: "book" is taken to mean "do something with the book" (Pustejovsky, 1995). An alternative account is that the unexpressed meaning ("do something") is inserted into the meaning of the sentence to satisfy the selectional restrictions of the verb (Jackendoff, 1997). As a result, "begin the book" is understood as "begin 'doing something with' the book" (for a review, see Pylkkänen & McElree, 2006).

It has been argued that this process of coercion should incur a processing cost. And, indeed, behavioral studies using self-paced reading (Traxler, Pickering, & McElree, 2002; McElree, Traxler, Pickering, Seely, & Jackendoff, 2001), eye tracking (Frisson & McElree, 2008; McElree, Frisson, & Pickering, 2006; Pickering, McElree, & Traxler, 2005; Traxler, McElree, Williams, & Pickering, 2005; Scheepers, Mohr, Keller, & Lapata, 2004; Traxler et al., 2002), and speed accuracy trade-off (McElree, Pylkkänen, Pickering, & Traxler, 2006) procedures report that coerced complement NPs are harder to process than noncoerced NPs, even when they are matched for plausibility (Traxler et al., 2002; McElree et al., 2001). Importantly, these costs are not incurred on all NPs after eventive verbs, such as those denoting activities (e.g., "begin the work"), but rather appear to arise from the particular combination of an eventive verb with an entity NP (Traxler et al., 2002).

One possibility is that the increased processing costs associated with coerced versus noncoerced NPs reflect a second step of developing a full specific interpretation—the filling out or retrieval of the details of "do something" based on real-world knowledge and context (e.g., "reading a book," "writing a book"; Jackendoff, 1997; Pustejovsky, 1995). This account, however, seems unlikely because such costs are still present when the activity is explicitly provided in the immediately preceding discourse context

[1]Tufts University, [2]Massachusetts General Hospital

(Traxler et al., 2005).[1] Finally, these costs are unlikely to be due to the process of selecting between retrieved alternative activities (resolving ambiguity): In a recent eye-movement study, Frisson and McElree (2008) reported equal costs for processing complement NPs that were highly constrained for a single interpretation (e.g., "The student started the essay," usually interpreted as "writing") and complement NPs that were less constrained (e.g., "The director started the script," which could be interpreted as depicting "reading," "directing," "filming," or other actions). Similarly, Scheepers, Keller, and Lapata (2008) reported data from a visual world paradigm that support a serial account of coercion in which a single dominant interpretation, rather than multiple interpretations, is pursued during processing. Together, these observations have been interpreted as supporting the view that the processing cost of coercion reflects the building of a complex nonsyntactic representation of the complement.

In a recent study, Pylkkänen and McElree (2007) used magneto-encephalography (MEG) to contrast activity to coerced complements, noncoerced complements, and complements that violated the selection restrictions (animacy-based) of the preceding verb (e.g., "The journalist began/wrote/astonished the article…"). The animacy-violated complement NPs were associated with a significant MEG effect, relative to the noncoerced complements, from 300 to 400 msec, localizing to a left temporal source. In contrast, the coerced NPs were associated with a significant anterior midline effect, relative to both the noncoerced and the animacy-violated complement NPs, between 350 and 500 msec, which localized to a ventromedial prefrontal source. Because there was no difference in activity at this source between the noncoerced and the animacy-violated complements, the authors interpreted these observations as evidence that complement coercion engages neurocognitive processes distinct from those engaged in detecting lexical mismatch, semantic predictability, or semantic implausibility. A similar anterior midline effect has been described by the same group in association with other forms of coercion where it has been interpreted as a more general neural signature of enriched composition (Pylkkänen, Martin, McElree, & Smart, 2009; Brennan & Pylkkänen, 2008).

Taken together, this series of studies provides compelling evidence that complement coercion entails a behavioral and neural cost. The present study used another technique—ERPs—to seek converging evidence and further information on the neurocognitive processes engaged during complement coercion.

In ERP studies, the component that has been most closely linked to semantic processing is the N400—a negative-going waveform observed approximately between 300 and 500 msec after words that are incongruous (vs. congruous) with their preceding word (Bentin, McCarthy, & Wood, 1985; Rugg, 1985), sentence (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kutas & Hillyard, 1980, 1984), or discourse (Van Berkum, Hagoort, & Brown, 1999) contexts. Although it has sometimes been assumed that the N400 during sentence processing is a reflection of semantic anomaly or implausibility per se, it has been recognized for some time that an N400 effect is evoked by words that are plausible but relatively unexpected with respect to their preceding context (Kutas & Hillyard, 1984) and that the N400 amplitude is modulated by a host of factors that can influence plausibility but that can be theoretically dissociated from this construct. These include fine-grained associative relationships between individual words (Van Petten, 1993), coarser-grained categorical relationships between entities sharing common features (Federmeier & Kutas, 1999), and, in the case of verb–argument structures, selection restriction-based relationships (Friederici & Frisch, 2000) and animacy-based relationships (Paczynski & Kuperberg, 2009; Frisch & Schlesewsky, 2001; Weckerly & Kutas, 1999).

The N400 evoked by a word reflects the cost of semantically processing that word. Its amplitude is modulated by a three-way dynamic interaction between (a) its semantic features, (b) semantic relationships within that word's context, and (c) semantic relationships stored at various grains of representation within semantic memory (Lau, Phillips, & Poeppel, 2008; Kutas, Van Petten, & Kluender, 2006; Kutas & Federmeier, 2000). Predictions as to the nature of the incoming critical word may be generated before it has been presented (see Federmeier, 2007; DeLong, Urbach, & Kutas, 2005; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005), or matching processes may take place only once the critical word has been encountered ("semantic integration"; Hagoort, 2005; Holcomb, 1993; for a recent discussion, see Van Berkum, 2009). In the case of verb–argument processing, we have contrasted these types of "semantic memory-based processes" with independent but interacting "combinatorial" stream(s) of processing that come up with full propositional representations or interpretations that may be plausible or implausible with respect to real-world knowledge (Kuperberg, 2007). Such combinatorial process(es) are often syntactically driven (e.g., through assigning thematic roles to arguments).

The present study compared the ERP responses evoked by coerced, noncoerced, and animacy-violated complement NPs. Similar to the MEG study by Pylkkänen and McElree (2007), participants carried out an acceptability judgment task as they viewed these sentences. On the basis of this MEG study, we predicted that, relative to noncoerced NPs, coerced NPs would evoke increased activity within the N400 time window (300–500 msec). Although this effect could theoretically reflect the interpretative process of type shifting the meaning of the entity NP to an event, an alternative possibility, given the discussion of the N400 above, is that it might simply reflect the mismatch between the semantic properties of the eventive verb and the semantic properties of the incoming NP (an entity).

On the basis of previous studies demonstrating an N400 effect to object NPs that violate the selection restriction properties of their preceding verbs (Friederici & Frisch, 2000), we predicted that, relative to the noncoerced NP, the animacy-violated complement NPs would also produce an N400 effect. Rather than reflecting the implausibility of the proposition produced by syntactically and thematically combining the animacy-violated NP with the verb, this N400 might, once again, reflect a mismatch—this time between the animate selection restrictions of the preceding verb and the semantic features of the direct object NP argument (an inanimate entity).

One question was whether the N400 effects evoked by the coerced and animacy-violated complement NPs, each relative to noncoerced NPs, would have distinct scalp distributions. As noted above, this was the case in the MEG study by Pylkkänen and McElree (2007). Although, the spatial resolution of ERPs is inferior to that of MEG because of its sensitivity to the effects of intervening tissues that smear the EEG patterns measured on the scalp (Cooper, Winter, Crow, & Walter, 1965; Delucchi, Garoutte, & Aird, 1962; Geisler & Gerstein, 1961), differences in the spatial distribution on the surface of the scalp of the N400 effects evoked across different experimental conditions have been well documented in other studies (e.g., Sitnikova, West, Kuperberg, & Holcomb, 2006; West & Holcomb, 2002; Holcomb, Kounios, Anderson, & West, 1999; Kounios & Holcomb, 1992). On the basis of the MEG findings, we therefore hypothesized that the N400 effect to coerced (vs. noncoerced) NPs would show a more anterior distribution than to that evoked by the animacy-violated (vs. noncoerced) NPs.

A second question was whether the neural costs associated with processing coerced complements would be sensitive to any ambiguity in their interpretation. To examine this possibility, we followed Frisson and McElree (2008) by carrying out separate ratings that were used to categorize the coerced sentences into those with dominant interpretations (e.g., "The student started the essay…") and those with multiple possible interpretations (e.g., "The director started the script…"). If any N400 effect evoked by the coerced (vs. noncoerced) NPs reflected a process of selecting from multiple possible activities, then it should be modulated by this parameter.

A third question was whether ERPs would be modulated across conditions in the time window after the N400, affecting the P600—a centro-parietally distributed positive-going component observed approximately between 500 and 900 msec. The P600 has been classically associated with syntactic anomalies and ambiguities (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992) but, more recently, has been described in association with certain types of semantic anomalies under particular circumstances (for a review, see Kuperberg, 2007). Unlike the N400 effect, the semantic P600 effect is not usually seen to words that are simply unpredictable in plausible sentences but primarily to words that are semantically very implausible or impossible—often violations of animacy. It is more likely to be evoked when the verb or the wider context is semantically constraining and when participants are asked to make acceptability judgments and may reflect continued combinatorial processes as participants attempt to make sense of the sentences. In the present study, we therefore predicted that a P600 effect would be observed to the animacy-violated complement NPs, but not to the coerced complements (each relative to the noncoerced complements).

Finally, in addition to examining neural activity at the point of the complement NP, we examined ERPs to subsequent words in the sentence. Self-paced reading and eye-tracking studies have reported behavioral costs of coercion at one or two words after the complement (Frisson & McElree, 2008; Traxler et al., 2002, 2005; McElree et al., 2001), but have not examined processing beyond this point. Pylkkänen and McElree's (2007) MEG study focused on neural costs at the complement itself but did not examine past this point. We therefore aimed to determine whether the neural costs of coercion are primarily met at the complement NP or whether additional costs are incurred at subsequent words, particularly on sentence-final words (SFWs) where "wrap-up" of sentence meaning is thought to take place.

## METHODS

### Development and Pretesting of Materials

Six hundred thirty sentences (210 scenarios, each with three sentence types) were developed and expanded from 70 scenario triplets originally used by Pylkkänen and McElree (2007) in their MEG study. In this original set of 70 triplets, a sentence in a given triplet contained one of three types of verbs—noncoercive (entity selecting), coercive (event selecting), and object experiencer—followed by the same inanimate critical complement NP that rendered the sentences noncoerced, coerced, and animacy violated, respectively (Table 1). The direct object complement was followed by between three and five words, followed by the sentence final word (SFW). The original set was then expanded threefold to counterbalance the identical critical nouns across three lists as follows: for each of the 70 original triplet scenarios, three additional sentences were created using the same verbs but new subjects and objects. This replacement of subjects and object NPs was repeated once more, yielding nine sentences per scenario: three sentences with a given coercive verb, three with a given noncoercive verb, and three with a given object-experiencer verb, but with the same subject and object NPs appearing only once with each type of verb. In half of the original scenarios used by Pylkkänen and McElree (2007), the clause containing the critical verb and the complement noun was embedded within a relative clause (e.g., "The staff was shocked that the journalist began the article before his coffee break…"). This varied the length of the sentences and the position at

**Table 1.** Stimuli Examples and Parameters for Each Sentence Type

| *Sentence Type (Example)* | *Naturalness/Plausibility of Entire Sentence* | *LSA: At Point of CN* | *Cloze Probability of CN* | *No. Letters of Verb* | *Frequency of Verb* |
|---|---|---|---|---|---|
| *Noncoerced* | | | | | |
| The journalist *wrote* the <u>article</u> before his coffee break. | 3.8 (0.5) | 0.17 (0.14) | 0.14 (0.27) | 6.3 (1.7) | 59.4 (109.05) |
| *Coerced* | | | | | |
| The journalist *began* the <u>article</u> before his coffee break. | 3.8 (0.6) | 0.14 (0.10) | 0.06 (0.15) | 7.0 (1.6) | 102.5 (95.79) |
| *Animacy Violated* | | | | | |
| The journalist *astonished* the <u>article</u> before his coffee break. | N/A | 0.12 (0.10) | – | 7.8 (1.3) | 15.7 (17.64) |

Data are presented as mean (SD). In the examples of each sentence type, the verb is shown in italics and the critical complement NP (to which ERPs were measured) is underlined. CN = complement noun.

which the coercion and animacy violations were introduced, thereby introducing variety into the material. To keep the overall experimental stimuli similar to the MEG study, we maintained this manipulation in the present stimulus set. However, we did not expect that it should systematically influence ERP responses to any of our main manipulations and therefore did not include this as a factor in our main series of analyses.

We then carried out a norming study with this initial set of coerced and noncoerced sentences (randomized across three lists) in order to screen out the more unnatural or implausible coerced sentences. We excluded the animacy violated sentences in this norming study. Twelve undergraduates from the Tufts University (four for each list), who did not participate in the ERP study and who gave written informed consent before participation, judged the likelihood that they might encounter each sentence (presented as a whole) in the real world using a scale from 1 (*the sentence did not make sense and/or it sounded unnatural*) to 5 (the sentence made sense and seemed natural). On the basis of these initial ratings, 30 scenarios with an average rating of less than 3 were discarded, leaving a final set of 180 scenarios.

To generate the final stimulus lists used in the ERP experiment, we reinserted the appropriate animacy-violated sentences for each scenario. This yielded 540 sentences in total, counterbalanced across three lists, each list with 180 sentences, 60 of each sentence type. Across all lists, each subject and NP combination was paired with all three types of verbs (i.e., seen in all three sentence types), but within any given list, the same combination (pairing) of subject and complement NPs was not viewed with more than one type of verb (i.e., in more than one condition) except on two occasions. Eighty-four of 180 scenarios contained a relative clause. In each list, test sentences were pseudorandomized among 158 filler sentences, 50 that contained semantic incongruities. The incongruous filler sentences contained a variety of different types of incongruities ranging from animacy violations (e.g., "The whistler trained the chapstick so his lips wouldn't chap."), other types of selection restriction violations (e.g., "The congressman smoldered the meeting until the food ran out."), and pragmatic real-world incongruities (e.g., "The girl smiled at the parking meter to make sure she had enough time."). Six of 158 fillers contained object-experiencer verbs. Thus, in total, each list contained 338 sentences and approximately 33% of these contained semantic incongruities.

This final stimulus set was further characterized in terms of several metrics and by conducting a cloze study. In the cloze study, the coercive and noncoercive sentence frames (without the critical words) were presented on a computer to 30 undergraduates at the Tufts University (10 per list) who did not participate in the ERP experiment or any other rating study. Participants gave written, informed consent before participation and were asked to type in the most likely next word in the sentence.

The results of all norming and stimulus characterizations are shown in Table 1. Coercive verbs were more frequent than the noncoercive verbs, $t(335) = 3.86$, $p < .01$, that were in turn more frequent than the object-experiencer verbs, $t(296) = 4.71$, $p < .01$. The object-experiencer verbs were slightly longer than the coercive verbs, $t(358) = 5.57$, $p < .01$, which were longer than the noncoercive verbs, $t(358) = 3.84$, $p < .001$. A latent semantic analysis (LSA; a measure of lexical co-occurrence)—calculated using pairwise comparisons of semantic similarity values term by term between each complement noun and all content words that preceded it (Landauer, Foltz, & Dumais, 1998; Landauer & Dumais, 1997)—yielded very slightly greater values in the noncoerced complement nouns (0.17) than the coerced sentences (0.14), $t(358) = 2.06$, $p < .04$. Cloze probabilities of both the coerced and the noncoerced complement nouns were low (less than 15%) but greater in the noncoerced sentences (0.14) than that in the coerced sentences (0.06) by subjects, $t(29) = 8.25$, $p < .01$, and by items, $t(179) = 4.62$, $p < .01$. Plausibility ratings of the entire coerced and noncoerced sentences, gathered during the development of stimuli as described earlier, did not differ significantly by subjects, $t(11) = 0.312$, $p = .76$, or by items, $t(418) = 1.34$, $p = .18$.

### *Subdivision of Coerced Sentences*

Following Frisson and McElree (2008), we carried out an additional rating study to examine the precise interpretations of the activities implied by the coerced NPs in each sentence. Thirty undergraduates from the Tufts University (10 per counterbalanced list) were given the coerced sentences used in the ERP experiment with a blank space in between the verb and the complement NP (e.g., "The journalist began ___ the article before his coffee break."). Participants were asked to fill in the blank with one or two words describing the activity that best fit their interpretation of the sentence. We identified the number of unique interpretations for each sentences. Sentences in which the same interpretation was given in 80% or more of all responses were categorized as having a strongly preferred (dominant) interpretation ($n = 89$); all others (70% or less) were categorized as having weakly preferred interpretations ($n = 91$). We also examined three other measures identified by Frisson and McElree (2008): (1) the number of different verbs generated for the sentence; (2) the number of unique interpretations generated for the sentence; and (3) the ratio of the most frequent interpretation of the sentence to the second-most frequent interpretation. The data are shown in Table 2.

## ERP Experiment

### *Participants*

Twenty-six undergraduates (9 men, 17 women) from Tufts University aged 18 to 22 years (mean age = 19.6 years) initially participated, and 24 subjects (9 men, 15 women, mean age = 19.5 years) were included in the final analysis (see below). All selected participants were right-handed, native American English speakers who had not learned to speak another language fluently before the age of 5 years. Participants were not taking any medication, had normal or corrected-to-normal vision, and had no history of a reading disability or of neurological or psychiatric disorders. Written consent was obtained from all subjects before participation according to the established guidelines of Tufts University.

### *Stimulus Presentation*

Participants were randomly assigned to one of the three counterbalanced lists. They sat in a comfortable chair in a dimly lit room separate from the experimenter and computers. Sentences were presented word by word on a computer monitor. Each trial (one sentence) began with a fixation point ("+") at the center of the screen for 450 msec, followed by a 100-msec blank screen, followed by the first word of the sentence. Each word appeared on the screen for 450 msec with an ISI of 100 msec separating the words. The final word of each sentence appeared with a period and was followed by a 750-msec blank-screen interval and then a question mark (?). This cue remained on the screen until the participant made his or her response, at which point the next trial started. The participant's task was to decide whether each sentence made sense by pressing one of two buttons on a response box with either the left or the right thumb (counterbalanced across participants). Participants were instructed to wait until the "?" cue before responding. This delayed response was designed to reduce any contamination of the ERP waveform by response-sensitive components such as the P300 (Donchin & Coles, 1988). After subjects registered their responses, the word "begin" was displayed until they pressed a button to begin the next trial. Each participant was given 12 practice trials at the beginning of the experiment.

### *Electrophysiological Recording*

Twenty-nine active tin electrodes were held in place on the scalp by an elastic cap (Electro-Cap International, Inc.,

**Table 2.** Parameters Used to Subdivide the Coerced Sentences into Those with One Dominant Interpretation and Those with Multiple Possible Interpretations

| *Measure* | *Dominant Interpretation (n = 89)* | *Multiple Interpretations (n = 91)* |
|---|---|---|
| % use of the dominant interpretation | 90.1% (80–100%) | 54.3% (30–70%) |
| No. different verbs generated | 3.5 (1–7) | 5.56 (2–9) |
| Average number of different interpretations | 1.82 (1–3) | 3.97 (2–8) |
| Ratio of the most frequent interpretation to the second-most frequent interpretation | 13:1 | 2.4:1 |

The range of values for each measure is shown in brackets.

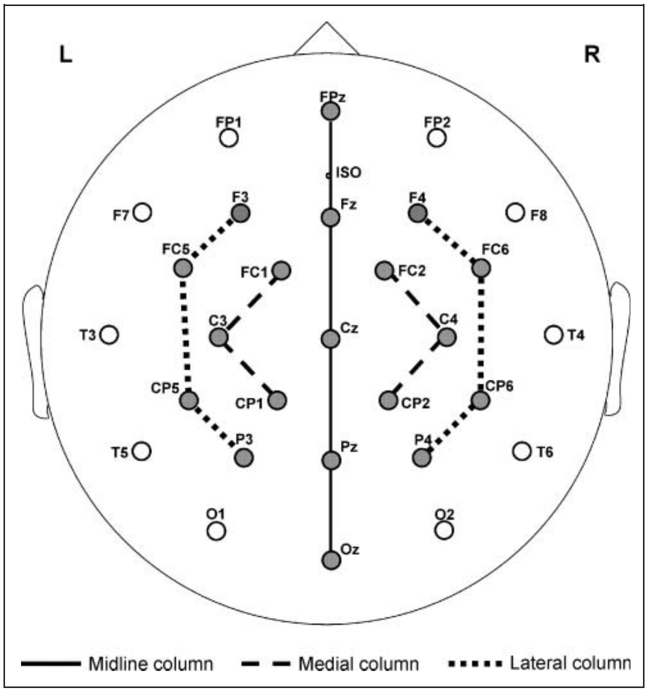



**Figure 1.** Electrode montage. ANOVAs based on a priori hypotheses were conducted at each of the three columns shown (midline, medial, and lateral).

Eaton, OH; see Figure 1). Electrodes were also placed below the left eye and at the outer canthus of the right eye to monitor vertical and horizontal eye movements and on the left and right mastoids. Impedance was kept less than 5 kΩ for all scalp electrode sites, at 2.5 kΩ for mastoid electrode sites, and less than 10 kΩ for the two eye channels. The EEG signal was amplified by an Isolated Bioelectric Amplifier System Model HandW-32/BA (SA Instrumentation Co., San Diego, CA) with a band-pass filter of 0.01 to 40 Hz and was continuously sampled at 200 Hz by an analogue-to-digital converter. The stimuli and the behavioral responses were simultaneously monitored by a digitizing computer.

*Data Analysis*

Accuracy was computed as the percentage of correct responses. A correct response was a judgment of acceptable for the noncoerced and coerced sentences and unacceptable for the animacy-violated sentences.

Averaged ERPs, time locked to target words, were formed off-line from trials free of ocular and muscular artifact and were quantified by calculating the mean amplitude (relative to a 100-msec prestimulus baseline) in time windows of interest. Because of our a priori hypotheses, we proceeded straight to planned pairwise comparisons between conditions of interest (coerced vs. noncoerced vs. animacy violated). We conducted ANOVAs at a midline column, containing five electrode sites and two lateral columns, each containing three (medial column) or four (lateral column) electrodes (see Figure 1). Within-subject factors were Sentence Type, Anterior–Posterior (AP) Distribution (with the number of levels corresponding to electrode sites along the AP axis), and, for the lateral analyses, Hemisphere (two levels).

Our series of ANOVAs yielded statistical information about differences in the distribution of effects along the AP axis of the scalp and across the two hemispheres. Main effects and interactions involving sentence type, which were of most theoretical interest, were followed up using appropriate simple effects ANOVAs. The N400 at the complement noun was quantified from 300 to 500 msec, and the P600 was quantified between 600 and 900 msec (to avoid overlap with the N400 effect). For each of the three words after the complement noun, 400- to 600-msec time windows were used for analyses. For the SFW, a 300- to 700-msec time window was used for analyses.

In all these ANOVAs, the Greenhouse–Geisser correction was used in cases with more than one degree of freedom in the numerator (Greenhouse & Geisser, 1959) to protect against Type 1 error resulting from violations of sphericity. In these cases, we reported the original degrees of freedom with the corrected $p$ value. In all analyses, an alpha level of .05 was used because we were testing *a priori* hypotheses. Linearly interpolated voltage maps showing the scalp distribution of differences in ERPs elicited by critical words between the three conditions within the time windows of interest were produced by the EEGLab program (MatLab; MathWorks, Natick, MA).

## RESULTS

### Behavioral Data

One participant was excluded on the basis of a clear behavioral response bias. One other participant was excluded because of ERP artifact. Of the 24 remaining participants, accuracy on the acceptability judgment task was high: coerced and noncoerced sentences were correctly identified as acceptable on 91.3% ($SD$ = 5.4) and 92.1% ($SD$ = 4.1) of trials, respectively. Animacy-violated sentences were correctly identified as unacceptable on 95.8% ($SD$ = 5.4) of trials. Accuracy judgments significantly differed between sentence types, $F(2, 46) = 8.78$, $p < .01$, because of more accurate judgments to the animacy-violated sentences than to both the coerced and the noncoerced sentences ($p < .001$ and $p < .01$ for pairwise comparisons, respectively).[2] There was no significant difference in accuracy between the coerced and the noncoerced sentences ($p > .1$).

### ERP Data

Across the 24 participants included in the analysis, approximately 11% of the critical trials were rejected because of the artifact. All ERP analyses reported are based on correctly answered trials. However, analyses were repeated, including all responses, and yielded qualitatively similar findings.

### *ERPs on the Complement Noun*

Grand average ERPs elicited by the complement nouns for all sentence types at selected electrode sites are presented in Figure 2. There were no significant differences in the N1–P2 complex over the first 250 msec after the onset of the critical word across conditions (no main effects or interactions involving sentence type, *p*s > .05).

*300–500 msec: The N400.* A significantly more negative N400 was observed to both the coerced and the animacy-violated complement nouns than noncoerced nouns (Table 3). The amplitude of the N400 to the coerced and animacy-violated complement nouns did not differ significantly from each other (Table 3 and Figure 2). N400 effects to both coerced and violated (relative to noncoerced) complement nouns were fairly widely distributed across the scalp (no interactions between sentence type and AP distribution).[3] An analysis that included a subset of the main stimulus set in which LSA and cloze probability at the point of the complement noun were all fully matched between the coerced and the noncoerced sentence types revealed a similar set of findings (see note to Table 3 and http://www.nmr.mgh.harvard.edu/kuperberglab/publications/materials/ComplementCoercion_suppl_figures.pdf). A secondary analysis that included the presence or absence of a relative clause (in just less than 50% of stimuli) as an additional within-subject variable confirmed that this did not interact significantly with sentence type at any electrode column, all *p*s > .05.

*600–900 msec: The P600.* The P600 was larger to animacy violated than to both noncoerced complement nouns (Table 3B and Figure 2, right bottom) and coerced complement nouns (Table 3C). This P600 effect was generally more positive posteriorly than anteriorly and had a slight left-lateralized distribution (main effects of sentence type and/or sentence type by AP distribution or hemisphere interactions at all columns). In contrast, there was no difference in the amplitude of the P600 evoked by coerced and noncoerced complement nouns (no significant main effects or interactions involving sentence type in any column; Table 3A and Figure 2, right top).

### *ERPs to Words after the Complement NP*

In comparing the coerced and the noncoerced sentences, the waveforms evoked by each of the three words that followed the complement nouns did not diverge from one another (Figure 3 and Table 4). In contrast, the positivity evoked by the animacy-violated complement

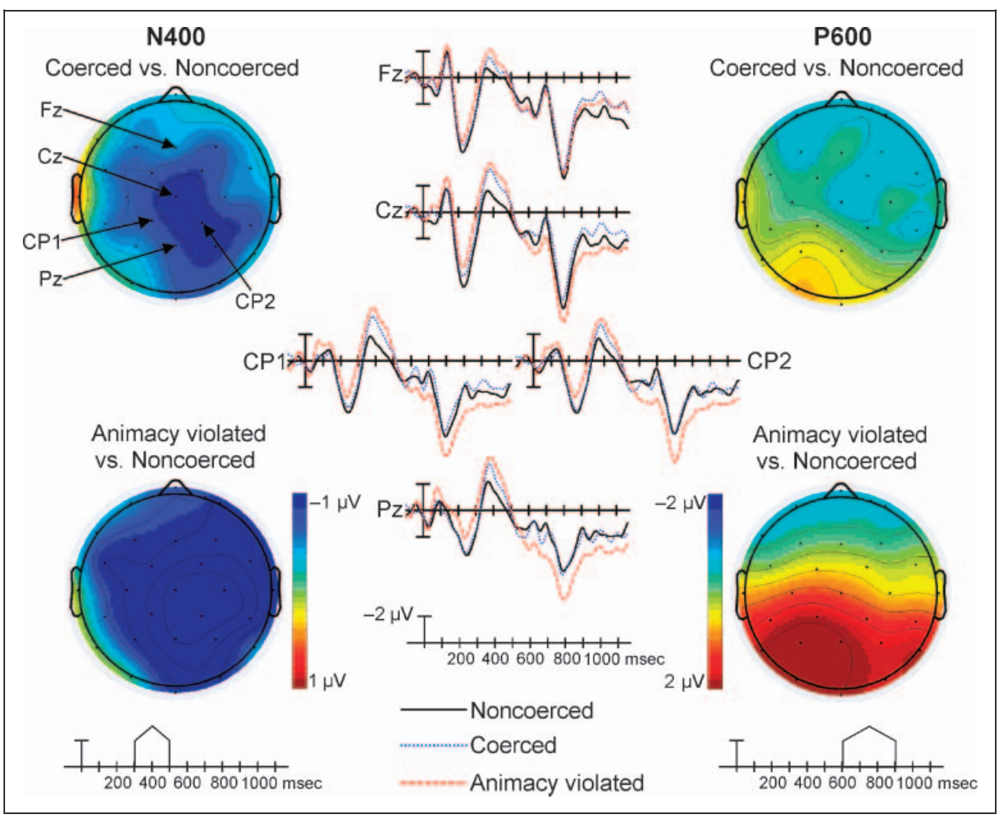



**Figure 2.** Grand-averaged waveforms to complement NPs in all three sentence types. Voltage maps comparing ERPs evoked by the complement noun between 300 and 500 msec—the N400 effect (left) and between 600 and 900 msec—the P600 effect (right). Note that to best illustrate the full scalp distribution of the ERP effects, the scale used for the voltage maps of the N400 effects (left) was half of that used for the voltage maps of the P600 effect (right). Grand-averaged waveforms to a cloze-matched data set demonstrate similar results and are available at http://www.nmr.mgh.harvard.edu/kuperberglab/publications/materials/ComplementCoercion_suppl_figures.pdf.

**Table 3.** ANOVAs Comparing ERPs to Complement Nouns across the N400 (300–500 msec) and the P600 (600–900 msec) Time Windows

| | | *N400: 300–500 msec* | | *P600: 600–900 msec* | |
|---|---|---|---|---|---|
| *Effect* | | *F (df)* | *p* | *F (df)* | *p* |
| *A. Coerced versus Noncoerced* | | | | | |
| Midline | ST | 5.71 (1, 23) | .025 | 0.09 (1, 23) | .763 |
| | ST × AP | 1.02 (4, 92) | .362 | 2.32 (4, 92) | .114 |
| Medial | ST | 5.85 (1, 23) | .024 | 0.44 (1, 23) | .514 |
| | ST × H | 0.64 (1, 23) | .433 | 0.87 (1, 23) | .361 |
| | ST × AP | 0.71 (2, 46) | .492 | 0.55 (2, 46) | .566 |
| Lateral | ST | 3.84 (1, 23) | .062 | 0.38 (1, 23) | .545 |
| | ST × H | 0.97 (1, 23) | .335 | 1.35 (1, 23) | .257 |
| | ST × AP | 0.31 (3, 69) | .651 | 1.70 (2, 69) | .195 |
| *B. Animacy Violated versus Noncoerced* | | | | | |
| Midline | ST | 8.72 (1, 23) | .007 | 4.37 (1, 23) | .048 |
| | ST × AP | 2.29 (4, 92) | .105 | 7.15 (4, 92) | .004 |
| Medial | ST | 9.23 (1, 23) | .006 | 4.35 (1, 23) | .048 |
| | ST × H | 13.61 (1, 23) | .001 | 5.98 (1, 23) | .023 |
| | ST × AP | 1.24 (2, 46) | .294 | 7.05 (2, 46) | .009 |
| Lateral | ST | 7.82 (1, 23) | .010 | 4.64 (1, 23) | .042 |
| | ST × H | 3.19 (1, 23) | .087 | 0.25 (1, 23) | .619 |
| | ST × AP | 0.36 (2, 69) | .656 | 14.02 (2, 69) | .0002 |
| *C. Coerced versus Animacy Violated* | | | | | |
| Midline | ST | 2.41 (1, 23) | .135 | 4.49 (1, 23) | .045 |
| | ST × AP | 0.61 (4, 92) | .562 | 5.82 (4, 92) | .008 |
| Medial | ST | 2.41 (1, 23) | .135 | 5.66 (1, 23) | .026 |
| | ST × H | 2.37 (1, 23) | .138 | 0.25 (1, 23) | .625 |
| | ST × AP | 0.65 (2, 46) | .493 | 2.85 (2, 46) | .075 |
| Lateral | ST | 1.59 (1, 23) | .219 | 6.07 (1, 23) | .022 |
| | ST × H | 0.54 (1, 23) | .470 | 0.25 (1, 23) | .624 |
| | ST × AP | 0.40 (2, 69) | .647 | 7.74 (2, 69) | .002 |

ST = main effect of sentence type; ST × H = Sentence Type × Hemisphere interaction; ST × AP = Sentence Type × Anterior–Posterior distribution.

A subanalysis performed on a cloze-matched data set produced comparable results, with significant differences in N400 modulation between the coerced and the noncoerced NPs at medial and lateral columns, $ps < .05$, but no significant differences in the N400 evoked by coerced and animacy-violated NPs ($ps > .1$). Please refer to the supplementary figure demonstrating these effects at http://www.nmr.mgh.harvard.edu/kuperberglab/publications/materials/ComplementCoercion_suppl_figures.pdf.

nouns, relative to the other two conditions, remained evident at the first word after the complement noun (Figure 3 and Table 4, CN+1). At the second and third words after the animacy-violated complement noun, however, the polarity of this effect reversed such that the waveforms to these words were more negative than that in the coerced and noncoerced sentences (Figure 3 and Table 4, CN+2, CN+3).

At the SFW, the coerced sentences evoked a more positive (less negative) waveform than the noncoerced sentences at anterior sites, and the animacy-violated sentences evoked a more negative waveform than the noncoerced sentences at posterior sites (see Figure 4 and Table 5).

### *ERPs in the Coerced Sentences: Effects of Interpretational Ambiguity*

ERPs were separately averaged in the coerced sentences with dominant ($n$ = 30 per list, on average) and with multiple interpretations ($n$ = 30 per list, on average; for parameters of this subdivision, see Table 2). After artifact rejection, there remained, on average, 24 trials in each of these two conditions. As shown in Figure 5, there appeared to be no divergence at all in the waveforms evoked by these two types of sentences either at the complement noun or at the SFW. This was confirmed by ANOVAs conducted between 300 and 500 msec after the onset of complement nouns and between 300 and 700 msec after the onset of SFWs, which showed no significant main effects or interactions involving sentence type (all $F$s < 2.52 and all $p$s > .12).

## DISCUSSION

This study aimed to examine the electrophysiological correlates of processing coerced complement NPs that violated the semantic structural specifications of their preceding verbs. NPs denoting entities (e.g., "book") that were preceded by verbs that selected for complements denoting activities (e.g., "began") evoked a larger N400 than when the same NPs were preceded by entity-selecting verbs (e.g., "wrote"). An N400 effect of the same magnitude was evoked by entity NPs that violated the animacy selection restrictions of their preceding verbs (e.g., "pleased the book"). Unlike the coerced NPs, the animacy-violated NPs were highly implausible and also evoked a robust later positivity—a P600 effect.

The neural response across the three conditions also differed as the sentences unfolded word by word after the complement NP. On the word after the animacy-violated complement NP, the positivity effect was still present, but on the subsequent word, the waveform flipped to a posteriorly distributed negativity effect (relative to both other conditions) that continued up to and including the SFW. In contrast, there was no divergence in the waveform to words after the coerced and the

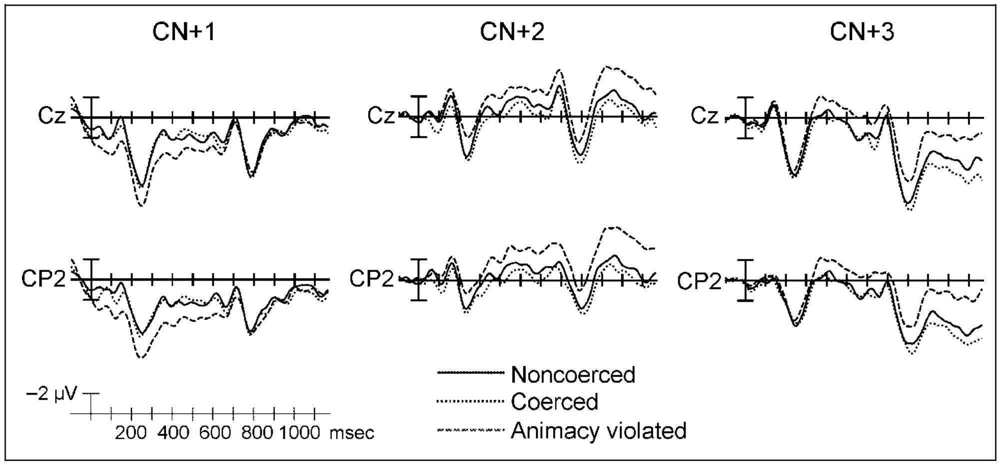



**Figure 3.** Grand-averaged waveforms at one, two, and three words after the complement noun (CN), comparing all three sentence types.

noncoerced complement NPs until the end of the sentence: relative to the SFWs of the noncoerced sentences, the SFWs of the coerced sentences produced a prolonged anteriorly distributed positivity effect.

In the following discussion, we considered each of these effects in relation to previous studies examining complement coercion and in relation to what we know more generally from ERP studies about their functional significance.

## Modulation of the N400 on the Complement NP

Our demonstration of increased neural costs to entity NPs after event-selecting verbs is consistent with the series of reading time studies mentioned in the Introduction section, which also reported processing costs in association with complement coercion (Frisson & McElree, 2008; McElree, Frisson, et al., 2006; McElree, Pylkkänen, et al., 2006; Pickering et al., 2005; Traxler et al., 2002, 2005; Scheepers et al., 2004; McElree et al., 2001). Our finding of neural modulation primarily between 300 and 500 msec is also consistent with Pylkkänen and McElree's (2007) MEG study, which reported neuromagnetic modulation between 350 and 500 msec in this contrast (although, as discussed further below, the ERP and MEG effects differed in their scalp distribution). Finally, these findings are consistent with a very recent study examining ERP correlates of complement coercion using similar stimuli to those used here and which also found an N400 effect to coerced (vs. noncoerced) complement NPs (Baggio, Choma, van Lambalgen, & Hagoort, 2010).

As in previous studies, these costs are unlikely to be fully accounted for by systematic differences between the coerced and the noncoerced NPs in their cloze probabilities or their semantic co-occurrences with their preceding content words (LSA values; Scheepers et al., 2004; McElree et al., 2001). In the present study, although these values did differ slightly between the coerced and the noncoerced conditions, the differences were small (much less than those between the animacy-violated and noncoerced NPs), and the N400 effect remained significant when we reanalyzed our data using a subset of the stimuli in which these factors were fully matched.

In previous self-paced reading and eye-movement studies, the processing cost on coerced complements has often been interpreted within the theoretical framework proposed by Pustejovsky (1995); that is, as reflecting the semantic work of type-shifting the complement from an entity to an activity (e.g., "book" to "reading a book") so as to reach a plausible interpretation of the event (Frisson & McElree, 2008; McElree, Pylkkänen, et al., 2006; Pickering et al., 2005; Traxler et al., 2005; Scheepers et al., 2004; McElree et al., 2001). Here, we suggested a slightly different interpretation: that, rather than indexing the work of type shifting, the N400 to the coerced complement reflected the mismatch between the semantic properties of the verb and those of the complement.

On this account, verbs such as "begin" and "finish" are stored in association with their particular semantic argument structures—their selection for events rather than entities. When an argument that matches this semantic argument structure is encountered, processing is facilitated, leading to an attenuation of the N400, compared with when arguments are encountered that mismatch this argument structure. This attenuation might result from a "preactivation" of eventive semantic frames, leading to the "prediction" of the upcoming argument as an activity rather than an entity (for evidence that the N400 can reflect the result of such predictive processing, see Federmeier, 2007; DeLong et al., 2005; Van Berkum et al., 2005), or it might result from facilitation after the presentation of the complement ("semantic integration"; Hagoort, 2005; Holcomb, 1993).

The mismatch between the verb and the complement may have been associated with implicit attempts to retrieve information from memory (inferencing), and this may have also contributed to N400 modulation (for a related interpretation, see Baggio et al., 2010). We suggested that such implicit memory-based inferencing was fairly course grained and limited to retrieving a general event schema (e.g., "begin 'doing something with' the book"; for a theoretical account, see Jackendoff, 1997,

**Table 4.** ANOVAs Comparing ERPs to 1, 2, and 3 Words after the Complement Noun (400- to 600-msec Time Window)

| | | *CN+1: 400–600 msec* | | *CN+2: 400–600 msec* | | *CN+3: 400–600 msec* | |
|---|---|---|---|---|---|---|---|
| *Effect* | | *F (df)* | *p* | *F (df)* | *p* | *F (df)* | *p* |
| *A. Coerced versus Noncoerced* | | | | | | | |
| Midline | ST | 0.74 (1, 23) | .399 | 1.89 (1, 23) | .182 | 0.32 (1, 23) | .575 |
| | ST × AP | 3.18 (4, 92) | .052 | 0.54 (4, 92) | .593 | 0.68 (4, 92) | .680 |
| Medial | ST | 0.37 (1, 23) | .548 | 1.64 (1, 23) | .213 | 0.17 (1, 23) | .686 |
| | ST × H | 0.14 (1, 23) | .715 | 0.95 (1, 23) | .340 | 0.02 (1, 23) | .900 |
| | ST × AP | 2.16 (2, 46) | .136 | 1.98 (2, 46) | .160 | 0.06 (2, 46) | .912 |
| Lateral | ST | 0.48 (1, 23) | .494 | 1.28 (1, 23) | .270 | 0.24 (1, 23) | .627 |
| | ST × H | 0.77 (1, 23) | .389 | 3.60 (1, 23) | .071 | 0.39 (1, 23) | .541 |
| | ST × AP | 2.84 (3, 69) | .084 | 0.94 (3, 69) | .405 | 0.71 (3, 69) | .581 |
| *B. Animacy Violated versus Noncoerced* | | | | | | | |
| Midline | ST | 1.03 (1, 23) | .321 | 3.96 (1, 23) | .059 | 5.20 (1, 23) | .032 |
| | ST × AP | 7.89 (4, 92) | .002 | 3.13 (4, 92) | .064 | 1.83 (4, 92) | .164 |
| Medial | ST | 4.16 (1, 23) | .053 | 4.51 (1, 23) | .045 | 4.34 (1, 23) | .049 |
| | ST × H | 3.99 (1, 23) | .058 | 0.95 (1, 23) | .340 | 0.08 (1, 23) | .784 |
| | ST × AP | 10.13 (2, 46) | .0003 | 5.68 (2, 46) | .018 | 0.74 (2, 46) | .435 |
| Lateral | ST | 2.24 (1, 23) | .148 | 2.34 (1, 23) | .140 | 3.75 (1, 23) | .065 |
| | ST × H | 4.65 (1, 23) | .042 | 0.95 (1, 23) | .341 | 1.90 (1, 23) | .182 |
| | ST × AP | 6.93 (3, 69) | .005 | 4.08 (3, 69) | .034 | 0.07 (3, 69) | .902 |
| *C. Coerced versus Animacy Violated* | | | | | | | |
| Midline | ST | 2.60 (1, 23) | .121 | 6.64 (1, 23) | .017 | 9.34 (1, 23) | .006 |
| | ST × AP | 5.38 (4, 92) | .006 | 5.11 (4, 92) | .011 | 1.41 (4, 92) | .254 |
| Medial | ST | 4.29 (1, 23) | .049 | 6.99 (1, 23) | .014 | 10.84 (1, 23) | .003 |
| | ST × H | 3.19 (1, 23) | .087 | 4.65 (1, 23) | .042 | 0.02 (1, 23) | .883 |
| | ST × AP | 6.08 (2, 46) | .008 | 8.68 (2, 46) | .001 | 0.46 (2, 46) | .611 |
| Lateral | ST | 2.98 (1, 23) | .098 | 4.26 (1, 23) | .050 | 8.70 (1, 23) | .007 |
| | ST × H | 2.10 (1, 23) | .161 | 10.86 (1, 23) | .003 | 1.02 (1, 23) | .322 |
| | ST × AP | 0.90 (3, 69) | .380 | 8.84 (3, 69) | .001 | 0.69 (3, 69) | .470 |

ST = main effect of sentence type; ST × H = Sentence Type × Hemisphere interaction; ST × AP = Sentence Type × Anterior–Posterior distribution.

All analyses presented in this table were repeated with the waveforms time locked to the complement NP itself (rather than to the prestimulus baselines of each successive word). These analyses revealed a similar overall pattern of findings.

2002) rather than retrieving and selecting the specific event or events implied by a particular combination of verb and complement (e.g., "begin 'writing/reading' the book"). Consistent with this idea, the amplitude of the N400 to complements in sentences such as "The student started the essay…," where there was only one dominant interpretation (started writing the essay), did not differ from that of the N400 to complements in sentences such as "The director started the script…" where there were many possible interpretations (e.g., reading the script, marking the script, examining the script, etc.).[4] Like Frisson and McElree (2008), we take this as evidence that the cost of processing coerced complements does not reflect the cost of selecting between alternative specific interpretations.

The amplitude of the N400 effect evoked by coerced complement nouns did not differ from that evoked by the animacy-violated (vs. noncoerced) nouns. We suggest

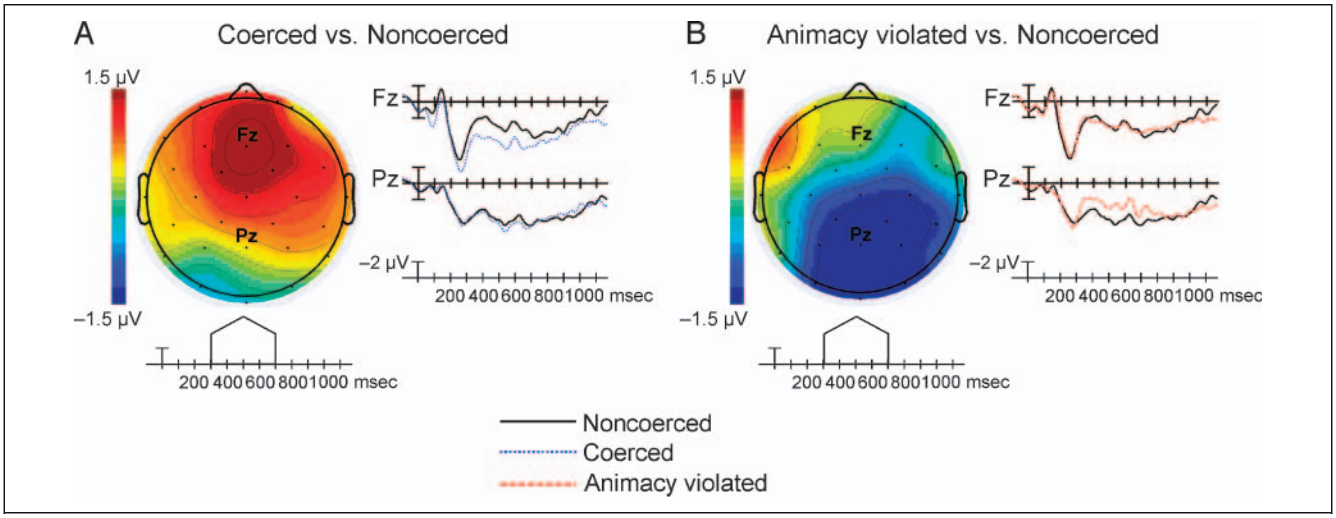



**Figure 4.** Grand-averaged waveforms and voltage maps to the SFW comparing ERPs to (A) coerced versus noncoerced sentence types and (B) animacy-violated versus noncoerced sentence types. Note that the scale used for the voltage maps in this figure is different from that used to illustrate the ERP effects at the point of the critical noun in Figure 2.

that the N400 effect produced by the animacy-violated complements also reflected semantic memory-based processes: matching between the requirements of the verb and the properties of the complement (and possibly attempts to retrieve additional information from semantic memory). Of course, the type of mismatch between the coerced and the noncoerced complements and between the animacy-violated and the noncoerced complements differed. In the case of the coerced complements, the mismatch was between the semantic eventive restrictions of the verb and the entity argument and any implicit retrieval of an event schema resulted in a plausible interpretation. In the case of the animacy violations, the mismatch was between the strict animacy-based restrictions of the object-experiencer verbs and any attempts to retrieve additional information failed to result in a plausible representation. However, these differences made little difference to the amplitude of the N400.

The account outlined earlier makes two related assumptions. The first is that, linguistically, a verb's (or class of verbs') semantic argument structure is represented at a distinct level from its syntactic argument structure and that a verb's semantic structural constraints are "invisible" to the syntax. This deviates from a fairly standard view that the selection restrictions and thematic constraints of a verb are both closely linked to its syntactic argument structure (Chomsky, 1981). In our view, syntactic and semantic structures are represented independently of one another (Culicover & Jackendoff, 2005; Jackendoff, 1997, 2002).[5] The second assumption is that semantic memory-based processes of matching and retrieval, reflected by the N400, are at least partially independent of processes that (syntactically) assign thematic roles to generate full propositions and that assess the plausibility of such propositions (Kuperberg, 2007). This view is supported by observations that the amplitude of the N400 does not necessary pattern with degree of implausibility of a word within a sentence (Van de Meerendonk, Kolk, Vissers, & Chwilla, 2010; Geyer, Holcomb, Kuperberg, & Perlmutter, 2006; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). This was very clear in the present study: There was no significant difference between the N400 amplitude to complement nouns in plausible coerced sentences and in highly implausible animacy-violated sentences (see also Baggio et al., 2010).

Of note, a semantic mismatch and retrieval account of complement coercion was considered by Traxler et al. (2005) but rejected mainly because the pattern of eye-movements observed to coerced (vs. noncoerced) complement nouns differs from the pattern of eye-movements that others have described to outright semantic violations: Although Traxler and others have consistently shown that the costs of coercion are confined mainly to complement NP itself (Pickering et al., 2005; Traxler et al., 2005; Scheepers et al., 2004), highly implausible sentences with outright selection restriction violations are often associated with additional downstream effects past the violated word (Warren & McConnell, 2007; Rayner, Warren, Juhasz, & Liversedge, 2004). Our current ERP data help reconcile these observations. They suggest that initially the coerced and-violated sentences were treated similarly (both evoked an N400 effect that may reflect semantic mismatch and memory-based retrieval). Only in the animacy-violated sentences, however, did the syntactic assignment of thematic roles lead to the generation of a highly implausible proposition. As discussed in the next section, we suggest that this implausibility triggered the P600 effect that continued downstream as a positivity to several words past the critical word. These downstream late positivity effects may map on to the downstream eye-movement effects previously seen in association with severe semantic implausibilities.

**Table 5.** ANOVAs Comparing ERPs to the SFW for the 300- to 700-msec Time Window

| | | SFW: 300–700 msec | |
|---|---|---|---|
| *Effect* | | *F (df)* | *p* |
| *A. Coerced versus Noncoerced* | | | |
| Midline | ST | 3.60 (1, 23) | .071 |
| | ST × AP | 5.95 (4, 92) | .003 |
| Medial | ST | 3.93 (1, 23) | .060 |
| | ST × H | 0.64 (1, 23) | .433 |
| | ST × AP | 5.72 (2, 46) | .013 |
| Lateral | ST | 3.27 (1, 23) | .083 |
| | ST × H | 3.18 (1, 23) | .088 |
| | ST × AP | 3.40 (3, 69) | .057 |
| *B. Animacy Violated versus Noncoerced* | | | |
| Midline | ST | 4.13 (1, 23) | .054 |
| | ST × AP | 6.49 (4, 92) | .002 |
| Medial | ST | 3.75 (1, 23) | .065 |
| | ST × H | 4.84 (1, 23) | .038 |
| | ST × AP | 6.73 (2, 46) | .009 |
| Lateral | ST | 2.52 (1, 23) | .126 |
| | ST × H | 5.96 (1, 23) | .023 |
| | ST × AP | 11.39 (3, 69) | .0001 |
| *C. Coerced versus Animacy Violated* | | | |
| Midline | ST | 16.78 (1, 23) | .0004 |
| | ST × AP | 3.93 (4, 92) | .028 |
| Medial | ST | 20.47 (1, 23) | .0002 |
| | ST × H | 7.98 (1, 23) | .010 |
| | ST × AP | 1.72 (2, 46) | .0002 |
| Lateral | ST | 12.80 (1, 23) | .002 |
| | ST × H | 14.12 (1, 23) | .001 |
| | ST × AP | 4.61 (3, 69) | .017 |

ST = main effect of sentence type; ST × H = Sentence Type × Hemisphere interaction; ST × AP—Sentence Type × Anterior–Posterior distribution.

The N400 effects evoked by the coerced and animacy-violated complement NP were not only similar in amplitude but, relative to the nonviolated NPs, they both had similar widespread scalp distributions (although the voltage map in Figure 2 suggests that the N400 effect to the animacy-violated complement nouns may have been more widespread than that to the coerced complement nouns, this difference was not statistically significant). These observations differ from the pattern of data described in the MEG study by Pylkkänen and McElree (2007) that used similar stimuli and the same acceptability judgment task. In that study, the effect to the coerced (vs. noncoerced) complement nouns was more anteriorly distributed (localizing to an anterior midline field) whereas that to the animacy-violated (vs. noncoerced) complement nouns had a more posterior distribution (localizing to temporal sources). The reasons for this discrepancy are unclear, but it is important to note that MEG and ERP measures at the surface of the scalp can be differentially sensitive to a given underlying neural source; for example, the contribution of radially oriented sources, prominent in EEG, is weak in MEG (Sharon, Hamalainen, Tootell, Halgren, & Belliveau, 2007; Baule & McFee, 1965).[6]

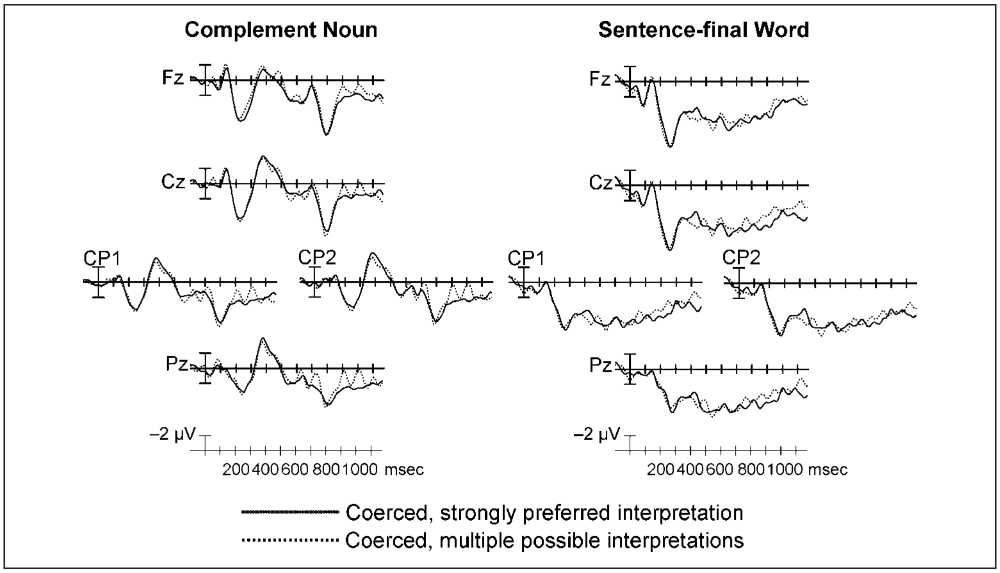
Because of the poor spatial resolution of ERPs and because the N400 is likely to composed of multiple underlying neural generators interacting over the same time scale (Marinkovic et al., 2003; Halgren et al., 2002; Dale et al., 2000; McCarthy, Nobre, Bentin, & Spencer, 1995; Halgren, Baudena, Heit, Clarke, Marinkovic, Chauvel, et al., 1994; Halgren, Baudena, Heit, Clarke, Marinkovic, & Clarke, 1994), we cannot deduce that these two contrasts necessarily engage identical neurocognitive systems. Future studies combining MEG and ERP methods (Sharon et al., 2007) will be able to shed further light on differences and similarities between the neurocognitive mechanisms engaged to coerced versus animacy-violated complement nouns.

## Effects after the N400 Effect

In our study, the ERP response that did clearly distinguish between the coerced and the animacy-violated complement NPs was the P600: Relative to noncoerced complements, the animacy-violated complements evoked a robust P600 effect, but the coerced complements failed to evoke this effect. This finding is also discrepant with that of the MEG study using similar stimuli and the same task and that did not report any modulation within the 500- to 900-msec time window to the animacy-violated NPs. Again, the reasons for this difference between the ERP and the MEG findings are unclear, but it does accord with others' observations that late positivity ERPs after the N400 component can sometimes be invisible to MEG (E. Lau, personal communication). Indeed, MEG is less sensitive than ERPs to the classic oddball P300 effect (Eulitz, Eulitz, & Elbert, 1997; Simpson et al., 1995; Okada, Kaufman, & Williamson, 1983), which shares some functional commonalities (Coulson, King, & Kutas, 1998) and some differences (Osterhout & Hagoort, 1999) with the P600.

The presence of a P600 ERP effect to the animacy violations in the present study is, however, consistent with a growing ERP literature documenting a P600 effect not only to syntactic violations where it was classically associated (Osterhout & Holcomb, 1992) but also to clear semantic implausibilities/impossibilities (reviewed by Kuperberg,

**Figure 5.** Grand-averaged waveforms comparing ERPs of coerced sentences with strongly preferred interpretations and multiple interpretations at the point of the critical noun (left) and the SFW (right).




2007). These include animacy violations falling on verbs (Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb, 2006; Kim & Osterhout, 2005; Hoeks, Stowe, & Doedens, 2004; Kuperberg et al., 2003) as well as animacy violations falling on arguments after verbs. These include inanimate NPs after animate-selecting object-experiencer verbs (Paczynski & Kuperberg, 2009; as in the present study), inanimate arguments after animate-selecting agent-patient verbs (Paczynski & Kuperberg, 2009; Nieuwland & Van Berkum, 2005), and animate arguments after inanimate-selecting agent-patient verbs (Paczynski & Kuperberg, 2009). There is also some evidence that a P600 can be evoked by other types of selection restriction violations (Geyer et al., 2006) and other types of severe implausibilities (Van de Meerendonk et al., 2010).

On the basis of a review of this "semantic P600 effect," Kuperberg (2007) suggested that this component reflects a continued combinatorial analysis (or reanalysis) that is triggered by a highly implausible or unlicensed proposition generated by a full interpretative combinatorial analysis, and that is particularly likely to be triggered in the presence of a conflict with a semantic memory-based analysis. This detection of conflict and reanalysis may draw upon more general executive functions, constituting an on-line "monitoring" process (see Kolk & Chwilla, 2007). Several factors, acting in combination, can bias toward increased conflict, including a strong semantic constraint of the context, a very implausible (as opposed a semi-implausible) final representation of meaning, and the performance of a plausibility judgment task (for further discussion, see Kuperberg, 2007). In the present study, we suggest that the combination of triggers of this effect was the semantic constraint imposed by the object-experiencer verbs, the highly implausible resulting proposition, together with the requirement to make explicit acceptability judgments.[7]

## Downstream ERP Effects after the Complement NP

The positivity evoked by the animacy-violated complement NP remained evident from 400 to 600 msec after the onset of the subsequent word. After this, however, the waveform flipped such that, relative to the other sentence types, a sustained negativity effect was seen on all words up until and including the SFW of the animacy-violated sentences. There have been several previous reports of prolonged negativity effects on SFWs after mid-sentence anomalies (both semantic and syntactic; Ditman, Holcomb, & Kuperberg, 2007; Hagoort, 2003; Hagoort & Brown, 2000; Hagoort et al., 1993; Osterhout & Holcomb, 1992, 1993), and its functional significance is debated. One suggestion has been that it reflects an ongoing difficulty in semantic integration, that is, a prolongation of the N400 or the result of multiple N400s (Osterhout & Holcomb, 1992). An alternative possibility is that it reflects a lack of processing relative to the nonviolated sentences. We are currently attempting to distinguish between these two accounts by combining ERPs with self-paced reading (Ditman et al., 2007) and using experimental tasks that vary in their requirements for participants to read until the end of the sentences.

Unlike the pattern of ERPs after the animacy-violated complements, the waveforms evoked by the three words after the coerced complement NPs did not diverge from those after the noncoerced complements. At the point of the SFW, however, the waveforms did differ significantly: SFWs in the coerced sentences evoked a robust anteriorly distributed sustained positivity effect relative to the SFWs of the noncoerced sentences. The functional significance of this effect is unclear (none of the previous behavioral or neural studies of complement coercion have examined effects at the point of the SFW). One possibility is that it reflects a frontally mediated active attempt to retrieve a specific unstated event (or possible set of events) in the coerced sentences to form a discourse-level representation. This interpretation links it to other types of frontal positivities that have been reported in various other situations where new information must be explicitly retrieved to build a coherent mental model (Filik, Sanford, & Leuthold, 2008; Dwivedi, Philips, Lague-Beauvais, & Baum, 2006; Coulson & Williams, 2005; Kaan & Swaab, 2003; Friederici, Hahne, & Saddy, 2002). For example, in a study by Dwivedi et al. (2006), a sustained frontal positivity was evoked to words such as "ends" in scenarios like "John is considering writing a novel. It ends quite abruptly," relative to "John is reading a novel. It ends quite abruptly". In the former case, the reader must infer that John wrote the novel to make full sense of the meaning. Similarly, in a recent study by Filik et al. (2008), a frontal positivity was evoked by the pronoun "she" versus "they" in scenarios such as "The in-flight meal I got was more impressive than usual. In fact, she/they courteously presented the food as well" where, again, the reader must make an inference that the in-flight meal was presented by a female.

There are, of course, important differences between the types of stimuli used in these previous studies and those used in the present investigation, and future studies will determine whether anterior positivities can, in fact, be linked to these types of explicit inferential processes. What is clear from the present data set is that the sentence-final anterior positivity in the coerced sentences did not reflect the resolution of ambiguity, as has been hypothesized for frontal positivities observed in other situations (e.g., Kaan & Swaab, 2003); similar to the earlier N400 effect evoked at the point of the complement NP, this sentence-final frontal positivity effect was not modulated by the dominance or number of possible interpretations of the coerced sentences. (Note that retrieval and selection processes can be neuroanatomically dissociated: There is fMRI evidence that they are mediated by distinct regions within the inferior frontal cortex [Wagner, Pare-Blagoev, Clark, & Poldrack, 2001; Thompson-Schill, D'Esposito, & Kan, 1999].)

## Conclusions

In sum, we have demonstrated a widespread N400 effect to entity complement NPs after verbs that selected for activities rather than entities. These findings are consistent with previous behavioral and MEG evidence indicating that the processing system registers such discrepancies between the semantic structure of verbs and arguments, although such violations are invisible to the syntax and do not lead to an implausible interpretation. In the present study, the amplitude of this N400 effect was very similar to that evoked by complements that violated the animacy-based selection restrictions of their preceding verbs. We have suggested that, in both cases, N400 modulation might reflect the registration of a mismatch between the semantics of the verb (whether this be its selection restrictions for events or features) and the semantic properties of the incoming complement and possibly implicit attempts to retrieve relevant information from semantic memory to "fill in" such mismatches. We also suggest that a delayed sustained anterior positivity on the SFW of the coerced sentences may reflect delayed more explicit efforts to retrieve the specific unstated activities implied by the verb–argument combination.

The interpretation of the N400 to the coerced complements outlined in this article is based on a growing literature suggesting that the modulation of this component is driven by semantic memory-based processes at several different levels and grains of representation. In this study, we have suggested that the N400 to coerced complements was modulated by a mismatch between the semantic properties of verbs that select for events and arguments that denote entities. However, it is unlikely that this specific type of mismatch between a verb and an argument is the only trigger to "coercion" or other types of inferencing in all situations. For example, within sentences, there is some evidence that coercion on complements can occur in the absence of verb–argument semantic mismatch (Frisson, McElree, & Thyparampil, 2005), and within discourse, on-line inferences can be generated even when semantic relationships between individual words are held constant (Paczynski, Ditman, Okano, & Kuperberg, 2007; Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006). We use many different types of stored information to comprehend language on-line, and the N400 is known to be sensitive to categorical feature-based relationships (Federmeier & Kutas, 1999), animacy-based relationships (Paczynski & Kuperberg, 2009; Frisch & Schlesewsky, 2001; Weckerly & Kutas, 1999), associative-based relationships (Van Petten, 1993) including those based on real-world expectations (Hagoort et al., 2004; Kuperberg et al., 2003), and pragmatic relationships (Nieuwland & Kuperberg, 2008; Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008). Mismatches between language input and any of these levels of representations could, in theory, be associated with attempts to retrieve unstated meaning that may, in some cases, lead to plausible representations. Future studies using complementary ERP, fMRI, and MEG methodologies will be necessary to examine the range of triggers and neural mechanisms engaged to retrieve unstated meaning to make full sense of language.

## Acknowledgments

The authors thank Phillip Holcomb and Tali Ditman for their insightful comments on the project. They are also grateful to Abigail Swain for her assistance in collecting data. This work was funded by the National Institute of Mental Health (RO1 MH02034 to G. R. K.) and the National Alliance for Research in Schizophrenia and Depression (NARSAD) with the Sidney J. Baer Trust (G. R. K.).

Reprint requests should be sent to Gina R. Kuperberg, Department of Psychology, Tufts University, 490 Boston Ave., Medford, MA 02155, or via e-mail: kuperber@nmr.mgh.harvard.edu.

## Notes

1. Processing costs are, however, attenuated when the full event, including the complement NP, is presented in the immediate discourse context (Traxler et al., 2005).
2. Participants' very high accuracy in classifying animacy-violated sentences as unacceptable is consistent with our previous studies (Kuperberg, Sitnikova, & Lakshmanan, 2008; Kuperberg et al., 2003, 2007; Kuperberg, Caplan, et al., 2006; for a discussion, see Kuperberg, Lakshmanan, et al., 2006).
3. These effects are unlikely to have been driven by differences before the point of the CN. Although the N400 to the coercive verbs, which had the lowest frequency, was less negative than to the noncoercive and object-experiencer verbs (coercive vs. noncoercive: $F$s $> 5.60$, $p$s $< .03$ at all columns; coercive vs. object-experiencer: $F$s $> 5.02$, $p$s $< .03$ at all columns), the waveforms converged by the point of the article (no effects of sentence type between 0 and 300 msec, between 300 and 500 msec, or between 600 and 900 msec after the onset of the article: all $F$s $< 2.49$, all $p$s $> .06$).
4. The average waveforms evoked to these two conditions lay completely on top of one another (Figure 5). Nonetheless, given that, after artifact rejection, the number of trials that went into the averaged waveforms of each of these two conditions was relatively small, it is possible that, if the effect size was small, this null result arose because of insufficient power to detect significant differences.
5. The linguistic accounts of Jackendoff (1997, 2002) and Pustejovsky (1995) both presume that coercion is a general process that should apply across the board to all aspectual verbs. However, some aspectual verbs are not acceptable in coerced contexts, for instance **stop the book* (cf. *stop reading the book*). This raises the possibility that coercion is verb specific: some verbs such as "begin" may be encoded in the lexicon with a disjunctive semantic argument structure, that is, (a) "begin Event" or (b) "begin to do something with Object," whereas other verbs such as "stop" only have structures like (a). Under such an analysis, however, the processor must still fill in the content of "do something" with an action appropriate to the object.
6. It is also possible that the posteriorly distributed P600 effect to the animacy violations observed in the present study (but not in the MEG study) overlapped spatially and temporally on the scalp surface with the earlier N400 effect, attenuating this effect at posterior sites. In other words, the N400 effect to the animacy-violated (vs. noncoerced) NPs may have actually been more posteriorly distributed had it not been masked by the overlapping centro-parietal positivity. The recent ERP data collected by Baggio et al. (2010), however, argue against this explanation: In that study, there was also no significant difference in the spatial distribution of the N400 effects evoked by the coerced and animacy-violated NPs despite the absence of a robust P600 effect to the animacy-violated NPs.
7. Baggio et al. (2010) failed to see a P600 effect on the animacy-violated complements. This may be because participants were not required to make explicit acceptability judgments. Other studies, particularly those using highly semantically constrained contexts, however, have reported semantic P600 effects to highly implausible NPs in the absence of acceptability judgment tasks (Van de Meerendonk et al., 2010; Nieuwland & Van Berkum, 2005). Still other studies do not find a P600 effect to less severe semantic implausibilities when participants carry out acceptability judgment tasks (Kuperberg et al., 2003, 2007; Kuperberg, Caplan, et al., 2006). This is why we believe that task is just one of many factors that act in consort to determine the likelihood that conflict between semantic memory-based and combinatorial streams of processing will be detected, leading to additional combinatorial processing and a P600 effect.

## REFERENCES


Baggio, G., Choma, T., van Lambalgen, M., & Hagoort, P. (2010). Coercion and compositionality. *Journal of Cognitive Neuroscience*, *22,* 2131–2140.

Baule, G., & McFee, R. (1965). Theory of magnetic detection of heart's electrical activity. *Journal of Applied Physics, 36,* 2066–2073.

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

Brennan, J., & Pylkkänen, L. (2008). Processing events: Behavioral and neuromagnetic correlates of aspectual coercion. *Brain and Language, 106,* 132–143.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Cooper, R., Winter, A. L., Crow, H. J., & Walter, W. G. (1965). Comparison of subcortical, cortical and scalp activity using chronically indwelling electrodes in man. *Electroencephalography and Clinical Neurophysiology, 18,* 217–228.

Coulson, S., King, J., & Kutas, M. (1998). Expect the unexpected: Event-related brain responses to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

Coulson, S., & Williams, R. F. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia, 43,* 128–141.

Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax*. New York: Oxford University Press.

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron, 26,* 55–67.

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.

Delucchi, M. R., Garoutte, B., & Aird, R. B. (1962). The scalp as an electroencephalographic averager. *Electroencephalography and Clinical Neurophysiology, 14,* 191–196.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology, 44,* 927–935.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Science, 11,* 355–372.

Dwivedi, V., Philips, N., Lague-Beauvais, M., & Baum, S. (2006). An electrophysiological study of mood, modal context, and anaphora. *Brain Research, 1117,* 135–153.

Eulitz, C., Eulitz, H., & Elbert, T. (1997). Differential outcomes from magneto- and electroencephalography for the analysis of human cognition. *Neuroscience Letters, 227,* 185–188.

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

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

Filik, R., Sanford, A. J., & Leuthold, H. (2008). Processing pronouns without antecedents: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 20,* 1315–1326.

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

Friederici, A., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic

complexity and syntactic repair. *Journal of Psycholinguistic Research, 31,* 45–63.
Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *NeuroReport, 12,* 3391–3394.
Frisson, S., & McElree, B. (2008). Complement coercion is not ambiguity resolution: Evidence from eye movement. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1–11.
Frisson, S., McElree, B., & Thyparampil, T. (2005). *Coercion without type-shifting: The role of the subject in enriched interpretations*. Paper presented at the 11th Annual Conference on Architectures and Mechanisms for Language Processing.
Geisler, C. D., & Gerstein, G. L. (1961). Surface EEG in relation to its sources. *Electroencephalography and Clinical Neurophysiology, 13,* 927–934.
Geyer, A., Holcomb, P., Kuperberg, G., & Perlmutter, N. (2006). Plausibility and sentence comprehension. An ERP Study. Paper presented at the conference for the Society of Cognitive Neuroscience.
Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95–112.
Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience, 15,* 883–899.
Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences, 9,* 416–423.
Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8*, 439–483.
Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared with reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia, 38,* 1531–1549.
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., et al. (1994). Spatio-temporal stages in face and word processing: II. Depth-recorded potentials in the human frontal and Rolandic cortices. *Journal of Physiology (Paris), 88,* 51–80.
Halgren, E., Baudena, P., Heit, G., Clarke, J. M., Marinkovic, K., & Clarke, M. (1994). Spatio-temporal stages in face and word processing: I. Depth-recorded potentials in the human occipital, temporal and parietal lobes. *Journal of Physiology (Paris), 88,* 1–50.
Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage, 17,* 1101–1116.
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.
Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30,* 47–61.
Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 721–742.
Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge, MA: MIT Press.
Jackendoff, R. (2002). *Foundations of language. Brain, meaning, grammar, evolution*. New York: Oxford University Press.
Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15,* 98–110.
Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52,* 205–225.
Kolk, H. H. J., & Chwilla, D. J. (2007). Late positivities in unusual situations. *Brain and Language, 100,* 257–262.
Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General, 121,* 459–479.
Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146,* 23–49.
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., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language, 100,* 223–238.
Kuperberg, G. R., Lakshmanan, B. M., Caplan, D. N., & Holcomb, P. J. (2006). Making sense of discourse: An fMRI study of causal inferencing across sentences. *Neuroimage, 33,* 343–361.
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.
Kuperberg, G. R., Sitnikova, T., & Lakshmanan, B. M. (2008). Neuroanatomical distinctions within the semantic system during sentence comprehension: Evidence from functional magnetic resonance imaging. *Neuroimage, 40,* 367–388.
Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends 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.
Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1994–2005. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 659–724). New York: Elsevier.
Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.
Landauer, T. K., Foltz, P. W., & Dumais, S. T. (1998). Introduction to latent semantic analysis. *Discourse Processes, 25,* 259–284.
Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience, 9,* 920–933.
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.
McCarthy, G., Nobre, A. C., Bentin, S., & Spencer, D. D. (1995). Language related field potentials in the anterior-medial

temporal lobe: I. Intracranial distribution and neural generators. *Journal of Neuroscience, 15,* 1080–1089.
McElree, B., Frisson, S., & Pickering, M. J. (2006). Deferred interpretations: Why starting Dickens is taxing but reading Dickens isn't. *Cognitive Science, 30,* 181–192.
McElree, B., Pylkkänen, L., Pickering, M. J., & Traxler, M. (2006). The time course of enriched composition. *Psychonomic Bulletin & Review, 13,* 53–59.
McElree, B., Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R. (2001). Reading time evidence for enriched composition. *Cognition, 78,* B17–B25.
Montague, R. (1970). Universal grammar. *Theoria, 36,* 373–398.
Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth isn't too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science, 19,* 1213–1218.
Nieuwland, M. S., & Van Berkum, J. J. A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Brain Research, Cognitive Brain Research, 24,* 691–701.
Okada, Y. C., Kaufman, L., & Williamson, S. J. (1983). The hippocampal formation as a source of the slow endogenous potentials. *Electroencephalography and Clinical Neurophysiology, 55,* 417–426.
Osterhout, L., & Hagoort, P. (1999). A superficial resemblence does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate. *Language and Cognitive Processes, 14,* 1–14.
Osterhout, L., & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.
Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes, 8,* 413–437.
Paczynski, M., Ditman, T., Okano, K., & Kuperberg, G. R. (2007). Drawing inferences during discourse comprehension: An ERP Study. Paper presented at the conference for the Society of Cognitive Neuroscience.
Paczynski, M., & Kuperberg, G. R. (2009). *Animacy hierarchy, but not thematic role type, influences direct object argument realization in English: Evidence from event-related potentials*. Paper presented at the CUNY Conference on Human Sentence Processing.
Pickering, M. J., McElree, B., & Traxler, M. J. (2005). The difficulty of coercion: A response to de Almeida. *Brain and Language, 93,* 1–9.
Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
Pylkkänen, L., Martin, A. E., McElree, B., & Smart, A. (2009). The anterior midline field: Coercion or decision making? *Brain and Language, 108,* 184–190.
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.). New York: Elsevier.
Pylkkänen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience, 19,* 1905–1921.
Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1290–1301.
Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22,* 642–647.
Scheepers, C., Keller, F., & Lapata, M. (2008). Evidence for serial coercion: A time course analysis using the visual-world paradigm. *Cognitive Psychology, 56,* 1–29.
Scheepers, C., Mohr, S., Keller, F., & Lapata, M. (2004). *The cost of enriched composition: Eye-movement evidence from German*. Paper presented at the CUNY Conference on Human Sentence Processing.
Sharon, D., Hamalainen, M. S., Tootell, R. B., Halgren, E., & Belliveau, J. W. (2007). The advantage of combining MEG and EEG: Comparison to fMRI in focally stimulated visual cortex. *Neuroimage, 36,* 1225–1235.
Simpson, G. V., Pflieger, M. E., Foxe, J. J., Ahlfors, S. P., Vaughan, H. G. J., Hrabe, J., et al. (1995). Dynamic neuroimaging of brain function. *Journal of Clinical Neurophysiology, 12,* 432–449.
Sitnikova, T., West, W. C., Kuperberg, G. R., & Holcomb, P. J. (2006). The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. *Biological Psychology, 71,* 326–340.
Thompson-Schill, S. L., D'Esposito, M., & Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron, 23,* 513–522.
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.
Van Berkum, J. J. A. (2009). The neuropragmatics of "simple" utterance comprehension: An ERP review. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory*. Basingstoke: Palgrave.
Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 443–467.
Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657–671.
Van Berkum, J. J. A., Van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20,* 580–591.
Van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience, 22,* 67–82.
Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. *Language and Cognitive Processes, 8,* 485–531.
Wagner, A. D., Pare-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31,* 329–338.
Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14,* 770–775.
Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology, 36,* 559–570.
West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research, 13,* 363–375.