Understanding Word and Sentence
G.B. Simpson (Editor)

Chapter 6

Electrophysiological Evidence for the Flexibility of Lexical Processing

Cyma Van Petten and Marta Kutas

University of California, San Diego

La Jolla, California U.S.A.

The frequency of a word's occurrence in common usage and the relationship of a word to a prior word have proven to be two of the most powerful determinants of performance in experimental studies of word recognition. In reaction time tasks such as pronunciation or lexical decision, subjects typically respond more slowly to rare than to common words; with tachistoscopic presentation, subjects require longer exposure durations to report rare words accurately (e.g., Rubenstein, Garfield, & Millikan, 1970; Solomon & Howes, 1951). Similarly, the results of a number of tasks, including pronunciation, lexical decision, verbal report of briefly presented or visually degraded words, naming the color of ink in which a word is printed, and monitoring for a target word, have led to the conclusion that processing of a single word is facilitated by the prior occurrence of a related word (Becker & Killion, 1977; Massaro, Jones, Lipscomb, & Scholz, 1979; Meyer & Schvaneveldt, 1971; Rouse & Verinis, 1962; Warren, 1974, 1977). Both of these findings can be explained by postulating a fairly simple self-contained lexicon wherein 1) commonly accessed entries are either closer to some threshold activation level (Morton, 1969) or have priority in a search procedure (Bradley & Forster, 1987) and 2) some entries are strongly linked such that they can influence each other via some type of "spreading activation."

The stimulus pairs used in these receptive language task have typically been judged as related via production norms (i.e., people tend to produce one when presented with the other). The pairs are often members of the same category of objects, or antonyms. It is unlikely that such pairs will often co-occur in the same sentence. When they do, they will comprise only a small subset of the words in the sentence. A priming effect between words that are semantically similar but not associatively related (e.g., "prince - boy") appears to be less robust; it has been observed to be either very small or dependent on the use of a lexical decision task (Fischler, 1977a; Huttenlocher & Kubicek, 1983; Lupker, 1984; Seidenberg, Waters, Langer, & Sanders, 1984; Warren, 1977). The extension from single word to sentence contexts is thus not obvious. Moreover, it can be argued that single word and sentence contexts must affect subsequent words in a fundamentally different way because lexical context effects can take place within a lexicon, whereas sentence-level processes, by definition, involve novel combinations of words that cannot be pre-stored in a lexicon.

There has been some doubt as to whether sentence-level semantic context effects occur in situations which approximate fluent reading, namely those wherein skilled readers quickly process words which are not highly predictable, not strongly associated to previous words, and not visually degraded (see Fischler & Bloom, 1979; Forster, 1981; Henderson, 1982, pp. 351-353; Mitchell & Green, 1978; Perfetti, Goldman, & Hogaboam, 1979). Of course, this position is not universally accepted and a number of recent studies clearly demonstrate the influence of sentence or sentence-like context (O'Seaghdha, 1989; Sanocki, Goldman, Waltz, Cook, Epstein, & Oden, 1985; Ratcliff, 1987; Simpson, Peterson, Casteel, & Burgess, 1989). Results from our lab have consistently shown a pervasive influence of sentence-level context as well; these will be presented in a later section of the present chapter.

Because sentence-level context effects are seen in the same experimental tasks that yield associative context and frequency effects, the question arises as to which, if any, of these effects originates in the lexicon. This leads in turn to the more general question concerning the constituents of a "lexical entry": abstract orthographic and/or phonemic information only (i.e. only the physical form of a word, perhaps with a function that allows some mapping between modalities), the syntactic category of the word, some basic semantic information in addition, or all of the above plus a complete encyclopedic listing? If we were to grant that the lexicon contained everything a person knows about words and their referents, the original conception of the lexicon as only one among several components in the language processing system would cease to be theoretically functional. Rather than postulate some division between operations "in" the lexicon and "post-lexical" processes, perhaps we should retreat to the most general issue of all, namely the sequence and organization of mental operations that yield the intended meaning of a sentence.

In the research to be described, we have focused primarily on only one aspect of this issue – how and when the specific connotation of a word in a sentence is singled out or constructed by the reader. Even a brief intuitive analysis will suggest that the potential meaning of a single word is broader and more diverse than that used in any one sentence, so that perceiving the author's intended meaning indeed poses a major task for the reader.

The semantic attributes of a word may be placed on a continuum ranging from the most basic, those which are frequently used and held in common by most users of the language, to those which are more encylopedic in that they are less often pertinent and perhaps belong only to the vocabulary of the specialist. For the word "bird," for example, "typically feathered" and "typically flies" are attributes which are often relevant to the interpretation of a sentence containing the word, and most English speakers will have little trouble accessing these attributes as needed. The facts that birds typically have hollow bones and high metabolic rates will less often be required to understand a sentence using the word "bird" and these attributes may be known by only a subset of English speakers. The basic and encyclopedic features of "bird" are not in conflict with one another; indeed there are logical entailments between them because having a high metabolic rate and hollow bones contribute to the ability to fly. The encylopedic features of the specialist's vocabulary simply reflect more extensive knowledge of the referent (see Langacker, 1987).

In contrast, other words may possess semantic attributes which have no necessary relationship, despite the fact that they may often be conflated in usage. To borrow Lakoff's (1987) example, "mother" may often imply a genetic, a birth, and a nurturance relationship all at once, but a single relationship can be picked out by the appropriate context as in "surrogate mother" (implying only the birth relationship) or "adoptive mother" (implying only the nurturance relationship). Unlike the basic-encyclopedic continuum, these different attributes of "mother" exist at the same level of detail, but may be more or less compatible with the intended meaning of a particular sentence.

A third manner in which the potential senses of a word may be related seems to be distinct from the previous two in that the different senses are always incompatible with one another. One may use either the spatial or temporal sense of "over," but not at the same time: a sentence context clearly selects one or the other (see Brugman & Lakoff, 1988 for an analysis of how even the spatial sense of "over" contains several distinct meanings or Lindner, 1981 for an analysis of the polysemy of "up" and "out"). These two senses of "over," however, do seem to be closely related to one another. In the most extreme case of polysemy even this family resemblance is lacking. For unsystematic homographs such as "spoke" (past tense of "speak") and "spoke" ("part of a wheel") it is merely an accident of linguistic change that the same physical form has come to be associated with completely different meanings. Like the lexical entry for "bird," both meanings of "spoke" may contain core and encyclopedic attributes, but there are two non-overlapping sets of core attributes, only one of which will be appropriate in any given sentence.

In the last ten years or so, one of the most influential views of language comprehension has been that the initial recognition of a word's meaning is context-invariant: while isolated word meanings may be inherently ambiguous, polysemous,or vague, the responsibility of selecting among the multiplicity of retrieved meanings for those that contribute to the ongoing discourse model is left to subsequent processes. A hierarchical processing sequence has prevailed as a means of explaining how the reader arrives at the final interpretation of a word. In broad outline, the work of a number of authors has suggested something like the following sequence of processes:

1) an orthographic (or perhaps phonemic) access code is derived from the visual input;

2) the code is used to tap into the mental lexicon and find a matching word;

3) the lexical entry for the word is accessed such that all of its basic core meanings are activated;

4) the appropriate core meaning is selected based on prior context;

5) inferential processes combined with knowledge of the world and the speaker's intentions are used to elaborate the relevant encyclopedic attributes of the word's meaning and integrate these with previous words to establish the meaning of the sentence.

In this scheme, the first three stages take place within the lexicon, while subsequent steps are carried out by distinct higher-level integrative mechanisms. Evidence for a sequential processing arrangement here, as in other domains of cognitive psychology, arises from two major sources. First, the presence or absence of interactions among different experimental effects: An effect which is constant despite the addition of other effects must arise at a stage of processing prior to that which yields the other effects. It has been argued that the influence of word frequency must arise prior to the completion of Stage 3 because it is impervious to lexical or sentential context manipulations (Bradley & Forster, 1987; Forster, 1981a, 1981b; Schuberth & Eimas, 1977; Schuberth, Spoehr, & Lane, 1981). However, this result has been inconsistent in that an attenuation of frequency effects by context effects has been observed in some experiments (Becker 1979; Grosjean & Itzler, 1984; Stanovich & West, 1983).

A second source of experimental support for placing contextual factors late in the processing sequence could come from measurements of the temporal onset and duration of different experimental effects. But because most dependent measures are temporally punctate, consisting of a single discrete motor response such as a button press, it has been possible to estimate these temporal factors only indirectly. It has been argued that manipulations of the stimulus-onset-asynchrony (SOA) between a stimulus word and a target can be used to limit the amount of time devoted to the stimulus word. Thus with a short SOA, responses to the target should be influenced only by the products of the initial, as-yet-incomplete analysis of the stimulus word. This type of experimental paradigm has been used primarily to contrast different varieties of semantic analyses by varying the nature of the semantic relation between the target word, the immediately preceding stimulus word, and the rest of the prior text. Results have suggested Stages 1 through 3 above can be separated from the "higher level" processes of Stages 4 and 5. So for instance, most of the plausible inferences supported by a text seem to be drawn relatively late. This follows from the finding that inference words do not show priming with a short SOA which does yield priming for semantic associates of the stimulus word (McKoon & Ratcliff, 1989; Swinney & Osterhout, 1990; Till, Mross, & Kintsch, 1988). To illustrate, following the sentence "Thinking of the amount of garlic in his dinner, the guest asked for a mint.", a word such as "candy" which is semantically associated to "mint" in isolation would show priming with a short SOA, while the inference probe "breath" would show priming with a longer SOA. This dissociation thus indicates that access to the core meaning of the stimulus word (and consequently priming of its associates) must precede inferential processes.

Manipulations of SOA have also led to the more remarkable conclusion that even the initial activation of the stimulus word's core semantic attributes is indiscriminate and context-invariant. When the prime word possesses mutually incompatible core meanings, as in the case of a homograph, semantic associates of both meanings show priming effects when the interval between the stimulus word and its associates is brief. In the example above, "money" would also show a priming effect if presented a short time after "mint." With a longer interval, only the associate of the sententially relevant meaning of the homographic word would continue to show a priming effect. This result has been taken as the final bit of evidence for the five-stage sequence outlined above in that it subdivides even core meaning processes into two discrete steps and appears to place the first of these beyond the reach of sentence context (Kintsch & Mross, 1985; Onifer & Swinney, 1985; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979; Till et al., 1988; see Tabossi, 1988a, 1988b for different results and conclusions). However, this conclusion is debatable; we will critique the experimental logic underlying it and present some contradictory data shortly.

Several experiments conducted in our laboratory have led us to question the hierarchical nature of the processing sequence above, particularly the postulate of a self-contained lexicon which is insensitive to sentence meaning but produces both word frequency and associative priming effects. In this chapter, we review several data sets which suggest that:

1) the process which yields frequency effects for words presented in isolation is neither mandatory nor immune to sentence-level context;

2) the influence of sentence-level context can be as powerful and act as early as that of a single lexically associated word; and

3) sentence context can be used to pick out the appropriate core meaning of an ambiguous word without first passing through an early stage of

indiscriminate semantic activation.

As all of the empirical evidence to follow consists of event-related brain potential data, some description of the technique and a review of its applicability to psycholinguistic research is in order.

Event-related Potentials

Electrodes placed on the scalp can be used to record voltage fluctuations known as the electroencephalogram (EEG). It is generally believed that the electrical activity recorded at the scalp is a summation of graded post-synaptic potentials (PSPs) generated by the depolarization and hyperpolarization of brain cells (see Wood & Allison, 1981 for a review of the neurophysiological basis of the EEG or Nunez, 1981 for a treatise on the physics of EEG). At any given moment the observed EEG is likely to reflect the activity of a number of functionally distinct neuronal populations. With the advent of computer averaging some two decades ago, it became possible to obtain an estimate of activity which is time-locked to an arbitrary point, such as the onset of a stimulus. Averaging many epochs of EEG following each of a set of similar stimuli tends to cancel the random background EEG, leaving a record of the evoked or event-related potentials (EPs or ERPs) which were synchronized to the time of stimulus presentation. Which stimuli are defined as "similar" depends on the goals of the experiment and is established *a priori* by the experimenter. The resulting waveform of voltage plotted against post-stimulus time typically includes a series of positive and negative peaks. Much ERP research has focused on the decomposition of these voltage fluctuations into experimentally dissociable "components" which can be linked to a specific physiological and/or cognitive process. Attempts to identify a functionally distinct component may include manipulations of the physical (e.g., size, luminance, pitch, etc.) or psychological (e.g., task-relevance, meaningfulness, predictability, etc.) attributes of the stimuli, or the physiological state of the subject (e.g., drug administration, selecting a population with a particular type of brain damage, etc.). Other factors called upon for component identification are voltage polarity, peak or onset latency and duration, distribution across the scalp, and general waveshape.

Knowledge of the neural generator can offer yet another criterion for the identity of a scalp-recorded component, one which is of interest in and of itself. A component which has been closely linked to language processing is the N400 (so called because it is negative-going in polarity and has a typical peak latency of 400 msec post-stimulus onset). To date, evidence from work with commissurotimized individuals, depth recordings from epileptic patients with temporarily-implanted electrodes, correlations between scalp recordings and regions of high glucose metabolism, and scalp recordings of evoked magnetic fields have all converged to suggest a left temporal lobe substrate for the N400

(Halgren, in press; Heit, Smith, & Halgren, 1988; Kutas, Hillyard, & Gazzaniga, 1988; McCarthy & Wood, 1984; Schmidt, Arthur, Kutas, George, & Flynn, 1989; Smith, Stapleton, & Halgren, 1986). However, it would be premature to conclude at this point that all N400 effects are produced by the same population of cells.

Sensitivity of the N400 to Psycholinguistic Manipulations

The N400 was first described in experiments which compared semantically predictable to semantically incongruent sentence completions. Subjects in this experiment, as in most of those here, read (silently) sentences as they were presented one word at a time on a CRT. Incongruous final words elicited a negative wave which was largest over posterior scalp locations and somewhat larger over the right than left hemisphere, whereas congruous words elicited a positive-going wave instead. The first separation between the congruous and incongruous waveforms occurred at about 200 msec after the onset of the visual word; the difference peaked at about 400 msec poststimulus (Kutas & Hillyard, 1980a, 1980b, 1980c, 1982; Kutas, Van Petten, & Besson, 1988). These first incongruity experiments were conducted in 1978. Since then it has become clear that the positivity for a wholly predictable word is the exceptional case: most words elicit an N400. Its amplitude and latency vary with experimental manipulation.

When lists of letter strings have been presented to subjects the following pattern of results has emerged: Unrepeated words which are semantically unrelated to previous words elicit the largest N400; orthographically legal, pronounceable nonwords (pseudowords) also elicit large N400s; and unpronounceable nonwords elicit little or no N400 activity (Bentin, 1987; Bentin, McCarthy & Wood, 1985; Holcomb, 1988; Rugg & Nagy, 1987; Smith & Halgren, 1987).[1] If the component were produced only after the meaning of a word had been accessed, there should be no N400 for pseudowords. On the other hand, if the N400 reflected simply the "wrongness" of a letter string, there should be a sizeable N400 for illegal nonwords. The results from the two classes of nonwords thus suggest that the N400 reflects some of the earlier processes in visual word recognition, wherein illegal nonwords can be quickly rejected but pseudowords require some additional processing to determine that they are not, in fact, words.

The frequency of a word's usage also modulates the amplitude of the N400 elicited by words presented in lists. Whether the task is lexical decision or detection of repeated words, rare words elicit larger N400s than do common words (Rugg, in press; Smith & Halgren, 1987). Finally, the amplitude of the N400 elicited by words in lists is reduced if the same word, or a semantically associated word occurred earlier in the list (Bentin et al., 1985; Holcomb, 1988; Kutas, 1985; Rugg, Furda, & Lorist, 1988). A similar reduction of N400 amplitude is apparent for words occurring in repeated sets of sentences, (Besson, Kutas, & Van Petten, in press) and for words which are reused in a text although they occur in different sentences (Van Petten, Kutas, Kluender, & McIsaac, unpublished observations).

Negative results are also important in establishing that the N400 component is specific enough to be useful as a measure of language processing. There has been some controversy among ERP researchers as to the relationship between the N400 and another ERP component, the N2, and whether or not the N400 should be regarded as "language-specific" (see Kutas & Van Petten, 1988; Ritter, Ford, Gaillard, Harter, Kutas, Naatanen, Polich, Renault, & Rohrbaugh, 1984; Stuss, Picton, & Cerri, 1986; Stuss, Sarazin, Leech, & Picton, 1983 for some discussion of this issue). However, it is clear that N400s are not universally elicited by any stimulus which fails to fit into a previously established context. So, for instance, ending a sentence with a word which is semantically congruous but of a larger typeface than the preceding sentence fragment, or presenting a slide of a complex abstract drawing rather than a word, does not yield an N400 component but rather a late positive component of the P300 family (Kutas & Hillyard, 1980a, 1984b). Similarly, altering a note in a well-known melody, or interrupting a progression of geometric figures which have been ordered by increasing size, do not produce N400s (Besson & Macar, 1987).

In experiments using sentences, the amplitude of the N400 elicited by the final word is sensitive not only to whether the word is, roughly, congruous or incongruous with the preceding fragment, but *how* congruous the terminal word is. An *a priori* metric of the amount of semantic constraint imposed on a terminal word by the preceding fragment can be obtained via the off-line technique of cloze probability, e.g., what proportion of subjects will fill in a particular word as being the most likely completion of a sentence fragment (Taylor, 1953). Cloze probability proportions and N400 amplitude have been shown to be inversely correlated at a level above 90%. It is important to note, however, the subtle distinction between the cloze probability of a terminal word and the contextual constraint of the sentence fragment *per se*. For example, the sentence fragment "The bill was due at the end of the ..." is of high contextual constraint in that most people will fill in "month" while "He was soothed by the gentle ..." is of low contextual constraint because there are a number of acceptable endings, no one of which is clearly preferred over the others (Bloom & Fischler, 1980). But both fragments can be completed by words of equal (low) cloze probability as in "The bill was due at the end of the *hour*." and "He was, soothed by the gentle *wind*." The results of experiments which crossed several levels of contextual constraint with several levels of cloze probability showed that the N400 was correlated with the cloze probability of the final word but generally independent of the contextual constraint of the preceding sentence fragment (Kutas & Hillyard, 1984a; Kutas, Lindamood, & Hillyard, 1984). This result was critical in establishing that N400 amplitude does not index the violation of previously established expectancies for a particular word which was not presented, but rather is sensitive to the degree to which the sentence fragment prepared the way for the word which actually followed. Note however that in the absence of an explicit attempt to dissociate cloze probability and contextual constraint, the two factors are generally correlated. In what follows we will use the term "contextual constraint" in reference to this more typical situation.[2]

Sentence Context and Word Frequency

We have conducted a number of experiments wherein subjects' primary task was to read a number of unrelated sentences. Across experiments, we have assigned a variety of secondary tasks in order to ensure subjects' continued alertness. These have included detecting an occasional repeated sentence, indicating whether or not a subsequent probe word had occurred in the preceding sentence, and indicating whether or not a particular letter of the alphabet occurred in a single word presented subsequent to each sentence. These tasks all were constructed in such a way that subjects were neither required to make overt responses, while reading nor, with the exception of the repeated sentence task (wherein the repeated sentences were excluded from analysis), to recognize task-related stimuli while reading (in the recognition probe and letter detection tasks, subjects did not know which word or which letter would be their target until some time after the completion of the sentence). We have seen little variability in the pattern of ERP results dependent on which secondary task was used, and thus believe that the data reflect general mechanisms of word recognition and sentence comprehension rather than task-specific factors.[3]

The cloze probability results reviewed above suggest that the amplitude of the N400 is a sensitive index of the semantic constraints imposed by a sentence fragment on the processing of its final word. More recently, we have evaluated whether or not this holds for the intermediate words of sentences as well. With a set of unrelated stimulus sentences, the subject/reader must begin each sentence with no information concerning the topic, but as it progresses he or she should begin building a mental model of the concept expressed by the sentence and have more information available concerning what sorts of words may occur next. Accordingly, we sorted intermediate open class words according to sentence position and observed a decrement in N400 amplitude with increasing position (Kutas et al., 1988). The linear decrease in N400 amplitude has been observed in an experiment in which all of the stimulus materials were normal congruent sentences, as shown in Figure 1. Moreover, no behavioral task which could have fostered differential levels of preparedness was assigned to the subjects (Van Petten & Kutas, in press).

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, word position is a simple index of the semantic and structural links which differentiate a sentence from a string of unconnected words. Thus we have interpreted the decrement in N400 amplitude across the course of a sentence as a consequence of the build-up of contextual constraints. A similar word position effect has been reported for word-monitoring times with speech materials (Marslen-Wilson & Tyler, 1980). The N400 word position effect during silent reading supports the the existence of sentence-level context effects in the visual modality.

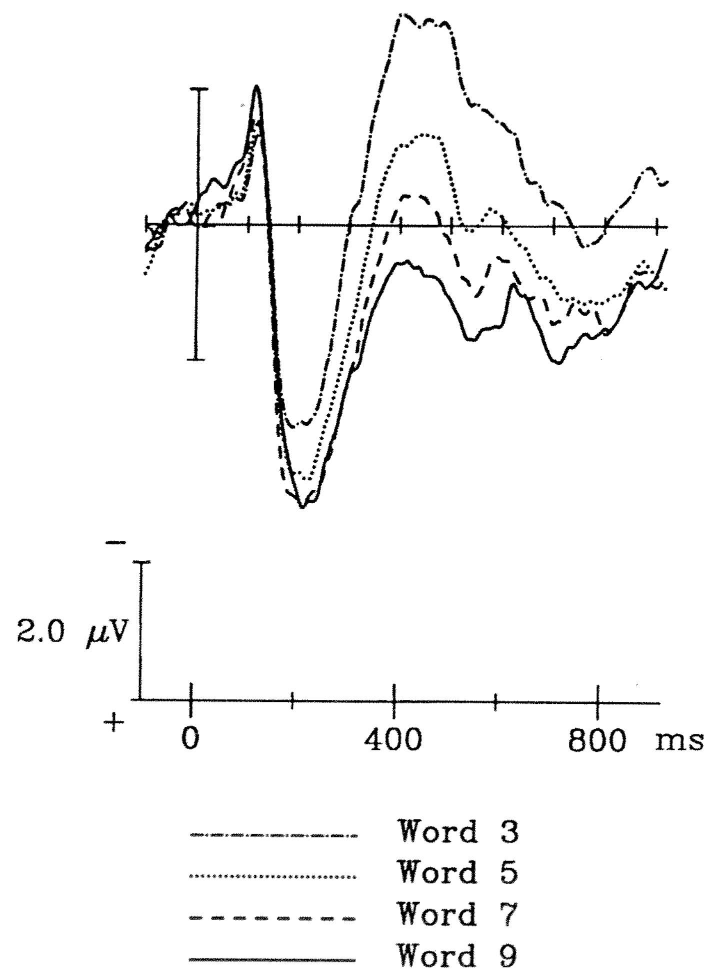

Figure 1. Grand average ERPs elicited by intermediate open-class words in four different sentence positions. Recorded at a central midline scalp site (Cz). Note that while the late negative components elicited at each word position all reach peak amplitude around 400 msec, the differentiation between word positions is apparent as early as 200 msec post stimulus onset (data from Van Petten & Kutas, in press).

In our initial experiment examining word position effects in detail, we also subdivided open class words according to their frequency of usage (Van Petten & Kutas, in press). Six categories of frequency were used, ranging from less than ten to over 450 per million in the Francis and Kučera count (1982). The relatively fine-grained frequency analysis mandated a coarser breakdown of word position in order to have a sufficient number of trials in each frequency-by-position category to form an ERP average with an adequate signal-to-noise ratio.[4] Thus we examined the ERPs elicited by the first open class word of each sentence (typically the second or third ordinal position), all of the intermediate open class words, and the final word at each of the six levels of frequency. For the category of "first open class words" we observed a frequency effect consisting of a larger N400 for rare words as shown in Figure 2. Although the data evidenced a gradient of N400 amplitude across the six frequency categories, the only statistically significant difference was between words of less-than versus greater-than 30/million. We have adopted this as our cutoff point for defining high versus low frequency words in subsequent experiments. The more important result of this experiment was that there was no hint of a frequency effect for intermediate or final open class words.

We have observed the same interaction between word frequency and the position of a word in its sentence in two subsequent experiments. The subsequent experiments answered some additional questions about the interaction. The first of these was actually a reservation as to whether the absence of a frequency effect for intermediate words should be considered a true interaction. When the ERPs for all intermediate words are averaged together as they were in the first experiment, the resulting N400 is fairly small because the incremental effect of sentence context drives amplitude down with each new word. It may thus have been possible that the absence of an observed frequency effect for intermediate words reflected a "floor effect" in N400 amplitude; that is, the influence of sentence context was so great that it swamped a persisting frequency factor. By examining a greater number of word position categories we were able to rule out this possibility. Figure 3 shows that the influence of word frequency is eliminated quite early in a sentence, although N400 amplitude continues to decline with increasing word position (Van Petten, 1989). We can thus regard the absence of a frequency effect late in a sentence despite its presence earlier as a true interaction between word frequency and sentence context.

A second question concerned whether the semantic or syntactic aspects of sentential context, or both, were responsible for attenuating the word frequency effect. In addition to offering some semantic clues about upcoming words, a sentence fragment can also place form class and agreement constraints on words which are legal continuations. We might have suspected from the first experiment that these syntactic constraints are not adequate to suppress the word frequency effect since the first open class words in those stimulus sentences

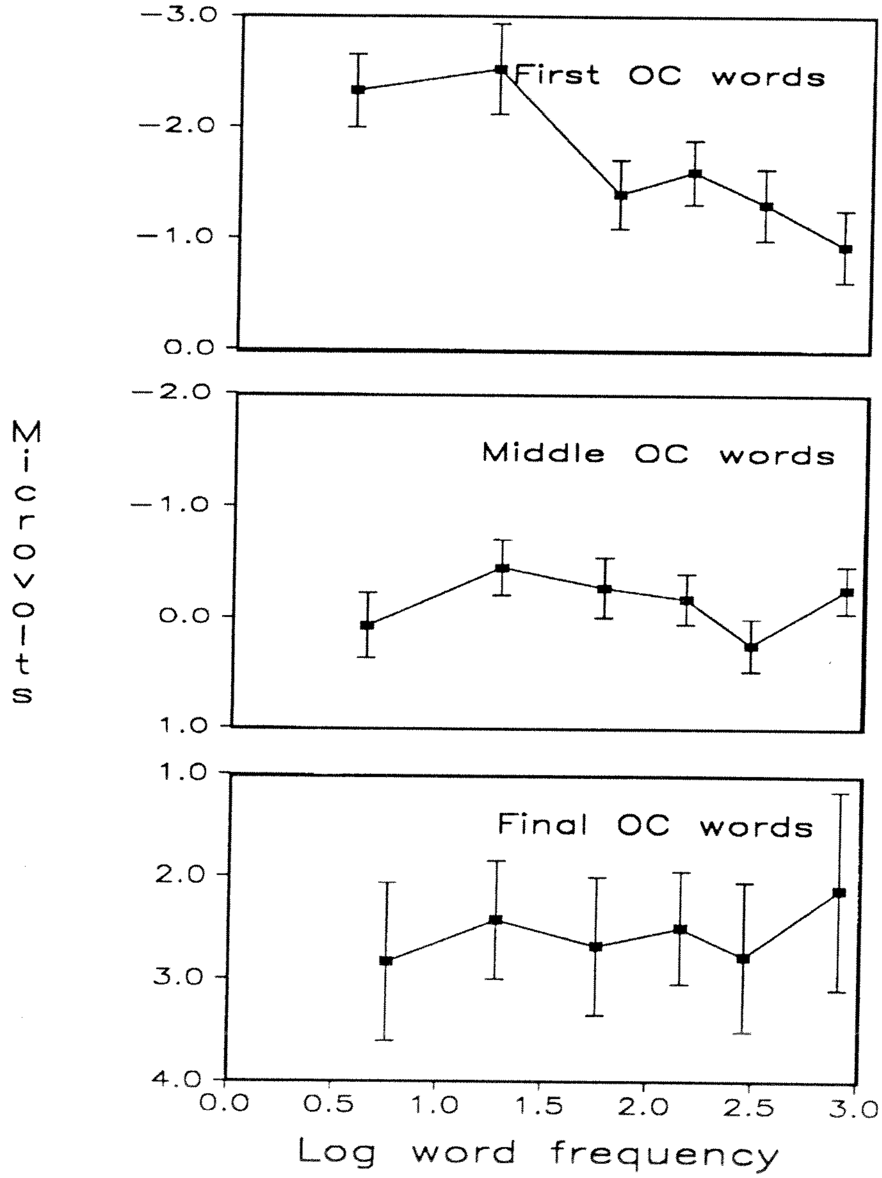

Figure 2. Mean voltage during the peak latency range of the N400 (350-500 msec post-stimulus onset) relative to a 100 msec pre-stimulus baseline, collapsed across recording site. Word frequency has been broken down into six categories for each of three sentence positions: the first open-class words, other intermediate open-class words, and sentence-final open-class words. The error bars represent the standard error across subjects (data from Van Petten & Kutas, in press).

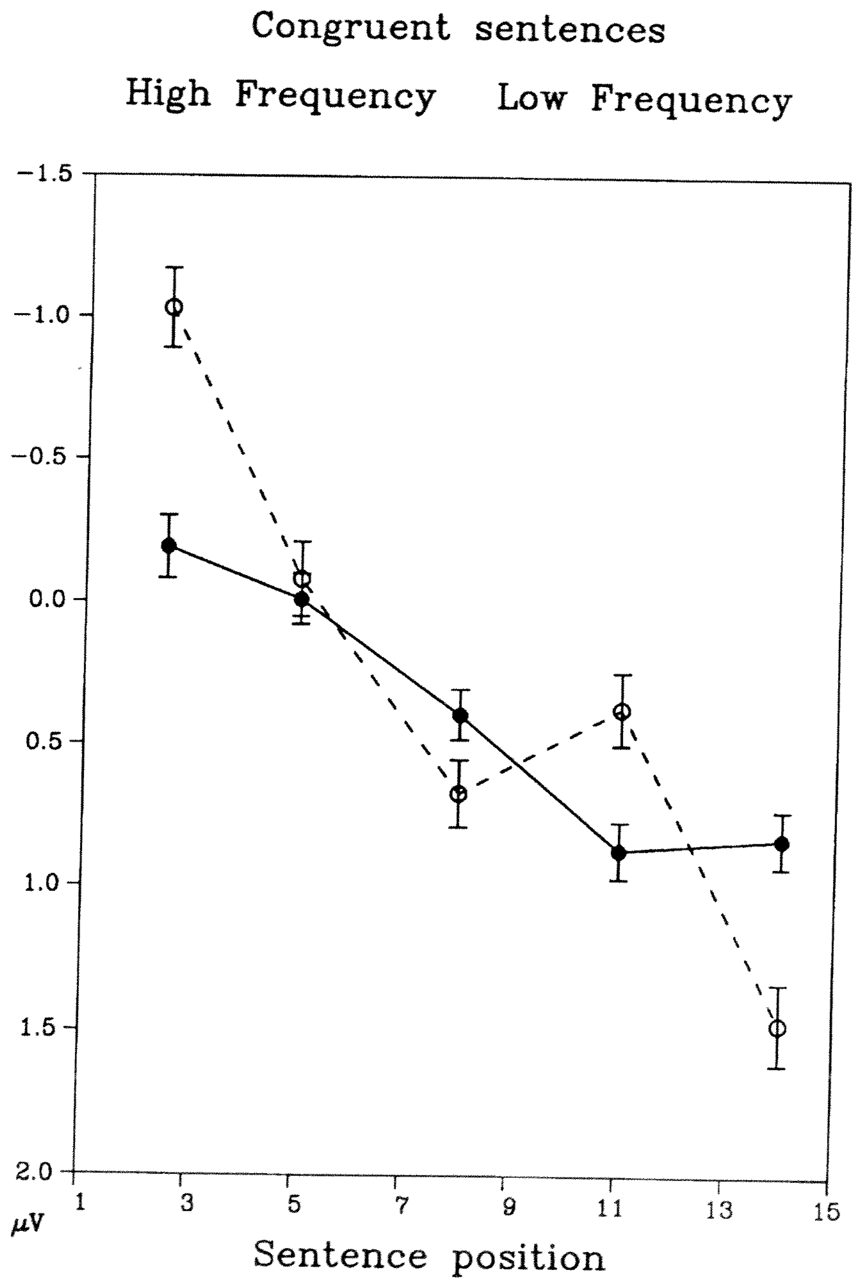

Figure 3. Mean voltage in the N400 peak latency range, plotted against word position for intermediate open-class words in congruent sentences from a different experiment from that shown in Figures 1 and 2 (data from Van Petten, 1989).

were preceded by one or two closed class words. In most cases the first open class words were nouns preceded by articles, or main verbs preceded by auxiliaries so that the syntactic constraints were fairly strong. However, we conducted a second experiment to explicitly test whether or not a legal syntactic structure alone would produce either a word position effect or a position by frequency interaction for open class words (Van Petten, 1989; Van Petten & Kutas, submitted). The materials were modelled on those used by Marslen-Wilson and Tyler in their (1980) word monitoring experiments, consisting of normal congruent sentences, "syntactic" strings which were constructed by replacing each of the open class words in a normal sentence with one of equivalent form class and frequency, and random strings wherein the open class words were again replaced and the closed class words re-ordered within each sentence.[5] Thus, the same closed-class items were present in the Random as in the Syntactic condition, but they were arranged in a non-meaningful way. Examples of each sentence type are shown in Table 1.

Table 1
Sample sentences

Congruent Sentences

The tenants were evicted when they did not pay the last two months rent.
It is supposed to bring seven years bad luck to break a mirror.
Most new drugs are tested on white lab rats.
He was so wrapped up in the past that he never thought about the present.
Everything she owned was in a brown paper bag.

Syntactic Sentences

He ran the half white car even though he couldn't name the raise.
The necklace pulled the certain cat and borrowed the spoon.
He went out of right food and had to go to the black bed.
In the wet levels fathers were smoking by congress.
They married their uranium in store and cigarettes.

Random Sentences

To prided the bury she room she of peanut the had china.
Into thumb cable male the effort his into group rowboat.
She which had jazz anchor a she to straight couldn't gun.
Was reason and ash the angry with technician.
Every opened the gripped they stepping kind steel pine.

The results of this experiment were clear in indicating that semantic processes alone are responsible for both the word position effect and the suppression of the frequency effect. As seen in Figure 4, a frequency effect was apparent for all of the open class words in the Syntactic and Random strings, but only for early words in the Congruent sentences. The usual decrement of N400 amplitude was present across the course of meaningful sentences but absent in both classes of meaningless strings. Of course, there are two alternative interpretations for this pattern of results. On the one hand, the experiment may have demonstrated that a syntactic structure provides very little information concerning upcoming words. On the other hand, we might suppose that such information is present and used by the reader, but that semantic and syntactic information are handled by different brain processors and the N400 measure is insensitive to the use of one class of information. We were able to rule out the second alternative by an examination of the ERPs elicited by the closed class words in the three sentence types.

Although much smaller than the open class N400, the amplitude of the N400 to closed class words showed a three-way difference across the conditions. It was smallest when the closed class words were embedded in congruent sentences, of intermediate amplitude in the Syntactic strings, and largest in the Random strings. We thus have an indication that the N400 can serve as an index of the utilization of both sources of information.

Several studies have demonstrated that, when confronted with a "fill-in-the-blank," or cloze, procedure, subjects are much more accurate in predicting function than content words (Aborn, Rubenstein, & Sterling, 1959; Gough, 1983; Smith-Burke & Gingrich, 1979). The present results suggest that this predictive power derives from both semantic and syntactic constraints. In contrast to closed class words, the similarity of the open class ERPs between the Syntactic and Random conditions indicates that syntactic structure alone places few constraints on these words. However, in meaningful sentences, N400 amplitude begins to decline after only a few words, suggesting that even the most incomplete semantic structure is quite powerful (see Tyler & Wessels, 1983; Tyler & Marslen-Wilson, 1986 for similar conclusions based on a different dependent measure).

The foregoing experiments established that the word position effect for open class words was a consequence of semantic analyses. Thus the absence of a word frequency effect late in a sentence must likewise be attributed to semantic processes. We can now consider what these results might tell us about the respective loci and timing of frequency and meaning-based effects on lexical processing. Word frequency effects have received such a great deal of research attention over the last thirty years that a number of sources have been proposed for their existence, including each of the stages in the five-stage sequence listed above. We first will describe those proposals which we can eliminate from consideration as explanations of the N400 frequency effect, and then discuss how suppression of this frequency effect by context might be interpreted in various models of word recognition.

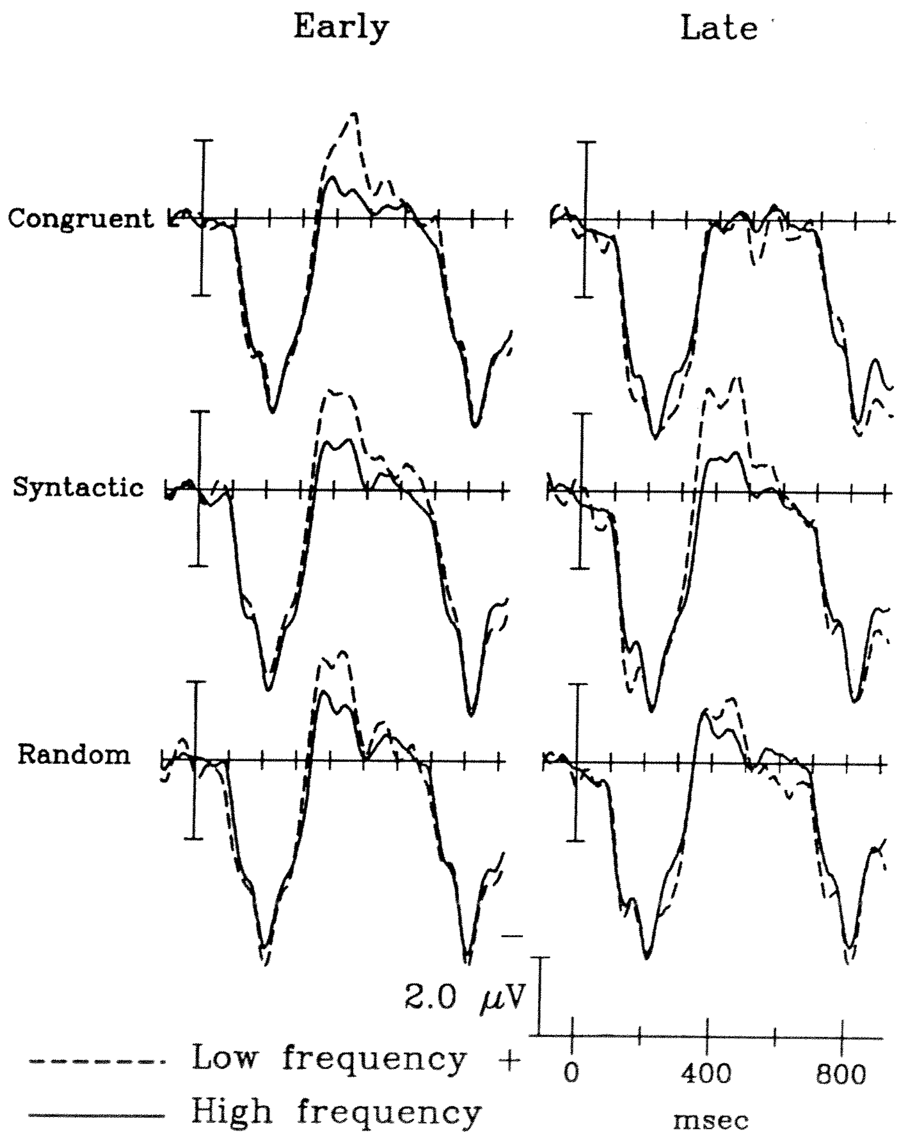

Figure 4. ERPs elicited by open-class words in three sentence types. Early and late refer here to approximately the first and second halves of the sentences, excluding the initial and final words (data from Van Petten, 1989; Van Petten & Kutas, submitted).

Visual encoding

After demonstrating that the letter bigrams and phonemes comprising low frequency words are themselves rare, Landauer and Streeter (1973) put forth the following proposal: "Suppose that some phonemes or graphemes are easier to produce or perceive than others. Then those words composed of good units should have an advantage over those composed of poorer units" (p. 121). In other words, frequency might have an impact at a perceptual stage prior to any expressly lexical processes, as in the first stage of the five-stage sequence outlined above. In support of this idea, Landauer and Streeter demonstrated that words containing low frequency phonemes were less intelligible in noise than words composed of high frequency phonemes even when the two sets of words were matched for both printed and spoken frequency. However, Gernsbacher (1984) found that bigram frequency did not contribute to word frequency effects for printed words: when factorially combined with experiental familiarity, bigram frequency did not influence lexical decision times whereas familiarity did.

Sentence integration

A proponent of the sequential processing scheme outlined earlier might argue that some late phase of sentence comprehension, as in Stage 5 (e.g., beyond the lexicon), is the sole source of the the larger N400 to low frequency words. By this view, there may be an "earlier" frequency-sensitive phase of lexical access, but this is simply not indexed by N400 amplitude. Accordingly, the observed interaction between frequency and sentence context would not reveal fundamental processes in word recognition but rather some increased difficulty during the integration of low frequency words with an established context. We find this argument untenable for two reasons. First, other laboratories have reported that low frequency words elicit larger N400s than high when the words are seen in lists for lexical decision, clearly a situation where there is little call for integrative processes (Rugg, in press; Smith & Halgren, 1987). Furthermore, within a sentence context, we have repeatedly observed word frequency effects for the first open class of a sentence and find the argument that it would be more difficult to integrate a low frequency noun than a high frequency one with a preceding article (e.g., "the squirrel" versus "the rock") a bit strained.

Task specific stages

Frequency-sensitive mechanisms have also been proposed for the opposite end of the perceptual-motor continuum, near the output stages of the most common tasks employed in word recognition research. The decision stage of the lexical decision task may be prolonged, or in naming tasks it may take longer to assemble the articulatory code for low frequency words (Balota & Chumbley, 1984, 1985; McCann, Besner, & Davelaar, 1988; Theios & Muise, 1977). Attributing the word frequency effect to these task-specific factors in its entirety has been a controversial idea (see Monsell, Doyle, & Haggard, 1989) but one which we need not consider here. The N400 frequency effects we have observed occurred several seconds or minutes before a motor response was required of our subjects.

Models of Word Recognition which Predict Additivity of Frequency and Context

Having ruled out very early and very late operations as accounts of the N400 frequency effect, we can conclude that it arises sometime during the intermediate stages of lexical processing, recognition and meaning access, loosely speaking. In fact, this has also been the standard assumption of most models of word recognition of how frequency influences behavioral responses (for summaries, see Norris, 1986; or Van Petten & Kutas, in press). However, among these models are some which postulate that word frequency and context effects should be additive, either because frequency of usage imposes a stable difference in the recognition criterion for a word (Morton, 1969) or because higher frequency entries are examined first in a serial search process (Bradley & Forster, 1987; Forster, 1981a, 1981b). These models thus cannot accommodate an interaction between frequency and context because frequency is presumed to have had its impact sometime prior to the start of meaning access.

We should next consider some models which allow semantic analyses to overlap in time with frequency-sensitive processes. We will see that while this stipulation allows for the possibility of an interaction between frequency and context, it is more difficult to describe the form of the interaction we have observed within extant models, namely, a complete elimination of a frequency effect by sentence context.

The Checking Model

In Norris's Checking model (1986), slower responses to low frequency words are accounted for by a higher recognition criterion much as in the Logogen model. The distinction between the two models is that the recognition criterion can be dynamically altered by reccurring contextual plausibility "checks." The initial high criterion for low frequency words allows them to accrue more "checks" and consequently more criterion reductions than high frequency words. The model thus predicts a disproportionate influence of context on low frequency words. However, because both high and low frequency words are subject to such criterion reductions on the basis of context, the criterion for low frequency words can approach but never reach that of high frequency words. This form of interaction is what has been observed in the lexical decision and naming latency data which Norris attempts to explain (Becker, 1979; and Stanovich & West, 1983), but not what we have observed. It is possible that the concept of "criterion" is too bound up with the notion of passing a threshold for the elicitation of an all-or-none behavioral response to be relevant to situations where no motor response is called for.

The Cohort Model

In the most recent formulation of the Cohort model it is supposed that perceptual recognition processes (e.g., Stage 2 above) proceed more quickly for high frequency words, so that they can have access to semantic processes earlier than low frequency words. As it is a model of spoken-word recognition, it is possible to collect data more directly related to this claim than in the visual modality. In her dissertation work, Zwitserlood (1989, see also Marslen-Wilson, 1987) presented visual probe words for lexical decision at different time points during an auditory word embedded in a sentence. High and low frequency words were interrupted at the same temporal positions, thus yielding data which is more difficult to obtain in the visual modality. At the earlier probe points, the acoustic signal had not yet unambiguously specified the identity of the eventual word. For instance, when her subjects heard "The men mourned the loss of their cap...", both "captain" (the actual word) and "captive" were equally consistent with the acoustic input. At this time point, visual probe words related to both possible words displayed a priming effect in lexical decision times, but the larger priming effect was for probes related to the higher frequency member of the competing pair. This result does therefore suggest that high frequency words might begin to participate in semantic processes earlier than low frequency words during auditory word recognition. This frequency advantage proved to be transient, at later probe points lexical decision times for associates of high and low frequency words became more equivalent.

We have not observed a similar early difference between high and low frequency printed words to suggest that high frequency words gain access to meaning more quickly. Sentence context effects begin at the same time relative to the onset of the word in the ERP. Figure 5 shows a word position effect for high and low frequency words (data from Van Petten, 1989). The inconsistency between the two patterns of results may reflect a difference between the auditory and visual modalities of language comprehension. Or it may reflect a methodological difference: we have based our conclusions on the measurement of responses to words in sentences, Zwitserlood's were based on responses to probe words presented subsequent to sentence words. We will return to this methodological point later.

Frequency Effects in Connectionist Models

Word frequency effects arise naturally in connectionist models because the weights between units are modified with each presentation of a word; more frequently presented words have a greater impact on the strength of the connections between units in different layers (see Seidenberg & McClelland, 1989). In various models, semantic context effects have arisen in two ways: via the abstraction of patterns of word co-occurrences in sentences (Harris, 1989; McClelland, St. John, & Taraban, 1989), or via the explicit encoding of a "word" as a collection of semantic features which may be shared with other words (Kawamoto, 1988). The single (to our knowledge) connectionist model incorporating both word frequency and semantics used the latter scheme for the semantic representation (Sharkey, 1989). In this model, the strength of the connections between graphemic input features and semantic features increased with the frequency of the word, thereby allowing more rapid activation of corresponding semantic features for high frequency inputs. As prior semantic context brought the state of the network closer to that for the next word less adjustment was needed to reach the new target state when the next word was actually encountered. The model thus yielded an interaction between frequency and context because the rate of movement toward the target state (i.e. the effect of frequency) became less important as the distance to be travelled decreased. Nonetheless, this model likewise cannot explain an absolute elimination of the frequency effect by context: unless the network is already in the desired target state when a new word arrives, speed of movement will always be somewhat important.

The Verification Model

The sole proposal which we find able to accommodate the form of the frequency by context interaction observed in the ERP data is Becker's Verification model (1976, 1979, 1980). In this model, the initial sensory analysis of a word yields a set of candidate words which are ordered by frequency. However, even before the word is actually encountered, the reader begins to generate a "semantic set" of candidates based on the prior context: this set is ordered by strength of semantic relationship rather than frequency of usage. Because readers are presumed to search the "semantic set" first, frequency effects obtain only when the stimulus word is not included in this set (i.e., not related to the previous context) and the reader is forced to check the "sensory set." Thus, given appropriate context, no frequency effects would be observed. While we appreciate the priority of semantics offered by this model, the serial nature of the search process leads to some additional predictions which we find implausible. If the "semantic set" proves to be larger than the sensory set because prior context

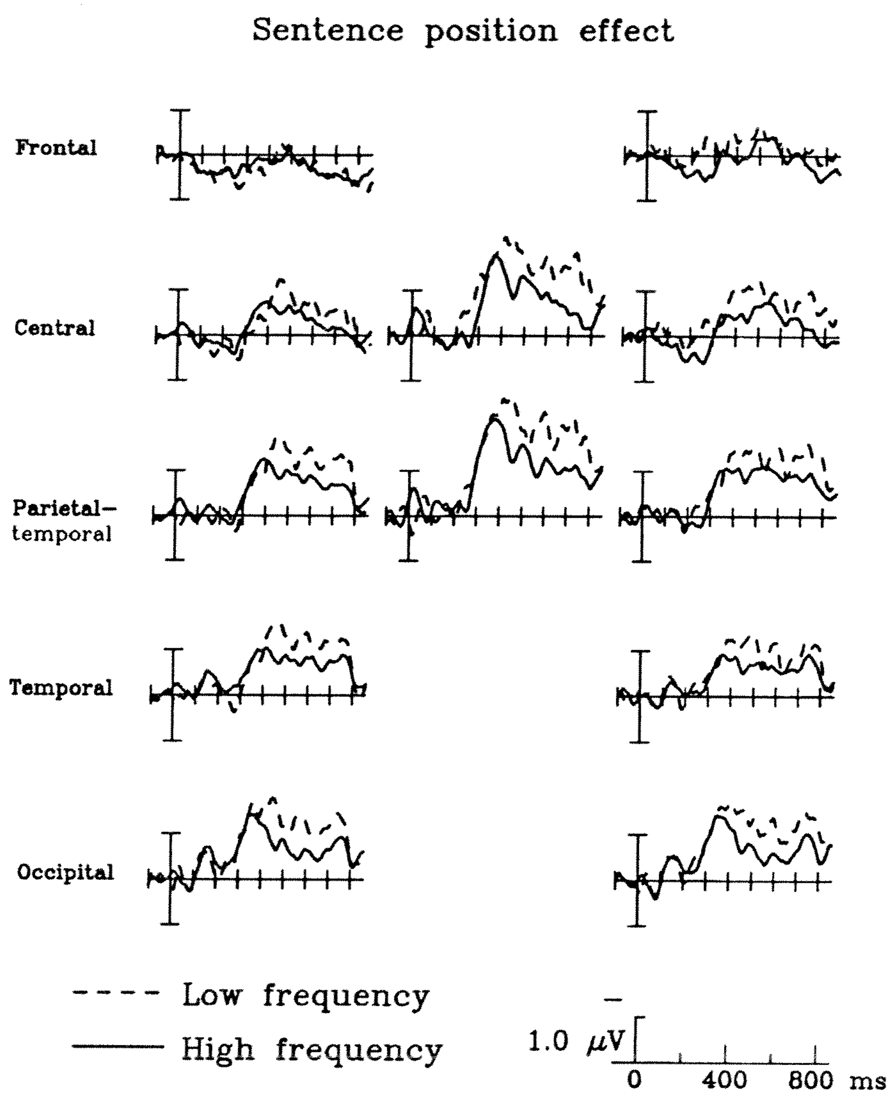

Figure 5. ERP difference waves formed by subtracting the responses to open-class words occurring late in congruent sentences from those occurring early. Data recorded from left hemisphere electrode sites is shown in the left-hand column, midline sites in the middle column, right hemisphere sites in the right-hand column. Note that the early (100 to 200 msec post-stimulus) negative wave seen at temporal and occipital sites is not part of the N400 effect. Rather, the difference wave reflects a refractory effect in this visual sensory component; amplitude of the "N180" falls off rapidly during a serial train of stimuli (see Hillyard, Munte, & Neville, 1985). Data from Van Petten (1989).

Table 2

Sample Sentences

Congruent Associated

When the MOON is full it is hard to see many STARS or the Milky Way.
There were advantages to living in a CITY but Martha moved to a small town for the peace and quiet.
After studying the map she realized they should have turned LEFT instead of RIGHT at the light.
She was glad she had brought a BOOK since there was nothing to READ in the waiting room.

Congruent Unassociated

When the INSURANCE investigators found out that he'd been drinking they REFUSED to pay the claim.
The biologist went to the desert every WEEK to collect a particular SPECIES of lizard that he hoped to study.
The union officials were worried about the long term health HAZARDS of breathing CHEMICAL fumes every day.
She picked up a wallet on the STREET and was honest enough to TRY to locate the owner.

Anomalous Associated

When the MOON is rusted it is available to buy Many STARS or the Santa Ana.
There was jewelry to drumming in a CITY but Martha turned to a grey TOWN for the lizard and scones.
After fixing the movie she found they should have killed LEFT instead of RIGHT at the pot.
She was glad she had waved a BOOK since there was everyone to READ in the security child.

Anomalous Unassociated

When the INSURANCE supplies explained that he'd been complaining they REFUSED to speak the keys.
The shirt went to the gun every WEEK to keep a good SPECIES of fumes that it hired to see.
She scrambled up an official black on the STREET and was deep enough to TRY to ring the glue.
The star hair was worried about the bared hard drinking HAZARDS of signing CHEMICAL boxes every town.

Note: The critical pairs of words are shown capitalized, although the subjects saw them in normal typeface.

only loosely constrains the current word, preferential search through this set would result in slower word recognition than if there had been no context whatsoever. As sentence fragments are rarely predictive of a single word or even a handful of words, it seems as if the Verification model would rarely stipulate a beneficial role for sentence context.

After reviewing a number of existing models of word recognition, we conclude that none provide an accurate account of the form of the context by frequency interaction we have observed in several experiments. In large part, this is due to the fact that each model presupposes that frequency and context act via different mechanisms. Even in the case where a formally identical mechanism is proposed to account for both (i.e., Norris's recognition criterion), its application would entail a different time course for the two: frequency is presumed to have pre-set the criterion long before the currently processed word is encountered, whereas context can alter the criterion only after the word's presentation.

Although we do not have a new model of word recognition to contrast to those above, we would like to call attention to one possible locus for the word frequency effect that has been relatively overlooked. Perhaps semantic analysis is the source of frequency effects. Where a semantically-based frequency effect has been considered, it has been restricted to quantitative dimensions such as concreteness or number of meanings, neither of which seems to contribute to frequency effects (Gernsbacher, 1984; see, however, Schreuder & Flores d'Arcais, 1989 for a discussion of the ease of accessing different types of semantic attributes). However, many low frequency words are little used because they possess semantic attributes that are not often called for. If the rarity of a word's semantic features rather than its physical form were responsible for frequency effects, we would expect the frequency effect to disappear with contexts that increased the predictability of these uncommon features. This speculation could be tested by comparing responses to low frequency words which possess high frequency synonyms to those which do not. If rarity of meaning is the basis of word frequency effects, then semantically isolated words should result in slower response times and larger N400s than words whose meanings overlap with more commonly used words.

Lexical versus Sentential Contexts

Lexical priming has often been assigned a different status from sentential context effects. A common theoretical stance views lexical-associative priming as a reflection of pre-stored connections within an autonomous lexicon and sentence-level context effects as a consequence of subsequent comprehension processes that act on the output of the lexicon (Fodor, 1983; Forster, 1981b; Kintsch, 1988; Seidenberg et al., 1982; Swinney, 1979). Recently, we have completed an experiment which was designed to compare these two types of context effects, taking advantage of the temporally continuous dependent measure offered by the ERP as a way of evaluating their relative time courses (Van Petten, 1989).

The materials for this experiment consisted of four types of sentences, each containing a critical pair of words. The same associated pairs of words (such as SALT-PEPPER) were embedded in both congruent and syntactically legal but semantically anomalous sentences ("Congruent Associated" and "Anomalous Associated" respectively). Likewise, pairs of words which were not particularly related to one another outside of a sentence context also were embedded in both congruent and anomalous sentences ("Congruent Unassociated" and "Anomalous Unassociated"). The overall lengths of the sentences, the positions of the two members of a critical pair in their sentences and the number and lexical class (open or closed) of the words intervening between the two members of the pair were matched across the four sentence types. Examples of the four sentence types are shown in Table 2. Outside of the critical associated pairs, an attempt was made to avoid including sentence words which seemed clearly associated to one another.[6] With this design, the second word of a pair could thus benefit from 1) both sentential and lexical context in the Congruent Associated condition, 2) lexical context alone in the Anomalous Associated condition, 3) sentential context alone in the Congruent Associated condition, or 4) neither in the Anomalous Unassociated condition.

The ERPs elicited by the critical pairs are shown in Figure 6. As expected, a reduction in N400 amplitude from the first to the second word of a pair was observed in all but one of the conditions, the Anomalous Unassociated. The overall word position effect from this experiment is shown in Figure 3. We have taken this more general word position effect as a demonstration of the influence of sentence context on N400 amplitude, and view the decrement in amplitude from the first to the second members of the Congruent Unassociated pairs as a subset of the general trend across a sentence. The associative priming effects were not surprising either, given that a number of laboratories have demonstrated modulations of N400 amplitude in word pair studies (see citations earlier). However, the present experiment allows a comparison of the two types of context effects when obtained in similar circumstances from the same group of subjects. Our analysis will thus focus on the timecourse of the two context effects, taking the difference between the first and second words of the Anomalous Associated pairs as a purely lexical context effect, and the difference between the first and second words of the Congruent Unassociated to be a purely sentential effect.[7]

The time course of a priming or context effect can logically be divided into two parameters: rise time and decay rate. In behavioral experiments, both are usually inferred from the effects of varying the interval between a context and a target word. Rise time is described by the shortest context-target interval that affords priming, and decay rate by the longest interval which sustains significant priming effects. However, both rise time and decay rate can be further subdivided. For rise time, one of these is the length of time it takes a subject to process previous words to the point where they become a potential source of context. It is this component which can be estimated by manipulating the SOA between a current word and prior context. A second component of rise time reflects the exact moment at which this processed information is actually

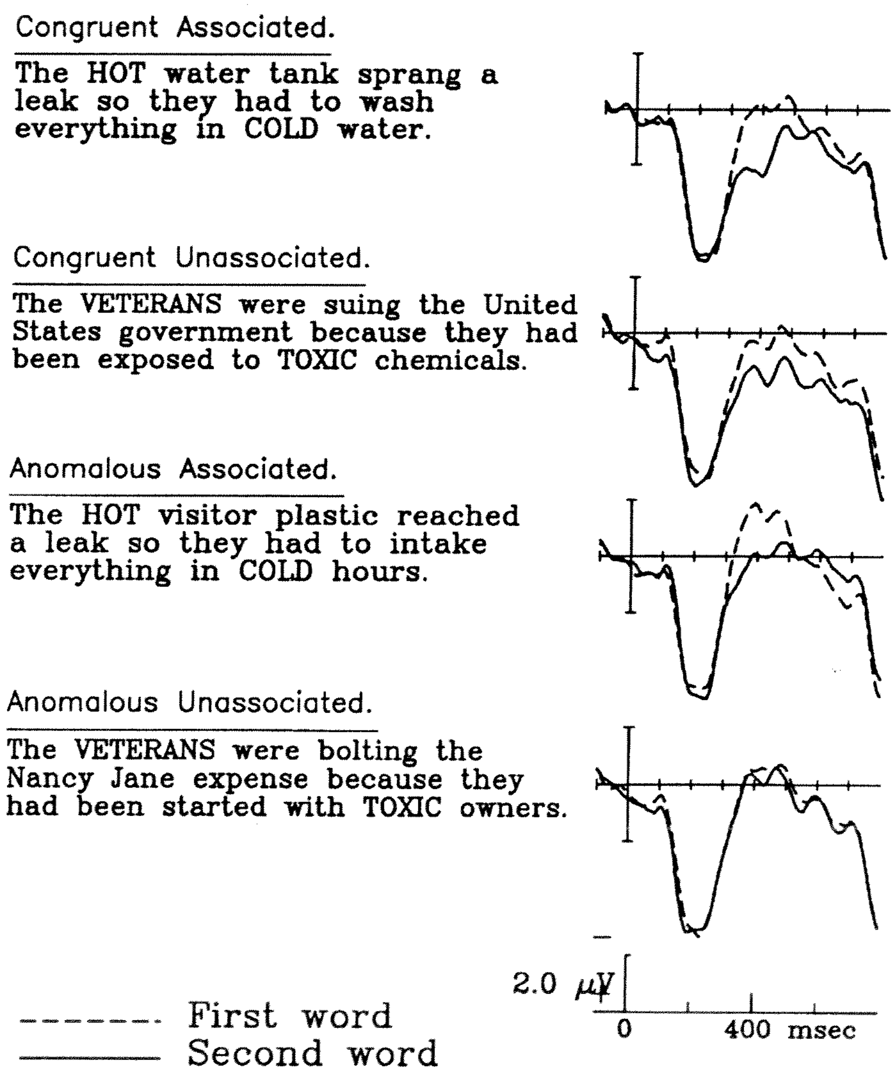

Figure 6. ERPs to the critical pairs of words (shown capitalized) in each of the four sentence types used to contrast sentential and lexical contexts.

applied to the analysis of the current word. Speech researchers have been sensitive to the distinction between these different components of rise time because the sequential nature of the acoustic signal allows measurements to be taken at different time points during the input signal (see Zwitserlood, 1989 or Grosjean, 1980). Because printed words are not spread out across time, these two rise time parameters have often been conflated in studies of visual word recognition. There is little question that it takes longer to extract meaning from a sentence fragment than a single word. But this fact has little bearing on the issue of whether or not word recognition draws on all potentially available semantic constraints, or if some information has privileged access.

The decay rate parameter can similarly be subdivided. On the one hand, we can estimate the temporal extent of a context effect during the processing of a single word. On the other hand, we can wonder how many words downstream from the relevant context are still susceptible to its influence. In the present experiment we were able to evaluate three of these four temporal factors.

Rise Time

A comparison of the onset latencies of the lexical and sentential context effects offers a window onto the second component of "rise time," namely when a context effect is evident during the processing of a given word. As these onset times were indistinguishable, we found no evidence for the view that there is a distinct stage in the processing of a word which can be influenced by associative links, but not by sentence-level context.[8] The first component of "rise time," the time it takes for a sentence context to become available (as opposed to the time it takes to be used, once available) can only be evaluated by varying the rate at which sentence words are presented. Throughout this experiment, words were presented at a single rate of one word every 600 msec. However, we are currently conducting an experiment wherein the same stimulus materials are presented at a rate of one word every 300 msec; comparison between the two experiments will allow for an evaluation of this temporal factor. Given that this faster rate approximates that of fluent reading, similar results will provide strong support not only for the theoretical position that associative and sentential context *can* have the same timecourse, but also for the practical position that readers are able to avail themselves of sentence context during the processing of individual words in more natural situations.

Decay Rate

For sentential context effects, we have already seen that there is essentially no limit to the number of words downstream which may be influenced by prior context. The general word position effect indicates that sentential context is incremental, each new word benefitting from the entire previous fragment. Lexical context effects have been thought to be more prone to disruption by unrelated words or long delays between the two members of the associated pair. This seems to be true when words are presented in lists (Foss, 1982; Gough, Alford, & Holley-Wilcox, 1981; Neely, 1977; Warren, 1972), but not when they occur in congruent sentences (Carrol & Slowiaczek, 1986; Foss, 1982; Simpson et al., 1989). The anomalous sentences used here fall somewhere between the two extremes in being more structured than a random word list, but less so than a congruent sentence. We divided the associated pairs occurring in anomalous sentences into two equal-sized classes of "near" and "far" pairs. The members of "near" pairs occurred either immediately adjacent to one another, or had one word intervening. "Far" pairs had more than one word intervening, with an average of 4.8 words. No difference in the magnitude of the lexical priming effect was observed as a function of number of intervening words. Together with the general finding that associative priming effects do fall off with distance in random lists, the present null result is consistent with the suggestion of O'Seaghdha (1989) that a syntactic structure licenses the more effective use of lexical associations. It is also consistent with our previous suggestion that subjects are probably more actively engaged or attentive when reading sentence-like materials than when reading random word strings (Van Petten, 1989; Van Petten & Kutas, submitted).

Evaluation of the other component of decay time yielded interesting results in that we observed a difference between lexical and sentential context effects. While both effects had the same onset latency at about 300 msec relative to stimulus presentation, the sentential priming effect proved to have a longer duration as is apparent in Figure 6. The sentence context effect persisted late in the recording epoch (500 to 700 msec post-stimulus) while there was no lexical effect in this latency window. A broad duration sentential effect was observed for the congruent sentences both with and without lexical associates (see Van Petten, 1989 for a lengthier discussion of this result).

We have found across several experiments that all of the N400 effects which are due to sentence context (i.e. comparisons of responses to words occurring early versus late in congruent sentences, words in congruent versus anomalous sentences, congruent versus incongruent words in otherwise normal sentences) have a broader temporal duration than N400 effects arising from lexical-associative priming or differences in word frequency (see Figures 1, 4 and 6). This durational difference may indicate that while sentence context word processing in the same time frame as these other factors, it also demands more extended processing of single words. Within the present comparison of associative and sentence contexts, it seems likely that comprehension of a sentence encourages a more detailed semantic analysis of each word than does lexical context alone. In isolation, or in an anomalous sentence, one might analyze the relationship between "salt" and "pepper" rather quickly and superficially as "those black and white substances in the twin shakers." Compare this to the senses of "salt" and "pepper" evoked by one of the congruent sentences used here: "He was trying to cut down on his SALT intake so he used a lot of PEPPER and other spices." This sentence suggests not only the white substance in the shaker, but also the fact that it may lead to high blood pressure, and perhaps the fact that salt contains sodium. On reaching "pepper," the reader is invited to infer that it has not been associated with high blood pressure, can serve as an alternative flavoring agent, etc. Accordingly, the longer duration of sentence context effects in the ERP may reflect prolonged processing of more encyclopedic semantic attributes of words as they are generally needed for sentence comprehension. Such a speculation could be tested by comparing ERPs to the same words embedded in contexts which refer to their core or peripheral senses.

Sentence Context and Lexical Ambiguity

Above, we speculated that the later portion of the N400 effects we observed in the context of meaningful sentences might reflect the encyclopedic elaboration of single word meanings and/or inferential processes. However, the results also indicated that some aspects of sentence context operate as quickly as lexical-associative context; perhaps these aspects of sentence context aid in focusing on the relevant core semantic attributes of the current word. We began this chapter by outlining a five-stage sequence of events occurring during sentence comprehension, one which we take to be a widely-accepted view, but one which much of our results have been inconsistent with. In this view, sentence context cannot select the relevant semantic attributes of a word until all of its core attributes are generated by a prior lexical processor. This early phase of generating meanings has been held to be mandatory, indiscriminate, and exhaustive (Fodor, 1983).

One of the key findings in supporting this view has arisen from the literature on lexical ambiguity. The typical experimental paradigm in this literature consists of embedding a homographic word in a sentence which clearly disambiguates it, and recording responses to target words which follow the homograph. It is usually found that, with a brief SOA between homograph and target, targets related to either meaning of the homograph show a priming effect. With a longer SOA, priming is observed only for the associates of the sententially appropriate sense of the homograph (Kintsch & Mross, 1985; Onifer & Swinney, 1985; Seidenberg et al., 1982; Swinney, 1979; Till et al., 1988). Priming for both appropriate and inappropriate associates is observed with SOAs of less than 200 msec or so, whereas longer intervals between the two words yields selective priming. This finding has been the primary evidence for a contrast between an early stage of word processing wherein all core (at least) meanings are activated (and lexical-associative priming occurs) and a later phase wherein sentence context is used to select the relevant meaning from among the candidates. The result has been so persuasive that even Kintsch, who has otherwise described his most recent model of language comprehension as being one where "word identification is ... deeply embedded into the process of discourse understanding" (1988, pp 173), has included a stage involving random sampling of associated words prior to any sentence integration.

When a single experimental finding supports so much theoretical infrastructure, it is reasonable to scrutinize both the data and the interpretation carefully. We believe there are both empirical and logical flaws in the conclusion. On the empirical side, the postulate that the initial access to a word's meaning is context-invariant should predict not only that priming be found for inappropriate associates of ambiguous words, but that it be equivalent in all respects to the priming for appropriate associates. Thus, greater priming for the appropriate probe at a short SOA would violate the modularity of the activation process and suggest an early influence of the preceding sentence context. This is indeed what has been reported in the majority of ambiguity studies (see Blutner & Sommer, 1988; Oden & Spira, 1983; Onifer & Swinney, 1981; Seidenberg et al., 1982; Simpson, 1984; Till et al., 1988; Swinney, 1979; Van Petten & Kutas, 1987, Exp. 1). In some cases this difference has been little noted because it did not reach statistical significance. Recently, however, St. John (1988) conducted a meta-analysis in which the results of these different experiments were analyzed jointly and found that the difference between appropriate and inappropriate associates was statistically significant. There are thus good empirical grounds to question the conclusion of context-invariant meaning activation.

Likewise, we can question the conceptual underpinnings of the experimental procedure. The logic of the homograph/probe-word paradigm is that, due to spreading activation, a response to the probe can reveal which meanings of the preceding word were activated and in what sequence. It is assumed that the processing of the probe itself is of no consequence. Although they are presented very close together in time, their is an assumption of strict seriality in the processing of the two words.[9] There has been an additional assumption that setting the SOA between the homograph and the probe to a particular interval is equivalent to taking a "snapshot" of the processing state of the homograph at that time point. For instance, based on manipulations of SOA, Till and colleagues have stated that: 1) "associates of a priming word are equally facilitated ... whether or not they are appropriate to the discourse context - at least for the first 300 msec of processing," 2) "Sense activation is achieved within 400 msec" and 3) "inference words are not activated by the discourse context ... until the prime word has been processed for more than 500 msec" (Till et al., 1988). These time estimates for the processing of prime words were based on reaction times to subsequent probe words, reaction times which exceeded 600 msec.

The danger of equating SOA to processing time is most clear if we pursue the logic to its extreme. Associative priming effects have been observed with an SOA of 16 msec (Simpson & Burgess, 1985), but we surely don't want to conclude that the meaning of the first word has been accessed within 16 msec, placing the processes of interest somewhere in the retina or the lateral geniculate. With interstimulus intervals as short as this, some processing of the prime clearly occurs after the presentation of the target. In this situation, the priming effect must be due to coincident processing of the two words rather than to a preactivation of the second word by the first.

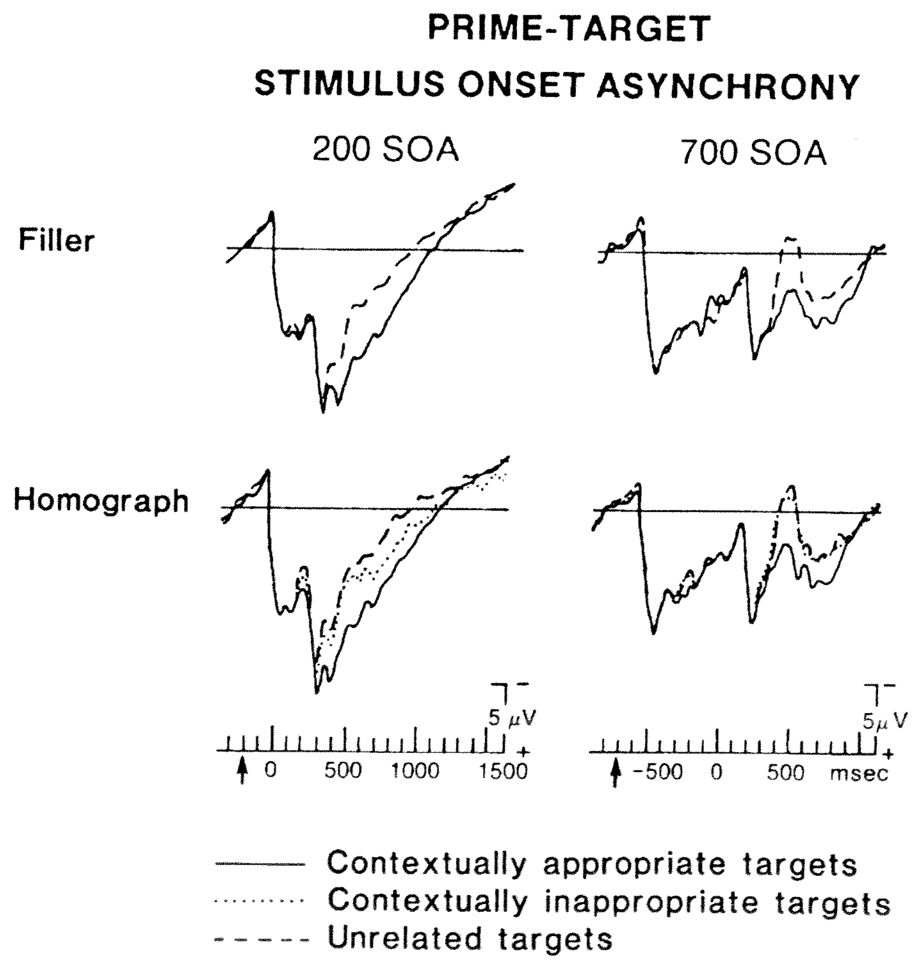

Figure 7. ERPs elicited by sentence-terminal words and subsequent probe words at a central midline scalp site (Cz). "Filler" sentences ended with unambiguous words. In all four plots, the sentence-terminal words were presented at the time indicated by the arrow and probe words at time 0.

The standard interpretation of the ambiguity literature is that readers or listeners simultaneously access all meanings of homographic words upon seeing or hearing them, and this yields the subsequent priming effect for associates of all meanings. However, if there is temporal overlap in the processing of prime and target words in the short SOA conditions of lexical ambiguity experiments, then the usual interpretation would be invalidated. The facilitated response to the inappropriate probe could reflect mutual priming between the homograph and the target; this situation would only arise in the laboratory setting. The reasonableness of this interpretation is suggested by the fact that, independent of the priming condition, reaction times are often slower following short SOAs (by some 160 msec in Exp 2 of Till et al.). This pattern of results would suggest that processing of the prime word is not *truncated* by presentation of the target, rather the target word is *added* to the subject's processing load, allowing for interactive processing of the two current words. We are then presented with an empirical question that must be resolved before the ambiguity results can be interpreted unequivocally: what is the longest separation between two words which will result in temporal overlap? There has been a paucity of data that is directly applicable to this question. We will describe that which exists shortly, but first digress to describe the results of an ERP ambiguity study which underscores our belief that this is a critical question.

While the probe word technique itself has some built-in interpretative complications, the most severe of these arise in conjunction with the use of a discrete dependent measure. Some years ago, we adopted the lexical ambiguity probe word technique from the behavioral literature but used the ERP to provide a more detailed picture of the timecourse of context/probe processing (Van Petten & Kutas, 1987). We presented sentences which biased one reading of an ambiguous word, followed by a target word that was either 1) related to the sententially appropriate sense, 2) related to the inappropriate sense, or 3) unrelated. With a 700 msec interval (onset to onset) between ambiguity and target, the appropriate target elicited a smaller N400 than did the unrelated. As seen in Figure 7, this N400 difference was apparent beginning about 300 msec after the onset of the probe word. In contrast, the inappropriate target elicited an N400 which was equivalent to the unrelated target. Given this long interval, no priming was apparent for the inappropriate target. However, with a short SOA between ambiguity and target (200 msec onset-to-onset), we observed a priming effect even for inappropriate targets. The ERP data thus demonstrated the same sensitivity to the temporal interval between ambiguity and target as reaction time data. However, the priming effect observed for inappropriate probes in the short SOA condition was substantially delayed relative to that for the appropriate probe; it was not apparent until some 500 msec after target onset.[10] These data therefore did not support the view that both meanings of the ambiguous words were simultaneously activated. In line with the principle of parsimony, we attributed the different time courses of the two priming effects seen in the short SOA condition to the different times at which contextual information was delivered. In one case, the preceding sentence served as context to interpret the ambiguous word thereby producing an early priming effect for the subsequent probe related to this meaning. In the other case, the inappropriate probe itself took on a double role not only as probe but as context for the alternative interpretation of the ambiguous word.

The ERP data are conclusive as an indication that the sententially appropriate and inappropriate senses of ambiguous words are not equally weighted. The proposed mechanism for how the inappropriate meaning of the ambiguous word is activated in studies of this type is more speculative. It hinges largely on the extent to which mutual priming between (slightly) temporally separated words is a genuine phenomena. The first report that lexical decision times could be influenced by a prime word presented *after* its target was that of Kiger and Glass (1983) who reported speeded RTs at SOAs of less than 130 msec. When a prime has been visually masked, "backward" priming effects on accuracy of reporting the prime (Dark, 1988) or lexical decision times (Briand, den Heyer, & Dannenbring, 1988) have been obtained with SOAs of up to 1 second. However, these three studies used word pair stimuli because they were not designed to investigate the issue of lexical ambiguity. As pointed out by Peterson and Simpson (1989), sentence materials would bear more directly on the results in the ambiguity literature. These investigators found that both lexical decision and naming latency times to unambiguous words were speeded by the presentation of a related word 200 msec later, but only if the words were presented in pairs and not when the prime words occurred at the ends of sentences. Because the ambiguity results have originated from sentence paradigms, Peterson and Simpson thus concluded that "backward" priming could not have accounted for the context-invariant priming observed in those studies.

We find some of the issues raised by Peterson and Simpson to be intriguing. For instance, they suggested that backward priming was observed in word pairs but not sentences because the preceding sentence fragment allowed faster processing of the prime thereby allowing less temporal overlap with the subsequent target. Additionally, these authors noted that their data "do not rule out the possibility that backward priming might occur for target words that have, in fact, received some forward priming" (1989, pp. 1028). They do not believe this idea to be inconsistent with the postulate of context-insensitive meaning retrieval since by "forward" priming they mean forward priming of the target word's meaning. However, the possibility that "backward" priming could be supported by "forward" priming could be at odds with the context-insensitive view if we consider other varieties of forward priming. This idea will become more clear as we describe some of our experiments in progress.

Like Peterson and Simpson, we think that the backward priming paradigm needs to be brought closer to the standard ambiguity paradigm if it is to have any relevance. In our attempt to do so without using ambiguous words, we have focused on the defining attribute of homographic words, namely that two distinct meanings are represented by the same orthographic form. When a homograph is used to terminate a sentence, one of its meanings will be congruent with the sentence context and one incongruent although the shared visual form will match the sentence context in both cases. We have tried to incorporate these aspects of lexical ambiguity in an experiment by starting with pairs of words which have quite distinct meanings but are orthographically similar, such as SOUP and SOAP, or BREAD and BROAD. Sentence contexts which suggested one member of the pair were generated, as in "He ate a bowl of hot chicken ..." Subjects were presented with such sentences terminated by the incongruent but visually similar member of the pair (e.g., SOAP) mixed in with normal congruent sentences (e.g., "Water and sunshine help plants grow"), and sentences with incongruent final words (e.g., "He ordered french fries with his cable"). As in the ambiguity paradigm, each sentence was followed by a word which was either related or unrelated to the meaning of the sentence terminal word, with an SOA of 200 msec. The condition including a "visually similar" sentence completion followed by a related word approximates the condition in an ambiguity experiment of a homograph followed by a related word in that the subsequent word is related not to the semantically congruent sentence completion, but instead to another word which shares some visual characteristics of the congruent completion.

We recorded naming latencies to the final words of the sentences in this experiment in order to assess the influence of the subsequent word. No backward priming was observed for the final words of either the congruent or the incongruent sentences. However, in the "visually similar" condition, we found that both error rate and reaction time were sensitive to the relationship between the sentence terminal word (the target) and the subsequent word. When the subsequent word was unrelated, subjects often committed the error of pronouncing the word suggested by the sentence fragment (e.g., SOUP) rather than the word actually presented (e.g., SOAP). This type of error was significantly reduced when the sentence was followed 200 msec later by a related word. In addition mean reaction time for correct pronunciations was reduced by the presence of a subsequent related word. Although the backward priming effect on RTs was only 10 msec, it was statistically significant. This compares favorably to the magnitude of priming effects reported for contextually inappropriate probes of ambiguous words in experiments using a naming latency measure and SOAs close to 200 msec. This backward priming effect was about half the magnitude of the forward effect observed when we asked a separate group of subjects to pronounce the subsequent words rather than the final words of the sentences. We are currently pursuing the impact of including subsequent words that are related to the congruent completions, increasing the similarity between the congruent and "visually similar" words, and other manipulations. However, even as it stands, our data indicate that it is possible to obtain a backward priming effect in the sentence/naming latency paradigm that has been a mainstay of the ambiguity literature.

The Five-stage Sequence Reconsidered

We began this chapter by outlining a common view of word recognition in which the interpretation of words in sentences is inevitably delayed by passage through a lexicon which is both ordered by the frequency of usage and unresponsive to the weight of sentence context. We set out to test several of the predictions of this view, primarily with the ERP experiments summarized above. The results of these have persuaded us that much of the interpretation of words in sentences is fairly immediate. In retrospect, the concept of the lexicon as a passive store of information which is acted upon by subsequent processes should have been suspect for a variety of logical considerations. Firstly, the distinction between memory and process is clearly a metaphor drawn from computer science which should not be applied seriously to biological systems, where information processing transactions are conducted by the same neural elements which are modified by experience. Secondly, while it was suggested that focusing on processes within the lexicon would make the research questions more tractable (see Fodor, 1983), this has had the result of making those "post-lexical" processes required for sentence comprehension a somewhat more vague and intractable issue. Gerrig (1986) has noted that any adequate theory of comprehension should "specify both the information that is derived from the lexicon and the processes that operate on that information. Since trade-offs of process and products lead to the same end result for comprehension, the validity of a theory that considers either aspect alone is indeterminate" (pp 188-189). Autonomous models of lexical processing have focused only on the first half of this equation.

The non-trivial nature of the work that remains has been argued by Clark (1983) in his accounts of the ubiquity of "contextual expressions," or phrases that will not yield their intended meaning via a simple concatenation of their lexical components. Such expressions include a number of common constructions such as 1) compound nouns ("finger cup," "tea garden") where the relationship between the two nouns is not specified by the words themselves, 2) indirect or shorthand noun phrases ("one water" taken to mean "one glass of water," "a Beethoven" taken to mean "symphony written by Beethoven," or "a talented composer"), 3) possessive nouns (in the appropriate context "John's dog" could refer to "the dog that attacked John yesterday" or "the dog John is standing is front of"), 4) pro-act verbs ("Alice did the lawn" might indicate that she planted it or mowed it or fertilized it, etc), 5) eponymous adjectives ("Churchillian" might mean "with a face like Churchill," "smoking a cigar like Churchill" or "with a speaking style like Churchill"). Clark's criteria for a contextual expression are that its possible senses are not denumerable, and that the intended sense depends on coordination between speaker and addressee: "who uttered the expression, where, what he was pointing at, who had just been mentioned in the conversation, what his addressee knew and didn't know" (1986, pp 301).

When contextual expressions are used innovatively, the intended sense of their lexical components will not be discovered in static elements of long term memory. Thus a first-pass activation process like that proposed in Stage 3 of the five-stage sequence would not yield the correct candidate for subsequent selection by context-sensitive processes. A context-invariant activation process would similarly fail to produce the correct sense for a novel metaphor which does not possess a pre-stored lexical representation.

Of course, as pointed out by both Gerrig (1986) and Clark (1983) both contextual expressions and metaphors do draw on some pre-stored information for their interpretation. In Gerrig's example "The chimney belched forth soiled wisps of cotton," there is something about the perceptual quality of "cotton" which lends itself to a metaphorical identification with "smoke" much better than, say, "wool." So while we should not expect a lexical retrieval system to be able to deliver a fully specified, contextually appropriate definition of every word it encounters, we would like it to yield some relevant bits of information which can then be amplified and elaborated by continued processing. This sort of flexibility and susceptibility to subsequent processing is what Stage 3 of the original sequence was supposed to provide. We believe that the failure of that original model is due primarily to the rigidity of the initial activation phase; as it has been characterized, failure of the activation phase to find an existing, apt definition would be catastrophic. A more efficient and less failure-prone system would allow prior context to highlight the relevant aspects of a word's meaning without requiring the retrieval of a fully-specified definition.

Acknowledgments

We are grateful to Cathy Harris and Rob Kluender for helpful discussions and comments on a previous version of this paper. During some of the time this research was conducted, Cyma Van Petten was supported by an NSF graduate fellowship, and Marta Kutas by a Research Scientist Development Award (MH00322). The research was supported by a grant from NIH (HD22614).

Notes

[1]In various studies using different materials, the pseudoword N400 has been either somewhat larger or somewhat smaller than that elicited by real words. It is not yet clear how the amplitude of the pseudoword N400 will compare to the largest possible N400 elicited by real words, namely that to a set of words which are unrepeated, unrelated to previous words, and low in fre-

quency of usage.

[2]All of the experiments cited above were conducted in the visual modality. The sentence incongruity results have been replicated in both speech and American Sign Language (Kutas, Neville, & Holcomb, 1987; Karniski, Vanderploeg, Diehl, Lease, 1988). Associative priming in word lists and repetition priming effects have also been obtained in the auditory modality as well as in an ideographic writing system (Holcomb, 1985; Katayama, Teraji, & Yagi, 1989; McCallum, Farmer, & Pocock, 1984; Feldstein, Smith, & Halgren, 1987).

[3]Another reason for explicitly designing tasks which prevent the subject from making task-related decisions while reading is that a large positive ERP component (the P300) generally appears in any task requiring a binary decision, as in go/no-go tasks or choice RT tasks. The P300 occurs in the same latency range as the N400 which makes it difficult to determine which of the two overlapping components is affected by an experimental manipulation (see Donchin, 1981; Johnson, 1988; Kutas & Hillyard, 1989; Pritchard, 1981).

[4]The "noise" in an ERP average is ongoing EEG activity which is unsynchronized to stimulus presentation. Residual EEG activity in an ERP declines with the square root of the number of trials in the average (see Regan, 1989). In practice, the noise level can be evaluated by examining the difference between two averages during the prestimulus baseline. We have found 20 to be a minimally adequate number of trials for most N400 experiments. However, there is an inverse relationship between the size of the experimental effect which one hopes to obtain (i.e., one hopes to rise above the noise level) and the number of trials required. The difference in N400 amplitude between adjacent word positions may be as small as 1 uV and demands a somewhat larger number of trials.

[5]Care was taken to maintain verb argument structure in the replacement process. In addition, only "ly" adverbs were replaced; quantifiers such as "some" and "many" were not (see Cowart, 1982).

[6]There were two cases where this was difficult to avoid: sentence terminal words, and two-word proper names such as "United States." However, the critical word pairs never occurred as either the first or last word of a sentence, nor were they proper names, so these factors will not influence the analysis of responses to the critical pairs.

[7]Note that our analysis could focus on only the second words of each pair, regarding the Anomalous Unassociated as the baseline condition to define both the lexical and sentential effects. This strategy yields similar conclusions concerning the timecourse of the two effects as that pursued here.

[8]It should be noted that because neither the cellular events underlying the generation of the N400 nor those underlying word recognition are known in any detail, we do not make the claim that the ERP provides a real-time record of those events. However, the goal of the present experiment was to compare the latencies of two ERP context effects. Given that the waveshape and scalp distributions of these were the same, we found it parsimonious to assume that they

bore the same temporal relationship to the underlying neurophysiology.

[9]Interactive processing of words that are close together in time or space has fallen out of several recent network models of word recognition which did not set out to explicitly incorporate this factor. Mozer found that a network trained to identify words presented at varying spatial locations suffered "crosstalk" or letter migration errors when multiple words were presented close together, as do human subjects (McClelland & Mozer, 1986; Mozer, 1987). Masson (1989) has modelled mutual priming between words which share some semantic features.

[10]Both the early and the late priming effects began prior to the average reaction times we obtained in a behavioral version of an experiment using the same stimuli.

References

Aborn, M., Rubenstein, H., & Sterling, T.D. (1959). Sources of contextual constraint upon words in sentences. *Journal of Experimental Psychology, 57,* 171-180.

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 and 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 and Language, 24,* 89-106.

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

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

Becker, C.A., & Killion, T.H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 389-401.

Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and Language, 31,* 308-327.

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

Besson, M., Kutas, M., and Van Petten, C. (in press). ERP signs of semantic congruity and word repetition in sentences. In C.H.M. Brunia, A.W.K.

Gaillard, A. Kok, G. Mulder, M.N. Verbaten (Eds.), *Proceedings of the ninth international conference on event-related potentials of the brain.* Tilburg, Germany: Tilburg University Press.

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.

Blutner, R., & Sommer, R. (1988). Sentence processing and lexical access: The influence of the focus-identifying task. *Journal of Memory and Language, 27,* 359-367.

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

Brugman, C., & Lakoff, G. (1988). Cognitive topology and lexical networks. In S.L. Small, G. Cottrell, and M.K. Tanenhaus (Eds.), *Lexical ambiguity resolution* (pp 477-508). San Mateo, CA: Morgan Kaufmann.

Carroll, P., & Slowiaczek, M.L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition, 14,* 509-522.

Clark, H.H. (1983). Making sense of nonce sense. In G.B. Flores d'Arcais and R.J. Jarvella (Eds.), *The process of language understanding* (pp 297-332). New York: John Wiley and Sons.

Cowart, W. (1982). Autonomy and interaction in the language processing system: A reply to Marslen-Wilson and Tyler. *Cognition, 12,* 109-117.

Dark, V.J. (1988). Semantic priming, prime reportability, and retroactive priming are interdependent. *Memory & Cognition, 16,* 299-308.

den Heyer, K., Briand, K., & Dannenbring, G.L. (1988). Retroactive semantic priming in a lexical decision task. *Quarterly Journal of Experimental Psychology, 40A,* 341-259.

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. (1977). Associative facilitation without expectancy in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 18-26.

Fischler, I.S., & Bloom, P.A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. *Journal of Verbal Learning and Verbal Behavior, 5,* 1-20.

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

Forster, K.I. (1981a). Lexical access and lexical decision: Mechanisms of frequency sensitivity. *Journal of Verbal Learning and Verbal Behavior, 22,* 24-44.

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. (1982). A discourse on semantic priming. *Cognitive Psychology, 14*, 590-607.

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

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.

Gerrig, R. J. (1986). Process and products of lexical access. *Language and Cognitive Process, 1*, 187-195.

Gough, P.B., Alford, J.A., Jr., & Holley-Wilcox, P. (1981). Words and contexts. In O.J.L. Tzeng and H. Singer (Eds.), *Perception of print: reading research in experimental psychology.* Hillsdale, NJ: Erlbaum.

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.

Halgren, E. (in press). Insights from evoked potentials into the neuropsychological mechanisms of reading. In A. Scheibel and A. Weschler (Eds.), *Neurobiological substrates of higher cognitive function.* New York: Guilford.

Harris, C.L. (in press). Connectionism and cognitive linguistics. *Connection Science.*

Heit, G., Smith, M.E., & Halgren, E. (1988). Neural encoding of individual words and faces by hippocampus and amygdala. *Nature, 333*, 773-775.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

Hillyard, S.A. & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34*, 33-61.

Hillyard, S.A., Munte, T.F., & Neville, H.J. (1985). Visual-spatial attention, orienting, and brain physiology. In M.I. Posner and O.S. Marin (Eds.), *Attention and Performance XI: Mechanism of Attention* (pp. 63-84). Hillsdale, NJ: Erlbaum.

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 and Language, 35*, 66-85.

Huttenlocher, J., & Kubicek, L.F. (1983). The source of relatedness effects on naming latency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 486-495.

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, and M.G.H. Coles (Eds.), *Advances in Psychophysiology, Vol. 3* (pp. 69-138).

Greenwich, CT: JAI Press.
Karniski, W., Vanderploeg, R.D., Diehl, S., & Lease, L. (1988). Analysis of the N400 event-related potential component in a simple auditory sentence processing task. *Psychophysiology, 25*, 460-461. [Abstract].
Katayama, J., Teraji, M., & Yagi, A. (1989). N400 in orthographic decision tasks. *Psychophysiology, 26*, S38. [Abstract].
Kawamoto, A.H. (1988). Distributed representations of ambiguous words and their resolution in a connectionist network. In S.L. Small, G.W. Cottrell, and M.K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence.* San Mateo, CA: Morgan Kaufmann.
Kiger, J.I., & Glass, A.L. (1983). The facilitation of lexical decisions by a prime occurring after the target. *Memory & Cognition, 11*, 356-365.
Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.
Kintsch, W., & Mross, E.F. (1985). Context effects in word identification. *Journal of Memory and Language, 24*, 336-349.
Kutas, M. (1985). ERP comparisons of the effects of single word and sentence contexts on word processing. *Psychophysiology, 22*, 575-576. [Abstract].
Kutas, M. (1988). Review of event-related potential studies of memory.In M.S. Gazzaniga (Ed.), *Perspectives in memory research* (pp 181-217). Cambridge, MA: MIT Press.
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 senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*, 203-205.
Kutas, M., & Hillyard, S.A. (1980c). Reading between the lines: Event related brain potentials during natural sentence processing. *Brain and Language, 11*, 354-373.
Kutas, M., & Hillyard, S.A. (1982). The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia, 20*, 579-590.
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. (1984a). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163.
Kutas, M. & Hillyard S.A. (1984b). Event related brain potentials (ERPs) elicited by novel stimuli during sentence processing. In R. Karrer, J. Cohen, & P. Tueting (Eds.), *Brain and information: Event-related potentials, Annals of the New York Academy of Sciences*, (Vol. 425, pp. 236-241). New York: New York Academy of Sciences.
Kutas, M., & Hillyard, S.A. (1989). An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience, 1*, 38-49.

Kutas, M., Hillyard, S.A., & Gazzaniga, M.S. (1988). Processing of semantic anomaly by right and left hemispheres of commissurotomy patients: Evidence from event-related brain potentials, *Brain, 111*, 553-576.

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 and Clinical Neurophysiology Supplement 39*, 325-330.

Kutas, M., & Van Petten, C. (1988). Event-related brain potential studies of language. In P.K. Ackles, J.R. Jennings, and 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 and Clinical Neurophysiology, 69*, 218-233.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Landauer, T.K. & Streeter, L.A. (1973). Structural differences between common and rare words: Failure of equivalence assumptions for theories of word recognition. *Journal of Verbal Learning and Verbal Behavior, 12*, 119-131.

Langacker, R.W. (1987). *Foundations of cognitive grammar, Vol.1: Theoretical prerequisites*. Stanford, CA: Stanford University Press.

Lindner, S. (1981). *A Lexico-semantic analysis of verb-particle constructions with UP and OUT.* Doctoral dissertation, University of California at San Diego.

Lupker, S.J. (1984). Semantic priming without association: A second look. *Journal of Verbal Learning and Verbal Behavior, 23*, 709-733.

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

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

Massaro, D.W., Jones, R.D., Lipscomb, D., & Scholz, R. (1978). Role of prior knowledge on lexical decisions with good and poor stimulus information. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 498-512.

Masson, M.E.J. (1989). Lexical ambiguity resolution in a constraint satisfaction network. *Program of the eleventh annual conference of the Cognitive Science Society*, 757-764. Hillsdale, New Jersey: Erlbaum.

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

McCann, R.S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word frequency effects reflect lexical access? *Journal of Experimental Psychology: Human Perception and Performance, 14,* 693-706.

McCarthy, G., & Wood, C.C. (1984). Intracranially recorded event-related potentials during sentence processing. *Society for Neuroscience Abstracts, 10,* 847.

McClelland, J.L., St. John, M., & Taraban, R. (1989). Sentence comprehension: A parallel distributed processing approach. *Language and Cognitive Processes, 4,* 287-336.

McClelland, J.L., & Mozer, M.C. (1986). Perceptual interactions in two-word displays: Familiarity and similarity effects. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 18-35.

McKoon, G., & Ratcliff, R. (1989). Semantic associations and elaborative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 326-338.

Meyer, D.E., & Schvaneveldt, R.W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227-234.

Mitchell, D.C., & Green, D.W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology, 30,* 609-636.

Monsell, S., Doyle, M.C., & Haggard, P.N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43-71.

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.

Mozer, M.C. (1987). Early parallel processing in reading: A connectionist approach. In M. Coltheart (Ed.), *Attention and Performance XII, The Psychology of Reading* (pp 83-104). Hillsdale, New Jersey: Erlbaum.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226-254.

Nenov, V.I., Halgren, E., Smith, M.E., Badier, J., Ropchan, J.R., Blahd, W.H., & Mandelkern, M. (submitted). Metabolic localization of brain potentials to words.

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.

O'Seaghdha, P. (1989). The dependence of lexical relatedness effects on syntactic connectedness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 73-87.

Oden, G.L., & Spira, J.L. (1983). Influence of context on the activation and selection of ambiguous word senses. *Quarterly Journal of Experimental Psychology, 35A,* 51-64.

Onifer, W., & Swinney, D.A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9,* 225-236.

Perfetti, C.A., Goldman, S.R. & Hogaboam, T.W. (1979). Reading skill and the identification of words in discourse context. *Memory & Cognition, 7,* 273-282.

Peterson, R.R., & Simpson, G.B. (1989). Effects of backward priming on word recognition in single word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1020-1032.

Ratcliff, J.E. (1987). The plausibility effect: Lexical priming or sentential processing. *Memory & Cognition, 15,* 482-496.

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

Ritter, W., Ford, J.M., Gaillard, A.W.K., Harter, M.R., Kutas,M., Naatanen, R., Polich, J., Renault, B. & Rohrbaugh, J. (1984). Cognition and event-related potentials: I. The relation of negative potentials and cognitive processes. In R. Karrer, J. Cohen, & P.Tueting (Eds.), *Brain and information: Event related potentials,* (pp 24-38). New York: New York Academy of Sciences.

Rouse, R.O., & Verinis, J.S. (1962). The effect of associative connections on the recognition of flashed words. *Journal of Verbal Learning ad Verbal Behavior, 88,* 454-462.

Rubenstein, H. Garfield, L., & Millikan, J.A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487-494,

Rugg, M.D. (in press). Event-related potentials dissociate repetition effects of high and low frequency words. *Memory & Cognition.*

Rugg, M.D., Furda, J., & Lorist, M. (1988). The effects of task on the modulation of event-related potentials by word repetition. *Psychophysiology, 25,* 55-63.

Rugg, M.D., & Nagy, M.E. (1987). Lexical contribution to nonword-repetition effects: evidence from event-related potentials. *Memory & Cognition, 15,* 473-481.

St. John, M. (1988). *Hitting the right pitch: A meta-analysis of the processing of ambiguous words in context.* Unpublished manuscript, Cognitive Science Department, University of California, San Diego.

Sanocki, T., Goldman, K., Waltz, J., Cook, C., Epstein, W., & Oden, G. C. (1985). Interaction of stimulus and contextual information during reading: Identifying words within sentences. *Memory & Cognition, 13*, 145-157.

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 [Abstract].

Schuberth, R.E., & Eimas, P.D. (1977). Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and 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.

Schreuder, R., & Flores d'Arcais, G.B. (1989). Psycholinguistic issues in the lexical representation of meaning. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 409-436). Cambridge: MIT Press.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Seidenberg, M.S., Tanenhaus, M.K., Lieman, J.M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*, 489-537.

Sharkey, N.O. (1989). The lexical distance model and word priming. *Program of the Eleventh Annual Conference of the Cognitive Science Society*, 860-867. Hillsdale, NJ: Erlbaum.

Simpson, G.B. (1984). Lexical ambiguity and its role in models of word recognition. *Psychological Bulletin, 96*, 316-340.

Simpson, G.B, & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 28-39.

Simpson, G.B., Peterson, R.R., Casteel, M.A., & Burgess, C. (1989). Lexical and sentence context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 88-97.

Smith-Burke, M., & Gingrich, P.S. (1979). The differential role of function words and lexical items in narrative and expository text. In: M. Kamil and A.J. Moe (Eds.), *Reading research: Studies and applications*. Clemson, SC: National Reading Conference.

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

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

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

Stuss, D.T., Picton, T.W. & Cerri, A.M. (1986). Searching for the names of pictures: An event-related potential study. *Psychophysiology, 23*, 215-223.

Stuss, D.T., Sarazin, F.F., Leech, E.E. & Picton, T.W. (1983). Event-related potentials during naming and mental rotation. *Electroencephalography and Clinical Neurophysiology, 56*, 133- 146.

Swinney, D.A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645-659.

Swinney, D.A., & Osterhout, L. (1990). Inference generation during auditory language comprehension. In A. Graesser and G.H. Bower (Eds.), *The psychology of learning and motivation*. New York: Academic Press.

Tabossi, P. (1988a). Effects of context on the immediate interpretation of unambiguous nouns. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 153-162.

Tabossi, P. (1988b). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27*, 324-340.

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

Theios, J., & Muise, J.G. (1977). The word identification process in reading. In N.J. Castellan, Jr., D.B. Pisoni, and G.R. Potts (Eds.), *Cognitive theory*, Vol. 2. Hillsdale, NJ: Erlbaum.

Till, R.E., Mross, E.F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition, 16*, 283-298.

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., & Marslen-Wilson, W. (1986). The effects of context on the recognition of polymorphemic words. *Journal of Memory and Language, 25*, 741-752.

Tyler, L.K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception and Psychophysics, 34*, 409-420.

Van Petten, C. (1989). *Context effects in word recognition: Studies using event-related brain potentials*. Doctoral dissertation, University of California at San Diego.

Van Petten, C. and Kutas, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language, 26*, 188-208.

Van Petten, C., & Kutas, M. (submitted). Influences of semantic and syntactic context on open and closed class words.

Van Petten, C. & Kutas, M. (in press). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition.*

Warren, R.E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology, 94,* 90-100.

Warren, R.E. (1974). Association, directionality, and stimulus encoding. *Journal of Experimental Psychology, 102,* 151-158.

Warren, R.E. (1977). Time and the spread of activation in memory. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 458-466.

Wood, C.C., & Allison, T. (1981). Interpretation of evoked potentials: A neurophysiological perspective. *Canadian Journal of Psychology, 35,* 113-135.

Zwitserlood, C.M.E. (1989). *Words and sentences: the effects of sentential-semantic context on spoken-word processing.* Doctoral dissertation, Max-Planck Institute for Psycholinguistics, Nijmegen, the Netherlands.