LANGUAGE AND COGNITIVE PROCESSES, 1993, 8 (4) 413-437

Event-related Potentials and Syntactic Anomaly: Evidence of Anomaly Detection During the Perception of Continuous Speech

Lee Osterhout

Department of Psychology, University of Washington, Seattle, Washington, USA

Phillip J. Holcomb

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

Event-related brain potentials (ERPs) were recorded from 13 scalp electrodes while subjects listened to sentences containing syntactic ambiguities. Words that were inconsistent with the "preferred" sentence structure elicited a positive-going wave (the P600 effect), similar to that elicited by such words during reading (Osterhout & Holcomb, 1992). These results suggest that (1) ERPs recorded during the comprehension of spoken sentences are sensitive to the syntactic anomaly engendered by disambiguating material following erroneous analysis of a syntactically ambiguous string (the "garden-path" effect), (2) the parsing strategies employed during sentence comprehension are (in some circumstances) constant across modalities, and (3) syntactic analysis of spoken sentences is temporally close to the acoustic input.

INTRODUCTION

For all but a limited set of simple sentences, sentence comprehension requires the comprehender to determine the structural relations that hold between words in the sentence; that is, the sentence must be parsed. Recent efforts to investigate sentence parsing have been largely restricted

Requests for reprints should be addressed to Lee Osterhout, Department of Psychology, NI-25, University of Washington, Seattle, Washington 98195, USA.

We thank Susan Garnsey and several anonymous reviewers for comments on previous drafts of the manuscript, Tim O'Rourke for running subjects, John Kounios for help in stimulus creation, and NIDCD Grant DCO 1947 to the first author and NICHD Grant HD25889 to the second author for financial support.

to examinations of these processes during reading (Ferreira & Clifton, 1986; Frazier & Rayner, 1982; Holmes, Stowe, & Cupples, 1989; Mitchell & Holmes, 1985; Rayner, Carlson, & Frazier, 1983; for review, see Frazier, 1987). These studies provide compelling support for certain classes of parsing models over others (cf. Clifton, Speer, & Abney, 1991) and are frequently cited as evidence that specific parsing strategies are applied with broad generality during sentence comprehension (e.g. Frazier, 1987). However, it remains uncertain whether the parsing characteristics uncovered to date are specific to reading or are general to all modes of language comprehension. In the experiment reported here, we examined sentence parsing during the comprehension of sentences presented as continuous, natural speech.

Efforts to examine syntactic analysis during reading have focused on instances of local (i.e. temporary) syntactic ambiguity, such as that in (1):

1. The broker persuaded ...
   (a) the investor to sell the stock.
   (b) to sell the stock was sent to jail.

Two grammatical structures can be assigned to the string *The broker persuaded;* the proper analysis cannot be determined until disambiguating material is encountered. A simple active analysis, consistent with continuation (a), would attach the verb *persuade* to the main clause (see Fig. la). The alternative analysis, consistent with continuation (b), involves passivising the verb and attaching it to a reduced relative clause embedded within the main clause (meaning roughly "The broker who was persuaded to sell the stock . . .") (Fig. 1b).

A number of researchers have examined the response to such ambiguity by measuring eye fixations and regressions during reading. The general finding has been that readers exhibit longer fixations and/or more regressions upon encountering continuations like (b), relative to continuations like (a) (Frazier, 1987; Frazier & Rayner, 1982; Rayner et al., 1983). Frazier and her colleagues have proposed two important interpretations of these data. First, they posit that readers rapidly construct a single "preferred" syntactic representation of the sentence (the serial parsing model), as opposed to building all possible structures in parallel (cf. Gorrell, 1989) or waiting until the proper syntactic roles can be assigned with certainty (cf. Marcus, 1980). Second, they argue that the preferred analysis is generally the simplest analysis consistent with the phrase structure constraints of the grammar, as defined by syntactic complexity (i.e. the number of nodes in the phrase structure tree); they refer to this as the *minimal attachment principle.* In sentence (1), continuation (a) is consistent with the simple active minimal attachment analysis. However, continuation (b) is not; the parser, "garden-pathed" in

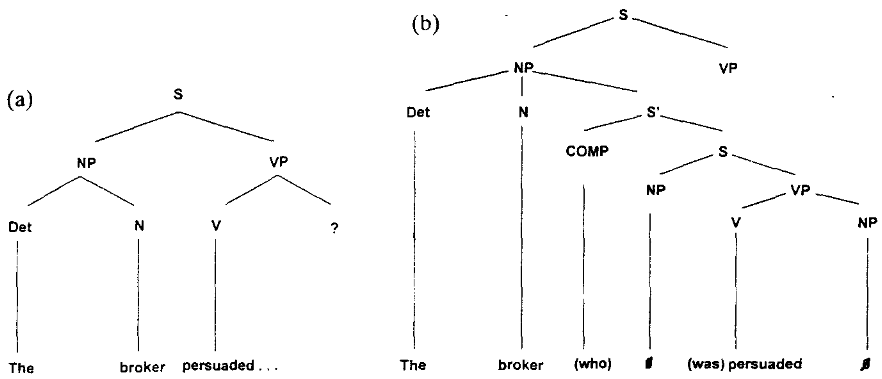

FIG. 1 Two possible interpretations of the word string "The broker persuaded . . .". (a) A simple active interpretation; (b) a passivised, reduced relative clause interpretation.

its initial pass through the sentence, must backtrack and reanalyse the earlier regions of the sentence. Data from experiments using other methods, such as word-byword or phrase-by-phrase reading time, have also been claimed to support this view (Ferreira & Henderson, 1990; for dissenting views and evidence, see Holmes, 1987; Holmes et al., 1989; Mitchell & Holmes, 1985).

Another measure that has been shown to be sensitive to the syntactic analysis of sentences is the recording of event-related brain potentials (ERPs) elicited during reading. ERPs are patterned voltage changes in the ongoing electroencephalogram that are time-locked to the onset of a sensory, motor or cognitive event (Hillyard & Picton, 1987). Scalp-recorded ERPs consist of a series of positive and negative voltage peaks (or "components") which are distributed across time. Importantly, certain late-occurring ("endogenous") ERP components appear to be sensitive to specific aspects of language comprehension. In a series of pioneering experiments, Kutas and her colleagues (Kutas & Hillyard, 1980a; 1980b; 1980c; for review, see Kutas & Van Petten, 1988) demonstrated that semantically anomalous words in sentences elicit a negative-going wave with a peak amplitude around 400 msec after the onset of the word (the N400 component). More recently, Neville et al. (1991) have reported that certain outright violations of syntactic constraints elicit a measurable ERP response. Two types of anomaly (violations of phrase structure and subjacency constraints; cf. Radford, 1988) were followed by a late positivegoing wave with an onset around 500 msec after the anomaly.

ERPs also appear to be sensitive to anomaly that results from garden-path effects (Osterhout, 1990; Osterhout & Holcomb, 1990; 1992; Osterhout & Swinney, 1989). For example, Osterhout and Holcomb (1992) examined the ERP response to sentences similar to (1)-(4) in Table 1. A minimal attachment parser would initially attempt a simple active analysis

TABLE 1

Examples of Sentences Used by Osterhout and Holcomb (1992) and in the Current Study. Words of Interest are Underlined

1. **The broker hoped to sell the stock**
2. ***The broker persuaded to sell the stock**
3. ***The broker hoped to sell the stock was sent to jail**
4. **The broker persuaded to sell the stock was sent to jail**

of all four sentences in Table 1. This analysis would result in anomaly when the infinitival marker to is encountered in sentences (2) and (4), since the subcategorisation properties of the verb *persuade* do not allow a constituent beginning with the word to (i.e. an infinitival clause or a prepositional phrase) to be directly attached to the verb, when the verb is used in its active form. In contrast, the subcategorisation properties of *hope* allow an infinitival clause to be directly attached to the verb in sentences (1) and (3). Osterhout and Holcomb found that the word to elicited a positive-going wave (which they labelled the "P600 effect") in sentences like (2) and (4), relative to the identical words in sentences like (1) and (3) (Fig. 2a). At some electrode sites, the onset of the P600 effect occurred approximately 300 msec subsequent to presentation of the word to. Assuming that the P600 is an electrophysiological marker of syntactic anomaly, this result is consistent with a minimal attachment parser that rapidly applies verb subcategorisation information.

Also of interest is the response to the auxiliary verb (was) in sentences (3) and (4). If the parser initially attempts a simple active analysis of (3), it would encounter difficulty attaching this word to the syntactic structure assigned to the preceding sentence fragment. Given a simple active analysis of the main clause, the phrase structure rules of English prevent the words was lying from being attached to the preceding sentence fragment; hence, the auxiliary verb represents a violation of the phrase structure constraints. In contrast, a parser attempting an analysis of (4) should recognise the need for a passivised reduced relative clause interpretation prior to encountering the auxiliary verb. The relative clause analysis allows the auxiliary verb to be attached as part of the main clause ("The broker was sent to jail"), thereby avoiding the syntactic anomaly engendered by the auxiliary verb in (3). Osterhout and Holcomb (1992) observed a large P600 effect to auxiliary verbs in sentences like (3), relative to the ERPs elicited by the same words in sentences like (4) (Fig. 2b).

Taken together, the data reviewed above appear to provide strong support for a parsing model in which simple structures are initially attempted. However, there is still the question of whether or not this model applies across modalities. There are compelling reasons for doubt. One

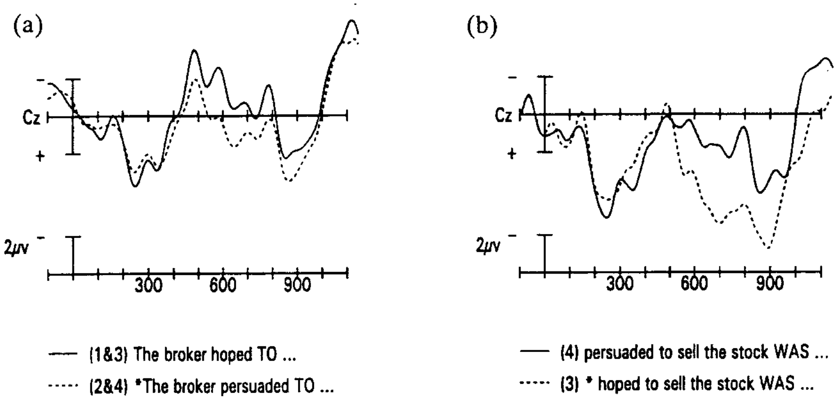

FIG. 2 Grand average ERPs (across subjects and exemplars) from site Cz. Data are from Osterhout and Holcomb (1992), who presented sentences visually. Onset of target is at 0 msec (vertical calibraticn bar). Each mark on the horizontal axis represents 100 msec. (a) ERPs to the infinitival marker *to* in sentences with "hope" verbs and "persuade" verbs, as shown in Table 1; (b) ERPs to auxiliary verbs (e.g. *was*) in sentences like (3) and (4) in Table 1.

non-trivial difference between modalities is the existence of prosodic cues to phrase structure within spoken language (e.g. Cooper & Paccia-Cooper, 1980; Cooper & Sorensen, 1981; Klatt, 1975). Although these cues (changes in fundamental frequency, syllable duration and pause duration) have been shown to maintain at least a loose correlation with syntactic structure (and particularly clause boundaries), little is known about the influence such cues might have on the on-line parsing of sentences. The minimal attachment model discussed above predicts that the parser's initial analyses will remain unaffected by prosodic cues, since such information is phonological rather than syntactic. More highly interactive models (e.g. Marslen-Wilson, 1975) predict that prosodic information can influence the initial parse of a sentence. One implication of the interactive model is that the presence of prosodic information should direct the parser to the appropriate analysis in the presence of syntactic ambiguity, thereby avoiding garden paths. In the present study, we wanted to determine whether or not the sentence parser operates similarly on auditory and visual input, when the auditory input is devoid of strong prosodic cues to sentence structure. Subsequent work can then determine the influence of prosodic information on these processes.

A second important difference between modalities is that reading is self-paced and allows re-examination of previously encountered words; that is, the comprehender typically controls the timing and sequence of stimulus presentation. In contrast, the timing and sequence of stimulus presentation during spoken language comprehension is controlled by the speaker. This

lack of control over stimulus presentation, coupled with the rapidity of fluent speech (Cooper & Paccia-Cooper, 1980; Klatt, 1975; Lenneberg, 1967), places severe temporal constraints on comprehension that are not present during reading. Hence, while a strategy that entails frequent backtracking and re-analysis might work well during reading (where the comprehender can re-read previously encountered material during a second-pass parse), such a strategy might be less successful during spoken language comprehension (where stimulus re-examination is not possible).

The present research extends the examination of syntactic analysis into an investigation of these processes during the on-line comprehension of spoken sentences by attempting to replicate with auditory stimuli the effects reported by Osterhout and Holcomb (1992). The current study is focused on three issues. First, we hoped to determine whether or not measurable electrophysiological correlates of garden-path effects are elicited during spoken language comprehension. Second, we wished not only to discover whether these correlates exist and are equivalent to those previously observed during reading, but also to determine the efficacy with which they could be applied to the study of the syntactic analysis of spoken sentences. Specifically, by monitoring for the occurrence of ERP correlates of the garden-path effect, we hoped to provide a preliminary answer to the following question: In the absence of strong intonational cues to phrase structure, are the parsing strategies employed during spoken language comprehension similar to those employed during reading? And, finally, the methodology employed by Osterhout and Holcomb (1992), in which words were presented sequentially at a rate of 650 msec per word, is rather far removed from "normal" comprehension environments. We wanted to determine if the results observed during the visual experiments would also be obtained under more naturalistic comprehension conditions.

METHOD

Subjects

Twenty-four undergraduates participated for class credit or as volunteers. All of them were right-handed native speakers of English, with normal or corrected-to-normal vision.

Materials

Four versions of 120 "root" sentences were constructed, as exemplified by (1)-(4) in Table 1. (Each set of four versions will be referred to as a "sentence set".) Sentences similar to (1) and (3) contained verbs that accept an infinitival clause complement (e.g. *hope)*. Sentences similar to

(2) and (4) contained verbs that do not typically accept complements beginning with the word to (infinitival clauses and prepositional phrases) when used in an active form and without a noun phrase acting as direct object (e.g. persuade). Fifteen verbs of each type were selected; verb selection was governed by the authors' intuitions concerning allowable complement structures of verbs. Each verb appeared equally often in acceptable and unacceptable sentences. The two verb classes did not significantly differ in mean frequency (Kucera & Francis. 1967), either in their present-tense form ["hope" verbs = 74, "persuade" verbs = 64, $t(14) < 1$] or in their past-tense form ["hope" verbs = 80, "persuade" verbs = 37, $t(14) = 2.00$, $P > 0.05$]. The mean length (in past-tense form) of "hope" and "persuade" verbs was 7.6 and 7.7 letters, respectively.

The experimental sentences were counterbalanced across four stimulus lists in a Latin square design, such that each list contained only one version of each sentence, and 30 exemplars of each sentence type. These materials were pseudorandomly mixed prior to presentation. The entire set of experimental sentences can be found in Osterhout and Holcomb (1992, appendix 2). Given the length and number of these sentences, we did not include any filler items. We were concerned that increasing the number of trials would decrease the quality of our obtained data, due to subject boredom and fatigue. However, since active and reduced relative clause (and "acceptable" and "unacceptable") sentences appeared equally often, we did not consider filler sentences to be a necessity. Furthermore, the results reported by Osterhout and Holcomb (1992) remained constant across experiments, even though filler items were present in only one of the experiments they report.

The four sentences in each sentence set were tape-recorded by a practised male speaker (who read at a normal speaking rate of approximately five syllables per second, and with neutral intonation)[1] and subsequently digitised (16 kHz, 12-bit resolution, 6-kHz Butterworth filter). These recorded versions were not always used as the experimental sentences. Instead, experimental sentences were created via signal processing software in the following manner. For each sentence set, one sentence was chosen to act as the "frame" sentence. All sentences in the sentence set were then created by digitally modifying the frame sentence (i.e. words were digitally removed from and spliced into the frame sentence). For example, to create the experimental version of sentence (1) in Table 1 from the "frame" sentence (4), the verb persuaded was

[1]The speaker was instructed to avoid pitch changes, pauses and syllable lengthenings that might allow the subject to predict upcoming structure. Recorded sentences were screened by the experimenters for obvious intonational cues to sentence structure. Those containing obvious instances of such cues were re-recorded.

replaced by the verb hoped from the recorded version of sentence (1), and the fragment stock was sent to jail was replaced by the sentence-final word stock from the recorded version of (1). (The word stock was replaced in order to accommodate sentence-ending intonational patterns.) Word boundaries were identified visually and confirmed auditorily. In some cases, word boundaries were more ambiguous than in others. For example, the initial plosive sound in the word to was always easily identifiable, whereas the semi-vowel or fricative word-initial sounds in the auxiliary verbs (e.g. was or had) were not always as clearly identifiable. Hence, it is probable that more temporal variation existed in word boundary identification for some classes of words than for others. In a few cases, coarticulation effects made it necessary to also replace the final syllable of the pre-verbal noun (e.g. broker) with the corresponding syllable from the recorded version of the sentence being created. The sentence version acting as the frame sentence was counterbalanced, such that across all sentence sets all versions acted as the frame with equal frequency. This procedure ensured that the critical regions of the sentences were acoustically identical across the four versions of each sentence within a sentence set, and also mitigated the influence of any subtle intonational cues that might have predicted subsequent sentence structure. Mean durations (in msec) of the critical words across all sentences were as follows: the word to, 154; auxiliary verbs (e.g. was), 313; final words in short sentences, 657; final words in long sentences, 603. Examples of the recorded and experimental versions of the sentences were chosen at random and played for several naive listeners, who were at chance levels in determining which sentences had been recorded and which had been digitally modified.

Procedure

All of the stimuli were generated and controlled by an IBM-PC computer. Each trial consisted of the following events. A fixation cross appeared in the centre of the CRT, and remained for the duration of the sentence. The digital representation of the sentence was then output through a digital-to-analogue converter and binaurally presented to the subject over headphones at a comfortable listening level (approximately 65 dBsl). A 1450 msec blank screen interval followed each sentence, after which a prompt appeared on the CRT asking subjects to decide if the previous sentence was an "acceptable" or "unacceptable" sentence. Acceptable sentences were defined as semantically coherent and grammatically correct; unacceptable sentences were defined as those which were semantically incoherent or bizarre, or which were judged as being ungrammatical. The subjects were provided with a few examples of syntactically and semantically anomalous sentences. No sentences presented during the experiment were used as examples. The subjects were asked if they understood the criteria for acceptability, and additional examples were provided as needed. The subjects responded by pressing one of two buttons. The buttons used to indicate "acceptable" or "unacceptable" (left or right hand) were counterbalanced across subjects. The subjects were tested in one session, which lasted 1-2 h, during which they were seated in a comfortable chair situated in a sound-attenuating chamber.

Recording System. EEG activity was recorded from 13 scalp locations, using tin electrodes attached to an elastic cap (Electrocap International). Electrode placement included the International 10-20 system locations (Jasper, 1958) at homologous positions over the left and right occipital (O1, 02) and frontal (F7, 178) regions, and from the frontal (Fz), central (Cz) and parietal (Pz) midline sites. In addition, several non-standard sites over posited language centres were used, including Wernicke's area (WL, WR: 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz), posterior-temporal (PTL, PTR: 33% of the interaural distance lateral to Cz) and anterior-temporal (ATL, ATR: half the distance between F7 and T3, and between F8 and T4). Vertical eye movements and blinks were monitored by means of an electrode placed beneath the left eye, and horizontal eye movements were monitored by an electrode positioned to the right of the right eye. The above 15 channels were referenced to an electrode placed over the left mastoid bone, and were amplified with a bandpass of 0.01 to 100 Hz (3 dB cut-off) by a Grass Model 12 amplifier system. Activity over the right mastoid bone was actively recorded on a sixteenth channel in order to determine if there were lateral asymmetries associated with the left mastoid reference.

Data Summary

Continuous analogue-to-digital conversion of the EEG and stimulus trigger codes was performed on-line by a Data Translation 2801-A board and an AT-compatible computer, at a sampling frequency of 200 Hz. ERPs were quantified by computer as the mean voltage within a latency range time-locked to the onset of words of interest, relative to the 100 msec of activity preceding onset of the word of interest. Trials characterised by excessive eye movement (vertical or horizontal) or amplifier blocking were rejected prior to signal averaging. Less than 10% of the trials were removed due to artifact. Data acquired at the midline and lateral sites were treated separately during the analyses, to allow for quantitative analysis of hemispheric differences. To protect against excessive type I error due to violations of the assumption of equal variances of differences between

conditions of within-subject factors (Huynh & Feldt, 1976), the Geisser-Greenhouse correction (Geisser & Greenhouse, 1959) was applied when evaluating effects with more than one degree of freedom. In all the analyses reported below, ERP averaging was performed without regard to the subjects' behavioural response. Analyses based on response-contingent ERP averages were also examined, and did not differ in any noticeable way from the averages reported below.

RESULTS

Behavioural Data

The subjects judged the four sentence types in Table 1 to be acceptable on the following percentages of trials: sentences of type (1), 95%; sentences of type (2), 13%; sentences of type (3), 7%; sentences of type (4), 71%. The relatively low percentage of "acceptable" judgements to the long sentences containing "persuade" verbs [sentences of type (4)] is not unexpected. Although these sentences are technically grammatical, they are complex in structure.

Event-related Potentials

Due to the rapid temporal sequencing of words within each sentence (causing partial overlap of components associated with each word), the general shape of the obtained waveforms deviated from that typically observed with presentation of language stimuli. In some cases, the early "exogenous" components generally observed within the first 250 msec of activity [e.g. the negative-positive complex (Nl-P2)] were not clearly visible.

Response to the Word "to". The garden-path parsing model described above predicts that the infinitival marker to will produce a garden-path effect when it follows a "persuade" verb, but not when it follows a "hope" verb. Responses to the word *to* following "persuade" and "hope" verbs

---

[2]**The early-occurring components of the ERP waveform (e.g. the N1-P2 complex) were reduced for a number of reasons. Most importantly. given the nature of the stimuli (continuous speech), ERPs to successive words overlapped to a substantial degree. This caused some of the smaller components to "wash out" during signal averaging. Furthermore, it has been demonstrated that during the perception of continuous speech, the N1-P2 complex becomes refractory, due to the rapid presentation of words (Holcomb & Neville, 1991). The endogenous components of primary interest in this study were sufficiently large to remain visible after signal averaging.**

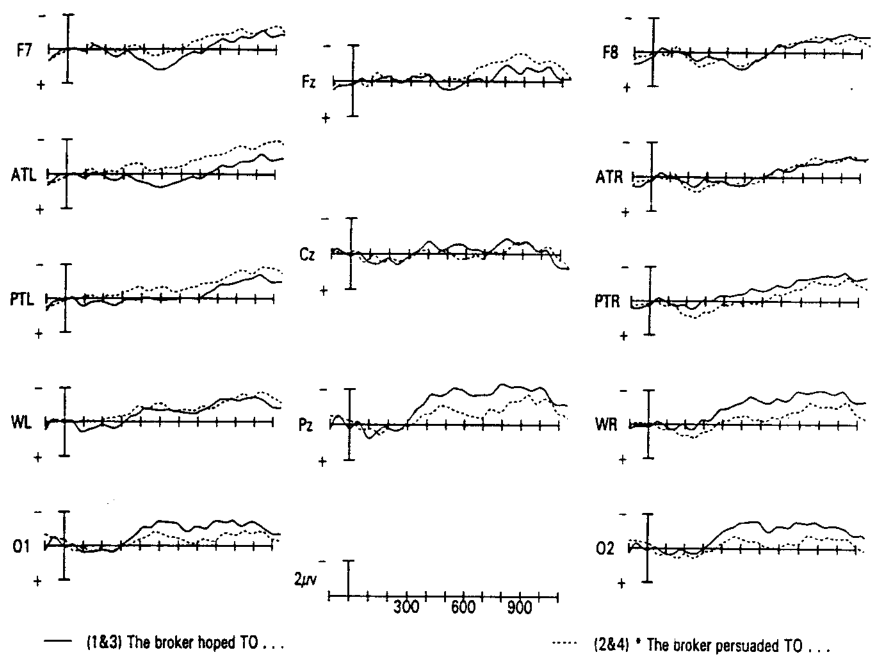

FIG. 3 Grand average ERPs (across all subjects and exemplars), from 13 scalp sites, to the infinitival marker *to* in the four sentence types shown in Table 1. Each plot represents averages made over approximately 1300 trials.

are shown in Fig. 3. ERPs to infinitival markers in sentences like (1) and (3) in Table 1 were averaged together, as were ERPs to the infinitival markers in (2) and (4).

An examination of Fig. 3 shows that ERPs to the word to in the two sentence types began to diverge early in the epoch, around 50 msec at some [3] After close inspection, we found no effect of the experimental variables on the right mastoid recording. Therefore, in this and all other analyses, a left mastoid reference was used. Also, in describing ERP effects elicited during critical regions of the sentences, we on occasion state that the ERP effect is elicited following the presentation of some critical word, e.g. the P600 was elicited following presentation of the critical word to in the ungrammatical sentences. We avoid claiming that the ERP effect is *elicited by* the critical word, since in the present study we cannot claim this with certainty given the rapid presentation of words. However, we believe that the critical word is in fact eliciting the ERP effects of interest, for the following reasons: (1) the effects are highly similar to those reported by Osterhout and Holcomb *(1992),* who visually presented words at a slow rate, and where the effects were clearly elicited by the critical words; (2) the ERP waveforms in the "anomalous" and "nonanomalous" conditions tend to diverge very rapidly subsequent to presentation of the critical words; and (3) our ERP averages are time-locked to the onset of the critical words themselves, and not to the onset

of words that appeared subsequent to the critical words. electrode sites. These differences extended throughout the epoch. At the posterior and right hemisphere sites, ERPs to infinitival markers that followed "persuade" verbs were more positive-going than those to the same words that followed "hope" verbs. In addition, the response to the word to following "persuade" verbs was associated with a negative-going wave at the parietal and frontal regions of the left hemisphere. In order to quantify these differences between sentence types, analyses were performed on the mean amplitude within three windows of activity: 50-300, 300-500 and 500-800 msec. These windows were chosen because they roughly correspond to the latency ranges of the N1-P2 complex, N400 component and P3/P600 slow wave components typically reported in cognitive ERP studies.

Reliable differences between conditions were found at all three windows. In the 50-300 msec window, no reliable main effects or interactions were found at the midline sites. However, the lateral site ANOVA revealed significant interactions between sentence type and hemisphere [$F(1,23) = 13.89$, $P < 0.001$] and between sentence type, hemisphere and electrode site [$F(4,92) = 5.70$, $P < 0.01$]. These interactions reflected the greater positivity to infinitival markers following "persuade" verbs at the right hemisphere sites, and greater negativity to these words at the parietal and anterior left hemisphere sites, relative to the ERPs to the same words following "hope" verbs.[4]

[4]Given the close temporal relationship between onset of the main verbs in these sentences and subsequent onset of the infinitival markers, and given that different verbs preceded the infinitival markers in the two sentence types, one explanation for this early effect is that it reflects a differential response to the verbs preceding the infinitival marker, rather than a response to the infinitival markers themselves. For example, if the onset of the infinitival markers occurred prior to resolution of the N400 component elicited by the verbs, then differences in N400 amplitude across verb type could have resulted in spurious differences in the ERPs to infinitival markers. Alternatively, if the mean durations of the two verb classes were substantially different, then the late endogenous components (e.g. N400) elicited by verbs would have been desynchronous across sentence types, perhaps resulting in apparent differences in the ERP response to the infinitival markers even if, for example, N400 amplitude was equivalent across verb types. To examine these possibilities, we first measured the duration from the onset of each verb to the onset of the plosive indicating the beginning of the infinitival marker. The mean durations of the "hope" and "persuade" verbs were 637 (SE = 13 msec) and 644 (SE = 11 msec), respectively, a difference which did not approach significance [$t(29) = -0.551$. These durations suggest that the onset of the infinitival markers occurred after the usual N400 peak, and that the late endogenous components to verbs were reasonably synchronised across sentence types. Second, we examined the ERPs following the two verb types. Visual inspection indicated that the ERPs to "persuade" verbs were slightly more negative-going between about 300 and *500* msec at the posterior regions (the N400 component), but that differences between verb types resolved at about *500* msec. No noticeable differences in the waveforms were evident later in the epoch. We performed

The differences observed within the early window extended into the 300-500 msec window. Differences also began to appear at Pz within this window, where the ERPs to infinitival markers following "persuade" verbs were more positive-going than those following "hope" verbs. The midline ANOVAs revealed a significant interaction between sentence type and electrode site [$F(1,23) = 5.95$, $P < 0.01$], reflecting the larger differences between conditions at posterior sites. The ANOVA on data acquired at the lateral sites found significant interactions between sentence type and hemisphere [$F(1,23) = 16.85$, $P < 0.001$], sentence type and electrode site [$F(4,92) = 7.24$, $P < 0.01$], and between sentence type, hemisphere and electrode site [$F(4,92) = 5.91$, $P < 0.01$], again reflecting the greater positivity at the posterior and right hemisphere sites and greater negativity at the more anterior left hemisphere sites to infinitival markers following "persuade" verbs.

Similar effects were found between 500 and 800 msec. At the midline sites, differences between sentence types were again restricted to Pz [sentence type x electrode site: $F(2,46) = 8.08$, $P < 0.01$]. ERPs to the word to were more positive-going following "persuade" verbs than following "hope" verbs at the posterior regions in the left hemisphere and at the posterior and parietal sites in the right hemisphere. The reverse was true at the parietal and frontal regions in the left hemisphere [sentence type x hemisphere: $F(1,23) = 22.23$, $P < 0.001$; sentence type x electrode site: $F(4,92) = 7.92$, $P < 0.01$; sentence type x hemisphere x electrode site: $F(4,92) = 5.54$, $P < 0.01$].

Response to Auxiliary Verbs. Also of interest was the ERP response to auxiliary verbs in sentences like (3) and (4) in Table 1. The garden-path parsing model predicts that the auxiliary verb in sentences like (3) should elicit syntactic anomaly, while the same words in sentences like (4) should not. ERPs to the auxiliary verbs in both sentence types are shown in Fig. 4. As was the case with infinitival markers, ERPs to the auxiliary verbs predicted to produce syntactic anomaly were associated with a late positive-going wave, largest posteriorly and more widely distributed in th right than in the left hemisphere; however, the onset of the positivity appeared to be later following auxiliary verbs than following the anoma-

ANOVAs on data within two windows: 300-500 msec and 500-700 msec. The first window encompassed the N400 differences (confirmed by visual inspection). The second window encompassed the 100 msec of activity that acted as the pre-stimulus baseline for the infinitival markers, and roughly the first 50 msec of activity after presentation of the infinitival markers. No reliable differences were observed in any of the midline or lateral site ANOVAs (P > 0.2 in all analyses). Hence, it seems unlikely that differences in the ERPs to the infinitival marker were actually caused by differences in the ERPs to the preceding verbs.

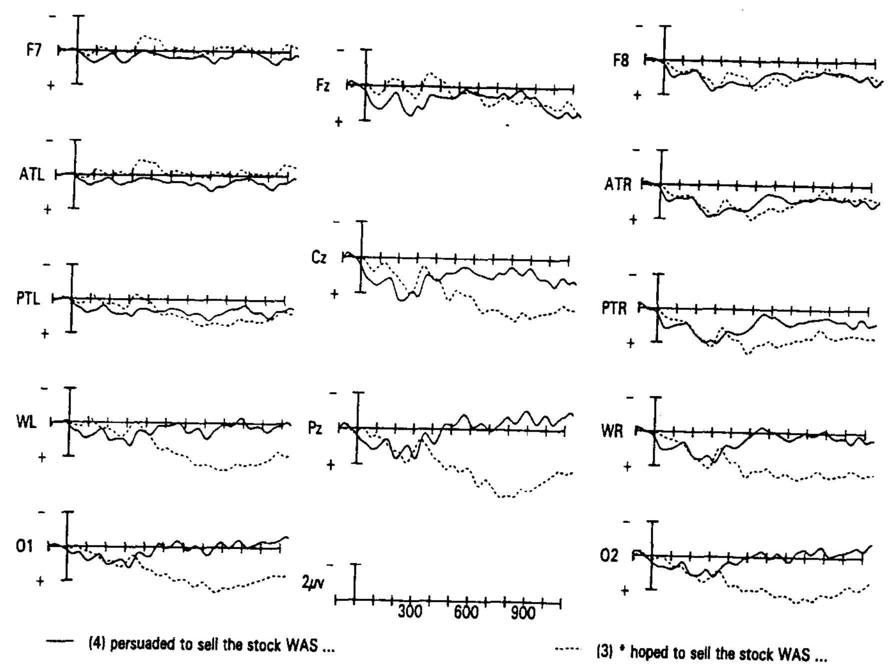

FIG. 4 Grand average ERPs to auxiliary verbs in sentences of type (3) and (4) in Table 1. In this and subsequent figures, each plot represents between 600 and 700 trials.

lous infinitival markers. Additionally, ERPs to auxiliary verbs in sentences like (3) were more negative-going at parietal and frontal left hemisphere sites, particularly between about 300 and 500 ms. This effect was also evident at Fz.

These observations were confirmed by statistical analyses. Although small differences between conditions were visible, no reliable effects were found in the 50-300 msec window. In the 300-500 msec window, ERPs to auxiliary verbs in sentences like (3) were more positive-going at Cz and Pz and more negative-going at Fz than ERPs to the same words in sentences like (4) [sentence type x electrode position: $F(2,46) = 5.51$, $P < 0.02$]. At the lateral sites, differences between conditions were larger in the right hemisphere than the left [sentence type x hemisphere: $F(1,23) = 5.37$, $P < 0.05$], and differences were larger at the posterior sites [sentence type x electrode site: $F(4,92) = 4.74$, $P < 0.02$]. However, the three-way interaction between sentence type, hemisphere and electrode site was not reliable [$F(4,92) = 1.26$], suggesting that ERPs to auxiliary verbs in "persuade'-verb sentences were not reliably more negative-going at the parietal and frontal sites in the left hemisphere than ERPs to "hope"-verb sentences.

Similar differences between conditions were observed within the 500800 msec window. At most electrode sites, ERPs to auxiliary verbs in sentences like (3) were more positive-going than those to auxiliary verbs in sentences like (4) [midline: $F(1,23) = 9.41$, $P < 0.01$; lateral: $F(1,23) = 10.20$, $P < 0.01$]. Differences were largest at the posterior regions [sentence type x electrode site, midline: $F(2,46) = 6.28$, $P < 0.02$; lateral: $F(4,92) = 8.02$, $P < 0.01$], and in the right hemisphere [sentence type x hemisphere: $F(1,23) = 4.69$, $P < 0.05$].

In summary, words which were predicted to elicit syntactic anomaly were associated with a late positive-going wave ("P600"), largest over the posterior regions and more widely distributed in the right than in the left hemisphere, and a negative-going wave over the parietal and anterior regions of the left hemisphere. This general result was consistent across both "anomalous" instances of to (i.e. those that indicated an apparent subcategorisation error) and "anomalous" auxiliary verbs (i.e. those that indicated an apparent phrase structure violation). Furthermore, these results are by and large consistent with the findings of previous studies that have involved the word-by-word visual presentation of sentences (Neville et al., 1991; Osterhout & Holcomb, 1992; Osterhout & Swinney, 1989). However, interesting differences existed in the response to the "anomalous" to and auxiliary verbs. The positivity associated with the "anomalous" to had an earlier onset at most electrode sites than that associated with the "anomalous" auxiliary verbs, although the P600 effect was larger for auxiliary verbs than for infinitival markers. Also, the left-hemisphere negativity to the "anomalous" words was more reliable and of longer duration following the word to than following auxiliary verbs.

Response to Sentence-ending Words. Osterhout and Holcomb (1992) reported an N400-like effect to final words in sentences typically judged to be unacceptable, relative to the response to final words in sentences typically judged to be acceptable. Figure 5 plots ERPs to sentence-ending words in short sentences [e.g. sentences (1) and (2) in Table 1]. Analyses were performed on ERPs within two time windows: 300-500 msec (capturing the area typically associated with the N400 effect) and 500-800 msec (since differences between conditions extended past 500 msec). Within the 300-500 msec window, ERPs to final words in sentences with "persuade" verbs were clearly more negative than those to final words in sentences with "hope" verbs [main effect for sentence type at midline sites: $F(1,23) = 5.77$, $P < 0.03$]. Additionally, differences between conditions were largest at the posterior regions [sentence type x electrode site, midline: $F(2,46) = 3.87$, $P < 0.05$; lateral: $F(4,92) = 4.03$, $P < 0.05$]. Although these differences extended into the remainder of the epoch, no reliable differences between conditions were

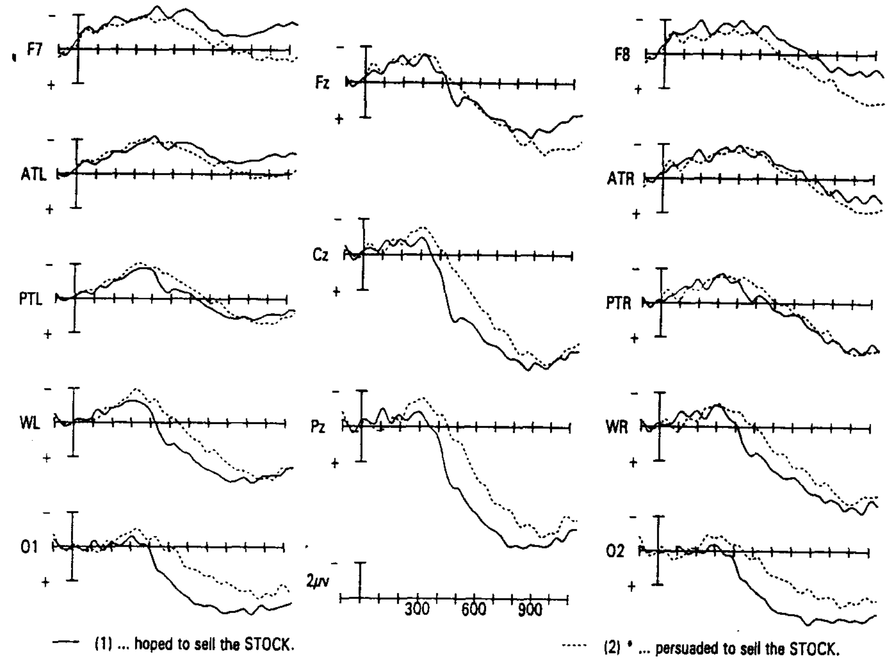

FIG. 5 Grand average ERPs to final words in short sentences, e.g. (1) and (2) in Table 1.

found at the midline sites in the 500-800 msec window. However, a significant interaction between sentence type and electrode site at the lateral sites [$F(4,92) = 3.36$, $P < 0.05$] reflected the posterior distribution of differences, with ERPs to final words in sentences with "persuade" verbs being more negative than those to final words in sentences with "hope" verbs.

ERPs to final words in the two types of long sentences [e.g. sentences (3) and (4) in Table 1] are shown in Fig. 6. ERPs to final words in sentences like (3) were more negative at word onset than those to the final words in sentences like (4), and differences between conditions extended throughout the epoch at most sites. Analyses were performed on mean amplitude within the 0-300, 300-500 and 500-800 msec windows. Differences between conditions in the early window were reliable [midline: $F(1,23) = 10.51$, $P < 0.01$; lateral: $F(1,23) = 4.15$, $P < 0.05$], and were larger at the posterior regions [sentence type x electrode site, midline: $F(2,46) = 3.68$, $P < 0.05$; lateral: $F(4,92) = 4.58$, $P < 0.05$].

Between 300 and 500 msec, ERPs to final words in sentences like (3) were more negative-going than those to final words in sentences like (4) at all midline sites [$F(1,23) = 5.77$, $P < 0.03$], and differences between conditions were larger at

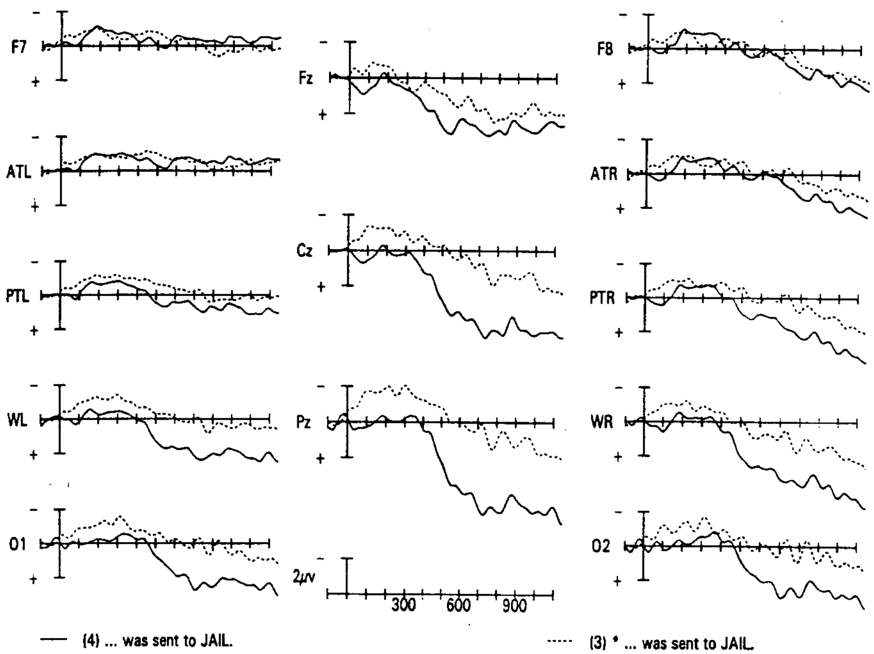

FIG. 6 Grand average ERPs to final words in long sentences, e.g. (3) and (4) in Table 1.

the posterior regions both at the midline and lateral sites [sentence type x electrode site, midline: $F(2,46) = 3.87$, $P < 0.06$; lateral: $F(4,92) = 4.03$, $P < 0.05$]. Reliable main effects for sentence type [midline: $F(1,23) = 17.05$, $P < 0.001$; lateral: $F(1,23) = 10.02$, $P < 0.01$] and reliable interactions between sentence type and electrode position [midline: $F(2,46) = 8.96$, $P < 0.001$; lateral: $F(4,92) = 11.27$, $P < 0.001$] were also found within the 500-800 msec window.

DISCUSSION

Words that were inconsistent with the putative "favoured" structural interpretation of spoken sentences elicited a positive-going wave, which for convenience we will refer to as the "P600 effect". This effect was similar (although not identical) to the ERP response to such words encountered during reading (Osterhout, 1990; Osterhout & Holcomb, 1992; Osterhout & Swinney, 1989). These results are consistent with the claim that the P600 is an electrophysiological marker of the syntactic "garden-path" effect, one which occurs both during reading and during auditory language comprehension.

This research appears to support two additional and relatively straightforward conclusions. First, the P600 effect was elicited by the same words in both modalities, suggesting that the comprehension system (under certain conditions) employs a similar parsing strategy regardless of the modality of the input. Specifically, these data are consistent with a parsing model in which a single syntactic structure is initially computed, even in the presence of syntactic ambiguity. In the cases examined here, the subjects appeared initially to pursue the simple active analysis over a passivised, reduced relative clause analysis. This finding is consistent with (but not exclusively supportive of) predictions that follow from the use of a minimal attachment strategy (cf. Frazier, 1987).

Second, the electrophysiological response to apparent syntactic anomaly occurs very rapidly during auditory language comprehension. The differentiation between well-formed and anomalous (or apparently anomalous) structures, as indexed by diverging waveforms, appears to begin within 50-300 msec in the case of apparent subcategorisation violations (i.e. the "anomalous" infinitival markers), and within 300-500 msec for apparent phrase structure violations (the "anomalous" auxiliary verbs). These findings suggest a close temporal link between the acoustic input on the one hand, and the output of the processes that compute syntactic structure on the other.[5,6]

---

[5]We should also note that the early onset of the P600 component to apparent subcategorisation violations suggests a rapid use of verb-specific information. Since the grammar allows structures of the form V-S', a sentence fragment such as *The broker persuaded to* would be perceived as syntactically anomalous only if the subcategorisation properties of the verb *persuade* had been consulted by the parser; the verb *persuade,* when used in its active form, does not accept complements beginning with the word *to*. The rapidity of P600 onset following presentation of the infinitival marker in such sentences brings into question recent claims that detailed verb-specific information is *not* consulted during the earliest moments of structural analysis (Ferreira & Henderson, 1990; Mitchell, 1989). Specifically, it has been suggested that the parser's initial decisions about sentence structure are influenced only by item-independent phrase structure knowledge; the parser is presumed later to modify these structures to conform to the constraints imposed by verb-specific knowledge. If such revisions were occurring during the passage of the sentences presented in this study, then they seemingly must have occurred within the temporal bounds suggested by the data reported here, i.e. within roughly 300 msec or so subsequent to the onset of the infinitival marker.

[6]A reviewer raised the possibility that the response to the "anomalous" infinitival markers could be interpreted as a *reprocessing* effect, rather than as a response to anomaly. Such an interpretation would be consistent with the claim that the parser delays its use of verb subcategorisation information. Specifically, when faced with the string *The broker hoped to* or *The broker persuaded to,* a minimal attachment parser that delays use of subcategorisation information would attach *to* as the beginning of a prepositional phrase, since the structure needed for a PP is simpler than that needed for an infinitival clause. The garden-path effect would then occur at the verb *sell* for both sentences, leading to a reprocessing of both sentences. Reprocessing could involve verb subcategorisation information. The effects

The basic similarity of the "auditory" P600 to the P600 elicited by visual stimuli (reported in Osterhout & Holcomb, 1992) suggests that these components belong to a common family of endogenous components. Nonetheless, there were important modality-related differences. The "auditory" P600 appeared to have an earlier onset than its visual counterpart. Furthermore, there are slight differences in the scalp distribution of the P600 across modalities. For both the auditory and the visual P600, the effect tended to be largest over the midline and right hemisphere sites. However, differences between conditions tended to be more restricted over the posterior scalp in the auditory study than in the visual study. Furthermore, the left hemisphere negativity observed following syntactically anomalous words was somewhat more pronounced and of longer duration with auditory sentences than with visual sentences. There were also differences in the positive-going wave elicited by infinitival markers and auxiliary verbs in the present study. The positivity elicited by syntactically anomalous auxiliary verbs had a somewhat later onset and was much larger in amplitude than that elicited by the syntactically anomalous infinitival markers.[7] Additionally,

---

observed following the infinitival marker to could reflect more costly reprocessing for the reduced relative analysis in the *persuade* sentences, relative to the simple active analysis in the *hope* sentences. However, we believe that the effects we observed following the infinitival markers are in fact elicited by the infinitival markers themselves, and not by the verbs presented subsequent to the infinitivals, for two reasons. First, similar effects were observed in the visual analogue of the current study (Osterhout & Holcomb, 1992), under conditions in which the infinitival markers were unambiguously responsible for the P600 effect. (The onset of the P600 occurred before the onset of the verb following the infinitival marker.) Second, the onset of the P600 following infinitival markers in the current study was extremely rapid (between 50 and 300 msec), and preceded onset of the verb subsequent to the infinitival on most trials (see footnote 3).

[7]Although it is tempting to interpret the more rapid onset of the P600 following to relative to the auxiliary verbs as reflecting some processing factor (i.e. apparent violations of subcategorisation are recognised more rapidly than apparent violations of phrase structure),. several caveats prevent us from doing this. Most importantly, as we noted earlier, word-initial boundaries were much more easily identified for to than for the auxiliary verbs; hence, the observed differences in P600 onset might be related to greater temporal jitter in the temporal locking of the ERP to the word-initial sound of the auxiliary verbs, relative to the lesser jitter in the case of to. Similarly, one might attempt to give a processing explanation to the observation that the "time of P600 onset" values of between 50 and 300 msec were shorter in the auditory study than during the word-by-word reading study reported by Osterhout and Holcomb (1992), in which the onset rested around 500 msec. The difference in onset might reflect differences in the rate at which words are encountered. In the present study, subjects encountered roughly five syllables per second; in the visual studies, words have been presented at slower rates ranging from 500 to 650 msec per word. However, this modality difference must be interpreted with caution, for two reasons. First, the relevant comparisons involve contrasts between different groups of subjects. Second, given the effects of coarticulation and related phenomena, "word onset" is much more ambiguous for words presented in continuous speech than for visually presented words.

differences in the response to the infinitival markers across sentence types showed a clear reversal in polarity across the anterior and posterior regions; that is, the response was positive-going over the posterior regions and negative-going over the anterior regions, relative to the non-anomalous control condition. This polarity inversion was not as noticeable in the response to auxiliary verbs.

At present, we cannot identify with confidence the functional or cognitive significance of these differences across modality and anomaly type, primarily because we do not know enough about the cognitive events underlying these ERP components. One possibility is that the positivities observed here and in Osterhout and Holcomb (1992) are members of the P300 family of waves. The P300 (or "P3b") is often observed following unexpected, task-relevant stimuli (Donchin, 1979, 1981; Johnson, 1989; Ritter & Vaughan, 1969; Ruchkin et al., 1990). If one assumes that readers anticipate grammaticality when reading or listening to sentences, then "perceived ungrammaticality" might act as an "unexpected event", thereby eliciting a P300 component. Alternatively, the positivities observed here might have been elicited in response to encountering a word from an unexpected grammatical category, e.g. a word from a category that typically does not follow a specific verb. Additional grounds for positing that the P600 is a member of the P300 family derive from the fact that the syntactically anomalous words in these sentences provided information relevant to the subjects' acceptability judgements following each sentence. Ruchkin et al. (1990) have shown that P300 amplitude can be strongly influenced by the amount of information that is delivered by the stimulus, even when this information is not related to expectancy. Specifically, stimuli that provide full information concerning the outcome of a trial (and hence determine the correct response to the trial) elicit larger P300s than stimuli that do not. In the present study, the syntactically anomalous words obviously provided subjects with information concerning the "outcome" of the current trial, by bearing on the acceptability judgements of the subjects. A third reason for suspecting that the P600 is a member of the P300 family concerns the modality-related differences in scalp distribution. In a study contrasting P300s elicited by visual and auditory stimuli, Johnson (1989) reports that the P300 elicited by visual stimuli was larger over the frontal and central scalp than the P300 elicited by auditory stimuli, a modality-dependent shift in distribution that is fully consistent with modality-related differences described above.

These observations, however, by no means conclusively demonstrate that the P600 is another manifestation of the P300 family. Indeed, a recent study by Hagoort, Brown and Groothusen (this issue) observed positivities, similar to those reported here, in response to syntactic violations when no explicit judgement of sentence acceptability or grammaticality was required of the subjects. Since the P300 is generally elicited by task-relevant stimuli, and since the influence of "outcome information" can only be relevant if a judgement task is involved, this finding suggests that the "P600 effect" might in fact be distinct from the P300 family of positivities.

A related question is whether or not the positivities elicited in this study and in the study reported by Osterhout and Holcomb (1992) represent a single ERP response to a unitary cognitive process or event. In particular, although the P600 to anomalous auxiliary verbs showed the classic posterior-maximal scalp distribution associated with the P300 (P3b) component, the P600 effect to the anomalous infinitival markers had the classic "slow wave" characteristic of being positive-going posteriorly and negative-going anteriorly (cf. Johnson, 1989). It has been argued that the P300 and slow-wave positivity reflect distinct cognitive and neural events (Johnson, 1989; Ruchkin et al., 1990).

The questions of whether or not the P600 is "just another" P300, and whether or not the P600s observed here are instances of a single set of cognitive/neural processes, are of obvious importance. But we believe that the utility of the P600 for examining language processes does not critically hinge on these questions (cf. Donchin & Coles, 1988). Use of the P600 as an electrophysiological marker of syntactic anomaly becomes feasible once the P600 is shown to *co-vary systematically* with manipulations that are presumed to influence syntactic well-formedness. Hence, if the P600 is shown to occur reliably whenever comprehenders are presumed to encounter syntactic anomaly, then the P600 is a useful tool *even in the absence of certain knowledge concerning the cognitive events underlying the component* (cf. Osterhout & Holcomb, 1992).

As in several earlier studies in which sentences were presented visually (Osterhout, 1990; Osterhout & Holcomb, 1990; 1992), ERPs to final words in sentences typically judged to be unacceptable were more negative-going (most notably between about 300 and 800 msec) than those to the final words in sentences typically judged to be acceptable. This negativity (particularly that associated with final words in the short "unacceptable" sentences) was similar in its temporal and distributional characteristics to the N400 effect elicited by semantically inappropriate words in both visually (Kutas & Hillyard, 1980a; 1980b; 1980c) and auditorily (Holcomb & Neville, 1991; McCallum, Farmer, & Pocock, 1984) presented sentences. Osterhout and Holcomb (1992) suggest that uninterpretable sentences, in general, elicit an N400-like effect, regardless of the cause of uninterpretability (e.g. a semantically or syntactically anomalous word). The present findings appear at first glance to support this claim. However, the rapid rate at which words appeared in the present study might have resulted in a temporal overlap of the ERP components of interest. This is particularly true in the case of the negative-going activity elicited by the

sentence-final words in the long "unacceptable" sentences. Differences in ERPs to final words in the two types of "long" sentences [e.g. sentences (3) and (4) in Table 1] were apparent extremely early, i.e. at about 0 msec after onset of the word. One possible explanation for this early onset implies that the differences in ERPs to the sentence-final words in the long sentences are partially or fully spurious. Specifically, the positive-going activity (P600) elicited by the anomalous auxiliary verbs may not have fully resolved prior to the 100 msec of activity preceding the sentence-final words. That is, the waveform in sentences with anomalous auxiliary verbs might have been distinctly negative-going just prior to the onset of the sentence-final words, as the positive-going activity returned to the amplitude existing prior to the anomaly; this negative-going resolution would not exist in sentences with non-anomalous auxiliary verbs, because these verbs did not elicit large positive-going waves. Since the 100 msec of activity preceding the final words was used as a baseline to align activity across conditions, any inequalities between conditions within the 100 msec baseline window (e.g. a negative-going waveform in one condition but not in the other) would introduce spurious differences in the epoch of interest. In the present case, such a discrepancy would result in "negative-going" activity in the unacceptable sentences with an "onset" of 0 msec. This issue raises the general problem of "overlapping" ERP components. In general, as the temporal distance between stimuli of interest decreases (as is inevitable in studies presenting spoken or written sentences at "natural" rates of presentation), the difficulty in disentangling overlapping components increases.

Perhaps most importantly, these results point to the promise of ERPs as tools for investigating auditory language comprehension, and for examining aspects of comprehension that have been difficult to explore with other methods. As just one example, consider the influence of intonation on sentence comprehension. A plethora of research on speech production has shown that the grouping of speech elements into units (e.g. phrases and clauses) is conveyed by certain intonational cues. For example, the syllable preceding a major syntactic boundary is often characterised by a falling fundamental frequency contour, a lengthening in duration and a reduction in amplitude, and is generally followed by a measurable pre-boundary pause (Cooper & Sorensen, 1977; Klatt, 1975; Scott, 1982). Although there is evidence that these cues are sometimes used in the eventual interpretation of sentences (e.g. Scott, 1982; Streeter, 1978), there have been few investigations of how these cues are employed during the on-line processing of spoken sentences, or of the interaction between these cues and other informational sources (e.g. structural representations and knowledge) (for one recent exception, see Beach, 1991). It is currently unclear, for example, whether the parser continues to rely on a structurally based strategy (e.g. minimal attachment) even when intonational cues to sentence structure are available. The results of the present study, having provided evidence that ERPs are sensitive to temporal and functionally relevant aspects of the syntactic analysis of spoken sentences, allow one to speculate that such issues will soon be successfully addressed.

Manuscript received December 1991
Revised manuscript received March 1993

REFERENCES

Beach, C.M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language, 30,* 644 -663.

Clifton, C., Speer, S., & Abney, S. (1991). Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. *Journal of Memory and Language, 30,* 251-271.

Cooper, W.E., & Paccia-Cooper, J. (1980). *Syntax and speech.* Cambridge, MA: MIT Press.

Cooper, W.E., & Sorensen, J.M. (1977). Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America, 61,* 1314-1320.

Cooper, W.E., & Sorensen, J.M. (1981). *Fundamental frequency in sentence production.* New York: Springer-Verlag.

Donchin, E. (1979). Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), *Evoked brain potentials and behavior.* New York: Plenum Press.

Donchin, E. (1981). Surprise! ... Surprise? *Psychophysiology, 18,* 493-513.

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11,* 343-427.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Ferreira, F., & Henderson, J. (1990). Use of verb information during syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. ***Journal*** *of Experimental Psychology: Learning, Memory and Cognition, 16,* 555-568.

Frazier, L. (1987). Sentence processing: A tutorial. In M. Coltheart (Ed.), *Attention and performance XII.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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.

Geisser, S., & Greenhouse, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95-112.

Gorrell, P. (1989). Establishing the loci of serial and parallel effects in syntactic processing. *Journal of Psycholinguistic Research, 18,* 61-71.

Hillyard, S.A., & Picton, T.W. (1987). Electrophysiology of cognition. In F. Plum (Ed.), *Handbook of physiology. Section 1: Neurophysiology.* New York: American Physiological Society.

Holcomb, P.J., & Neville, H. (1991). Speech analysis during language processing: An eventrelated brain potential study. *Psychobiology, 19,* 286-300.

Holmes, V.M. (1987). Syntactic parsing: In search of the garden-path. In M. Coltheart (Ed.), *Attention and performance X11.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc. Holmes, V.M., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement-verb sentences. *Journal of Memory and Language, 28,* 668-689.

Huynh, H., & Feldt, L.S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1,* 69-82.

Jasper, H.H. (1958). Report to the committee on methods of clinical examination in electroencephalography. Appendix: The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10,* 371-375.

Johnson, R. Jr. (1989). Auditory and visual P300s in temporal lobectomy patients: Evidence for modality dependent generators. *Psychophysiology, 26,* 651-667.

Klatt, D.H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics, 3,* 129-140.

Kucera, H., & Francis. W.N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

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

Kutas, M., & Hillyard, S.A. (1980b). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11,* 99-116.

Kutas, M., & Hillyard, S.A. (1980c). Reading between the lines: Event-related brain potentials during natural sentence processing. *Brain and Language, 11,* 354-373.

Kutas, M., & Van Petten, C. (1988). Event-related potential studies of language. In P.K. Ackles, J.R. Jennings, & M.G.H. Coles (Eds), *Advances in psychophysiology.* Greenwich, CT: JAI Press.

Lenneberg, E.H. (1967). *Biological foundations of language.* New York: John Wiley. Marcus, M.P. (1980). *A theory of syntactic recognition for natural language.* Cambridge, MA: MIT Press.

Marslen-Wilson, W.D. (1975). Sentence perception as an interactive parallel process. Sci*ence, 189,* 226-228.

McCallum, W.C., Farmer, S.F., & Pocock, P.V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology,* 59, 477-488.

Mitchell, D. (1989). Verb-guidance and other lexical effects in parsing. *Language and Cognitive Processes, 4,* 123-154.

Mitchell, D., & Holmes, V.M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguities. *Journal of Memory and Language,* 5, 542-559.

Neville, H., 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.

Osterhout, L. (1990). *Event-related brain potentials elicited during sentence comprehension.* Unpublished doctoral dissertation, Tufts University, Medford, MA.

Osterhout, L., & Holcomb, P. (1990). Event-related potentials elicited by grammatical anomalies. In C.H.M. Brunia, A.W.K. Gaillard, & A. Kok (Eds), *Psychophysiological brain research.* Tilburg: Tilburg University Press.

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

Osterhout, L., & Swinney, D.A. (1989). On the role of the simplicity heuristic in language processing: Evidence from structural and inferential processing. *Journal of Psycholinguistic Research, 18,* 553-562.

Radford, A. (1988). *Transformational grammar.* Cambridge: Cambridge University Press.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 358-374.

Ritter, W., & Vaughan, H.G. Jr. (1969). Average evoked responses in vigilance and discrimination. *Science, 164,* 326-328.

Ruchkin, D.S., Johnson, R. Jr, Canoune, H.L., Ritter, W., & Hammer, M. (1990). Multiple sources of P3b associated with different types of information. *Psychophysiology, 27,* 157-176.

Scott, D. (1982). Duration as a cue to perception of phrase boundary. *Journal of the Acoustical Society of America, 71,* 996-1007.

Streeter, L.A. (1978). Acoustic determinants of phrase boundary perception. *Journal of the Acoustical Society of America, 64,* 1582-1592.