Psychobiology
1991, 19 (4), 286-300

Natural speech processing: An analysis using event-related brain potentials

PHILLIP J. HOLCOMB
Tufts University, Medford, Massachusetts

and

HELEN J. NEVILLE
The Salk Institute, La Jolla, California

In two experiments, event-related brain potentials were collected as subjects listened to spoken sentences. In the first, all words were presented as connected (natural) speech. In the second, there was a 750-msec interstimulus interval (ISI) separating each of the words. Three types of sentence-ending words were used: best completions (contextually meaningful), unrelated anomalies (contextually meaningless), and related anomalies (contextually meaningless but related to the best completion). In both experiments, large N400s were found for the unrelated and related anomalies, relative to those found for the best-completion final words, although the effect was earlier and more prolonged for unrelated anomalies. The auditory N400 effect onset earlier in the natural-speech experiment than it did in either the 750-msec ISI experiment or previous visual studies.

The role of context in language processing continues to be a central topic in much of the psycholinguistic literature (e.g., Ferreira & Clifton, 1986; Fischler & Bloom, 1985; Schustack, Enrlich, & Rayner, 1987; Tyler & Marslen-Wilson, 1982; Van Petten & Kutas, 1990). While there is considerable debate as to the range and source of contextual influences, particularly with regard to word recognition, there are few theorists that would deny that prior linguistic information plays at least some role in the processing of currently presented words. However, until very recently, the effects of context on *spoken*-language processing have been largely ignored (see Tyler & Frauenfelder, 1987).

Early psychological models of language comprehension processes were primarily concerned with reading, and they assumed that the same set of operations are involved in spoken-language comprehension (e.g., Forster, 1979; Morton, 1969). The reasons for this concentration on written language are at least partially practical, in that visual stimuli are easier to present, control, and manipulate. However, from another perspective, the concentration on reading is curious, in that spoken language has certainly had more time to have an impact on the evolution of the language processing system and continues to be the most widely used language medium (e.g., Walker, 1987). A number of researchers have recently attempted to rectify this imbalance by focusing their efforts on deriving models of spoken-language processes (e.g., Cole & Jakimik, 1980; Ellman & McClelland, 1987; Marslen-Wilson, 1987).

This research was supported by NIH Grants DC00128, DC00481, NS22343, and HD25889. Correspondence should be sent to either Phillip J. Holcomb, Department of Psychology, Tufts University, Medford, MA 02155, or Helen J. Neville, The Salk Institute, Box 85800, San Diego, CA 92138.

Marslen-Wilson (1980, 1987) has proposed one of the most widely cited theories of spoken-language processing. In his *cohort* model, incoming acoustic-phonetic information from the beginning of a word makes contact with all of the compatible entries in the listener's lexicon. Marslen-Wilson refers to this set of activated, or "accessed," entries as the *word-initial cohort*. Over time, as more and more information becomes available, the initial cohort is pruned until a single entry remains, at which time the word is said to be "selected." It is only after selection that semantic and syntactic information stored with the lexical entry becomes available for "integration" with the ongoing discourse.

Part of the appeal of the cohort model has been that it makes strong predictions about the time course of spoken-word processing. The point in time at which a word presented in isolation is selected (recognized) will be determined not by its absolute duration, but rather by the existence (or more accurately, nonexistence) of other words that share the same initial sounds. Marslen-Wilson and colleagues (Marslen-Wilson, 1980, 1984; Marslen-Wilson & Tyler, 1980; Marslen-Wilson & Welsh, 1978) have demonstrated that subjects' response latencies to words spoken in isolation are best predicted by each word's point of uniqueness; the first point at which the incoming acoustic-phonetic pattern is different from all but one word. However, evidence from tasks in which target items were embedded in sentences has shown that selection can occur even earlier than predicted by the point of uniqueness (in one study, the average recognition time was 200 msec for words in a sentence context; see Marslen-Wilson & Tyler, 1980). This finding supports the hypothesis that spoken-word processing is facilitated by sentence-level contextual factors.

One problem with the conclusions from the Marslen-Wilson studies is that the techniques he and others have used (including word monitoring, shadowing, lexical decision, and gating) are inherently intrusive and may thus require a different strategy or mode of word processing than that used in normal spoken-language comprehension. In addition, the dependent variables used by these researchers (primarily reaction time) usually yield information several hundreds of milliseconds after presentation of the critical stimulus, thus making them relatively "offline" measures of the cognitive processes under study. The goal of the first experiment to be reported here was to engage subjects in a more natural spoken-language task and to unobtrusively monitor ongoing linguistic processes in real time. This was accomplished by recording event-related brain potentials (ERPs) to the onset of spoken words placed in a sentence context.

EVENT-RELATED POTENTIALS

By placing electrodes on the scalp of human subjects, it is possible to record the ongoing electrical activity of the brain. ERPs are stimulus-locked perturbations in this activity, which have been demonstrated to be sensitive to both sensory and cognitive processes (see Regan, 1989, for a review).

In several recent reports, ERPs have been used to study the effects of context on linguistic stimuli. A number of these studies have reported changes in a negative component (the individual waves that make up an ERP are referred to as *components*), which onsets as early as 200 msec poststimulus onset and peaks near 400 msec. Kutas and her colleagues (e.g., Kutas & Hillyard, 1980, 1984) have reported that this "N400"[1] component is large (more negative) to sentence-final words that are anomalous (e.g., He takes cream and sugar in his *attention*.) and is small or nonexistent to highly probable "best-completion" sentence endings (e.g., "He takes cream and sugar in his *coffee*"). In contrast, manipulation of the physical parameters of the final word (e.g., using a different type font) results in a variability in an ERP positivity that peaks around 600 msec.

In a related study, Kutas, Lindamood, and Hillyard (1984) demonstrated that N400 amplitude was a monotonic function of the cloze probability of sentence-final words; N400 was greater in amplitude to less predictable words (e.g., "Captain Sheir wanted to stay with the sinking *raft*") and smaller to more predictable words (e.g., "She called her husband at his *office*").[2] In a second experiment, Kutas et al. (1984) replicated an earlier reaction time study by Kleiman (1980). Kleiman demonstrated that sentences with best-completion endings (e.g., "The king of beasts is the *lion*") are responded to faster than are sentences with related anomalous endings (e.g., "The king of beasts is the *roar*" [*roar* is semantically related to *lion*]), which are in turn responded to faster than are sentences with completely anomalous endings (e.g., "The king of beasts is the *work*"). In their ERP study, Kutas et al. (1984) demonstrated that N400 amplitude followed a similar pattern, and was largest to unrelated and anomalous sentence endings, was next largest to anomalous but related endings, and was smallest to best-completion endings. As pointed out by Kutas et al., the finding of faster responses and smaller N400s to related but anomalous endings, relative to unrelated anomalies, is consistent with a spreading-activation model of final-word processing. According to this view, the semantic context provided by the sentence (prior to the final item) primes the lexical entry for the best-completion ending, which in turn spreads activation to semantically related items, some of which may be anomalous in the context of the sentence. The spread of activation facilitates the processing of this otherwise contextually anomalous item. N400 was intermediate in amplitude presumably because these items received some priming due to spreading activation (this tends to reduce N400 amplitude), but they were anomalous in the context of the sentence (this tends to increase N400 amplitude).

The above studies used visual letter strings as stimuli. Recently, Holcomb and Neville (1990) compared and contrasted semantic priming in the visual and auditory modalities. Subjects participated in two versions (one visual, one auditory) of a lexical decision task in which stimuli were word pairs consisting of a prime word followed by a semantically related word, an unrelated word, or a nonword. N400s were larger to unrelated words than to related words in both modalities. However, this ERP "priming effect" began earlier in the auditory modality than in the visual modality. In addition, the distribution over the scalp differed in the two modalities, with the visual priming effect slightly larger over the right hemisphere and with the auditory effect slightly larger over the left hemisphere. Holcomb and Neville concluded that there may be overlap in the priming processes that occur in each modality but that these processes are not identical. In particular, they noted that the earlier onset of the N400 in the auditory modality was consistent with the Marslen-Wilson (1987) view that auditory word processing can begin prior to the arrival of all of the acoustic information in a spoken word and that the time course of this processing can be influenced by semantic properties of a prior word (i.e., context).

McCallum, Farmer, and Pocock (1984) performed an auditory replication of the original Kutas and Hillyard (1980) visual sentence study. They included sentences spoken by a male with best-completion endings (e.g., "The pen was full of blue *ink*"), semantically anomalous endings (e.g., "At school he was taught to *snow*"), and best-completion endings that had a physical deviation (the final word spoken by a female). As in previous visual studies, ERPs to the anomalous endings produced a negative component (peak latency 456 msec), whereas ERPs to appropriate endings produced a relatively flat response between 200 and 600 msec and ERPs to physically deviant endings produced a large positivity in this latency band. McCallum et al., noted that while their auditory N400s had a similar scalp distribution to those recorded by Kutas and Hillyard (1980) in the visual modality, they were also somewhat less peaked.

EXPERIMENT 1

Given the findings of Holcomb and Neville (1990) of a significantly earlier onset of the auditory N400 effect for word pairs, it is curious that McCallum et al. (1984) did not notice a similar early onset in their sentence study. Although McCallum et al. did not systematically examine the time course of the N400 (other than to note that the peak was somewhat later than in the visual modality), examination of their figures does not suggest much of a difference from previous visual studies (e.g., Kutas & Hillyard, 1980). One possibility for the apparent contradiction between the Holcomb and Neville study and the McCallum et al. work is that McCallum et al. did not use natural connected speech stimuli for the words of interest. Rather than using the actual final words spoken at the end of each experimental sentence, these investigators spliced in final words spoken in other sentence contexts. Although this seems like a reasonable procedure for looking at the "pure" effects of semantic context on the processing of sentence-final words, it might have introduced conflicting or contradictory nonsemantic contextual cues that are normally part of connected natural speech, such as prosodic or across-word coarticulatory factors.

Prosody refers to the rhythmic patterns that occur both within words (lexical prosody) and across the words of an utterance (metrical prosody; see Cutler, 1989). From the listener's perspective, coarticulation refers to changes in speech sounds that are due to the influence of sounds coming prior to and after the current sound. For example, the *a* in *cat* is pronounced differently from the *a* in *hat* because of the different influences of the *h* and *c* sounds on the *a* sound. Likewise, the *c* in *cat* is pronounced differently from the *c* in *cold* due the differential influence of the following *a* and *o* sounds. Coarticulatory factors also play a role across word boundaries in natural speech. So, for example, the final sounds in *her* (as in the sentence: "The bird built a nest in which to lay *her* . . .) will differ depending on whether the next word is *nest* or whether it is *car*.

It seems reasonable that the onset of the N400 effect may have been delayed in the McCallum et al. study because of inconsistent across-word cues between the penultimate and sentence-final words. The primary purpose of the first experiment to be reported here was to replicate and extend the McCallum et al. study by using more natural speech with normal semantic and nonsemantic cues and to more closely examine the time course of the ERP differences between the various types of sentence-final words. One prediction was that the inclusion of nonsemantic speech cues would shift the temporal distribution of the N400 effect earlier than in the McCallum et al. study. A second and related purpose of this experiment was to examine the generality of the finding in the visual modality (Kutas et al., 1984) that anomalous words that are related to best-completion words produce smaller N400s than do unrelated anomalous words.

Finally, a considerable amount of theorizing and speculation has been directed at the problem of speech segmentation (e.g., Cole & Jakimik, 1980; Cutler, 1987; Frazier, 1987; Pisoni & Luce, 1987). It has been known for many years that the speech signal is not composed of a series of discrete words, but rather is a continuous flow of sound with few breaks or pauses (e.g., Pisoni & Luce, 1987). Two important questions arise from this observation: (1) How do listeners segment the complex spoken signal into individual units? and (2) What are the basic units of speech perception? The ERP technique permits the recording of brain-wave activity time locked to the onset of each word in a sentence. Numerous studies have shown a consistent pattern of early "sensory" components in ERPs recorded to isolated nonlinguistic (e.g., Picton, Hillyard, Krausz, & Galambos, 1974) and linguistic (e.g., Hansen, Dickstein, Berka, & Hillyard, 1983; Holcomb & Neville, 1990) sounds. A reasonable preliminary question then is, Do words in connected speech generate a consistent pattern of sensory ERP components?

Method

Subjects. Twelve young adults (7 male, 5 female; mean age = 21.6 years) were paid $5 per hour to serve as subjects. All were right-handed (three had at least one left-handed relative in the immediate family) and native speakers of English.

Stimuli and Procedure. The stimuli for Experiment 1 were generated from 135 highly constrained sentences (final-word cloze probability > 0.8), ranging from 6 to 13 words in length. Forty-five of the sentences were randomly selected to be used in the *best-completions* condition. These sentences were all completed with a final word that fit with the previous context of the sentence (e.g., "December is the last month of the *year*"). Forty-five other sentences were randomly selected to be used in the *unrelated-anomaly* condition. In these sentences, the best-completion final words were replaced with words that made no sense given the prior context (e.g., "The bird built a nest in which to lay her *cars*"). The remaining 45 sentences were selected to be used in the *related-anomaly* condition. In these sentences, best-completion final words were replaced with semantically related words that were anomalous given the sentence context (e.g., "The sink was so clogged they called a *pipe*").[3] Unrelated anomalies and related anomalies had a cloze probability of zero. They also always maintained the inflection of the original best-completion endings and, with the exception of four items in each condition, did not share any word-initial sounds with the original best-completion words.[4] There were no significant differences between the three conditions in the duration of final words (mean = 561 msec, range 318 to 901), nor were there any differences in the number of words in the stem sentences or the contextual constraint of the stems (as measured by cloze probability to the original sentences). There were no significant differences in log frequency between the best completions (log frequency = 1.77) and unrelated anomalies (log frequency = 1.62) ($F = 0.58$); however, related anomalies (log frequency = 1.42) were significantly less frequent than were best completions and unrelated anomalies [$F(2,132) = 3.43$, $p < .035$].

Each sentence was spoken by a female member of our research team at a normal speaking rate (mean = two words per second) and with normal pitch and intonation. Sentences were first recorded on analogue tape and were then digitized (12 KHz, 12-bit resolution) and broken up into word-sized pieces, which were stored in separate disk files. Using this technique, the onset envelope of each word was aligned with the beginning of its digital disk file. Pauses and natural silent periods between words were maintained by placing them (when they occurred) at the ends of files (note that this procedure results in all of the information in each sentence being preserved in the digital representation). During the experiment, the stimulus-presentation software reassembled each sentence in real time (through the use of a circular buffer) and, simultaneous with the onset of each word, a code was output to the computer, which digitized the EEG. This procedure resulted in natural-sounding sentences that were indistinguishable from those recorded on analogue tape.

The experiment was self-paced, each trial beginning when the subject pressed a button on a small panel resting in his/her lap. Two seconds later the outline of a white rectangle appeared on a video monitor in front of the subject (6° × 3°). The subject was told not to move or blink during the time the rectangle was on the screen. One second after the onset of the rectangle, the sentence was played (binaurally over headphones, 60 dB SL). The rectangle was turned off 1.5 sec after the end of the last word of the sentence and was replaced by a message to respond. The subject pressed a button indicating whether or not the sentence made sense. Because the response was delayed 1.5 sec (to prevent the motor response from contaminating the ERP to the final word), accuracy of response, rather than speed, was emphasized. Response hand was counterbalanced across subjects. Each subject engaged in 10 practice trials prior to the run of 135 experimental sentences. The three types of sentence endings (best completions, unrelated anomalies, and related anomalies) were randomly intermixed.

ERP recording. The EEG was recorded from 14 scalp electrodes including 8 placed over standard International 10-20 System locations—left and right occipital (O1, O2), posterior temporal (T5, T6), frontal (F7, F8), and the midline (Cz, Pz)—and 6 nonstandard locations—Wernicke's region and the right hemisphere homologue (WL and WR), left and right temporal (TL and TR), and anterior temporal left and right (ATL and ATR)[5]—attached to an elastic cap (Electro-Cap) and referenced to linked mastoids. The electro-oculogram (EOG) was recorded from an electrode attached beneath the left eye (mastoid reference) and from a bipolar montage of electrodes attached just lateral to the two eyes. All impedances were maintained below 5 kΩ. Grass 7P511 amplifiers (bandpass 0.01 to 100 Hz) were interfaced to a 16-channel 12-bit A/D converter, and the EEG was digitized on line and stored on digital tape.

Off line, separate ERPs (100 msec prestimulus baseline) were averaged for each subject at each electrode site from trials free of EOG artifact. For the sentence-final words, a separate average was made for each condition (best completions, unrelated anomalies, and related anomalies). Only trials on which the subject responded correctly were included. Separate ERPs were also formed for all middle words from each of the sentences (first and final words were excluded). However, due to excessive eye artifact to middle words, 2 subjects' data were removed from middle-word analyses.

ERPs to the final words were quantified by measuring the mean amplitude in several latency windows (200-500 msec, 500-1,000 msec, and 1,000-1,400 msec) and by measuring the peak latency and amplitude of the N1 component (90-200-msec window). To look at the time course of the onset of the N400 effect (the point in time at which the best completions and the two anomalies differentiate), a series of smaller epoch measures were taken. These included nine sequential 50-msec mean amplitude measurements between stimulus onset and 450 msec.

Middle-word ERPs were quantified by measuring the peak amplitude and latency of the P1, N1, and P2 components (P1, latency window 14-100 msec; N1, 90-200 msec; P2, 150-300 msec) and by measuring the average areas between 200 and 700 msec and between 700 and 1,400 msec.

Repeated measures analyses of variance (ANOVAs) were used to analyze all measures. Factors in the analysis included electrode site (frontal, anterior temporal, temporal, parietal, posterior temporal, and occipital) hemisphere and, for final words, condition (related anomalies vs. unrelated anomalies vs. related anomalies).[6] In cases where specific predictions were made for final words, the overall ANOVA was followed up with a series of planned pairwise comparisons contrasting each of the three conditions. The correction recommended by Geisser and Greenhouse (1959) was applied to all variables with greater than two levels.

Results

Behavior. The subjects reported no difficulty in understanding the sentences and were equally accurate at judging whether or not each of the three sentence types made sense (97% for each condition).

ERPs from middle words. Plotted in Figure 1 are the ERPs from the average of all the middle words in the sentences. There is a very early positivity that appears to begin prior to stimulus onset and peaks at 45 msec poststimulus onset (P45). P45 can be seen at all sites, but is smaller at the most posterior sites (T5, T6, O1, O2). This observation was confirmed by the ANOVA on P45 peak amplitude [main effect of electrode site, $F(5,45) = 5.63$, $p < .0004$]. The analysis of the latency of P45 indicated that this component peaked slightly earlier in the left hemisphere [40 vs. 50 msec; main effect of hemisphere, $F(1,9) = 5.58$, $p < .043$]. Because of its prestimulus onset, this component may reflect the summation of activity from the prior stimulus and the P45 component of the time-locked stimulus.

A prominent feature early in the waveforms was a negative component that peaked near 140 msec (N140). Although N140 can be seen as far posterior as the WL/WR sites, it was largest at the most anterior electrode locations (frontal, anterior temporal, and temporal sites) and was virtually nonexistent at the most posterior sites for both middle words [main effect of electrode site, $F(5,45) = 17.8$, $p < .0001$] and final words [$F(5,55) = 15.72$, $p < .0004$]. This scalp distribution is consistent with that observed to other auditory stimuli and suggests that this component belongs to the N1 family of negativities reported in numerous previous studies (Regan, 1989). Following N140 was a positive-going component that peaked near 200 msec (P200). Although positive-going, the P200 to the middle words did not cross the zero baseline. P200 was most apparent at the most anterior electrode positions and was small or nonexistent at the most posterior sites [main effect of electrode, $F(5,45) = 3.78$, $p < .043$]. This positivity seems most likely to be related to the P2 component frequently reported in auditory studies (see Regan, 1989). Following P200, there was a slow negative shift that continued to increase in amplitude up to the end of the recording epoch at the three anterior sites but was more peaked at the three posterior locations. At all sites, the right hemisphere was more negative than the left [200-700-msec measure; main effect of hemisphere, $F(1,9) = 6.70$, $p < .029$]. In the post 600-msec epoch, the anterior sites continued to show a slow negative wave,

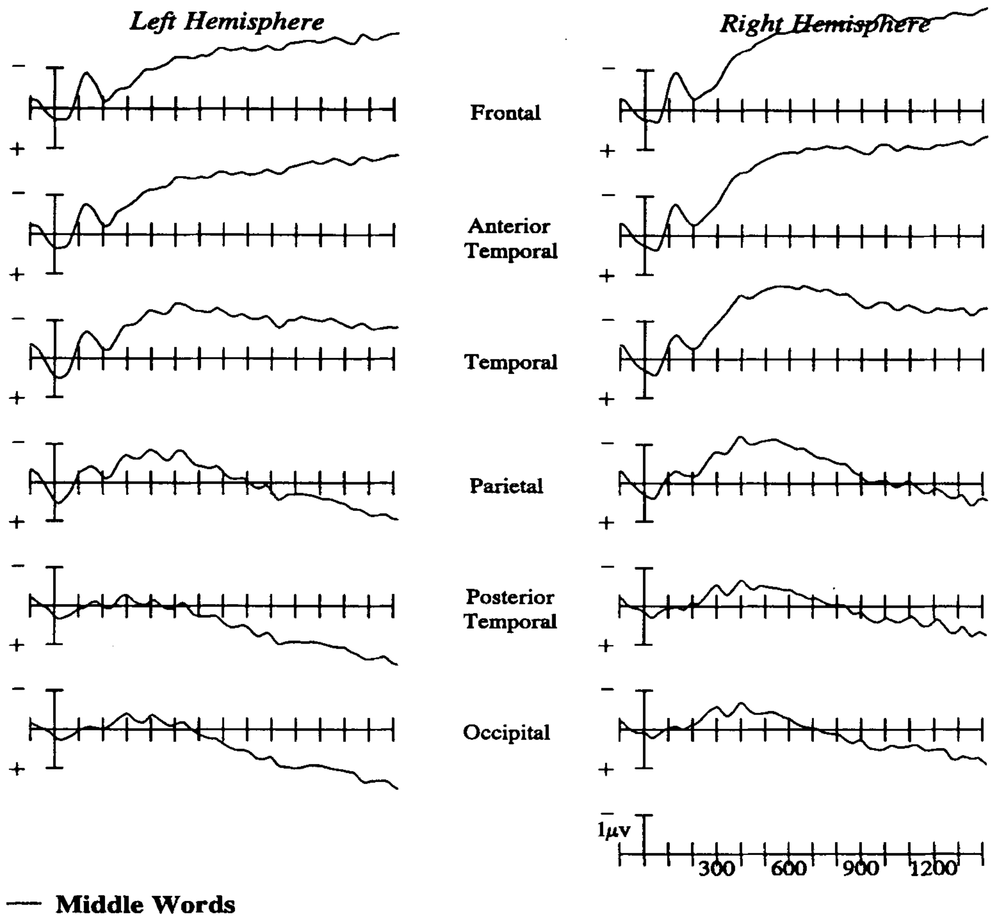


Figure 1. Grand mean ERPs from all the middle words (i.e., the second word through the second to last word) in each sentence of Experiment 1. The vertical calibration bar marks stimulus onset.

whereas the more posterior sites displayed a positive shift that was larger over the left hemisphere [700–1,400-msec measure; main effect of hemisphere, $F(1,9) = 13.01$, $p < .006$].

ERPs from final words. Figure 2 displays the ERPs elicited by the final words of each of the three sentence types. The final words also elicited an N140 component that was largest from over anterior sites. Moreover, analyses of the differences between the two hemispheres indicated that the N140 was more negative over the left hemisphere [main effect of hemisphere final words, $F(1,11) = 9.31$, $p < .011$].

As seen in Figure 2, the ERPs in the best-completion condition tended to go positive after the peak of N140, whereas ERPs in the unrelated-anomaly condition and, to a lesser extent, the related-anomaly condition continued to go negative for up to an additional 700 msec. At most sites, this negative-going wave was quite broad, but, at a few sites, it had a discernible peak around 300 or 400 msec. Its time course and response to final-word conditions suggest that this wave was related to the N400 component. By 600 msec, all but the most anterior sites (F7, F8, ATL, ATR) crossed the baseline to become positive and remained positive until the end of the recording epoch. Like the preceding negativity, this late positivity had a broad base, but, at the more posterior locations, there was a consistent peak between 800 and 1,000 msec in all three conditions. This pattern, together with the requirement that subjects respond to final words, suggests that this positive wave was related to the P3 component (see Regan, 1989).

The large negativity seen in the unrelated-anomaly and related-anomaly conditions and the positivity seen in the best-completion condition were quantified with an average area measure between 200 to 500 msec. The ANOVA on this measure confirmed that there was a main effect of conditions [$F(2,22) = 15.72, p < .0001$]. Planned contrasts revealed that waveforms in the unrelated-anomaly and related-anomaly conditions were more negative than those in the best-completion condition [$F(1,11) = 21.35$, $p < .0007$, and $F(1,11) = 16.74, p < .0018$, respectively], but that those in the unrelated-anomaly condition were only marginally more negative than those in the related-anomaly condition [$F(1,11) = 3.87, p < .075$].

A significant condition × electrode site interaction [$F(10,110) = 6.08, p < .0002$] indicated that the differ-

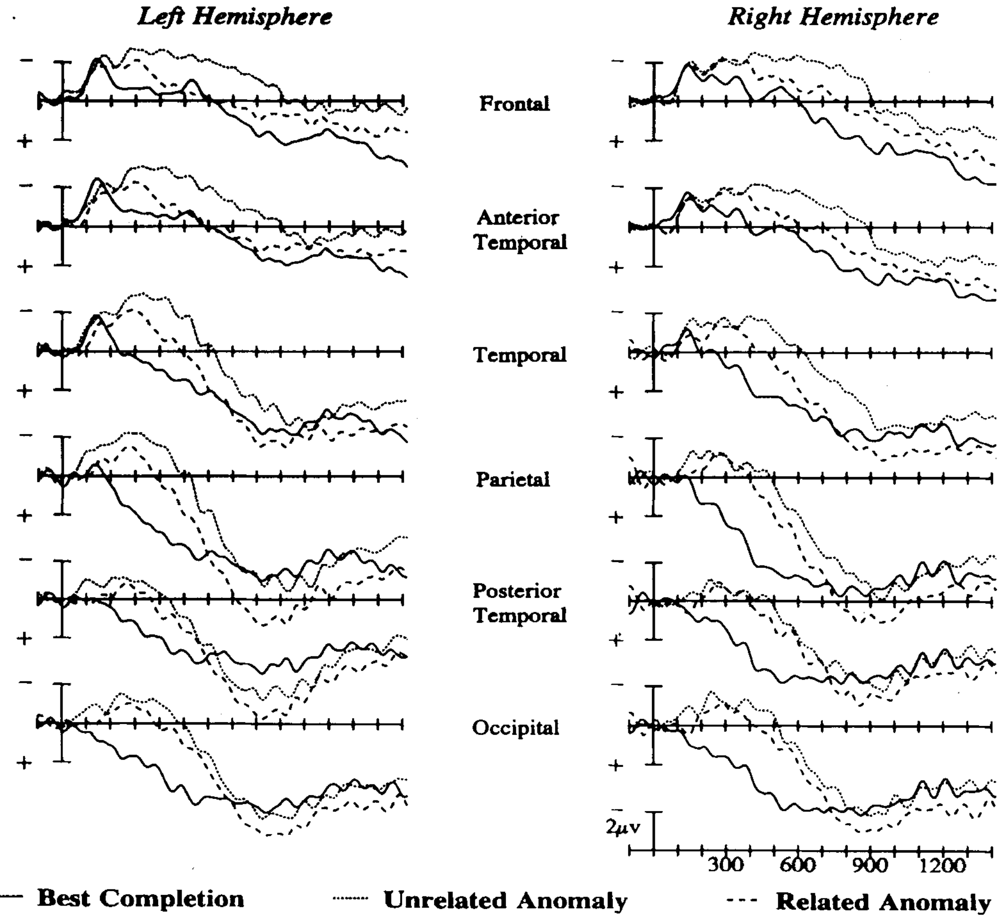


Figure 2. Grand mean ERPs from the three final-word conditions of Experiment 1. The vertical calibration bar marks stimulus onset.

ences between best completions and the two anomalous endings were largest over occipital, Wernicke's, and posterior temporal sites. This negative area displayed a marginal hemisphere × electrode site interaction [$F(5,55) = 2.82, p < .08$], indicating that the Wernicke's and temporal sites were asymmetric. Since previous studies of both auditory and visual word processing have reported asymmetries for negativities in this latency band over these sites (e.g., Holcomb & Neville, 1990; Neville, Kutas, Chesney, & Schmidt, 1986; Neville, Kutas, & Schmidt, 1982), a follow-up analysis of the Wernicke's and posterior temporal sites was performed. This analysis showed the left hemisphere was more negative than the right [$F(1,11) = 5.11, p < .045$] and that this effect was equivalent for all three final-word conditions. In other words, while overall the left hemisphere tended to be more negative than the right, the effects of context were symmetrical across the hemispheres.

The end of the negativity and the beginning of the subsequent positivity were quantified with an average area measure from 500 to 1,000 msec. Analyses of this measure also showed that the three types of sentence endings differed [main effect of condition, $F(2,22) = 6.40, p < .0064$], but more so at the anterior temporal and frontal electrodes [condition × electrode site interaction, $F(10,110) = 3.35, p < .027$]. Planned comparisons revealed that ERPs in the unrelated-anomaly condition remained more negative than ERPs in the best-completion and related-anomaly conditions [$F(1,11) = 11.19, p < .007$, and $F(1,11) = 8.83, p < .013$, respectively], but the related-anomaly and best-completion conditions no longer differed.

The area at the end of the recording epoch (1,000 to 1,400 msec) produced a significant condition × electrode site interaction [$F(10,110) = 2.92, p < .037$], indicating that, at posterior sites, there was very little difference between the conditions but that, at more anterior sites, ERPs in the unrelated-anomaly condition remained more negative than ERPs in either the best-completion condition or the related-anomaly condition.

Onset of the condition effects. As can be seen in Figure 2, differences between best completions and unrelated anomalies appear to onset quite early at some electrode sites. Analyses of the nine consecutive 50-msec epochs starting at stimulus onset confirmed this observation. The earliest significant condition × electrode site interaction

Table 1
N400 Onset Latencies

	BC vs. UA		BC vs. RA	
	LH	RH	LH	RH
Experiment 1 (Natural Speech)				
frontal	200	350	300	350
anterior temporal	200	350	300	350
temporal	200	300	250	250
parietal	150	150	250	250
posterior temporal	200	200	250	200
occipital	50	100	150	200
Experiment 2 (750-msec ISI)				
frontal	250	350	300	400
anterior temporal	200	300	250	250
temporal	200	250	250	250
parietal	200	200	250	300
posterior temporal	200	250	250	300
occipital	250	250	300	300

Note—N400 onset latency = beginning latency of the first significant ($p < .05$) 50-msec epoch. BC = best completions, UA = unrelated anomalies, RA = related anomalies, LH = left hemisphere, and RH = right hemisphere.

occurred for the 50 to 100-msec measure, showing the ERPs in the unrelated-anomaly condition to be more negative than in the best-completion condition [$F(5,55) = 5.05$, $p < .008$]. Table 1 reports the first time window during which each of the electrode sites differentiated between conditions (best completion vs. unrelated anomaly and best completion vs. related anomaly). As can be seen in Table 1, more posterior locations revealed differences earlier than did anterior sites, and the best-completion condition differentiated from the unrelated-anomaly condition earlier than it differentiated from the related-anomaly condition. Over anterior regions, the best-completion and unrelated-anomaly conditions were differentiated 100 to 150 msec earlier from the left hemisphere than from the right hemisphere (see Figure 2).

Discussion

Even though the average duration of final words was 561 msec (range 318-901 msec), the electrical activity of the brain (at posterior electrode sites) reliably registered a difference between sentence-final words that were contextually appropriate and those that were contextually anomalous as early as 50 msec after word onset. This result is consistent with Marslen-Wilson's (1987) claim that sentence-level contextual factors can have an effect on auditory word processing prior to the point at which all the acoustic information about a given word is available. This finding is particularly striking when compared with similar data from procedurally similar visual studies (e.g., Neville et al., 1986). In most visual studies, where all of the information about a given final word is available at stimulus onset, the difference between best-completion and anomalous final words typically does not start prior to 200 msec. It is noteworthy that while these early posterior effects of context were bilaterally symmetrical, in line with evidence for a greater role of the left hemisphere in speech processing, more anterior sites revealed an earlier difference between best completions and the two anomalies in the left hemisphere than in the right hemisphere.

The latency of the posterior context effects are 100-150 msec earlier than the onset of the first effects visible in the McCallum et al. (1984) auditory study. A likely possibility for this finding is the presence of nonsemantic contextual cues (i.e., prosody and across-word coarticulation) in the natural speech stimuli from the current experiment since McCallum et al. spliced in final words spoken in other contexts.

The results of this study more generally replicate McCallum et al.'s (1984) findings of an enhanced late negativity (N400) to semantically anomalous final words of spoken sentences and extend their results by demonstrating that final words—which, although anomalous, are semantically related to the best-completion word—generate an auditory N400 effect. However, the related-anomaly effect was both smaller and more restricted in latency than was the effect found for unrelated anomalies, suggesting that semantic contextual factors played a role in reducing the N400 effect to these otherwise anomalous items. This is the same basic pattern previously reported by Kutas et al. (1984) for similar materials presented visually and is consistent with their spreading-activation interpretation. However, Kutas et al. did not report a similar difference in the time course of the unrelated- and related-anomaly effects, although examination of their waveforms (see Figure 11.5, in Kutas et al., 1984) suggests that such differences may have been present at some scalp sites.

There are at least two other possibilities for the above pattern of results, both of which assume differences in the characteristics of the three types of final words. First, it is possible, that related anomalies had, on average, a later point of uniqueness than did best completions. This might have delayed the onset of differences between the best completions and related anomalies. Although we did not attempt to match our final words along this dimension, we feel that such differences are an unlikely source of the observed differences, because Marslen-Wilson (1987) has shown that the point of uniqueness becomes less critical in determining when a word is recognized (selected) when the word is placed within a sentence context. A second possibility is that the larger amplitude N400 in the related-anomaly condition, relative to the best-completion condition, could be due to the related anomalies' being of slightly lower frequency (1.77 vs. 1.42 log frequency). Van Petten and Kutas (1990) and Rugg (1990) have shown that lower frequency words have a larger N400 than do higher frequency words. However, such frequency effects are typically much smaller than context effects and tend to disappear by the end of contextually constrained sentences (Van Petten & Kutas, 1990). Therefore, it seems unlikely that small frequency differences seen in the present experiment could be contributing to the observed differences in final word ERPs.

Finally, the ERPs to the middle words in the sentences, although small in overall amplitude, revealed a clear set of early ERP components. That these waves represent the P1, N1, and P2 components frequently seen with other auditory stimuli is supported by their scalp distributions, latencies, and relative positions in the ERPs. In previous studies, these components have been shown to be closely tied to the physical parameters of the stimulus (see Regan, 1989). For example, previous work using pure tone stimuli has shown that both the N1 and P2 components are attenuated at short interstimulus intervals (ISIs) and do not begin to regain their full amplitude with ISIs shorter than several seconds (e.g., Gastaut, Gastaut, Roger, Corriol, & Naquet, 1951; Knight, Hillyard, Woods, & Neville, 1980). The smaller amplitude of the N1 and P2 to middle words in the present experiment may have been due to the relatively rapid rate of word presentation and the short duration between the end of one word and the beginning of the next.

The ERPs to words in the middle of sentences also generated a slow negative wave at anterior sites that started quite early, possibly overlapping the N140 and P200 (Figure 1). At posterior sites, there was a more peaked negativity with a latency of about 400 msec and a slow return to baseline. In both the 200–700-msec and the 700–1,400-msec measurements, anterior sites were more negative than posterior; however, both the anterior and the posterior sites were more negative from over the right hemisphere. At posterior locations, the more peaked nature of the negativity and its right hemisphere predominance suggest a similarity to the visual N400 component. Kutas, Van Petten, and Besson (1988) have looked at negativities across various positions in visual sentences in a variety of studies. Their work has shown a greater right than left distribution for N400 in almost every case.

EXPERIMENT 2

In Experiment 1, the words within each sentence were presented in a continuous stream of natural-sounding speech. This is the typical manner in which normal discourse is conducted, but, as pointed out, it may confound the effects of semantic factors with other between-word contextual influences. In Experiment 2, we attempted to limit these other sources by breaking up the natural flow of words in spoken sentences. This was accomplished by introducing a constant 750-msec ISI between the words of each sentence from Experiment 1. Pilot work with several subjects indicated that adding this time between words did not hinder subjects' comprehension, but that it did make the sentences sound less than natural. We assumed that, by adding this interval, semantic contextual influences would not be significantly altered, but that the effect of cues, which depend on the temporal/acoustic relationships between words in the sentence (prosody and between-word coarticulation), would be disrupted. We reasoned that, by comparing the pattern of ERP responses from the present experiment and Experiment 1, it should be possible to determine if this between-word information contributes to the ERP differences in the three final-word conditions.

A second purpose of Experiment 2 was to examine the ERPs to words in the middle of sentences when the interval between items was temporally extended and static. It was predicted that a similar pattern of early components to those seen in Experiment 1 would be obtained in Experiment 2, but that these would be less refractory (greater in amplitude) due to the longer and fixed interval.

Method

Subjects. Twelve young adults (7 male, 5 female; mean age = 21.6 years) were paid $5 per hour to serve as subjects. None had participated in Experiment 1. All were right-handed (4 had at least one left-handed family member within the immediate family) and native speakers of English.

Stimuli and Procedure. The stimuli were the same sequence of sentences used in Experiment 1. The only difference was that all words were presented with a 750-msec ISI. Prior to the experimental run, each subject was presented with 10 practice sentences using the same stimulus-presentation procedures.

Data analysis. The same components and measurement windows used for the ERPs in Experiment 1 were used in Experiment 2. Due to excessive artifact, the data from 2 subjects' middle words were not included in analyses.

In addition to the contrasts made within Experiment 2, several across-experiment comparisons were made on difference waveforms (formed by a procedure that involves subtracting ERPs from different conditions). One advantage of this procedure is that factors contributing to static morphological differences in the ERPs of the two experiments (e.g., physical differences in the stimulus conditions that might differentially affect components such as the N1) are removed, leaving only the effects of the three final-word conditions.

Two sets of "difference" ERPs were formed by subtracting the ERPs from the best-completion words from the ERPs to the unrelated anomalies and the ERPs to the best completions from the related anomalies. The mean amplitude between 0 and 200 msec, 200 and 400 msec, 400 and 600 msec, and 600 and 1,000 msec, as well as the peak latency between 200 and 800 msec, were then measured. For within-experiment comparisons, the factors were electrode site, hemisphere, and condition (unrelated anomaly-best completion vs. related anomaly-best completion). Comparison of the two experiments added one between-subject factor (experiment number).

Results

Behavior. The subjects reported no difficulty in understanding the sentences in Experiment 2 (best completions = 96% correct, unrelated anomalies = 98% correct, and related anomalies = 96% correct). There were no significant differences in percent correct scores between Experiment 2 and Experiment 1.

ERPs from middle words. Plotted in Figure 3 are the ERPs from the middle words of Experiment 2. Apparent are the larger and more clearly defined early components to words presented at this rate, relative to the natural-speech results (Figure 1). An early positivity was apparent just after stimulus onset and peaked between 40 and 50 msec (P45). Note that, unlike Experiment 1, P45 did not appear to be overlapped by components from the prior

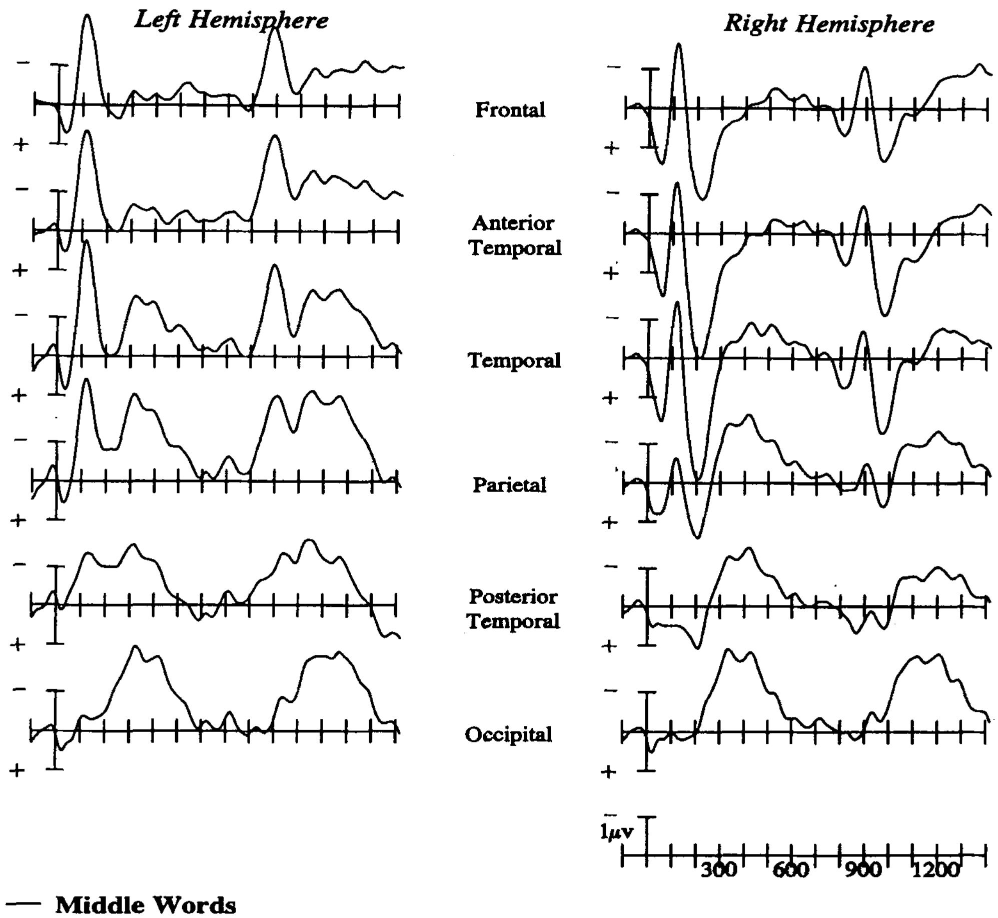


Figure 3. Grand mean ERPs from all the middle words (i.e., the second word through the second to last word) in each sentence of Experiment 2. The vertical calibration bar marks stimulus onset.

word. P45 can be seen at all sites and, although there was no significant difference along the anterior-posterior dimension, there was a significant hemisphere effect [middle words, $F(1,9) = 9.88, p < .012$], indicating that the right hemisphere was more positive than the left. The analysis of the latency of P45 indicated that, as in Experiment 1, this component peaked earlier in the left hemisphere for both middle and final words (36 vs. 48 msec middle words and 47 vs. 55 msec final words) [main effect of hemisphere: middle, $F(1,9) = 9.90, p < .012$; final, $F(1,11) = 10.07, p < .009$].

Similar to Experiment 1, the most prominent feature early in the waveforms of both middle and final words was a negative component that peaked between 112 and 130 msec (because of its similarity to the N140 in Experiment 1, we will continue to refer to this negativity by this label). N140 can be seen as far posterior as T5/T6, but as in Experiment 1, it was larger at the more anterior electrode locations [main effect of electrode site: middle words, $F(5,45) = 7.87, p < .01$]. As in Experiment 1, the amplitude of N140 was larger over the left hemisphere than over the right hemisphere [electrode site × hemisphere interaction: final words, $F(5,55) = 4.0, p < .005$; middle words, $F(5,45) = 6.68, p < .006$].

As in Experiment 1, following the N140, ERPs to middle words were characterized by a positivity, which peaked at about 200 msec (P200) and was largest over right frontal and temporal sites [hemisphere × electrode site interaction, $F(5,45) = 6.02, p < .009$]. Unlike in Experiment 1, ERPs to middle words did not produce a large frontal negative wave following the P200, but they did reveal a negativity between 300 and 600 msec over temporal and posterior sites [main effect of electrode site, $F(5,45) = 5.55, p < .032$]. Note that because of the constant ISI between words in Experiment 2, the N140/P200 to the next word can be seen starting at about 800 msec (Figure 3).

ERPs from final words. The ERPs to the final words in the three conditions are shown in Figure 4. As in Experiment 1, an N140 can be seen in the ERPs of all three final-word conditions; however, in contrast to Experiment 1, there do not appear to be any differences between the conditions in this time period. Between 200 and 300 msec, the ERPs in the best-completion condition were markedly

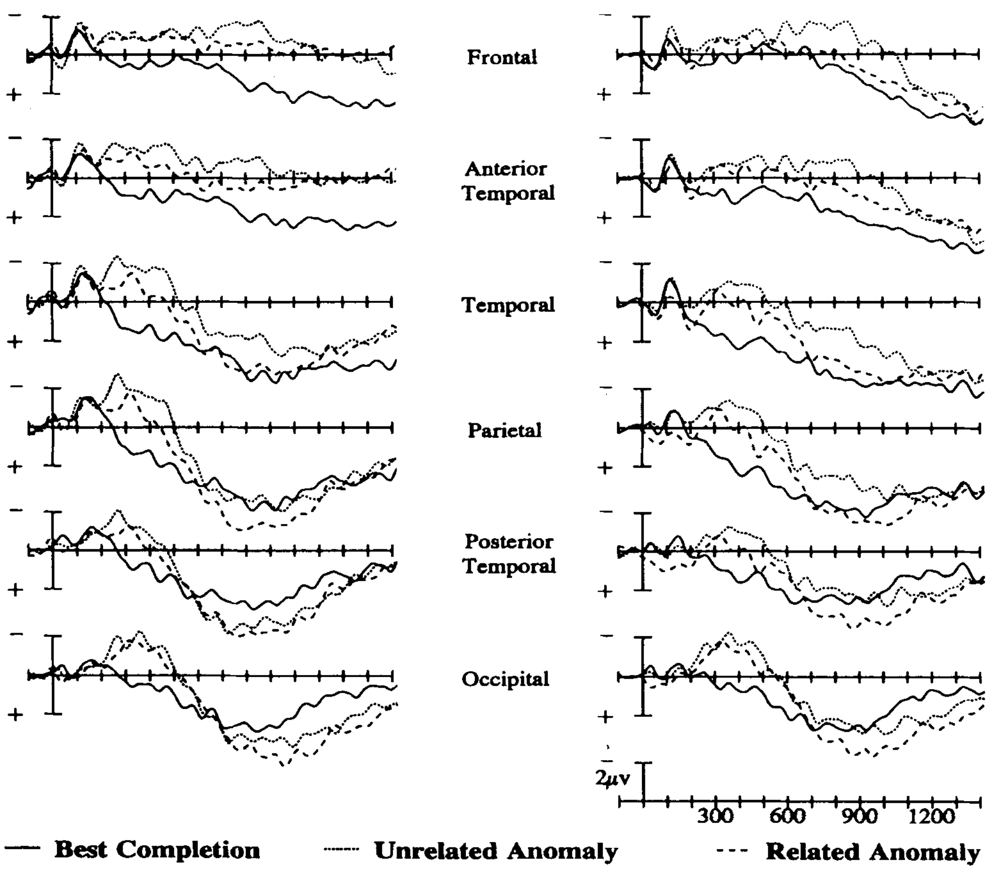


Figure 4. Grand mean ERPs from the three final-word conditions of Experiment 2. The vertical calibration bar marks stimulus onset.

divergent from those in the unrelated- and related-anomaly conditions at most sites. As in Experiment 1, the best-completion condition displayed a positive shift, whereas both anomaly conditions were characterized by a negative component. As in Experiment 1, by 600 msec, all but the most anterior sites crossed the baseline to become positive and remained positive until the end of the recording epoch.

Analyses on the 200 to 500-msec measure confirmed that the three conditions differed [main effect of condition, $F(2,22) = 21.12, p < .0001$]. Planned contrasts revealed a similar pattern of effects as was observed in Experiment 1: the unrelated- and related-anomaly conditions produced more negative-going ERPs than did the best-completion condition [$F(1,11) = 41.51, p < .0001$, and $F(1,11) = 18.24, p < .0013$, respectively], but the unrelated-anomaly condition produced only marginally more negative-going ERPs than did the related-anomaly condition [$F(1,11) = 4.78, p < .051$]. As in Experiment 1, there was a hemisphere × electrode site interaction [$F(5,55) = 4.02, p < .006$], which indicated that at frontal, temporal, and parietal sites, the left hemisphere was more negative than the right. There was not, however, a significant interaction between hemisphere and the three conditions, suggesting that, as in Experiment 1, the condition effects were symmetrical across the hemispheres.

The 500–1,000-msec area measure was used to quantify the end of the negativity and the beginning of the subsequent late positivity. Measures of this epoch revealed a condition × electrode site interaction [$F(10,110) = 3.77, p < .047$], indicating that, over posterior sites (O, P, T, W), there were no significant differences between the three conditions but that, at the more anterior sites (T, AT, and F), the unrelated-anomaly condition was more negative than the best-completion condition, and the related-anomaly condition was intermediate between the other two. Planned comparisons confirmed this observation. The unrelated-anomaly/best-completion contrast produced a significant electrode site × condition interaction [$F(5,55) = 8.80, p < .004$; unrelated anomalies more negative than best completions]. The unrelated-anomaly/related-anomaly contrast produced a main effect of condition [$F(1,11) = 10.54, p < .008$; unrelated anomalies more negative than related anomalies]. However, the related-anomaly/best-completion contrast did not produce either effect.

The area at the end of the recording epoch (1,000–1,400 msec) also produced a significant condition × electrode site interaction [$F(10,110) = 8.77$, $p < .002$], indicating that, at the anterior temporal and frontal sites, ERPs in the unrelated-anomaly condition and, to a lesser extent, those in the related-anomaly condition, were more negative than those in the best-completion condition. A significant hemisphere × electrode site interaction indicated that, over anterior sites, the left hemisphere was more negative than the right [$F(5,55) = 6.79$, $p < .002$].

Onset effects. We compared condition effects for each of nine 50-msec epochs (beginning at word onset) to determine the onset of the priming effects. The earliest reliable differences between conditions occurred from 200 to 250 msec (i.e., 150 msec later than in Experiment 1; see Table 1). Also, the onset of the differences between conditions was more similar from anterior and posterior locations of the scalp than in Experiment 1. However, the same tendency for the effects to onset earlier in regions of the left hemisphere than in regions of the right hemisphere is evident. Additionally, as in Experiment 1, the best-completion/unrelated-anomaly differences occurred earlier than did the best-completion/related-anomaly differences.

Difference waves and comparison of Experiments 1 and 2. Figure 5 shows the difference waves (Figure 5A, unrelated anomalies minus best completions; Figure 5B, related anomalies minus best completions) from Experiments 1 and 2. In the natural-speech experiment, but not in the 750-msec experiment, there were clear context effects present in the first 200 msec following word onset, but only over posterior locations [electrode site × experiment interaction, $F(5,110) = 3.40$, $p < .05$; see Figure 5A]. In both experiments, from 200 to 400 msec, context effects were asymmetrical over anterior sites and were larger from the left hemisphere [hemisphere × electrode site interaction, $F(5,110) = 3.05$, $p < .04$].

In the natural-speech experiment, the maximum negativity occurred between 400 and 600 msec and was larger from over posterior regions [main effect of electrode site, $F(5,110) = 7.2$, $p < .004$]. The final phase (600–1,400 msec) was largest from anterior electrodes [main effect of electrode site, $F(5,110) = 30.8$, $p < .001$]. Beyond 200 msec, the results from Experiment 2 were similar in many respects. The overall amplitude and the anterior/posterior and lateral distribution of the difference waves were similar to those observed in Experiment 1. As noted above, the effect from 200 to 400 msec was larger from over anterior regions of the left hemisphere than the right hemisphere. The later (>600 msec) effects were larger from anterior regions [main effect of electrode site, $F(5,110) = 13.0$, $p < .001$]. Over temporal and parietal regions, these effects were asymmetrical, larger from over the right hemisphere [hemisphere × electrode site interaction, $F(5,110) = 5.50$, $p < .007$].

In the interval between 200 and 800 msec, the mean peak latency of the largest negativity was 444 msec in Experiment 1 and 474 msec in Experiment 2. However, neither this difference nor its interaction with other variables was significant. Across experiments, the latency of the peak negativity was earlier at posterior sites than at anterior sites [main effect of electrode site, $F(5,110) = 15.73$, $p < .0001$] and occurred earlier in the left hemisphere than in the right hemisphere, but most notably in the unrelated-anomaly–best-completion waves [condition × hemisphere interaction, $F(1,22) = 6.20$, $p < .021$].

Discussion

The condition effects on the ERPs to final words in Experiment 2 were similar to those observed in Experiment 1, the major difference being that they began later. Unlike in Experiment 1, the differences between the final-word conditions in Experiment 2 did not start in the time frame of the N140 component, but, as indicated by the time-course analysis, they started later (220–300 msec), particularly at posterior sites. The later phase of the negativity (N400) did not appear to differ between the experiments.

As in Experiment 1, the ERPs in the unrelated- and related-anomaly conditions were significantly more negative between 200 and 500 msec than those in the best-completion condition; however, between 500 and 1,000 msec, the unrelated-anomaly condition was more negative than either the best-completion or the related-anomaly condition.

The longer interval between stimuli in Experiment 2 resulted in a similar, but more clearly differentiated, set of early components, relative to that seen in Experiment 1 (P45, N140, and P200). In both experiments, P45 could be seen with words in the middle of sentences and was earlier in peak latency in the left hemisphere. But asymmetries in amplitude of early components (P45, N140, and P200), where the left hemisphere was more negative than the right, occurred only in Experiment 2. We have previously reported the tendencies for the left hemisphere to be more negative than the right during the first 300 msec following word onset in both visual and auditory studies (e.g., Holcomb & Neville, 1990; Neville et al., 1986; Neville et al., 1982). Each of these studies involved relatively long ISIs. In Experiment 1, presentation rate averaged two words per second, with very little time between the offset of words. The absence of an asymmetry to the middle words of Experiment 1 may have been due to the presence of an overlapping component that was more negative over the right hemisphere (i.e., two opposite asymmetries may have canceled). The 200–700-msec and 700–1,400-msec epochs in Experiment 1 revealed a slow anteriorly distributed negative wave that was larger over the right hemisphere.

GENERAL DISCUSSION

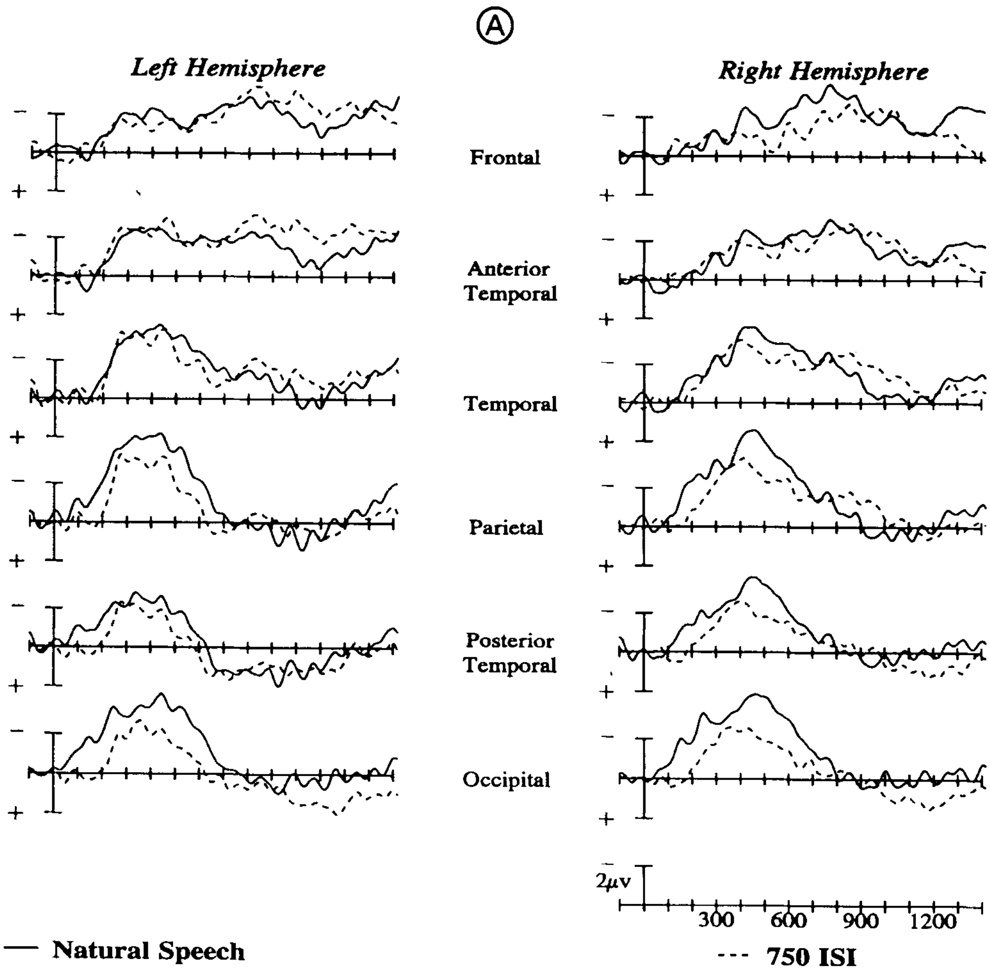
Middle Words

The presence of a slow right-hemisphere negativity in Experiment 1 and its absence in Experiment 2 suggest that this wave has something to do either with the rate of stimulus presentation or with the "naturalness" of the sentence. A number of studies have indicated that the right hemisphere plays a role in processing prosodic information (e.g., Behrens, 1988; Ross, 1981; Ross, Edmondson, Seibert, & Homan, 1988). Presumably, prosodic cues were diminished or removed by the artificial ISI introduced in Experiment 2. The right-more-negative-than-left asymmetry at anterior sites in Experiment 1 and the lack of a similar symmetry in Experiment 2 suggest that this negativity may reflect the greater role of the right hemisphere in processing prosodic cues in natural speech. However, the presence of similarly distributed effects in visual sentences (Kutas et al., 1988) casts some doubt on this hypothesis. Another possibility is that latency-jittered N400s from subsequent middle words contributed more to right-hemisphere sites (remember that subsequent middle words were not systematically time locked to the current middle word in Experiment 1 due to the natural speaking rate). However, this implies that N400s were, in general, more negative over the right hemisphere, as they typically are in visual studies. This was not the case in either of the present experiments. These and other speculations about middle word auditory ERPs need to be tested more directly in future studies.

The existence of a relatively normal set of early ERP components time locked to the onset of words in natural continuous speech is consistent with the hypothesis that speech segmentation occurs on line and at a relatively early point in the processing of speech stimuli. However, whether these ERP findings directly reflect the segmentation process at work or whether they indicate that segmentation is complete prior to the occurrence of these early waves will have to await the results of future research. For example, it would be interesting to see if continuous speech produces discernible "middle latency" ERPs (those with a latency between 10 and 50 msec). It would also be interesting to see if time locking to other features within words (e.g., stressed vs. unstressed syllables) results in identifiable ERP components.

Final Words

The earlier onset of the N400 effect (Figure 5) in Experiment 1, relative to that in Experiment 2, is consistent with the hypothesis that nonsemantic between-word contextual cues facilitated the processing of sentence-final words beyond that produced by semantic context effects. Evidence for this position comes from at least three sources. First, studies using reading tasks (e.g., Kutas & Hillyard, 1980), where these types of between-word



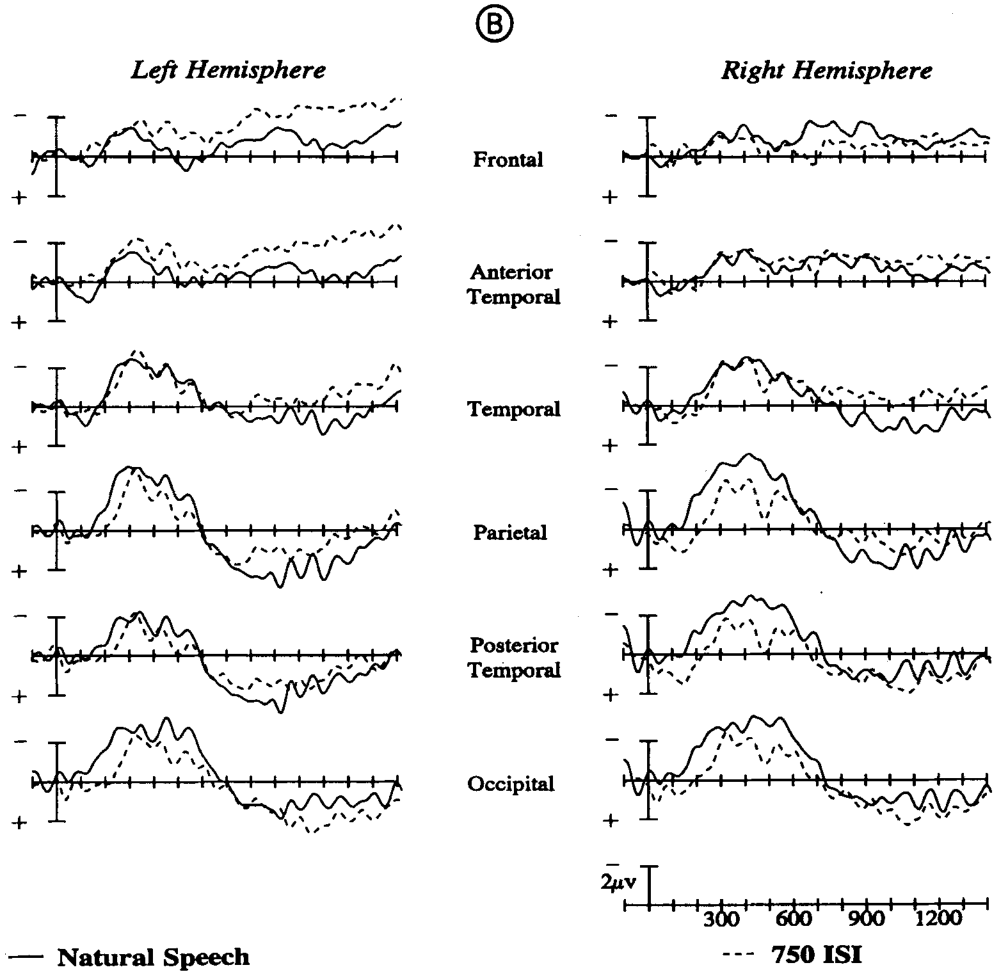


Figure 5. Grand mean difference ERPs from Experiment 1 (solid line) and Experiment 2 (dashed line). Panel A displays the unrelated-anomalies-minus-best-completions waves. Panel B displays the related-anomalies-minus-best-completions waves. The vertical calibration bar marks stimulus onset.

cues are not present, have shown a later onset for the N400.[7] Second, McCallum et al. (1984) removed or attenuated these between-word cues and their N400s onset substantially later than did those in Experiment 1. Finally, Holcomb and Neville (1990) found differences in ERPs recorded to auditory and visual word pairs. However, although their auditory semantic-priming effects were relatively early (200–290 msec auditory and 300–360 msec visual), they were not as early as those found in Experiment 1 (50 to 100 msec). One difference between their study and the present experiments is that their words were presented (and originally pronounced) in isolation. Therefore, the prime word (the first word in each pair) could not have provided prosodic or coarticulatory cues to facilitate the processing of the target word (the second word of each pair). Taken together, these results suggest that there was an interaction between the semantic contextual information and certain nonsemantic between-word cues and that, together, these sources provided rapid information about the final words in the present sentence paradigm after a relatively small amount of the word had been heard.

It should be pointed out that it is unlikely that the between-word nonsemantic cues alone would have been sufficient to determine whether the sentence-final word was the correct item (and therefore drive the time course of the early ERP effect), since even in the unrelated-anomaly sentences, the final two words were spoken together (naturally) and were always a perfectly plausible pair of words for ending a sentence (however, not the given sentence).

There is another possible explanation for the early onset of the context effect. In addition to providing between-word nonsemantic contextual cues, the stimuli in this study were also presented at a relatively high rate (an average of two words per second, which is a relatively normal speaking rate and is only high in comparison with most other ERP language studies). One line of evidence in support of an early rate effect comes from a recent visual study (Neville, Pratarelli, & Forster, 1989). Briefly, under conditions in which related and unrelated word pairs were presented very rapidly, an early (100–200 msec) priming effect was evident. However, this early effect displayed an anterior/posterior and lateral distribution very different from the later, typical visual N400 effect. Another study casts doubt on the rate hypothesis. Kutas (1987) presented visual sentences to subjects at either 10 words per second or one word every 700 msec. The peak latency and the time course of the ERP difference between anomalous final words and best-completion final words were substantially earlier in the slower presentation condition than in the 10-per-second condition.

Clearly, more work is needed in this area. For instance, it is important to know the relative contribution of both prosodic and coarticulatory cues. An experiment in which one or both of these variables could be factorially manipulated would be helpful. Also further studies looking at variations in speaking rate are critical to determine whether this variable does in fact interact in any way with the effects of context.

REFERENCES

Behrens, S. J. (1988). The role of the right hemisphere in the production of linguistic stress. *Brain & Language*, **33**, 104-127.

Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 133-163). Hillsdale, NJ: Erlbaum.

Cutler, A. (1987). Speaking for listening. In A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading and writing* (pp. 23-40). New York: Academic Press.

Cutler, A. (1989). Auditory lexical access: Where do we start? In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 342-356). Cambridge, MA: MIT Press.

Ellman, J. L., & McClelland, J. L. (1987). Speech perception as a cognitive process: The interactive activation model. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 337-374). New York: Academic Press.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory & Language*, **25**, 75-87.

Fischler, I., & Bloom, P. A. (1985). Effects of constraint and validity of sentence contexts on lexical decisions. *Memory & Cognition*, **13**, 128-139.

Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Copper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 27-85). Hillsdale, NJ: Erlbaum.

Frazier, L. (1987). Structure in auditory word recognition. *Cognition*, **25**, 157-189.

Gastaut, H., Gastaut, Y., Roger, A., Corriol, J., & Naquet, R. (1951). Étude electroencephalographique du cycle d'excitabilité cortical. *Electroencephalography & Clinical Neurophysiology*, **3**, 401-428.

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

Hansen, J. C., Dickstein, P. W., Berka, C., & Hillyard, S. A. (1983). Event-related potentials during selective attention to speech sounds. *Biological Psychology*, **16**, 211-224.

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

Kleiman, G. M. (1980). Sentence frame contexts and lexical decisions: Sentence-acceptability and word-relatedness effects. *Memory & Cognition*, **8**, 336-344.

Knight, R. T., Hillyard, S. A., Woods, D. L., & Neville, H. J. (1980). The effects of frontal and temporal lesions on the auditory evoked potential in man. *Electroencephalography & Clinical Neurophysiology*, **50**, 112-124.

Kutas, M. (1987). Event-related brain potentials (ERPs) elicited during rapid serial visual presentation of congruous and incongruous sentences. In R. Johnston, J. W. Rohrbaugh, & R. Parasuraman (Eds.), *Current trends in event-related potential research* (pp. 406-411). New York: Elsevier.

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

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

Kutas, M., Lindamood, T., & Hillyard, S. A. (1984). Word expectancy and event-related brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 217-238). Hillsdale, NJ: Erlbaum.

Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during reading of sentences. *Electroencephalography & Clinical Neurophysiology*, **69**, 218-233.

Marslen-Wilson, W. D. (1980). Speech understanding as a psychological process. In J. C. Simpson (Ed.), *Spoken language generation and understanding* (pp. 39-67). Dordrecht: Reidel.

Marslen-Wilson, W. D. (1984). Function and process in spoken word recognition: A tutorial review. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance: Control of language processes* (pp. 125-150). Hillsdale, NJ: Erlbaum.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**, 71-102.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1-71.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, **10**, 29-63.

McCallum, W. C., Farmer, S. F., & Pocock, P. K. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography & Clinical Neurophysiology*, **59**, 447-488.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, **76**, 165-178.

Neville, H. J., Kutas, M., Chesney, G., & Schmidt, A. (1986). Event-related brain potentials during initial encoding and recognition memory of congruous and incongruous words. *Journal of Memory & Language*, **25**, 75-92.

Neville, H. J., Kutas, M., & Schmidt, A. (1982). Event-related potentials studies of cerebral specialization during reading: II. Studies of congenitally deaf adults. *Brain & Language*, **16**, 316-337.

Neville, H. J., Pratarelli, M. E., & Forster, K. I. (1989). Distinct neural systems for lexical and episodic representations of words. *Society for Neuroscience Abstracts*, **15**, 246.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials: I. Evaluation of components. *Electroencephalography & Clinical Neurophysiology*, **36**, 179-190.

Pisoni, D. B., & Luce, P. A. (1987). Acoustic-phonetic representation in word recognition. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 21-52). Cambridge, MA: MIT Press.

Regan, D. (1989). *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*. New York: Elsevier.

Ross, E. (1981). The aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology*, **38**, 561-589.

Ross, E., Edmondson, J. A., Seibert, G. B., & Homan, R. W. (1988). Acoustic analysis of affective prosody during right-sided Wada test: A within-subjects verification of the right hemisphere's role in language. *Brain & Language*, **33**, 128-145.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory & Cognition*, **18**, 367-379.

Schustack, M. W., Enrlich, S. F., & Rayner, K. (1987). Local and global sources of contextual facilitation in reading. *Journal of Memory & Language*, **26**, 322-340.

Tyler, L. K., & Frauenfelder, U. H. (1987). The process of spoken word recognition: An introduction. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 1-20). Cambridge, MA: MIT Press.

Tyler, L. K., & Marslen-Wilson, W. D. (1982). Speech comprehension process. In J. Mehler, E. Walker, & M. Garrett (Eds.), *Perspective on cognitive representations* (pp. 169-184). Hillsdale, NJ: Erlbaum.

Van Petten, C., & Kutas, M. (1990). Interaction between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, **18**, 380-393.

VAN PETTEN, C., & KUTAS, M. (1991). Electrophysiological evidence for the flexibility of lexical processing. In G. Simpson (Ed.), *Word and sentence* (pp. 129-174). Amsterdam: North-Holland.

WALKER, S. (1987). The evolution and dissolution of language. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 3, pp. 5-47). Hillsdale, NJ: Erlbaum.

NOTES

1. By convention, ERP components are labeled with a prefix indicating their polarity (N or P) and a suffix that indicates either the peak latency of the component (poststimulus onset) or the relative position of the component in the waveform. For example, P250 is a positive component with a peak latency at 250 msec, and N2 is the second negative component in the waveform.

2. The cloze procedure involves having a separate group of subjects fill in the final word of each sentence and then calculating the percentage that used the same word.

3. In 29 of the sentences from this condition, the related but anomalous final word was also semantically related to one or more words within the prior sentence context. In the remaining 16 sentences, the final word was related to only the appropriate best-completion word (e.g., "The game was called when it started to *umbrella*"). A similar mixing of sentences with one or more associative primes and sentences in which only the best completion was related to the anomalous ending was also employed by Kleiman (1980) and Kutas et al. (1984). Given the relatively small number of the later type of sentence, we were unable to examine this as a separate subcondition. However, recent data from a study by Van Petten and Kutas (1991) suggest that these two types of sentences should not produce large differences in the N400.

4. In the eight cases of anomalies that shared a sound with the original best completion, this was always the word-initial phoneme (e.g., "He had a toothache so he called the *drum*," where *drum* and the original best completion [*dentist*] share the same first, but no subsequent, phonemes).

5. WL and WR were located 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz. Left and right posterior temporal were located 33% of the interaural distance lateral to Cz. Left and right anterior temporal were located one half of the distance between F7 or F8 and T3 or T4.

6. Although ERPs were also collected from two midline sites (Cz and Pz), these data will not be discussed here because of the almost doubling of the number of analyses that would have to be presented (midline sites should be analyzed separately from lateral sites). By and large, the results from these sites were consistent with those from the lateral electrodes.

7. It should be noted that we do not have a good feel for the effects of natural reading on the time course of the N400 (or any other ERP). This is because of the great difficulties encountered in trying to control (or remove) eye-movement artifacts during reading. It might well be the case that a similar early N400 onset would be apparent in normal reading.

(Manuscript received April 2, 1991;
revision accepted for publication July 25, 1991.)