Memory & Cognition
1990, 18 (4), 380-393

Interactions between sentence context and word frequency in event-related brain potentials

CYMA VAN PETTEN and MARTA KUTAS
University of California, La Jolla, California

Event-related brain potentials (ERPs) were recorded as subjects silently read a set of unrelated sentences. The ERP responses elicited by open-class words were sorted according to word frequency and the ordinal position of the eliciting word within its sentence. We observed a strong inverse correlation between sentence position and the amplitude of the N400 component of the ERP. In addition, we found that less frequent words were associated with larger N400s than were more frequent words, but only if the eliciting words occurred early in their respective sentences. We take this interaction between sentence position and word frequency as evidence that frequency does not play a mandatory role in word recognition, but can be superseded by the contextual constraint provided by a sentence.

The frequency of a word's occurrence in common usage and the relationship of a word to prior context have proven to be two of the most powerful determinants of performance in experimental studies of word recognition. In the laboratory, subjects typically require either more time or more information to respond to rare words than they do to respond to common words (e.g., Rubinstein, Garfield, & Millikan, 1970; Solomon & Howes, 1951). Similarly, subjects are faster or more accurate to respond to words that are congruent with the preceding context (Fischler & Bloom, 1979; Kleiman, 1980; Morton, 1964; Stanovich & West, 1981, 1983; Tulving & Gold, 1963; Underwood & Bargh, 1982). These two facts have been accorded a prominent place in current models of word recognition because, to date, they have provided the most replicable clues to the organization of memory for words and/or the processing strategies applied to text and speech.

The relationship between frequency and context has, however, been handled differently by various theorists. Serial search models, by their essence, postulate that either word frequency or context, but not both, can be utilized at any given moment during the word recognition process. By contrast, parallel models allow the possibility that both types of information can be simultaneously effective. Within each of these two broad classes of models are those that predict an additive relationship between frequency and context in experimental situations, as well as those that predict an interactive relationship.

The present study focuses on the relationship between sentence context and word frequency during silent reading. We begin by summarizing the theoretical and empirical work that served as a framework for our investigations.

A brief report of this study was presented at the Fourth International Conference on Cognitive Neuroscience, Dourdan, France, in June 1987. The work was supported by NSF Grant BNS83-09243 and NICHD Grant HD 22614. C. Van Petten was supported by an NSF Graduate Fellowship. M. Kutas is supported by NIH Research Career Development Award MH 00322. Correspondence may be addressed to Cyma Van Petten, Department of Neurosciences, M-008, University of California, San Diego, La Jolla, CA 92093.

Models of Frequency and Context

In Forster's serial search model, a partial analysis of sensory input is used to search a number of candidate words for a match with the to-be-recognized item. The search is conducted along an invariant path proceeding from the most frequent to the least frequent of the candidate words. Neither the meaning of the word in question nor its relationship to prior context enters into this procedure. Contextual integration is presumed to occur subsequent to the location of the appropriate lexical entry and, furthermore, takes place in a different "processor." In early versions of this model (see Forster, 1981a, 1981b), there was a provision for contextual influence on the search path under very limited circumstances; strong lexical associates (e.g., "doctor" if "nurse" was the target word) could short-circuit the frequency-dependent mechanism. A later version of this model has excluded this feature (Bradley & Forster, 1987). In either case, this serial search model draws a distinction between comprehension and sentence processing on the one hand, and lexical processing on the other. In their most recent formulation, Bradley and Forster (1987) state that "the assumption of a frequency-ordered examination of candidate representations should be seen as the keystone of search theory: to abandon this assumption amounts to abandoning the whole notion of a search retrieval system" (p. 119). Thus, the model predicts additivity of frequency effects and context effects.

In contrast, Becker (1976, 1979, 1980) has also outlined a serial search model that predicts interactions between word frequency and semantic context. This is accomplished by postulating that the search path may be ordered *either* by word frequency or by strength of relationship to prior context. As in Forster's model, the initial sensory analysis of a letter string yields a set of candidate words that are ordered by frequency. However, before the letter string is actually encountered, the reader is able to begin forming a "semantic set" of candidates based on the prior context: this is ordered by strength of semantic relationship rather than by frequency of usage. Because readers search the semantic set first, frequency effects will obtain only when the stimulus word is not included (i.e., not related to the previous context) and the reader is forced to check the "sensory set." Note that if the context is not predictive of a specific word, preferential search through the semantic set could result in slower word recognition than if there had been no context (Becker, 1976). The plausibility of this model seems dependent on the assumption that, in general, semantic sets will include fewer words than will sensory sets. If this were not the case, Becker's model would be inefficient wherever the range of meaningful subsequent words is only moderately restricted by preceding words, as in reading connected text.

The *logogen* model of Morton (1969) differs from search theories in postulating a large number of passive word-detector elements (logogens) that can accrue information or activation from a number of sources in parallel. Frequency and context both act to reduce the amount of stimulus information required to exceed a logogen's threshold: frequency by lowering the threshold, and prior context by raising the level of activation. Because the two factors act independently to influence a common final mechanism, this model predicts that context and frequency are additive.

Finally, there are two additional "parallel" models that allow for interactions of context and frequency. Both the revised *cohort* model of Marslen-Wilson (1987) and the *checking* model of Norris (1986) specify that low-frequency words approach their recognition thresholds more slowly than do high-frequency words. This prolonged recognition process for low-frequency words thus allows more time for contextual constraints to come into play. Despite the differences in their terminologies and hypothetical loci of context and frequency effects, the two models resemble each other in that they both propose that baseline differences in processing speed make low-frequency words more susceptible to context effects. Stanovich and West (1981) proposed a similar account of the context × frequency interactions on naming latency observed in their studies of sentence contexts.

Combined Effects of Frequency and Context

Empirical results concerning the combined effects of frequency and context have been too inconsistent to constrain theoretical models. Cairns and Foss (1971) first suggested that sentence structure might play a role in determining the importance of word frequency after demonstrating that phoneme monitoring times were prolonged following low-frequency words only when these words were adjectives, and not when they were nouns or verbs. Cairns and Foss raised the possibility that selectional restrictions imposed during their sentences made nouns and verbs more predictable than adjectives and thereby less subject to frequency effects.

More recent studies have focused on the overall constraints, both semantic and syntactic, that a sentence imposes on its final word. Stanovich and West (1981) recorded naming latencies for *difficult* (i.e., lower frequency, longer, less predictable) and *easy* (i.e., higher frequency, shorter, more predictable) words, which formed congruous or incongruous completions to sentence fragments. They found a substantially larger congruity effect for difficult words than for easy words. However, Forster (1981b) did not replicate this interaction with the same stimulus materials.

Similarly, Schuberth and Eimas (1977) observed purely additive effects of sentence congruity and sentence-terminal word frequency, whereas Becker (1979) observed frequency × context interactions for word-pair stimuli. Becker (1979) criticized the Schuberth and Eimas study for inadequate semantic rating procedures and repetition of sentence contexts. In response, Schuberth, Spoehr, and Lane (1981) reported two follow-up experiments that corrected these deficiencies.

In the first of these experiments, words of varying frequencies were used to complete sentence fragments that had been rated as exercising a high, moderate, or low degree of constraint on their completions. The effect of frequency on lexical decision times was estimated by linear regressions in the three conditions. No differences were found in the slopes of these lines. This analysis, however, is based on the assumption that the effect of word frequency is linear across the broad range of frequencies used (approximately 0 to 385 occurrences per million in the Kučera and Francis, 1967, count). Given that there were only 28 sentences in each context condition, the frequency extremes were probably represented by very few terminal words. A disproportionate effect of context on very low-frequency words relative to moderate- and high-frequency words might thus have gone undetected in such an analysis. The second experiment of this investigation included a stronger comparison: 21 sentence/terminal word combinations factorially arranged by word frequency (high or low) and sentence congruity. For these sentences, the congruity effect for low-frequency words was nearly double that for high-frequency words—an effect that was statistically significant in an analysis across subjects, but not across items.

The relationship between frequency and context has also been examined in auditory word recognition. These experiments employed a gating procedure in which subjects heard successively larger fragments of a word and attempted to identify it. These researchers thus determined how much sensory input was required to identify high- and low-frequency words given different amounts of preceding semantic context. Grosjean (1980) initially reported that, although low-frequency words required more sensory input, there was no interaction between frequency and contextual constraint. In a subsequent study (Grosjean & Itzler, 1984), this experiment was expanded to include more constraining contexts. In the latter experiment, it was found that the difference between high- and low-frequency words was much reduced under higher levels of context. Tyler (1985) compared a condition of no semantic context to one of minimal context and reported no interaction, suggesting that fairly constraining contexts are required if frequency × context interactions are to be found.

The Present Study

It is difficult to base strong conclusions on such mixed results. Hence, the empirical question of whether contextual constraint and word frequency have additive or interactive effects on word recognition is in need of more data. Ideally, such data should be obtained via techniques that are sensitive to the psycholinguistic variables of interest and relatively insensitive to other factors. Balota and Chumbley (1984, 1985) have demonstrated that the two reaction-time techniques utilized in the studies cited above, pronunciation and lexical decision, are both associated with task-specific frequency effects that are not likely to occur in normal language processing. Since all techniques in the study of language processing may be plagued by specific factors, a number of convergent measures, each with its unique drawbacks, are more likely to yield a valid outcome than is any single measure. In the present study, we employ the event-related brain potential (ERP) technique. ERPs recorded during silent reading may also be subject to technique-specific factors extraneous to the cognitive operations involved in lexical access or word recognition, but these are likely to be distinct from the confounds introduced by the decision stage of the lexical decision task or the production stage of the pronunciation task.

A second desirable criterion for experimental results bearing on the relationship between word frequency and context is that they derive from large sets of experimental stimuli. This is particularly important when one relies on normative frequency counts to estimate an underlying variable, such as a particular subject's familiarity with a particular word. Both Gernsbacher (1984) and Gordon (1985) have found that subjective familiarity is more predictive of response speed than is printed word frequency. The larger a set of experimental stimuli, the more likely this subjective variable will approach the normative frequency. Larger sets of experimental stimuli are also less likely to yield idiosyncratic results peculiar to a few words in the English language and more likely to reveal general factors in word recognition.

Limitations on the number of data points in any one experiment have generally arisen from the fact that subjects cannot be asked to respond to every word they read, although each word is associated with some level of contextual constraint and some frequency of usage. The possibility of obtaining a metric of frequency and context effects in the absence of an overt behavioral task is thus attractive. The measurement of eye fixations during natural reading holds promise in this regard, because both predictability and frequency decrease gaze duration (Inhoff, 1983; Just & Carpenter, 1980). However, to our knowledge, investigators using this technique have not discussed the relationship between these two variables. The analysis of ERPs, like that of eye movements, can provide a record of subjects' responses to every word within an experiment without the intrusive requirement of a manual or verbal response. The present study exploits this possibility to obtain a relatively large data base to examine the relationship between context and frequency.

The N400 component of the ERP has proved to be a sensitive index of contextual factors in visual and auditory language processing (see Holcomb, 1985; Kutas, Neville, & Holcomb, 1987; McCallum, Farmer, & Pocock, 1984). This component characterizes the ERP elicited by open-class words (nouns, verbs, adjectives, and some adverbs) in contrast to closed-class words (articles, prepositions, pronouns, quantifiers, etc.) that elicit little N400 activity (Kutas & Hillyard, 1983; Kutas, Van Petten, & Besson, 1988). The amplitude of the N400 elicited by a given word is reduced by preceding semantic context, whether this takes the form of a sentence fragment, phrase, or single word (Bentin, McCarthy, & Wood, 1985; Kutas & Hillyard, 1980a, 1980b, 1980c, 1983; Neville, Kutas, Chesney, & Schmidt, 1986). The amplitude of the N400 is inversely proportional to off-line measures of semantic constraint, such as a cloze-probability rating (Kutas & Hillyard, 1984; Kutas, Lindamood, & Hillyard, 1984). Word repetition and phonological priming also have been reported to influence late negative components in the ERP that bear a resemblance to the N400 elicited by semantically unrelated words (Kramer & Donchin, 1987; Rugg, 1985; Smith & Halgren, 1987).

Recently, we determined that the amplitude of the N400 can be influenced by context *within*, as well as at the end of, sentences. In particular, we found that frequency-matched words elicit larger N400s when they occurred early, as opposed to late, in sentences (Kutas, Van Petten, & Besson, 1988). Analogous word-position effects have been observed in the latencies to detect a particular phoneme, word, or nonword embedded within a sentence (Foss, 1969; Marslen-Wilson & Tyler, 1980; Sanocki et al., 1985). We have interpreted the amplitude decrement of the N400 across the course of a sentence as an on-line reflection of the buildup of constraints imposed by the sentence upon individual succeeding words. In the present paper, we expand upon our earlier results by analyzing the ERPs to intermediate sentence words on a position-by-position basis. Moreover, we use the word-position effect as a vehicle to examine the relationship between sentence structure and word frequency.

EXPERIMENT 1

Methods

Materials. Stimulus materials consisted of 338 unrelated sentences of variable length (M = 9.3 words, SD = 2.6). Terminal words were predictable or semantically constrained, as assessed by a cloze-probability procedure (Taylor, 1953) using a minimum of 25 subjects per sentence. The mean cloze probability was 0.83 ($SD = 0.17$). A subset of the sentences are included in Appendix 1; others were drawn from Bloom and Fischler (1980).

Subjects. The subjects were 43 young adults (26 males, 17 females), ranging in age from 18 to 29 years, who were paid for their participation. All were native English speakers with normal or corrected-to-normal vision. Most were students or staff at the university. Two were left-handed; 5 of the right-handed subjects had a left-handed relative in their immediate families.

Procedure. The subjects were tested individually in a single session that lasted between 3 and 3.5 h. Each subject was seated in a comfortable chair in a sound-attentuated chamber facing a CRT screen controlled by an Apple II microcomputer. The CRT screen was approximately 100 cm away; at this distance, each letter in a word subtended 0.36° of visual angle horizontally. Sentences were presented one word at a time in the center of the screen for a duration of 200 msec. A blank interval of 700 msec between words resulted in a total stimulus-onset asynchrony (SOA) of 900 msec. The subjects were asked to confine eye and body movements to the 6-sec interval provided between sentences.

Sentences were presented in sets of 20, with 4- to 8-min breaks between sets and a 20- to 30-min break halfway through the experiment. The subjects were instructed to read each sentence for meaning. In addition, they were told that a few duplicate sentences had been included and that, as a check on their alertness over the course of the experiment, they should report whether or not a set had included a repeat upon its termination. The subjects did quite well at detecting the 10 repeated sentences. None of these sentences were included in the frequency or word-position analyses.

Following the end of each sentence, a probe word and a single letter of the alphabet were presented sequentially. For 28 subjects, the probe word was presented 200 msec after the sentence-terminal word; for the remaining subjects, it was presented 700 msec after the sentence-terminal word. In each case, the single letter followed the probe word by 2 sec. For both groups of subjects, the task was to press one of two response buttons, on the basis of whether or not the letter had been present in the probe word. Neither probe-word nor letter-search data are directly relevant to the present analysis and will be discussed elsewhere. However, of concern here is that, for the short-SOA group, it was not possible to analyze frequency effects in the ERP elicited by the terminal word due to the overlap of the ERP elicited by the probe. Thus, our analyses of terminal-word ERPs were restricted to the 15 subjects in the long-SOA group.

Recording system. The electroencephalogram (EEG) was recorded from scalp sites defined by the 10-20 system (Jasper, 1958), with tin electrodes mounted in an elastic cap (Electro-Cap International, see Polich & Lawson, 1985). Midline frontal (Fz), central (Cz), and parietal (Pz) recording sites were used, along with lateral pairs of electrodes over the frontal (F3, F4), central (C3, C4), parietal (P3, P4), temporal (T5, T6), and occipital (O1, O2) scalp, each referred to the left mastoid. Vertical eye movements and blinks were monitored via an electrode placed below the right eye referred to the left mastoid. Horizontal eye movements were monitored via a right-to-left bipolar montage at the external canthi. Activity over the right mastoid was also recorded.

The EEG was amplified by a Grass Model 12 Neurodata Acquisition System with half-amplitude cutoffs of 0.01 and 30 Hz, digitized on-line at a sampling rate of 250 Hz, and stored on magnetic tape along with stimulus codes for subsequent averaging.

Data analysis. Open-class words (nouns, verbs, adjectives, and "ly" adverbs) were categorized according to sentence position and word frequency.[1] For the purpose of analyzing interactions between sentence structure and word frequency, these were initially divided into three broadly defined categories based on their positions within the sentence: *first*, *middle*, and *final* OC words. The *first OC word* category contained the first open-class word from each sentence. Since all sentences began with a closed-class word, words in the *first OC word* category typically occupied the second or third word position of the sentences. The *middle OC word* category included open-class words from all other sentence positions except the last. Sentence-terminal words comprised the *final OC word* category. Each sentence-position category was subdivided into six groups, according to the summed frequency of usage of all regularly inflected word forms (Francis & Kučera, 1982): 0 to 9, 10 to 29, 30 to 99, 100 to 199, 200 to 450, and greater than 450 occurrences per million. The frequency of irregular verb forms (paid, hung, ate, etc.) was taken as the frequency of occurrence of this form. Each of the categories, including the lowest frequency, consisted of words that anyone with some college education could be expected to know (see Appendix 2). The mean frequency, length in letters, and number of words included in each sentence position × word frequency subgroup are detailed in Table 1.

For a more fine-grained analysis of the effect of word position independent of frequency, middle OC words were divided further, according to their actual position within sentences (e.g., third, fourth, etc.).

Finally, intermediate open-class words (middle OC words) were subdivided into those that had immediately followed another open-class word and those that had followed immediately upon a closed-class word. The aim of this categorization scheme was to determine whether any observed effects of word position might be due to physiological refractoriness of the N400 component of the ERP rather than to cognitively mediated effects of sentence structure. Since closed-class words typically elicit little or no N400 activity, it makes sense that the N400 response to an open-class word would be more vulnerable to refractory period effects if it followed another N400-eliciting word than if the brain cells that generate this response had not been active in response to the immediately preceding (closed-class) word.[2] Comparing the responses to open-class words con-

Table 1
Characteristics of the Words in Experiment 1

Frequency Range	Mean Frequency	Mean Log Frequency	Mean No. of Letters	No. of Words
		First		
0–9	3.5	0.55	6.8	39
10–29	17	1.23	6.4	25
30–99	64	1.81	6.2	70
100–199	145	2.16	5.5	54
200–450	318	2.50	5.3	65
over 450	756	2.88	5.0	53
		Middle		
0–9	4.4	0.64	6.3	73
10–29	19	1.28	5.8	95
30–99	59	1.77	5.8	168
100–199	143	2.16	5.3	141
200–450	295	2.47	5.2	138
over 450	841	2.92	4.7	116
		Final		
0–9	5.8	0.76	5.0	24
10–29	19	1.28	5.0	59
30–99	58	1.76	4.8	98
100–199	144	2.16	4.8	85
200–450	287	2.46	4.5	40
over 450	817	2.91	4.6	33

Note—First = the first open-class word in each sentence; Middle = all of the other intermediate open-class words; Final = the sentence-terminal words. The column headed "No. of Words" lists the number of words in each frequency × position category.

tingent on the lexical class of the preceding word thus allowed us to determine whether the N400 was refractory at the SOAs used in the current experiment.

ERP averages were formed for each of the stimulus categories described, beginning 100 msec before word onset, and ending 700 msec after word onset. Trials containing excessive eye movements, amplifier blocking, or other artifacts were rejected prior to averaging (approximately 17% of the trials were lost; there was little variability in rejection rate across the experimental conditions).

Results and Discussion

The ERPs elicited by the first open-class words, intermediate open-class words, and sentence-terminal words at a central midline site are shown in Figure 1. Visual inspection of the waveforms indicates the presence of a broad negative-going wave peaking around 400 msec in response to words in the *first* and *middle* open-class categories, but not to those in the *final OC* category. The first open-class word of each sentence elicited a larger N400 than did the later intermediate words. The absence of an N400 for sentence-terminal words replicates past results showing that the terminal words of high cloze-probability sentences elicit a late positivity rather than an N400 (Kutas & Hillyard, 1984; Kutas, Lindamood, & Hillyard, 1984).

Sentence position and word frequency: Amplitude effects. The amplitude of the N400 was quantified by measuring the mean voltage in the peak latency range (350-500 msec poststimulus) relative to the 100-msec baseline preceding stimulus onset. The amplitude measures were taken at central, parietal, temporal, and occipital electrode sites for each subject and word category. These were analyzed by a three-way repeated measures analysis of variance (ANOVA), using word position (first or middle), frequency (6 levels), and electrode site (10 levels) as factors. There were significant main effects of word position [$F(1,42) = 115.6$, $p < .001$] and frequency [$F(5,210) = 6.9$, $p < .001$], as well as an interaction between position and frequency [$F(5,210) = 6.4$, $p < .001$]. The interaction suggests that word frequency influenced N400 amplitude for first OC words, but not for middle OC words (see Figure 2).

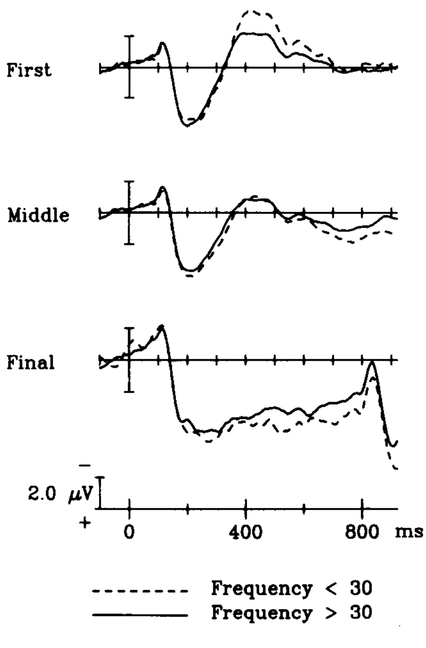

Figure 1. ERPs elicited by high- and low-frequency words in different sentence positions. Responses to first and middle open-class words are shown for a grand average of 43 subjects; the sentence-final ERPs are from a grand average of 15 subjects. All recordings were taken from a central midline scalp site (Cz) in Experiment 1.

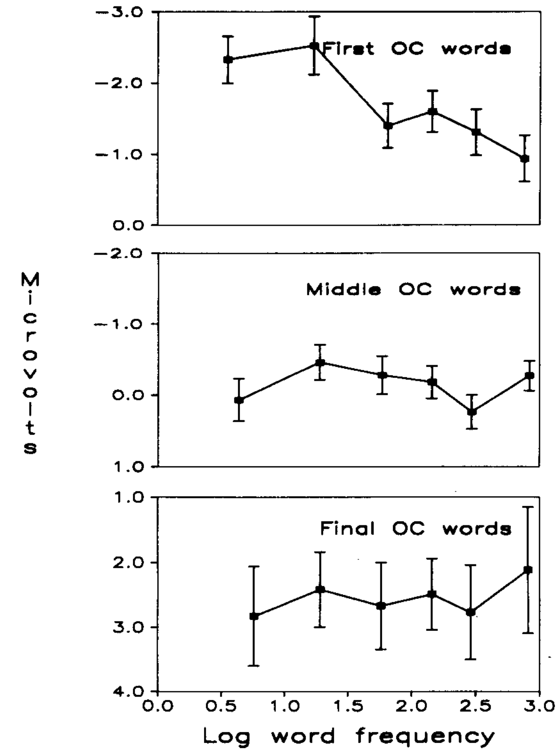

Figure 2. Mean voltage during the peak latency range of the N400 (350–500 msec poststimulus onset) relative to a 100-msec prestimulus baseline, collapsed across recording site in Experiment 1. Word frequency has been broken down into six categories for each of three sentence positions: the first open-class words, other intermediate open-class words, and sentence-final open-class words. The error bars represent the standard error across subjects.

Trend analyses were used to determine if N400 amplitude, collapsed over electrode site, was influenced by frequency in a linear fashion. The mean log frequencies of the six word-frequency subdivisions were used in this analysis, since several experiments have shown that reaction times vary with log, rather than raw, word frequency (Gordon, 1985; Hudson & Bergman, 1985; Whaley, 1978). For first OC words, the linear component was significant [$F(1,42) = 20.9$, $p < .001$]. However, the data presented in Figure 2 suggest that the influence of frequency on N400 amplitude may not be truly linear over the entire frequency range; rather, there appears to be a difference between the two lowest frequency divisions (less than 30 occurrences per million) and the four

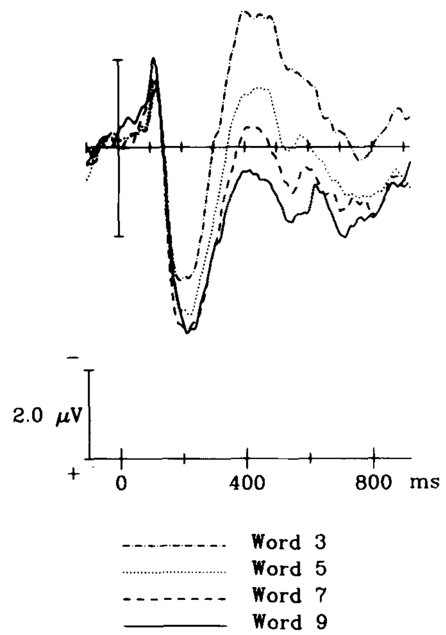

Figure 3. Grand average ERPs elicited by intermediate open-class words in four different sentence positions. Recorded at a central midline scalp site (Cz) in Experiment 1. Note that, although the late negative components elicited at each word position all reach peak amplitude around 400 msec, the differentiation between word positions is apparent as early as 200 msec poststimulus onset.

highest frequency divisions (greater than 30 occurrences per million). In fact, there was no significant linear trend across the four highest frequency divisions [$F(1,42) = 1.92$, n.s.]. A post hoc Scheffé test confirmed that there was a difference between word frequencies of greater and less than 30 ($p < .001$). For intermediate open-class words, neither a trend analysis over the six frequency divisions nor the Scheffé comparison revealed any significant effects of frequency.

Separate analyses were performed on the ERPs from the 15 subjects for whom the ERP to sentence-terminal words could be isolated. The small number of subjects did not allow a detailed breakdown of word frequency. On the basis of the outcome of our preceding analyses, we compared ERP responses to words with frequencies greater than 30 with those of less than 30. There were significant main effects of position [$F(2,28) = 8.25, p < .001$] and frequency [$F(1,14) = 5.97, p < .03$], and an interaction between the two factors [$F(2,28) = 4.83, p < .02$]. Separate ANOVAs showed that the frequency effect was restricted to the first open-class words [first OC words, $F(1,14) = 16.2, p < .001$; middle OC words, $F(1,14) = 0.35$, n.s.; final OC words, $F(1,14) = 0.46$, n.s.].

The interaction between frequency and position in the preceding analyses allows for the possibility that the main effect of position was due solely to the low-frequency words. However, an analysis restricted to high-frequency words (frequency greater than 30) also yielded a main effect of sentence position [$F(1,42) = 56.7, p < .001$]. A comparison of the first, middle, and sentence-terminal words in the subset of 15 subjects where sentence-terminal word ERPs could be analyzed yielded a similar main effect of position [$F(1,14) = 33.5, p < .001$].

Sentence position within intermediate words. We examined the effect of word position on N400 amplitude in greater detail by subdividing the *middle OC word* category into eight groups based on each word's actual position in the sentences from positions 3 to 10. These groups were collapsed across frequency. The amplitude of the N400 for intermediate word positions is shown in Figures 3 and 4. A trend analysis showed a significant

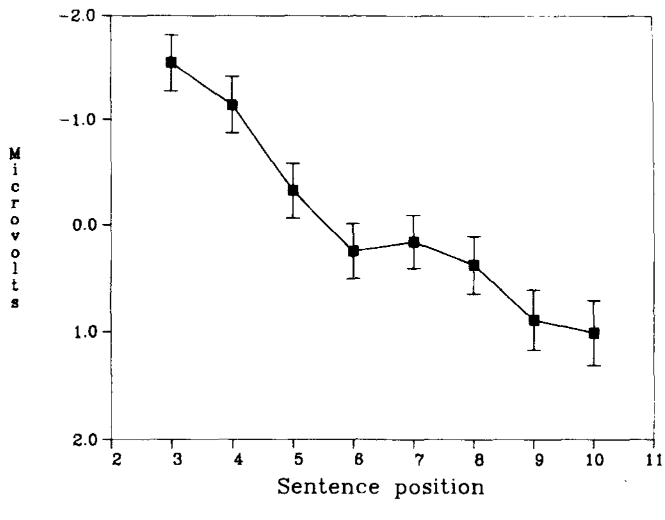

Figure 4. Mean voltage during the peak latency range of the N400 (350–500 msec poststimulus onset) relative to a 100-msec prestimulus baseline, collapsed across recording site in Experiment 1. The measures were taken from ERPs to intermediate open-class words based on sentence position. Neither the first nor the final open-class of the sentences are included.

linear component to N400 amplitude as a function of sentence position [$F(1,42) = 240.0$, $p < .001$].

Refractory effects. We examined the amplitude of the N400 elicited by intermediate open-class words as a function of the lexical class (open or closed) of the immediately preceding word. This analysis was conducted on open-class words occurring in sentence positions 6, 7, and 8, because these were approximately equally divided in terms of the lexical class of the preceding word and represent the flattest portion of the word-position curve. As seen in Figure 5, N400 amplitude was unaffected by lexical class of the preceding word [$F(1,42) = 0.67$, n.s.].

This negative result indicates that we can dismiss one possible technique-specific confound in the interpretation of the word-position effects observed here and in a previous study (Kutas et al., 1988). Some ERP components are highly sensitive to the temporal interval between eliciting stimuli. In the present experiment, some of the middle OC words followed other N400-eliciting words at an interval of only 0.9 sec, whereas others (those immediately preceded by a closed-class word) followed at an interval of at least 1.8 sec. If these timing differences were important in determining N400 amplitude, then the N400 response to the intermediate open-class words immediately following other open-class words would have been relatively depressed. That this was not the case indicates that the N400 has a relatively short refractory period, ruling out one explanation for the amplitude decrement across the course of a sentence.[3]

Word-length effects. The covariation of word frequency with word length has been a persistent problem in attempts to find pure frequency effects in English (Hudson & Bergman, 1985). Because we did not attempt to control word length in Experiment 1, the low-frequency words contained more letters than high-frequency words. The low-frequency words in the *first OC word* category also tended to be longer than the low-frequency words in the *middle OC word* category (see Table 1). Word length was therefore a possible confound for the observed interaction between frequency and sentence position. We thus examined another set of data to obtain independent estimates of frequency effects on the ERP by comparing responses to words that were matched for length. The data reported in Experiment 2 were derived from a previous experiment (Van Petten & Kutas, 1987), in which ERPs elicited by sentence-terminal words have been reported.

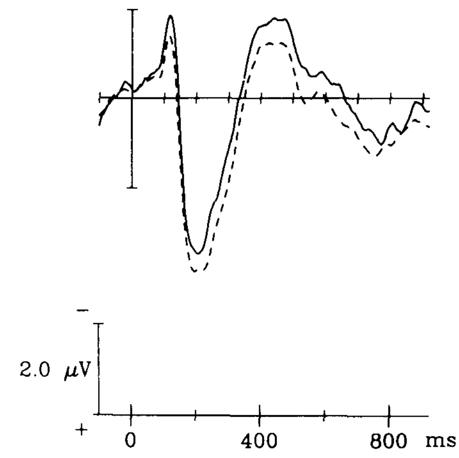

Figure 5. Grand average ERPs elicited by intermediate open-class words according to the lexical class of the immediately preceding word. Recorded at a midline central site (Cz) in Experiment 1.

EXPERIMENT 2

Methods

Materials. The stimuli consisted of 240 unrelated sentences of variable length ($M = 9.6$ words, $SD = 3.0$).

Subjects. The subjects were 31 young adults (20 males, 11 females), ranging in age from 18 to 29 years, who were paid for their participation. All were native English speakers with normal or corrected-to-normal vision. All were right-handed; 5 had left-handed family members. Two additional subjects participated in Experiment 2, but their data were excluded from analysis due to eye-movement artifacts in more than 50% of the EEG epochs following intermediate sentence words. None of the subjects had participated in Experiment 1.

Procedure. The details of stimulus presentation, task, and instructions were very similar to those of Experiment 1. The one exception was that there were no repeated sentences. Rather, the subjects were encouraged to read for meaning; they were told that there would be a questionnaire concerning the sentence contents at the end of the experiment.

The subjects sat approximately 100 cm from the CRT screen so that each letter in a word subtended 0.36° of visual angle horizontally. The shortest open-class words (three letters) therefore occupied about 1°; an eight-letter word would subtend approximately 3°.

Recording system. EEG activity was recorded from 10 silver-silver chloride electrodes, each referred to an average of the activity recorded at the left and right mastoids. Eight electrodes were placed according to the International 10-20 system at frontal (Fz), central (Cz), parietal (Pz), and occipital (Oz) midline locations, as well as at frontal and central lateral sites (F3, F4, C3, C4). Symmetrical posterior temporal electrodes were placed lateral and posterior to the vertex (30% of the interaural distance lateral, and 12.5% of nasion-inion distance posterior). These were approximately over Wernicke's area (WL) and its right hemisphere homologue (WR). Eye movements were monitored as in Experiment 1.

The midline and EOG recordings were amplified with Grass 7P122 preamplifiers (system bandpass 0.01 to 35 Hz, half-amplitude cutoff). The EEG from the lateral scalp leads was amplified with Grass 7P511 preamplifiers modified to have an 8-sec time constant (high-frequency half-amplitude cutoff = 60 Hz).

Data analysis. Open-class words of six or seven letters were categorized by two levels of sentence position (first and middle OC words) and two levels of word frequency (greater or less than 30 per million). Words of six or seven letters were selected for this comparison because they were the most common. Although responses to both longer and shorter words were averaged, many of these individual length × sentence position × word frequency categories contained too few trials to form reliable ERPs. Word-length effects were examined by averaging intermediate open-class

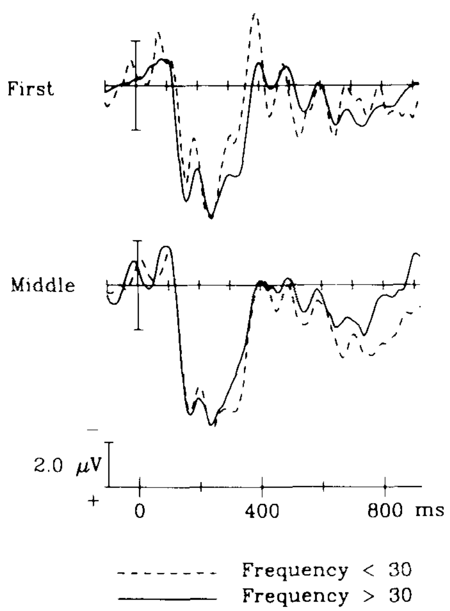

Figure 6. Grand average ERPs elicited by high- and low-frequency words occurring as the first or intermediate open-class words in sentences. All words were six or seven letters long. Recorded at a midline parietal site (Pz) in Experiment 2.

words according to number of letters: three letters, four or five, six or seven, and eight or more.[4]

Analog-to-digital conversion of the EEG, EOG, and stimulus trigger codes was performed on-line by a PDP 11/45 computer at a sampling rate of 250 Hz. A 1,024-msec epoch of EEG, beginning 100 msec before the onset of word presentation, was averaged. Trials characterized by excessive eye movement or amplifier blocking were rejected (approximately 15% of the trials).

Results and Discussion

When matched for length, low-frequency words occurring as the first open-class words of their respective sentences elicited larger N400s than did high-frequency words. In contrast, there was no frequency effect for middle open-class words (see Figure 6).

As in Experiment 1, the N400 was quantified as the mean voltage from 350–500 msec poststimulus, relative to a 100-msec prestimulus baseline at central, parietal, and temporal electrode sites (Cz, Pz, C3, C4, WL, and WR). There was a significant main effect of frequency for first OC words [$F(1,30) = 5.10$, $p < .05$]. In contrast, there was no frequency effect for middle OC words [$F(1,30) = 1.91$, n.s.].

ERPs elicited by words of different lengths are shown in Figure 7. There were two distinct effects of increasing word length on the ERP: (1) the enhancement of a negativity, peaking around 200 msec after stimulus onset at parietal and temporal electrodes (i.e., N200), and (2) a broad positivity, beginning at about 250 msec and extending throughout the recording epoch at some scalp sites. The latter effect was most evident at anterior scalp sites.

The peak amplitude of the N200 was quantified as the most negative point in a latency range of 150–225 msec poststimulus at the posterior electrode sites where it occurred (Pz, WL, WR). For middle open-class words of frequency greater than 30, an ANOVA with length (4 levels) and electrode site (3 levels) as factors showed a significant effect of length [$F(3,54) = 6.2$, $p < .001$]. The inference that physical size, rather than some other aspect of word length, was the critical factor in the N200 enhancement is consistent with the spatial locus of the N200 over visual cortical areas and with other studies suggesting that the posterior negativity reflects visuoperceptual, rather than linguistic, operations (Mangun & Hillyard, 1987; Neville & Lawson, 1987; Rugg, Milner, & Lines, 1985).

The broad positivity apparent in the ERPs elicited by long words was quantified as the mean voltage from 250–600 msec poststimulus at all electrode sites. An ANOVA with length (4 levels) and electrode site (10 levels) as factors showed a significant effect of length [$F(3,90) = 8.7$, $p < .001$] and an interaction between length and electrode site [$F(27,810) = 3.1$, $p < .001$]. The interaction reflects the gradient of the length effect across the anterior–posterior dimension; it was largest at frontal and central scalp sites, intermediate at parietal and temporal sites, and absent at the occipital site (see Table 2). This scalp distribution is important in that it is distinct from that of other ERP components that occur in the same latency range. The anterior maximum resembles

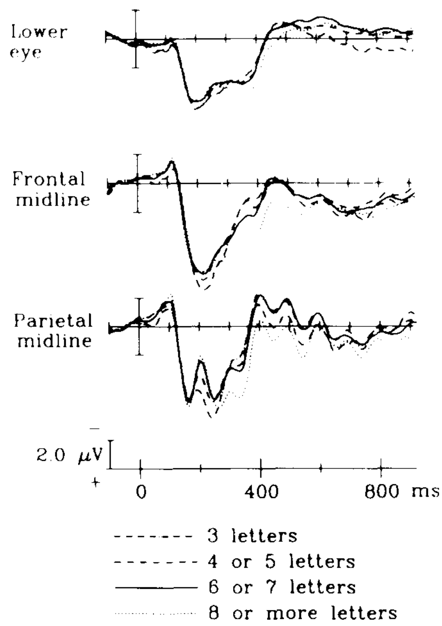

Figure 7. Grand average ERPs elicited by high-frequency (greater than 30 per million) intermediate open-class words of different lengths in Experiment 2. Shown are three different recording sites: below the eye, frontal midline (Fz), and parietal midline (Pz). It is important to note that the broad positivity elicited by the longest words is also apparent at the lower eye site. This rules out the possibility that it was generated by electroculographic, rather than brain, activity; vertical eye movements create potentials of opposite polarity at lower eye and frontal sites with the recording montage used here.

Table 2
Effect of Word Length at Different Scalp Sites

	Number of Letters							
	3		4-5		6-7		8 or more	
Scalp Site	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Midline								
Frontal	2.3	0.4	1.8	0.4	2.2	0.5	3.9	0.5
Central	2.7	0.5	1.7	0.5	2.0	0.5	3.8	0.7
Parietal	1.9	0.7	1.4	0.4	1.6	0.5	2.9	0.7
Occipital	−0.1	0.6	−0.6	0.5	−0.2	0.5	−0.1	0.7
Left								
Frontal	1.7	0.4	1.7	0.4	1.9	0.5	3.3	0.5
Central	2.4	0.4	2.0	0.4	2.2	0.4	3.8	0.5
Wernicke	1.9	0.5	1.6	0.3	1.7	0.3	2.6	0.5
Right								
Frontal	1.7	0.4	1.8	0.4	2.2	0.5	3.5	0.5
Central	2.5	0.6	1.9	0.4	2.1	0.4	3.6	0.7
Wernicke	1.6	0.6	1.1	0.4	1.5	0.4	2.1	0.6

Note—Shown are means and standard errors for middle open-class words with frequency of usage greater than 30/million in Experiment 2. The dependent variable was the mean voltage (in microvolts) of the ERP in the latency window of 350–500 msec poststimulus.

neither the N400 nor another much-studied cognitive ERP component, the P300 (see Johnson, 1988, and Kutas & Van Petten, 1988, for reviews).

Figure 7 shows that the effect of word length on the broad positivity was specific to the longest words (containing eight or more letters) of Experiment 2. This visual impression was verified by the absence of a significant length effect in an ANOVA restricted to the ERPs elicited by words in the other three length categories [$F(2,60) = 1.1$].

At present, we have no firm basis for identifying the broad positivity with any previously discussed ERP component or for suggesting why the response to words of eight or more letters differed from that to the other three length categories. A number of possibilities exist. The effect may be specific to our method of word presentation, given the difficulty of reading long words on a CRT while maintaining fixation on a central point. Alternatively, the broad positivity may be related to some linguistic characteristic of long words, such as number of syllables, likelihood of being an inflected or compound word, and so forth. For present purposes, it is sufficient to know that word length had no effect on the N400. In fact, given a negative correlation between length and frequency of usage, the broad positivity elicited by long words would tend to obscure frequency effects on the N400 and lead to an underestimate of these effects when length is uncontrolled.

GENERAL DISCUSSION

The results of the two experiments can be summarized as follows: The amplitude of the N400 elicited by open-class words declined across the course of a sentence, and word frequency influenced N400 amplitude only for words occurring early in a sentence.

Sentence Position

The sequential position of a word within a sentence, in and of itself, is not likely to be a critical variable in language comprehension. However, we believe that word position can serve as a metric of the semantic and structural links that differentiate a sentence from a string of unconnected words. There are both logical and empirical reasons for supposing that the decline in N400 amplitude across the course of a sentence is an on-line reflection of these sentential constraints. In independent unrelated sentences, such as those used here, the first open-class word of each sentence must, by definition, be recognized without the benefit of prior semantic context. These words elicited the largest N400s. In contrast, our off-line cloze-probability measures indicated that the subjects were capable of "filling in" the final words of these sentences on the basis of prior context alone. These highly predictable final words elicited no N400s. Although we have no independent measure of the degree of contextual constraint imposed upon intermediate open-class words, it is reasonable to suppose that, as a group, they were more constrained than were initial content words, but less so than were terminal words. Unlike sentence-terminal words, these intermediate words elicited N400s, but they were of smaller amplitude than were those elicited by first open-class words. The observation that N400 amplitude decreased with numerical position *within* this group of intermediate words strongly suggests that each increment of sentence context influenced the analysis of subsequent words.

In Experiment 1, we evaluated one alternative explanation of the word-position effect: the possibility that it was due to the physiological refractory period of the N400. The possibility that such a refractory effect might exist, and be dissociated from the cognitive operations reflected by the N400, was suggested by analogy to another ERP component, the N100 of the auditory evoked response. In experimental paradigms that include a train of identical auditory stimuli, the amplitude of the N100 to a single stimulus is a function of the temporal interval since the most recent stimulus. N100 amplitude may be reduced by the presence of preceding stimuli within a period as long as 10 sec without any obvious changes in the perceptual quality of the eliciting stimuli (Davis, Mast, Yoshie, & Zerlin, 1966; Naatanen & Picton, 1987). Such refractory effects are likely to be restricted to early modality-specific ERP components, such as the auditory N100, rather than to later ERP components that are less closely tied to the physical parameters of the eliciting stimulus (see Woods, Hillyard, Courchesne, & Galambos, 1980a, 1980b). Our finding that N400 amplitude was unaffected by the time elapsed since the last N400-eliciting stimulus indicated that the word-position effect could not be attributed to refractory cycle. Additional support for the interpretation of our word-position effect as a consequence of surrounding context derives from the finding that words in intermediate sentence positions *do* elicit very large N400s if they are semantically anomalous (e.g., "Other turtles eat leaves, *radios*, and various plant parts"; Kutas & Hillyard, 1983).

The present results are thus consistent with other reports of word-position effects in supporting the existence of sentence-level context effects on the recognition of individual words (Marslen-Wilson & Tyler, 1980; Sanocki et al., 1985). The modulation of N400 amplitude observed here as a function of intermediate sentential position possesses no obvious morphological or latency characteristics to distinguish it from the amplitude difference observed in experiments using related versus unrelated word pairs (see Kutas, 1985; Kutas & Hillyard, 1988; Kutas & Van Petten, 1988).[5] This similarity is inconsistent with a division between low-level autonomous lexical priming effects and higher level sentence-integration effects (Fodor, 1983; Forster, 1981b). However, we will not press this argument too strongly on the basis of the present data, since we made no attempt to determine which aspect of the sentences was responsible for the word-position effect. Each sentence possessed a legal syntactic structure, some general semantic constraints based on world knowledge and, occasionally, lexically associated words, all of which might have contributed to the observed word-position effect. The issue of syntactic versus semantic sentence structure has been addressed in a subsequent experiment (Van Petten & Kutas, 1989). The question of whether sentence-position effects can be obtained in the absence of any lexically associated words is the focus of ongoing work.

The Interaction Between Word Frequency and Sentence Position

Smith and Halgren (1987) first reported that isolated low-frequency words elicit larger N400s than do high-frequency words when presented for lexical decision. They also noted an interaction between frequency and repetition (the N400 frequency effect was absent for repeated words), an interaction previously noted in reaction-time literature (Scarborough, Cortese, & Scarborough, 1977). In the present experiments, we observed an interaction between frequency and sentence context: a frequency effect was present for the first open-class words of the sentences, but not for later words. It may be tempting to speculate that the absence of a frequency effect for intermediate words does not reflect a true interaction between frequency and context; rather, the context effect is so large that N400 amplitude is "at floor," and no modulations due to frequency can be observed. The data in Figure 1 indicate that this speculation must be incorrect. While the N400 elicited by middle open-class words was significantly reduced relative to the first words, there remained a substantial negative component, unlike the ERPs to sentence-terminal words, which clearly lack such a component.

The sentence-position categories used for the analysis of frequency effects were somewhat coarse in that all of the intermediate open-class words subsequent to the first one were averaged together. The lack of a frequency effect for the entire category of "middle OC words" should not be interpreted to mean that, beginning at the second open-class word, frequency had no effect. The magnitude of the frequency effect may well have followed a gradient across the course of sentence, which was lost by collapsing across intermediate word position. Whatever the exact form of the word position × frequency function may be, the interaction between these two factors indicates that the operations involved in word recognition can be modified by context.

It is also true that the experimental scenarios used here were somewhat artificial in that the temporal delay between each word (900 msec) was longer than is normally the case in fluent reading. This factor may well have exaggerated the sentence-position effect, because the subjects had time to process and integrate each word fully before being presented with a subsequent word. In self-paced reading, there may be more overlap in the processing of contiguous words, such that recognition of one word begins before processing of the previous word is complete. The present experiments thus illustrate an extreme case in which context was allowed to play as large a role as possible. However, this extreme case has been valuable in showing that word frequency does not play a mandatory role in word recognition. It will be of some interest to determine, in future experiments, whether context can still reduce the word-frequency effect when the rate of stimulus presentation more closely approximates that of fluent reading.

The elimination of frequency effects by prior context argues against the logogen model's description of frequency as a stable additive factor in the recognition of each word (Morton, 1969). This interaction between frequency and context also argues against the concept of discrete serial stages in word recognition in which a preliminary stage is influenced by frequency and later stages utilize contextual information (as in Bradley & Forster, 1987).

The lack of a frequency effect for intermediate and sentence-terminal words is consistent with Becker's (1976, 1979, 1980) serial search model, which postulates that contextual information is used first and a frequency-ordered search is initiated only when the candidate word is not included in the semantic set. Despite the congruity between the present results and some of the predictions of Becker's serial search model, we had some conceptual difficulty in trying to apply the model to our experimental scenarios in a straightforward way. As formulated, the model does not allow the reader or listener the option of choosing between the use of a semantic set or a sensory set; the semantic set always receives priority. The model specifies a relatively inflexible search path, because there is no provision for a top-down control process that determines whether or not there is any relevant context on which to base a semantic set. Thus, it was difficult to imagine what the contents of the semantic set would be in the instance of the first open-class word of an independent isolated sentence.

Given that subjects in experimental scenarios such as ours are aware that each sentence begins a new topic, it is reasonable to believe that they delay using a semantic set until they have some useful information from which to derive it. We can imagine the addition of a time-decay factor that would preserve the principle of an inflexible ordered search but discard the use of a semantic set if more than a few seconds had elapsed since the most recent content word. But even such an addition would not adequately capture many real-life situations in which people can distinguish and resolve unconnected streams of discourse even though they occur in close temporal proximity. Consider sitting at a party and intermittently switching your attention between two ongoing conversations, changing between two radio stations, or being interrupted by a talkative visitor during a telephone conversation. In each case, people seem to be able to utilize pragmatic cues indicating a break in one linguistic context and the beginning of another one.

Although the present data fit Becker's (1979) prediction that semantic cues to a word's identity can reduce the importance of frequency, we believe that a purely serial application of strategies (semantic relationships first, word frequency second) yields an implausible view of how people read a series of independent sentences. An alternative view is that people are able to recognize single words quickly and efficiently by using the best information available at any point in time. Overall frequency of usage is a poor clue to the identity of any given word, but, in the absence of context, adopting a bias toward more common words is an efficient processing strategy. The structure of a sentence provides more specific cues to the identity of subsequent words; the lack of a frequency effect for intermediate words suggests that these more powerful constraints have their effects in real time.

Recent models of speech recognition have described the language-processing system as being complex enough to perform a number of computations in parallel and to use each bit of information as it becomes available (Marslen-Wilson, 1987; McClelland & Elman, 1986). These models have had some success in dealing with issues peculiar to speech, such as cue tradeoffs and lexical effects in phoneme identification, by continuously combining information from different dimensions of the input. The present results suggest that similar trading relations and interactions between different information sources can occur in reading and that semantic relations may prove to be one of the most powerful sources of information in the processing of individual words.

REFERENCES

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical and/or production frequency? *Journal of Memory & Language*, **24**, 89-106.

Becker, C. A. (1976). Allocation of attention during visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 556-566.

Becker, C. A. (1979). Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 252-259.

Becker, C. A. (1980). Semantic context effects in visual word recognition: an analysis of semantic strategies. *Memory & Cognition*, **8**, 493-512.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials associated with semantic priming. *Electroencephalography & Clinical Neurophysiology*, **60**, 343-355.

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, **24**, 14-25.

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, **8**, 631-642.

Bradley, D. C., & Forster, K. I. (1987). A reader's view of listening. *Cognition*, **25**, 72-103.

Cairns, H. S., & Foss, D. J. (1971). Falsification of the hypothesis that word frequency is a unified variable in sentence processing. *Journal of Verbal Learning & Verbal Behavior*, **10**, 41-43.

Davis, H., Mast, T., Yoshie, N., & Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: Recovery process. *Electroencephalography & Clinical Neurophysiology*, **21**, 105-113.

Feldstein, P., Smith, M. E., & Halgren, E. (1987, June). *Cross-modal repetition effects on the N4.* Paper presented at the Fourth International Conference on Cognitive Neuroscience, Dourdan, France.

Fischler, I., & Bloom, P. A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning & Verbal Behavior*, **18**, 1-20.

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

Forster, K. I. (1981a). Frequency blocking and lexical access: One mental lexicon or two? *Journal of Verbal Learning & Verbal Behavior*, **20**, 190-203.

Forster, K. I. (1981b). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology*, **33A**, 465-495.

Foss, D. J. (1969). Decision processes during sentence comprehension: Effects of lexical time difficulty and position upon decision times. *Journal of Verbal Learning & Verbal Behavior*, **8**, 457-462.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113**, 256-281.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. *Journal of Memory & Language*, **24**, 631-645.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, **28**, 267-283.

Grosjean, F., & Itzler, J. (1984). Can semantic constraint reduce the role of word frequency during spoken-word recognition? *Bulletin of the Psychonomic Society*, **22**, 180-182.

Harbin, T. J., Marsh, G. R., & Harvey, M. T. (1984). Differences in the late components of the event-related potential due to age and to semantic and non-semantic tasks. *Electroencephalography & Clinical Neurophysiology*, **59**, 489-496.

Holcomb, P. J. (1985). Unimodal and multimodal models of lexical memory: An ERP analysis. *Psychophysiology*, **22**, 576. (Abstract)

Holcomb, P. J. (1988). Automatic and attentional processing: An event-related brain potential analysis of semantic processing. *Brain & Language*, **35**, 66-85.

Hudson, P. T. W., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory & Language*, **24**, 46-58.

Inhoff, A. W. (1983). Attentional strategies during the reading of short stories. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 181-192). New York: Academic Press.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography & Clinical Neurophysiology*, **10**, 371-375.

JOHNSON, R., JR. (1988). The amplitude of the P300 component of the event-related potential: Review and synthesis. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 3, pp. 69-138). Greenwich, CT: JAI Press.

JUST, M. A., & CARPENTER, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, **87**, 329-354.

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

KRAMER, A. F., & DONCHIN, E. (1987). Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 76-86.

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

KUTAS, M. (1985). ERP comparisons of the effects of single word and sentence contexts on word processing. *Psychophysiology*, **22**, 575-576. (Abstract)

KUTAS, M. (1986). Event-related potentials (ERPs) elicited during rapid serial visual presentation of congruous and incongruous sentences. *Electroencephalography & Clinical Neurophysiology*, Suppl. 40, 406-411.

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 & Language*, **11**, 354-373.

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

KUTAS, M., & HILLYARD, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, **11**, 539-550.

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

KUTAS, M., & HILLYARD, S. A. (1988). An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience*, **1**, 38-49.

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., NEVILLE, H. J., & HOLCOMB, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening, and signing. *Electroencephalography & Clinical Neurophysiology*, Suppl. 39, 325-330.

KUTAS, M., & VAN PETTEN, C. (1988). Event-related brain potential studies of language. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychophysiology* (Vol. 3, pp. 139-187). Greenwich, CT: JAI Press.

KUTAS, M., VAN PETTEN, C., & BESSON, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography & Clinical Neurophysiology*, **69**, 218-233.

MANGUN, G. R., & HILLYARD, S. A. (1987). The spatial allocation of visual attention as addressed by event-related brain potentials. *Human Factors*, **29**, 195-212.

MARSLEN-WILSON, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, **25**.

MARSLEN-WILSON, W., & TYLER, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, **8**, 1-71.

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**, 477-488.

MCCLELLAND, J. L., & ELMAN, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**, 1-86.

MORTON, J. (1964). The effects of context on the visual duration threshold for words. *British Journal of Psychology*, **55**, 165-180.

MORTON, J. (1969). The interaction of information in word recognition. *Psychological Review*, **76**, 165-178.

NAATANEN, R., & PICTON, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, **24**, 375-425.

NEVILLE, H. J., KUTAS, M., CHESNEY, G., & SCHMIDT, A. (1986). Event-related brain potentials during the initial encoding and subsequent recognition memory of congruous and incongruous words. *Journal of Memory & Language*, **25**, 75-92.

NEVILLE, H. J., & LAWSON, D. (1987). Attention to central and peripheral visual space in a movement detection task: An event-related potential and behavioral study: I. Normal hearing adults. *Brain Research*, **405**, 253-267.

NORRIS, D. (1986). Word recognition: Context effects without priming. *Cognition*, **22**, 93-136.

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

POLICH, J., & LAWSON, D. (1985). Event-related potential paradigms using tin electrodes. *American Journal of EEG Technology*, **25**, 187-192.

RUBINSTEIN, H., GARFIELD, L., & MILLIKAN, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning & Verbal Behavior*, **9**, 487-494.

RUGG, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, **22**, 642-647.

RUGG, M. D., MILNER, A. D., & LINES, C. R. (1985). Visual evoked potentials to lateralised stimuli in two cases of callosal agenesis. *Journal of Neurology, Neurosurgery & Psychiatry*, **48**, 367-373.

SANOCKI, T., GOLDMAN, K., WALTZ, J., COOK, C., EPSTEIN, W., & ODEN, G. (1985). Interaction of stimulus and contextual information during reading: Identifying words within sentences. *Memory & Cognition*, **13**, 145-157.

SCARBOROUGH, D. L., CORTESE, C., & SCARBOROUGH, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 1-17.

SCHMIDT, A. L., ARTHUR, D. L., KUTAS, M., GEORGE, J., & FLYNN, E. (1989). Neuromagnetic responses evoked during reading meaningful and meaningless sentences. *Psychophysiology*, **26**, S6.

SCHUBERTH, R. E., & EIMAS, P. D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 27-36.

SCHUBERTH, R. E., SPOEHR, K. T., & LANE, D. M. (1981). Effects of stimulus and contextual information on the lexical decision process. *Memory & Cognition*, **9**, 68-77.

SMITH, M. E., & HALGREN, E. (1987). Event-related potentials during lexical decision: Effects of repetition, word frequency, pronounceability, and concreteness. *Electroencephalography & Clinical Neurophysiology*, Suppl. 40, 417-421.

SMITH, M. E., STAPLETON, J. M., & HALGREN, E. (1986). Human medial temporal lobe potentials evoked in memory and language tasks. *Electroencephalography & Clinical Neurophysiology*, **63**, 145-149.

SOLOMON, R. L., & HOWES, D. H. (1951). Word frequency, personal values and visual duration thresholds. *Psychological Review*, **58**, 256-270.

STANOVICH, K. E., & WEST, R. F. (1981). The effect of sentence processing on ongoing word recognition: Tests of a two process theory. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 658-672.

STANOVICH, K. E., & WEST, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, **112**, 1-36.

TAYLOR, W. L. (1953). "Cloze" procedure: A new tool for measuring readability. *Journalism Quarterly*, **30**, 415-417.

TULVING, E., & GOLD, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*, **66**, 319-327.

TYLER, L. K. (1985). The structure of the initial cohort: Evidence from gating. *Perception & Psychophysics*, **36**, 417-427.

UNDERWOOD, G., & BARGH, K. (1982). Word shape, orthographic regularity, and contextual interactions in a reading task. *Cognition*, **12**, 197-209.

VAN PETTEN, C., & KUTAS, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning course of meaning activation. *Journal of Memory & Language*, **26**, 188-208.

VAN PETTEN, C., & KUTAS, M. (1989). *Influences of semantic and syntactic context on open and closed class words*. Manuscript submitted for publication.

WHALEY, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning & Verbal Behavior*, **17**, 143-154.

WOODS, D. L., HILLYARD, S. A., COURCHESNE, E., & GALAMBOS, R. (1980a). Electrophysiological signs of split-second decision making. *Science*, **207**, 655-657.

WOODS, D. L., HILLYARD, S. A., COURCHESNE, E., & GALAMBOS, R. (1980b). Recovery cycles of event-related potentials in multiple detection tasks. *Electroencephalography & Clinical Neurophysiology*, **50**, 335-347.

NOTES

1. Note that apportioning stimuli into experimentally defined categories is an essential precondition to forming an ERP. The brain activity elicited by a single stimulus is accompanied by spontaneous ongoing EEG that is larger in amplitude (see Nunez, 1981). Stimulus-elicted activity is extracted from the EEG by the simple signal-processing technique of averaging single trials, allowing activity that is time-locked to stimulus presentation to summate and allowing activity that is random with regard to the stimuli to cancel. In the present study, we sorted stimuli so as to yield a minimum of 24 trials per category. On the basis of previous experience, we have found this number to yield an adequate signal-to-noise ratio for the observation of variations in N400 amplitude.

2. The brain system responsible for the scalp-recorded N400 is largely unknown, although a similar potential has been recorded within the medial temporal lobe. The additional finding that patients with left temporal lobectomies lack repetition priming effects on the N400 suggests at least a modulatory role for the temporal lobe on the processes underlying the elicitation of an N400 (Feldstein, Smith, & Halgren, 1987; Smith, Stapleton, & Halgren, 1986). Recent neuromagnetic data from a semantic priming paradigm are also consistent with a source in the left temporal lobe (Schmidt, Arthur, Kutas, George, & Flynn, 1989).

3. Kutas (1986) has similarly reported that the amplitude of the N400 to semantically incongruous sentence completions is similar whether the sentences are presented at a rate of one word every 100 msec or one every 700 msec.

4. A "reliable" average is one in which the "signal," or stimulus-elicited activity, is sufficiently greater than the "noise," or spontaneous ongoing EEG activity, to allow its visualization and accurate measurement. A simple means of gauging whether or not the signal-to-noise ratio is adequate for the purposes of a specific comparison is to compare the prestimulus baselines across the conditions to be compared. Given adequate experimental control over the events preceding the stimulus conditions one wishes to compare, there should be little difference in these prestimulus baselines. If, in fact, these baseline differences are as large as the stimulus-elicited effects one might reasonably expect to obtain, then there are not enough trials in the averages to produce a reliable signal.

5. It should be noted that, unlike the present study, most published studies of ERPs elicited by related and unrelated word pairs have required an immediate behavioral response from the subject (Bentin, McCarthy, & Wood, 1985; Harbin, Marsh, & Harvey, 1984; Holcomb, 1988). In these cases, the N400 overlaps with a decision-related positivity (the P300), which makes the related/unrelated difference appear somewhat different from the word-position effect reported here. The papers cited in the text report an experiment in which no immediate task-related decision was possible; the related/unrelated difference in these waveforms closely resembles the sentential word-position effect.

APPENDIX 1
Examples of Sentences Used in Experiment 1

The movie was so jammed they couldn't find a single seat.
When the two met one of them held out his hand.
Everything she owned was in a brown paper bag.
The bill was due at the end of the month.
The fire was burning low so he went out to chop more wood.
He had the flu and needed to stay home and rest.
Most shark attacks occur very close to shore.
The uranium miner had been exposed to dangerous levels of radiation.
She finally got up her nerve to ask her boss for a raise.
They decided to have a picnic at the county park.
The rock climber lost her grip and started to fall.
Their car was searched when they tried to cross the border.
The sweater was knitted from blue and grey wool.
The magician took out his hat and made a rabbit appear.
It is supposed to bring seven years bad luck to break a mirror.
In the Old West stagecoaches were pulled by horses.
He refused to register for the draft.
He lost all his chips in the last hand of poker.
Once he makes a decision he never changes his mind.

APPENDIX 2
Words of Frequency Less than 30 per Million Used in Experiment 1

First Open-Class Words

amateur	dough	magician	shark
ambitious	dyed	mole	shaved
awoke	eskimo	mosquito	skiing
babysitter	fasting	movers	sparrow
basketball	firewood	mustard	squirrel
bathroom	florist	necklace	stamp
bet	flu	pacemaker	stung
bitten	freshman	pint	swam
boxer	gambler	pizza	sweater
caddy	gardener	plumber	tenants
chargers	gotten	rainstorm	toothache
chocolate	graders	robbed	trainer
cub	hungry	sailor	uranium
dentist	invested	scraped	wealthy
dictator	jeweler	seasoned	weatherman
docked	loosened	sewed	wrapped

APPENDIX 2 (continued)

Middle Open-Class Words

airport	den	landlord	shingle
almonds	diamond	lap	shirt
anchor	diet	lawn	shoved
anthem	digital	leaky	sidewalk
apple	dipped	lifter	slice
bake	downhill	lion	sofa
barbed	downstream	loaves	sore
bathroom	dripping	marbles	spilled
bedtime	easter	miner	sprigs
bike	evicted	mows	sprung
bite	exit	ms	stagecoaches
blunt	faucet	neighbor	steak
boss	fielder	nervous	steeple
bottles	flip	oak	stolen
bowls	fountain	ordained	stoplight
brake	frosted	outfit	storm
brand	fry	pancakes	stove
burnt	fudge	patch	stray
button	gardener	paw	streak
cake	golfer	pentagon	stuck
campers	gotten	physic	subtract
canoed	grammatical	picnic	sunny
cans	grocery	pie	sunshine
cardboard	gypsy	pleaded	surfboard
carved	halftime	poker	swab
cereal	ham	polar	sweater
cheese	hamburger	purse	sweaty
cherry	humid	quart	telescope
chips	hurried	quit	teller
chocolate	icy	rabbit	thermometer
chop	illegal	rake	tile
chorus	jail	raspberry	toast
climber	jammed	razor	triangle
clock	juice	recover	turkey
coin	juleps	relish	vodka
collapsed	ketchup	repay	wax
conditioner	knitted	roommate	zip
couch	knockout	royalty	
curled	lab	rubies	
curtsy	lame	scrambled	

Final Words

bail	exam	organ	rye
bark	fog	oven	sandwich
bee	fur	pen	shave
bluff	ghost	penny	shed
bow	glasses	pepper	sole
bowl	guilty	photo	sour
broom	hamburger	pit	spade
bug	hen	pitcher	spider
cabinet	honey	poker	squash
calf	hungry	port	stall
carpet	jam	punch	stamp
cavity	lap	quart	straw
chase	lawn	rainbow	swallow
cheese	leak	ramp	syrup
chemistry	lion	rat	thread
closet	litter	refrain	tick
cream	loaf	ribbon	toast
deed	mint	robber	toll
dentist	mustache	rope	towel
dessert	neat	ruler	wool
diamond	orange	rung	

(Manuscript received December 1987;
revision accepted for publication October 16, 1989.)