LANGUAGE AND COGNITIVE PROCESSES
0000, 00 (00), 1–40

Psychology Press
Taylor & Francis Group

The linguistic processes underlying the P600

Ana C. Gouvea

Department of Communication Sciences and Disorders, Florida International University, Miami, FL, USA

Colin Phillips

Department of Linguistics, Neuroscience & Cognitive Science Program, University of Maryland, College Park, MD, USA

Nina Kazanina

Department of Experimental Psychology, University of Bristol, Bristol, UK

David Poeppel

Department of Linguistics, Neuroscience & Cognitive Science Program, University of Maryland, College Park, MD, USA and Department of Psychology, New York University, New York, NY, USA

The P600 is an event-related brain potential (ERP) typically associated with the processing of grammatical anomalies or incongruities. A similar response has also been observed in fully acceptable long-distance *wh*-dependencies. Such findings raise the question of whether these ERP responses reflect common underlying processes, and what might be the specific mechanisms that are shared between successful processing of well-formed sentences and the detection and repair of syntactic anomalies. The current study presents a comparison of the ERP responses elicited by syntactic violations, garden path sentences, and long-distance *wh*-dependencies, using maximally similar

Correspondence should be addressed to Ana C. Gouvea, Department of Communication Sciences and Disorders, Florida International University, HLS 146, 11200 SW 8th St, Miami, FL 33199. E-mail: agouvea@fiu.edu

We would like to thank Virginie van Wassenhove, Anthony Boemio, and Anna Salajegheh for their help with EEG recordings, Anna Salajegheh and Beth Rabbin for their assistance in the preparation of materials, and Brian Dillon for his help with data analyses. This work was supported in part by NSF BCS-0196004, NSF DGE-0801465 and Human Frontiers Science Program RGY-0134 to CP, by NIH R01 DC 05660 to DP and by a fellowship to DP from the Wissenschaftskolleg zu Berlin.

http://www.psypress.com/lcp DOI: 10.1080/01690960902965951

materials in a within-subjects design. Results showed that a P600 component was elicited by syntactic violations and garden path sentences, but was less robustly elicited in the long-distance *wh*-dependency condition. Differences in the scalp topography, onset and duration of the P600 effects are characterised in terms of the syntactic operations involved in building complex syntactic structures, with particular attention to retrieval processes, which control the latency of the P600, and structure building processes, which control its duration and amplitude.

Keywords: ambiguity; Event-related potentials; Left anterior negativity (LAN); P600; Sentence processing; Syntactic dependencies.

INTRODUCTION

A primary goal of sentence processing research is to understand how speakers are able to successfully and accurately build up representations of sentences in real time. An important step in this research program is the attempt to characterise the detailed time-course of operations involved in successful parsing. Event-related brain potentials (ERPs) hold a great deal of promise in this regard, since they reflect neuronal activity related to language processing with millisecond accuracy. Moreover, previous research has shown that electrophysiological responses differ reliably in timing, amplitude, and scalp distribution as a function of different linguistic manipulations involving phonology, syntax, and semantics, to name but a few.

Much has been learned from the fact that linguistic anomalies involving semantics, morphology or syntax elicit a series of different characteristic ERP response components. The fact that ERPs measure activity that is time-locked to the presentation of eliciting stimuli may account for their particular sensitivity to the processing of events that are unexpected. Studies of syntactic and morphological anomalies, for instance, have demonstrated qualitatively different ERP responses to different types of violations. A wide variety of syntactic anomalies due to ungrammaticality or temporary misanalysis ('garden path sentences') elicit a broad posterior positive ERP component known as the P600 or as the Syntactic Positive Shift (SPS) (Friederici, Pfeifer & Hahne, 1993; Hagoort, Brown, & Groothusen, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992; Osterhout, Holcomb, & Swinney, 1994). A narrower set of anomalies, primarily involving ungrammaticality, have been shown to elicit a somewhat earlier ($\sim$ 300–500 ms latency) negative deflection known as the (Left) Anterior Negativity ((L)AN), due to its characteristic scalp distribution (Coulson, King, & Kutas, 1998a; Friederici et al., 1993; Gunter, Friederici, & Schriefers, 2000; Hagoort, Wassenaar, & Brown, 2003a) (e.g., *Every Monday he mow the lawn*). An even narrower class of anomalies, typically characterised as involving 'phrase structure violations' such as *The students enjoyed Bill's of review the play* is argued to elicit an Early Left Anterior

Negativity (ELAN), a component with a latency of 125–250 ms (Friederici et al., 1993; Hahne & Friederici, 1999; Lau, Stroud, Plesch, & Phillips, 2006; Neville et al., 1991).

However, we cannot know a priori whether the brain responses associated with the detection and diagnosis of different kinds of linguistic anomalies are related to the processes involved in successful processing of well-formed linguistic input. The link between the processing of violations and the processing of well-formed sentences must be established through empirical studies and explicit models. In the case of the ERP activity associated with semantic anomalies, this link is now well established. It is now understood that the N400 response associated with semantic anomalies is a response component that is elicited by all content words (Kutas, 1997; Kutas & Hillyard, 1980), with an amplitude that varies with a number of lexical and discourse properties of the incoming word, including cloze probability (Hagoort & Brown, 1994; Kutas & Hillyard, 1984; Kutas & van Petten, 1994) and similarity to an expected word (Federmeier & Kutas, 1999; for reviews see Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008). In the case of responses associated with syntactic anomalies, a small number of studies have begun to forge a similar link between anomaly detection and successful processing of well-formed input.

Development of a link between studies of anomaly detection and models of successful processing has been restricted by two factors. First, there have been only a limited number of ERP studies that have presented side-by-side comparisons of anomalous and well-formed sentences that elicit a P600 (Kaan, Harris, Gibson, & Holcomb, 2000; Osterhout et al., 1994). Second, there are few detailed cognitive or computational accounts of how syntactic anomalies might be detected in real-time, and what processes this might share with the processing of well-formed sentences, making it difficult to generate specific predictions about the connection between ERP responses to the two sentence types. A notable exception is a neurocognitive model presented by Hagoort (2003) that is directly related to an implemented computational model of sentence parsing (Vosse & Kempen, 2000). In this article, we aim to contribute to both of these areas by further investigating the parallels between the processing of anomalous and well-formed sentences, and by considering the question of what might be the common processes underlying the processing of well-formed sentences and anomaly detection. The focus of the study is the response (or family of responses) known as the P600, although we also address the difference between the P600 and the LAN response. We present a comparison of the ERPs elicited by closely matched sentences involving *wh*-dependency formation, ungrammaticality, and garden path sentences, and we propose a theoretical approach to understanding the variation in the responses to these different sentence types. In particular, we argue that the P600 component reflects the processes involved in creating (and destroying) syntactic relations. We propose that structural relations cannot be formed until the participants in those relations are available, and consequently that the onset latency of the P600 is a function of recognition and retrieval times, and the amplitude/duration of the P600 is a function of the structure building operations themselves. We begin by reviewing previous findings about P600 responses elicited by anomalous and well-formed sentences.

P600-type responses elicited by syntactic anomalies

The P600 or Syntactic Positive Shift (SPS) was first reported in cases of syntactic and morphological violations. For example, Hagoort, Brown, and Groothusen (1993a) observed a P600 response in cases of number agreement mismatch between the subject and verb of sentences in Dutch such as (1). A positive-going deflection in the ERP response was observed starting around 500 ms after the presentation of the anomalous verb, peaking around 600 ms, and lasting for at least 500 ms.

(1) * Het verwende kind *gooien* het speelgoed op de grond.
'The spoilt child *throw* the toys on the ground.'

There has been interesting debate about whether the P600 reflects language-specific processes or is a special case of the more general-purpose P300 family of ERP components. This debate has spanned discussions of whether the P600 to linguistic materials exhibits properties of the P300 (Coulson et al., 1998a; Coulson, King, & Kutas, 1998b; Osterhout & Hagoort, 1999; Osterhout, McKinnon, Bersick, & Corey, 1996), and discussions of late positivities elicited by anomalies in areas other than natural language syntax, such as spelling errors (Münte, Heinze, Matzke, Wieringa, & Johannes, 1998), music (Patel, Gibson, Ratner, Besson, & Holcomb, 1998), mathematical sequences (Martín-Loeches, Casado, Gonzalo, de Heras, & Fernández-Frías, 2006; Núñez-Peña & Honrubia-Serrano, 2004), and word stress (Domahs, Wiese, Bornkessel-Schlesewsky, & Schlesewsky, 2008). Many studies have shown that late positivities are not confined to syntactic anomalies, but there is also interesting evidence from patients with basal ganglia lesions for dissociation of the P300 and P600 components (Frisch, Kotz, von Cramon, & Friederici, 2003). These findings are not incompatible, as it is possible that the P600 reflects processes that are not specific to language, while still being functionally distinct from the P300. The current study is not well suited to test this issue, since its focus is on linguistic manipulations, but in the Discussion we offer brief remarks on how our account of fine-grained linguistic processes might generalise to other domains.

The P600 has been observed in response to a variety of different syntactic violations, including phrase-structure violations (Hagoort et al., 1993; Neville et al., 1991; Osterhout & Holcomb, 1992), subcategorisation violations (Ainsworth-Darnell, Shulman, & Boland, 1998; Osterhout & Holcomb, 1992; Osterhout et al., 1994), violations of number, tense, gender, and case agreement (Allen, Badecker, & Osterhout, 2003; Coulson et al., 1998a; Gunter, Stowe, & Mulder, 1997; Hagoort et al., 1993; Münte, Szentkui, Wieringa, Matzke, & Johannes, 1997; Nevins, Dillon, Malhotra, & Phillips, 2007) and violations of constraints on long-distance dependencies (Kluender & Kutas, 1993; McKinnon & Osterhout, 1996; Neville et al., 1991).

The P600 has also been elicited in cases of syntactic garden-path effects, i.e., syntactic anomalies that result from misanalysis of an ambiguity rather than from ungrammaticality (Kaan & Swaab, 2003a; Osterhout & Holcomb, 1992, 1993; Osterhout et al., 1994). For example, Osterhout et al. (1994: Experiment 1) examined the structures illustrated in (2) below. In (2a) the verb *charge* allows either an NP direct object or a clausal complement, and therefore the NP *the defendant* is initially ambiguous between a direct object and embedded clause subject analysis. The preference to initially analyse this NP as a direct object, which may reflect either general structural biases (Frazier & Rayner, 1982) or lexically specific constraints (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Trueswell, Tanenhaus, & Kello, 1993), leads to a garden-path effect when the disambiguating auxiliary *was* is parsed. Osterhout and colleagues showed that the ERP response to the auxiliary in (2a) elicits a P600, relative to the response to the unambiguous control in (2b).

(2) a. The lawyer charged the defendant was lying.
b. The lawyer charged that the defendant was lying.

Another important aspect of the Osterhout et al. (1994) study is that it manipulated the strength of the lexical bias of the main verb to take a following NP as a direct object and showed that the amplitude of the P600 response varied as a function of the lexical bias of the main verb. Although there is a good deal of variation in the amplitude of the P600 response observed in different studies and in response to different syntactic anomalies, this is one of only few studies to have shown parametric manipulation of P600 amplitude. This reflects in part the scarcity of continuous measures of 'syntactic fit', in contrast to the readily available continuous measures of semantic fit that have been exploited extensively in studies of the N400.

Syntactic violations and syntactic garden paths have been shown to elicit a late positive ERP component with a similar latency. It has been suggested that there might be systematic differences in the topographic distribution of the P600 elicited by violations and garden paths (Hagoort, Brown, & Osterhout, 1999), since some studies have observed a more strongly posterior scalp distribution of the P600 in response to violations than in response to garden paths (Osterhout & Holcomb, 1992). However, even among studies that have focused on syntactic violations, there has been some variation in the overall posterior scalp distribution of the P600, with some studies showing more lateralisation (e.g., Osterhout & Holcomb, 1992) and other studies reporting a more symmetrical distribution (e.g. Neville et al., 1991). Thus, it remains unresolved whether scalp differences in the P600 reflect systematic differences in the underlying cognitive processes (cf. Coulson et al., 1998a).

Attempts to understand systematic variation in the timing of the late positive response have been more successful. Within an approach that views the P600 as the reflection of syntactic reanalysis and repair processes (e.g., Friederici, 1995; Osterhout & Holcomb, 1992), it has been proposed that the latency of the P600 reflects the difficulty of completing the diagnosis or reanalysis of the anomalous structure. Friederici and Mecklinger (1996) and Friederici, Mecklinger, Spencer, Steinhauer, and Donchin (2001) compared ambiguous subject and object relative clauses and also compared ambiguous subject-first and object-first complement clauses in German. In both comparisons, a late positivity was elicited when the clause-final verb disambiguated in favour of the dispreferred object relative clause or object-first complement clause analysis. However, the late positivity began earlier in the object relative clause condition, a difference that the authors attribute to the greater ease of reanalysis in German relative clauses. The earlier part of the late positivity is treated as a distinct response component called the 'P345', and has been associated with the diagnosis stage of structural reanalysis, following a model of reanalysis proposed by Fodor and Inoue (1994).

P600-type responses elicited by congruous sentences

In contrast to the many studies that have reported P600-type responses to syntactic anomalies, a smaller number of recent studies have observed that similar responses are elicited in fully well-formed sentences, at the point of completion of a long-distance structural dependency (Felser, Clahsen, & Münte, 2003; Fiebach, Schlesewsky & Friederici, 2002; Kaan et al., 2000; Phillips, Kazanina, & Abada, 2005). In an influential study, Kaan and colleagues compared sentences like (3a-b) that contain fronted direct object *wh*-phrases (i.e., 'who' and 'which pop star') with sentences in which the direct object appears in its canonical position (3c), and showed that a P600 response was elicited at the verb position (underlined) in the *wh*-fronting conditions, relative to the control.

(3) a. Emily wondered who the performer in the concert had <u>imitated</u> for the audience's amusement.

b. Emily wondered which pop-star the performer in the concert had imitated for the audience's amusement
c. Emily wondered whether the performer in the concert had imitated a pop star for the audience's amusement.

Kaan and colleagues suggested that the P600 observed at the verb reflects the increased structural processing that occurs at that position in the *wh*-fronting conditions. They also make the more specific claim that the P600 reflects the demands on working memory of linking the *wh*-phrase to the verb across several intervening words, suggesting that the amplitude of the P600 is an index of the 'syntactic integration difficulty' metric proposed by Gibson (1998).

Although the polarity, latency, and scalp distribution of this ERP effect resembled the P600 effect widely reported in studies of syntactic anomaly, Kaan and colleagues sought to more directly test the parallels in a follow-up study that independently manipulated the presence of wh-fronting and the presence of a subject-verb agreement violation, as in the examples in (4).

(4) a. Emily wonders whether the performers in the concert imitate(s) a pop star for the audience's amusement.
b. Emily wonders who the performers in the concert imitate(s) for the audience's amusement

Results from the follow-up study showed that the *wh*-fronting manipulation and the agreement violation manipulation elicited broadly similar P600-type responses. Both responses became significant in the 500–700 ms time interval, and reached maximum in the 700–900 ms interval, and both displayed peak amplitudes at mid-posterior scalp locations. Although certain statistical differences were found, these were insufficient to warrant the conclusion that the two effects were generated by independent sources.

Small differences may have arisen in the study by Kaan and colleagues due to lexical differences between the agreement violation conditions and the *wh*-dependency conditions. The P600 elicited by processing of the verb appeared during the presentation of the following word, which was a determiner ('the') in the agreement violation conditions and a preposition in the *wh*-dependency conditions. Another potential concern is that the completion of the *wh*-dependency in this study also involved a type of structural ambiguity resolution, since the case and thematic role of a fronted NP remained uncertain until the critical verb was processed. One of the goals of the current study was to control for these issues.

Some other studies have reported that completion of a syntactic dependency elicits a P600 response component, although they have not directly compared this response with the P600 response elicited by ungrammatical or garden path sentences (Fiebach et al., 2002; Phillips et al., 2005). Our study compares the P600 generated by well-formed long-distance dependencies, garden-path and ungrammatical sentences in order to examine the differences and similarities between anomaly detection and successful processing. We also discuss psycholinguistic models that could explain our results.

Some additional ERP studies have reported P600s elicited by sentences that are strictly well-formed, but that are likely temporarily perceived as ill-formed, due to errors in agreement or reference resolution that are similar in nature to garden paths (e.g., Kaan & Swaab, 2003b; Osterhout, Bersick, & McLaughlin, 1997; van Berkum, Koornneef, Otten, & Nieuwland, 2007).

A further class of recent studies have shown P600s elicited by grammatically well-formed but semantically anomalous sentences in which the subject is a poor agent of the verb and is either unrelated to the verb, as in *Every morning at breakfast the eggs would plant* . . . (Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Stroud & Phillips, 2008) or is semantically related to the verb, as in *The hearty meal was devouring the kids* (Kim & Osterhout, 2005; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). Similar effects have been observed in studies in Dutch involving inappropriate or reversed argument phrases (Hoeks, Stowe, & Doedens, 2004; Kolk, Chwilla, van Herten, & Oor, 2003; van Herten, Kolk, & Chwilla, 2005). These findings are striking, in the respect that they involve sentences that may be viewed as grammatically well-formed but semantically ill-formed. The findings have attracted a number of different accounts, which have attributed the P600 either to the detection of a conflict between the outputs of different analysis streams in sentence comprehension (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kim & Osterhout, 2005; van Herten et al., 2005) or to the detection of an anomaly in thematic role assignment (Hoeks et al., 2004; Stroud, 2008), or to both of these factors (Kuperberg, 2007). Importantly, there is a consensus that the P600 observed in these studies should not be viewed as qualitatively or functionally distinct from the P600 observed in studies of grammatical violations or garden paths, although this does not imply a consensus on the functional interpretation of the P600 (for reviews, see Bornkessel-Schlesewsky & Schlesewsky, 2008; Kuperberg, 2007; Stroud, 2008).

The present study

The primary empirical goal of the current study was to present a carefully controlled side-by-side comparison of three types of structural manipulations that have figured prominently in previous ERP studies on sentence processing and late positivities, specifically (i) ungrammaticalities, (ii) syntactic garden-paths, and (iii) the completion of *wh*-dependencies. To our knowledge, no previous study has made this 3-way comparison. The primary theoretical goal of the study was to use the results of this comparison to constrain models of the syntactic processes underlying the P600 and the substantial variability seen in previous reports of this component. Our working hypothesis was that variation in the timing of the P600 response does not reflect qualitatively different neurocognitive processes that operate at different times, but rather that it reflects variation in the enabling conditions that govern the onset and conclusion of the P600 response. We propose the P600 itself reflects the creation (and in some cases also the destruction) of syntactic relations, and that the latency of the P600 onset reflects the time needed for recognition and retrieval of the elements that participate in those relations, whereas the duration and amplitude of the P600 reflects the structure building operations themselves. This predicts that different types of structural and lexical manipulation should modulate the P600 in different ways, depending on whether they impact recognition and retrieval processes or the complexity of the structure building operations that are undertaken.

The design of the current study was motivated by the possibility that variation in the P600 seen in previous studies might have been due in part to lexical and contextual differences in the materials used. Consequently, the materials were designed to minimise variation in critical sentence regions and in the preceding and following words. A sample set of stimulus items for the five experimental conditions is shown in Table 1. In all conditions, the critical word was the same verb (e.g., ‘showed’ in Table 1) and the six preceding words and the five following words were identical. The use of fronted prepositional *wh*-phrases (e.g., ‘to whom’) in the *wh*-dependency conditions made it possible to match the words following the critical verb, because all

TABLE 1
Sample set of stimulus materials. Critical verbs are underlined.

(a) Control	The patient met the doctor *while* the nurse with the white dress <u>showed</u> the chart during the meeting.
(b) Ungrammatical	The patient met the doctor *while* the nurse with the white dress <u>show</u> the chart during the meeting.
(c) Wh-dependency	The patient met the doctor *to whom* the nurse with the white dress <u>showed</u> the chart during the meeting.
(d) Ungrammatical + Wh-dependency	The patient met the doctor *to whom* the nurse with the white dress <u>show</u> the chart during the meeting.
(e) Garden path	The patient met the doctor *and* the nurse with the white dress <u>showed</u> the chart during the meeting.

verbs were followed by a direct object NP. Previous psycholinguistic studies have shown that even if the gap position associated with a prepositional *wh*-phrase is not directly adjacent to the verb, the dependency between the *wh*-phrase and the verb is constructed immediately upon processing the verb (Pickering & Barry, 1991; for review see Phillips & Wagers, 2007). The use of a prepositional *wh*-phrase also minimised the ambiguity associated with the *wh*-phrase, since the case (dative) and the thematic role (goal/recipient) of the fronted phrase is already clear before the verb is processed. This is important, since the resolution of syntactic ambiguity is independently known to elicit a late positivity (e.g., Osterhout et al., 1994). Therefore, it should be possible to more clearly determine the contribution of syntactic dependency formation to ERP responses, as opposed to lexical differences and ambiguity of case or thematic roles, which was a possible confounding factor in previous studies on English (Kaan et al., 2000; Phillips et al., 2005).

Our guiding hypotheses about the P600 led to a number of specific predictions about the effect of our experimental manipulation on the P600. First, we predicted that the three types of constructions (ungrammatical, syntactic garden-path, and completion of wh-dependencies) would elicit a P600 with a similar topographic distribution, based on the hypothesis that common mechanisms underlie the P600 in each case. Second, based on the hypothesis that the P600 reflects the building of syntactic relations we predicted that the posterior positivity elicited by wh-dependency completion should be attenuated relative to previous reports. If confirmed, this would suggest that the previous findings of P600s associated with *wh*-dependencies reflected combined effects of long-distance dependency formation and ambiguity resolution. Third, we predicted that errors that could be detected using more reliable retrieval cues should elicit earlier P600 onsets. This suggests that the P600 might begin earlier in the garden path condition than in the ungrammatical condition, because correct agreement (garden path condition) may provide a more reliable cue for retrieval of the subject noun phrase than does incorrect agreement (ungrammatical condition). Finally, we predicted that the P600 should be smaller or more short-lived when the parser encounters 'resolvable' problems such as mild-to-medium garden paths and *wh*-dependencies than when it encounters unresolvable problems such as ungrammaticalities.

We should note that the general approach that we pursue here could be accommodated within a number of existing neurochronometric models of sentence processing (e.g., Friederici, 2002; Hagoort, 2003), provided that those models were elaborated to incorporate the detailed parsing operations invoked here. Also, our hypotheses are broadly related to various existing models of sentence parsing and comprehension difficulty (e.g., Gibson, 1998; Lewis, Vasishth, & van Dyke, 2006; Vosse & Kempen, 2000). The goal of this research is less to decide among competing extant models, and more to test the feasibility of developing specific models of variation in the P600, such that it can guide theory building in the same way that N400 variation has guided models of how words are semantically integrated into a sentence (Federmeier, 2007; Kutas & Federmeier, 2000; Lau et al., 2008; van Petten & Luka, 2006).

MATERIALS AND METHODS

Participants

Twenty American English native speakers (8 women; age range 19–28 years; mean age 21.7) participated in the experiment. The data from 2 participants were excluded from the analysis due to recording artifacts; the results reported thus include 18 participants (7 women). All participants were undergraduate or graduate students at the University of Maryland, had normal or corrected to normal vision, were strongly right-handed as measured by the Edinburgh Handedness Inventory (Oldfield, 1971), and had no known history of neurological impairment. The subjects received payment for their participation and provided written informed consent before the experiment.

Stimuli

Table 1 contains a sample set of stimulus items. There were five experimental conditions, and the critical word in all conditions was a ditransitive verb in the second clause (e.g., ‘showed’). All items were identical except for one region at the beginning of the second clause. In the control condition and the ungrammatical condition the second clause was a temporal modifier clause headed by *while*. In the *wh*-dependency conditions the second clause was a relative clause headed by the prepositional *wh*-phrase *to whom*, and in the garden path condition the second clause was a conjoined clause headed by *and*. In all conditions the critical verb was preceded by a six-word NP containing a prepositional phrase modifier (e.g., ‘the nurse with the white dress’). This NP was made long in order to ensure that the sequence of words preceding the critical verb was identical across conditions. In the control condition the critical verb appeared with a past tense suffix, and in the two ungrammatical conditions the critical verb appeared as an uninflected root form (e.g., ‘show’) that failed to agree with the 3rd person singular subject NP. In the *wh*-dependency conditions the processing of this verb allowed completion of the filler-gap dependency between the fronted prepositional *wh*-phrase and the verb. In the garden-path conditions the processing of this verb indicated the need for reanalysis. The NP preceding the critical verb allowed an initial analysis as part of a conjoined NP (e.g., ‘The patient met the doctor *and* the nurse with the white dress'). Previous studies suggest that this is the most commonly preferred analysis (Frazier, 1985; Frazier & Clifton, 1996). The verb therefore signalled that this initial analysis must be revised in favour of a clausal conjunction analysis. Due to the choice of verb, it was also possible to match the five words following the verb across conditions. In all cases the ditransitive verb allowed optional omission of the indirect object PP. In all conditions the verb was immediately followed by a direct object NP. This is important, since P600 effects elicited by the verb are typically observed during presentation of the following word, and thus represents an improvement over earlier ERP studies of *wh*-dependency formation in English (Kaan et al., 2000; Phillips et al., 2005). The indirect object PP never appeared after the verb, either because it was fronted (*wh*-dependency conditions) or because it was omitted (control, ungrammatical, garden path conditions). In all conditions the sentences ended with a modifier PP that was included in order to reduce the possibility of contamination of the results by wrap-up effects.

A total of 180 sets of five sentences were created using 90 different critical verbs. The sentences were distributed among five lists in a Latin Square design, such that each list contained 36 instances of each condition but only one version of each item. The 180 target sentences were interspersed with 360 fillers of comparable length and complexity, yielding a 2:1 filler-to-target ratio. Due to the large number of trials, the experiment was divided into two sessions of 270 trials each, separated by an interval of at least 2 days. The stimulus lists were organised such that each of the 90 critical verbs appeared only once during each recording session. The target stimuli and fillers were pseudo-randomised. A complete set of materials used in this experiment is available from the second author's web site.

Procedure

Participants were comfortably seated in an armchair facing a computer screen at an approximate distance of 1 m. Sentences were presented visually in the centre of the screen in an RSVP paradigm, at an SOA of 500 ms (300 ms per word, 200 ms blank screen) using black letters on a white background. Punctuation and the use of upper and lower case letters were normal. Each trial began with a fixation point at the centre of the screen. The participant pressed a button on a button box to start the sentence. In order to ensure that participants attended to the content of the sentences all sentences were followed by a yes/no comprehension question. The question remained on the screen until the participant answered it by pressing a button. Feedback on accuracy was given on all trials. Participants were asked to restrict blinks and other movements to the interval when the fixation point was displayed on the screen between each trial and were also asked to answer the questions as quickly and as accurately as possible. Each experimental session was preceded by a short practice session to familiarise the participant with the task. Each session consisted of five blocks of 54 sentences, each lasting approximately 15 minutes and followed by a short break.

EEG recordings

Continuous EEG was recorded from 30 Ag/AgCl electrodes, mounted in an electrode cap (Electrocap International) and arranged in the following modified 10–20 configuration: midline: Fz, FCz, Cz, CPz, Pz, Oz; lateral: FP1/2, F3/4, F7/8, FC3/4, FT7/8, C3/4, T7/8, CP3/4, TP7/8, P4/5, P7/8, O1/2. Horizontal eye movements were monitored by additional electrodes placed on the left and right outer canthus and vertical eye movements by electrodes placed above and below the left eye. The recordings were referenced to linked mastoid electrodes[1] and the AFZ electrode served as ground. The EEG and EOG recordings were amplified by a SynAmps™ Model 5083 EEG amplifier, using a DC to 70 Hz low-pass filter, and digitised at a frequency of 500 Hz. Impedances were kept below 5 kΩ per channel.

Data analysis

Trials with eye movements or other artifacts were rejected, affecting around 11% of the trials (range: 88.7–90.0% across conditions). Since the data showed a slow drift common to DC recordings, a detrending algorithm similar to the one reported in Fiebach et al. (2002) was used to correct for a common linear component. Raw data files were segmented into 11 s time intervals, subject to the constraint that target sentences always fell within a single interval. Within each interval, a linear regression was computed for each electrode, and subtracted from the original recording. After preprocessing the data, event related potentials were computed separately for each participant in each experimental condition for a 1300 ms interval time-locked to the onset of the critical verb relative to a 100 ms prestimulus baseline.

[1] There is some controversy in the ERP literature over the selection of appropriate reference electrodes, and some have argued that a linked mastoid reference risks distortion of the scalp topography of observed effects (Luck, 2005; but see Davidson, Jackson, & Larson, 2000 for a different viewpoint). However, a survey of ERP sentence processing studies that have used linked and unlinked reference electrodes suggests that this choice has little impact upon the major language-related ERP components (linked reference: Donaldson & Rugg, 1999; Friederici, Hahn, & Mecklinger, 1996; Hoeks et al., 2004; Mills, Prat, Zangl, Stager, Neville, & Werker, 2004; Neville et al., 1991; Phillips et al., 2005; unlinked reference: Friederici et al., 2001; Hagoort et al., 2003b; Kaan et al., 2000; Kim & Osterhout, 2005). Moreover, the choice of reference should not impact the timing differences that are the focus of the current study, nor should it create spurious topographic differences among conditions, since we used a fully within-subjects design.

Statistical analyses were performed on unfiltered data on the mean amplitude relative to baseline within six time windows: 0–300 ms, 300–500 ms, 500–700 ms, 700–900 ms, 900–1100 ms, 1100–1300 ms. These intervals were chosen based on the previous literature and on visual inspection of the grand averages. Statistical analyses were performed on 18 electrodes, which were distributed among six regions of interest according to two topographic factors, laterality and anterior/posterior. The regions of interest were left anterior (FT7, F3, FC3), midline anterior (FZ, FCZ, CZ), right anterior (F4, FC4, FT8), left posterior (TP7, P3, CP3), midline posterior (PZ, CPZ, OZ), and right posterior (P4, CP4, TP8). For each time interval two types of repeated-measures ANOVA were performed. One ANOVA included 4 of the 5 experimental conditions in a 2×2 factorial design based upon the factors *grammaticality* (2: grammatical vs. ungrammatical) and *wh-dependency* (2: *wh*-dependency vs. no-dependency), in addition to the topographic factors. This ANOVA excluded the garden path condition. A second ANOVA consisted of a series of planned pairwise comparisons between the control condition and each of the four other conditions, including the garden path condition. Follow-up ANOVA analyses within specific topographic regions were conducted in order to further examine the source of main effects or interactions in the overall ANOVAs and to test possible differences suggested by visual inspection. The same electrodes and time intervals were used in the follow-up analyses as in the main ANOVAs. Results are reported for all main effects and interactions involving at least one condition factor. However, due to the large number of possible interactions in the experimental design, we report as significant only those interactions for which subsequent analyses yielded significance within the levels of the interacting factors. For all effects involving more than one degree of freedom, the Greenhouse–Geisser correction was applied (Greenhouse & Geisser, 1959).

RESULTS

Comprehension accuracy

Overall accuracy in answering the comprehension questions was 85% for all experimental conditions combined. Mean accuracy was 85% (*SD* 8.3%) in the control condition, 87% (*SD* 10.7%) in the ungrammatical condition, 83% (*SD* 8.5%) in the *wh*-dependency condition, 84% (*SD* 8.5%) in the ungrammatical *wh*-dependency condition, and 85% (*SD* 10.5%) in the garden path condition. These mean scores showed no reliable differences between conditions ($F < 1$).

Event related potentials

We report results from a series of analyses that test the similarities and differences between the ERPs elicited by ungrammaticality, *wh*-dependency completion, and syntactic garden paths, together with analyses that test for possible topographic differences or effects of experiment-specific strategies. For the conditions involving the grammaticality and *wh*-dependency factors we focus on the results of the 2×2 ANOVA shown in Table 2, drawing on the separate pairwise comparisons of conditions only where this impacts the reliability or interpretation of the findings. For the garden path condition we focus on the pairwise comparison with the control condition. Figures 1–5 show comparisons of pairs of grand average ERPs (control vs. ungrammatical, *wh*-dependency vs. ungrammatical *wh*-dependency, control vs. *wh*-dependency, control vs. ungrammatical *wh*-dependency, and control vs. garden path, respectively), averaged across the electrodes in each of six topographic regions. Figure 6 shows topographic scalp maps for each successive interval for the voltage difference between the control condition and the four other conditions. Grand average waveforms were treated using a 10 Hz low-pass filter for visualisation purposes, but all analyses were conducted on unfiltered data.

As predicted, all conditions elicited a late positivity relative to the control condition. However, the main interest of the study lies in the differences that were observed in the amplitude and timing of the late positivity, and in the other ERP components elicited by the different conditions.

Grammaticality effects

The effects of grammaticality in the 2×2 ANOVA reflected the combined effects of the ungrammatical condition and the ungrammatical *wh*-dependency conditions. The earliest effect of grammaticality was a negativity in the 300–500 ms interval that was more pronounced at anterior and midline sites, although it also extended to left posterior sites. At all subsequent analysis intervals from 500–1300 ms there was a posterior positivity characteristic of the P600 elicited by ungrammatical materials in other studies. In the overall ANOVA these effects appeared as main effects of grammaticality, as grammaticality $\times$ anteriority interactions, or both. Separate analyses of anterior and posterior regions showed that in the 300–500 ms interval the effect of grammaticality was reliable at anterior regions but was only marginally significant at posterior regions. In contrast, from 500–1300 ms the effect of grammaticality was reliable at posterior regions only. From 700–1100 ms there was also a significant interaction of grammaticality with laterality, which reflected the fact that the positivity was more pronounced at right hemisphere than at left hemisphere and midline channels.

TABLE 2
Summary of ANOVA *f*-values for analysis involving the condition factors *grammaticality* and *wh-dependency* at successive latency intervals relative to the embedded verb.

Overall ANOVA (dfs)	*0–300 ms*	*300–500 ms*	*500–700 ms*	*700–900 ms*	*900–1100 ms*	*1100–1300 ms*
gram (1, 17)	–	5.03*	3.30†	–	5.53*	5.11*
wh (1, 17)	–	4.03†	3.35†	–	–	–
gram × *wh* (1, 17)	–	–	–	–	–	–
gram × *ant* (1, 17)	–	–	7.34*	8.41**	6.79*	–
gram × *lat* (2, 34)	–	4.11†	–	10.25**	6.05*	–
gram × *ant* × *lat* (2, 34)	–	5.80*	2.84†	–	–	–
wh × *ant* (1, 17)	–	–	–	–	–	–
wh × *lat* (2, 34)	–	2.91†	–	–	–	–
wh × *ant* × *lat* (2, 34)	–	–	–	–	–	–
gram × *wh* × *ant* (1, 17)	–	–	–	–	–	–
gram × *wh* × *lat* (2, 34)	–	–	–	–	–	–
gram × *wh* × *ant* × *lat* (2, 34)	–	–	–	–	–	–
Anterior regions only						
gram (1, 17)	–	5.53*	–	–	–	–
wh (1, 17)	–	5.96*	–	–	–	–
gram × *wh* (1, 17)	–	–	–	–	–	–
Posterior regions only						
gram (1, 17)	–	3.89†	10.42**	11.83**	9.70**	6.95*
wh (1, 17)	–	–	3.14†	–	–	–
gram × *wh* (1, 17)	–	–	–	–	–	–
Left anterior						
gram (1, 17)	–	5.79*	–	–	–	–
wh (1, 17)	–	7.69*	1.84†	–	–	–
Midline anterior						
gram (1, 17)	–	6.32*	–	–	–	–
wh (1, 17)	–	5.29*	–	–	–	–
Right anterior						
gram (1,17)	–	3.61†	–	–	5.89*	4.45*
wh (1, 17)	–	3.85†	–	–	–	–
Left posterior						
gram (1, 17)	–	5.63*	–	–	–	–
wh (1, 17)	–	3.71†	–	–	–	–
Midline posterior						
gram (1, 17)	–	–	18.55***	13.77**	9.34**	9.01**
wh (1, 17)	–	–	3.60†	–	–	–
Right posterior						
gram (1, 17)	–	3.26†	9.09**	17.81**	14.03**	7.11**
wh (1, 17)	–	–	–	–	–	–

Factors: *gram* – grammaticality; *wh* – wh-dependency; *ant* – anterior/posterior; *lat* – laterality.
† $.1 > p > .05$; * $.05 > p > .01$; ** $p < .01$.

We conducted separate pairwise comparisons of the control and ungrammatical conditions (Figure 1) and of the grammatical and ungrammatical *wh*-dependency conditions (Figure 2) in order to test the robustness of the effect of grammaticality. The effects of the anterior negativity and the posterior positivity were observed in both comparisons, but the anterior negativity in the 300–500 ms interval was more pronounced in the comparison of the grammatical and ungrammatical *wh*-dependency conditions, whereas the later posterior positivity was more pronounced in the comparison of the control and ungrammatical conditions. The anterior negativity appeared only as a condition × laterality × anteriority effect in the overall ANOVA for the control vs. ungrammatical comparison, $F(2, 34) = 4.73$, $p < .05$, with no reliable effects found in subsequent region-by-region comparisons. In the comparison of the grammatical and ungrammatical *wh*-dependency conditions (Figure 2) the anterior negativity in the 300–500 ms interval yielded a main effect of condition, $F(1, 17) = 4.48$, $p < .05$, and region-by-region analyses showed significant effects at left anterior, $F(1, 17) = 6.02$, $p < .05$, anterior midline, $F(1, 17) = 5.61$, $p < .05$, and left posterior sites, $F(1, 17) = 5.01$, $p < .05$. In contrast, the later posterior positivity was reliable and long-lasting in the control vs. ungrammatical comparison (Figure 1), yielding a main effect of condition at all intervals from 500 ms to 1300 ms (all Fs > 5, $p < .05$), and condition × anteriority interactions at all intervals from 500 ms to 1100 ms (all Fs > 5, $p < .05$). Separate analyses of anterior and posterior regions showed that the

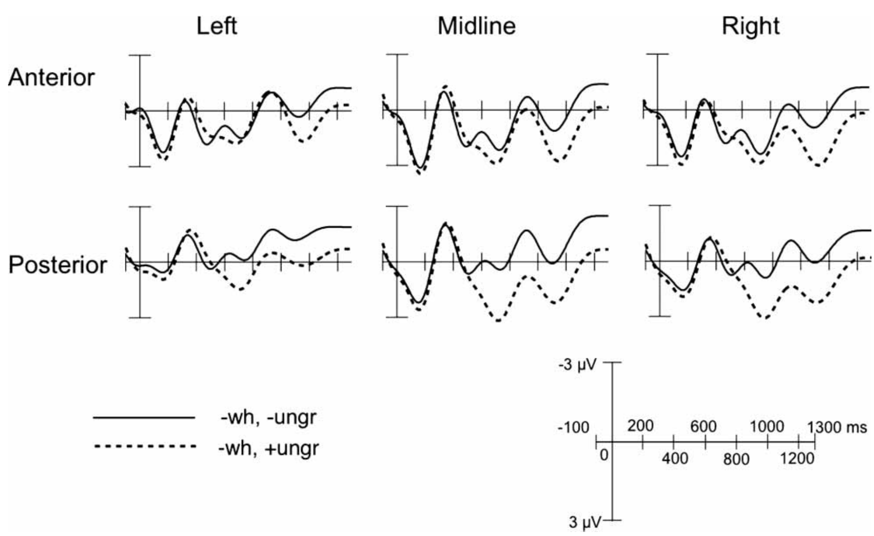

Figure 1. Grand average ERPs at six groups of topographically arranged electrode sites in the control and ungrammatical conditions (examples (a) and (b) in Table 1).

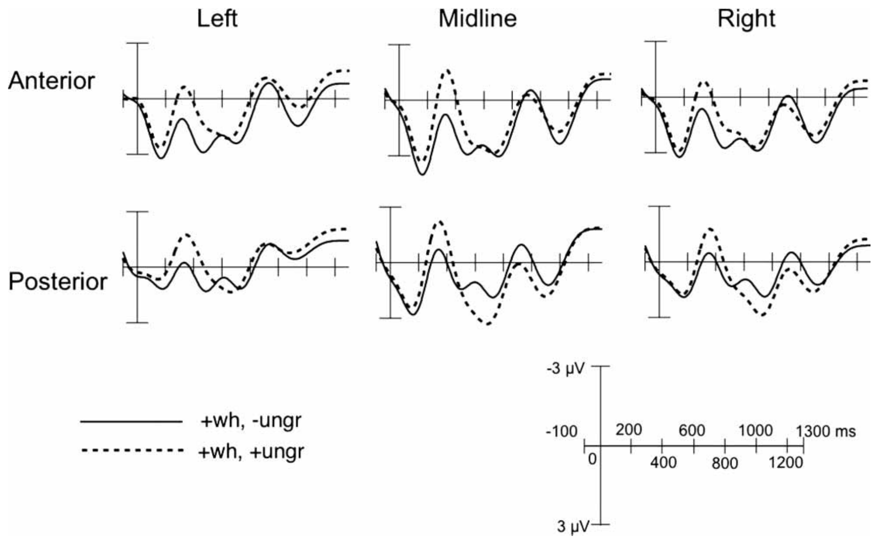

Figure 2. Grand average ERPs at six groups of topographically arranged electrode sites in the *wh*-dependency and ungrammatical *wh*-dependency conditions (examples (c) and (d) in Table 1).

positivity was highly reliable at posterior sites at all intervals from 500 ms to 1300 ms (all *F*s > 12, $p < .01$), and at anterior sites only at the 1100–1300 ms interval, $F(1, 17) = 5.93$, $p < .05$. In the comparison of the grammatical and ungrammatical *wh*-dependency conditions the posterior positivity yielded a marginally significant condition × anteriority interaction at the 500–700 ms interval, $F(1, 17) = 3.94$, $p < .07$, and a significant effect at posterior midline electrodes in the same interval, $F(1, 17) = 4.83$, $p < .05$.

Effects of wh-dependency formation

A primary goal of the *wh*-dependency manipulation in the current study was to determine whether the ERP effects of *wh*-dependency formation found in previous studies would be altered by closer lexical matching of conditions and reduction in the ambiguity of the syntactic role of the *wh*-phrase. In the overall ANOVA the effects of the *wh*-dependency factor were relatively weak and were confined to the 300–700 ms interval. At the 300–500 ms interval there was a main effect of the *wh*-dependency factor and an interaction of the *wh*-dependency factor with laterality. This was due to a positivity that was strongest at left anterior sites but that extended to additional anterior and left hemisphere sites. At the 500–700 ms interval, on the other hand, the main effect of the *wh*-dependency factor was due to a positivity that was strongest over posterior sites, although this effect was only marginally significant in the overall ANOVA. Subsequent planned comparisons involving pairs of conditions showed that the posterior positivity yielded reliable differences in the comparison of the control vs. *wh*-dependency conditions, but not in the comparison of the ungrammatical vs. *wh*-ungrammatical conditions.

The pairwise comparison of the control condition with the grammatical *wh*-dependency condition (Figure 3) revealed that the positivity was already weakly present in the 0–300 ms interval, where there was a marginally significant main effect of condition, $F(1, 17) = 3.82$, $p < .07$ and a marginally significant condition × anteriority interaction $F(1, 17) = 3.42$, $p < .09$. Inspection of averaged ERPs to the preceding word confirmed that this effect was not an artifact of pre-existing differences between the two conditions. Analyses at individual regions showed that the effect of condition was significant at anterior channels, $F(1, 17) = 4.74$, $p < .05$, but not at posterior channels, and was also significant at left posterior channels, $F(2, 34) = 5.69$, $p < .05$. At the 300–500 ms interval there was a main effect of condition, $F(1, 17) = 6.61$, $p < .05$, and a marginally significant condition × laterality interaction, $F(1, 17) = 3.53$, $p < .08$, reflecting the fact that the positivity was stronger at left hemisphere channels than at right hemisphere channels, although the effect of condition was significant at anterior channels and posterior channels alike: anterior, $F(1, 17) = 6.79$, $p < .05$; posterior, $F(1, 17) = 4.79$, $p < .05$. By the 500–700 ms interval the positivity remained, but it was now stronger at posterior than at anterior channels. There was a main effect of condition, $F(1, 17) = 5.86$, $p < .05$, but separate analyses at anterior

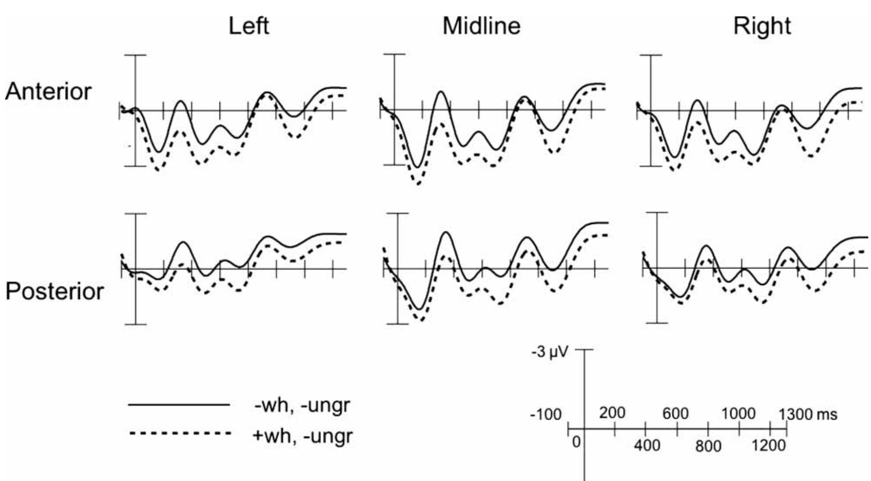

Figure 3. Grand average ERPs at six groups of topographically arranged electrode sites in the control and grammatical *wh*-dependency conditions (examples (a) and (c) in Table 1).

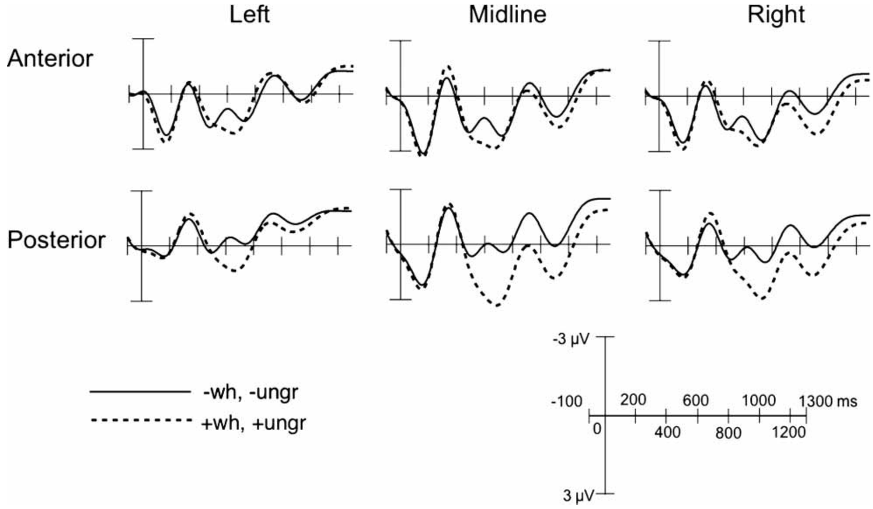

Figure 4. Grand average ERPs at six groups of topographically arranged electrode sites in the control and ungrammatical *wh*-dependency conditions (examples (a) and (d) in Table 1).

and posterior regions showed that the effect of condition was significant at posterior channels, $F(1, 17) = 5.92$, $p < .05$, and only marginally significant at anterior channels, $F(1, 17) = 4.4$, $p < .06$. At the 700–900 ms interval there was a marginally significant interaction of condition with the anterior/posterior factor, $F(1, 17) = 3.5$, $p < .08$, but since analyses within each level of the anterior/posterior factor yielded no significant or marginally significant differences, this effect is not considered further. There were no main effects or interactions involving the *wh*-dependency factor at subsequent intervals. In the comparison of the ungrammatical and ungrammatical *wh*-dependency conditions there were no reliable differences at any interval or region.

Thus, although in this study we find evidence for a posterior positivity associated with *wh*-dependency formation in the global ANOVA, subsequent more detailed analyses suggest that this effect is less robust than was the case in other ERP studies. We return to this issue in more detail in the Discussion section.

Garden path condition

Since the garden path condition was not included in the main 2×2 ANOVA, its effects were tested using a separate series of ANOVAs that compared the garden path condition with the control condition (Figure 5). This analysis revealed a posterior positivity that lasted from the 300–500 ms interval until the 1100–1300 ms interval, as shown in Table 3. A further analysis that focused on ERPs at electrode PZ, where the P600 was maximal, showed that

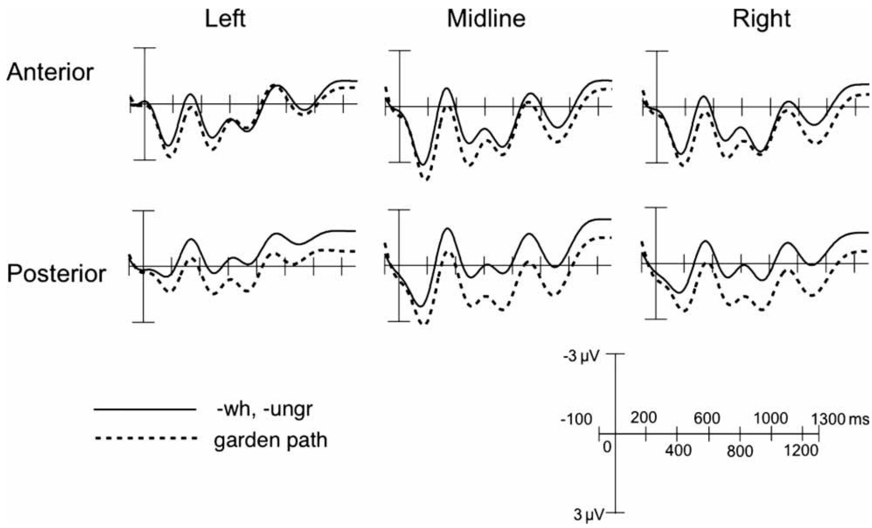

Figure 5. Grand average ERPs at six groups of topographically arranged electrode sites in the control and garden path conditions (examples (a) and (e) in Table 1).

the positivity was already significant in a 200–300 ms interval, $F(1, 17) = 6.71$, $p < .05$, but not at earlier intervals. Although the ANOVA involving all regions of interest showed variation across time in the reliability of the main effect of condition and the condition × anterior/posterior interaction, each interval showed a significant effect of condition at posterior sites and no corresponding effect at anterior sites.

TABLE 3
Summary of ANOVA *f*-values for comparison of the garden path and control conditions at successive latency intervals relative to the embedded verb.

	0–300 ms	*300–500 ms*	*500–700 ms*	*700–900 ms*	*900–1100 ms*	*1100–1300 ms*
cond (1, 17)	–	5.49*	4.39†	–	3.61†	3.16†
cond × ant (1, 17)	–	3.62*	11.68**	28.07**	10.86**	–
cond × lat (2, 34)	–	–	–	–	5.62*	–
cond × ant × lat (2, 34)	–	–	–	–	–	–
Anterior regions only						
cond (1,17)	–	–	–	–	–	–
Posterior regions only						
cond (1, 17)	–	10.22**	13.84**	11.07**	8.47**	6.17*

Factors: *cond* – condition; *ant* – anterior/posterior; *lat* – laterality. † $.1 > p > .05$; * $.05 > p > .01$; ** $p < .01$.

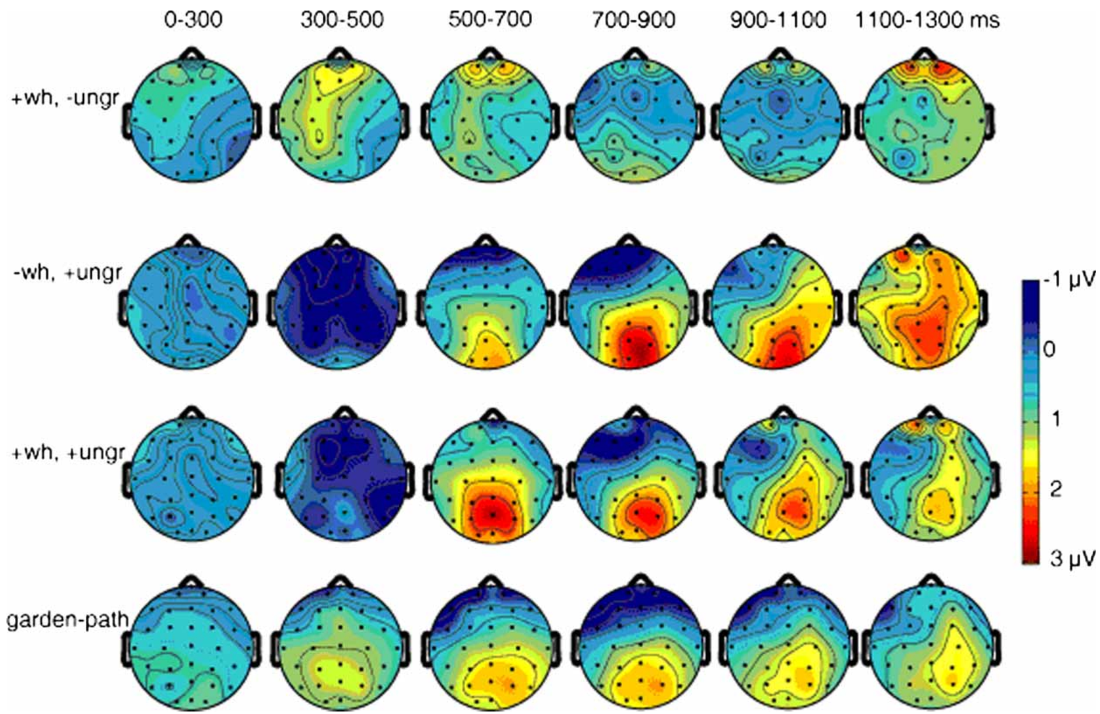

Figure 6. Topographic scalp maps for the comparison of the control condition with the four other conditions at six successive time intervals, showing the amplitude, temporal extent, and scalp distribution of the P600 and other effects. To view this figure in colour, please visit the online version of this issue.

Comparison of conditions

Having found that the ERP responses elicited by ungrammaticality, garden paths, and *wh*-dependency formation differ in terms of timing, amplitude, and in some cases scalp topography, we next turn to further analyses that help to identify the origin of these differences.

The 2 × 2 ANOVA showed no interactions involving both the grammaticality factor and the *wh*-dependency factor at any time interval. Nor were such interactions found in follow-up analyses in any region of interest. This finding suggests that the effects of ungrammaticality and *wh*-dependency formation were additive, and therefore that the ERP responses to the two manipulations were independent. However, it would be premature to conclude from this that the P600 responses to ungrammaticality and *wh*-dependency formation are generated by distinct sources, as statistical independence may arise in a single generator.

As an alternative probe for possible variation in the source of the P600 we conducted an additional ANOVA on the difference waves to test for topographic differences across conditions, using the presence of condition × electrode interactions as a measure of topographic differences. In order to avoid spurious interactions we scaled the difference waves using a method based upon relations of signal amplitudes (Jing, Pivik, & Dykman, 2006) instead of the widely used vector-based method (McCarthy & Wood, 1985) that has come under criticism in recent studies (Haig, Gordon, & Hook, 1997; Urbach & Kutas, 2002). This analysis showed no condition × electrode interactions, except at anterior channels in the 300–500 ms interval, where topographic differences were observed in the ANOVAs reported above. Therefore, despite differences in amplitude and timing across conditions, these analyses found no evidence that the posterior P600 showed reliable topographic differences across conditions. This is consistent with the notion of a common generator, although it certainly does not entail this conclusion.

Finally, we considered the possibility that ERP differences between the conditions might be the consequence of task-specific strategies developed by the experimental participants across the course of the study. Although the careful matching of experimental materials and the large numbers of filler items reduced the likelihood of strategic processing, this possibility must be taken seriously. We therefore conducted a new set of ANOVAs on the four difference waves, including a new *Block* factor that distinguished the trials in the first half of each experimental session from the trials in the second half of each session. We hypothesised that if the differences between conditions reflect condition-specific strategies developed over the course of the experiment, then we should encounter condition × block interactions. This is to be distinguished from effects of practice or fatigue that affect all conditions similarly, which should present as main effects of the block factor. In order to increase the chance of finding condition × block interactions we scaled the difference waves in each condition such that the mean amplitude (combining both blocks) was matched across conditions, while the ratio of the two blocks within any experimental condition was left unchanged. This analysis revealed that although P600 amplitudes reduced overall between the first and second blocks, there were no significant condition × block interactions, suggesting that the ERP differences between conditions were not the result of experiment-specific strategies.

DISCUSSION

Summary

Event related potential studies of sentence processing have revealed that a number of different types of syntactic configurations elicit the P600 response component, also known as the 'Syntactic Positive Shift'. The specific conditions eliciting this response include ungrammaticality, garden paths, and the completion of long-distance dependencies. This naturally raises the question of what properties these phenomena have in common. However, it has been difficult to establish the extent to which these phenomena elicit parallel ERP responses, since previous findings have often been based on different materials and/or different participants. In order to address this issue we used a within-subjects design with maximally similar materials for each of three structural configuration-types that have been argued to elicit a P600 response. Under these controlled circumstances a late positivity was elicited in all three structural environments, but with reliable variation across conditions in the latency and the duration of the positivity, and limited variation in the scalp distribution of the late positivity. After briefly summarising the main findings of the study and some goals for a theory of the processes that underlie the P600, we then discuss each of the main findings of the study in turn and their implications for neurocognitive models of sentence comprehension.

First, the scalp topography of the P600 was very similar across conditions. This is consistent with the possibility of a common mechanism underlying the P600 elicited by syntactic garden paths, grammatical violations, and completion of long-distance dependencies. The notable exception to this generalisation is that the *wh*-dependency condition elicited a positivity that initially had a more anterior scalp distribution (300–500 ms) before shifting to the more standard posterior distribution. The second difference between conditions was that the P600 elicited by the *wh*-dependency was smaller than the P600 elicited by garden paths and ungrammaticality. Further, the P600 elicited by the dependency involving a fronted PP was less robust than the effect elicited by dependencies involving fronted NPs in previous studies (Kaan et al., 2000; Phillips et al., 2005). The third difference between conditions was that the P600 in the garden path and ungrammatical conditions contrasted in two respects, despite their similar scalp topography and long duration: the ungrammatical condition elicited an additional negativity with a weak anterior focus, and the posterior P600 had a later onset latency in the ungrammatical conditions (500–700 ms) than in the garden path condition (300–500 ms).

Before discussing the implications of each of these findings in turn, we consider alternative ways in which variability in the P600 may be understood. Relatively few studies have explored systematic variation in the P600, but it is increasingly clear that the P600 varies along a number of different parameters, including latency, duration, amplitude, and scalp distribution. This contrasts with the N400, for which there is substantial documentation of systematic amplitude variation, but the existence of latency and topographic variation remains less well established. The multi-dimensional variation in the P600 response provides valuable clues for identifying the specific cognitive processes and computations that underlie the component.

Since the P600 was first identified it has been widely assumed that the P600 observed in anomalous or difficult sentences reflects syntactic processes. However, it has been less clear whether the P600 reflects syntactic processes in general, or a more specific subset of syntactic processes. It has also been unclear whether the P600 is properly understood as reflecting a cohesive set of processes, or disparate syntactic processes that happen to elicit similar ERP effects at the scalp. We suggest that the P600 may be understood as reflecting a common set of processes that occur on a *just-in-time* basis, beginning as soon as sufficient information has been accrued to initiate the processes. Under this view, when the P600 occurs at different latencies it reflects the same underlying processes, with latency variation reflecting the time needed to complete the processes that trigger the P600. This view of P600 timing is compatible with discussions of the P600 by Friederici et al. (2001) and Hagoort (2003), among others.

If it is true that the P600 reflects a common set of processes that may occur at different latencies, then this calls for an account that distinguishes between those processes that are directly reflected in the P600, and those processes that modulate the latency of the P600 without directly contributing to the P600 component itself. Such an account may be possible, drawing upon the distinction between the retrieval of elements that participate in syntactic relations and the creation of the syntactic relations themselves. If the P600 reflects the creation (and possibly also the destruction) of syntactic relations, then it follows that the latency of the P600 should reflect the time needed for retrieval of the elements that participate in those relations, whereas the duration and amplitude of the P600 should be a function of the structure building processes themselves. We therefore predict that different structural and lexical manipulations should impact the P600 differently. Manipulations that impact retrieval processes should change the latency of the P600, whereas manipulations that impact the number and type of syntactic relations that are attempted should change the amplitude and/or duration of the P600.

This account of the processes that underlie the P600 may be applied to the three types of structural phenomena that are the focus of the current study. First, completion of a long-distance filler-gap dependency requires retrieval of the filler from memory and a verb from the linguistic input, followed by the updating of the syntactic representation to encode the relation between the filler and the verb. Second, recovery from a syntactic garden path requires retrieval of candidate words and phrases in memory that may enter into syntactic relations with the incoming word; it typically requires dismantling of previously built syntactic relations, and in some cases it may involve retrieval of irrelevant items and creation of inappropriate structures. Third, processing of ungrammatical input requires a search for words and phrases in memory that are candidates for forming a structural relation with the incoming word, even if that relation is not entirely well-formed.

Our suggestion about how different parsing processes affect the P600 may be related to a number of different models of sentence parsing and sentence comprehension difficulty (e.g., Gibson, 1998; Lewis et al., 2006; Vosse & Kempen, 2000) and with neurochronometric models of sentence processing (e.g., Friederici, 2002; Hagoort, 2003). We further explore the relation between our proposal and these other models after discussing each of the main findings of our study in turn.

Timing and onset of P600

One contrast in the responses to the ungrammatical and garden path conditions was the earlier onset of the P600 in the garden path condition (300–500 ms interval) than in the ungrammatical condition (500–700 ms interval). This difference must be treated with some caution, since it is possible that an earlier onset for the P600 in the ungrammatical condition was masked by the relatively broad AN in the 300–500 ms interval. However, the late onset of the response to agreement violations in this study is consistent with results from other studies of agreement violations (e.g., Hagoort et al., 1993; Hagoort & Brown, 1994; Lau et al., 2006). Given the theoretical importance of onset latency variation in the P600, plus good evidence from other studies for such variation, we suggest an account of the P600 latency variation found in this and previous studies.

Friederici et al. (2001) and Phillips et al. (2005) both demonstrate variation in P600 onset latencies in closely matched materials. Friederici and colleagues tested two types of garden paths involving unexpected object-subject word order in German, showing an earlier P600 onset in relative clauses, which are easier to reanalyse, and a later P600 onset in complement clauses, which are harder to reanalyse. They highlight the importance of the finding that more difficult reanalysis delays the onset of the P600, rather than merely increasing its amplitude or duration. Friederici and colleagues suggest that the P600 onset latency reflects the completion of a 'diagnosis' stage that precedes reanalysis of anomalous sentences. Phillips et al. (2005) compare P600s at the completion of shorter and longer *wh*-dependencies, showing earlier P600 onset latencies in the short dependency conditions. They propose that the latency difference reflects the longer time needed to retrieve a more distant *wh*-phrase from memory. Whereas these two previous accounts of P600 latency variation each apply to one type of syntactic phenomenon, we suggest that the variation can be understood in more general terms that encompass the processing of syntactic violations, garden paths, and long-distance dependencies. As outlined above, we suggest that P600 amplitude and duration directly reflect structure-building (and dismantling) operations, whereas the retrieval processes that are needed to initiate structure building are reflected only in the onset latency of the P600.

The proposal that P600 onset latency reflects retrieval times straightforwardly captures the dependency length effect in Phillips et al. (2005), since it plausibly takes longer to reactivate a more distant *wh*-filler above a threshold level. (This claim is not incompatible with the argument from speed-accuracy tradeoff paradigms that access to filler phrases is unaffected by distance (McElree, Foraker, & Dyer, 2003). Filler phrases may take longer to be reactivated to a threshold level due to distance-based decay rather than slower access times.) This account may also extend to the garden path materials tested by Friederici and colleagues. Participants in that study had to use an agreement cue from a sentence-final auxiliary to recognise that the immediate pre-verbal NP was a subject NP rather than the object NP that it was initially assumed to be. In order to do this the parser had to rescind its initial assignment of the subject role. This task was easier in the relative clause condition (5a) than in the complement clause condition (5b), and we suggest that a delay in retrieving the correct pre-verbal subject NP in the complement clause condition may have been responsible for the delay in the onset of the P600. Retrieval of the correct subject NP in the complement clause condition may have been slower because of difficulty in inhibiting the initial assignment of the clause-initial NP to the subject role. In the relative clause condition, on the other hand, the pre-verbal subject NP may have been retrieved more rapidly because of reduced interference from the gap that was initially assigned to the subject role.

(5) a. Das ist die Direktorin, die die Sekretärinnen gesucht haben.
that is the director that the secretaries sought have.pl

b. Er wußte, daß die Sekretärin die Direktorinnen gesucht haben.
he knew that the secretary the directors sought have.pl

Turning to the current study, if it is true that the P600 onset was delayed in the ungrammatical condition relative to the garden path condition, this can also be accounted for in terms of retrieval processes, by assuming that correct subject verb agreement provides a more effective retrieval cue than incorrect agreement. In the garden path condition presentation of the verb may have triggered a search for an appropriately agreeing subject NP. Once an appropriate subject NP was identified, structure building could commence and hence a P600 effect was observed. In the ungrammatical condition, on the other hand, the incorrect agreement on the verb may have delayed retrieval of the relevant subject NP, as the parser may have initially searched for an NP that matched the agreement features of the verb.

Wh-dependency

Although completion of the *wh*-dependency elicited a posterior P600 effect that was significant in the 2×2 ANOVA, this was a relatively weak effect and it was only marginally significant in the comparison of the control condition with the grammatical *wh*-dependency condition. The positivity was smaller and more short-lived than the P600 elicited in the garden path and ungrammatical conditions in the current study and the *wh*-dependency conditions in previous studies (e.g., Kaan et al., 2000; Phillips et al., 2005).

We suggest that an important difference between the current study and previous ERP studies of *wh*-dependency completion in English involves the information carried on the *wh*-phrase itself. In the *wh*-dependency condition in the current study participants read dative-marked *wh*-phrases (e.g., *to whom*) that included a fronted preposition. This made it possible to closely match the words following the critical verb, but also meant that participants were able to identify the case and thematic properties of the *wh*-phrase already at the beginning of the relative clause. Thus, at the point of processing the verb, the *wh*-phrase had to be integrated with the verb, identifying the specific predicate that it is an argument of, but its case and thematic properties had already been determined. In contrast, in previous studies in English the *wh*-phrase consisted of a noun phrase that remained ambiguous with regard to case and thematic properties until the point of processing the verb. Thus, the processing of the verb led to disambiguation of the properties of the *wh*-phrase in addition to integration of the *wh*-phrase with the verb. In the terms of our proposal about the processes underlying the P600, the use of a dative-marked *wh*-phrase reduced the number of structural relations that needed to be constructed at the verb position, and therefore reduced the amplitude and duration of the P600 relative to earlier studies.

Topographic effects

All three structural manipulations (*wh*-dependency, garden path, agreement violation) elicited a posterior positivity with a similar scalp distribution. The earlier aspect of this positivity, extending to the 700–900 ms interval, was maximal around posterior midline electrode PZ. In conditions where the positivity extended beyond this interval, it showed greater right lateralisation at later intervals. Statistical analyses found no evidence for reliable topographic differences in this component, and therefore our findings are compatible with accounts that propose a common currency underlying the P600. Nevertheless, our topographic analyses are relatively coarse and we cannot rule out the possibility that the P600 is the consequence of a disparate set of processes that happen to elicit topographically similar responses.

In addition to the posterior positivity, the initial response to the completion of the *wh*-dependency was an anterior positivity in the 300–500 ms interval, and in fact this effect was more robust than the later posterior positivity elicited by the *wh*-dependency. The anterior positivity contrasts with the more uniformly posterior positivity elicited by *wh*-dependency completion in previous studies (Kaan et al., 2000; Phillips et al., 2005). This contrast may reflect the different demands of processing the relative clauses in the current study and the indirect *wh*-questions in the earlier studies. In a *wh*-question the filler forms a syntactic dependency with the verb, and may also encode the semantic content of the fronted argument, as in a sentence like *The patient asked which doctor the nurse showed* . . . . In contrast, in the relative clauses used in the current study the verb forms a syntactic dependency with the fronted PP, but the semantic content of that argument is provided by the head of the relative clause, as in *The patient met the doctor to whom the nurse showed* . . . . However, this is just one among a number of properties of relative clauses that may be responsible for the anterior positivity.

Anterior negativity in ungrammatical conditions

Although the ungrammatical and garden path conditions both elicited a long-lasting posterior positivity, a difference between the two conditions was the presence of a negativity in the 300–500 ms interval in the ungrammatical condition that was absent from the garden path condition. The negativity showed a broad scalp distribution, but it was more reliable at anterior electrodes, and thus it is a plausible counterpart of the anterior negativity (AN) elicited by morphosyntactic anomalies in many previous studies (e.g., Coulson et al., 1998b; Friederici et al., 1993; Hagoort et al., 2003a; Osterhout & Mobley, 1995). However, an important caveat is that the negativity in the 300–500 ms interval turned out not to be reliable in the pairwise comparison of the control and ungrammatical conditions, and the reliable difference found in the comparison of the grammatical and ungrammatical wh-dependency conditions could be related to the anterior positivity observed in the grammatical wh-dependency condition. Furthermore, a number of other studies of morphosyntactic anomalies have failed to observe an AN or have observed it only selectively (e.g., Gunter et al., 1997; Hagoort et al., 1993; Lau et al., 2006; Nevins et al., 2007; Vos, Gunter, Kolk, & Mulder, 2001). The lack of AN in the garden path condition is consistent with a number of previous ERP studies of syntactic garden paths (e.g., Friederici et al., 1996Friederici et al., 2001; Osterhout & Holcomb, 1992; Osterhout et al., 1994), but it contrasts with the AN elicited by very similar garden path materials in one recent study (Kaan & Swaab, 2003a). Taken together, these findings raise the question of why there is variability across studies in the presence of the AN response.

Kaan and Swaab examined garden paths caused by ambiguous conjunctions similar to those tested in our study, as in (6a), and observed a left anterior negativity and a P600 at the disambiguating auxiliary verb (underlined), relative to unambiguous control sentences like (6b) (Kaan & Swaab, 2003a).

(6) a. The man is painting the house and the garage is already finished.
b. The man is painting the house but the garage is already finished.

We speculate that a LAN may have been elicited in Kaan and Swaab's study but not in the current study because of the form of the disambiguating verb. In the current study the disambiguating verb was a main verb appearing in past tense form (e.g., *showed*), and thus provided no information about subject-verb agreement. On the other hand, the disambiguating auxiliary in Kaan and Swaab's study was an auxiliary marked with number agreement. If the parser's initial analysis of the conjunction in (6a) was as a plural noun phrase, and if the parser's first attempt to repair the misanalysis involves an attempt to link that noun phrase to the auxiliary verb, this should trigger a mismatch between the plural NP and the singular auxiliary. Verb agreement mismatches are commonly associated with AN responses in ERP studies, and therefore this may have been the source of the early negativity in Kaan and Swaab's study.

A next question involves the issue of why syntactic violations elicit the AN in some ERP studies but not in others (for a review of the variability see Vos et al., 2001). Although a detailed review of findings about the AN lies beyond the scope of this discussion, we suggest that the AN does not reflect morphosyntactic violations in general, but rather arises when a specific morphological feature is predicted by a highly constraining context or by a marked feature on a subject NP, e.g., the marked number feature [+ plural]. This contrasts with violations that arise when a verb form conflicts with an unmarked singular subject NP, where an AN is typically not observed (Hagoort et al., 1993; Lau et al., 2006; Nevins et al., 2007). The selectivity of the AN may be explained by either of two mechanisms. A first possibility is that the AN is associated with a mismatch between the features of the incoming word and the features of a predicted word. Under this account the AN reflects a process of comparing experience with expectations rather than an attempt to create a new syntactic relation, and the selective appearance of the AN reflects the selective use of structural prediction in parsing. This is compatible with other evidence for the role of predictions in anterior negativities (Lau et al., 2006), and it is possible that the sensitivity of the AN to processing load (Gunter et al., 1997; Vos et al., 2001) ultimately reflects the sensitivity of predictive mechanisms to processing load. This account is also compatible with behavioural evidence that inflectional predictions are selectively generated for 'marked' morphological features (Gurjanov, Lukatela, Moskovljevic, Savic, & Turvey, 1985; Lukatela, Moraca, Stojnov, Savic, Katz, & Turvey, 1982). A second possibility is that the AN reflects retrieval processes, specifically a mismatch between an incoming word and a marked feature (e.g., [+ plural]) on a partially compatible item in the context. The assumption that plurals are syntactically marked whereas singulars are not syntactically marked is consistent with findings from many behavioural studies of agreement processing that show that plural NPs induce local agreement attraction errors whereas singular NPs do not (e.g., Bock & Miller, 1991; Hartsuiker, Schriefers, Bock, & Kikstra, 2003; Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, in press). This second approach does not require predictive mechanisms, and therefore is also easier to reconcile with the findings of Kaan and Swaab (2003a).

In sum, the selective elicitation of an AN in the current study is consistent with the finding from previous studies that the AN is elicited only by selective morphosyntactic anomalies. This selectivity in the presence of the AN poses a challenge for accounts of the AN that link it to general processes of morphosyntactic error detection (Friederici, 2002) or 'failure to bind' (Hagoort, 2003), since those accounts predict that the AN should be more consistently elicited by all morphosyntactic violations. Friederici and Weissenborn (2007) present an interesting alternative account of variability in LAN effects across languages, based on the functional importance of agreement for identifying grammatical relations in some languages. We cannot rule out a role for this factor, but it is unlikely to be sufficient to account for the variation in LAN effects observed within languages such as Dutch and English, since an account based on overall properties of agreement and word order in individual languages does not predict systematic variation within languages. Such an account also does not capture the absence of a LAN effect in a study of agreement violations in Hindi, a language with rich person, number, and gender agreement (Nevins et al., 2007).

P600s elicited by agreement anomalies

Based on the foregoing discussion, it is now possible to understand the results of two previous studies that have examined the properties of the P600 response elicited by different types of agreement violations. Starting from the premise that the P600 reflects attempts to create or repair syntactic relations, both in fully well-formed sentences and in cases of ungrammaticality or garden paths, we have argued that the onset of the P600 should vary systematically as a function of the time needed to begin structure building, and that the amplitude/duration of the P600 should vary as a function of the required structure building operations. For the specific case of processing anomalous subject-verb agreement, we therefore predict that the onset latency of the P600 should be a function of the time needed to recognise and analyse the incoming verb form, access the relevant features of the subject noun phrase, and detect a mismatch. We suggest that the P600 reflects unsuccessful revision processes that onset after a mismatch has been detected.

Kaan (2002) presents findings from a study that manipulated a number of properties of subject-verb agreement in Dutch. Most relevant for current purposes, she found that manipulation of the linear distance (in words) between a subject and a verb had no impact on the timing (or the amplitude) of the P600. As Kaan points out, this insensitivity to subject-verb distance can be captured by any account in which access to the subject of a clause is insensitive to the subject-verb distance. This distance insensitivity could be captured by hierarchical search-and-diagnosis mechanisms (e.g., Fodor & Inoue, 1994), content-addressable memory (McElree, 2000), or predictive construction of agreement features (e.g., Wagers et al., in press).

In contrast, Nevins and colleagues manipulated the number and type of incorrect agreement features in a study of Hindi subject-verb agreement, and found that violations involving an incorrect person feature elicited an earlier P600 response than other agreement violations (Nevins et al., 2007). Person violations are likely more salient for Hindi speakers, for orthographic and cognitive reasons, and hence recognition of the incoming verb form and detection of a mismatch with the features of the subject noun phrase may have been faster for person violations, leading to an earlier onset of the P600 component.

Comparison with other models

Although previous neurocognitive accounts of variation in the P600 have tended to focus on fewer constructions or parameters of variation, the distinction that we draw here between retrieval processes and structure-building processes is compatible with a number of models of sentence processing. The distinction is emphasised in the ACT-R based model of Lewis and colleagues (Lewis et al., 2006) and in Gibson's model of sentence processing difficulty (Gibson, 1998). In Gibson's model a number of different parameters contribute to the 'syntactic integration difficulty' of each incoming word, including retrieval, the number of relations created, and their locality, and this composite measure has been used to model reading-time data (Grodner & Gibson, 2005). Kaan et al. (2000) suggest that the same composite measure of difficulty may be reflected in the amplitude of the P600. In contrast, Phillips et al. (2005) argue that retrieval and integration contribute to the P600 in different ways, based on the finding that locality of filler-gap dependencies affects the latency but not the amplitude of the P600. Nevertheless, these findings are compatible with Gibson's model, provided that it is possible to distinguish the effects of syntactic structure building from locality effects.

A computational model by Vosse and Kempen (2000) forms the basis of a neurochronometric sentence processing model by Hagoort (2003). Vosse and Kempen's model focuses on capturing the difficulty of different types of garden path sentences. Vosse and Kempen assume a parallel parser in which multiple analyses may be pursued simultaneously, and therefore the role of explicit restructuring and reanalysis operations in serial models is replaced in their model by the dynamics of competition and lateral inhibition. Hagoort (2003) proposes that the P600 reflects the duration and the amount of competition among competing unification links for incoming words. If we further assume that a P600 is only observed when the amount of competition crosses some threshold, then this model may also be able to predict variation in the onset latency of the P600. On the basis of the outlines that Hagoort provides it is possible to make predictions about the ERP responses to violations, garden paths and long-distance dependencies observed in the current study and elsewhere. The P600 elicited by *wh*-dependency completion could be derived from competition between two alternative verb frames, one containing a *wh*-gap and one without a gap. This approach to the *wh*-dependency effect might successfully predict the effect of dependency length on the P600 onset latency (Phillips et al., 2005), due to the impact of activation decay on distant *wh*-phrases. On the other hand, it is less clear under this approach why the presence of clear case and thematic role information on the *wh*-phrase should attenuate the P600, as observed here. The model is able to predict that garden paths and grammatical violations lead to greater competition, and hence to a P600, although it is less clear how to capture the variation across constructions in P600 latency observed here and elsewhere. Nevertheless, there is a good deal of overlap between Hagoort's model and the current proposal, and the main differences reflect assumptions about the parser architecture (e.g., use of predictive structure building) rather than the linking hypotheses that relate parser properties to ERP components.

Another class of sentence processing models characterise the cost of integrating new words into a sentence in information theoretic terms, based on the *surprisal* of an incoming word (i.e., its negative log probability: Hale, 2001; Levy, 2008) or the word's contribution to *entropy reduction* (i.e., change in uncertainty about sentence completions: Hale, 2003). An advantage of these models is that they provide explicit metrics that can be used to predict the cost of garden path phenomena and long-distance dependency creation, and these metrics might be used as predictors of P600 amplitude or duration. Such models could readily account for our finding that the P600 elicited by completion of a *wh*-dependency is attenuated when the case and thematic role of the *wh*-phrase is known in advance. However, it is less clear how these models could predict the ERP consequences of detecting different types of syntactic violations. In particular, it is difficult to capture the finding that certain syntactic manipulations delay the P600 response, whereas others impact the amplitude or duration of the P600. More generally, since the information-theoretic models generally present one-dimensional measures of syntactic processing cost, it is difficult to capture the multi-dimensional variation in ERP responses to different types of syntactic manipulations.

Our proposed account of variation in the P600 and AN components shares a number of properties with models proposed by Friederici (1995, 2002). Friederici's model, like our own, characterises sentence processing mechanisms in terms of the operations of a largely serial parser. This approach is conducive to multi-dimensional accounts of how the P600 and AN are impacted by different structural manipulations, including an account of delays in the onset latency of specific components (Friederici et al., 1996, 2001). However, our account goes beyond Friederici's account in capturing the commonalities among the processing of garden paths, syntactic violations, and long-distance dependencies. Nevertheless, our characterisation of retrieval process in reanalysis could be viewed as an operationalisation of the notion of 'diagnosis' in Friederici's model, one that also naturally extends to other syntactic phenomena. By focusing on the distinction between retrieval and structure-building operations we are able to characterise each of these three domains in closely related terms.

Our discussion here has focused on the possible sources of variability in the P600 elicited by fine-grained manipulation of linguistic materials. As such, the current study is not well-suited to addressing the debate over the language-specificity or domain-generality of the P600, or its relation to the P300 component (e.g., Coulson et al., 1998a; Frisch et al., 2003; Martín-Loeches et al., 2006; Osterhout & Hagoort, 1999; Patel et al., 1998). Nevertheless, our account could be extended to a domain-general account of the P600. To the extent that processing of non-linguistic anomalies may be characterised in terms of retrieval and relation-forming processes, the predictions tested here for syntactic materials could be extended to non-linguistic materials.

CONCLUSIONS

The aim of the current study was to compare three different types of syntactic manipulations that previous studies have shown to elicit a P600-like ERP effect, using a within-subjects design and maximally similar materials. The experiment showed that ungrammaticality detection, resolution of garden paths, and completion of well-formed *wh*-dependencies all elicit a P600-like effect relative to a control condition, but with substantial differences in latency and duration across conditions, and in the case of the *wh*-dependency condition with a scalp distribution that differed in some respects from the other conditions. Thus, the first conclusion from our study is that the P600 is differentially affected by different syntactic sub-processes. The study also provided little reason to question the widespread assumption that the P600 elicited by syntactic anomaly detection and by syntactic garden paths reflects a common underlying source. With regard to the more controversial question of what *wh*-dependency formation might have in common with disambiguation and anomaly detection, we suggested that it is important to distinguish two subparts of *wh*-dependency completion in English, namely disambiguation and the formation of new syntactic relations ('syntactic integration'). The verb that allows completion of a *wh*-dependency typically also disambiguates the case and thematic role of the *wh*-phrase. This disambiguating function may straightforwardly be related to other types of disambiguation that elicit the P600 response, and hence the elimination of this ambiguity in the current study may account for the reduced P600 component. We suggested that the multidimensional variation in ERP responses to syntactic processes, and variation in the P600 response in particular, may be understood in terms of the prediction, retrieval, and structure building operations needed to create syntactic structure. We suggested that the latency of the P600 reflects the time needed to retrieve the elements that participate in a structural relation, and that the amplitude and duration of the P600 is a function of the assembly (and disassembly) of syntactic relations. These proposals suggest a ‘common currency’ for understanding the computations that underlie the electrophysiology of syntactic processing.

Manuscript received May 2007
Revised manuscript received March 2009
First published online Month/Year

REFERENCES

Ainsworth-Darnell, K., & Shulman, H., & Boland, J. (1998).Disssociating brain responses to syntactic and semantic anomalies: evidence from event-related potentials *Journal of Memory and Language*, *38*, 112–130.

Allen, M., & Badecker, W., & Osterhout, L. (2003).Morphological analysis in sentence processing: An ERP study *Language and Cognitive Processes*, *18*, 405–430.

Bock, J. K., & Miller, C. A. (1991).Broken agreement *Cognitive Psychology*, *23*, 45–93.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008).An alternative perspective on “semantic P600” effects in language comprehension *Brain Research Reviews*, *59*, 55–73.

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

Coulson, S., & King, J. W., & Kutas, M. (1998b).ERPs and domain specificity: Beating a straw horse *Language and Cognitive Processes*, *13*, 653–672.

Davidson, R. J., & Jackson, D. C., & Larson, C. L. (2000).Human electroencephalography In J. T. Cacioppo, & L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (2nd edn).

Domahs, U., & Wiese, R., & Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008).The processing of German word stress: evidence for the prosodic hierarchy *Phonology*, *25*, 1–36.

Donaldson, D., & Rugg, M. (1999).Event related potential studies of associative recognition and recall: Electrophysiological evidence for context dependent retrieval process *Brain Research: Cognitive Brain Research*, *8*, 1–16.

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

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

Felser, C., & Clahsen, H., & Münte, T. F. (2003).Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and *wh*-movement in German *Brain and Language*, *87*, 345–354.

Fiebach, C. J., & Schlesewsky, M., & Friederici, A. D. (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.

Fodor, J. D., & Inoue, A. (1994).The diagnosis and cure of garden paths *Journal of Psycholinguistics Research*, *23*, 519–560.

Frazier, L. (1985).Syntactic complexity In D. Dowty, & L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing*.

Frazier, L., & Clifton, C., Jr. (1996). *Construal*.

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. D. (1995).The time course of syntactic activation during language processing: a model based on neurological and neurophysiological data *Brain and Language*, *50*, 259–281.

Friederici, A. D. (2002).Towards a neural basis of auditory sentence processing *Trends in Cognitive Sciences*, *6*, 78–84.

Friederici, A. D., & Hahne, A., & Mecklinger, A. (1996).The temporal structure of syntactic parsing: Early and late event-related potential effects elicited by syntactic anomalies *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1219–1248.

Friederici, A. D., & Mecklinger, A. (1996).Syntactic parsing as revealed by brain responses: First-pass and second-pass processes *Journal of Psycholinguistic Research*, *25*, 157–176.

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*, 305–323.

Friederici, A. D., & Pfeifer, E., & Hahne, A. (1993).Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations *Cognitive Brain Research*, *1*, 183–192.

Friederici, A. D., & Weissenborn, J. (2007).Mapping sentence form onto meaning: The syntax-semantic interface *Brain Research*, *1146*, 50–58.

Frisch, S., & Kotz, S. A., & von Cramon, D. Y., & Friederici, A. D. (2003).Why the P600 is not just a P300: the role of the basal ganglia *Clinical Neuropsychology*, *114*, 336–340.

Garnsey, S., & Pearlmutter, N. J., & Myers, E., & Lotocky, M. A. (1997).The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences *Journal of Memory and Language*, *37*, 58–93.

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.

Grodner, D., & Gibson, E. (2005).Consequences of the serial nature of linguistic input *Cognitive Science*, *29*, 261–291.

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.

Gurjanov, M., & Lukatela, G., & Moskovljevic, J., & Savic, M., & Turvey, M. T. (1985).Grammatical priming of inflected nouns by inflected adjectives *Cognition*, *19*, 55–71.

Hagoort, P. (2003).How the brain solves the binding problem for language: a neurocomputational model of syntactic processing *Neuroimage*, *20*, S18–S29.

Hagoort, P., & Brown, C. M. (1994).Brain responses to lexical ambiguity resolution and parsing In C. Clifton Jr., & L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing*.

Hagoort, P., & Brown, C. M., & Groothusen, J. (1993).The syntactic positive shift as an ERP measure of syntactic processing *Language and Cognitive Processes*, *8*, 439–483.

Hagoort, P., & Brown, C. M., & Osterhout, L. (1999).The neurocognition of syntactic processing In C. M. Brown, & P. Hagoort (Eds.), *The neurocognition of language*.

Hagoort, P., & Wassenaar, M., & Brown, C. (2003a).Syntax-related ERP-effects in Dutch *Cognitive Brain Research*, *16*, 38–50.

Hagoort, P., & Wassenaar, M., & Brown, C. (2003b).Real-time semantic compensation in patients with agrammatic comprehension: electrophysiological evidence for multiple-route plasticity *Proceedings of the National Academy of Sciences USA*, *100*(7), 4340–4345.

Hahne, A., & Friederici, A. D. (1999).Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes *Journal of Cognitive Neuroscience*, *11*, 193–204.

Haig, A. R., & Gordon, E., & Hook, S. (1997).To scale or not to scale: McCarthy & Wood revisited *Electroencephalography and Clinical Neurophysiology*, *103*, 323–325.

Hale, J. (2001).A probabilistic Earley parser as a psycholinguistic model *Proceedings of NAACL*, *2*, 159–166.

Hale, J. (2003).The information carried by words in sentences *Journal of Psycholinguistic Research*, *32*, 101–123.

Hartsuiker, R. J., & Schriefers, H. J., & Bock, K., & Kikstra, G. M. (2003).Morphological influences on the construction of subject-verb agreement *Memory and Cognition*, *31*, 1316–1326.

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.

Jing, H., & Pivik, R. T., & Dykman, R. A. (2006).A new method for topographic scaling of event-related potentials *Journal of Neuroscience Methods*, *151*, 239–249.

Kaan, E. (2002).Investigating the effects of distance and number interference in processing subject-verb dependencies: An ERP study *Journal of Psycholinguistic Research*, *31*, 165–193.

Kaan, E., & Harris, A., & Gibson, E., & Holcomb, P. J. (2000).The P600 as an index of syntactic integration difficulty *Language and Cognitive Processes*, *15*, 159–201.

Kaan, E., & Swaab, T. Y. (2003a).Electrophysiological evidence for serial sentence processing: A comparison between non-preferred and ungrammatical continuations *Cognitive Brain Research*, *17*, 621–635.

Kaan, E., & Swaab, T. Y. (2003b).Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation *Journal of Cognitive Neuroscience*, *15*, 98–110.

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

Kluender, R., & Kutas, M. (1993).Subjacency as a processing phenomenon *Language and Cognitive Processes*, *8*, 573–633.

Kolk, H. H. J., & Chwilla, D. J., & van Herten, M., & Oor, P. (2003).Structure and limited capacity in verbal working memory: a study with event-related potentials *Brain and Language*, *85*, 1–36.

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

Kuperberg, G. R., & Kreher, D. A., & Sitnikova, T., & Caplan, D., & Holcomb, P. (2007).The role of animacy and thematic relationships in processing active English sentences *Brain and Language*, *100*, 223–238.

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

Kutas, M. (1997).Views on how the electrical activity that the brain generates reflects the functions of different language structures *Psychophysiology*, *34*, 383–398.

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

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

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

Kutas, M., & van Petten, C. (1994).Psycholinguistics electrified: Event-related brain potential investigations In M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics.*

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

Lau, E., & Stroud, C., & Plesch, S., & Phillips, C. (2006).The role of structural prediction in rapid syntactic analysis *Brain and Language*, *98*, 74–88.

Levy, R. (2008).Expectation-based syntactic comprehension *Cognition*, *106*, 1126–1177.

Lewis, R. L., & Vasishth, S., & van Dyke, J. A. (2006).Computational principles of working memory in sentence comprehension *Trends in Cognitive Sciences*, *10*, 447–454.

Luck, S. J. (2005). *An introduction to the event-related potential technique.*

Lukatela, G., & Moraca, J., & Stojnov, D., & Savic, M., & Katz, L., & Turvey, M. T. (1982).Grammatical priming effects between pronouns and inflected verb forms *Psychological Research (Psychologische Forschung)*, *44*, 297–311.

Martín-Loeches, M., & Casado, P., & Gonzalo, R., & de Heras, L., & Fernández-Frías, C. (2006).Brain potentials to mathematical syntax problems *Psychophysiology*, *43*, 579–591.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62*, 203–208.

McElree, B. (2000).Sentence comprehension is mediated by content-addressable memory *Journal of Psycholinguistic Research*, *29*, 111–123.

McElree, B., & Foraker, S., & Dyer, L. (2003).Memory structures that subserve sentence comprehension *Journal of Memory and Language*, *48*, 67–91.

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.

Mills, D. L., & Prat, C., & Zangl, R., & Stager, C. L., & Neville, H. J., & Werker, J. F. (2004).Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds *Journal of Cognitive Neuroscience*, *16*, 1452–1464.

Münte, T. F., & Szentkui, A., & Wieringa, B. M., & Matzke, M., & Johannes, S. (1997).Human brain potentials to reading syntactic errors in sentences of different complexity *Neuroscience Letters*, *235*, 105–108.

Münte, 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*, 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.

Nevins, A. I., & Dillon, B. W., & Malhotra, S., & Phillips, C. (2007).The role of feature-number and feature-type in processing Hindi verb agreement violations *Brain Research*, *1164*, 81–94.

Núñez-Peña, M. I., & Honrubia-Serrano, M. L. (2004).P600 related to rule violation in an arithmetic task *Cognitive Brain Research*, *18*, 130–141.

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

Osterhout, L., & Bersick, M., & McLaughlin, J. (1997).Brain potentials reflect violations of gender stereotypes *Memory and Cognition*, *25*, 273–285.

Osterhout, L., & Hagoort, P. (1999).A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King, & Kutas (1998) in the P600/SPS–P300 debate *Language and Cognitive Processes*, *14*, 1–14.

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

Osterhout, L., & Holcomb, P. J. (1993).Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech *Language and Cognitive Processes*, *8*, 413–438.

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., & McKinnon, R., & Bersick, M., & Corey, V. (1996).On the language-specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience*, *8*, 507–526.

Osterhout, L., & Mobley, L. A. (1995).Event-related brain potentials elicited by failure to agree *Journal of Memory and Language*, *34*, 739–773.

Patel, A. D., & Gibson, E., & Ratner, J., & Besson, M., & Holcomb, P. J. (1998).Processing syntactic relations in language and music: An event-related potentials study *Journal of Cognitive Neuroscience*, *10*, 717–733.

Pearlmutter, N. J., & Garnsey, S. M., & Bock, J. K. (1999).Agreement processes in sentence comprehension *Journal of Memory and Language*, *41*, 427–456.

Pickering, M., & Barry, G. (1991).Sentence processing without empty categories *Language and Cognitive Processes*, *6*, 229–259.

Phillips, C., & Kazanina, N., & Abada, S. (2005).ERP effects of the processing of syntactic long-distance dependencies *Cognitive Brain Research*, *22*, 407–428.

Phillips, C., & Wagers, M. (2007).Relating structure and time in linguistics and psycholinguistics In G. Gaskell (Eds.), *The Oxford handbook of psycholinguistics.*

Stroud, C. (2008). *Structural and semantic selectivity in the electrophysiology of sentence comprehension*.

Stroud, C., & Phillips, C. (2008). *The structural and semantic selectivity of the ''thematic'' P600 in sentence comprehension*.

Trueswell, J., & Tanenhaus, M., & Kello, C. (1993).Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 528–553.

Urbach, T. P., & Kutas, M. (2002).The intractability of scaling scalp distributions to infer neuroelectric sources *Psychophysiology*, *39*, 791–808.

van Berkum, J. J. A., & Koornneef, A. W., & Otten, M., & Nieuwland, M. S. (2007).Establishing reference in language comprehension: An electrophysiological perspective *Brain Research*, *1146*, 158–171.

van Herten, M., & Kolk, H. H. J., & Chwilla, D. J. (2005).An ERP study of P600 effects elicited by semantic anomaly *Cognitive Brain Research*, *22*, 241–255.

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

Vos, S. H., & Gunter, T. C., & Kolk, H. H. J., & Mulder, G. (2001).Working memory constraints on syntactic processing: An electrophysiological investigation *Psychophysiology*, *38*, 41–63.

Vosse, T., & Kempen, G. (2000).Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar *Cognition*, *75*, 105–143.

Wagers, M. W., Lau, E. F., & Phillips, C. (in press). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language.*