

Neural correlates of processing syntactic, semantic, and thematic relationships in sentences

Gina R. Kuperberg

*Department of Psychiatry, Massachusetts General Hospital,
Charlestown, MA, USA, and*

Department of Psychology, Tufts University, Medford, MA, USA

David Caplan

*Department of Neurology, Massachusetts General Hospital,
Boston, MA, USA*

**Tatiana Sitnikova, Marianna Eddy, and
Phillip J. Holcomb**

Department of Psychology, Tufts University, Medford, MA, USA

Event-related potentials were measured as subjects read sentences presented word by word. A small N400 and a robust P600 effect were elicited by verbs that assigned the thematic role of Agent to their preceding noun-phrase argument when this argument was inanimate in nature. The amplitude of the P600, but not the N400, was modulated by the transitivity of the critical verbs and by plausibility ratings of passivised versions of these sentences (reflecting the fit between the critical verb and the inanimate noun-phrase as the verb's

Correspondence should be addressed to Gina R. Kuperberg, Psychiatry Neuroscience Program, CNY-2, Bldg 149, MGH-East, 13th Street, Charlestown, MA 02129, USA. Email: kuperber@nmr.mgh.harvard.edu.

Gina R. Kuperberg is supported by NIMH (K23 MH02034), NARSAD (with the Sidney Baer Trust), and a Claflin Distinguished Scholars Award from Massachusetts General Hospital. Part of this work was supported by grants from NIDCD to David Caplan (DC02146) and NICHD to Phillip Holcomb (HD25889) and by the Institute for Mental Illness and Neuroscience Discovery (MIND). We thank Karin Blais for help with running statistical analyses and making figures, and Donna Kreher for her help with running and analysing data for the behavioural study. We are also grateful to Neal Perlmutter, Ted Gibson, and two anonymous reviewers for their helpful comments on the manuscript. The Brandeis verb lexicon database was supported by a grant NSF IST-81-20403 awarded to Brandeis University.

Theme). The P600 was similar in scalp distribution although smaller in amplitude, than that elicited by verbs with morphosyntactic violations. Pragmatically unlikely verbs that did not violate thematic constraints elicited a larger N400 but no P600 effect. These findings support the theory that the cost of syntactic processing on a verb is influenced by the precise thematic relationships between that verb and its preceding arguments.

INTRODUCTION

One focus of current psycholinguistic research is when and how semantic and syntactic information are used during sentence processing. It is generally agreed that the lexical representation of a verb contains both semantic and syntactic information. This information includes a specification of its grammatical category (the fact that it is a verb), its meaning(s), its morphological features as well as information about how it can be combined with other words, i.e., what types of complements (arguments) it can take.

How these different types of information are used during online sentence processing is more controversial. Classic studies by Tyler and Marslen-Wilson, measuring online reaction times, established that semantic, syntactic and pragmatic information held within the verb are all used immediately and incrementally to constrain the construction of verb-argument structures during online processing (Marslen-Wilson et al., 1988; Tyler, 1992). More recent reaction time and eye-movement studies examining the effects of semantic and pragmatic information on the resolution of local syntactic ambiguities, support the idea that there is rapid, word-by-word communication between syntactic and semantic processes (Boland & Tanenhaus, 1991; Ferreira & Clifton, 1986; Frazier & Rayner, 1982; MacDonald et al., 1994; Stowe, 1989; Tanenhaus & Carlson, 1989; Trueswell & Tanenhaus, 1994), although there remains debate about whether there exists a distinct first stage of syntactic processing that is not subject to the influence of lexico-semantic factors (see Frazier and Rayner, 1982 and Trueswell, Tanenhaus, and Garnsey, 1994, for opposing views and MacDonald et al., 1994, for a review).

Traditional behavioural measures have offered valuable insights into the interactions between semantic and syntactic processes during sentence comprehension but are limited in that they are indirect indices of underlying neurocognitive processes. A difference in reaction time between sentence types A and B could arise either as a result of more extensive processing of sentence type B than A by the same neural processor, or from additional processing in condition B by a qualitatively distinct neural processor.

SYNTACTIC AND SEMANTIC RELATIONSHIPS

Event related potentials (ERPs) are multidimensional measures that can help establish qualitative or quantitative similarities or differences between processing different types of linguistic information. In this study, we measured ERPs associated with the presentation of verbs that violated semantic and grammatical constraints, established by previous parts of a sentence, in an effort to shed light on the use of these types of information during online sentence processing. We were primarily concerned with the modulation of two distinct ERPs elicited by such verbs—the N400 and the P600—that have been associated with semantic and syntactic processes respectively.

The N400 is a negative-going waveform that occurs approximately 400 ms following the onset of contextually inappropriate words within sentences (Kutas & Hillyard, 1980). Its amplitude is also sensitive to semantic associations between single words (Bentin, McCarthy, & Wood, 1985), the organisation of semantic memory during online sentence processing (Federmeier & Kutas, 1999; Kutas & Federmeier, 2000), the cloze probability of a word in normal sentences (Kutas & Hillyard, 1984) and contextual incongruities within a discourse (van Berkum, Hagoort & Brown, 1999). The N400 is thought to reflect the difficulty of semantically integrating a word into its preceding context (Holcomb, 1993).

The P600 is a late positivity that is observed in association with both recoverable and irrecoverable violations of syntax (Friederici, Hahne, & Mecklinger, 1996; Hagoort et al., 1993; McKinnon & Osterhout, 1996; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992, 1993; Osterhout, Holcomb, & Swinney, 1994; Osterhout & Mobley, 1995). The P600 is thought to reflect the cost of syntactic processing (Osterhout et al., 1994). More specifically, it has been proposed that it reflects syntactic reanalysis and repair processes (Friederici, 1995), and/or processes of syntactic integration (Fiebach, Schlesewsky, & Friederici, 2002; Kaan, Harris, Gibson, & Holcomb, 2000).

Although the N400 and the P600 are highly correlated with semantic/pragmatic and syntactic parameters respectively, this association is not perfect. Under some circumstances, semantic parameters appear to modulate the amplitude of the P600 (Brown, van Berkum, & Hagoort, 2000; Friederici & Frisch, 2000; Gunter, Friederici, & Schriefers, 2000; Gunter, Stowe, & Mulder, 1997; Hoeks, Stowe, & Doedens, 2004; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003b; Munte, Heinze, Matzke, Wieringa, & Johannes, 1998; Weckerly & Kutas, 1999). The effect of semantic parameters on the P600 has been explained in different ways. On the one hand, it has been argued that, because the P600 is not only seen in association with syntactic parameters, it is not a reliable index of syntactic processing at all (Munte et

al., 1998). On the other hand, the influence of semantic parameters on the P600 has been taken as evidence that, under certain circumstances, lexico-semantic information can have a direct and immediate influence on the assignment of syntactic structure (Brown et al., 2000; Friederici & Frisch, 2000; Weckerly & Kutas, 1999). This latter interpretation is consistent with the online behavioural studies outlined above.

A striking example of the influence of semantic parameters on the P600 is our recent demonstration that a robust P600 was elicited by verbs that violated the animacy constraints of their preceding noun-phrases (NPs) in non-ambiguous English sentences (Kuperberg et al., 2003b). Specifically, we showed that a P600 effect (and a small but non-significant N400 effect) was elicited by verbs such as “eat” in sentences such as, “*At breakfast, the eggs would eat toast and jam*” (see sentence 2 in Table 1). In contrast, a much larger N400 effect but no P600 effect was elicited by conceptually unlikely verbs that did not violate the animacy constraints of their preceding animate NP, e.g., “bury” in “*For breakfast the boys would only bury . . .*” (see sentence 3 in Table 1) (Kuperberg et al., 2003b). We explained our observation of a larger P600 to the animacy violated relative to non-violated verbs by the fact that the P600 is sensitive to violations of thematic role: the animacy violated verbs assigned a thematic role of

TABLE 1
Types of linguistic violation

<i>Linguistic violation</i>	<i>Explanation</i>	<i>Example</i>
(1) None	Baseline condition against which the other conditions are evaluated.	“ <i>For breakfast the boys would eat toast and jam.</i> ”
(2) Thematic role animacy violation	The animate NP that is assigned the role of Agent by the critical verb, is replaced by an inanimate NP. This makes the sentence pragmatically/semantically implausible.	“ <i>For breakfast the eggs would eat toast and jam.</i> ”
(3) Non-thematic role pragmatic violation	The critical verb is replaced by another verb taken from another sentence scenario. This makes the sentence pragmatically/semantically implausible.	“ <i>For breakfast the boys would bury toast and jam.</i> ”
(4) Morphosyntactic violation	The verb is changed either to violate subject-verb agreement or by using a finite in place of an infinitival verb.	“ <i>For breakfast the boys would eats toast and jam.</i> ”

SYNTACTIC AND SEMANTIC RELATIONSHIPS

Agent to their preceding inanimate arguments that was highly unlikely given the lexical (inanimate) properties of this NP.

The aim of the current study was to further examine the nature of the P600 elicited by thematic role animacy violations. We addressed two questions. First, does the functional process reflected by the P600 elicited by thematic role animacy violations have anything in common with the functional process reflected by the P600 elicited by unequivocal syntactic violations? To address this question, we compared ERPs elicited by thematic role animacy violations to those elicited by morphosyntactic violations (between the NP, an auxiliary when one was present, and the critical verb, see sentence 4 in Table 1). On the basis of previous studies, we expected that the morphosyntactic violations would elicit a robust P600 effect.¹ We reasoned that, if the P600 elicited by animacy thematic role violations reflected a similar syntactic process to that engaged during the processing of morphosyntactic violations, then the waveforms of the P600 effects elicited by these two different violations should be similar. In addition, we reasoned that their scalp distributions should be similar, i.e., that there would be no interaction between sentence type (animacy violated vs. morphosyntactically violated sentences) and electrode site.

The second question that the current study addressed was whether the P600 elicited by the thematic role animacy violations was modulated by the precise thematic relationship between the critical verb and its preceding inanimate NP argument.

In all the animacy violated sentences used in the current experiment, the specific thematic role requirements that the verb imposed on its preceding subject NP were violated. In a subset of the animacy violated sentences presented, however, the NP could have plausibly occupied an alternative thematic role around the verb, had the syntax allowed. For example, in the sentence, “*When John arrived at the restaurant, the food would order . . .*”, “*food*” is an unlikely Agent but a likely Theme for “*order*”, i.e., the sentence, “*When John arrived at the restaurant, the food would be ordered . . .*” is plausible. In a subset of other animacy violated sentences

¹ There is disagreement about whether morphosyntactic agreement violations also elicit a left anterior negativity (LAN) that has been associated with function words and a variety of syntactic anomalies (Friederici, 1995; Kluender & Kutas, 1993; Neville et al., 1991). The LAN may reflect a first stage of syntactic parsing (Friederici, 1995) or working memory operations (Fiebach et al., 2002; King & Kutas, 1995; Kluender & Kutas, 1993). Some studies (Osterhout & Mobley, 1995), although not all (Gunter & Friederici, 1999; Hagoort, Brown, & Groothusen, 1993; Osterhout & Nicol, 1999), report a significant LAN in response to morphosyntactic violations. We did not expect a LAN to be elicited by the morphosyntactic violations presented in the current study as these stimuli did not elicit this effect in a previous ERP study in a separate group of participants (Kuperberg et al., 2003a).

presented, however, the NP could not have plausibly occupied an alternate thematic role around the verb. For example, in the sentence, “*When they greeted the Queen of England the trumpets would curtsey . . .*”, “*trumpets*” is not only an unlikely Agent but also an unlikely Theme for “*curtsey*”, i.e., the sentence, “*When they greeted the Queen of England the trumpets would be curtseyed . . .*” is implausible.

We hypothesised the amplitude of the P600 elicited by a verb would be sensitive to whether there was a possible thematic fit between that verb and its preceding NP argument.² Specifically, we hypothesised that the cost of processing a verb that showed a close possible thematic fit with its preceding NP would be smaller, leading to a smaller P600 effect, than the cost of processing a verb that showed no possible thematic fit with its preceding NP.

We conceptualised the possible thematic fit between a verb and a NP argument as being dependent on at least two factors: first, the subcategorisation properties of that verb (Chomsky, 1981) and second, the finer-grained lexico-semantic properties of the verb and its preceding NP (Ladusaw & Dowty, 1988). In order to investigate the influence of subcategorisation properties of animacy violated verbs (without finer-grained lexico-semantic information) on the cost of their processing, we subdivided the critical verbs in animacy violated sentences into those that were (strictly or preferentially) transitive and those that were (strictly or preferentially) intransitive and compared the ERPs elicited by these two types of verbs. We predicted that the cost of processing transitive verbs (e.g., “*order*”) that can take inanimate NPs as direct objects (e.g., “*food*”) would be less than processing intransitive verbs (e.g., “*curtsey*”) that cannot take inanimate NPs as direct objects, and that this would be reflected by a smaller P600 to transitive than to intransitive critical verbs.

In order to quantify how finer-grained lexico-semantic properties of both animacy violated critical verbs and their preceding NPs influenced the potential thematic fit between these specific verbs and their preceding inanimate NPs, we conducted a rating study using passivised versions of these sentences to determine the likelihood that a particular inanimate NP could act as a Theme or Patient of a particular verb. A group of participants who did not participate in the ERP experiment was presented with passive versions of the animacy violated sentences, presented up to the point of the verb (e.g., “*When John arrived at the restaurant, the food would be ordered . . .*”; “*When they greeted the Queen of England the*

² We are defining “possible thematic fit” as the thematic relationships (generalisable across different verb or sentence meanings) between NPs in a sentence (e.g., Agent, Goal, Instrument, Patient, Location) that are independent of the actual syntactic order of these NPs.

SYNTACTIC AND SEMANTIC RELATIONSHIPS

trumpets would be curtseyed . . .”). They were asked to rate the plausibility of these passivised “sentence beginnings”. These “thematic fit” ratings were used as an independent variable in the ERP experiment. We predicted that the cost of processing critical verbs that had a relatively close “thematic fit” with their preceding NP argument (i.e., that were rated as plausible in their passivised form) would be less, leading to a smaller P600, than that of processing verbs that had a relatively poor “thematic fit” with their preceding NP argument (i.e., that were rated as implausible in their passivised form).

To summarise, the current study aimed first to determine whether there are similarities in the way the brain processed thematic role animacy violations and morphosyntactic violations, and second, to determine whether the processing of thematic role animacy violations could be modulated by the precise thematic relationship between the critical verb and its preceding inanimate NP argument. To address these questions, we measured ERPs to critical words as subjects viewed the same types of sentences as in our previous study (for examples, see sentences 1, 2, and 3 in Table 1 and for a detailed description of stimuli, see Kuperberg et al., 2003b) and, in addition, included a fourth condition that introduced a morphosyntactic violation between the NP, an auxiliary when one was present, and the critical verb (see sentence 4 in Table 1). We first compared the waveform and scalp-distributions of the P600 elicited by critical verbs of thematic role animacy violated and morphosyntactically violated sentences, and second, determined whether the P600 elicited by critical verbs of animacy violated sentences was modulated by the transitivity of these verbs and by independent plausibility ratings of passivised versions of these sentences.

METHODS

Participants

Twenty (sixteen male and four female; mean age: 41) subjects were recruited by advertisement. All participants were right-handed as assessed using the modified Edinburgh Handedness Inventory (Oldfield, 1971; White & Ashton, 1976). All participants had normal or corrected-to-normal vision, were native speakers of English, were not taking any medication, and were screened to exclude the presence of psychiatric disorders and neurological disorders. Written consent was obtained from all subjects before participation according to the established guidelines of the Massachusetts General Hospital and the Tufts New England Medical Center Institutional Review Boards.

Construction and classification of stimuli

Two-hundred and forty verbs were chosen as critical words and sentences providing a constraining context were constructed for each of them (see Table 1). All verbs assigned the role of Agent or Experiencer to the subject of an active sentence, e.g., “*Mary*” in “*Mary kicked the table*” or “*Bill*” in “*Bill feared the mugger*”. Each sentence contained an animate NP that was separated from the critical verb by one (and occasionally more than one) intermediate word. This was to minimise potential ERP carry-over effects of the NP onto the critical verb of interest. None of the critical verbs were sentence-final words: in all sentences the critical verb was followed by two to six words.

The thematic role animacy violated sentences were constructed by replacing the animate NP with an inanimate NP. In all these sentences, the sentence became anomalous on the critical verb. In order that this inanimate NP would not constitute an anomaly in its own right, it was chosen to be in keeping with the overall preceding context of the sentence (although direct semantic relationships with the target verb were avoided). Morphosyntactically violated sentences were constructed by introducing a morphosyntactic violation, either by violating subject-verb agreement or by using a finite in place of an infinitival verb. The non-thematic role pragmatically violated sentences were constructed by replacing the critical verbs with verbs that were chosen pseudorandomly from sentences of another list (see below).

So that no participant would encounter the same word more than once (leading to repetition priming effects) but that, across all subjects, all critical verbs would be seen in all four conditions, the sentences were divided into four lists that were counterbalanced between subjects. This excluded the possibility that any differences found between conditions were due to differences in participants’ recognition of different words.

Thus, in each list, there were 240 test sentences (60 of each of the four experimental conditions: normal sentences, non-thematic role pragmatically violated sentences, thematic role animacy violated sentences, and morphosyntactically violated sentences).

A rating study that we have previously described (Kuperberg et al., 2003b) established that the subject nouns and the critical verbs from the normal sentences (and morphosyntactically violated sentences) sentences were rated as more related in meaning than those from the sentences with thematic role animacy violations, $t_1(24) = 15.2, p < .0001$; $t_2(239) = 14.5, p < .0001$, that were, in turn, rated as more related in meaning than those from the sentences with non-thematic role pragmatic violations $t_1(24) = 2.9, p < .01$; $t_2(239) = 2.9, p < .004$.

Subdivision of animacy violated sentences by transitivity of critical verbs and by plausibility ratings of passivised versions of sentences

The 240 thematic role animacy violated sentences were classified (a) by whether or not their critical verbs were transitive or intransitive and (b) according to a set of independent ratings of passivised versions of the same sentences.

(a) *Classification of transitive versus intransitive verbs*

Verbs were categorised as obligatorily/preferred intransitive, obligatorily/preferred transitive, or both/other as follows: First, for all verbs that appeared in the descriptive norms described by Connine, Ferreira, Jones, Clifton, and Frazier (1984) (20% of all verbs used in the current study), a measure of transitive bias was computed by dividing the number of transitive completions (categories 9–14) by the number of transitive plus intransitive completions (categories 1–6 plus 9–14). Verbs with a transitive bias of less than 0.5 were classified as obligatorily/preferred intransitive; verbs with a transitive bias of more than 0.5 were classified as obligatorily/preferred transitive. Verbs used in the current study but that did not appear in Connine et al. (1984), but that did appear in the Brandeis verb lexicon database (26% of all verbs in the current study) were classified as follows: those verbs with an intransitive code but not a direct-object code were classified as obligatorily/preferred intransitive; those verbs with a direct-object code but not an intransitive code were classified as obligatorily/preferred transitive. The remainder of verbs that were used in the current study but that did not appear in either Connine et al. (1984) or in the Brandeis verb lexicon database were provisionally classified on the intuition of the first author as either obligatorily transitive or obligatorily intransitive or both. The subset of verbs that were provisionally classified as either obligatorily transitive or obligatorily intransitive (54% of verbs used in the current study) were then substituted into the sentence “*This is the man John expected to ____*” (see Connine et al., 1984) and three people were asked whether John or the man carried out the action described by the verb. For example, for the verb, “*sleep*”, the sentence read “*This is the man John expected to sleep*”. The raters were asked “Who slept?”. When all three raters replied “the man”, the verb was classified as intransitive. When all three raters replied “John”, the verb was classified as transitive.

In total, 63 verbs (26.3%) were classified as obligatorily/preferred intransitive, 82 verbs (34.2%) were classified as obligatorily/preferred transitive, and 95 verbs (39.3%) were classified as both/other.

(b) Ratings of passivised versions of the animacy violated sentences

Twelve Tufts undergraduates who did not participate in the ERP experiment participated in the rating study of passivised versions of the animacy violated sentences. Each of the animacy violated sentences was reconstructed in its passive form such that the inanimate agent NP of the sentence became the Theme or Patient in the sentence. As the purpose of the rating study was to determine how likely these nouns would act as Themes of the critical verbs, the passivised sentences ended with the critical verb; three periods after the critical verbs in the passivised sentences were used to indicate that the sentence could continue after the critical verb and participants were told that they were rating “beginnings of sentences”. For example, “*To greet the Queen of England the trumpets should curtsey with respect.*” became “*To greet the Queen of England the trumpets should be curtseyed . . .*” and “*When John arrived at the restaurant food had already ordered from her.*” became “*When John arrived at the restaurant food had already been ordered . . .*”. Participants were asked to give plausibility ratings from 1 through 7, 1 indicating that the sentence described something that would be very unlikely to occur in the real world and 7 indicating that the sentence described something that would be very likely to occur in the real world. They were given the following specific rating instructions followed by some examples:

“In this survey, we’d like you to rate some beginnings of sentences for how good they sound to you, in terms of whether they describe something that’s likely to happen in the real world. So, good sentence beginnings should be ones that describe events that are likely to happen, while bad sentence beginnings should be ones that describe unlikely events.”

A mean rating across the 12 subjects for each of the 240 passivised versions of the sentences (each corresponding to an animacy violated sentence) was computed. This mean rating was used to subdivide the 240 animacy violated sentences into four quartiles. Sixty-two sentences (25.8%) had ratings that were equal to or less than the 25th percentile (mean rating 1.7) and 60 sentences (25%) had ratings that were greater than the 75th percentile (mean rating 5.0).

Of the animacy violated sentences that were rated as most implausible when passivised (lowest quartile), 48% contained critical verbs that were strictly/preferentially intransitive, 24% contained critical verbs that were strictly/preferentially transitive and 27% contained critical verbs that did not fall into either of these categories. Of the animacy violated sentences that were rated as most plausible when passivised (highest quartile), 5% contained critical verbs that were strictly/preferentially intransitive, 43% contained critical verbs that were strictly/preferentially transitive and

SYNTACTIC AND SEMANTIC RELATIONSHIPS

52% contained critical verbs that did not fall into either of these categories.

(c) Ratings of semantic associations between critical verbs and their preceding content words in animacy violated sentences

The degree of semantic association between critical verbs and their preceding content words could have potentially influenced plausibility ratings of the passivised versions of the animacy violated sentences. In order to determine whether this was the case, we conducted an additional rating study in which we extracted just the critical verbs and their preceding content words from the animacy violated sentences. These word lists (with the critical verb underlined) were presented to an additional 12 volunteers who did not participate in either the ERP study or the passivised sentence plausibility ratings. Again, the stimuli were divided into three lists (each seen by four of the volunteers) to allow for counterbalancing of experimental conditions between subjects. These volunteers were asked to give ratings on a 7-point scale according to “how bad or how good the fit is between the last underlined word and the previous words”, with a 7 indicating a good fit and a 1 indicating a bad fit. Analysis of these data revealed that these semantic association ratings were unrelated to the plausibility ratings of the passivised versions of the sentences. The mean semantic association between the critical verb and previous content words of the animacy violated sentences that were rated as most implausible when passivised (lowest quartile) was 4.18 ($SD = 1.42$). The mean semantic association between the critical verb and previous content words of the animacy violated sentences that were rated as most plausible when passivised (highest quartile) was 3.85 ($SD = 1.5$). These were not significantly different, $t(120) = 1.22, p = .22$.

ERP procedure

Each subject was given 12 practice trials at the start of the experiment. Experimental participants were randomly assigned to one of the four lists used for counterbalancing between participants. Participants 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 presentation of a fixation point at the center of the screen for 450 ms followed by a 100 ms blank screen followed by the first word. Each word appeared on the screen for 450 ms with an interstimulus interval (ISI) of 100 ms separating words. The final word of each sentence appeared with a full stop. A 750 ms blank-

screen interval followed the final word in each sentence followed by a “?”. This cue remained on the screen until the subject made his/her response at which point the next trial started. The subject’s task was to decide whether or not each sentence made sense by pressing one of two buttons on a response box with either the left or right thumb. Subjects were told that sentences may not make sense in different ways and that if sentences seemed at all “odd” or unlikely or if it had an obvious grammatical error, they should indicate that it doesn’t make sense. “Yes” and “No” response hands were counterbalanced across subjects. Subjects 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).

Electrophysiological recording

Twenty-nine active tin electrodes were held in place on the scalp by an elastic cap (Electro-Cap International, Inc., Eaton, OH), see Figure 1. Several of these electrodes were placed in standard International 10–20 System locations. These included five sites along the midline (FPz, Fz, Cz, Pz, and Oz) and eight lateral (four over each hemisphere) sites (F3/F4, C3/C4, T3/T4, and P3/P4). Eight extended 10–20 system sites were also used (FC1/FC2, FC5/FC6, CP1/CP2, and CP5/CP6) and eight additional 10–20 sites were altered to form a circle around the perimeter of the scalp. These altered sites included FP1’/FP2’ (33% of the distance along the circle between T3/T4), F7’/F8’ (67% of the distance between FPz and T3/T4), T5’/T6’ (33% of the distance between T3/T4 and Oz), and O1’/O2’ (67% of the distance between T3/T4 and Oz). Electrodes were also placed below the left eye and at the outer canthus of the right eye to monitor vertical and horizontal eye movements. Electrodes were also placed on the left and right mastoids.

The EEG signal was amplified by an Isolated Bioelectric Amplifier System Model H&W-32/BA (SA Instrumentation Co., San Diego, CA) with a bandpass of 0.01–40 Hz and was continuously sampled at 200 Hz by an analogue-to-digital converter. The stimuli presented to participants and their behavioural responses were simultaneously monitored by the digitising computer.

ERP data analysis

Accuracy was computed as the percentage of correct responses. A correct response was a judgement of acceptable for the normal sentences and unacceptable for the anomalous sentences.

Averaged ERPs were formed off-line from trials free of ocular and muscular artifact. The mean of the right and left mastoids was subtracted

SYNTACTIC AND SEMANTIC RELATIONSHIPS

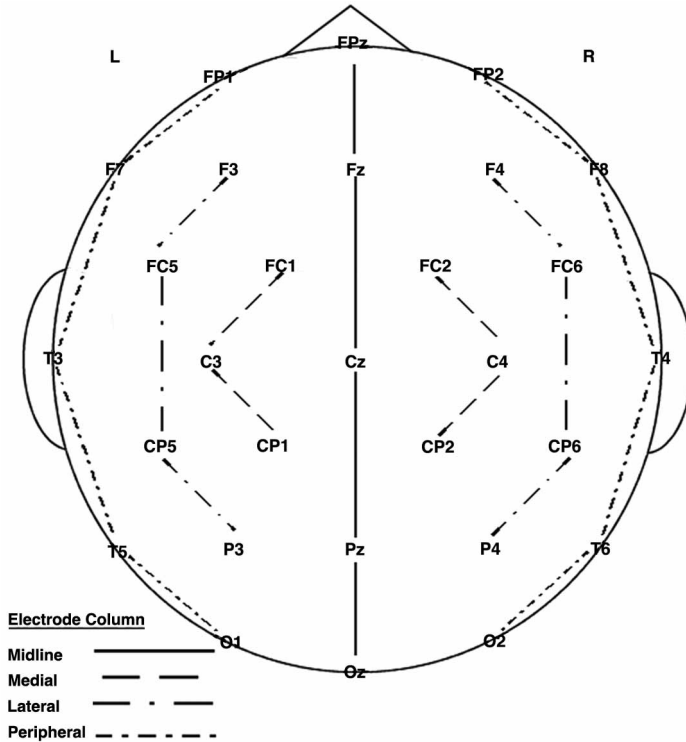


Figure 1. Electrode montage. Analyses of variance were conducted at each of the four parasagittal columns of scalp electrodes shown (see Methods).

from all electrode sites to produce algebraically linked mastoids as the reference. The averaged ERPs elicited by critical verbs were quantified by calculating the mean amplitude values (relative to a 100 ms prestimulus baseline) for the voltage points in three time windows (0–300 ms, 300–500 ms and 500–800 ms after stimulus onset). The latter two time windows roughly correspond to the windows associated with the N400 and P600 components, often reported in studies using linguistic stimuli.

Linearly interpolated voltage maps showing the scalp distribution of differences in ERPs elicited by critical verbs between violated and non-violated sentences within the 300–500 ms and 500–800 ms time windows across the scalp were produced by the EMSC Data Editor program (Source Signal Imaging).

The resulting data for each time window were analysed with analyses of variance (ANOVAs) for repeated-measures. ERPs were analysed separately for midline and lateral electrode sites. The ANOVA for midline electrodes included an electrode site factor (from anterior to posterior:

FPz, Fz, Cz, Pz, Oz). In order to examine distributional differences across the scalp in both left–right and anterior–posterior dimensions, we carried out ANOVAs comparing right and left-sided parasagittal columns: medial, lateral, and peripheral. The medial analysis had three levels of electrode site (FC1/FC2, C3/C4, CP1/CP2) and two levels of hemisphere. The lateral analysis had four levels of electrode site (F3/F4, FC5/FC6, CP5/CP6, P3/P4) and two levels of hemisphere. The peripheral analysis had five levels of electrode site (FP1'/FP2', F7'/F8', T3/T4, T5'/T6', O1'/O2') and two levels of hemisphere.

In cases when we aimed to determine whether effects differed topographically (see hypotheses above), if an interaction between sentence type and topographic (hemisphere and electrode site) factors was significant, we repeated the analysis after z -normalising the data. This normalisation procedure equalises the mean amplitudes across conditions (McCarthy & Wood, 1986). In such cases, we report F and p values for the z -normalised data. It has recently been argued that normalisation does little to improve our ability to distinguish distinct underlying neural generators on the basis of distinct scalp topographies (Urbach & Kutas, 2002). We therefore also report instances where the raw data but not the z -transformed data showed significant (at $p < .05$) sentence type by electrode site interactions. When only non-normalised data were significant, we are very careful about its interpretation.

To protect against Type 1 error resulting from violations of sphericity, the Geisser-Greenhouse correction (Greenhouse & Geisser, 1959) was applied to all repeated measures with more than one degree of freedom in the numerator. In these cases, we report the original degrees of freedom and the corrected probability level. In reporting whether or not a particular effect was significant, we used a standard significance level of $\alpha = .05$ as, in all cases, we were testing specific *a priori* hypotheses. In addition, we report significance at $p < .01$, $p < .001$, and $p < .0001$ for each ANOVA in the results tables.

Effects on ERPs of the four main sentence types

Sentence type (four levels: normal, thematic role animacy violated, non-thematic role pragmatically violated, morphosyntactically violated) were within-subjects variables in initial ANOVAs (in addition to the hemisphere and/or electrode site within-subject variables described above). Significant main effects and interactions in these ANOVAs were followed up by planned simple effects ANOVAs that allowed comparisons between each sentence type.

Effect of transitivity and of plausibility ratings of passivised versions of sentences on critical verbs in the animacy violated sentences

To determine the effect of transitivity on ERPs elicited by the critical verbs of the animacy violated sentences, transitivity (strictly/preferentially transitive versus strictly/preferentially intransitive) was used as a within-subject variable (in addition to the within-subject variables described above) in ANOVAs at the four electrode columns.

To determine the effects of the passivised sentence plausibility ratings on the ERPs elicited by the critical verbs of the animacy violated sentences, a mean rating across the twelve subjects for each of the 240 passivised versions of the sentences (each corresponding to an animacy violated sentence) was computed and this mean rating was used to subdivide the 240 animacy violated sentences into four quartiles. Average ERPs elicited by critical animacy violated verbs in each quartile were computed. Quartile was used as a within-subjects variable in repeated-measures ANOVAs.³

RESULTS

ERP data: Comparisons between the four main sentence types

Nine per cent of the trials were rejected for artifact and there were no significant differences in the percentage of rejected trials across experimental conditions, $F(3, 57) = 0.822, p = .44$. Only trials that were answered correctly were included in the ERP analysis.

Grand-average ERPs elicited by the critical verbs in the four main conditions are shown in Figure 2 at Fz, Cz and Pz and at all electrode sites in Figures 3A, B and C. A clear negative-positive complex is seen in the first 300 ms following word onset (the N1–P2 complex). There appeared to be no differences in the waveforms elicited by the different types of verbs within this early time window, as reflected by a failure to find significant

³ In determining the effects of transitivity and passivised sentence plausibility ratings, we compared different verbs across conditions. Factors such as frequency, concreteness and the number of letters could therefore potentially confound the results. We therefore re-averaged the ERP data after systematically excluding a subset of verbs (5–10% per condition) so as to match frequency (Kucera & Francis, 1967), concreteness (Pavio et al., 1968), and number of letters across transitive and intransitive conditions and across the four quartiles. We then repeated all ANOVAs on these re-averaged data. We report F and p values from the original analyses that included all verbs in each condition, but the results were essentially the same after matching for these potentially confounding parameters.

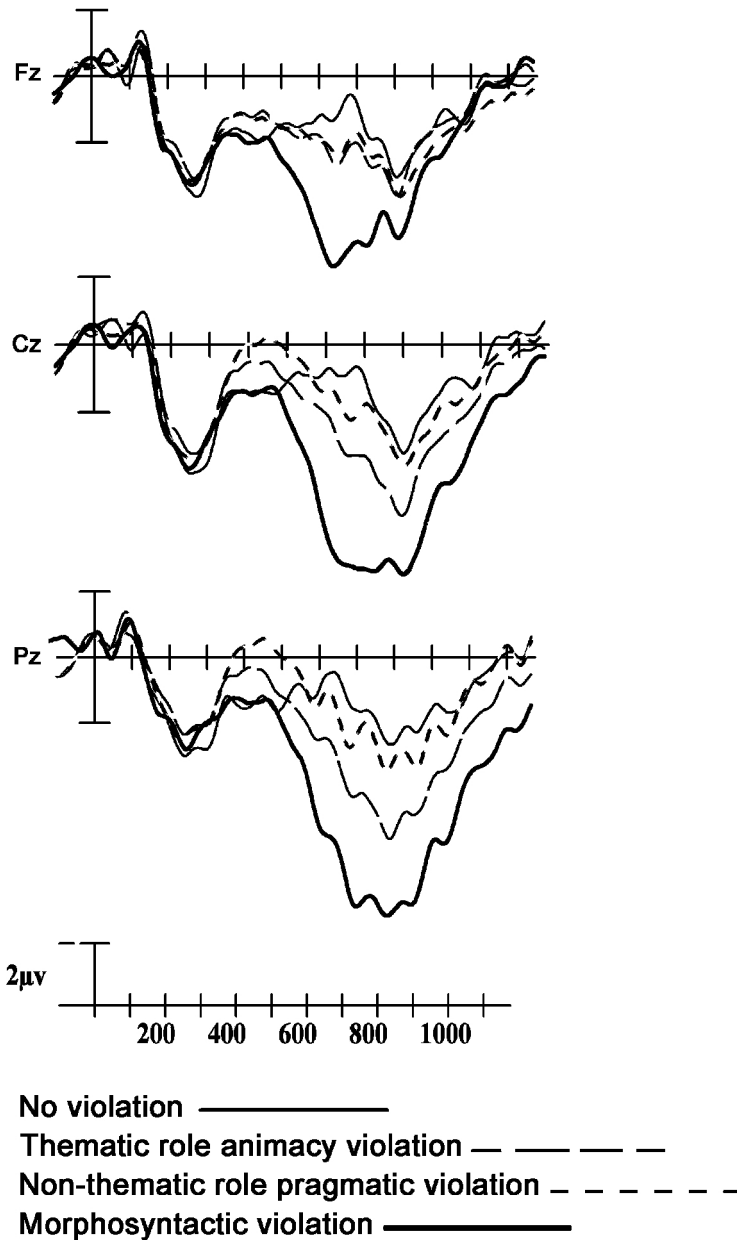


Figure 2. Averaged waveforms elicited by critical verbs at Fz, Cz, and Pz electrode sites. Non-violated verbs (thin solid) vs. thematic role animacy violated verbs (long-dashed) vs. non-thematic role pragmatically violated verbs (dotted) vs. morphosyntactically violated verbs (bold solid).

SYNTACTIC AND SEMANTIC RELATIONSHIPS

main effects of sentence type ($F_s < 0.50$; $p_s > .65$) or significant interactions between sentence type and electrode site ($F_s < 1.2$; $p_s > .3$) in the 0–300 ms time-window at any of the electrode columns.

The N1–P2 complex was followed by a negative-going component with a peak amplitude around 425 ms (the N400). The N400 appeared to be most negative-going to the non-thematic role pragmatically violated verbs, less negative-going to the thematic role animacy violated verbs and least negative-going to the non-violated verbs. In some experimental conditions, this was followed by a large late positive-going wave from approximately 500–900 ms (the P600). The amplitude of the P600 elicited by the thematic role animacy violated verbs appeared to be more positive-going than the P600 elicited either by the non-thematic role pragmatically violated verbs or the non-violated verbs. The amplitude of the P600 elicited by the morphosyntactically violated verbs, however, appeared to be more positive-going than the P600 elicited by the thematic role animacy violated verbs. Below we report the results of statistical analyses comparing ERPs elicited by the critical verbs within time windows capturing the N400 and P600 effects.

300–500 ms: the N400

During the 300–500 ms epoch, there was no overall main effect of sentence type (Table 2, left) but there were significant interactions between sentence type and electrode site at all electrode columns except the medial column (Table 2, right). There appeared to be no clear difference between the amplitude of the N400 waveform elicited by non-violated verbs and the morphosyntactically anomalous verbs. Planned simple-effects analyses were carried out to determine whether the amplitude of

TABLE 2
ERP analyses of parasagittal columns of scalp electrodes: main effect of sentence type and sentence type by electrode site interactions (four levels) at 300–500 ms

<i>Analysis</i>	<i>Main effect of sentence type</i>		<i>Interactions between sentence type and electrode site</i>	
	<i>Degrees of freedom</i>	<i>F value</i>	<i>Degrees of freedom</i>	<i>F value</i>
Midline	3, 57	2.94	12, 228	2.85*
Medial	3, 57	2.86	6, 114	2.33
Lateral	3, 57	2.31	9, 171	3.85**
Peripheral	3, 57	0.77	12, 228	2.63*

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

the N400 elicited by each of these sentence types differed significantly from one another.

As expected, there were no significant differences between the amplitude of the N400 elicited by the morphosyntactically anomalous verbs and the non-violated verbs (main effect of sentence type and interactions involving sentence type: all F s < 1.34 and all p s $> .217$), see Figure 3C.

A significantly larger N400 was elicited by non-thematic role pragmatically violated verbs than by non-violated verbs (significant effects of sentence type at the midline, medial, and lateral electrode columns, Table 3, left). There were also significant interactions between sentence type and electrode site at all electrode columns except the medial column⁴ (Table 3, right) reflecting a larger N400 at more central and posterior than anterior electrode sites (Figure 3A).

In comparing the thematic role animacy violated and normal sentences (Table 4, left), there was no significant main effect of sentence type and no significant interactions with electrode site at any of the electrode columns (F s < 1.65 , p s $> .2$). There were, however, significant interactions between sentence type and hemisphere at the medial, lateral, and peripheral electrode columns (Table 4, right) that arose because the N400 elicited by thematic role animacy violations was more negative-going at right than left electrode sites (Figure 3B).

In comparing ERPs elicited by non-thematic role pragmatically violated verbs and thematic role animacy violated verbs, there was a significant main effect of sentence type at the medial column (Table 5, left) and significant interactions with electrode site at all columns except the lateral column (Table 5, right), arising because the N400 to non-thematic role

⁴ At the medial electrode column, the raw but not the z-transformed data showed a significant (at $p < .05$) sentence type by electrode site interaction.

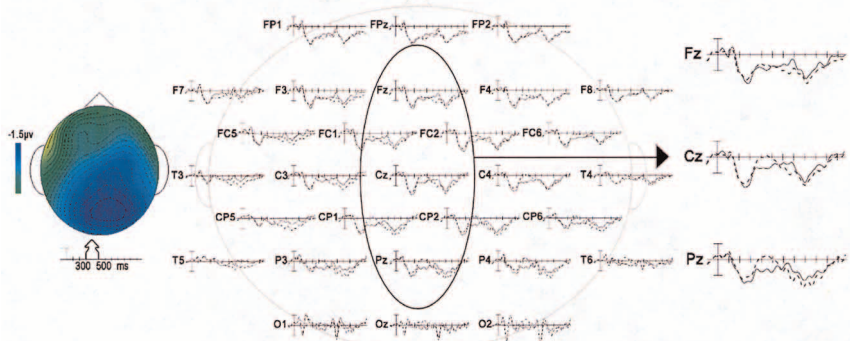
Figure 3 (opposite). Averaged waveforms elicited by non-violated verbs vs. different types of violated verbs within sentences. For each comparison, ERPs elicited by non-violated (solid) and violated (dotted) critical verbs are shown at all electrode sites; voltage maps showing the spatial distribution of effects over all electrodes at the scalp surface are shown on the far left; waveforms elicited at Fz, Cz, and Pz are shown at higher magnification on the far right.

Panel A. ERPs to normal versus pragmatically violated verbs. Far left, blue: the spatial distribution of the N400 effect (the voltage difference between 300 and 500 ms).

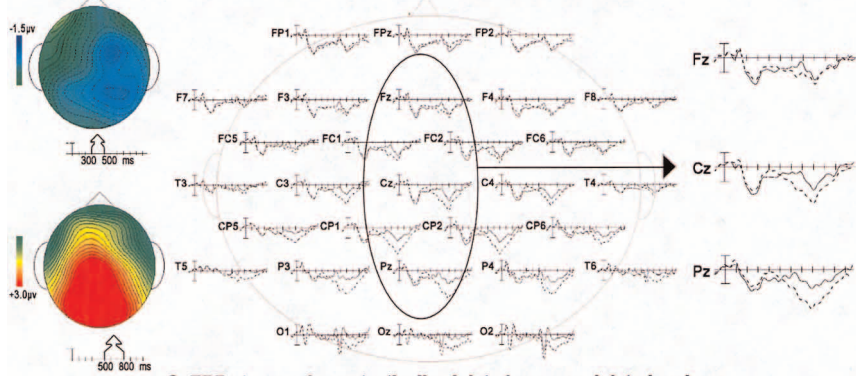
Panel B. ERPs to normal versus thematic role animacy violated verbs. Far left top, blue: the spatial scalp distribution of the N400 effect (the voltage difference between 300 and 500 ms). Far left bottom, red: the spatial scalp distribution of the P600 effect (the voltage difference between 500 and 800 ms). Note that, because the N400 effect (blue) is smaller in magnitude than the P600 effect (red), it is plotted at a smaller scale on the voltage maps.

Panel C. ERPs to normal versus morphosyntactically violated verbs. Far left, red: the spatial distribution of the P600 effect (the voltage difference between 500 and 800 ms).

A. ERPs to pragmatically violated vs. non-violated verbs



B. ERPs to animacy violated vs. non-violated verbs



C. ERPs to morphosyntactically violated vs. non-violated verbs

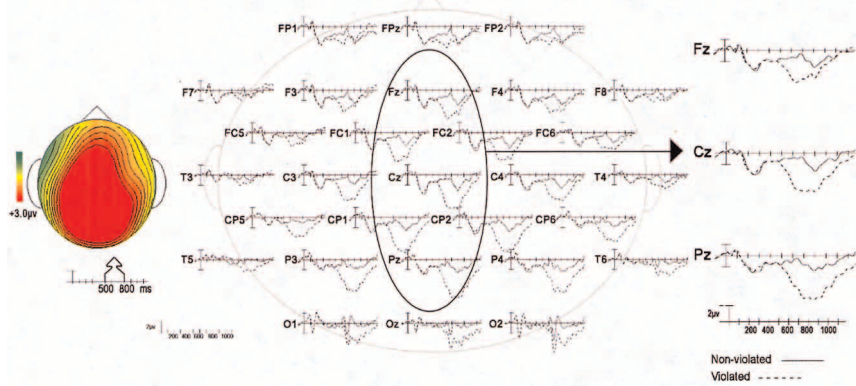


TABLE 3

ERP analyses of parasagittal columns of scalp electrodes: 300–500 ms Epoch. Main effects and interactions in the simple effects ANOVAs: normal vs. non-thematic role pragmatically violated sentences

<i>Analysis</i>	<i>Main effect of sentence type</i>		<i>Interactions between sentence type and electrode site</i>	
	<i>Degrees of freedom</i>	<i>F value</i>	<i>Degrees of freedom</i>	<i>F value</i>
Midline	1, 19	5.78*	4, 76	4.47*
Medial	1, 19	7.67*	2, 38	3.48
Lateral	1, 19	6.54*	3, 57	8.36**
Peripheral	1, 19	1.85	4, 76	4.26*

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

TABLE 4

ERP analyses of parasagittal columns of scalp electrodes: 300–500 ms Epoch. Main effects and interactions in the simple effects ANOVAs: normal sentences vs. thematic role animacy violated sentences

<i>Analysis</i>	<i>Main effect of sentence type</i>		<i>Interactions between sentence type and hemisphere</i>	
	<i>Degrees of freedom</i>	<i>F value</i>	<i>Degrees of freedom</i>	<i>F value</i>
Midline	1, 19	0.87	NA	NA
Medial	1, 19	1.16	1, 19	7.09*
Lateral	1, 19	1.63	1, 19	9.38**
Peripheral	1, 19	1.16	1, 19	14.73***

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

pragmatically violated verbs was more negative-going than to thematic role animacy violated verbs, particularly at central-posterior electrode sites.

500–800 ms: the P600

The overall ANOVA within the 500–800 ms time window revealed clear differences in the amplitude of the P600 between sentence types, with main effects of sentence type (Table 6, left) and significant interactions between sentence type and electrode site at all electrode columns (Table 6, right). Planned simple-effects analyses were carried out to determine which sentence types differed significantly from each of the other sentence types.

SYNTACTIC AND SEMANTIC RELATIONSHIPS

TABLE 5

ERP analyses of parasagittal columns of scalp electrodes: 300–500 ms Epoch. Main effects and interactions in the simple effects ANOVAs: non-thematic role pragmatically violated sentences vs. thematic role animacy violated sentences

Analysis	Main effect of sentence type		Interactions between sentence type and electrode site	
	Degrees of freedom	F value	Degrees of freedom	F value
Midline	1, 19	3.00	4, 76	6.50**
Medial	1, 19	4.99*	2, 38	4.31*
Lateral	1, 19	2.01	3, 57	3.11
Peripheral	1, 19	0.21	4, 76	3.83*

* $p < .05$ ** $p < .01$.

As expected, in the 500–800 ms time window, there were generally no significant differences in the amplitude of waveforms elicited by non-thematic role pragmatic violations in comparison with non-violated verbs (no significant electrode site by condition interactions, $F_s < 0.7$, $p_s > .4$, at any electrode columns; no significant main effects of condition at any columns, $F_s < 2.5$; $p_s > .13$, except approaching significance at the peripheral column, $F = 4.13$, $p = .056$) (Figure 3A).

The P600 elicited by thematic role animacy violations was more positive-going than that elicited by non-violated verbs (main effects of sentence type at all electrode columns, Table 7, left). There were significant sentence type by electrode site interactions at all electrode columns (Table 7, right), reflecting a larger P600 effect at posterior than anterior electrode sites (Figure 3B).

TABLE 6

ERP analyses of parasagittal columns of scalp electrodes: sentence type main effect and sentence type by electrode site interactions (four levels) at 500–800 ms

Analysis	Main effect of sentence type		Interactions between sentence type and electrode site	
	Degrees of freedom	F value	Degrees of freedom	F value
Midline	3, 57	25.66****	12, 228	4.56**
Medial	3, 57	21.18****	6, 114	3.83*
Lateral	3, 57	16.77****	9, 171	2.62*
Peripheral	3, 57	12.48****	12, 228	5.51***

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

TABLE 7

ERP analyses of parasagittal columns of scalp electrodes: 500–800 ms Epoch. Main effects and interactions in the simple effects ANOVAs: normal vs. thematic role animacy violated sentences

Analysis	Main effect of sentence type		Interactions between sentence type and electrode site	
	Degrees of freedom	F value	Degrees of freedom	F value
Midline	1, 19	11.93**	4, 76	6.61**
Medial	1, 19	11.04**	2, 38	6.09*
Lateral	1, 19	7.54**	3, 57	3.86*
Peripheral	1, 19	5.49*	4, 76	9.43**

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

The P600 elicited by morphosyntactic violations was also more positive-going than that elicited by non-violated verbs as reflected by significant main effects of sentence type at all columns (Table 8, left). There were also significant sentence type by electrode site interactions at all electrode columns (Table 8, right), reflecting a larger P600 effect at posterior than at anterior sites (Figure 3C).

The P600 elicited by the morphosyntactically violated verbs was significantly more positive-going than that elicited by thematic role animacy violated verbs with significant main effects at all electrode columns (Table 9, left).⁵ There were no significant interactions between sentence type and electrode site (Table 9, right).^{6,7}

⁵ This was the case even when the data were re-normalised to a 300-500msec baseline: there remained significant main effects of sentence type at all electrode columns (all F s > 10.8 , all p s $< .004$). This makes it unlikely that any difference between the amplitude of the P600 to the morphosyntactic and thematic role animacy violations simply arose because of temporal overlap between the N400 and P600, i.e. as a result of differences between these two conditions in the preceding N400.

⁶ At the midline electrode column, although the z -transformed data failed to show a significant interaction, the raw data showed an interaction between sentence type and electrode site that just reached significance, $p < .05$.

⁷ It has been argued that analyses at multiple electrode sites may be at the expense of statistical power (Oken & Chiappa, 1986). In order to protect against such Type II error, we also carried out an ANOVA on four quadrant regions of interest (ROIs), each comprising five electrodes: upper left (FP1, F7, F3, FC5, FC1); lower left (CP1, CP5, T5, O1, P3); upper right (FP2, F8, F4, FC2, FC6), and lower right (CP2, CP6, P4, O2, T3). This analysis also failed to show a significant sentence type by quadrant interaction before and after z -normalisation, $F(3, 57) = 1.3$, $p = .28$.

SYNTACTIC AND SEMANTIC RELATIONSHIPS

TABLE 8

ERP analyses of parasagittal columns of scalp electrodes: 500–800 ms Epoch. Main effects and interactions in the simple effects ANOVAs: normal sentences vs. morphosyntactically violated sentences

<i>Analysis</i>	<i>Main effect of sentence type</i>		<i>Interactions between sentence type and electrode site</i>	
	<i>Degrees of freedom</i>	<i>F value</i>	<i>Degrees of freedom</i>	<i>F value</i>
Midline	1, 19	45.65****	4, 76	4.26*
Medial	1, 19	36.88****	2, 38	5.74**
Lateral	1, 19	30.79****	3, 57	4.55*
Peripheral	1, 19	23.16****	4, 76	4.34*

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

TABLE 9

ERP analyses of parasagittal columns of scalp electrodes: 500–800 ms Epoch. Main effects and interactions in the simple effects ANOVAs: morphosyntactically anomalous sentences vs. thematic role animacy violated sentences

<i>Analysis</i>	<i>Main effect of sentence type</i>		<i>Interactions between sentence type and electrode site</i>	
	<i>Degrees of freedom</i>	<i>F value</i>	<i>Degrees of freedom</i>	<i>F value</i>
Midline	1, 19	16.02***	4, 76	2.58
Medial	1, 19	12.66**	2, 38	3.0
Lateral	1, 19	11.30**	3, 57	0.83
Peripheral	1, 19	9.31**	4, 76	2.21

* $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

Summary

Non-thematic role pragmatic violations elicited a significant centro-posteriorly distributed N400 effect and thematic role animacy violations elicited a significant right lateralised N400 effect. The amplitude of the N400 elicited by the non-thematic role pragmatic violations was significantly greater than that elicited by the thematic role animacy violations, particularly at central and posterior sites.

Both thematic role animacy and morphosyntactic violations elicited significant P600 effects that were widespread but maximal at posterior electrode sites. The amplitude of the P600 effect elicited by the morphosyntactic violations was significantly greater than that elicited by the thematic role animacy violations but there were no significant differences in their scalp distribution.

ERPs elicited by transitive vs. intransitive thematic role animacy violated verbs

We classified the 240 thematic role animacy violated sentences according to whether or not their critical verbs were strictly/preferentially transitive or strictly/preferentially intransitive.

Grand-average ERPs elicited by transitive and intransitive thematic role animacy violated verbs at Fz, Cz, and Pz are shown in Figure 4 (left). There were no significant effects of transitivity at any of the electrode columns between 0–300 ms ($F_s < 2.4$; $p_s > .137$) or between 300–500 ms ($F_s < 0.21$; $p_s > .657$).

Figure 4 (left) suggests that the amplitude of the P600 elicited by intransitive animacy violated verbs was more positive going than that elicited by transitive animacy violated verbs and that the time-interval that maximally distinguished the two sentence types was between 500 and 700

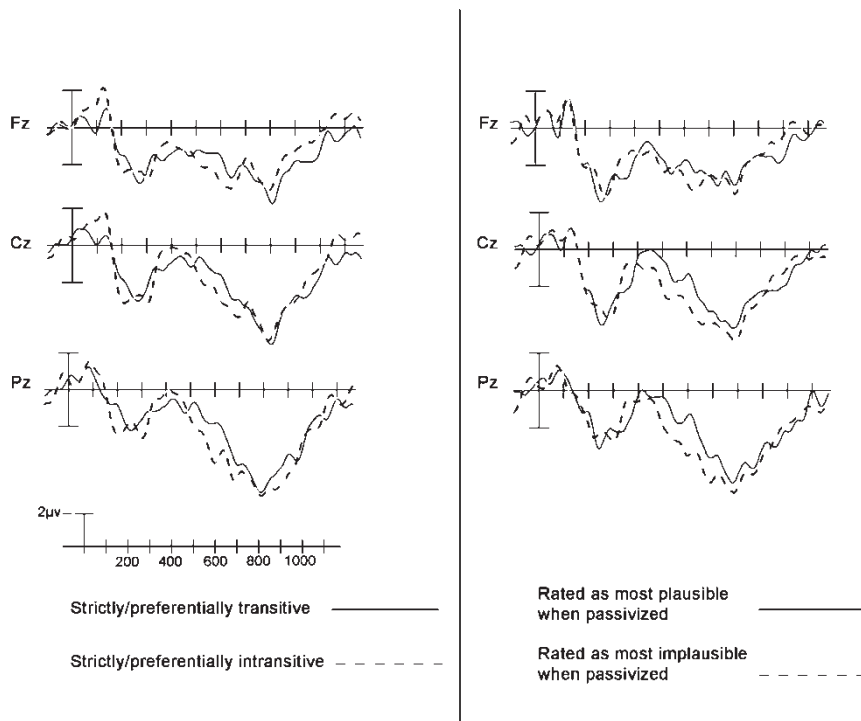


Figure 4. Grand-average ERPs at Fz, Cz, and Pz elicited by obligatorily/preferred transitive (solid) vs. obligatorily/preferred intransitive (dotted) thematic role animacy violated verbs (left) and by critical verbs in animacy violated sentences that, when passivized, were given ratings falling into the highest (solid) and lowest (dotted) quartile (i.e., rated as most plausible and most implausible respectively).

SYNTACTIC AND SEMANTIC RELATIONSHIPS

TABLE 10
ERP analyses of parasagittal columns of scalp electrodes: main effect of transitivity (two levels) and transitivity by electrode site interactions (at 500–700 ms)

Analysis	Main effect of transitivity type		Interactions between transitivity and electrode site	
	Degrees of freedom	F value	Degrees of freedom	F value
Midline	1, 19	4.18*	4, 76	1.53
Medial	1, 19	5.82*	2, 38	0.06
Lateral	1, 19	7.24**	3, 57	0.37
Peripheral	1, 19	3.49 [^]	4, 76	0.65

[^] $p < 0.1$. * $p < .05$ ** $p < .01$. *** $p < .001$. **** $p < .0001$.

ms. ANOVAs between 500–700 ms showed significant main effects of transitivity at all columns except the peripheral column where there was a trend towards significance, see Table 10, left. There were no significant transitivity by electrode site interactions in this time window (Table 10, right).⁸

ERPs to critical verbs in thematic role animacy violated sentences in relation to plausibility ratings when these sentences were passivised

We classified the 240 thematic role animacy violated sentences into four quartiles, according to independent ratings of how plausible they became when passivised. ERPs elicited by the animacy violated verbs were averaged according to whether they fell in quartile 1, 2, 3, or 4. ERPs elicited by critical verbs in animacy violated sentences that, when passivised, were given ratings falling into the lowest and highest quartile (rated as most implausible and most plausible respectively) are shown in Figure 4 (right). Between 0 and 400/500 ms (depending on the electrode site), there appeared to be no systematic differences in the ERPs elicited by verbs with ratings that fell into the lowest quartile in comparison with the highest quartile. Repeated-measures ANOVAs confirmed that there were no significant main effects of quartile (four levels) at any electrode column in the 0–300 ms time window ($F_s < 2.39$; $p_s > .081$) or in the 300–500 ms time-window ($F_s < 1.15$; $p_s > .63$).

⁸ In the 500–800msec time window, effects of transitivity reached significance at the lateral electrode column ($F = 4.5$, $p < .046$) and approached significance at the medial column ($F = 3.49$, $p < .077$) but effects were non-significant at the midline and peripheral columns ($F_s < 2.19$, $p > .15$).

Figure 4 (right) suggests that the P600 elicited by animacy violations that, when passivised, were rated as most implausible (lowest quartile ratings) was more positive-going in comparison with animacy violations that, when passivised, were rated as most plausible (highest quartile ratings). Again, the time-interval that maximally distinguished between sentence types was between 500 and 700 ms. In this time window, there were significant main effects of plausibility rating at all electrode columns, Table 11, left, but no plausibility by electrode site interactions, Table 11, right. Planned simple effects ANOVAs revealed significant differences in comparing quartile 1 with quartile 3 (across all electrode columns, $F_s > 6.37$; $ps < .01$), in comparing quartile 2 with quartile 3 (across all electrode columns, $F_s > 6.06$; $ps < .01$) and in comparing quartile 1 and quartile 4 (main effects of quartile across midline, medial, and lateral columns, $F_s > 6.06$; $ps < .02$, and a quartile by electrode interaction at the peripheral column, $F(4, 76) = 4.73$, $p < .023$), but no significant differences between quartiles 1 and 2 and between quartiles 3 and 4 ($p > .05$) suggesting that the decrease in the amplitude of the P600 across quartiles 1 through 4 was not completely linear.

In order to take better advantage of the full range of plausibility ratings and the full range of ERP amplitudes elicited by critical verbs in the animacy violated sentences, we adopted a second method to determine whether passivised plausibility ratings predicted the amplitude of ERPs. The mean rating across the 12 subjects for each of the 240 passivised versions of the sentences was used to subdivide the 240 animacy violated sentences into 40 groups of six sentences. We calculated the average voltage of ERPs elicited by animacy violated verbs across participants at central (averaged across C3, Cz, and C4), centro-parietal (averaged across CP5, CP1, CP2, and CP6) and parietal (averaged across P3, Pz, and P4) electrode sites in each of the 40 sentence groups in the 0–300 ms, the 300–

TABLE 11
ERP analyses of parasagittal columns of scalp electrodes: main effect of passivised quartiles (four levels) and quartile by electrode site interactions at 500–700 ms

Analysis	Main effect of passivised quartiles type		Interactions between quartile and electrode site	
	Degrees of freedom	F value	Degrees of freedom	F value
Midline	3, 57	5.63**	12, 228	1.94
Medial	3, 57	4.69*	6, 114	0.76
Lateral	3, 57	5.07*	9, 171	1.58
Peripheral	3, 57	3.82*	12, 228	2.15

* $p < .05$ ** $p < .01$.

500 ms and the 500–700 ms time windows. We conducted correlational analyses to determine whether the mean rating within each of these 40 sentence-groups predicted the average voltage (across all participants) of ERPs elicited by critical verbs at central, centro-parietal, and parietal sites in each of these 40 sentence groups.

The mean passivised plausibility ratings failed to predict the average voltages elicited by the animacy violated verbs across these 40 sentence groups in the 0–300 ms and the 300–500 ms time windows (Spearman's $r_s < -0.08$, $p_s > .1$). However, in the 500–700 ms time window, there were significant inverse correlations between the plausibility ratings of the passivised versions of the sentences and the average voltages across the 40 groups (more positive ERPs associated with lower ratings) at central (Spearman's $r = -0.37$, $p < .02$), parietal (Spearman's $r = -0.34$, $p < .03$) sites and centro-parietal sites (Spearman's $r = -0.39$, $p < .01$).⁹

Summary

The amplitude of the P600 elicited by verbs in thematic role animacy violated sentences was modulated by how plausible the sentences became when passivised as well as by whether the critical verbs were transitive or intransitive, particularly between 500–700 ms. Critical verbs in animacy violated sentences that were intransitive elicited significantly more positive P600 effects than transitive critical verbs. Critical verbs of animacy violated sentences that were rated as implausible when passivised elicited significantly more positive P600 effects than critical verbs of sentences that were rated as plausible when passivised.

Behavioural accuracy data in ERP experiment

The accuracy rate (averaged over all subjects) for the normal sentences, non-thematic role pragmatically violated sentences, thematic role animacy violated sentences and morphosyntactically violated sentences was 86.3%, 85%, 96%, and 89%, respectively. The difference in accuracy rate between the four sentence types was significant, $F(3, 57) = 4.14$, $p < .02$, with post-hoc pair-wise comparisons showing significantly greater accuracy to the thematic role animacy violated sentences than to the non-violated sentences, $t(19) = 4.17$, $p < .0001$, the non-thematic role pragmatically violated sentences, $t(19) = 3.70$, $p < .001$, or the morphosyntactically violated sentences, $t(19) = 2.33$, $p < .031$.

⁹ In the 500–800msec time window, there were significant inverse correlations between the plausibility ratings of the passivised versions of the sentences and the average voltages across the forty groups (more positive ERPs associated with lower ratings) at central (Spearman's $r = -0.34$, $p < .03$) and centro-parietal (Spearman's $r = -0.34$, $p < .03$) sites and inverse correlations that approached significance (Spearman's $r = -0.27$, $p < .09$) at parietal sites.

There were no differences in subjects' accuracy in comparing animacy violated sentences with transitive (97%) vs. intransitive (98%) verbs, $t(19) = 0.49$, $p = .62$. There were also no differences in subjects' accuracy between animacy violated sentences that, when passivised, were rated as most plausible through most implausible (Quartile 4: 97%; Quartile 3: 96%; Quartile 2: 97%; Quartile 1: 98%), $F(3, 58) = 0.37$, $p = .7$.

BEHAVIORAL STUDY: SPEEDED ANOMALY DETECTION

Introduction

The increased amplitude of the P600 elicited by morphosyntactic vs. animacy violations, intransitive vs. transitive animacy violated verbs and by critical verbs in animacy violated sentences that were rated as implausible vs. plausible when passivised, might have been driven by differences in whether participants perceived these sentences as anomalous on the critical verbs. Specifically, if a subgroup of participants on a subgroup of trials failed to register an anomaly on the critical verb, this might have led to a reduced cost of syntactic processing on that verb and a smaller P600 in the grand average across individual trials and participants because of poorer time-locking of the ERP to the underlying neurocognitive event.

To investigate this possibility, we conducted a study in which we measured subjects' reaction times to detect each type of anomaly. We predicted that, if the modulation of the P600 was driven by differences in when subjects perceived the anomaly across the different sentence types, then anomaly detection reaction times (RTs) would be inversely proportional to the amplitude of the P600, i.e., shortest RTs to the morphosyntactic violations, longer to intransitive or implausible-when-passivised critical verbs in animacy violated sentences, and longest to transitive or plausible-when-passivised critical verbs in animacy violated sentences.

Methods

Sixteen volunteers (seven women), aged between 19 and 33 years (average: 23.5 years) who had not taken part in the ERP experiment participated in this behavioural experiment. They were presented with the same sentences, counterbalanced across the same experimental lists with the same presentation parameters as in the ERP experiment. Subjects were asked to read the sentences carefully and to press a button labeled "no" as soon as they detected an anomaly. They were told that, if they did not detect an anomaly, they should press a second button labeled "yes" at the

end of the sentence, indicating that the sentence was correct. Participants then pressed a third button to move on to the next sentence trial.

Results

Accuracy

In some cases, subjects failed to detect the anomaly and pressed a button indicating that the sentence made sense. In other cases, subjects pressed a button to indicate that there was an anomaly when the sentence made sense. The error rates (averaged over all subjects) for the normal sentences (false negatives), non-thematic role pragmatically violated sentences (false positives), thematic role animacy violated sentences (false positives), and morphosyntactically violated sentences (false positives) were 11.1%, 10.6%, 4.3%, and 4.4% respectively.

An ANOVA revealed a significant main effect of sentence type, $F(3, 45) = 4.89, p < .005$, that arose because participants were significantly more accurate in judging the acceptability of both the animacy violated and the morphosyntactically violated sentences than the non-violated sentences (animacy vs. normal: $t(15) = 1.6, p < .003$; morphosyntactic vs. normal: $t(15) = 3.03, p < .008$) and the pragmatically violated sentences (animacy vs. pragmatic: $t(15) = 2.63, p < .019$; morphosyntactic vs. pragmatic: $t(15) = 2.08, p < .05$). There were no significant differences in accuracy between the animacy violated and the morphosyntactically violated sentences, $t(15) = 0.16, p = .87$.

Five per cent of the transitive animacy violated verbs and 6% of intransitive animacy violated verbs were false positives. An items analysis in which the accuracy to each thematic role animacy violated sentence was averaged across all subjects, revealed no significant differences in error rates between animacy violated sentences with intransitive vs. transitive verbs, $t(144) = 0.8, p < .42$. There were also no differences in subjects' error rates between animacy violated sentences that, when passivised, were rated as plausible vs. implausible: Quartile 4 (most plausible when passivised: 2% false positives), Quartile 3 (6% false positives), Quartile 2 (7% false positives) and Quartile 1 (least plausible when passivised: 3% false positives), $F(3, 237) = 2, p = .1$.

Reaction times

Figure 5 (top) summarises the reaction times to correctly detect anomalies within the three types of violated sentences. RTs were longest to detect pragmatic violations (1622 ms after the onset of the critical verb, $SD = 363$), shorter to detect the morphosyntactic violations (1380 ms, $SD = 412$), and slightly shorter still to detect the animacy violations (1347 ms, $SD = 299$). ANOVAs using both subjects (F_1) and items (F_2)

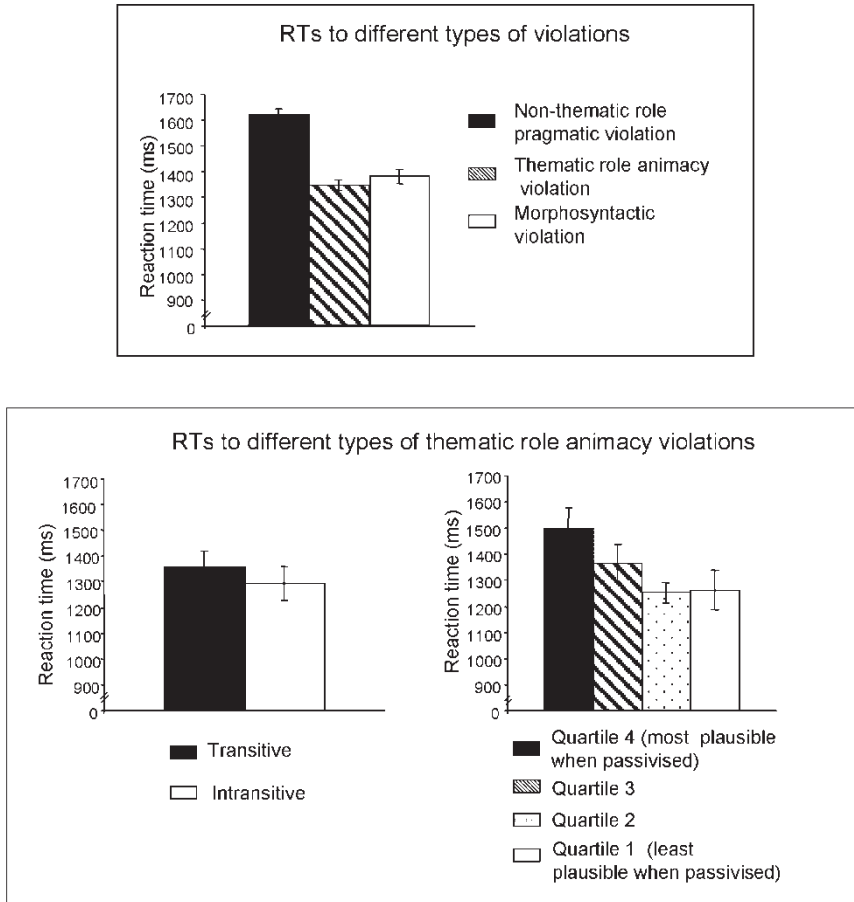


Figure 5. Mean reaction times to correctly detect anomalies in different types of violated sentences. Error bars indicate standard errors. Panel A: reaction times across the three main types of violated sentences: pragmatically violated, thematic role animacy violated, and morphosyntactically violated. Panel B, left: reaction times across thematic role animacy violated sentences containing transitive versus intransitive animacy violated verbs. Panel B, right: reaction times across thematic role animacy violated sentences in relation to independent ratings of how plausible these sentences became when passivised.

as random effects revealed an overall significant main effect of anomaly type, $F_1(2, 30) = 14.64, p < .001$; $F_2(2) = 40.6, p < .001$. Follow-up comparisons revealed significantly longer RTs to detect the pragmatic violations in comparison with the animacy violations, $t_1(15) = 7.0, p < .001$; $t_2(476) = 15.16, p < .001$, and in comparison with the morphosyntactic violations $t_2(15) = 4.10, p = .001$; $t_2(476) = 12.5, p < .001$, but no significant differences between RTs to detect the morphosyntactic and animacy violations, $t(15) = 1.3, p = .21$.

As shown in Figure 5 (bottom left), average RTs were slightly longer to detect transitive (1356 ms after the onset of the critical verb, $SD = 282$) than intransitive animacy violated verbs (1293 ms, $SD = 283$), although this difference was not significant on an items analysis in which the dependent variable was the RT to each thematic role animacy violated sentence averaged across all subjects, $t(38) = 0.7$, $p = .49$.¹⁰ Figure 5 (bottom right) summarises mean reaction times to correctly detect critical verbs in animacy violated sentences in relation to independent ratings of how plausible these sentences became when passivised. An items analysis revealed significant differences in RT across the four plausibility rating quartiles, $F(3) = 2.84$, $p < .04$. RTs to detect animacy violations in sentences that, when passivised, fell into quartile 4 (rated as most plausible: 1497 ms after the onset of the critical verb, $SD = 80$) were slightly longer (but not significantly, $t(29) = 1.2$, $p = .23$) than RTs to critical verbs in Quartile 3 (1364 ms, $SD = 74$), and significantly longer than RTs to critical verbs in Quartile 2 (1252 ms, $SD = 37$; $t(30) = 3$, $p < .005$) and Quartile 1 (1262 ms, $SD = 75$; $t(23) = 2.1$, $p < .04$).¹¹

Conclusion

Anomaly detection times were longer to pragmatic violations than to morphosyntactic or animacy violations but there were no differences in RTs to morphosyntactic or animacy violations. Thus, differences in when subjects perceive the violation cannot explain the larger P600 to morphosyntactic relative to animacy violations.

On the other hand, the shorter detection times for intransitive (vs. transitive) or implausible-when-passivised (vs. plausible-when-passivised) critical verbs in animacy violated sentences are consistent with the larger P600 effect elicited by these types of violation. They suggest that subjects detected intransitive or implausible-when-passivised violations sooner than transitive and plausible-when-passivised violations respectively, leading to a greater cost to syntactic processing on the verb itself and hence better time-locking of the P600 response to the critical verb across individual trials and subjects.

In conclusion, temporal differences in the way these different types of anomalies were perceived may explain differences in the amplitude of the P600 elicited by the different types of animacy violations but cannot explain difference in the amplitude of the P600 elicited by animacy violations and morphosyntactic violations. These findings are discussed further below.

¹⁰ In the analysis in which all responses were included (rather than correct responses only), this difference reached significance, $p < .05$.

¹¹ All these differences became more significant after covarying for word frequency.

DISCUSSION

We demonstrated that verbs violating the thematic role animacy constraints of their preceding NPs elicited a significant P600 effect. This replicates the findings of our previous study in a different group of subjects (Kuperberg et al., 2003b).

The main aim of the current study was to further investigate the nature of the P600 elicited by these animacy thematic role violations. There were two main findings. First, the P600 elicited by the animacy thematic role violations was similar in waveform, duration and scalp distribution but smaller in amplitude than the P600 elicited by the morphosyntactic violations. Second, critical verbs of animacy violated sentences that were transitive or that were rated as plausible when passivised elicited a less positive-going P600 than critical verbs of animacy violated sentences that were intransitive or that were rated as implausible when passivised. This remained the case when the frequency, concreteness, and the number of letters of critical verbs were matched across conditions.

As in our previous study, the thematic role animacy violated verbs elicited an N400 effect that was significantly smaller than the N400 elicited by non-thematic role pragmatically violated verbs. In the current study, the N400 elicited by the critical verbs in the animacy violated sentences was not modulated by the transitivity of these verbs or by plausibility ratings of passivised versions of these sentences.

A P600 to thematic role animacy violations

Our demonstration that a P600 and a relatively small N400 were evoked to thematic role animacy violations is consistent not only with our own previous findings (Kuperberg et al., 2003b) but also with three recent reports of a significant P600 effect to conceptually violated verbs in sentences. First, Kolk et al. (2003) demonstrated a significant P600 to verbs such as “fled” (that, in Dutch, occurs after its arguments, “cat” and “mice”) in sentences such as “The cat that from the mice fled ran across the room”/“De kat die voor de muizen vluchtte, rende door de kamer” (see also recent follow-up study by van Herten, Kolk, and Chwilla, 2005). Second, Hoeks et al. (2004) describe a significant P600 elicited by verbs such as “thrown” (again occurring after its arguments “javelin” and “athletes”) in sentences such as “The javelin has the athletes thrown”/“De speer heeft de atleten geworpen”. Third, Kim and Osterhout (2005) reported a P600 effect to verbs such as “devouring” in sentences such as, “The meal was devouring . . .”.

We suggest that what the sentences in these studies have in common is that, while the critical verb assigned a thematic role to its preceding arguments that rendered the overall meaning implausible, the semantic

SYNTACTIC AND SEMANTIC RELATIONSHIPS

features of the noun, in combination with semantic-thematic features of the verb, biased towards an alternative syntactic structure.¹² In other words, in all these studies, the semantics biased towards an interpretation that was contradicted by the syntax, leading to a cost of integrating the verb into its ongoing thematic structure. We suggest that this cost was manifest by a P600 effect. This interpretation is consistent with numerous online behavioural studies showing an influence of semantic and pragmatic information on syntactic ambiguity resolution (Boland & Tanenhaus, 1991; Ferreira & Clifton, 1986; MacDonald et al., 1994; Stowe, 1989; Tanenhaus & Carlson, 1989; Trueswell & Tanenhaus, 1994). It supports the theory that semantic and, at least some syntactic information, are activated in parallel and that there is online, rapid, word-by-word communication between a process that evaluates potential semantic relationships between verbs and their preceding arguments (whether or not there exist plausible thematic relationships between these items) and a process of online syntactic attachment and assignment of thematic roles.¹³

The results of the current study suggest two features of the process indexed by the P600. The first feature is that its trigger appears to be a *discrepancy or mismatch* between the most likely thematic roles suggested by the semantic relationship between a verb and its context and the implausible thematic roles assigned to arguments by the actual syntactic structure of a sentence,¹⁴ rather than by incompatibilities between those

¹² This bias may have arisen because of a relatively high semantic fit between critical verbs and their preceding content words (considered in isolation, out of the context of the sentence, e.g. thrown-javelin-athletes, cat-mice-fled). For example, it has recently been suggested that participants use a plausibility heuristic, checking semantic input against their world-knowledge (van Herten, Kolk, & Chwilla, 2005). The P600 may also have resulted in some of these studies, although not in Kolk et al. (2003) or van Herten et al. (2005), because the subject NP was inanimate in nature, biasing towards its interpretation as a Theme rather than an Agent.

¹³ Note, however, that there is some disagreement about whether or not semantic and syntactic information come together during a first or second stage of parsing. Our data do not address this question. In some models, lexico-semantic information is thought to influence a second stage of processing (Ferreira & Clifton, 1986). Others suggest that lexical semantic and other all types of information (including contextual discourse information) information can influence immediate parsing decisions (Boland & Tanenhaus, 1991; MacDonald et al., 1994; Stowe, 1989; Tanenhaus & Carlson, 1989; J. C. Trueswell & Tanenhaus, 1994).

¹⁴ An open question is whether a P600 is elicited only by verbs in such situations, or whether it can also be elicited by nouns that are unlikely to occupy the thematic roles assigned to them by the syntax but whose meaning either biases towards an alternative thematic roles or is related to preceding individual content words, e.g. by NPs that are unlikely Agents in passive constructions (e.g. “eggs” in “For breakfast the toast was eaten by the eggs every day”), or NPs that are unlikely Instruments (e.g. ‘freezer’ in ‘The boy ate his icecream with the freezer after dinner’).

arguments and the context in which the entire set of thematic roles occurs. The evidence for this is that pragmatically violated non-thematic role violated verbs in the current study did not elicit a P600. These pragmatically violated verbs assigned plausible thematic roles to their preceding argument; the pragmatic violations arose at the level of relating those thematic roles to the entire preceding context.

The second feature of the process reflected by the P600 suggested by the current study is that it is likely to be syntactic in nature. The evidence for this is that the P600 elicited by thematic role violations was similar to that elicited by the morphosyntactic violations.¹⁵ First, the overall morphology of the P600 elicited by morphosyntactic violations was similar to the P600 elicited by animacy thematic role violations. Second, in directly comparing the P600 to these two types of violations, there were no significant interactions between sentence type and electrode site or hemisphere, suggesting that the P600 effects elicited by both these violations had similar scalp distributions. Finding that two effects have similar scalp distributions does not necessarily imply that they have identical underlying sources. Nonetheless, these similarities do suggest that processing these two types of violations engage similar syntactic neurocognitive processes that are qualitatively distinct from the process reflected by the N400. This interpretation is consistent with theories that the P600 indexes the cost of syntactic processing (Osterhout et al., 1994) and, given the discussion above, that integrating a verb into its ongoing thematic structure is a syntactic process (Fiebach et al., 2002; Kaan et al., 2000).

Modulation of the amplitude of the P600

The P600 elicited by the morphosyntactic violations was more positive than the P600 elicited by the animacy violations. Within the animacy violated sentences, critical verbs that were intransitive or that were rated as implausible when passivised elicited a more positive-going P600 than critical verbs that were transitive or that were rated as plausible when passivised. Below we consider two explanations for this pattern of P600

¹⁵ Of note, the morphosyntactic violations in this study did not elicit a LAN (see footnote 1), replicating our previous findings (Kuperberg et al., 2003a). It has been proposed that the LAN reflects a stage of first-pass parsing (the establishment of an initial syntactic structure) that only later interfaces with semantic information). If this is the case, our findings are consistent with the idea that morphosyntactic information does not play a significant role in the establishment of such an initial structure (Rayner et al., 1983). Future studies will need to address the question of why a LAN is elicited by morphosyntactic violations in some studies but not others.

SYNTACTIC AND SEMANTIC RELATIONSHIPS

modulation across these three conditions: first, differences across conditions in lexico-semantic relationships between critical verbs and their preceding content words, and, second, differences across conditions in the precise thematic fit between critical verbs and their preceding argument(s). Neither explanation alone can account for all our findings.

Lexico-semantic associations between critical verbs and preceding context. The morphosyntactically violated sentences were derived from highly contextually constrained normal sentences. The critical verbs and preceding NPs in these highly constrained normal sentences were more closely semantically associated than in the animacy thematic role violated sentences (see ratings described in Methods). It may therefore be that the closer the semantic association between the critical anomalous word and individual words in its preceding context, the greater the P600 effect. In other words, the amplitude of the P600 may be dictated by the degree of discrepancy between the close semantic associations between critical verbs and their preceding context and the incongruous thematic roles assigned by the syntax. By this account, the P600 reflects the cost of syntactically integrating the verb into an ongoing thematic structure once semantic structures have been activated in parallel (Gibson, 1998; Kaan et al., 2000): the greater the discrepancy between semantic and syntactic information that have been activated, the harder to syntactically integrate the verb and the larger the P600. This explanation is consistent with the findings of Gunter, Friederici, and Schriefers (2000) who demonstrated that the P600 elicited by morphosyntactic anomalies was greater when such anomalies occurred within sentences of high vs. low semantic constraint.

An account by which the amplitude of the P600 is determined entirely by the degree of discrepancy between syntactic and semantic relationships between critical verbs and their preceding context, however, does not alone explain the pattern of findings in the current study. It cannot account for why the amplitude of the P600 elicited by transitive or plausible-when-passivised animacy violated verbs was smaller than the P600 elicited by intransitive or implausible-when-passivised animacy violated verbs. This is because our rating study (in which we extracted just the critical verbs and their preceding content words from the animacy violated sentences) demonstrated that the semantic associative fit between the critical verb and its preceding content words was equally high in animacy violated sentences that were transitive versus intransitive and that were rated as plausible vs. implausible-when-passivised. We therefore suggest that, once the processing system has detected a discrepancy between semantic and syntactic information, it attempts to repair the anomaly by reassigning thematic roles and the difficulty of this reassignment can modulate the amplitude of the P600.

Thematic relationships between critical verbs and preceding arguments. The amplitude of the P600 was modulated by the transitivity of critical verbs within animacy violated sentences and by plausibility ratings of passivised versions of these sentences. As discussed in the Introduction, we conceptualised both transitivity and these plausibility ratings as reflecting the potential fit between the critical verb and its preceding inanimate NP if that NP were to play an alternative thematic role around the verb in these sentences.¹⁶ The P600 elicited by transitive and plausible-when-passivised verbs within animacy violated sentences was less positive than the P600 effect elicited by intransitive and implausible-when-passivised verbs within animacy violated sentences. In other words, for a given semantic associative fit between the critical verb and its preceding content words, the more implausible the potential thematic fit, the more positive the P600 effect. We interpret this observation as evidence that participants attempted to reassign thematic roles to make sense of the sentence and that the P600, in part reflected this process of thematic reanalysis/syntactic restructuring: the more positive the P600 on a verb, the more effort required to reassign thematic roles so that the sentence would make sense. The degree to which participants engaged in thematic reanalysis/syntactic restructuring at the point of the verb may have been influenced by the time-point at which they recognised the anomaly and/or by the intensity with which they engaged in repair once the anomaly was recognised. We consider each of these accounts below.

A less positive P600 on verbs with a plausible (vs. an implausible) possible thematic fit with their preceding arguments might have arisen because of a relative *delay* in syntactic restructuring due to a failure of participants to recognise the anomaly at the point of the critical verb. This may have, in turn, have led to worse time-locking of ERP response to the critical verb across individual trials and participants. The findings of our behavioural study support this latter explanation. In an anomaly detection

¹⁶ As reported in the Methods section, the subcategorisation properties (transitivity) of the critical verbs partly but not completely correlated with the plausibility ratings of the passivised sentences. Of the animacy violated sentences that were rated as most implausible when passivised (lowest quartile), nearly half contained critical verbs that were strictly/preferentially intransitive, nearly a quarter contained critical verbs that were strictly/preferentially transitive and just over a quarter contained critical verbs that did not fall into either of these categories. Of the animacy violated sentences that were rated as most plausible when passivised (highest quartile), only 5% contained critical verbs that were intransitive, nearly a half contained critical verbs that were strictly/preferentially transitive and just over a half contained critical verbs that did not fall into either of these categories.

task, participants found it easier (shorter RTs) to detect intransitive than transitive critical verbs and implausible-when-passivised than plausible-when-passivised critical verbs in animacy violated sentences. This suggests that intransitive and implausible-when-passivised violations on the critical verb were perceived as more obvious, leading to a greater cost to syntactic processing on the verb itself and hence better time-locking of the P600 response to the critical verb across individual trials and subjects.¹⁷

A more positive P600 to intransitive and implausible-when-passivised relative to transitive and plausible-when-passivised verbs might have also arisen because it was more difficult to repair the former than the latter structures in the animacy violated sentences presented here. Although we cannot say exactly what syntactic structures were constructed as part of the effort to re-assign thematic roles, we suggest that the parser may have attempted to create the simplest possible syntactic structure to make sense of these sentences. This is likely to be a structure in which the actual subject is assigned the grammatical role of object of the verb or the object of a preposition. In the case of verbs with a plausible (vs. an implausible) possible thematic fit with their preceding arguments, participants may have made fewer attempts to syntactically restructure such sentences (leading to a smaller P600) before discarding the revised syntactic structure and interpreting the sentence as determined by its actual syntax.¹⁸ By this account, transitive verbs elicited a less positive P600 than intransitive verbs because, in assigning the actual subject to the role of object during syntactic restructuring, there was no need to look outside the existing thematic grid to make sense of the sentence. In the case of the intransitive verbs, however, in order to make sense of the sentence, the participant needed to go outside the existing thematic grid by inserting a preposition before the new object Theme.

¹⁷ We thank an anonymous reviewer for suggesting this explanation and the behavioral experiment. Note that this argument cannot explain the more positive P600 elicited by the morphosyntactic violations in comparison with the animacy violations: in the ERP study, there were, in fact, more errors to the morphosyntactically violated sentences than to the animacy violated sentences and, in the behavioral study, anomaly detection RTs were longer to morphosyntactic violations than to animacy violations. Taken together, these data suggest that animacy violations were easier to detect than morphosyntactic violations, predicting a larger P600 to the former than the latter violations.

¹⁸ This is evidenced by the fact that participants almost always came to the correct conclusion that these thematic role animacy violated sentences were anomalous, i.e. that the inanimate subject was the Agent (see MacWhinney, Bates, and Kliegl, 1984, and Ferreira, 2003, for further evidence that participants correctly identify the Agent in implausible simple sentences).

To conclude, we interpret the more positive P600 elicited by critical verbs that were intransitive or that were within sentences that were rated as implausible when passivised, as reflecting a more prolonged and/or greater attempt to reanalyse and repair such sentences across individual trials and subjects. These findings may be specific to the task participants performed in this study: to judge whether the sentences made sense. It will be interesting to determine whether the same pattern of ERP findings are observed if participants do not perform an overt plausibility judgement task.

Summary and implications of findings for models of language processing

To summarise, the findings discussed above suggest that the P600 elicited by thematic role animacy violations reflects similar syntactic processes to those elicited by morphosyntactic violations and that its modulation was influenced by the precise relationship between the verb and its preceding NP: its amplitude was *increased* by a discrepancy between the close semantic association between verbs and their arguments and the incongruous thematic roles assigned by the syntax, but *attenuated* when there was a potential for the critical verb and its preceding argument(s) to fit together thematically and the violation was less obvious. We have interpreted the P600 to a discrepancy between semantic and syntactic information as reflecting the cost of syntactically integrating the verb into an ongoing thematic structure once these semantic and syntactic structures have been activated in parallel (Gibson, 1998; Kaan et al., 2000). We have suggested that the closer semantic associations between critical words and their preceding arguments in the morphosyntactically violations vs. the thematic role animacy violations leads to a greater discrepancy between semantic and syntactic information in the morphosyntactic violations than the thematic role violations, and that this may account for the more positive P600 to the former than the latter violations, although this hypothesis will require systematic follow-up. In addition, we have also suggested that, once a discrepancy between semantic and syntactic incoming information has been detected and a cost in syntactic integration is incurred, there is an attempt to make sense of the sentences through syntactically restructuring and the reassignment of thematic roles (Friederici, 1997), at least when participants are asked to make plausibility judgements, and that this too can modulate the amplitude of the P600. The easier it is to repair a sentence through reassignment of thematic roles, the smaller the P600, accounting for the modulation of the P600 within the different types of thematic role animacy violated sentences. These interpretations are consistent with the idea that making sense of sentences

engages both parallel and serial processes. They are also consistent with the idea that the P600 reflects more than one syntactic process.¹⁹

Conclusions

In conclusion, this study provides further evidence that the P600 reflects a syntactic process that can be modulated by semantic factors, supporting the idea that, under certain circumstances, the neural systems supporting syntactic and semantic processing may be linked (Kuperberg et al., 2003b). Our data suggest that this interaction can occur both at a stage when both semantic and syntactic structures have been activated in parallel, as well as at a stage of attempted syntactic reanalysis and repair. On the other hand, semantic processes reflected by the N400 were not modulated by the transitivity of critical verbs or by the precise thematic relationships between the critical verb and its preceding arguments, suggesting that some distinction between the semantic and syntactic processing systems is respected in the brain.

Manuscript received May 2003

Revised manuscript received February 2005

PrEview proof published online August 2005

REFERENCES

- Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*, 343–355.
- Boland, J., & Tanenhaus, M. K. (1991). The role of lexical representation in sentence processing. In G. B. Simpson (Ed.), *Understanding word and sentence* (pp. 331–366). Amsterdam: North Holland.
- Brown, C. M., van Berkum, J. J., & Hagoort, P. (2000). Discourse before gender: An event-related brain potential study on the interplay of semantic and syntactic information during spoken language understanding. *Journal of Psycholinguistic Research*, *29*, 53–68.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Connine, C., Ferreira, F., Jones, C., Clifton, C., & Frazier, L. (1984). Verb frame preferences: Descriptive norms. *Journal of Psycholinguistic Research*, *13*, 307–319.

¹⁹ Aspects of the P600 that was modulated by transitivity and plausibility ratings differed from the aspects of the P600 that was modulation by animacy and morphosyntactic violations in relation to non-violated verbs. Differences in the amplitude of the P600 elicited by intransitive versus transitive verbs as well as by implausible-when-passivised versus plausible-when-passivised verbs, appeared in the earlier portion of the P600 time window (between 500-700msec) and were relatively equally distributed along the anterior-posterior axis of the scalp (i.e. there were no significant electrode site by condition interactions). The P600 effect elicited by animacy or morphosyntactic violations (versus non-violated verbs) was more posteriorly distributed. It remains unclear, however, which aspects of syntactic processing are mediated by different subcomponents of the P600 complex (see Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001; Hagoort & Brown, 2000; Kaan & Swaab, 2003 for debate).

- Donchin, E., & Coles, M. G. H. (1988). Is the p300 component a manifestation of context updating? *Behavioral and Brain Science*, *11*, 355–372.
- 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.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.
- Ferreira, F., & Clifton, C., Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.
- Fiebach, C. J., Schlesewsky, M., & Friederici, A. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German wh-questions. *Journal of Memory and Language*, *47*, 250–272.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally-ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.
- 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. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–281.
- Friederici, A. D. (1997). Neurophysiological aspects of language processing. *Clinical Neuroscience*, *4*(2), 64–72.
- Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology, Learning, Memory and Cognition*, *22*(5), 1219–1248.
- Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research*, *11*(2), 305–323.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.
- Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.
- Gunter, T. C., & Friederici, A. D. (1999). Concerning the automaticity of syntactic processing. *Psychophysiology*, *36*, 126–137.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: Erps reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*, 556–568.
- Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*, 660–676.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an ERP measure of syntactic processing. In S. M. Garnsey (Ed.), *Language and cognitive processes. [Special issue: Event-related brain potentials in the study of language]* (Vol. 8(4), pp. 439–483). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: The p600/sps to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, *38*, 1531–1549.
- 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.

SYNTACTIC AND SEMANTIC RELATIONSHIPS

- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The p600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*, 159–201.
- 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 combinatorial semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, in press.
- King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level erps to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*, 376–395.
- Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from erps on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196–214.
- Kolk, H. H., Chwilla, D. J., van Herten, M., & Oor, P. J. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*, 1–36.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present day American English*. Kuperberg, G. R., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003a). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*, 272–293.
- Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003b). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, *17*, 117–129.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *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.
- Ladusaw, W. A., & Dowty, D. (1988). Towards a non-grammatical account of thematic roles. In W. Wilkins (Ed.), *Syntax and semantics* (Vol. 21: Thematic relations). London: Academic Press.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.
- MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Memory and Language*, *23*, 127–150.
- Marslen-Wilson, W. D., Brown, C., & Tyler, L. K. (1988). Lexical representations in language comprehension. *Language and Cognitive Processes*, *3*, 1–17.
- McCarthy, G., & Wood, C. C. (1986). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208.
- McKinnon, R., & Osterhout, L. (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Language and Cognitive Processes*, *11*, 495–523.
- Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition*, *23*, 477–494.
- Munte, T. F., Heinze, H. J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia*, *36*(3), 217–226.
- Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3*, 151–165.

- Oken, B. S., & Chiappa, K. H. (1986). Statistical issues concerning computerized analysis of brainwave topography. *Annals of Neurology*, *19*(5), 493–497.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- 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. In S. M. Garnsey (Ed.), *Language and cognitive processes*. [Special issue: Event-related brain potentials in the study of language] (Vol. 8, pp. 413–437). Hove, UK: Lawrence Erlbaum Associates Ltd.
- Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology, Learning, Memory and Cognition*, *20*, 786–803.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34*, 739–773.
- Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, *14*, 283–317.
- Pavio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, *76* (3, part 2).
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358–374.
- Stowe, L. A. (1989). Thematic structures and sentence comprehension. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structures in language processing* (pp. 318–357). Dordrecht: Kluwer Academic.
- Tanenhaus, M. K., & Carlson, G. (1989). Lexical structure and language comprehension. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 529–561). Cambridge, MA: MIT Press.
- Trueswell, J. C., & Tanenhaus, M. K. (1994). Towards a lexicalist framework of constraint-based syntactic ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 155–179). Hillsdale, NJ: Lawrence Erlbaum Associates Ltd.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318.
- Tyler, L. K. (1992). Verb-argument structures. In L. K. Tyler (Ed.), *Spoken language comprehension: An experimental approach to disordered and normal processing* (pp. 107–124). London: MIT Press.
- Urbach, T. P., & Kutas, M. (2002). The intractability of scaling scalp distributions to infer neuroelectric sources. *Psychophysiology*, *39*(6), 791–808.
- 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 Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*, 241–255.
- Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, *36*(5), 559–570.
- White, K., & Ashton, R. (1976). Handedness assessment inventory. *Neuropsychologia*, *14*(2), 261–264.