On the Language Specificity of the Brain Response to Syntactic Anomalies: Is the Syntactic Positive Shift a Member of the P300 Family?

Lee Osterhout, Richard McKinnon, Michael Bersick, and Vicka Corey
University of Washington

Abstract

■ Event-related brain potentials (ERPs) were recorded from 13 scalp electrodes while subjects read sentences, some of which contained either a verb that disagreed in number with the subject noun (syntactic anomaly) or a word in uppercase letters (physical anomaly). Uppercase words elicited the P300 complex of positivities, whereas agreement violations elicited a late positive shift with an onset around 500 msec and a duration of several hundred msec. These effects differed in their morphology, temporal course, amplitude, and scalp distribution. Furthermore, manipulations of the probability-of-occurrence and task relevance of the anomalies had robust effects on the response to uppercase words, but not on the response to agreement violations. Finally, these anomalies had additive effects when agreement-violating uppercase (doubly anomalous) words were presented. These results are taken to be an initial indication that the positive shift elicited by agreement violations is distinct from the P300 response to unexpected, task-relevant anomalies that do not involve the violation of a grammatical rule. ■

INTRODUCTION

Contemporary formal theories of language structure distinguish among several levels of representation—phonological, syntactic, semantic, pragmatic, etc. (cf. Chomsky, 1981, 1986). One fundamental claim of such models is that each representational level is characterized by its own set of grammatical rules, and that the rules at each level can be specified independently; for example, that the rules of syntax can be specified independently of the rules of semantics. Another claim is that grammatical knowledge is uniquely linguistic and does not directly interact with nonlinguistic knowledge (e.g., the speaker's "real-world" knowledge; cf. Chomsky, 1986). Whether or not such descriptions accurately apply to the processes underlying *language comprehension* remains a matter for debate. A standard assumption underlying much recent psycholinguistic work is that a relatively direct mapping exists between the levels of representation posited within linguistic theories and the cognitive and neural processes underlying comprehension (Bock & Kroch, 1989; Fodor, 1983; Forster, 1979). Distinct language-specific processes are thought to interpret a sentence at each level of analysis, and distinct representations are thought to result from these computations. However, other theorists, most notably those working in the neural net modeling domain, deny that such language-specific processes and representations exist (Ades & Steedman, 1982; Elman, 1990; MacWhinney, Bates, & Kliegel, 1984).

One means for addressing this issue involves the recording of event-related brain potentials (ERPs) elicited during comprehension. ERPs are patterned voltage changes in the ongoing electroencephalogram that are time-locked to the onset of a sensory, cognitive, or motor event (Garnsey, 1993; Hillyard & Picton, 1987). Scalp-recorded ERPs consist of a series of positive and negative voltage peaks (or "components") that is distributed over time. ERP components are multidimensional; they can vary in latency, amplitude, polarity, and scalp distribution. Furthermore, certain late-occurring ERPs (i.e., those with a latency of greater than 100 msec) appear to be highly sensitive to specific changes in cognitive state (e.g., attentional state; Hillyard, Münte, & Neville, 1985; Hillyard & Picton, 1987). Assuming that cognitively distinct processes are mediated by neurally distinct brain systems, evidence that events occurring at different linguistic levels elicit distinct patterns of brain response could be construed as evidence that separable processes exist (Osterhout & Holcomb, 1992). Furthermore, evidence that language-sensitive ERP effects are in fact *language specific* (i.e., sensitive only to language-related manipulations) would count as evidence that language-specific processes exist.

Prior work has demonstrated that certain variations in the ERP are sensitive to language-related events. In a series of seminal experiments, Kutas and Hillyard (1980a, 1980b, 1980c) discovered that semantically inappropriate words (e.g., "John buttered his bread with *socks*") elicit an enhanced negative-going wave with a peak amplitude at about 400 msec (the *N400 effect;* for a review, see Kutas & Van Petten, 1988). Recent studies examining the ERP response to syntactic violations have reported a variety of robust effects (Friederici, Pfeifer, & Hahne, 1993; Hagoort, Brown, & Groothusen, 1993; McKinnon & Osterhout, 1995; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout, 1990; Osterhout & Holcomb, 1992, 1993; Osterhout, Holcomb, & Swinney, 1994; Osterhout & Mobley, 1995; Rösler, Friederici, Pütz, & Hahn, 1993). Critically, none of these effects resembles the N400 effect. Of particular relevance to the current paper is the finding that a disparate set of syntactic violations, including violations of constraints on phrase structure (e.g., "The scientist criticized Max's *of* proof the theorem"), verb subcategorization ("The broker persuaded *to* sell the stock"), subject–verb agreement ("The elected officials *hopes* to succeed"), and reflexive pronoun-antecedent agreement ("The successful woman congratulated *himself*"), elicits a large-amplitude positive-going wave, variously labeled the *P600* (Osterhout & Holcomb, 1992) and the *syntactic positive shift* (Hagoort et al., 1993). Typically, this positive wave has a centroposterior distribution, an onset around 500 msec, and a duration of several hundred msec (for a review, see Osterhout, 1994, or Osterhout & Holcomb, 1995).

These findings indicate that ERPs are highly sensitive to the syntactic and semantic well-formedness of the sentence, and that the response to a variety of syntactic anomalies is quite distinct from the response to semantic/pragmatic anomalies.[1] However, the fundamental question of whether or not these language-sensitive ERP effects are in any sense language specific remains unanswered. This question has been asked recently with respect to the N400 (Deacon, Breton, Ritter, & Vaughan, 1991; Polich, 1985). For example, researchers have attempted to determine whether the N400 is a longer-latency manifestation of a more generally sensitive component, the N2 (cf. Deacon et al., 1991). The question of whether the syntactic positive shift is uniquely related to language processing, and in particular the processes underlying the syntactic analysis of sentences, has not yet been systematically addressed. This question is particulary compelling in that a wide variety of evidence seems to indicate that separable syntactic processes exist. For example, psycholinguistic evidence indicates that syntactic analysis is largely uninfluenced by nonsyntactic types of information (e.g., semantic or pragmatic contextual information; Ferreira & Clifton, 1986; Frazier & Rayner, 1982), and neuropsychological evidence from brain-damaged populations has been taken to indicate that such damage can lead to highly selective syntactic deficits (Grodzinsky, 1986; Schwartz, Saffran, & Marin, 1980). The question, then, is whether the syntactic positive shift is uniquely a manifestation of these putative specialized processes.

We addressed this question by investigating a salient alternative possibility, namely, the possibility that the syntactic positive shift is a member of the P300 family of late positive components elicited by a wide variety of (linguistic and nonlinguistic) events. The P300 is a centroparietal-maximal positivity of variable onset that is elicited by attended, task-relevant stimuli (for reviews, see Donchin, 1981; Hillyard & Picton, 1987; Pritchard, 1981). P300 amplitude is a function of the probability, task relevance, and informational content of the eliciting stimulus (Ruchkin, Johnson, Canoune, Ritter, & Hammer, 1990), with the most improbable, task-relevant, and informative stimuli eliciting the largest-amplitude P300s. The peak latency of the P300 varies as a function of stimulus complexity and ranges from about 300 to 800 msec (Fabiani, Gratton, Karis, & Donchin, 1987). Although early descriptions of the P300 treated the effect as a monolithic component reflecting a single neural source (e.g., Sutton, Braren, Zubin, & John, 1965), more recent work has indicated that a variety of positivities combine to form a "late positive complex" (Sutton & Ruchkin, 1984). These positivities include the frontocentral P3a (Squires, Squires, & Hillyard, 1975), the centroparietal P3b (the classic P300), and a longer latency, longer duration late positive slow wave (Donchin, Ritter, & McCallum, 1978; Ruchkin, Munson, & Sutton, 1982; Squires et al., 1975). Evidence exists indicating that these effects are generated by distinct neural sources (Johnson, 1989, 1993; Ruchkin et al., 1990). However, all three effects have been shown to be sensitive to the probability-of-occurrence and informational content of the stimulus, and at least the P3b and slow wave appear to be sensitive to the task relevance of the eliciting stimulus (see, e.g., Donchin et al., 1978; Ruchkin et al., 1990; Squires et al., 1975). Furthermore, in the vast majority of experiments the most prominent component of the ERP elicited by unexpected, task-relevant stimuli is the centroparietal P3b (Donchin, 1979; Donchin et al., 1978).[2]

The observation that the P300 is not a unitary phenomenon generated by a single neural source complicates the question of whether the syntactic positive shift is another instantiation of the P300. For example, it could be that the positive shift is a manifestation of the neural/cognitive events underlying just one of the subcomponents of the P300 complex, or of one or more of these subcomponents plus other effects that do not contribute to the P300. This complexity is increased even more by recent evidence that the subcomponents of the P300 complex might themselves be composites of activity in an indeterminate number of neural generators. In particular, recent work suggests that the P3b reflects the simultaneous activation of at least several independent neural systems (Johnson, 1993). Hence, it is conceivable that some but not all of the systems that contribute to the P3b also contribute to the syntactic positive shift. One implication is that the question of whether the positivity to syntactic violations is a P300 is no longer sufficiently precise. A better way to phrase the question might be to ask whether the brain response to anomalies that involve formal, rule-governed aspects of language is partially or entirely distinct from the response to unexpected, task-relevant anomalies that do not involve the grammar. Henceforth, we use the terms "syntactic positive shift" and "P300" as convenient labels for the ERP responses to these two categories of anomaly.

The distinctiveness of these two brain responses can be assessed in several ways. First, at least a rough estimate of the similarity in the neural events underlying the responses is provided by comparing the component characteristics of these effects, e.g., their morphology, temporal characteristics, and scalp distributions. In particular, it is generally agreed that ERP components with distinct scalp distributions are necessarily generated by distinct brain systems (cf. Johnson, 1993). Correspondingly, evidence that the syntactic positive shift and the P300 have distinct scalp topographies would be consistent with the claim that the two effects are (at least partially) neurally distinct.[3] Although both of these effects tend to be largest in amplitude over centroposterior regions, a careful examination of this issue requires quantitative analysis of data acquired from a single set of subjects within one experiment. To date, no such experiment has been reported.

Second, one can determine whether the effects of stimulus and task manipulations known to modulate the P300 similarly affect the positive wave elicited by syntactic violations. If so, then one might conclude that these effects are functionally related. Based on the available evidence, one could reasonably infer that the amplitude of the syntactic positive shift, much like P300 amplitude, is a function of the probability and task relevance of the eliciting stimulus. For example, given that the vast majority of sentences encountered during reading are fully grammatical, syntactic violations are very likely perceived to be unexpected or rare events. Furthermore, in many of the experiments reporting a late positive shift to syntactic violations subjects were asked to make a "sentence-acceptability" judgment on each trial and to indicate their judgment by pushing one of two buttons ("acceptable" and "not acceptable") after each sentence (Osterhout & Holcomb, 1992, 1993; Osterhout et al., 1994). In these experiments the syntactic anomalies were therefore both directly task relevant and highly informative with respect to the outcome of the trial. However, the results of two recent experiments have been taken as indications that the syntactic positive shift might be functionally distinct from the P300. Hagoort et al. (1993) presented sentences containing violations of phrase structure, verb subcategorization, and subject-verb number agreement. Because subjects were asked to simply read the sentences and did not perform any secondary task, the anomalies were not as directly task relevant as they were in other studies in which subjects made sentence acceptability judgments. Nonetheless, phrase structure and agreement violations elicited a positive shift. Unfortunately, this evidence is suggestive but not conclusive, since several researchers have observed a P300-like wave to deviant attended stimuli that were not directly task relevant (Courchesne, Courchesne, & Hillyard, 1978; Squires, Donchin, Squires, & Grossberg, 1977). One might also argue that the syntactic anomalies were implicitly relevant to the task of understanding the sentence. A second type of evidence is reported by Osterhout et al. (1994). These researchers manipulated the "severity" of syntactic anomalies encountered during sentence processing by contrasting the ERP response to violations of verb subcategorization and subcategorization biases. Outright violations of subcategorization result in an irrevocable ungrammaticality, whereas violations of biases simply require structural reanalysis of the sentence. Correspondingly, the amplitude of the positive shift was much larger in response to violations of subcategorization compared to violations of subcategorization biases. The fact that the amplitude of the positivity varied as a function of syntactic variables could be construed as indicating that the positivity manifests syntactic processes. However, these results can also be explained under a P300 account, e.g., violations of verb subcategorization biases are more frequent in text and speech than are outright violations of verb subcategorization constraints. At present, then, there is very little evidence that would allow one to determine whether or not the syntactic positive shift is functionally distinct from the P300 family.

A third method for assessing the distinctiveness of the P300 and the syntactic positive shift is to determine whether these components have additive effects. This can be accomplished by presenting stimuli that are expected to elicit both components simultaneously and comparing the response to the "doubly anomalous" words to the response to each type of anomaly in isolation from the other. This approach follows from Helmholtz's Rule of Superposition, which maintains that electrical fields propagating through a conductive medium summate where they intersect. Evidence of additivity in such situations strongly implies independence of the underlying neural sources (cf. Kutas & Hillyard, 1980a).[4]

The purpose of the current study was to begin to investigate the question of whether the ERP response to syntactic violations is distinct from that elicited by unexpected, task-relevant anomalies that do not involve a violation of a grammatical rule. We did this by comparing the component characteristics of these two responses (Experiments 1-3), evaluating their relative sensitivity to manipulations of task relevance and probability (Experiments 1 and 2), and determining whether the responses are additive (Experiment 3). We should be clear about what is at stake in these experiments. Evidence that the positive wave elicited by syntactic anomalies is related to the P300 family would *not* imply that the syntactic positive shift can no longer be viewed as a useful tool for examining sentence comprehension (cf. Garnsey, 1993; Osterhout & Holcomb, 1992, 1993; Osterhout et al., 1994). The sensitivity of this effect to syntactic manipulations exists regardless of whether it is a member of the P300 family. Nonetheless, evidence that the syntactic positive shift is "just another" P300 would indicate that it is not a direct manifestation of the syntactic processes themselves but rather a manifestation of domain-general processes that are perhaps correlated with but indeterminately removed from these processes. Conversely, evidence that the syntactic positive shift is distinct from the P300 family would not in and of itself conclusively demonstrate that it is in any sense language-specific; however, such evidence would eliminate a salient alternative to the language specificity hypothesis.

EXPERIMENT 1

In Experiment 1, we presented three types of sentences, as exemplified by sentences (1)-(3) in Table 1: nonanomalous controls, sentences containing a verb that disagreed in number with the subject noun, and sentences that contained a word in uppercase letters. Prior work has shown that physically anomalous words elicit a "P300" (Kutas & Hillyard, 1980a), whereas agreement violations elicit a "syntactic positive shift" (Hagoort et al., 1993; Osterhout & Mobley, 1995). One goal was to directly compare the component characteristics of these two responses. A second goal was to examine the sensitivity of these responses to the task relevance and informational content of the anomaly. To this end, half of the subjects made sentence-acceptability judgments about each sentence and the remaining subjects passively read the sentences for comprehension. In the sentence-acceptability condition, the anomalies were explicitly and directly task relevant and completely informative with respect to trial outcome; in the passive reading condition, the anomalies were less directly task relevant and contained no information about trial outcome.

Table 1. Examples of Sentences Presented during Experiments 1-3

(1) The doctors *believe* the patient will recover.
(nonanomalous control)

(2) The doctors *believes* the patient will recover.
(agreement violation)

(3) The doctors *BELIEVE* the patient will recover.
(uppercase word)

(4) The doctors *BELIEVES* the patient will recover.
(doubly anomalous word)

Results and Discussion

Sentence-Acceptability Judgments

Subjects in the sentence-acceptability condition judged nonanomalous sentences, sentences containing agreement violations, and sentences containing an uppercase word to be acceptable on 90, 6, and 10% of the trials, respectively, $F(2,22) = 169, p < 0.0001$.

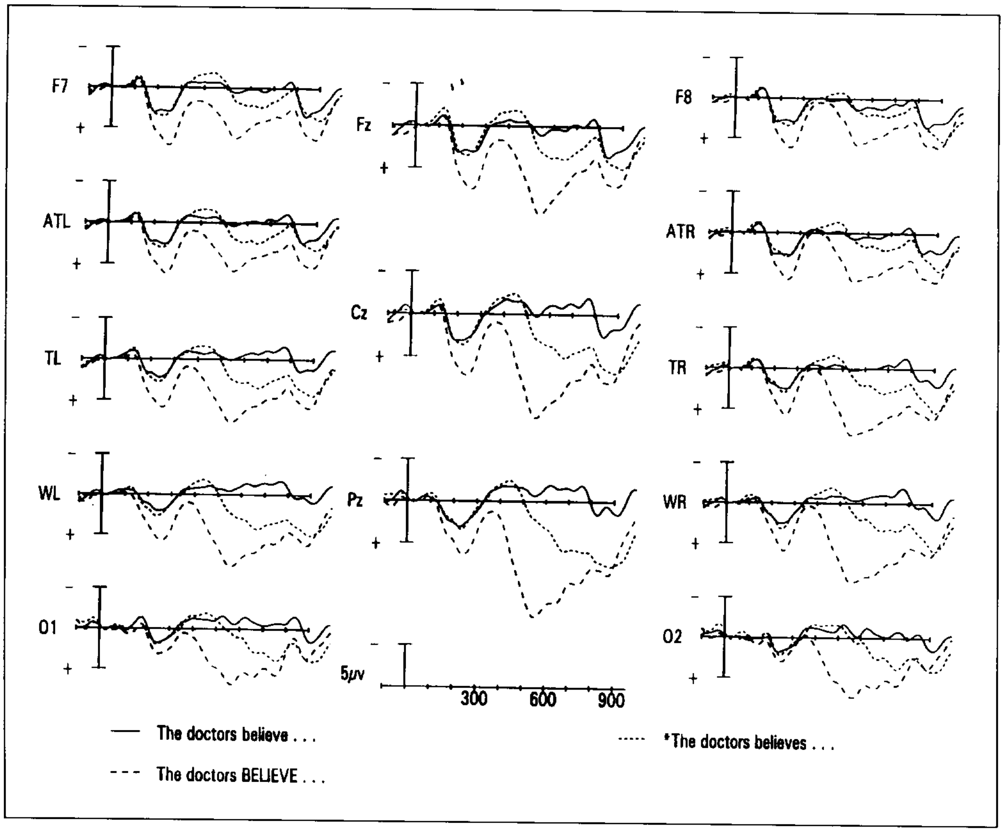
Event-Related Potentials

Sentence-Acceptability Judgment Condition. ERPs to critical words in the three sentence types, averaged over subjects and items, are shown in Figure 1. (Seventeen percent of the trials, approximately evenly distributed across conditions, were rejected for artifact.) In this and subsequent figures, the general shapes of the waveforms were consistent with those reported in other experiments involving language stimuli (cf. Neville, Kutas, Chesney, & Schmidt, 1985; Osterhout & Holcomb, 1992). A clear negative-positive complex was visible in the first 300 msec after word onset. The negative component ("N1") was largest anteriorly and peaked at about 120 msec. The positive component ("P2/P3a") peaked at about 250 msec; the amplitude of this component varied across conditions. A negative-going deflection with a peak at about 400 msec (N400) was evident in the response to most words.

Effects of anomaly type were clearly evident. Uppercase words elicited a late positive complex quite similar to previous reports of the P300 complex, relative to the nonanomalous control words: a frontocentral positivity peaking between 200 and 300 msec (P2/P3a), a large, centroparietal positivity with a peak at about 500 msec (P3b), and a subsequent gradual return to baseline (slow wave). ERPs to agreement-violating words also differed from those to nonanomalous controls; however, the response to these words was clearly distinct from the response to the uppercase words. Agreement violations elicited a slightly larger N400 component at some sites followed by a large-amplitude late positive shift with a posterior distribution. This wave had an onset around 500 msec and a duration of several hundred msec and was highly similar to previous reports of the syntactic positive shift (Osterhout & Holcomb, 1992; Osterhout et al., 1994; Osterhout & Mobley, 1995).

Statistical analyses confirmed these observations. Within the 200-300 msec window, the ANOVAs revealed a main effect for sentence type [midline, $F(2,22) = 25.28$, $p < 0.001$; lateral, $F(2,22) = 20.89$, $p < 0.001$], reflecting the large-amplitude P2/P3a elicited by uppercase words. This effect was most pronounced over anterior sites [sentence type × electrode site: $F(8,88) = 5.46$,

Figure 1. Grand average ERPs over three midline and 10 lateral sites to nonanomalous (solid line), agreement-violating (small dashes), and uppercase (large dashes) words in the sentence-acceptability judgment condition, Experiment 1. Each hashmark represents 100 msec. Positive voltage is plotted down.



$p < 0.05$]. Simple effects analyses found that ERPs to uppercase words were more positive-going than those to the nonanomalous words [midline, $F(1,11) = 38.74$, $p < 0.001$; lateral, $F(1,11) = 34.05$, $p < 0.001$] and agreement-violating words [midline, $F(1,11) = 38.18$, $p < 0.001$; lateral, $F(1,11) = 28.35$, $p < 0.01$]. No reliable differences were found between agreement-violating words and nonanomalous words, $F < 1$ in all analyses.

Reliable differences between sentence types were also found between 300 and 500 msec [midline, $F(2,22) = 32.82$, $p < 0.001$; lateral, $F(2,22) = 22.88$, $p < 0.001$]. These differences were primarily due to the onset of the P3b component in the uppercase condition. Additionally, an interaction between sentence type, hemisphere, and electrode site reflected the fact that differences in the amplitude of the P2/P3a component across conditions resolved prior to onset of the P3b component in the right but not the left hemisphere, $F(8,88) = 4.10$, $p < 0.05$. Simple effects analyses revealed reliable differences between uppercase words and both nonanomalous controls [midline, $F(1,11) = 42.31$, $p < 0.001$; lateral, $F(1,11) = 29.73$, $p < 0.001$] and agreement violations [midline, $F(1,11) = 52.97$, $p < 0.0001$; lateral, $F(1,11) = 36.97$, $p < 0.001$]. However, despite a slight increase in N400 amplitude at anterior and temporal sites to the agreement violations relative to the control condition, no reliable differences were found between these two conditions, $F < 1$ in all analyses.

Between 500 and 800 msec, ERPs to both types of anomaly noticeably differed from those to controls [main effect of sentence type: midline, $F(2,22) = 27.49$, $p < 0.01$; lateral, $F(2,22) = 32.08$, $p < 0.001$], and these differences were larger at posterior sites [sentence type × electrode site: midline, $F(2,22) = 8.17$, $p < 0.01$; lateral, $F(8,88) = 8.47$, $p < 0.01$]. Furthermore, ERPs elicited by the critical words in the three sentence types also differed in scalp distribution within this window [sentence type × electrode site × hemisphere: $F(8,88) = 3.17$, $p < 0.05$]. Simple effects analyses found reliable differences between the uppercase and control conditions [midline, $F(1,11) = 43.56$, $p < 0.001$; lateral, $F(1,11) = 63.49$, $p < 0.001$] and between the agreement violations and controls [midline, $F(1,11) = 14.03$, $p < 0.01$; lateral, $F(1,11) = 10.74$, $p < 0.01$]. These differences were posteriorly distributed, both in the uppercase [midline, $F(2,22) = 25.98$, $p < 0.001$; lateral, $F(4,44) = 16.82$, $p < 0.01$] and in the agreement-violating conditions [midline, $F(2,22) = 5.01$, $p < 0.05$; lateral, $F(4,44) = 5.62$, $p < 0.05$]. Of particular interest were simple effects analyses on mean amplitude between 500 and 800 msec involving direct comparisons between the uppercase words and agreement violations. ERPs to uppercase words were more positive-going than those to agreement violations [midline, $F(1,11) = 15.37$, $p < 0.01$; lateral, $F(1,11) = 22.00$, $p < 0.001$]. Furthermore, the scalp distributions of these two effects within the 500–800 msec window were not identical; the response to agreement violations was slightly larger over the left than the right hemisphere throughout this window, whereas the response to uppercase words was larger over the right than left hemisphere, particularly at posterior sites [sentence type × hemisphere × electrode site: raw data, $F(4,44) = 5.77, p < 0.01$; normalized data, $F(4,44) = 4.50, p < 0.05$].

The latencies of the most positive-going peaks elicited by the two types of anomaly were computed for data acquired over midline sites, by determining the most positive-going point between 400 and 1000 msec. Peak latency was reliably later occurring in the agreement condition, $F(1,11) = 21.38, p < 0.001$.

Thus, although both the uppercase words and agreement violations elicited a large positive wave, these positivities had distinct morphologies, onsets, peak latencies, amplitudes, and, at least between 500 and 800 msec, distinct scalp distributions. However, comparisons of scalp distribution are complicated by the fact that the onsets and peak amplitudes of these effects did not temporally overlap. An alternative means for comparing scalp distributions involves computing mean averages over different time windows for the two effects, thereby capturing the onset, peak amplitude, and extent of each effect. To this end, we performed ANOVAs on mean amplitude between 400 and 800 msec for uppercase words and between 500 and 900 msec for agreement violations. These analyses revealed a reliable three-way interaction between sentence type, hemisphere, and electrode site [raw data, $F(4,44) = 4.95, p < 0.01$; normalized data, $F(4,44) = 4.19, p < 0.05$].

In summary, uppercase words elicited a complex of positivities quite similar to previous reports of the P300 complex: a frontocentral P2/P3a, a centroparietal P3b, and a late positive slow wave. By contrast, agreement violations elicited a posterior-maximal late positive shift beginning at about 500 msec. Importantly, the positivities elicited by the two anomalies had distinct scalp distributions over lateral electrode sites. The evidence from Experiment 1, then, indicates that the large positivities elicited by uppercase words and agreement violations are distinguishable in terms of morphology, onset, peak latency, amplitude, and scalp distribution.

Passive Reading Condition. ERPs elicited by the critical words in the passive reading condition are shown in Figure 2. (Approximately 17% of the trials were rejected for artifact.) As in the acceptability-judgment condition, uppercase words elicited an increase in the P2/P3a component, particularly at anterior sites. This was followed by a larger-amplitude wave at most sites (P3b). Agreement violations again elicited a late positive wave with an onset around 500 msec. However, these effects were smaller in amplitude than those observed in the sentence-acceptability judgment condition.

ANOVAs on mean amplitude within the 200–300 msec window revealed a main effect for sentence type [midline, $F(2,22) = 7.25$, $p < 0.01$; lateral: $F(2,22) =$

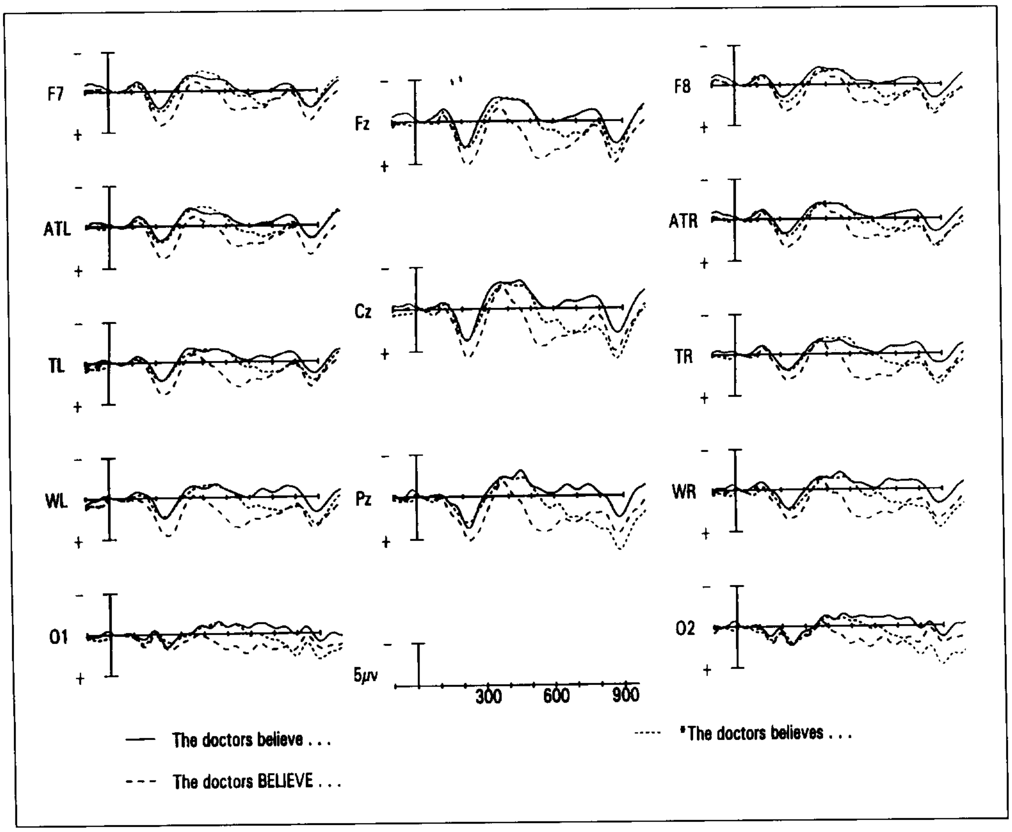


Figure 2. Grand average ERPs to nonanomalous (solid line), agreement-violating (small dashes), and uppercase (large dashes) words in the passive reading condition, Experiment 1.

5.98, $p < 0.01$]. These differences were slightly larger over the left than the right hemisphere, $F(2,22) = 3.51$, $p < 0.05$. Simple effects analyses again found reliable differences between the uppercase words and both the nonanomalous [midline, $F(1,11) = 10.88, p < 0.01$; lateral, $F(1,11) = 9.20, p < 0.02$] and agreement-violating conditions [midline, $F(1,11) = 7.99$, $p < 0.02$; lateral, $F(1,11) = 8.69, p < 0.05$]. No reliable differences were found between agreement-violating and control sentences. Reliable differences between sentence types were also found between 300 and 500 msec [midline, $F(2,22) = 3.91$, $p < 0.04$; lateral, $F(2,22) = 5.00$, $p < 0.02$]. Again, ERPs to uppercase words were more positive-going than those to nonanomalous [midline, $F(1,11) = 4.23$, $p = 0.06$; lateral, $F(1,11) = 5.34$, $p < 0.05$] and agreement-violating words [midline, $F(1,11) = 4.96, p < 0.05$; lateral, $F(1,11) = 7.13, p < 0.03$], reflecting the onset of the P3b component in the uppercase condition.

Differences between sentence types were also found within the 500–800 msec window [midline, $F(2,22) = 10.32$, $p < 0.001$; lateral, $F(2,22) = 11.04$, $p < 0.001$]. ERPs to uppercase words [midline, $F(1,11) = 23.40, p < 0.01$; lateral, $F(1,11) = 23.86, p < 0.001$] and agreement violations [midline, $F(1,11) = 9.61$, $p < 0.05$; lateral, $F(1,11) = 6.15, p < 0.04$] were more positive-going than those to nonanomalous controls. However, ERPs to uppercase words and agreement-violating words differed reliably only at lateral sites, $F(1,11) = 4.67, p = 0.01$. By contrast to the results from the sentence-acceptibility task condition, the distribution of ERPs to agreement violations and uppercase words did not differ reliably within this window.

Additional analyses were performed on mean amplitude between 500 and 900 msec in the agreement condition and between 400 and 800 msec in the uppercase condition. In contrast to the sentence-acceptability judgment condition, ERPs to the two anomaly types did not reliably differ in amplitude or distribution.

Cross-Group Analyses. To statistically evaluate the effects of task on ERPs to the anomalies, we computed difference waves by subtracting ERPs elicited by the critical words in control sentences from those elicited by each type of anomaly, in each task condition.[5] The resulting difference waves to uppercase words and agreement violations in each task condition are shown in Figures 3 and 4, respectively. These figures indicate that the amplitude of the positivity to both types of anomaly was greater in the sentence-acceptability judgment condition than in the passive reading condition, but that the effects of task were much more noticeable for uppercase words than for agreement violations. ERPs were quantified as mean amplitude within the 400–800 msec window for uppercase words and between 500 and 900 msec for the agreement violations. These windows included the onset and peak amplitude of the positive-going activity in each anomaly condition. ANOVAs with task condition as a between-subjects factor revealed main effects for anomaly type [midline, $F(1,22) = 10.30$,

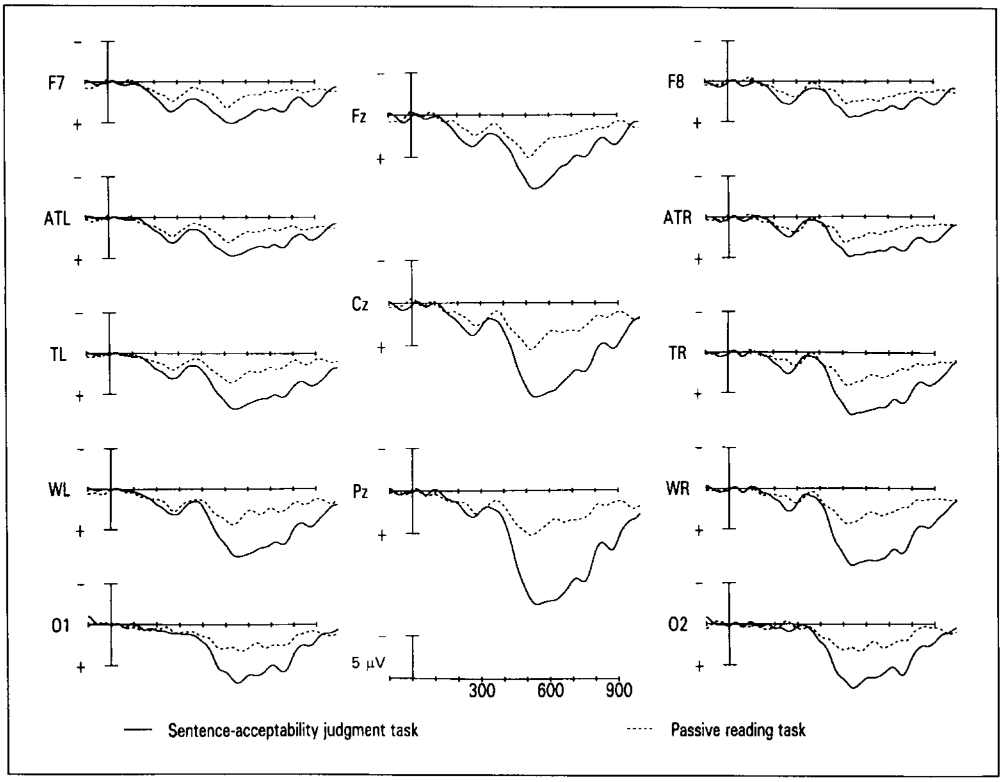


Figure 3. Difference waves formed by subtracting ERPs to nonanomalous controls from those to uppercase words in the sentence-acceptability judgment task (solid line) and passive reading (dashed line) conditions.

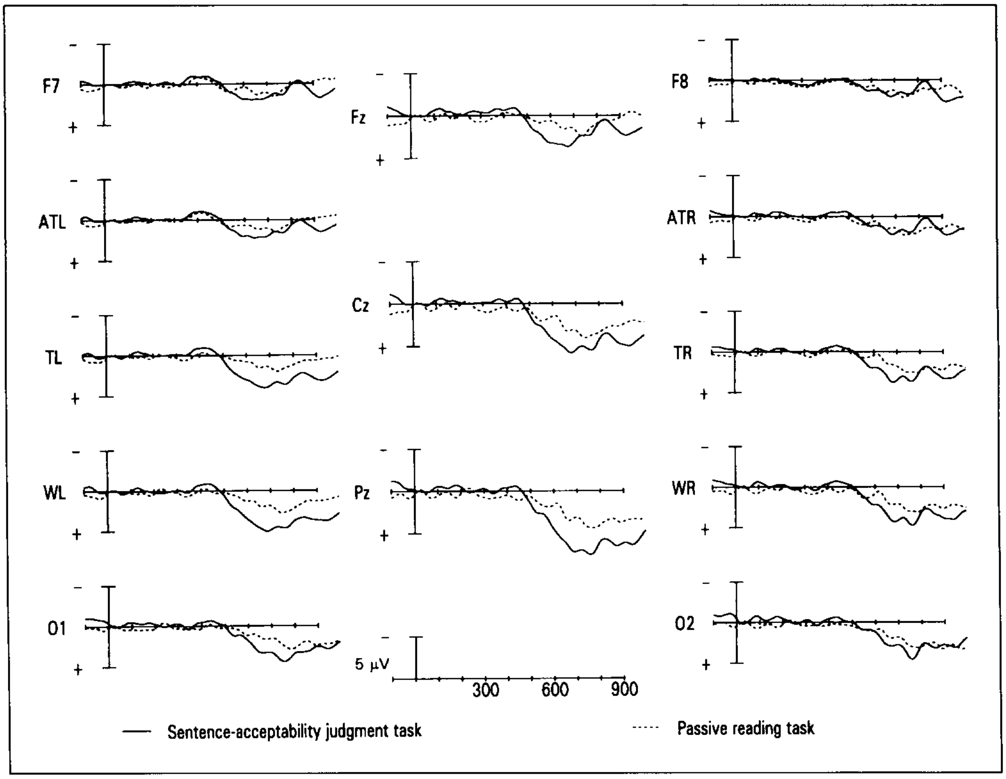


Figure 4. Difference waves formed by subtracting ERPs to nonanomalous controls from those to agreement violations in the sentence-acceptability judgment task (solid line) and passive reading (dashed line) conditions.

$p < 0.01$; lateral, $F(1,22) = 14.39, p < 0.01$] and task [midline: $F(1,22) = 5.92, p < 0.05$], and an interaction between anomaly type and task [midline, $F(1,22) = 7.00$, $p < 0.02$; lateral, $F(1,22) = 5.78, p < 0.03$]. Simple effects analyses revealed that the task effects were statistically reliable for uppercase words [midline, $F(1,22) = 6.56$, $p < 0.05$; lateral, $F(1,22) = 4.40, p < 0.05$] but not for agreement violations [midline, $F(1,22) < 1$; lateral, $F(1,22) < 1$). Furthermore, the scalp distributions of the two positivities were clearly distinct. ERPs to uppercase words were more negative-going over posterior portions of the right hemisphere, whereas those to agreement violations were slightly larger over left hemisphere sites [sentence type × hemisphere × electrode site: raw data, $F(4,88) = 7.81, p < 0.001$; normalized data, $F(4,88) = 4.98, p < 0.01$], and these differences were much larger in the sentence-acceptability judgment condition than in the passive reading condition [task × sentence type × hemisphere × electrode site: raw data, $F(4,88) = 4.56$, $p < 0.05$; normalized data, $F(4,88) = 3.04, p < 0.02$].

These results indicate that uppercase words and agreement violations elicit positive waves even when these anomalies are not directly task relevant and do not provide definitive information regarding the outcome of the trial (cf. Hagoort et al., 1993). Both effects were larger in amplitude in the sentence-acceptability judgment condition than in the passive reading condition. However, the change in task had more robust and reliable effects on the response to uppercase words than on the response to agreement violations.

EXPERIMENT 2

Previous experiments have shown that P300 amplitude is an inverse function of the probability of the eliciting event. In Experiment 2, we directly compared the effects of manipulating the probability of uppercase and agreement-violating words. Two lists were presented. In one list, 20% of the sentences contained an agreement violation whereas 60% of the sentences contained an uppercase word (the "20% agreement/60% uppercase" condition). In the second list, these percentages were reversed (the "60% agreement/20% uppercase" condition). The critical question was whether or not the positive wave elicited by agreement violations would prove to be sensitive to manipulations of probability-of-occurrence.

Results and Discussion

Sentence-Acceptability Judgments

Subjects judged the nonanomalous, agreement violation, and uppercase word sentences to be acceptable on 90, 10, and 6% of the trials, $F(2,22) = 850$.

Event-Related Potentials

Comparisons across Sentence Types. Figures 5 and 6 plot ERPs elicited by nonanomalous, uppercase, and agreement-violating words in the "60% agreement/20% uppercase" and in the "20% agreement/60% uppercase"

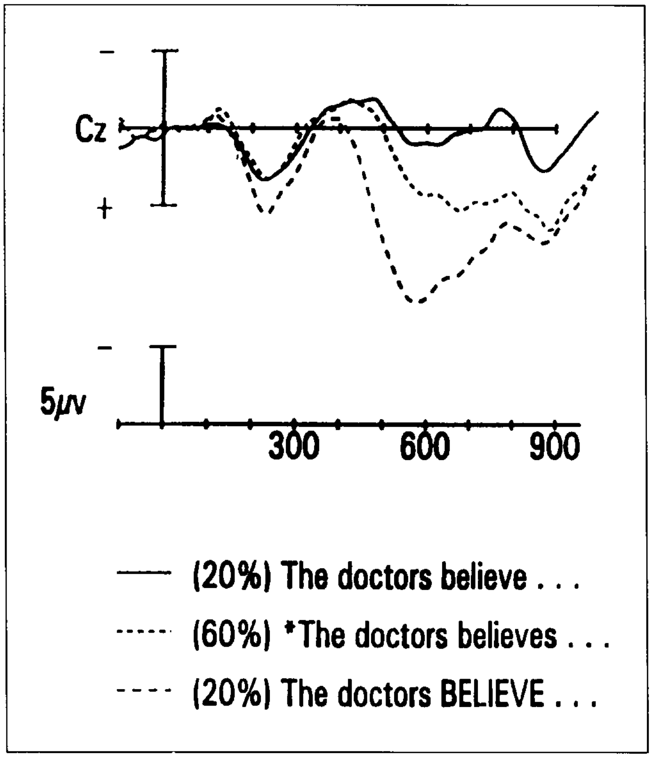


Figure 5. Grand average ERPs recorded from Cz to nonanomalous (solid line), agreement-violating (small dashes), and uppercase (large dashes) words in the "60% agreement/20% uppercase words" condition, Experiment 2.

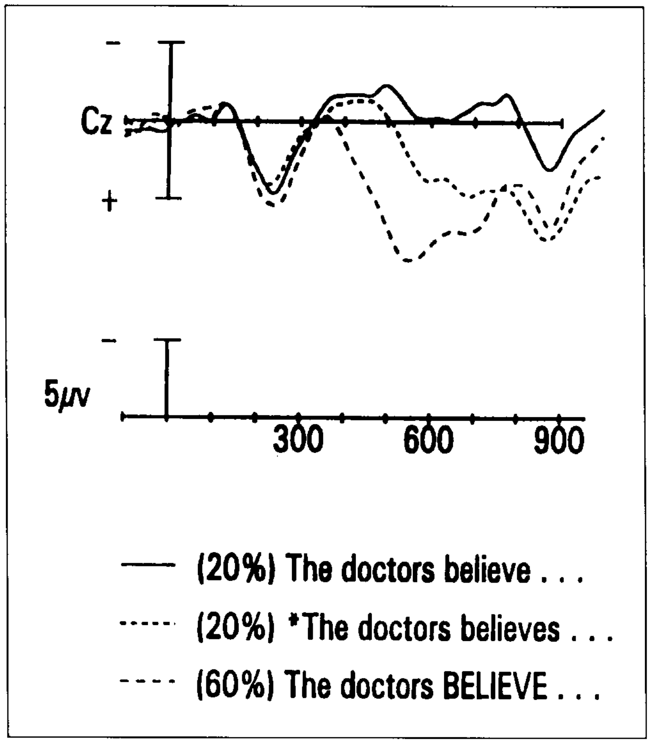


Figure 6. Grand average ERPs recorded from Cz to nonanomalous (solid line), agreement-violating (small dashes), and uppercase (large dashes) words in the "20% agreement/60% uppercase words" condition, Experiment 2.

conditions, respectively. In both conditions, uppercase words elicited the P300 complex (frontocentral P3a, centroparietal P3b with a peak at about 500 msec, late slow wave) observed in Experiment 1, and agreement violations elicited a positive shift with a posterior distribution and an onset around 500 msec that was quite similar to previous reports of the syntactic positive shift.

ANOVAs were performed treating list condition as a within-subjects factor. Within the 200-300 msec window, a main effect for sentence type [midline, $F(2,22) = 6.10$, $p < 0.01$; lateral, $F(2,22) = 7.90, p < 0.01$] and a reliable interaction between sentence type and electrode site (midline, $F(4,44) = 2.77$, $p < 0.05$; lateral, $F(8,88) = 7.20$, $p < 0.0001$) reflected the fact that the P2/P3a component to uppercase words was larger in amplitude than that elicited in the other two sentence type conditions, particularly at anterior sites. Furthermore, these differences were marginally larger in the "60% agreement/20% uppercase" condition than in the "20% agreement/60% uppercase" condition [list × sentence type × electrode: $F(8,88) = 1.98$, $p < 0.1$]. The reliable main effect for sentence type within the 300-500 msec window [midline: $F(2,22) = 43.53$, $p < 0.001$; lateral: $F(2,22) = 18.71, p < 0.001$] reflected the onset of the positive wave (P300) in the uppercase word condition.

As in Experiment 1, the largest differences between sentence types were found within the 500-800 msec window [main effect for sentence type: midline, $F(2,22) = 100.62$, $p < 0.0001$; lateral, $F(2,22) = 51.99$, $p < 0.0001$], and these differences were larger over posterior sites [sentence type × electrode site: midline, $F(4,44) = 6.63$, $p < 0.05$; lateral, $F(8,88) = 3.35$, $p < 0.06$]. Also as in Experiment 1, ERPs to uppercase words were slightly larger over posterior right hemisphere sites whereas ERPs to agreement violations were slightly larger in the left hemisphere, but this interaction was reliable only when the Geisser-Greenhouse correction was not applied [sentence type × hemisphere × electrode site: $F(8,88) = 2.48, p < 0.05$]. These differences in scalp topography were larger in the "60% agreement/20% uppercase" condition than in the "20% agreement/60% uppercase" condition, but again this effect was reliable only when the Geisser-Greenhouse correction was not applied [list × sentence type × hemisphere × electrode site, $F(8,88) = 2.63, p < 0.05$].

Comparisons across Probability Conditions. Figures 7 and 8 plot ERPs to the uppercase and agreement-violating words, respectively, in the two probability conditions. As expected, uppercase words elicited a larger positive wave when the proportion of such words in the list was relatively small. In contrast, changes in the proportion of agreement violations had much smaller effects on the amplitude of the positive wave elicited by the violations,

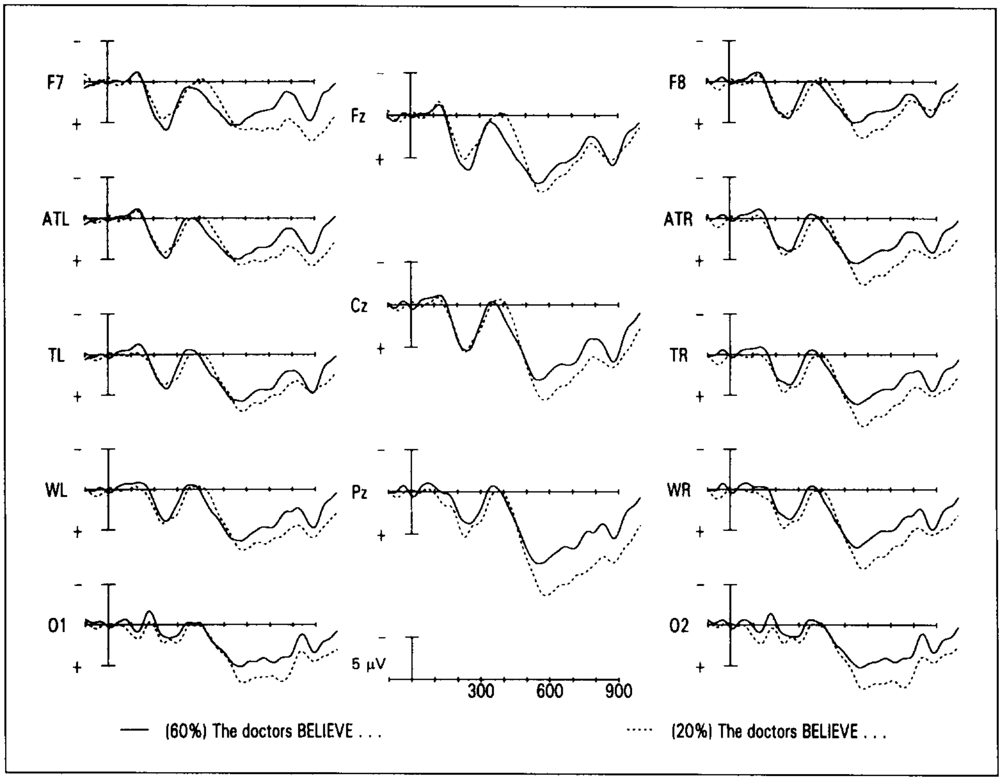


Figure 7. Grand average ERPs to uppercase words when they comprised 60% (solid line) and 20% (dashed line) of the trials, Experiment 2.

and at some sites (e.g., Pz) the positivity was larger when the proportion of agreement violations was relatively high. These effects were quantified by computing mean amplitude between 400 and 800 msec (uppercase condition) and 500 and 900 msec (agreement condition). These windows encompassed the onsets and peaks of the late positivities. An ANOVA with repeated measures on probability condition, sentence type, and electrode site revealed that mean amplitude was a function of sentence type, $F(1,11) = 14.43, p < 0.01$, and that the responses to the anomalous words were differentially affected by manipulations to the probability of the anomalies [probability × sentence type × electrode site: $F(2,22) = 4.45, p < 0.05$].

Separate analyses were performed on data from each anomaly condition. For uppercase words, ERPs in the low probability condition were larger over midline sites, particularly at posterior sites [probability condition × electrode site: $F(2,22) = 4.96, p < 0.05$]. Reliable differences were observed over the right [$F(1,11) = 4.11, p = 0.06$] but not over the left hemisphere [$F(1,11) = 1.20, p > 0.2$]. For the agreement condition, the difference in amplitude of the late positive shift across violating and control sentences was not reliable, $p > 0.2$ in all analyses.

Thus, the results of Experiment 2 indicate that although the amplitude of the positive-going response to uppercase words was a function of the probability of encountering an uppercase word within the list, the response to agreement violations was not reliably affected by the probability of encountering an agreement violation.

EXPERIMENT 3

In Experiment 3 we attempted to determine whether the positivities elicited by uppercase and agreement-violating words have additive effects. Subjects read sentences similar to those presented in Experiments 1 and 2, plus a fourth sentence type containing a word that was both agreement-violating and in uppercase letters (see Table 1). In the simplest case of complete independence, the "doubly anomalous" words might be expected to elicit a response that approximates a linear summation of the responses to the two types of anomaly when they occur in isolation (cf. Kutas & Hillyard, 1980a). However, this prediction is complicated by recent evidence that the P3b component reflects the summed activity of an indeterminate number of independent neural sources, each of which responds to a specific (and unique) aspect of the stimulus (e.g., the probability, task relevance, and informational content of the stimulus; Johnson, 1989, 1993; Ruchkin et al., 1990). If the responses to uppercase words and agreement violations share any of these independent sources, then the response to the doubly anomalous words will be less than the linear summation of the two responses in isolation.

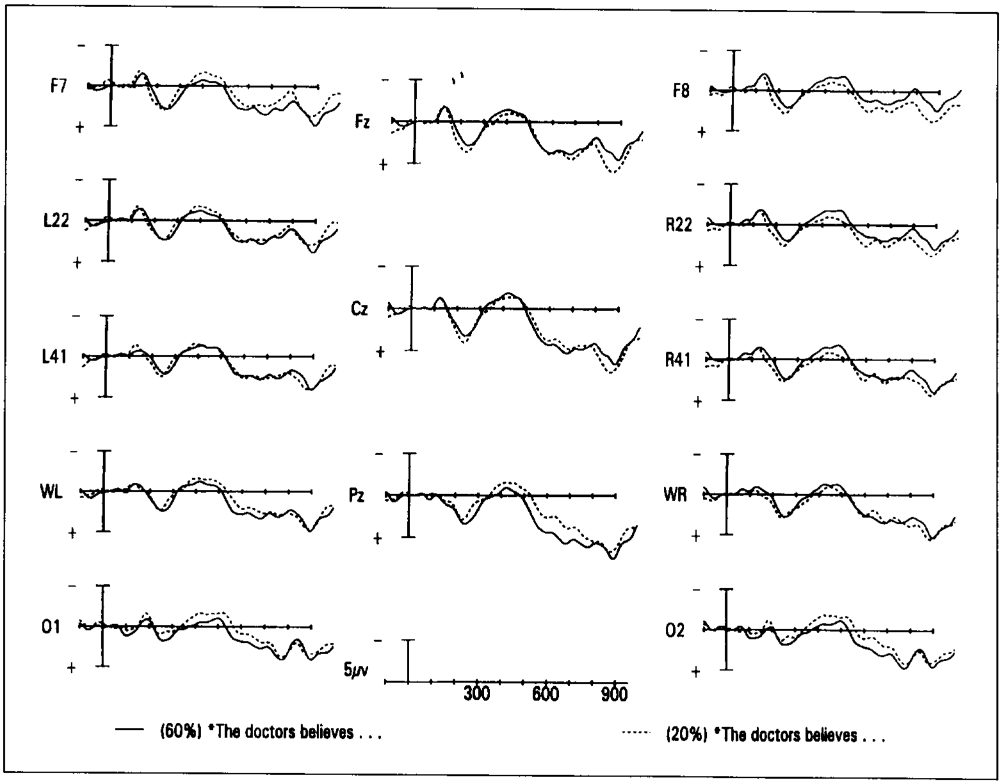


Figure 8. Grand average ERPs to agreement violations when they comprised 60% (solid line) and 20% (dashed line) of the trials, Experiment 2.

Results and Discussion

Sentence-Acceptability Judgments

Sentences in the nonanomalous, agreement-violating, uppercase, and doubly anomalous conditions were judged to be acceptable on 93, 14, 5, and 1% of the trials, respectively, $F(3,45) = 421$.

Event-Related Potentials

Grand average ERPs to the critical words in the four sentence types are shown in Figure 9. (Thirteen percent of the trials were rejected due to artifact.) As in Experiments 1 and 2, uppercase words elicited an increase in the P2/P3a component, a posterior-maximal P3b, and a late slow wave. Agreement violations elicited a positive shift with an onset around 500 msec. ERPs elicited by doubly anomalous words overplotted those to uppercase words from word onset to roughly 500 msec subsequent to word onset. At about 500 msec, ERPs to the doubly anomalous words diverged from those to uppercase words by becoming more positive-going, and remained so for the remainder of the recording epoch at most sites.

ANOVAs performed on mean amplitude within the 200–300 msec window revealed a main effect for sentence type [midline, $F(3,45) = 25$, $p < 0.0001$; lateral, $F(3,45) = 9.51$, $p < 0.001$], reflecting the increase in P2/P3a amplitude in the uppercase and doubly anomalous conditions. This effect was larger anteriorly [sentence type × electrode site: midline, $F(6,90) = 4.09$, $p < 0.01$; lateral, $F(3,45) = 3.76$, $p < 0.05$] and in the left hemisphere [sentence type × hemisphere: $F(12,180) = 13.75$, $p < 0.001$]. Simple effects analyses on midline data indicated that ERPs to both uppercase words and doubly anomalous words were reliably more positive-going than those to nonanomalous controls [uppercase: $F(1,15) = 34.98$, $p < 0.0001$; doubly anomalous: $F(1,15) = 42.22$, $p < 0.0001$] and agreement violations [uppercase: $F(1,15) = 26.09$, $p < 0.001$; doubly anomalous: $F(1,15) = 36.36$, $p < 0.001$]. No reliable differences were found between agreement violations and controls, or between uppercase and doubly anomalous words. Analyses on mean amplitude within the 300-500 msec window again revealed a main effect for sentence type [midline, $F(3,45) = 46.70$, $p < 0.001$; lateral, $F(3,45) = 33.50$, $p < 0.0001$], indicating the onset of the P3b component in the uppercase and doubly anomalous conditions. Differences between conditions were largest over posterior portions of the right hemisphere [sentence type × hemisphere × electrode site: $F(12, 180) = 3.27$, $p < 0.01$]. As in the earlier window, simple effects analyses indicated that both uppercase and doubly anomalous words were reliably more positive-going in this region than nonanomalous controls [uppercase: $F(1,15) = 75.52$, $p < 0.0001$; doubly anomalous: $F(1,15) = 85.93$, $p < 0.0001$] and agreement violations [uppercase: $F(1,15) = 54.60$, $p < 0.0001$; agreement violations:

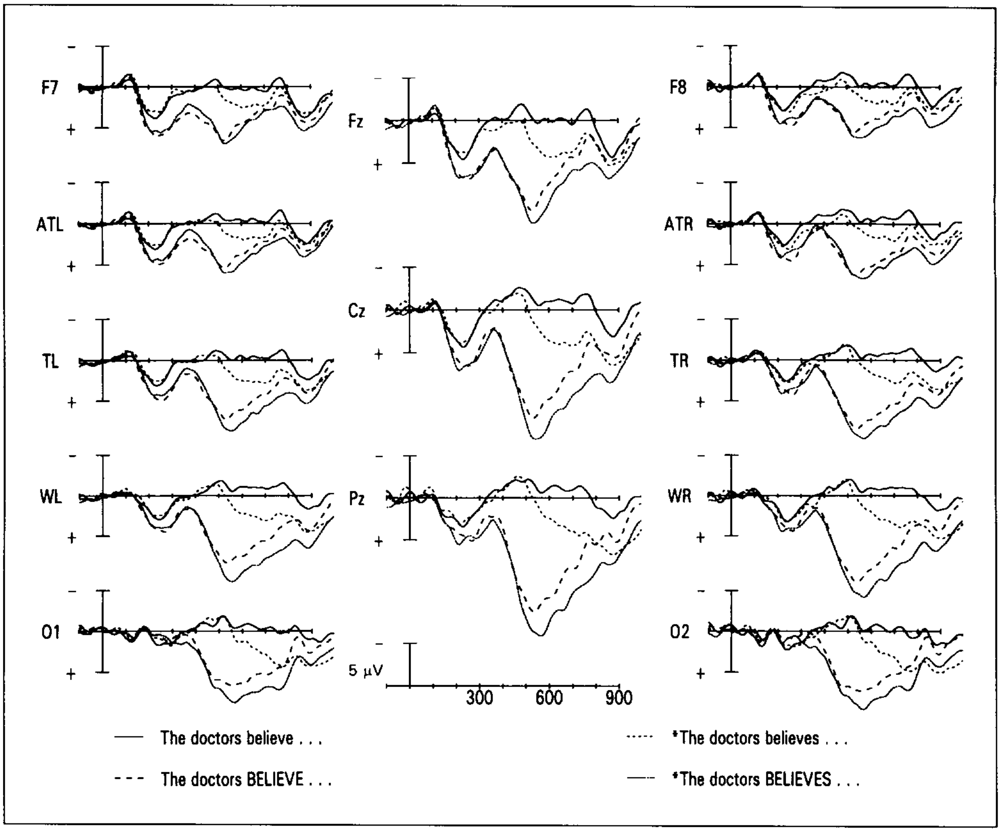


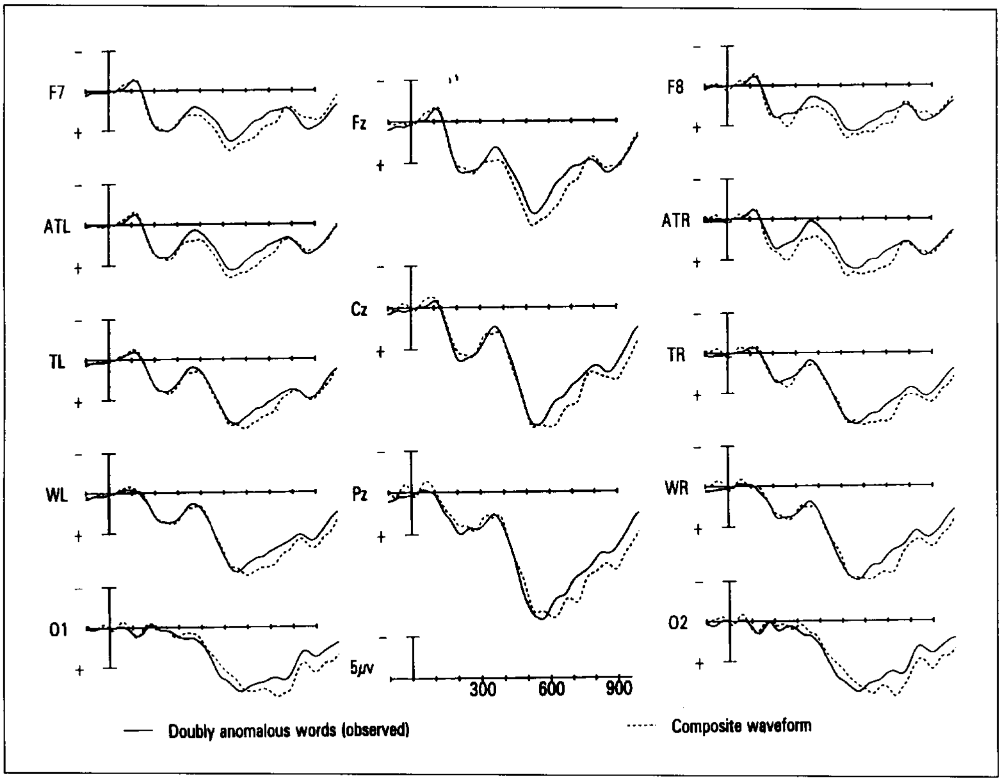
Figure 9. Grand average ERPs to nonanomalous (solid line), agreement-violating (small dashes), uppercase (large dashes), and doubly anomalous words (line-dash alternations), Experiment 3.

$F(1,15) = 63.19$, $p < 0.0001$], but did not differ from each other.

The most critical analyses involved mean amplitude within the 500–800 msec window, since it was within this window that both the agreement and case anomalies elicited ERPs distinct from those to controls; hence, it was in this window that evidence of additivity should be found. Again, the main effect for sentence type was reliable [midline, $F(3,45) = 61.97$, $p < 0.0001$; lateral, $F(3,45) = 61.51$, $p < 0.0001$]. Furthermore, there was a highly reliable interaction between sentence type and electrode site [midline, $F(6,90) = 17.96$, $p < 0.001$; lateral, $F(12,180) = 7.51$, $p < 0.01$], reflecting the posterior distribution of differences. Subsequent simple effects analyses revealed the following effects: ERPs to all three types of anomaly were reliably more positive-going than ERPs to controls [agreement violations: midline, $F(1,15) = 23.39$, $p < 0.001$; lateral, $F(1,15) = 21.28$, $p < 0.01$; uppercase words: midline, $F(1,15) = 94.50$, $p < 0.0001$; lateral, $F(1,15) = 96.05$, $p < 0.0001$; doubly anomalous words: midline, $F(1,15) = 160.05$, $p < 0.0001$; lateral, $F(1,15) = 155.63$, $p < 0.0001$]. More critical were direct comparisons of ERPs to the three anomaly types. ERPs to uppercase words were more positive-going than those to agreement violations [midline, $F(1,15) = 23.85$, $p < 0.0001$; lateral, $F(1,15) = 26.90$, $p < 0.0001$]. Furthermore, these effects had reliably different scalp distributions at lateral sites; as in Experiments 1 and 2, ERPs to uppercase words were largest over posterior regions of the right hemisphere, whereas ERPs to agreement violations were more evenly distributed across the hemispheres [sentence type × hemisphere × electrode site: raw data, $F(4,60) = 3.20$, $p < 0.05$; normalized data, $F(4,60) = 3.03$, $p < 0.05$]. Finally, ERPs to doubly anomalous words were more positive-going than those to agreement-violating [midline, $F(1,15) = 61.09$, $p < 0.0001$; lateral, $F(1,15) = 61.77$, $p < 0.0001$] and uppercase words [midline, $F(1,15) = 8.59$, $p < 0.01$; lateral, $F(1,15) = 7.15$, $p < 0.01$].

Thus, two predictions that follow from the independence hypothesis were supported in the above analyses: ERPs to the doubly anomalous words were reliably more positive-going between 500 and 800 msec than were ERPs to either of the anomaly types in isolation, and the onset of differences between the doubly anomalous and uppercase conditions occurred at about the same time as did differences between the agreement-violating and nonanomalous control words (i.e., at about 500 msec). We performed additional analyses to more precisely investigate these hypotheses. To obtain a more precise measure of the onset of differences between the agreement-violating and control conditions and between the doubly anomalous and uppercase conditions, we computed mean amplitude at midline sites within successive 50-msec windows, beginning at 200 msec and continuing until 600 msec. In both comparisons, reliable differences between conditions did not emerge until the

Figure 10. ERPs to doubly anomalous words (solid line) and a composite waveform (see text), Experiment 3.



500-550 msec window [agreement-violating vs. control: $F(1,15) = 11.62, p < 0.01$; doubly anomalous vs. uppercase: $F(1,15) = 4.48, p = 0.05$].

To more precisely evaluate the additivity of the responses to uppercase and agreement anomalies, we followed the procedures employed by Kutas and Hillyard (1980a), who examined the additivity of the N400 and P300 components. Under the assumption of simple additivity, ERPs to the words that were both uppercase and agreement violations (uc_V) should be equivalent to a composite waveform algebraically constructed from three components: the response to the lowercase nonviolating control words (lc_{NV}) plus the effects of uppercase alone (uc_{NV} minus lc_{NV}) and the agreement violation alone (lc_V minus lc_{NV}). That is,

$$uc_V = lc_{NV} + (uc_{NV} - lc_{NV}) + (lc_V - lc_{NV}) \qquad (1)$$

Figure 10 plots the observed response to the doubly anomalous words (solid line) and the constructed composite waveform (dashed line). At most sites, the observed waveform was reduced in amplitude relative to the composite waveform, particularly between about 600 and 800 msec. To determine whether the waveforms were reliably different within the critical region, we performed ANOVAs on mean amplitude within the 500-800 msec window. The two waves did not reliably differ, neither at midline, $F(1,15) = 2.23, p > 0.1$, nor at lateral sites, $F(1,15) = 2.05, p > 0.1$.

Note that Eq. (1) can be manipulated to form the following equation:

$$uc_V - uc_{NV} = lc_V - lc_{NV} \qquad (2)$$

That is, if the response to the agreement violations is completely independent of whether the word is in upper or lower case, the difference between the uppercase agreement-violating and nonviolating words should equal the difference between the lowercase agreement-violating and nonviolating words. We formed difference waves by subtracting ERPs to nonviolating words from those to violating words for each of the case conditions. The resulting waveforms are shown in Figure 11. These waveforms have strikingly similar onsets and scalp distributions, although they clearly differ (but not reliably) in amplitude within the region associated with the response to agreement violations, e.g., between 500 and 900 msec.

These observations indicate that although the effects of agreement violations and uppercase words appeared to be additive, this additivity was not a simple linear summation, although differences between the expected and observed waveforms were not statistically reliable. The lack of perfect additivity might be partly explained by recent indications that the P3b component is a composite of activation in a number of independent neural sources (Johnson, 1989, 1993; Ruchkin et al., 1990). In particular, the results of Experiment 1 indicated that the amplitude of the positivities elicited by the uppercase and agreement-violating words were both affected by a change in task relevance (although such effects were more robust and reliable in the uppercase condition). One reasonable possibility, then, is that in Experiment 3,

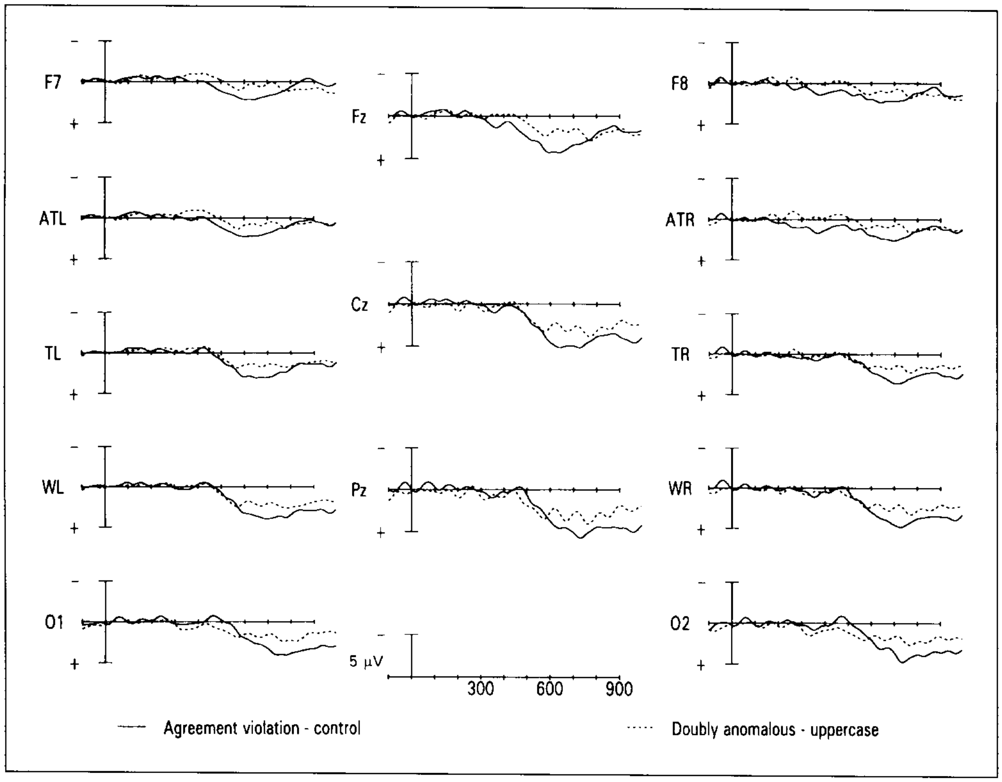


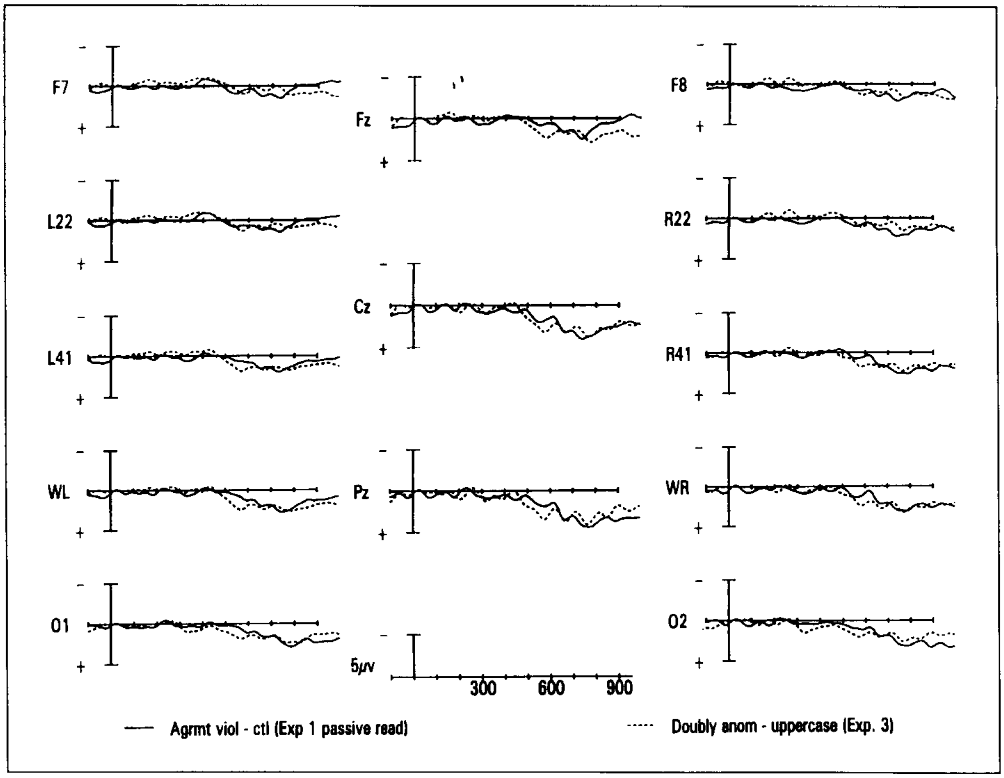
Figure 11. Difference waves formed by subtracting ERPs to nonanomalous control words from ERPs to agreement violations (solid line), and ERPs to uppercase words from ERPs to doubly anomalous words (dashed line), Experiment 3.

in which subjects made sentence-acceptability judgments, both types of anomaly activated (in common) the neural systems that respond specifically to the task relevance of the anomalies. If so, then one would *not* expect the contributions of these neural systems to the positive-going activity to summate when the stimulus is doubly anomalous, especially given that the uppercase, agreement-violating, and doubly anomalous words were all equally task relevant and informative with respect to the trial outcome. A preliminary indication of whether this factor can account for the lack of simple additivity is provided by a cross-experiment comparison. Note that some of the subjects in Experiment 1 were asked to read the sentences without making any sort of explicit judgment about sentence acceptability. Presumably, such conditions minimize the effects of the neural systems responding to the task-relevance and trial-outcome properties of agreement-violating words. Difference waves, formed by subtracting the response to the nonanomalous words from that to agreement violations in the passive reading condition in Experiment 2 (solid line) and the response to the uppercase words from the response to the doubly anomalous words in Experiment 1 (dashed line), are shown in Figure 12. Even though this comparison involves different subjects in each group (and different numbers of subjects and trials across conditions), the two waveforms were quite similar at most electrode sites in terms of onset and amplitude. These preliminary indications suggest that once the nonadditive effects of task relevance and trial-outcome information are taken into account, the effects of physical (uppercase) anomalies and grammar-violating (agreement) anomalies summate in an approximately linear fashion.

GENERAL DISCUSSION

The three experiments reported here comprised an initial effort to determine whether the positive shift elicited by syntactic violations is distinct from the P300 family of positivities elicited by unexpected, task-relevant events. We directly contrasted the response to agreement violations and uppercase words, which were predicted to elicit the syntactic positive shift (P600) and P300 responses, respectively. Several important findings were reported. First, whereas uppercase words elicited the classic P300 complex of positivities (a frontocentral P2/P3a peaking at about 300 msec, a centroparietal P3b peaking at about 500 msec, and a late slow wave), agreement violations elicited a late positive wave with a distinct morphology, time course, amplitude, and scalp topography. Second, although the amplitude of the response to uppercase words was reliably affected by changes in the task relevance and probability-of-occurrence of the anomaly, these factors had smaller and less reliable effects on the response to agreement violations. Third, the uppercase anomalies and agreement violations had additive effects, and preliminary results indicated that the additivity approximated a linear summation when the (putatively nonadditive) effects of the task

Figure 12. Difference waves formed by subtracting ERPs to nonanomalous words from those to agreement violations in the passive reading condition of Experiment 1 (solid line), and ERPs to uppercase words from those to doubly anomalous words (dashed line), Experiment 3.



relevance and informational content of the anomaly were taken into account.

All of these results are consistent with the claim that the response to at least one type of syntactic anomaly is not identical to the domain-general response to attended, unexpected events. It could be argued, however, that none of these results, in isolation, represents definitive evidence in support of this claim. For example, given that P3b latency is known to be a function of stimulus complexity, the differences in peak latency across conditions might reflect the likely possibility that uppercase words are more easily detected than are agreement violations. The reduction in peak amplitude in the agreement condition could result from increased latency jitter across trials and subjects. Differences in morphology and distribution across conditions might reflect differences in the relative timing of the P3b and slow wave components. However, such possibilities have difficulty accounting for all of the observed facts. If the syntactic effect is indeed a P3b, slow wave, or some combination of these components, then why did probability manipulations fail to affect the amplitude of the syntactic response as they affected P3b and slow wave amplitude in the response to uppercase words? The strength of the present study, in our view, is that the results reported here provide converging evidence that the ERP response to agreement violations is (at least partially) neurally and functionally distinct from the ERP response to nonsyntactic anomalies, whereas alternative explanations for particular results tend to be inconsistent with the results as a whole.

The possibilities discussed above presume that the P3b, slow wave, and syntactic positive shift are in large degree monolithic effects, i.e., that each effect reflects activity in a unique neural source or set of sources and (by implication) in a restricted, well-delineated set of cognitive processes. Perhaps a more realistic scenario is that these components are composites reflecting simultaneous activation in a number of more or less distinct neural and cognitive systems (Johnson, 1993; Ruchkin et al., 1990). Under such a scenario, the issue then becomes the *degree* of neural or cognitive similarity between two effects, rather than whether the syntactic positive shift is or is not a P3b or slow wave. For example, the amplitudes of both the syntactic positive shift and the P300 were affected by the task relevance of the anomaly (although the P300 seems to be much more sensitive to this variable than does the syntactic positive shift). This might indicate that both responses share a common neural source that is sensitive to task relevance. More generally, late-occurring endogenous components such as the syntactic positive shift might prove to be sensitive to a variety of factors reflecting the cognitive state of the subject (e.g., attentional, motivational, or strategic factors) even while being more selectively sensitive to certain aspects of cognitive function.

We should also note two recent reports indicating that the response to syntactic anomalies is in fact sensitive to probability manipulations, at least under certain experimental conditions (Coulson, King, & Kutas, 1995; Gunter, Vos, & Gulder, 1995). For example, Coulson et al. (1995) manipulated the probability (80% vs. 20%) of subject-verb number and pronoun-antecedent case agreement violations. The response to the agreement violations was significantly larger in the 20% condition than in the 80% condition. This finding seems to contradict the finding reported here (Experiment 2) that the response to agreement violations is not sensitive to probability manipulations. At present, it is unclear how to reconcile these conflicting reports. It might be that probability manipulations stronger than that used here will reveal reliable effects of probability on the response to syntactic anomalies. Regardless, it is our view that these reports provide insufficient grounds for concluding that the syntactic positive shift is another instantiation of the P300 family, for at least two reasons. First, it is entirely possible that both effects are probability-sensitive under certain conditions and yet remain largely distinct in terms of their underlying neural and cognitive events (e.g., it is possible that both effects are independently sensitive to probability manipulations, or that both effects share in common a probability-sensitive neural system but are independent in all other respects). Second, the critical issue is whether or not the responses to syntactic and nonsyntactic anomalies are *similarly affected* by probability manipulations, not simply whether or not the syntactic anomaly response is sensitive to probability. Importantly, neither Coulson et al. (1995) nor Gunter et al. (1995) presented nonsyntactic anomalies, thereby preventing an evaluation of the relative effects of probability on the responses to "syntactic" and "nonsyntactic" anomalies. However, the probability-related amplitude changes reported in these studies appeared to be much smaller in amplitude than previously reported changes in the P300 component, given a similar manipulation of probability (cf. Duncan-Johnson & Donchin, 1977).

The possibility that the ERP responses to agreement violations and uppercase words share a sensitivity to some subset of neural and cognitive processes is an important issue for subsequent study, as is the possibility that the syntactic positive shift will prove to be highly sensitive to probability and task manipulations under other experimental conditions. However, our results provide a preliminary indication that the ERP response to one type of syntactic anomaly is (at least to an interesting degree) both neurally and cognitively distinct from the response to one type of unexpected, task-relevant anomaly that does not involve the grammar. We should explicitly note that these results do not bring us closer to an understanding of the precise cognitive and neural events underlying these two brain responses. In particular, the current findings do not directly address the question of whether or not the syntactic positive shift is uniquely related to the processes underlying the syntactic analysis of sentences. Nonetheless, these results are largely inconsistent with the most salient alternative to the language specificity hypothesis, namely, the possibility that the syntactic positive shift is simply another manifestation of the domain-general response to improbable or unexpected events.

METHODS

Experiment 1

Subjects

Twenty-four undergraduates (15 females and 9 males) participated for course credit. In this and all subsequent experiments, all subjects were right-handed, native English-speaking students from the University of Washington who had normal or corrected-to-normal vision. Ages ranged from 18 to 22 (M = 20) years.

Materials

One hundred and fifty sentence triplets were constructed, as exemplified by the sentences in Table 1. Each sentence began with a plural noun phrase in subject position. One version of each sentence was grammatically well-formed and contained no anomalies of any type. A second version (agreement violation condition) contained a matrix verb that appeared in its singular form, engendering a number mismatch between the plural subject noun and the singular verb. In a third version (physical anomaly condition), the verb appeared in the correct plural form but was presented in uppercase letters. These materials were counterbalanced across three stimulus lists, such that each list contained one version of each sentence and 50 exemplars of each sentence type. Fifty well-formed, nonanomalous filler sentences were then added to each list. Hence, each subject saw 200 sentences, 100 of which contained an anomaly.

Procedure

Each trial consisted of the following events: A fixation cross appeared for 500 msec, after which a sentence was presented in a word-by-word manner, with each word appearing on the center of the screen for 350 msec. A blank-screen interval of 350 msec separated words. Sentence-ending words appeared with a period. A 1450-msec blank-screen interval followed each sentence. In the sentence-acceptability condition, the blank screen interval was followed by a prompt asking subjects to decide if the preceding sentence was "acceptable" or "unacceptable." Acceptable sentences were defined as those that were semantically coherent and grammatically well-formed. Subjects responded by pressing one of two buttons, which were counterbalanced (left and right) across subjects. Subjects in the passive reading condition simply pushed one button when they were ready for the next trial. Testing occurred in one 2-hour session, during which subjects were seated in a comfortable chair located in an isolated room.

Data Acquisition and Analysis

Continuous EEG was recorded from 13 scalp sites using tin electrodes attached to an elastic cap (Electrocap International). Electrode placement included International 10–20 system locations (Jasper, 1958) over homologous positions over the left and right occipital (O1, O2) and frontal (F7, F8) regions and from frontal (Fz), central (Cz), and parietal (Pz) midline locations. In addition, several nonstandard sites over posited language centers were used, including Wernicke's area and its right hemisphere homologue (WL, WR: 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz), temporal (TL, TR: 33% of the interaural distance lateral to Cz), and anterior temporal (ATL, ATR: one-half the distance between F7/F8 and T3/T4). Vertical eye movements and blinks were monitored by means of two electrodes, one placed beneath the left eye and one placed 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 cutoff) by a Grass Model 12 amplifier system. Activity over the right mastoid was actively recorded on a sixteenth channel to determine if there were any effects of the experimental variables on the mastoid recordings. No such effects were observed.

Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was performed by a Data Translation 2801-A board and a 486-based computer at a sampling frequency of 200 Hz. Epochs were comprised of the 100 msec preceding and the 1180 msec following presentation of individual words in the sentences. Trials characterized by excessive eye movement or amplifier blocking were removed prior to averaging. Analyses involving mean amplitude and peak amplitude were performed. Mean amplitude was quantified within a latency range following presentation of words of interest relative to a baseline of activity comprised of the 50 msec of activity following presentation of the word of interest.[6] In most cases, analyses of variance were performed on mean amplitude within three time windows: 200–300, 300–500, and 500–800 msec. These windows were chosen because they roughly correspond to the latency ranges of the P2/P3a, P3b, and syntactic positive shift/slow wave positivities. Peak latency was quantified as the largest positive voltage within a specific time window. Reliable main effects and interactions were followed when appropriate by simple effects analyses; the error term for these analyses was the within-groups mean square from the original analysis of variance (cf. Keppel, 1982). Data acquired at midline and lateral sites were treated separately to allow for quantitative analysis of hemispheric differences. On data acquired over midline sites, ANOVAs were performed with repeated measures on three levels of sentence type and three levels of electrode site. On data acquired over lateral sites, the ANOVAs involved repeated measures on sentence type, five levels of electrode site, and two levels of hemisphere. To protect against Type I error due to violations of the assumption of equal variances of differences between conditions of within-subject factors, the Geisser–Greenhouse correction (Greenhouse & Geisser, 1959) was applied when evaluating effects with more than one degree of freedom in the numerator. In such cases, the corrected p value is reported. For analyses involving significant interactions between sentence types and electrode sites in the presence of a reliable main effect of sentence type, two sets of analyses were reported: analyses on raw data, and analyses on data that have been normalized following the procedure described by McCarthy and Wood (1985). This normalization procedure was used because in certain cases spurious interactions can result if the experimental effects are of different overall amplitude. To minimize the number of reported analyses, analyses on normalized data are reported only when the effect is deemed to be theoretically important (i.e., when the distributions of the responses to two anomaly conditions are directly compared).

Experiment 2

Subjects

Twelve undergraduate and graduate students (8 males and 4 females) participated for course credit or for a small compensation. Ages ranged from 18 to 38 ($M = 22$) years. Each subject participated in two experimental sessions separated by at least 1 week.

Materials

One hundred and eighty experimental sentences were constructed. Each sentence contained a singular or plural noun in subject position of the matrix clause. (Some of these sentences had been used in Experiments 1 and 2.) These sentences were divided into two sets (A and B) of 90 sentences each. Three versions of each sentence were constructed: nonanomalous, sentences that contained a matrix verb that disagreed in number with the subject noun, and sentences that contained an uppercase word. As in Experiment 1, the critical word was always the same across all versions of a sentence. Within each set, the sentences were then combined in a latin square design to form three stimulus lists such that each list contained 30 sentences of each type with only one version of a particular sentence on a given list.

One hundred and twenty filler sentences were also constructed. These filler sentences were similar in form to the experimental sentences, and were used to manipulate the probability of occurrence of the agreement and case anomalies across lists. Sixty of the fillers were added to each of the three stimulus lists in Set A, the other 60 were added to Set B. Two versions of each filler sentence were created: sentences containing an agreement violation (of the type noted above) and 60 containing an uppercase word. Again, the critical word was always the same word across versions of a sentence. The agreement-violating versions were added to the experimental sentences to form lists containing 30 nonanomalous sentences, 90 sentences containing an agreement violation, and 30 sentences containing an uppercase word (the "60% agreement/20% uppercase" condition). The uppercase word versions of the fillers were added to the experimental sentences to form lists containing 30 nonanomalous sentences, 30 sentences containing agreement violations, and 90 sentences containing an uppercase word (the "20% agreement/60% uppercase" condition). Thus, there were a total of 12 stimulus lists, each of which contained 150 sentences, 120 of which were anomalous. In the "60% agreement/20% uppercase" condition, agreement and case anomalies occurred on 60 and 20% of the trials; those percentages were reversed in the "20% agreement/60% uppercase condition." Similar shifts in probability have been shown to elicit large changes in P300 amplitude (e.g., Duncan-Johnson & Donchin, 1977).

Procedure

These were identical to those used in Experiment 1, with the following exception. Each subject was run twice, once on a list with 90 agreement violations and 30 uppercase words, and once on a list with 30 agreement violations and 90 uppercase words. The order of list presentation was counterbalanced such that half of the subjects saw the "60% agreement/20% uppercase" list first, and half did not. Furthermore, subjects always saw sentences from Set A during one session and sentences from Set B in the second session, or the reverse; order of presentation was counterbalanced across subjects. Hence, no subject saw more than one version of a particular experimental or filler sentence during the entire two-session experiment. For all but one subject, the first and second sessions were separated by 1 week and occurred at the same time of day. For one subject, 11 days intervened between test sessions.

Experiment 3

Subjects

Sixteen undergraduates (8 females and 8 males) participated for course credit or for a small compensation. Ages ranged from 18 to 42 (M = 25) years.

Materials

Four versions of 120 sentences were constructed, as exemplified in Table 1. These sentences were modified versions of a subset of the experimental sentences used in Experiment 1. The critical words in the four versions were as follows: nonanomalous, uppercase, agreement-violating, or doubly anomalous, i.e., both in uppercase letters and an agreement violation. These materials were counterbalanced across four stimulus lists, such that each list contained one version of each sentence and 30 exemplars of each sentence type. In addition to the experimental sentences, 60 well-formed, nonanomalous filler sentences were also constructed and added to each stimulus list. These fillers consisted of a variety of sentence structures.

Procedures

Procedures were identical to those used during Experiment 2, except that ANOVAs involved four levels of sentence type.

Acknowledgments

We thank Judy McLaughlin, Sherri Sipe, and Marty Blount for help in running subjects and for comments on previous versions of this manuscript. We received financial support from Grant 5 R29 DC01947, National Institute for Deafness and Other Communication Disorders, National Institutes of Health. Portions of these data were presented at the Fifth Annual CUNY Conference on Human Sentence Processing, 1993, Amherst, MA, the Thirtieth Annual Meeting of the Psychonomics Society, 1993, Washington, D.C., and the First Annual Meeting of the Cognitive Neuroscience Society, 1994, San Francisco.

Reprint requests should be sent Lee Osterhout, Department of Psychology, University of Washington, Seattle, WA 98195.

Notes

1. Exceptions exist in the literature to the generalization that semantic anomalies elicit a monophasic negative-going wave (N400), whereas syntactic anomalies elicit a monophasic positive-going wave (syntactic positive shift). For example, Münte and Heinze (1994) reported that semantic anomalies encountered during the reading of text elicited both an N400 effect and a subsequent late positive wave, and in some reports syntactic anomalies have elicited a negativity over left anterior regions (Münte, Heinze, & Mangun, 1993; Rösler et al., 1993).
2. Identification of the precise cognitive events underlying the subcomponents of the P300 remains an issue for debate. Hypotheses concerning this question are typically stated as generalizations over the set of stimulus or task manipulations that alter the characteristics of the component (Donchin & Coles, 1988). The P3a has been posited to reflect an "orienting response" to novel stimuli (cf. Sutton & Ruchkin, 1984). The P3b has been associated with the updating of working memory (cf. Donchin, 1981; Donchin & Coles, 1988) or with processes that extract and evaluate information (Johnson, 1993; Ruchkin et al., 1982). Functional characterizations of the slow wave are more tentative, although some researchers have proposed that it reflects the mobilization of effort that occurs in response to task demands (Ruchkin et al., 1982).
3. The use of scalp distribution to determine whether two effects are generated by independent neural sources becomes somewhat problematic when the effects to be contrasted are composites of activity in multiple sources. For example, two effects that are composites of activity in several simultaneously

active neural sources might have distinct distributions, even if a subset of these neural sources contributes to both effects.
4. The Rule of Superposition assumes homogeneity of the conductive medium, which is clearly an idealization of the conductive properties of neural tissue (cf. Nunez, 1981).
5. We computed difference waves rather than directly compare ERPs in each task condition in order to isolate the effects of interest from general effects of the change in task and the use of separate groups of subjects in the two task conditions.
6. The conventional procedure would be to use a prestimulus baseline, i.e., a baseline comprised of activity preceding critical word onset. We used the poststimulus baseline to mitigate small differences between conditions that existed very early in the epoch (e.g., within the initial 50 msec) when a prestimulus baseline was used. Given their very early onset, it is highly unlikely that these small differences resulted from our experimental manipulations. In employing a poststimulus baseline, we hoped to mitigate the influence of these differences when assessing the effects of our manipulated variables. However, we also performed all of the analyses reported here with a baseline comprised of the 100 msec of activity preceding critical word onset. The results of these analyses were quite similar in most respects to those reported.

REFERENCES

Ades, A., & Steedman, S. (1982). On the order of words. *Linguistics and Philosophy, 6,* 517-558.

Bock, J. K., & Kroch, A. S. (1989). The isolability of syntactic processing. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing.* Dordrecht: Kluwer.

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

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Coulson, S., King, J., & Kutas, M. (1995). The late show: The syntactic positive shift meets the late positive component. Poster presented at the Eighth Annual CUNY Conference on Human Sentence Processing.

Courchesne, E., Courchesne, R. Y., & Hillyard, S. A. (1978). The effect of stimulus deviation on P3 waves to easily recognized stimuli. *Neuropsychologia, 16,* 189-199.

Deacon, D., Breton, F., Ritter, W., & Vaughan, H. G. (1991). The relationship between N2 and N400: Scalp distribution, stimulus probability, and task relevance. *Psychophysiology, 28,* 185-200.

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.

Donchin, E., Ritter, W., & McCallum, W. C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, & S. H. Koslow (Eds.), *Event-related brain potentials in man.* New York: Academic Press.

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation in event-related potentials with subjective probability. *Psychophysiology, 14,* 456-467.

Elman, J. L. (1990). Representation and structure in connectionist models. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing.* Cambridge, MA: MIT Press.

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). The definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 1). Greenwich, CT: JAI Press.

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

Fodor, J. A. (1983). *Modularity of mind.* Cambridge, MA: MIT Press.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett.* Hillsdale, NJ: Erlbaum.

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., 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.

Garnsey, S. (1993). Event-related brain potentials in the study of language: An introduction. *Language and Cognitive Processes, 8,* 337-356.

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

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language, 27,* 135-159.

Gunter, T. C., Vos, S. H., & Mulder, G. (1995). *Syntactic violations and ERPs: P600 or P3b?* Poster presented at the Eighth Annual CUNY Conference on Human Sentence Processing.

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

Hillyard, S. A., Münte, T. F., & Neville, H. J. (1985). Visual-spatial attention, orienting, and brain physiology. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI.* Hillsdale, NJ: Erlbaum.

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. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes, 5,* 281-312.

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,* 633-650.

Johnson, R., Jr. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology, 30,* 90-97.

Keppel, G. (1982). *Design and analysis: A researcher's handbook.* Englewood Cliffs, NJ: Prentice-Hall.

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

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

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

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.

MacWhinney, B., Bates, E. S., & Kliegel, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior, 23,* 127-150.

McCarthy, G., & Wood, C. C. (1986). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 62,* 203-208.

McKinnon, R., & Osterhout, L. (in press). *Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials.*

Münte, T. F., & Heinze, H. (1994). ERP negativities during syntactic processing of written words. In H. Heinze, T. F. Münte, & G. R. Mangun (Eds.), *Cognitive electrophysiology.* Boston, MA: Birkhauser.

Münte, T. F., Heinze, H., & Mangun, G. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience, 5,* 335-344.

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. L. (1985). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory and Language, 25,* 75-92.

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.

Nunez, P. L. (1981). *Electrical fields in the brain: The neurophysics of EEG.* New York: Oxford University Press.

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

Osterhout, L. (1994). Event-related brain potentials as tools for comprehending sentence comprehension. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing.* Hillsdale, NJ: Erlbaum.

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 brain 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. (1995). Event-related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related potentials and cognition.* Oxford: Oxford University Press.

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 786-803.

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

Polich, J. (1985). Semantic categorization and event-related potentials. *Brain and Language, 26,* 304-321.

Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin, 89,* 506-540.

Rösler, F., Friederici, A., Pütz, P., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience, 5,* 345-362.

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,* 157-175.

Ruchkin, D. S., Munson, R., & Sutton, S. (1982). P300 and slow wave in a message consisting of two events. *Psychophysiology, 19,* 629-642.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). The word order problem in agrammatism. I. Comprehension. *Brain and Language, 10,* 249-262.

Squires, K. C., Donchin, E., Herning, R. I., & McCarthy, G. (1977). On the influence of task relevance and stimulus probability on event-related potential components. *Electroencephalography and Clinical Neurophysiology, 42,* 1-14.

Squires, N. K., Donchin, E., Squires, K. C., & Grossberg, S. (1977). Bisensory stimulation: Inferring decision-related processes from the P300 component. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 299-315.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology, 38,* 387-401.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked potential correlates of stimulus uncertainty. *Science, 150,* 1187-1188.

Sutton, S., & Ruchkin, D. S. (1984). The late positive complex: Advances and new problems. In R. Karrer, J. Cohen, & P. Tueting (Eds.), *Brain and information: Event-related potentials.* New York: Academy of Sciences.