Effects of stimulus and contextual information on the lexical decision process

RICHARD E. SCHUBERTH Rice University, Houston, Texas 77001

and

KATHRYN T. SPOEHR
Brown University, Providence, Rhode Island 02912

and

DAVID M. LANE
Rice University, Houston, Texas 77001

Two experiments were conducted to investigate context effects on the lexical decision process. In both experiments, observers classified letter strings as words or nonwords following the presentation of context in the form of an incomplete sentence. In Experiment 1, the predictability of congruous word stimuli and their frequency of occurrence in printed English were varied. These two factors had independent and additive effects on decision latencies. Stimulus quality, word frequency, and semantic congruity (i.e., congruous vs. incongruous) between the context and the stimulus were varied in Experiment 2. The effects of semantic congruity and word frequency on decision latencies combined additively, as did the effects of semantic congruity and stimulus quality. Two complementary mechanisms were proposed within the framework of a modified version of Becker's verification model to account for the differential effects of single-word and sentence context priming on the lexical decision process.

The effects of semantic context on the process of word recognition currently provide the focus for much research. A primary goal of this line of research has been to provide an explanation of how visual and contextual information are combined during reading. As our knowledge of context effects has increased, a variety of theoretical word recognition models have been advanced to describe the interaction of visual and contextual information (e.g., Becker & Killion, 1977; Forster, 1976; Meyer, Schvaneveldt, & Ruddy, 1975; Morton, 1969; Stanovich, in press).

Several investigators have employed a lexical decision task to study context effects. This task consists of classifying a visually presented letter string as a word or a nonword, and decision latency is taken as the dependent measure. Using this paradigm, it has been found that context decreases decision latencies for congruous or semantically related words and increases decision latencies for incongruous or semantically unrelated words. These effects have been obtained

This research was supported in part by United States Public Health Service Research Grant 07085-09 to Brown University and by National Science Foundation Grant BNS 76-82337 to the second author. The authors wish to thank William Howell, James Neely, and an anonymous reviewer for their critical evaluation of an earlier draft of this article. Requests for reprints should be sent to Richard E. Schuberth, Department of Psychology, Rice University, Houston, Texas 77001.

whether single words (Becker, 1980; Neely, 1976, 1977) or incomplete sentences (Fischler & Bloom, 1979, 1980; Schuberth & Eimas, 1977; Stanovich, 1980) were used as context.

An adequate theoretical interpretation of these findings requires consideration of at least three additional results obtained in lexical decision experiments. The effect of the semantic relatedness of the context and the stimulus on decision latency has been found (1) to interact with the effect of the visual quality (Meyer et al., 1975) and intensity (Becker & Killion, 1977) of the stimulus and (2) to combine additively with the counted frequency of occurrence of the stimulus in the English language (Schuberth & Eimas, 1977; but see also Becker, 1979). In addition, the effect of word frequency on decision latency has been shown to combine additively with the effects of the visual quality (Stanners, Jastrzembski, & Westbrook, 1975) and intensity (Becker & Killion, 1977) of the stimulus.

One approach to explaining context effects on the lexical decision process is based upon the logogen model of Morton (1969, 1970), which provides an explicit account of the interaction of stimulus and contextual information in word recognition. Logogen-type lexical decision models are characterized by a memory system in which each lexical entry is represented by a detector unit called a logogen that contains a specification of the word's defining characteristics along auditory, visual,

and semantic dimensions. Logogens are accessed directly rather than through an active search process, and they function as counting devices during recognition in accepting inputs from sensory feature analyzers and contextual mechanisms. When the number of inputs in a logogen's defining set exceeds a criterion, the word that the logogen represents is recognized.

Logogen-type lexical decision models enjoy only limited success in accounting for those findings that must be considered in any theoretical interpretation of context effects. For example, the model of Schuberth and Eimas (1977) can accommodate both the facilitatory and the inhibitory effects of context, the interaction between context and stimulus encoding variables, and the additivity of context and word frequency effects. However, models of this type also incorrectly predict that word frequency and stimulus encoding variables will have interactive rather than additive effects on decision latencies (for a detailed derivation of the predictions arising from such models, see Becker & Killion, 1977, and Schuberth & Eimas, 1977). Since a directaccess model of the logogen type cannot easily be modified to account for the myriad established findings, an alternative approach is necessary.

Becker (1976, 1979, 1980; Becker & Killion, 1977) has recently advanced a stage model of the lexical decision process that accommodates data obtained when word stimuli are presented alone or following the presentation of context. The components of this verification model include a visual feature analyzer and a system of lexical detectors, and the two basic processes are feature extraction and verification. Incoming stimulus information enters a short-term sensory store, and at that time the visual feature analyzer extracts primitive features (e.g., lines, angles, curves) to yield a partial description of the stimulus. As each stimulus feature is identified, it is fed into the detector system, in which detectors representing words that contain the feature respond by incrementing their feature counts. Whenever a word is presented, several detectors are assumed to exceed their criterion counts for primitive features, and these features form a subset of lexical memory items from which sampling occurs during the verification process.

For a stimulus word presented in isolation, the verification process consists of a series of comparisons between words from the sensory feature-defined set and the contents of visual memory. Words from the sensory set are sampled one at a time, in an order determined by their frequency of occurrence in the language. A representation of the sampled word is constructed, and its relational features (e.g., how separate features are organized, where one feature is in relation to other features, and how individual features are to be connected) are used to predict the contents of visual memory. If the contents of visual memory do not match the predicted relational features, another word from the sensory set is sampled and submitted to verification. The

sampling of sensory set items continues until a match with stimulus information is obtained. At that point, the word that generated the predicted relational features is recognized.

In the verification model, the presentation of contextual information has the effect of raising several lexical detectors above their criterion counts prior to the presentation of the stimulus. Items in this semantic set are sampled for verification as soon as incoming stimulus information reaches the sensory store. If the stimulus matches an item in the semantic set, decision latency is rapid and less than the latency obtained when no context precedes the stimulus. If an exhaustive sampling of semantic set items fails to yield a successful verification, items are then sampled from the sensory set, which is defined during the verification of semantic set items. Sampling from the sensory set proceeds until a match with stimulus information is obtained. The exhaustive sampling of semantic set items is assumed to delay verification of items in the sensory set, thereby increasing decision latencies for incongruous words in comparison with latencies obtained for these stimuli when no context is presented. In this manner, the verification model accounts for both the facilitatory and the inhibitory effects of context.

The verification model is consistent with two additional findings reported by Becker and Killion (1977). As noted earlier, they found that the effect of context interacts with the effect of stimulus intensity and that the effects of word frequency and stimulus intensity combine additively. To accommodate the interaction of context and stimulus intensity effects, they assumed that decreases in intensity slow the rate of feature extraction but leave unaffected the rate of the verification process. Since a context that is semantically related to the stimulus functions to bypass the feature analyzer component of the model, a Context by Stimulus Intensity interaction effect necessarily ensues. That is, decreases in stimulus intensity yield a substantial increase in decision latency only when the context and the stimulus are semantically unrelated.

The finding that word frequency and stimulus intensity have additive effects on decision latency is consistent with the mechanism for the word frequency effect proposed in the verification model. Unlike previous investigators who attribute frequency effects to different criterion counts for word detectors (Morton, 1969), Becker (1976, 1979, 1980; Becker & Killion, 1977) assumes that all detectors have comparable criterion counts. The word frequency effect arises during the verification process, in which items are sampled in an order determined by their frequency of occurrence in the language. That is, high-frequency words are sampled for verification before low-frequency words. Since, in the verification model, stimulus intensity and word frequency influence the feature extraction and verification processes, respectively, the effects of these factors on decision latency should be additive.

Becker (1979) has recently cited one aspect of the data reported by Schuberth and Eimas (1977) as inconsistent with the verification model. As mentioned previously, Schuberth and Eimas reported that the semantic congruity (i.e., congruous vs. incongruous) between a sentence context and a target word and the counted frequency of the target word have independent and additive effects on decision latency. Based on the assumption that semantic set items (i.e., congruous words) are selected for verification according to the strength of their relationship to the context rather than according to word frequency, Becker argues that a strong Context by Word Frequency interaction should ensue. That is, since semantically incongruous (i.e., sensory set) items are sampled according to word frequency, there should be a substantial effect of word frequency on decision latency for these words, but little or no effect of word frequency on decision latency for congruous words.

Becker (1979) has criticized the Schuberth and Eimas (1977) study on the grounds that the words of differing frequency may not have been presented in equally strong semantic contexts and that the repeated presentation of the lexical decision stimuli may have altered the frequency effect for at least some of the stimuli (cf. Scarborough, Cortese, & Scarborough, 1977). Becker (1979) reports two experiments designed to correct for these potential shortcomings. In his Experiment 1, subjects provided ratings of the strength of the association between contexts and word stimuli (e.g., freezing-cold) that were more refined than were those obtained by Schuberth and Eimas. In Experiment 2, subjects performed a lexical decision task in which the following factors were varied: rated strength of the relationship between the contexts and stimulus word associates, stimulus word frequency, and semantic congruity (i.e., related vs. unrelated) of the contextstimulus pairing. The result of major interest was the presence of a significant Word Frequency by Semantic Congruity interaction effect. The frequency effect was 54 msec in the related pairing condition and 96 msec in the unrelated pairing condition.

The discrepancy between the results of Becker (1979) and those of Schuberth and Eimas (1977) can be interpreted in two ways. On one hand, it is possible that the additivity of frequency and context effects obtained by Schuberth and Eimas is an artifact of the stimulus materials and the methodology they employed. On the other hand, a more interesting possibility, with important theoretical consequences, is that the incomplete sentence contexts employed by Schuberth and Eimas affected the lexical decision process in a manner different from that of the single-word contexts employed by Becker. To determine which of the two interpretations is more valid, two experiments were conducted in which observers were required to make timed lexical decisions for word and nonword stimuli that appeared after the presentation of incomplete sentence contexts. In both studies, those aspects of the Schuberth and Eimas study that Becker has cited as potential contaminants of their data were remedies. In particular, the strength of the relationship between the context and the stimulus was measured by a cloze technique, and none of the contexts or lexical decision stimuli was presented more than once.

Becker's (1979) assumption that semantic set items are sampled for verification according to the strength of their relationship to the context rather than according to word frequency provides the basis for the prediction of a Word Frequency by Semantic Congruity interaction effect in the verification model. It seemed necessary, therefore, to first test the plausibility of Becker's assumption when incomplete sentences are used as primes in a lexical decision task. Previous studies of the effects of sentence context priming on the visual recognition thresholds of word stimuli (e.g., Morton, 1964; Tulving & Gold, 1963) and on lexical decision latency (Fischler & Bloom, 1979) indicate that the predictability of the target word in the context is a reliable determinant of task performance. Unfortunately, in none of these investigations was the effect of word frequency on task performance examined simultaneously.

Fredericksen (Note 1) has more recently reported that word frequency and the degree of constraint by a prior sentence context have independent and additive effects on pronunciation times for a group of highability readers. It is unclear, however, how Fredericksen's measure of contextual constraint, the number of words elicited by a context, relates to the predictability of any single word in a context. In addition, pronunciation and lexical decision tasks are believed to involve different aspects of word recognition (cf. Coltheart, 1978). Therefore, in Experiment 1 of this report, the predictability of congruous word stimuli in context and their frequency of occurrence in printed English were varied.

In Experiment 2, stimulus quality, word frequency, and semantic congruity between the context and stimulus were varied factorially. The purpose of the experiment was to replicate the additivity of word frequency and semantic congruity effects reported by Schuberth and Eimas (1977) and to examine the combined effects of stimulus quality and semantic congruity on decision latency. It has already been established that when single words provide the context in a lexical decision task, the effect of the semantic congruity between the context and the stimulus interacts with the effect of stimulus quality (Meyer et al., 1975). To find that the effects of stimulus quality and semantic congruity do not interact when incomplete sentences are used as context would provide additional evidence that the two types of context influence the lexical decision process in different manners.

EXPERIMENT 1

Method

Subjects. The subjects were 51 Rice University undergraduates (33 females and 18 males), who participated in one 45-min experimental session for course credit. All subjects had normal or corrected-to-normal vision.

Stimuli and Apparatus. One hundred and sixty-eight sentence contexts were initially constructed. Each consisted of a subject and a transitive verb followed by a blank space in which the direct object was to be filled in (e.g., THE TAILOR MENDED THE ______). One hundred and nine Brown University introductory psychology students were then asked to fill in the blank at the end of each context with the first word that came to mind that provided a logical completion. Each subject supplied one completion for each of the contexts. The transition probability (TP; i.e., the predictability) of each completion word was measured by the percentage of subjects who produced the word for a given context.

Eighty-four contexts and one stimulus word for each of the contexts were then chosen from the larger set of stimulus materials. The words and corresponding contexts were divided into three groups of 28 items according to the TP of the word stimuli in the contexts: high TP (range = .60-.91, mean = .74), medium TP (range = .30-.59, mean = .42), and low TP (range = .1-.29, mean = .13). Within each of the three groups, approximately equal ranges of word frequency values (cf. Kučera & Francis, 1967) were represented: high TP (range = 0-393, mean = 77), medium TP (range = 7-382, mean = 77), and low TP (range = 0-383, mean = 82). The word stimuli ranged from three to nine letters in length.

The remaining 84 sentence contexts from the larger set were paired with nonword stimuli. Half of the contexts were paired with items from the sample of orthographically and phonologically legal nonwords employed by Rubenstein, Lewis, and Rubenstein (1971). The remaining contexts were paired with nonpronounceable nonwords that were formed by scrambling the pronounceable nonwords. All nonwords were from four to seven letters in length.

All stimulus materials were typed in uppercase Gothic font with an IBM Selectric II typewriter. A fixation dot was typed .8 cm below the center of each context. Xeroxographic transparencies (XF-10) of the stimulus materials were made, and these were mounted in 24 x 36 cm slide binders. The slides were displayed on a white screen by means of two Kodak Ektagraphic slide projectors in a Gerbrands projection tachistoscope (Model 1170). One projector contained the context slides; the other contained the target slides. The images from the two projectors were aligned so that when a context and a target were displayed simultaneously, the target appeared centered on the fixation dot underneath the context. The images of the targets were 10.2 cm high and 22.3-66.9 cm wide. Since the subjects sat 264.2 cm from the screen, the target stimuli subtended a vertical visual angle of 2.4 deg and a horizontal visual angle of 5.0-14.9 deg.

Procedure. Prior to testing, the sequence of stimulus events and the task requirements were explained to the subject. When the experimenter said "ready," the subject initiated each trial by pressing (with the preferred hand) the middle button of three horizontally aligned panel buttons, which resulted in the appearance of a sentence context on the screen. The subject was directed to read the context aloud quickly and then immediately press the middle button again and to fixate on the dot underneath the context. When the subject pressed the middle button again, the sentence context disappeared, and after an 80-msec delay, a target stimulus appeared on the screen, triggering a millisecond clock. The stimulus remained in view until the subject had executed a lexical decision by pressing one of the two outside panel buttons, which stopped the millisecond clock. Each subject made "word" responses with his or her preferred hand. Following each lexical decision, the experimenter informed the subject whether or not the decision was correct and recorded decision latency to the nearest 1 msec. There was a delay of approximately 6 sec between trials.

An experimental session consisted of 12 (6 word and 6 non-word) practice trials followed by 168 (84 word and 84 nonword) test trials. None of the contexts and targets used in the practice trials appeared during test trials. The stimuli were presented in

three blocks of 60 trials each, and subjects received a rest period of at least 1 min between blocks. Three presentation orders of stimulus materials were generated randomly, and equal numbers of subjects received each order.

Results

The overall error rate for the experiment was 1.89% (1.96% for words, 1.82% for nonwords). A multiple-regression analysis (cf. Cohen & Cohen, 1975) was done on the error data, in which word frequency, TP, number of syllables, word length, and the cross product of word frequency and TP were used to predict the number of errors made in classifying each of the word stimuli. A hierarchical procedure was used, with the variables entered into the equation in the order specified above. The only significant effect was a decrease in the number of errors with increases in word frequency [F(1,78) = 6.99, p < .01]. No other effect even approached significance (p > .20 in all cases).

Mean decision latencies were calculated across subjects for correct responses to each of the 84 word stimuli. The same regression analysis performed on the error data was also performed on the latency data. Decision latency increased with number of syllables [F(1,78) = 6.83, p < .025] and with word length [F(1,78) = 6.48, p < .025] and decreased with TP [F(1,78) = 38.57, p < .01] and with word frequency [F(1,78) = 12.28, p < .01]. The cross product of word frequency and TP did not contribute significantly to the regression equation [F(1,78) < 1], indicating no reliable Word Frequency (linear) by TP (linear) interaction effect.

Figure 1 contains the best-fitting regression lines describing the relationship between word frequency and mean decision latency for the high-, medium-, and low-TP groups. It is important to note that while TP is treated as a discrete variable for purposes of illustration, it was treated as a continuous variable in the regression analysis. A power analysis was performed to determine the probability of detecting (at the .05 level) a Word Frequency by TP interaction effect that accounted for

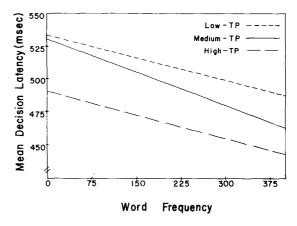


Figure 1. Best-fitting regression lines describing the relationship between word frequency and mean decision latency for the high- (.60-.91), medium- (.30-.59), and low- (.1-.29) transition probability (TP) stimulus groups.

5% of the variance. The squared multiple-correlation coefficient based on the four main effects was .45. If the population R^2 was .45, the probability of detecting an interaction effect that would raise the R^2 to .50 would be .80.

Discussion

The results of Experiment 1 indicate that the predictability of a congruous word stimulus in context and its frequency of occurrence in printed English affect decision latency reliably, and that these effects are additive. In terms of the verification model, the data thus cast doubt on the plausibility of Becker's (1979) assumption that the order in which semantic set items are sampled for verification is independent of word frequency. When sentence contexts are used as primes, both word frequency and predictability in context appear to influence the order of sampling from the semantic set (i.e., congruous words), whereas word frequency alone governs the order of sampling from the sensory set (i.e., incongruous words).

The additivity of word frequency and semantic congruity (i.e., congruous vs. incongruous) effects reported by Schuberth and Eimas (1977) is consistent with the notion that the word frequency effect does not differ for semantic set (i.e., congruous) and sensory set (i.e., incongruous) items. As mentioned previously, however, Becker (1979) has noted potential weaknesses in the Schuberth and Eimas study that may have biased their results. Therefore, Experiment 2 was conducted with improved procedures in an attempt to replicate the finding of Schuberth and Eimas. An additional purpose of Experiment 2 was to test the generality of the Stimulus Quality by Semantic Congruity interaction effect reported by Meyer et al. (1975). These investigators employed single-word contexts and degraded their stimuli with random dot patterns. There is evidence available that suggests that the interaction effect reported by Meyer et al. does not generalize to situations in which sentence material serves as context. For example, Mitchell and Green (Note 2) failed to find the effect in the data from a self-paced reading task (cf. Mitchell & Green, 1978) in which the visual quality of the test phrase and its semantic congruity with prior context were varied. Stanovich and West (1979) reported both the presence (Experiment 1 and Experiment 2, short response-stimulus interval) and the absence (Experiment 2, long response-stimulus interval) of a significant Stimulus Quality by Semantic Congruity interaction effect when subjects pronounced words that followed the presentation of sentence contexts. These authors discussed a number of differences in the nature of stimulus conditions in the two experiments that may have led to the discrepancy in results. Therefore, in Experiment 2 of the present research, the focus was also on the combined effect of stimulus quality and semantic congruity between context and stimulus on lexical decision latency.

EXPERIMENT 2

Method

Subjects. Sixty-four Brown University undergraduates, graduate students, and employees (45 females and 19 males) with normal or corrected-to-normal vision were paid \$3 for their participation. The subjects were tested individually in one 45-min session and were assigned randomly to either an intact or a degraded stimulus condition upon arrival in the laboratory.

Stimuli and Apparatus. Eighty-four sentence contexts from the pool described in Experiment 1 and 84 congruous word completions, ranging from three to nine letters in length, were selected. The contexts and corresponding word completions were divided into two groups of 42 sentences each according to the frequency score assigned to each stimulus word in the corpus of the Kučera and Francis (1967) count. Scores for highfrequency words ranged from 50 to 393, with an average score of 147. Low-frequency words had scores ranging from 0 to 29, with an average score of 13. The high- and low-frequency word groups were equated in the average TP score per item. TP scores for the high-frequency group ranged from .01 to .91, with an average score of .41; low-frequency items had TP scores ranging from .01 to .87, with an average score of .40. In addition, the high- and low-frequency word groups were equated in the average number of letters and syllables per item.

An additional 84 sentence contexts were selected and paired with the nonword stimuli described in Experiment 1. The word and nonword stimuli were equated in the average number of letters per item, and the pronounceable nonword and word stimuli were equated in the average number of syllables per item.

The stimulus materials were typed in the manner described in Experiment 1. For each word and nonword stimulus, an intact and a degraded version were made. Degradation was achieved by superimposing a black dot pattern (Letratone 10) directly over the typed stimulus. Contexts and stimuli were then photographed with Kodak high-contrast film, and the negatives were mounted in 24 x 36 cm slide binders. A Gerbrands projection tachistoscope (Model 1175) was used to display the contexts and stimuli on a screen positioned 287 cm in front of the subject. The word and nonword stimuli subtended .63 deg of vertical visual angle and 1.61-4.82 deg of horizontal visual angle.

Design. Three variables were manipulated in the experiment. Stimulus quality (two levels) served as a between-subjects variable, and word frequency (two levels) and semantic congruity between the context and stimulus (two levels) served as withinsubjects variables. The first two variables were described above. The third variable, semantic congruity, was manipulated by repairing some of the contexts and word stimuli such that the word stimuli provided semantically incongruous completions to the sentence contexts. Half of the stimuli at each level of word frequency were presented as congruous word stimuli (e.g., THE MODEL WORE THE GOWN) and half were presented as incongruous word stimuli (e.g., THE COOK BAKED THE PILLOW). One assignment of stimuli to congruity condition was used for one half of the subjects at each level of stimulus quality, and the opposite assignment was used for the other half of the subjects. That is, one half of the subjects saw the word GOWN as a congruous completion to the context THE MODEL WORE _, and the other half of the subjects saw the word THE GOWN as an incongruous completion to the context THE FARMER GREW THE _. The high-frequency and lowfrequency congruous word stimuli presented to each subject were equated in average TP score per item. In addition, at each level of word frequency, congruous and incongruous words were equated in average frequency score per item. To summarize, each subject classified 21 word stimuli in each of the four conditions defined by the factorial combination of word frequency (high vs. low) and semantic congruity (congruous vs.

incongruous). One half of the subjects classified intact stimuli, and the other half classified degraded stimuli.

Procedure. The sequence of events and task requirements were identical to those described in Experiment 1, with two exceptions. First, to initiate each trial, the experimenter said "ready" and pushed a control button that caused a sentence context to appear on the screen. Second, immediately after the subject had read the context aloud, he or she was directed to fixate on the dot underneath the context and to depress the foot pedal upon which his or her right foot was resting. When the subject depressed the foot pedal, the sentence context disappeared and, following a 100-msec delay, a target stimulus appeared on the screen. The subject executed a lexical decision by pressing one of two appropriately labeled panel buttons. Each subject executed word decisions with his or her preferred hand. Following each response, the experimenter informed the subject whether or not the response was correct and recorded decision latency to the nearest 1 msec. There was a delay of approximately 6 sec between trials.

An experimental session consisted of 20 (10 word and 10 nonword) practice trials followed by 168 (84 word and 84 nonword) test trials. None of the contexts and stimuli used in the practice trials appeared during test trials. The stimuli were presented in blocks of 55, 40, 40, and 33 trials, and subjects received a rest period of at least 1 min between trial blocks. Two presentation orders of stimulus materials were generated randomly, and equal numbers of subjects in intact and degraded stimulus conditions received each order.

Results

The errors made in response to word stimuli and the latency data for correct responses to word stimuli were analyzed according to the min F' procedure described by Clark (1973). The mean decision latency and error percentage for the subjects in each of the experimental conditions are presented in Table 1.

The analysis of the error data indicated that word frequency and semantic congruity had significant main effects. Fewer errors occurred in response to high-frequency words (2.31%) than to low-frequency words (4.58%) [min F'(1,123) = 6.08, p < .01], and fewer errors occurred in response to congruous words (2.01%) than to incongruous words (4.87%) [min F'(1,143) = 11.38, p < .01]. In addition, the effect of stimulus quality was marginally significant. Fewer errors were made in the intact stimulus condition (2.64%) than in the degraded stimulus condition (4.24%) [min F'(1,144) = 3.25, .05]. None of the interaction effects was significant (min <math>F' < 1 in all cases).

The analysis of the latency data for correct responses to word stimuli indicated that all three factors, word frequency, semantic congruity, and stimulus quality,

Table 1
Mean Decision Latency (in Milliseconds) and Percentage of Errors for Subjects in Each of the Word Conditions of Experiment 2

	Congruous				Incongruous			
Stimulus	Н	L	Н	L	Н	L	Н	L
Intact	484	.74	502	2.23	522	2.68	551	4.91
Degraded	584	1.49	602	3.57	632	4.32	673	7.59

Note -H = high frequency, L = low frequency.

had significant main effects. Overall, high-frequency words were classified 27 msec faster than were lowfrequency words [min F'(1,101) = 10.55, p < .01], congruous words were classified 50 msec faster than were incongruous words [min F'(1,142) = 41.35, p < .01], and intact stimuli were classified 108 msec faster than were degraded stimuli [min F'(1,70) = 24.14, p < .01]. The Word Frequency by Semantic Congruity interaction effect was significant in the subject (F₁) analysis [F(1,62) = 7.30, p < .01], but not in the word (F_2) analysis [F(1,82) = 1.60, p > .21], yielding a nonsignificant effect over both subject and item populations [min F'(1,116) = 1.31, p > .10]. The power of F₂ analysis to reject the null hypothesis at the .05 level was computed using a 42-msec Word Frequency by Semantic Congruity interaction effect (Becker, 1979) as the alternative hypothesis; the error term was based on the relevant MSe in the F₂ analysis. The power of the F₂ analysis was .89. The Stimulus Quality by Semantic Congruity interaction effect was significant in the F_2 analysis [F(1.82) = 5.54, p < .025], but not in the F_1 analysis [F(1,62) = 2.13, p = .15], yielding a nonsignificant effect over both subject and item populations [min F'(1,107) = 1.54, p > .10]. The power of the F₁ analysis to reject the null hypothesis at the .05 level was computed using a 33-msec Stimulus Quality by Semantic Congruity interaction effect (Meyer et al., 1975) as the alternative hypothesis; the error term was based on the relevant MSe in the F₁ analysis. The power of the F_1 analysis was .91. None of the remaining interaction effects was significant in either the F₁ or the F_2 analysis (min F' < 1 in all cases).

GENERAL DISCUSSION

To summarize, Experiments 1 and 2 provide new evidence concerning the effects of sentence context priming on the lexical decision process for word stimuli. In Experiment 1, the predictability of a congruous word stimulus in context and its frequency of occurrence in printed English had independent and additive effects on decision latency. Two major results were obtained in Experiment 2. First, stimulus word frequency and the semantic congruity between the context and the stimulus word had independent and additive effects on decision latency. This result confirms the finding of Schuberth and Eimas (1977) and, at the same time, conflicts with the significant Word Frequency by Semantic Congruity interaction effect reported by Becker (1979). Second, stimulus quality and semantic congruity also had independent and additive effects on decision latency. This finding is inconsistent with the significant Stimulus Quality by Semantic Congruity interaction effect reported by Meyer et al. (1975).

A comparison of the present results with the findings reported by Becker (1979) and by Meyer et al. (1975) suggests that priming with sentence contexts affects the lexical decision process in a way different from that of

priming with single-word contexts. Within the framework of the verification model, two distinct mechanisms may give rise to these differential effects. One possibility is the semantic set-size mechanism advanced by Becker (1980) to explain facilitation-dominant vs. interferencedominant effects of single-word contexts on lexical decision latency. Becker has hypothesized that contexts that allow subjects to predict the related target with a high degree of accuracy (e.g., hot-cold) result in the generation of a small semantic set and produce a dominance of facilitation for target words in related contexts. On the other hand, contexts that permit only a general expectation about the stimulus (e.g., dogcollie) result in the generation of relatively large semantic sets and produce a dominance of interference for target words in unrelated contexts. If one assumes that low-constraint (i.e., eliciting a large number of different responses) sentence contexts generate relatively large semantic sets, then the set-size mechanism provides a basis for interpreting the observed differences in the combined effect of word frequency and semantic congruity between the context and stimulus.

Consider our finding in Experiment 1 that the predictability of a congruous word stimulus in context and its frequency of occurrence in printed English have independent and additive effects on decision latency. Two additional assumptions are required in the verification model to accommodate this result. First, when semantic set items become available for verification, there is an initial assignment of set members to higher or to lower probability groups. Higher probability set members are sampled for verification before lower probability set members are. Second, within the higher and the lower probability groups, items are sampled according to word frequency. With low-constraint sentence contexts, it is likely that the higher and the lower probability semantic set groups contain several items, yielding substantial word frequency effects.

When single words are used as context, the order in which semantic set items are sampled should conform to the assumptions outlined above. However, if the semantic set consists of only a few items that are strongly associated with the context word, then the effect of word frequency should be negligible. Becker (1979) reported a word frequency effect of 53 msec across levels of strength of association to the context. To accommodate this finding, he suggested that as the strength of the relationship between the context and the stimulus decreases, the probability that the stimulus word will be included in the semantic set also decreases. Thus, according to this argument, the decision latencies for related words obtained by Becker (1979) may have reflected a mixture of decision latencies for words treated as related and decision latencies for words treated as unrelated. This explanation is consistent with the notion that when single words provide the context, the semantic set may consist of the few associates that are strongly related to the context.

The observed difference in the combined effect of stimulus word frequency and semantic congruity between the context and stimulus word follows logically from the proposed mechanism. In Experiment 2 of this report, high-frequency and low-frequency congruous word stimuli were equated in average TP per item. Word frequency should therefore have been the sole factor governing the order of sampling from both the semantic set (i.e., congruous words) and the sensory set (i.e., incongruous words). The obtained additivity of word frequency and semantic congruity suggests that a comparable number of verification cycles preceded the recognition of semantic set and sensory set items at the same level of word frequency. When a single word provides the context, the word frequency effect for semantically related words is smaller than the word frequency effect for semantically unrelated words (cf. Becker, 1979) because the sensory set is likely to contain more items. Thus, in the larger sensory set, a greater difference should exist between the average number of verification cycles needed to recognize high-frequency items and low-frequency items.

The hypothesis that differential priming effects can be attributed, in part, to the size of the semantic sets generated in response to sentence and single-word contexts yields two testable predictions. First, the word frequency effect for semantic set items should decrease with an increase in the level of contextual constraint associated with incomplete sentences. Fredericksen (Note 1) obtained results that failed to confirm this prediction. For good readers, the word frequency effect for congruous word stimuli did not differ across levels of contextual constraint. For poor readers, the word frequency effect was larger for high- than for lowconstraint sentence contexts. However, as we noted earlier, Fredericksen's subjects pronounced the stimuli, an activity that is thought to engage processing operations different from those engaged in making lexical decisions (cf. Coltheart, 1978). The second prediction is that the effects of word frequency and semantic congruity on decision latency should interact rather than add together when high-constraint sentence contexts are employed. Confirmation of these two predictions would lend strong support to the proposed mechanism.

The second mechanism through which single-word and sentence contexts may differentially affect the lexical decision process is rooted in the speed with which semantic set members become available for verification. As Fischler and Bloom (1979) have pointed out recently, the processing demands of a sentence context outweigh by far those of a single-word context. For example, there may be configural aspects of the sentence that cannot be obtained from the meaning of individual words (cf. Foss & Harwood, 1975). In addition, sentence contexts contain no single word that is strongly related to the stimulus. It seems reasonable to assume, then, that given a limited amount of time to process each type of context, semantic associates of a single-

word context should be activated more rapidly than words that provide congruous completions to a sentence context. Consistent with this notion are the results of Fischler and Goodman (1978) and Cosky and Gough (Note 3). Fischler and Goodman showed that priming effects can be obtained when a single-word context precedes the test stimulus by as little as 40 msec. In contrast, Cosky and Gough reported that a sentence context must be displayed for a considerably longer period of time to affect word recognition.

Our two postulated mechanisms, semantic set size and the speed with which semantic set items become available for verification, provide a basis for interpreting another difference in the effects obtained with the two types of context. In Experiment 2 of this report, the visual quality of the stimulus word and its congruity in context had additive effects on decision latency. In contrast, in the study by Meyer et al. (1975), a significant Stimulus Quality by Semantic Congruity interaction effect was obtained. To facilitate the discussion of why these different effects may occur, it is necessary to consider first whether or not the verification model, as described by Becker and Killion (1977), can accommodate the effects on decision latency that arise when context is presented and the stimulus word is degraded with a random dot pattern.

Becker and Killion (1977) reported that virtually all of the effect of decreases in stimulus intensity was limited to decision latencies for stimulus words that were semantically unrelated to the context. To accommodate their finding, they assumed that the reductions in stimulus intensity employed in their study slowed the rate of feature extraction but had no effect on the verification process. Decision latencies for stimulus words that were semantically related to the context therefore remained largely unaffected by decreases in stimulus intensity, because for these words, context functioned to bypass the feature analyzer component of the model.

In Experiment 2 of this report and in the study by Meyer et al. (1975), degrading the stimuli with a random dot pattern produced a substantial increase in decision latencies for congruous words and for incongruous words. In the verification model, the increase in decision latencies for congruous words would presumably arise during the verification of semantic set items because the feature analyzer component of the model is bypassed. However, since Becker and Killion (1977) conceptualize the verification process as serial and self-terminating in nature, word frequency and stimulus quality should most likely interact in their combined effect on decision latency (cf. Sternberg, 1969). That is, if one assumes that stimulus degradation increases the time needed to complete each cycle of the verification process, then the effect of stimulus degradation should be less pronounced for high-frequency words than for low-frequency words. This prediction follows logically from the assumption that fewer cycles of the verification process are required to recognize high-frequency words than to recognize low-frequency words. The results of our second experiment, in which the effects of word frequency and stimulus quality combined additively, fail to confirm this prediction.

It is apparent that Becker and Killion's (1977) version of the verification model must first be modified to accommodate the effects on decision latency associated with decreases in the visual quality of a stimulus via random dot degradation. We shall assume that stimulus degradation initiates a preprocessing operation that functions to reduce the visual noise superimposed on the stimulus percept (cf. Neisser, 1967; Sternberg, 1967). Upon completion of preprocessing, visual features are extracted at a normal rate from the stimulus representation and are sent to the system of lexical detectors. Thus, one effect of stimulus degradation is to delay the formation of the sensory set.

An additional, indirect effect of stimulus degradation occurs when context is presented. We shall make two additional assumptions here: (1) the formation of the entire semantic set precedes the verification of set members against the contents of the visual store, and (2) the verification of semantic set members begins prior to the completion of stimulus preprocessing. An increment in time will therefore be added to each verification that is undertaken while visual noise is superimposed on the stimulus percept. Once stimulus preprocessing has been completed, the rate of each subsequent verification cycle will be equivalent to the rate of verification when the stimulus representation is of normal visual quality. Thus, stimulus degradation also functions to retard the rate of verification for at least some members of the semantic set. Such an effect is indexed by the increase in decision latencies for congruous word stimuli obtained in our second experiment and in the experiment by Meyer et al. (1975).

Let us now consider how the modified verification model accommodates the finding that the combined effect of stimulus quality and semantic congruity on decision latency differs when sentence contexts and single-word contexts are employed. When sentence contexts are presented and test stimuli are degraded. decision latencies for congruous word stimuli will increase, because the verification of semantic set items begins before stimulus preprocessing has been completed. As stated previously, we have assumed that sentence contexts generate relatively large semantic sets that become available for verification at a slow pace. An exhaustive sampling of semantic set items is therefore assumed to extend beyond the time needed to preprocess the stimulus and to form the sensory set. Thus, when the stimulus is a degraded incongruous word, the verification of sensory set items is delayed due to the increase in time needed to verify semantic set items. The additivity of stimulus quality and semantic congruity effects obtained in our second experiment is compatible with this analysis.

When single-word contexts are employed and test stimuli are degraded, decision latencies for congruous word stimuli increase. This increase in decision latencies occurs because, as stated previously, the verification of semantic set items begins when visual noise still remains on the stimulus percept. We have assumed that singleword contexts generate relatively small semantic sets that rapidly become available to the verification of process. Therefore, it seems reasonable to assume that an exhaustive sampling of semantic set members will be completed prior to the formation of the sensory set, which is delayed substantially by stimulus preprocessing. Thus, when the stimulus is a degraded incongruous word, the verification of sensory set items is delayed by the stimulus preprocessing operation rather than by the increase in time needed to verify semantic set items. The significant Stimulus Quality by Semantic Congruity interaction effect reported by Meyer et al. (1975) is consistent with this analysis.

To summarize, we have proposed two mechanisms within the framework of a modified verification model by which single-word and sentence contexts may differentially affect the lexical decision process. The proposed mechanisms together provide a reasonable explanation of several sets of results, including those obtained in the present experiments. Moreover, because the verification framework is compatible with top-down approaches to reading comprehension (cf. Adams, 1979; Rumelhart, 1977), it appears to be a promising approach to the study of word recognition.

REFERENCE NOTES

- 1. Fredericksen, J. R. Word recognition in the presence of semantically constraining context. Paper presented at the annual meeting of the Psychonomic Society, San Antonio, November 1978.
- 2. Mitchell, D. C., & Green, D. W. Contextual effects in continuous reading. Paper presented at the British Psychological Society conference on reading, London, March 1980.
- 3. Cosky, M., & Gough, P. B. The effect of context on word recognition. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1973.

REFERENCES

- Adams, M. J. Models of word recognition. Cognitive Psychology, 1979, 11, 133-176.
- BECKER, C. A. Allocation of attention during visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 556-566.
- BECKER, C. A. Semantic context and word frequency effects in visual word recognition. *Journal of Experimental Psychology:* Human Perception and Performance, 1979, 5, 252-259.
- Becker, C. A. Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & 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.
- CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, 12, 335-359.

- COHEN, J., & COHEN, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N.J. Erlbaum, 1975.
- COLTHEART, M. Lexical access in simple reading tasks. In G. Underwood (Ed.), Strategies of information processing. New York: Academic Press, 1978.
- 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. Rapid processing of the meaning of sentences. *Memory & Cognition*, 1980, 8, 216-225.
- Fischler, I., & Goodman, G. O. Latency of associative activation in memory. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4, 455-470.
- FORSTER, K. I. Accessing the mental lexicon. In E. C. T. Walker & R. J. Wales (Eds.), New approaches to language mechanisms. Amsterdam: North Holland, 1976.
- Foss, D. J., & Harwood, D. A. Memory for sentences: Implications for human associative memory. *Journal of Verbal Learning* and Verbal Behavior, 1975, 14, 1-16.
- KUČERA, H., & FRANCIS, W. N. Computational analysis of present-day American English. Providence, R.I: Brown University Press, 1967.
- MEYER, D. E., Schvaneveldt, R. W., & Ruddy, M. G. Loci of contextual effects on visual word recognition. In P. M. A. Rabbitt & S. Dornic (Eds.), Attention and performance V. London: Academic Press, 1975.
- MITCHELL, D. C., & GREEN, D. W. The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, 1978, 30, 609-636.
- MORTON, J. The effects of context on the visual duration thresholds for words. British Journal of Psychology, 1964, 55, 165-180.
- MORTON, J. Interaction of information in word recognition. *Psychological Review*, 1969, 76, 165-178.
- Morton, J. A functional model for memory. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970.
- NEELY, J. H. Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition*, 1976, 4, 648-654.
- 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.
- Neisser, U. Cognitive psychology. New York: Appleton-Century-Crofts, 1967.
- Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 645-657.
- RUMELHART, D. E. Toward an interactive model of reading. In S. Dornic (Ed.), Attention and performance VI. Hillsdale, N.J. Erlbaum, 1977.
- Scarborough, D. L., Cortese, C., & Scarborough, H. S. Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, 1977. 3, 1-17.
- Schuberth, R. E., & Elmas, P. D. Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3, 27-36.
- STANNERS, R. F., JASTRZEMBSKI, J. E., & WESTBROOK, A. Frequency and visual quality in a word-nonword classification task. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14. 259-264.
- Stanovich, K. E. Attentional and automatic context effects in reading. In C. Perfetti & A. Lesgold (Eds.), *Interactive processes in reading*. Hillsdale, N.J: Erlbaum, 1980.
- Stanovich, K. E. Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, in press.
- STANOVICH, K. E., & WEST, R. F. Mechanisms of sentence

- context effects in reading: Automatic activation and conscious attention. Memory & Cognition, 1979, 7, 77-85.
- STERNBERG, S. Two operations in character recognition: Some evidence from reaction-time measurements. *Perception & Psychophysics*, 1967, 2, 45-53.
- STERNBERG, S. The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), Attention and performance II. Amsterdam: North Holland, 1969.
- Tulving, E., & Gold, C. Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*, 1963, 66, 319-327.

(Received for publication July 5, 1979; revision accepted July 30, 1980.)