Desmedt, J. E., & Godaux, E. Ballistic skilled movements: Load compensation and patterning of the motor commands. In J. E. Desmedt (Ed.), *Cerebral motor control in man: Long loop mechanisms*. Basel: Karger, 1978.

Dyhre-Poulsen, P. Perception of tactile stimuli before ballistic and during tracking movements. In G. Gordon (Ed.), *Active touch. The mechanism of recognition of objects by manipulation: a multidisciplinary approach*. Oxford: Pergamon Press, 1978.

Evarts, E. V., & Fromm, C. Sensory responses in motor cortex neurons during precise motor control. *Neuroscience Letters*, 1977, *5*, 267–272.

Fromm, C., & Evarts, E. V. Relation of motor cortex neurons to precisely controlled and ballistic movements. *Neuroscience Letters*, 1977, *5*, 259–265.

Ghez, C., & Lenzi, G. L. Modulation of sensory transmission in cat lemniscal system during voluntary movement. *Pflügers Archivs*, 1971, *323*, 273–278.

Ghez, C., & Pisa, M. Inhibition of afferent transmission in cuneate nucleus during voluntary movement in the cat. *Brain Research*, 1972, *40*, 145–151.

Hallett, M., & Marsden, C. D. Ballistic flexion movements of the human thumb. *Journal of Physiology*, 1979, *294*, 33–55.

Hugon, M. *Réflexes polysynaptiques cutanés et commandes volontaires*. Thèse d'Etat, Paris, 1967.

Hugon, M. Methodology of the Hoffmann reflex in man. In J. E. Desmedt (Ed.), *New developments in electromyography and clinical neurophysiology* (Vol. 3). Basel: Karger, 1973.

Le Bizec, J., Demairé, C., & Honoré, J. Facilitation spinale du réflexe de Hoffmann par une stimulation du nerf sural de faible intensité: étude chez le paraplégique. *Electroencephalography and Clinical Neurophysiology*, 1983, *55*, 162–164.

Paillard, J. *Réflexes et régulations d' origine proprioceptive chez l'Homme*. Paris: Arnette, 1955.

Pierrot-Deseilligny, E., Lacert, P., & Cathala, H. P. Amplitude et variabilité des réflexes monosynaptiques avant un mouvement volontaire. *Physiology & Behavior*, 1971, *7*, 495–508.

Robson, D. S. A simple method for constructing orthogonal polynomials when the independent variable is unequally spaced. *Biometrics*, 1959, *15*, 187–191.

Simon, J. N. Dispositif de contention des électrodes de stimulation pour l'étude du réflexe de Hoffmann chez l'homme. *Electroencephalography and Clinical Neurophysiology*, 1962, suppl. 22, 174–176.

Tsumoto, T., Nakamura, S., & Iwama, K. Pyramidal tract control over cutaneous and kinesthetic sensory transmission in the cat thalamus. *Experimental Brain Research*, 1975, *22*, 281–294.

Vallbo, A. B., Hagbarth, K. E., Torebjörk, H. E., & Wallin, B. G. Somatosensory, proprioceptive, and sympathetic activity in human peripheral nerves. *Physiological Review*, 1979, *59*, 919–957.

Wall, P. D. The laminar organization of dorsal horn and effects of descending impulses. *Journal of Physiology*, 1967, *188*, 403–423.

Winer, B. J. *Statistical principles in experimental design*. Tokyo: McGraw-Hill, International Student Edition, 1971.

11

Word Expectancy and Event-Related Brain Potentials During Sentence Processing

Marta Kutas,
Timothy E. Lindamood,
Steven A. Hillyard
Department of Neurosciences
University of California, San Diego
La Jolla, California

ABSTRACT

Event-related brain potentials (ERPs) were recorded from subjects as they read a series of sentences, for which Bloom and Fischler (1980) had established the cloze probability for alternative terminal words. The amplitude of a late negative component (N400) of the ERP to the terminal words was enlarged as an inverse function of their cloze probability. Even larger N400 waves were elicited by semantically incongruous words that were grafted onto the ends of some of the sentences. In a second experiment, the N400 to semantically anomalous terminal words was reduced in amplitude when those words were semantically related to the word best completing the sentence (i.e., having the highest cloze probability). These results indicate, first, that a semantic anomaly is not necessary for the production of a substantial N400 component, though it may be a sufficient condition, and second, that the N400 bears a reciprocal relationship to word expectancy in sentence contexts. It is suggested that the N400 may be an index of the degree of semantic priming or activation that a word receives from the prior context.

INTRODUCTION

It is now well established that a prior verbal context acts to facilitate word recognition during reading. One way to demonstrate these context effects is to measure the time required to classify a visually presented letter string as a word or non-word (that is, a lexical decision task). Using the lexical decision paradigm, it has been shown that a prior context decreases decision latencies for congruous or semantically related words while increasing latencies for seman-

tically unrelated words. These effects have been obtained for single words (Becker, 1980; Meyer & Schvaneveldt, 1971; Meyer, Schvaneveldt, & Ruddy, 1975; Neely, 1977; Schvaneveldt & Meyer, 1973) as well as for sentence contexts (Kleiman, 1980; Fischler & Bloom, 1979; Schuberth & Eimas, 1977; Schuberth, Spoehr, & Lane, 1981) and are similar to those reported for threshold recognition (Morton, 1964; Tulving & Gold, 1963) and pronunciation tasks (Stanovich, 1981; Stanovich & West, 1979, 1981; West & Stanovich, 1978). These investigations have been aimed at defining the locus as well as the direction (bottom-up versus top-down) of context effects. A number of theoretical models of word recognition have been advanced to describe the interaction of visual and contextual information at different levels of processing (Becker, 1980; Becker & Killion, 1977; Forster, 1976; Meyer et al., 1975; Morton, 1969).

A different technique for studying the role of context on linguistic expectancies during reading involves the recording of the electrical activity from the human scalp that is elicited in response to word presentations. This approach is based upon the sensitivity of certain components of these event-related brain potentials (ERPs) to fluctuations in both verbal and nonverbal expectancies (Courchesne, Hillyard, & Galambos, 1975; Duncan-Johnson & Donchin, 1977, 1980; Ford & Hillyard, 1981; Squires, Wickens, Squires, & Donchin, 1976; see Donchin, Ritter, & McCallum, 1978 for review). Inferences about the structure and time course of language comprehension processes can be derived by examining the morphology, timing, and scalp distribution of the ERPs elicited by linguistic stimuli.

The N400 Component and Contextual Deviations

For the past several years we have been investigating specific ERP components that are sensitive to the development and violation of linguistic and semantic expectancies during reading (Kutas & Hillyard, 1980a, 1980b, 1980c, 1981). In these studies, subjects were asked to read a series of different sentences presented one word at a time in the center of a screen and to be prepared to answer subsequent questions about their contents. In our initial experiments the majority of the sentences were simple and meaningful, while a small percentage (25%) of them were completed unpredictably by a semantically incongruous word that rendered the sentence nonsensical. Under such circumstances, the brain's responses to the semantically congruent and incongruent words were markedly different—semantic anomalies were characterized by a negative component between 300 and 600 msec post-stimulus (the N400 wave), which was not evident in response to the appropriate endings. We suggested that this N400 might be an electrophysiological sign of the interruption of sentence processing by a semantically inappropriate word and the attempt to reconcile the discrepant information so as to derive meaning from the sentence.

Further studies demonstrated that the N400 does not occur in response to all types of unexpected words in sentences. For example, the unpredictable occurrence of a semantically appropriate word in oversize, bold-face type was associated with an enhancement of a late positive complex (LPC) of waves that probably included the well-known P300 component (Kutas & Hillyard, 1980a). In fact, these two distinct ERP effects (the N400 and the LPC) were elicited concurrently by the same word when it was both semantically inappropriate and "surprisingly" large (Kutas & Hillyard, 1980c). The occasional presence of a complex, colorful picture at the end of some of these sentences was likewise associated with an LPC over the parietal regions, although the entire ERP configuration was different than that elicited by overly "large" words (Kutas & Hillyard, 1981).

Grammatical aberrations which did not involve semantic incongruity similarly elicited little or no N400 activity (Kutas & Hillyard, in press (a); Kutas & Hillyard, in press (b)). Violations of grammar such as incorrect noun-verb number correspondence or incorrect verb tense occurring at random in prose passages did not elicit an ERP pattern like that which followed semantic anomalies within the same material. The effect on the ERP of more flagrant syntactic alterations that do impact sentence comprehensibility, however, remains to be investigated. The fact that neither nonlinguistic deviations nor grammatical violations yield an N400 was taken as evidence that the N400 appears to be contingent upon some aspect of semantic analysis rather than being a general response to abberant words in text.

These experiments also revealed that the N400 follows a semantic anomaly whether it occurs at the end or at an intermediate position within the sentence. In addition, the N400 seems to be relatively insensitive to manipulations of the probability of occurrence of semantic anomalies, unlike the P300 response to other types of "surprising" stimuli (Duncan-Johnson & Donchin, 1977); the N400 had essentially the same amplitude and latency whether 50% or 25% of the sentences were anomalous. Finally, the N400 has a markedly posterior distribution over the scalp, with a slight but consistent right hemispheric predominance in amplitude and duration (Kutas & Hillyard, 1982).

The N400 and Semantic Processing

While we have restricted our study of the N400 to the reading of sentences one word at a time, others have observed similar late negative ERPs under different experimental conditions. For example, Fischler, Bloom, Childers, Roucos, and Perry (in press) recorded ERPs in an experiment where subjects made true/false validity judgments on simple propositions of the form: "A robin is a *bird*", or "A robin is not a *rock*." Under these circumstances they found that an enhanced negativity in the 300 to 500 msec region was elicited to the final word in the proposition whenever the subject and the object of the sentence were discrepant, regardless of the truth or falsity of the sentence as a whole. That is, an N400 was

elicited by affirmative sentences only if they were false and by negative sentences only if they were true. The authors suggested that the N400 may simply reflect a mismatch between the noun arguments of a sentence rather than its propositional validity; however, they also considered the possibility that the N400 may be a response to the intermediate anomaly that would be generated if the meanings of negative propositions were first dealt with in the affirmative. Hence, these findings suggest a possible reinterpretation of the N400 effect.

The results of several other recent studies similarly indicate that the "anomaly reprocessing" notion of the N400 component may be in need of revision. If the component described in each of these reports is indeed the same as the N400, then it would appear that neither a sentential context nor a semantic incongruity is necessary for its elicitation. While judgments about the equivalence of ERP components recorded under different experimental conditions are not without hazard, there is a striking similarity between the N400 and the late negativities observed in a number of experiments wherein subjects were required to read, name, or make a decision about a word or picture based on its semantic attributes. For example, in a study where subjects judged pairs of words to be the same or different by several criteria, the ERPs to semantic mismatches displayed an increased negativity around 400 msec (Sanquist, Rohrbaugh, Syndulko, & Lindsley, 1980). Boddy and Weinberg (1981) also reported ERP waveforms having larger N400-like components to words that did not belong to a previously named category (negative instances) than to those words that did (positive instances). Similar results were noted by Polich, Vanasse, and Donchin (1981) in experiments involving lexical decisions and semantic category judgments; they concluded that a late negativity (300 to 500 mesc) was elicited by stimuli that were outside the category expected by the subject. These findings, together with those of Stuss, Sarazin, Leech, and Picton (in press), who obtained an N400-like component in response to single words that were read orally or pictures that were named, suggest that a broad range of experimental manipulations may yield an N400 wave. The common element in all these studies is not the occurrence of semantic incongruity per se, but rather that a semantic analysis is required of words (or in some cases pictures) that are unpredictable in content.

While a semantic anomaly does not appear essential for the elicitation of the N400, its amplitude does appear to vary with the *degree* of semantic incongruity (Kutas & Hillyard 1980b). Thus, moderately incongruous sentences such as "Please put some milk in the *aquarium*" or "She punched him right in the *pocket*" were associated with N400 components, albeit significantly smaller than those to strongly incongruent words as in "Please put some milk in the *miracle*" or "She punched him right in the *aquarium*." Note that the moderately incongruous words on the whole formed meaningful propositions despite their relatively unlikely completions. This effect of the degree of semantic incongruity on N400 amplitude was observed in separate groups of subjects, and the assess-

ments of "moderate" and "strong" incongruity were based on the experiments' judgments with post hoc validations from subjects' ratings.

These findings suggested that the N400 may be a general index of word expectancy and that incongruous endings may represent one end of a continuum of highly unexpected words. We, therefore, conducted an experiment to evaluate the relationship between N400 amplitude and an operationally defined measure of semantic congruity. In so doing, we could test the suggestion that N400s were not specific to semantic anomalies but rather that their amplitudes systematically reflected gradations of word expectancy over a wide range.

EXPERIMENT I

The first experiment utilized a within-subjects design with a standardized set of sentences to investigate the relation between N400 amplitude, semantic incongruity, and word expectancy. For this purpose we selected 321 sentences for which Bloom and Fischler (1980) had established the degree of expectancy (cloze probability)[1] for alternative terminal words. The sentences were designed to produce a wide range of cloze probabilities for the terminal words, from very highly constrained (i.e., leading to a very predictable primary word choice) to more open-ended contexts that allowed a large number of alternative (and less probable) completions. The design called for comparisons between the ERPs to words varying in cloze probability with ERPs to wholly anomalous words that we grafted onto the ends of some of Bloom and Fischler's sentences.

METHODS

Of the 321 sentences presented to each subject, approximately 30% were highly constrained, 40% were of medium constraint, and 30% were of low contextual constraint. The degree of contextual constraint for sentence fragments up to the final word was defined in terms of the probability of the "dominant" or primary response. Approximately half of the high and low and one-third of the medium constraint sentences (N = 137) were completed by words that were the most probable responses or "best completions" (BCs) (e.g., "She called her husband

---

[1]The "cloze" procedure requires subjects to fill in the word, omitted from a text, that best completes the available context (Taylor, 1953). By collecting responses from a large number of subjects, one can calculate each word's cloze probability—that is, the probability that a particular word would be chosen to complete the given sentence context. This probability can be considered equivalent to the probability that a particular subject would predict or expect that word, given the context.

at his *office*"); the average cloze probabilities of the BCs were .92, .29 and .63 for the three types of sentences, respectively. The remaining half of the high and low constraint sentences (N = 85) were completed by words that were meaningful but relatively improbable (cloze probability < .05) (e.g., "Captain Sheir wanted to stay with the sinking *raft*"). The remaining two-thirds of the medium constraint sentences ended with words of medium cloze probability (N = 47, p = .23) or those that completed the sentences nonsensically (e.g., "George was fired but he couldn't tell his *fog*," N = 50).[2]

The sentences were presented one word at a time in the form of brightened dot matrices on a CRT controlled by an Apple II microcomputer. The sentences ranged in length from 5 to 13 words (X = 8). Words were flashed for a 132 msec duration at the rate of one every 700 msec. The final word in each sentence was accompanied by a punctuation mark to indicate sentence completion and was followed by a 2800 msec delay before the onset of the next sentence.

Subjects were tested while reclining in a comfortable chair in one session that lasted 2½ to 3 hours. They were informed that they would be presented with a series of simple English sentences of varying levels of comprehensibility. Their instructions were to read each of the sentences silently in order to answer questions about their contents at the end of the experiment. Subjects were exposed to the sentences in blocks of 40 each, separated by 2- to 3-minute rest periods.

EEG activity was recorded from eight scalp Beckman silver-silver chloride cup electrodes all referred to linked mastoids. Four were placed according to the International 10–20 convention at frontal (Fz), central (Cz), parietal (Pz), and occipital (Oz) midline locations. Symmetrical anterior temporal electrodes were placed halfway between F7 and T3 and F8 and T4 sites, respectively. Symmetrical posterior temporal electrodes were placed lateral (by 30% of the interaural distance) and 12.5% posterior to the vertex. In addition, eye movements were monitored via an electrode placed below the right eye and referred to the mastoids for vertical movements and blinks, and via a right to left canthal bipolar montage for horizontal movements. All trials on which the horizontal or vertical eye movements exceeded predetermined voltages were excluded from the ERP averages. In this way, it was possible to rule out contributions from electro-ocular artifacts to the ERP waveforms.

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

---

[2]The results of a previous experiment (unpublished) having a very similar design indicated that there were no significant differences in the amplitude of the N400 following low cloze probability words at the ends of low, medium, and high constraint sentences. We, therefore, chose to restrict semantic anomalies to the medium constraint sentences.

The EEG and EOG and stimulus trigger codes were recorded on FM tape and analog-to-digital conversion was performed off-line by a PDP 11/45 computer. A 1024 msec epoch of EEG data beginning 140 msec before the onset of each stimulus was analyzed at a sampling rate of 250 Hz.

Consistent with standard practice in our laboratory all measurements were made relative to a 100 msec prestimulus baseline. The ERP measures were similar to those we have used in our previous studies of the N400 wave; that is, base-to-peak amplitude and mean amplitude (or area) in a 100 or 150 msec window around the expected peak of the N400.

Prior to the EEG recording subjects were given a handedness questionnaire (Edinburgh Inventory) and administered the Level II spelling and pronunciation subtests of the Wide Range Achievement Test (WRAT). These latter tests were given to ensure that the subjects were not dyslexic and read at least at a high school level. Following the ERP recordings, subjects were given a memory/recognition questionnaire on the sentences. The first test listed all but the last words of 35 different sentences, 28 of which were selected from sentences just seen and 5 of which had not been shown. The subject's task was to check whether each sentence was "old" or "new" and to fill in the word they remembered as completing each of the "old" sentences. Next, the subjects were shown a list of all 321 of the completed sentences and were asked to rate on a 100-point scale how likely the final word was in that particular sentence context (1 = very unlikely; 100 = highly likely).

Data were collected from 13 young adults (six female, age range 18–32). Twelve of the subjects were right-handed according to self-report and the Edinburgh Inventory, and only one subject had a left-handed relative in his immediate family.

RESULTS

Memory/Recognition Questionnaire

The behavioral data from the memory/recognition questionnaire confirmed our assumption that the subjects had read and processed the sentences during the recording session. They correctly identified an average of 58% (range 25–90) of the "old" sentences and misclassified only 9% of the "new" sentences on the list as belonging to the set they had just seen. They correctly filled in an average of 30% (range 0–75%; two subjects failed to complete the questionnaire properly) of the final words for the sentences they had been shown; this broke down into approximately 13%, 21%, and 43% correct for the semantically anomalous, low cloze probability, and high probability words, respectively.

ERPS to Different Classes of Endings

Consistent with our previous findings, semantically anomalous words were associated with a broad negativity peaking between 300 and 600 msec poststimulus

(N400), which was appreciably more negative than the ERP to any of the words which completed the sentences appropriately. The N400 was largest over the posterior regions of the scalp. The nature of these ERP differences is shown in Fig. 11.1 for parietal and right posterior temporal electrode locations where the effects were most pronounced; the ERPs to words that best completed sentences with high contextual-constraint (average cloze probability, p = .92) and medium constraint (cloze p = .63) are compared against those elicited by semantically anomalous words occurring in medium-constraint sentences. The ERPs to the most probable words or best completions (BC) terminating the high and medium constraint sentences (solid lines) were not distinguishable from one another by a mean amplitude measure over 300–500 msec post-stimulus relative to 100 msec prestimulus baseline.
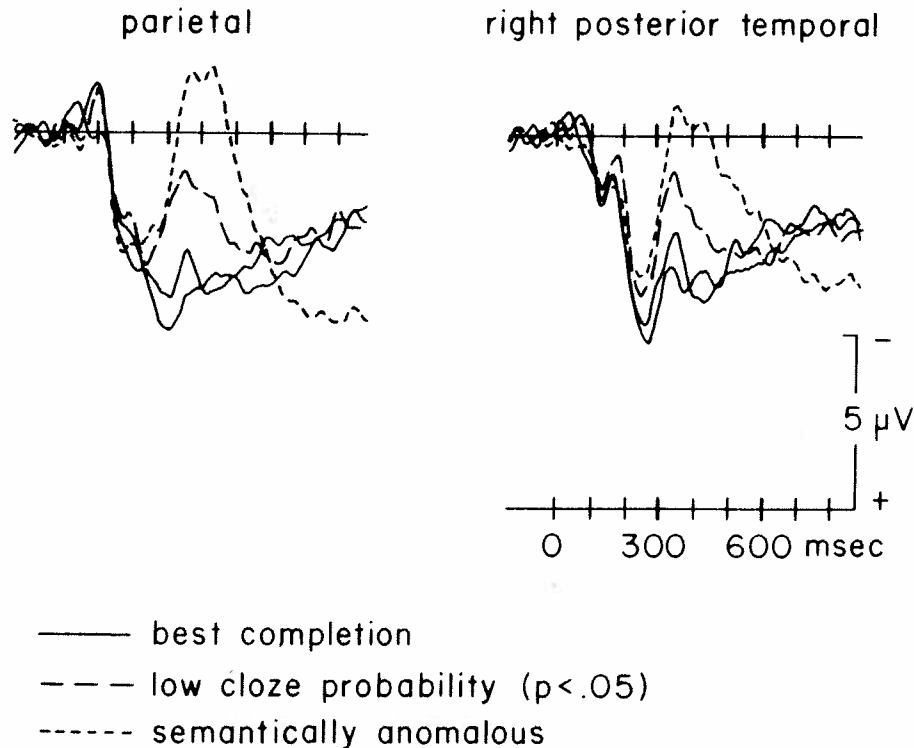

parietal   right posterior temporal

5 µV

0   300   600 msec

———— best completion

——— low cloze probability (p<.05)

------ semantically anomalous

FIG. 11.1.   The grand average ERPs (across all 13 subjects) elicited by final words which are either highly probable (BC), improbable (p < .05), or semantically anomalous in relation to the preceding sentence context. The two tracings for BC represent the ERPs to most probable words in highly constrained (p = .92) and medium constrained (p = .63) contexts. The low cloze probability words here completed highly constrained sentences while the semantic anomalies oc-

N400 components were elicited not only by the semantically anomalous words but also by the low probability, semantically congruent words at the ends of the highly constrained sentences (large dashed lines in Fig. 11.1). This parallels our finding of small N400s to the "moderate" semantic mismatches[3] in a previous study (Kutas & Hillyard, 1980a). While these meaningful but improbable terminal words were associated with a significant N400 relative to the response to the most probable or primary word choice within the same level of contextual constraint [main effect of ending probability $F(1,12) = 22.25$, p < .001], the N400s to these words were appreciably smaller than those following semantically anomalous words.

The broad, predominantly posterior distribution of the N400 effect can be seen in Fig. 11.2, where the ERPs elicited by three different types of endings of sentences with medium contextual constraints are compared. These data suggest that all terminal words other than the best completions (BC) were associated with some N400 activity. For example, if the sentence frame "The lawyer feared that his client was _____" was terminated by a low probability word such as "lying" (p = .14), there was enhanced negativity in the 300–500 region of the ERP in relation to its most probable completion, "guilty" (p = .73).

Using the mean amplitude over 300–500 msec as a measure, the N400 to semantically anomalous words was larger than that elicited by BC words or less probable completions for the medium-constrained sentences [main effect of ending type $F(2,24) = 31.24$, p < .001]. Moreover, the less probable completions were associated with somewhat more negativity in the same region than were the BCs [$F(1,12) = 6.73$, p < .02]. This sensitivity of the N400 amplitude to cloze probability was most evident over posterior electrode locations [ending type by electrode interaction $F(7,84) = 3.08$, p < .006]. Table 11.1 shows the amplitude of N400 to each of the different types of endings at the various scalp sites.

ERPs Averaged According to Cloze Probability

These results led us to re-average the data as a function of the cloze probability of the ending, collapsing across all values of contextual constraint (anomalous endings were excluded). The grand average ERPs in Fig. 11.3 demonstrate that the mean amplitude of the N400 in the 300–500 msec region was indeed a monotonic function of the probability of the terminal word, being most positive for high probability words, most negative for semantically anomalous completions, and intermediate in value for intermediate cloze probabilities.[4] This rela-

---

[3]Subjects ratings of the likelihood of these endings validated this distinction. The mean ratings for anomalies ranged from 1.0 to 18.3 (mean = 6.3) for different subjects and for low probability words in strong contexts from 42 to 81 (mean = 61).

[4]It should be noted that from these data one cannot distinguish between two possible aetiologies for the ERP to intermediate probability words: that is, whether all of them yielded an N400 of intermediate value or whether some trials were associated with a large N400 and others with a late
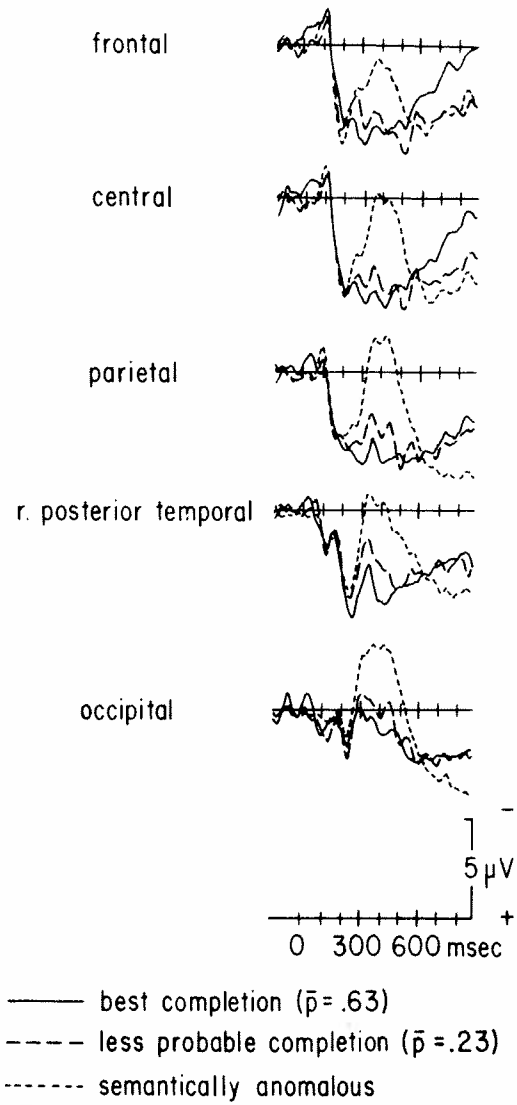
FIG. 11.2. The grand average ERPs (N = 13) elicited by the most probable (BC), less probable, and by semantically anomalous words in sentences having medium contextual constraint.

frontal

central

parietal

r. posterior temporal

occipital

5 μV

0   300 600 msec

——— best completion (p̄ = .63)
— — — less probable completion (p̄ =.23)
······· semantically anomalous

TABLE 11.1
Mean Amplitude 300–500 Measure (± std. error)

Sentence/Word Type	Frontal	Central	Parietal	Occipital	L. Anterior Temporal	R. Anterior Temporal	L. Posterior Temporal	R. Posterior Temporal
Hi constraint/ BC (p = .92)	4.27 (.69)	5.86 (.91)	4.43 (1.07)	1.01 (.92)	1.19 (.45)	1.43 (.36)	3.54 (.70)	4.00 (.92)
Hi constraint/ Low prob (p = .034)	4.23 (1.07)	4.34 (1.02)	1.96 (.84)	-1.43 (.59)	1.51 (.54)	1.34 (.51)	1.99 (.73)	2.39 (.80)
Med constraint/ BC (p = .63)	4.53 (.76)	5.24 (1.24)	4.37 (1.09)	0.90 (.86)	1.02 (.51)	1.24 (.41)	3.55 (.71)	4.07 (.90)
Med constraint/ Med prob (p = .23)	3.79 (.74)	4.07 (1.01)	2.83 (.92)	-0.09 (.65)	1.28 (.45)	1.08 (.40)	2.39 (.59)	2.64 (.80)
Med constraint/ Anomalous	1.38 (.74)	0.55 (.77)	-0.99 (.88)	-2.841 (.70)	0.02 (.47)	0.13 (.41)	-0.41 (.64)	0.01 (.57)
Low constraint/ BC (p = .29)	3.47 (.89)	3.38 (1.19)	2.34 (.96)	-0.73 (.82)	1.21 (.60)	1.01 (.54)	2.41 (.61)	2.24 (.87)
Low constraint/ Low prob (p = .025)	2.18 (.91)	1.05 (.97)	0.83 (.89)	-1.12 (.65)	1.19 (.56)	0.59 (.56)	0.92 (.64)	1.29 (.76)

occipital

5 μV

0   300   600 msec

——— cloze probability (p) >.7

——— p = .3 -.7

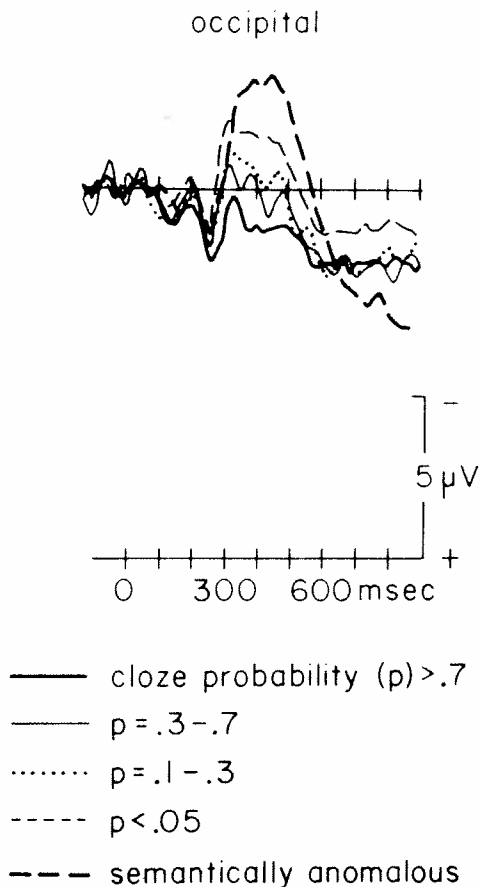········ p = .1 -.3

----- p < .05

— — — semantically anomalous

FIG. 11.3.   The grand average ERPs (N = 13) elicited by terminal words which were either semantically anomalous or within one of four ranges for cloze probabilities. The occipital site was chosen as a representative posterior recording location.

tionship was most pronounced at parietal, occipital, and lateral posterior temporal sites and was not evident at the left anterior temporal site.[5] However, since the semantically anomalous words were not associated with a cloze probability, it is difficult to determine whether the N400 elicited by them belong on the same continuum as the negativity elicited by the other improbable words or represents the addition of a different process.

N400 and Semantic Relatedness

The foregoing results indicated that the amplitude of the N400 might well be a good measure of word expectancy (operationally defined in terms of cloze proba-

[5]The correlations between cloze probability and the area in the 300- to 500-msec region of the grand average ERPs were Fz = .59, Cz = .77, Pz = .93, Oz = .94, l.AT = .56, r.AT = .62, l.PT = .89, and r.PT = .90.

bility). By extrapolation from the expectancy and priming explanations of many lexical decision task and word recognition results (Foss, 1982; Kleiman, 1980; Neely, 1977; Stanovich & West, 1981), we might expect that the sentence contexts establish expectancies not only for the best completions but also for semantically related words. Such expectancies might well be reflected in differential N400 amplitudes to improbable words which are or are not semantically related to the expected final word.

To test this prediction, we reanalyzed the ERPs to the semantically congruous, low probability words which completed highly constrained sentences according to whether or not they were semantically related (R) or unrelated (U) to the best completions for those sentences. The degree of semantic relatedness was determined by giving word pairs to 25 different subjects, each pair consisting of a low probability word actually presented together with the best completion of that particular sentence; subjects were asked to rate these pairs on a 5-point scale where "5" signified "very related" and "1" signified "totally unrelated." The mean ratings were 4.03 (s.d. = .63) for the "BC-R" word pairs (about half the sample) and 1.88 (s.d. = .75) for the "BC-U" word pairs. An example of a related low probability ending sentence would be: "The game was called when it started to *lightning*" (versus the best completion "*rain*"); by contrast, an example of an unrelated low probability sentence would be "When the two met, one of them held out his *badge*" (versus the best completion "*hand*"). The associated ERPs were then sorted and re-averaged according to these criteria.

The ERPs elicited by the best completions, and by the related and unrelated low probability completions are shown superimposed in Fig. 11.4. As previously noted, the low probability terminal words elicited a parieto-occipital negativity in relation to the BC words, peaking between 400 and 500 msec post-stimulus [main effect of ending type, F(2,24) = 14.07, p < .001]. There was also a definite trend for the N400 following low probability words that were semantically related to the best completion for that sentence to be smaller than those elicited by the unrelated words [base to peak N400 amplitude F(1,12) = 14.19, p < .002]. The effect of semantic relatedness, although evident at all the electrode sites, was most apparent over the posterior temporal areas. Subject's ratings indicated that there was a slight but nonsignificant tendency for final words which were related to the BC to be considered more likely completions than words which were unrelated.

EXPERIMENT II

The results of the first experiment indicated that the amplitude of the N400 to a low probability completion was reduced if the word was semantically related to the best completion of the sentence. This finding fits with the notion that a
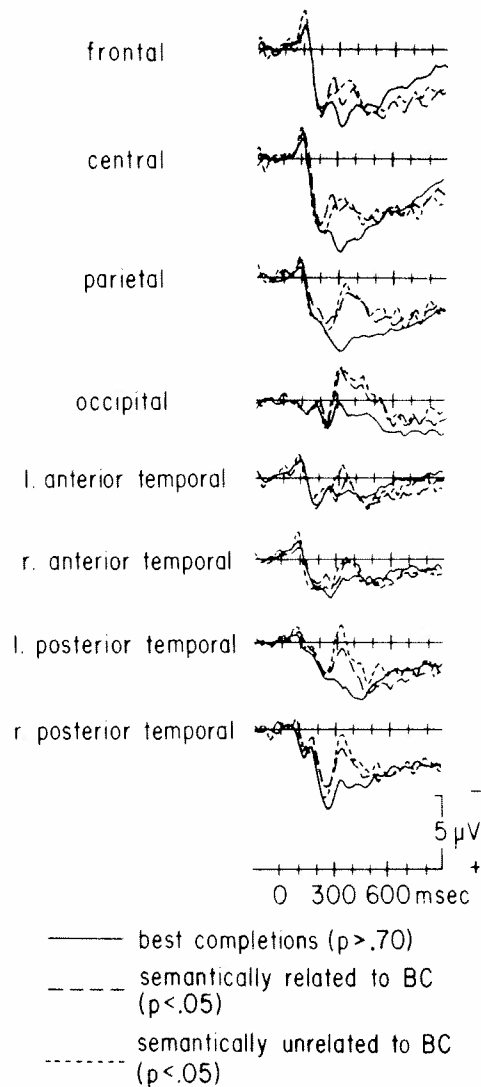
frontal

central

parietal

occipital

l. anterior temporal

r. anterior temporal

l. posterior temporal

r. posterior temporal

⎤ −
5 μV
⎦ +

0 300 600 msec

——— best completions (p>.70)

— — — semantically related to BC
(p<.05)

········ semantically unrelated to BC
(p<.05)

FIG. 11.4. The grand average ERPs (N = 13) elicited by highly probable (best completions) and relatively improbable words completing highly constrained sentences. The ERPs to words with low cloze probability were further subdivided according to whether they were or were not semantically related to the best completion of the sentences in which they occurred.

medium to strong sentence context sets up an expectancy for a particular word (the best completion), and activates or primes not only the best completions but also its semantic associates. It is possible, however, that semantic priming of this nature occurs only for acceptable, meaningful completions of a sentence, even if they are relatively improbable. The second experiment was designed as a stronger test of this "semantic relatedness" effect, using terminal words which completed sentences in a semantically anomalous fashion.

METHODS

This experiment was limited to highly constrained sentences (N = 160), which were terminated approximately 50% of the time by the most probable word as determined in a cloze procedure (BC) and 50% of the time by a semantically anomalous word. Of the semantic anomalies, approximately half were semantically related to the best completions (RA) and half were unrelated (UA). For example, "The game was called when it started to *umbrella*" (rather than the BC "*rain*") was an instance of a RA while "George had been fired but he couldn't tell his *fog*" (rather than the BC "*wife*") was an instance of an UA. The sentences varied in length from 5 to 14 words (X = 8.5).

The subjects for this experiment were 10 right-handed young adults (9 male, age range 18–30). Two of these subjects had sinistral relatives in their immediate family. The experimental procedures were similar to those described for the previous experiment with the following exceptions: (1) Stimulus exposure duration was now 50 instead of 100 msec and the delay interval between sentences was 3000 instead of 2800 msec;[6] (2) The memory/recognition test given to the subjects following the recording session consisted of 40 sentences, 10 of which were "new" and 30 of which were "old," selected from the sentences presented during the experimental session. Of the "old" sentences, 10 were BCs, 10 were RAs and 10 were UAs. (3) The subjects were next shown a list of all 160 completed sentences and were asked to rate how well the final word fit the preceding context on a 7-point scale (1—word doesn't fit at all, 7—word fits very well). If they did not consider the presented word to be the best completion, they were required to provide the word they thought most suitable; (4) subjects were asked to rate the relatedness of a list of word pairs on a 5-point scale (1—not at all related, 5—very related). One member of each word pair was the final word of one of the semantically anomalous sentences that had been shown, and the other member was the best completion for that same sentence. The mean relatedness ratings across 14 subjects (including the 10 for whom ERPs were recorded) were 4.25 (sd = .57) for the BC-RA word pairs and 1.18 (sd = .24) for the BC-UA word pair; (5) a 1024-msec epoch of EEG data beginning 100 msec before the onset of each stimulus was digitized on-line at a sampling rate of 250 Hz.

[6]Based on our previous studies (Kutas & Hillyard, 1980a, 1980b, 1982) and pilot work, we have found that the N400 is unaffected by variations in stimulus exposure duration (at least from 50 to 200 msec) or in inter-sentence interval (from 2000 to 4000 msec). The shortening of these intervals in Experiment 2 relative to those in Experiment 1 was dictated by our plan for comparing the effects of semantic relatedness at a number of different interword presentation rates while keeping stimulus duration and inter-sentence interval constant.

RESULTS

Figure 11.5 shows the superposition of the grand average ERPs elicited by the BC, RA, and UA terminal words. First, it is clear that whereas BC words were associated with a late positive complex with a parietal peak between 250 and 350 msec followed by a slow centro-parietal positivity, semantically anomalous words were characterized by a centro-parieto-occipital negativity (N400). Thus,
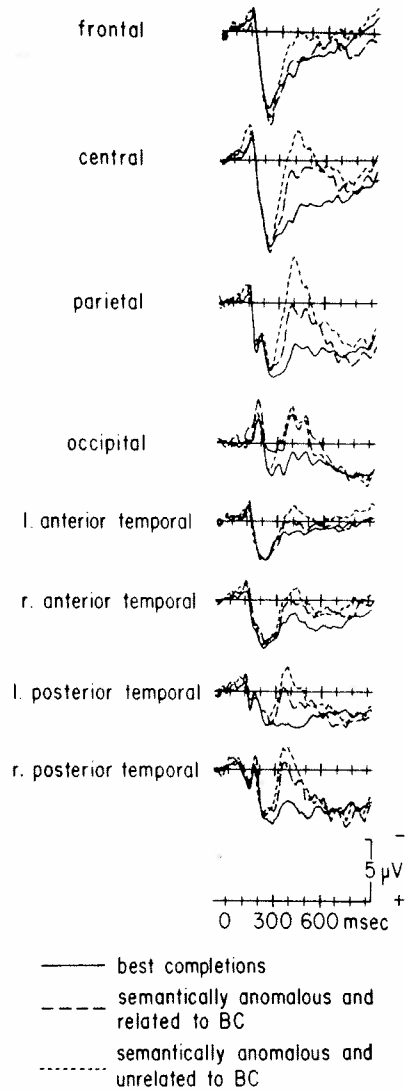

FIG. 11.5.   The grand average ERPs (N = 10) elicited by best completions (BC) and semantically anomalous words. The ERPs to the semantic anomalies were further subdivided according to whether or not they were related to the BCs of the sentences in which they occurred. All of the sentences in this experiment were highly contextually constrained.

the ERPs to BCs differed from those to the RA and UA words significantly in the 300–500 msec region [main effect of ending type $F(2,18) = 21.41$, $p < .001$]. This "semantic anomaly" effect was most pronounced at the midline parietal and lateral posterior temporal electrode sides [ending by electrode location interaction $F(14,126) = 5.04$, $p < .001$].

The amplitude of the N400 also reflected whether or not the anomalous word was semantically related to the BC. The N400 averaged 1.5 uV more negative for the semantically unrelated than related words over all electrodes. While there was some variability across subjects in the sensitivity of the N400 amplitude to this semantic relatedness factor, the N400s elicited by the two types of anomalous endings were significantly different by several measures [mean amplitude 350–450, $F(1,9) = 8.31$, $p < .018$; anomaly type by electrode interaction $F(7,63) = 2.25$, $p < .04$; base to peak N400 amplitude $F(1,9) = 7.76$, $p < .02$]. Consistent with other manipulations affecting N400 amplitude, the most reliable effects occurred at the right posterior temporal site [main effect of anomaly type, $F(1,9) = 12.64$, $p < .006$].

Behavioral data on the memory/recognition questionnaire also indicated reliable differences as a function of the type of sentence completion. The subjects correctly identified an average of 75% (range 56–100) of the "old" sentences and misclassified an average of 35% of the "new" sentences on the list as belonging to the set they had just seen. They correctly filled in an average of 25% (range 10–42) of the final words for the sentences they had been shown; segregated according to the type of completion, correct performance was 41% for BCs, 26% for RAs, and 5% for UAs.

Subjects were quite variable in the extent to which they rated the RA and UA words as "fitting" the preceding contexts. However, on the whole, consistent with the results in the previous experiment, anomalies that were related to the BC (range across subjects 1.0–4.6, mean = 2.8) were considered better "fits" than those words that were unrelated (range 1.0–2.2, mean = 1.5). This effect was significant in 6 of the 10 subjects (Mann-Whitney U, $p < .025$).

GENERAL DISCUSSION

The N400 component was found to be largest to the semantically incongruous sentence completions, but it also had a substantial amplitude following appropriate words having a low cloze probability. Moreover, there was a systematic decline in the N400 amplitude for terminal words as a function of increasing cloze probability. These findings cast considerable doubt on our previous interpretation of N400 as a sign of the "reprocessing" that is engendered by a semantic anomaly. It appears that semantic incongruity, while being a sufficient condition for N400 elicitation, is not a necessary condition.

If we assume that the cloze probability of a word is a measure of expectancy for its appearance, then it seems to be the case that the N400 amplitude is inversely proportional to word expectancy, at least for the natural sentence contexts used here. This leads to the hypothesis that the N400 may provide a measure of semantic priming of words being accessed from the internal lexicon; its amplitude in this case would be enhanced to the extent that a word is unprimed or nonactivated by the preceding context. This notion of automatic spreading activation is similar to that described by Posner and Synder (1975a, 1975b) in their cost-benefit theory of expectancy and attention. In their theory, also sub-scribed to by Stanovich and West (1981) for lexical decision tasks using sentence primes, the information in a sentence activates (through a fast-acting mechanism) not only the memory location for the expected word but also spreads automatical-ly to nearby semantically related memory locations.

Our second experiment provided evidence consistent with this interpretation of the N400. The finding of a smaller N400 amplitude for anomalous words that were related to the most probable (expected) sentence completions would be anticipated if the spreading activation process primed the semantic associates of the best completions (even when they were nonsensical). That spreading activa-tion does indeed operate in this fashion was suggested by the results of Kleiman (1980) in a lexical decision task where both the relatedness of the prime to the target and the degree to which the prime and the target formed an acceptable sentence were varied. Kleiman obtained large facilitation effects to best comple-tions, but he also found some facilitation for words that were related to the best completion and for words that formed logically acceptable completions to the sentences. His data further indicated that these two forms of facilitation were independent; that is, words that formed acceptable sentences with the prime were judged more rapidly even if they were unrelated to the best completion, and related words were responded to faster, even if they formed anomalous sentences with the prime. Kleiman invoked the concept of automatic spreading activation to account for the facilitation of both the best completions and the semantically related primes.

While the present experiment did not independently manipulate the proposi-tional and lexical/semantic incongruity of the terminal words, the results seem more in line with the proposal that N400 is more closely related to the latter type of discrepancy. This accords with the finding of Fischler et al. (in press) that the N400 in a sentence verification task is elicited by a semantic mismatch between the words used as subject and object in the sentence rather than by the overall truth/falsity of the proposition. They concluded that this negative ERP reflected a semantic mismatch among the components of a sentence, rather than a discrep-ancy between the sentence and previous knowledge.

Our tentative conclusion is that the amplitude of the N400 elicited under the present experimental conditions is a reliable measure of the degree to which a word has been primed by an activation process within the semantic networks. Accordingly, we would expect that its amplitude and latency will vary with many of the same factors that have been found to influence responses attributed to the facilitatory and inhibitory effects of sentence contexts (e.g., rate of presentation, reading skill, etc.). If this N400-semantic priming relationship proves to be valid, then this ERP can be utilized to examine the mechanism and timing by which sentence contexts facilitate or inhibit word processing during reading.

ACKNOWLEDGMENTS

REFERENCES

Becker, C. A. Semantic context effects in word recognition: An analysis of semantic strategies. Memory and Cognition, 1980, 8, 493–512.

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

Bloom, P. A., & Fischler, I. Completion norms for 329 sentence contexts. Memory and Cognition, 1980, 8, 631–642.

Boddy, J., & Weinberg, H. Brain potentials, perceptual mechanism and semantic categorization. Biological Psychology, 1981, 12, 43–61.

Courchesne, E., Hillyard, S. A., & Galambos, R. Stimulus novelty, task relevance, and the visual evoked potential in man. Electroencephalography and Clinical Neurophysiology 1975, 39, 131–143.

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

Duncan-Johnson, C. C., & Donchin, E. On quantifying surprise: the variation of event-related potentials with subjective probability. Psychophysiology, 1977, 14, 456–467.

Duncan-Johnson, C. C., & Donchin, E. The relation of P300 latency to reaction time as a function of expectancy. In H. H. Kornhuber & L. Deecke (Eds.), Motivation, motor and sensory pro-cesses of the brain: Electrical potentials, behavior and clinical use. Progress in Brain Research (Vol. 54). Amsterdam: Elsevier-North Holland, 1980.

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

Fischler, I., Bloom, P. A., Childers, D. A., Roucos, S. E., & Perry, N. W., Jr. Potentials related to sentence verification: Lexical versus sentential processes and the N400.Psychophysiology, in press.

Ford, J. M., & Hillyard, S. A. ERPs to interruptions of a steady rhythm. Psychophysiology, 1981, 18, 322–330.

Forster, K. I. Accessing the mental lexicon. In E. C. T. Walker & R. J. Wales (Eds.), *New approaches to language mechanism*. Amsterdam: North-Holland, 1976.

Foss, D. J. A discourse on semantic priming. *Cognitive Psychology*, 1982, *14*, 590–607.

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

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

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

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

Kutas, M., & Hillyard, S. A. Event-related brain potentials (ERPs) elicited by "novel" stimuli during sentence processing *Proceedings of EPIC VI*, Proceedings of the Sixth International Conference on Event-Related Slow Potentials of the Brain, Lake Forest, Illinois, 1981.

Kutas, M., & Hillyard, S. A. Event-related brain potentials and cognitive science. In M. Gazzaniga (Ed.), *Cognitive neuroscience*. New York: Plenum Press, in press. (a)

Kutas, M., & Hillyard, S. A. Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, in press. (b)

Kutas, M., & Hillyard, S. A. The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia*, 1982, *20*(5), 579–590.

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

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Loci of contextual effects on word recognition. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V*. New York: Academic Press, 1975.

Morton, J. The effects of context on the visual duration threshold for words. *British Journal of Psychology*, 1964, *55*, 165–180.

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

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

Polich, J., Vanasse, L., & Donchin, E. Category expectancy and the N200. *Psychophysiology*, 1981, *18*, 142.

Posner, M. I., & Snyder, C. R. R. Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975. (a)

Posner, M. I., & Snyder, C. R. R. Facilitation and inhibition in the processing of signals. *In P. M. A. Rabbit & S. Dornic (Eds.), Attention and performance V. New York: Academic Press*, 1975. (b)

Sanquist, T. F., Rohrbaugh, J. W., Syndulko, K., & Lindsley, D. B. Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology*, 1980, *17*, 568–576.

Schvaneveldt, R. W., & Meyer, D. E. Retrieval and comparison processes in semantic memory. In S. Kornblum (Ed.) *Attention and performance IV*. New York: Academic Press, 1973.

Schuberth, R. E., & Eimas, P. D. Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, *3*, 27–36.

Schuberth, R. E., Spoehr, K. T., & Lane, D. M. Effects of stimulus and contextual information on the lexical decision process. *Memory and Cognition*, 1981, *9*(1), 68–77.

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 1976, *193*, 1142–1146.

Stanovich, K. E. Attentional and automatic context effects in reading. *In A. M. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981.

Stanovich, K. E., & West, R. F. Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory and Cognition*, 1979, *7*, 77–85.

Stanovich, K. E., & West, R. F. The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, 1981, *7*, 658–672.

Stuss, D. T., Sarazin, F., Leech, E., & Picton, T. W. Evoked potentials during naming. In R. Karrer, P. Tueting, & J. Cohen (Eds.)., *Event-related potentials in man*, New York: New York Academy of Sciences, in press.

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

Tulving, E., & Gold, C. Stimulus information and contextual information as determination of tachistoiscopic recognition of words. *Journal of Experimental Psychology*, 1963, *66*, 319–327.

West, R. F., & Stanovich, K. E. Automatic contextual facilitation in readers of three ages. *Child Development*, 1978, *49*, 717–727.