

The effect of prose context on visual search for letters

ROBERT M. SCHINDLER
University of Rochester, Rochester, New York 14627

Subjects were given a simplified proofreading task in which they were instructed to circle every occurrence of a target letter in a prose passage or in a scrambled prose passage. It was found that the presence of a prose context enhanced the subjects' ability to find a target letter when the target letter was in a content word, but impaired the subjects' ability to find a target letter when it was in a function word. This interaction sheds light on a number of conflicting reports in the literature.

It has long been observed among people involved in proofreading that accurate search of prose for misspellings, word deletions, and other errors is an extremely difficult task (e.g., Crosland, 1924; Smith, 1903). There appears to be a persistent tendency to overlook at least some of the errors present. In the words of one author, "It certainly seems true that no matter how many times a set of proofs is read, more errors can be found" (Lee, 1965, p. 73).

Further, this phenomenon is not limited to search for errors, as can easily be demonstrated by asking someone to report the number of fs in the following sentence: "Finished files are the result of years of scientific study combined with the experience of many years." Most people report seeing only four or five of the six fs that occur in the sentence (Schindler & Jacobs, Note 1).

What is the cause of this tendency to overlook letters and words in prose? The first step in answering this question is to determine if it has anything to do with prose at all. It could be that these omission errors are simply due to brief lapses of attention that are likely to occur at random points during the performance of any prolonged task. Or, it is possible that these omission errors are due to properties of only letters or words. For example, fs might be overlooked because they appear similar to ts, or short words might always tend to be overlooked when placed in an array of longer words.

Several recent studies designed to test the effect of prose context on letter detection have produced conflicting results. Krueger (1970) instructed subjects to search for the single occurrence of a target letter in

a short passage or in a passage in which the words were rearranged to form scrambled prose. He found that search through prose was not only faster than search through scrambled prose, but it was also more accurate.

Healy (1976) observed the performance of subjects who searched a 100-word passage for all occurrences of the letter t. Like Krueger, she found that letter search was more rapid when the target-containing word was in a prose context than when it was in a scrambled prose context. But, unlike Krueger, Healy found that subjects made more errors when searching through prose than when searching through scrambled prose. In addition, Healy found that the t in the word "the" was especially likely to be missed; although only 28% of the ts in the prose passage were in the word "the," those ts comprised 62% of the ts that were overlooked. Further, the percentage of the ts in "the" that were missed in the scrambled prose passage did not differ significantly from the 62% missed in the prose passage. This suggests that, although there are fewer omission errors made in scrambled prose than in prose, the pattern of errors does not change when the meaningfulness of the context is destroyed.

Schindler and Jacobs (Note 1) presented subjects with the "Finished files..." sentence (see above), two additional sentences of similar structure, and scrambled versions of the three sentences. They found more errors occurred in the sentences than in the scrambled sentences, but that this difference was due entirely to the high probability of missing the fs in "of" and the ns in "in" and "on" when these words appeared in sentences. Although the sample of sentences used in this study was rather limited, this result does suggest that prose context increases the tendency to overlook the the letters in some words more than in others.

The present study was an attempt to explore further the role of prose context in the quantity and distribution of omission errors in a visual letter search task. In particular, an attempt was made to replicate the increase in search errors due to prose context,

This paper is based on a dissertation submitted in partial fulfillment of the requirements of the PhD degree at the University of Massachusetts. The author would like to thank Arnold Well, Alexander Pollatsek, Charles Clifton, Ralph Haber, and Keith Rayner for helpful comments and suggestions regarding this research. This paper was written while the author was supported by a postdoctoral fellowship from the National Institute of Mental Health. Requests for reprints should be addressed to Robert Schindler, Department of Psychology, University of Rochester, Rochester, New York 14627.

shown by Healy (1976) and by Schindler and Jacobs (Note 1), using a wider variety of passages and target letters. Also, an attempt was made to gauge the automaticity of these errors by varying the instructions in the task. Some subjects were told to read the passages and circle the target letters, while other subjects were told to find the target letters by searching the passages letter by letter. Finally, ratings of the perceived importance of the target-containing words were collected, using two different techniques, in order to test the hypothesis that the prose context increases errors only in words that are unimportant with respect to the meaning of the passage.

METHOD¹

Design and Materials

The letters a, e, i, o, f, h, n, s, t, and w were chosen as target letters. For each letter, a short passage that contained at least 20 occurrences of the letter was selected from *Reader's Digest*. These 10 passages were then modified (1) to contain exactly 60 words, and (2) to minimize the number of target letters occurring in consecutive words. The scrambled prose version of each of the 10 passages was created by rearranging the non-target-containing words of each prose passage to produce as meaningless and as syntactically illegal a word sequence as possible. With minor exception, the position of each target-containing word on the line and on the page was not altered by the scrambling procedure.

Each of the 10 prose and 10 scrambled prose passages was typed separately on a page. The original pattern of capitalization and punctuation used on the prose passages was retained in the scrambled prose passages as much as possible. Each passage was double spaced, and the target letter for the passage was printed (in both upper- and lowercase form) at the top of each page.

Each subject received five prose and five scrambled prose passages in alternating prose/scrambled prose order. For half of the subjects, a prose passage began the sequence, and for the other half, a scrambled prose passage occurred first. No subject received two versions of the same passage. For instance, if a subject received the prose passage using a as the target letter, he did not receive the scrambled prose passage that had a as the target letter. This design also insured that each subject searched for each of the 10 target letters.

Half of the subjects received "read" instructions ("read the sequence of words and circle the target letter every time it occurs"). The other half of the subjects received "search" instructions ("circle the target letter every time it occurs . . ." but do not read the words. Instead, "search each word letter by letter and totally ignore its meaning").

To assess the degree to which this instruction manipulation actually changed the subjects' search strategies, the subjects were given a surprise forced-choice recognition test immediately after they searched the 10 passages. The recognition test was constructed by selecting two distinctive non-target-containing words from each of the 10 passages and matching each of these 20 words with 2 words of equal length and comparable frequency of occurrence. The word taken from a passage and its two distractor words were typed in a row, so the recognition test consisted of 20 such rows. Assuming that reading the words of a passage is necessary for correct recognition of the words (see Craik & Lockhart, 1972), the score on this recognition test could be taken to measure the extent to which the subjects read the passages.

Two methods were used to obtain estimates for the perceived importance of the words in the passage. Half the subjects received estimation rating sheets: These consisted of triple-spaced prose passages, with a line under each word where the

subject could write a number corresponding to how important he felt the word was for the meaning of the passage. The other half of the subjects received telegraph rating sheets: These were simply double-spaced prose passages that the subject could use to cross out the least important words. Each subject was given five estimation and five telegraph rating pages that corresponded to the five prose passages he had searched. This rating method variable was crossed with the other between-subjects variable, read vs. search instructions, thus creating four groups of subjects.

Each subject received a questionnaire that asked his age, sex, dominant hand, year in school, and asked the subject to estimate how much he reads and how fast he reads. A second part of the questionnaire asked the subjects about their performance in the search task. They were asked to estimate (1) how often they had read the paragraphs, (2) how often they had read the scrambled paragraphs, (3) how often they had used the acoustic properties of the target letter, and (4) whether they found it more difficult to find the targets in the prose or the scrambled prose passages.

Procedure

The subjects were run in groups for one 45-min session. Each subject was given a booklet with the first part of the questionnaire as the cover sheet. After filling out the questionnaire, the subjects were given either read or search instructions. Subjects were given 1 min to search each page, since it had been determined from pilot subjects that this would be more than enough time for most people. The subjects were encouraged to check their work if they finished before the time was up. The interval between passages was approximately 10-15 sec.

Immediately after completing the tenth passage, the subjects were instructed to turn to the recognition test and to circle the one word in each row of three words that had occurred in the preceding passages. Subjects were given as much time as they needed to complete the recognition test. Following the test, they completed the second part of the questionnaire.

Next, instructions for the estimation or telegraph rating tasks were read. Estimation raters were told to use a 1-5 scale to rate the importance of each word to the meaning of the paragraph. The telegraph raters were told to pretend they had to send the 60-word passage as a telegraph message, but had only enough money for 40 words. Thus, their task was to cross out the least important 20 words. Subjects were given 5 min to complete the ratings of their five passages.

Subjects

Two hundred and four University of Massachusetts students served as subjects. Each subject was randomly assigned to one of the four groups: (1) read instructions, estimation rating method; (2) read instructions, telegraph rating method; (3) search instructions, estimation rating method; or (4) search instructions, telegraph rating method. Data from 4 subjects were discarded, leaving 50 subjects in each group. The discarded subjects included one who had problems with English, one who failed to follow instructions, and two who obviously failed to complete the searching of one or more of the passages. Subjects received course credit for their participation.

RESULTS

The 200 subjects made a total of 3,569 visual search errors. Since there was a total of 260 occurrences of target letters in the 10 passages, the average subject's error rate was 6.86%. Of the total number of errors, 3,554 were omission errors (failure to circle the occurrence of a target letter), and 15 were commission errors (circling a letter other than the target letter). The 15 commission errors were excluded from the

following analyses, and from here on, the terms "error" and "omission error" will be used synonymously.

Each subject's error rate for the prose and scrambled prose passages was computed, and the resulting mean error rates were 7.14% for the prose passages and 6.54% for the scrambled prose passages. A *t* test on the arc-sine transformation of these error rates² showed the difference between the prose and scrambled prose passages to be nonsignificant [$t(199) = 1.73, p < .10$].

Since there was no significant overall effect of passage meaningfulness (prose context vs. scrambled prose context), it was of interest whether there was an effect of passage meaningfulness for only those target letters which occurred in words which were relatively unimportant to the meaning of the passages. To test for this, the number of errors made when a target-containing word was in a scrambled prose context was subtracted from the number of errors made on the same word when it was in a prose context. This difference, the number of *prose-specific errors* (PSEs) for a target occurrence, was computed for each of the 260 target occurrences. If a target occurrence was more likely to be missed when it was in a prose context, then it had a positive number of PSEs. If a target occurrence was more often missed when surrounded by scrambled prose, then it had a negative number of PSEs.

Each target-containing word was rated for importance by two methods (100 subjects used each method). The estimation ratings ranged from 2.18 to 3.96 where 3.00 was the midpoint of the scale. The higher numbers indicate greater rated importance. The telegraph rating for a word equaled the number of subjects who crossed out that word in the telegraph rating task. These ratings ranged from 0 to 48, but were multiplied by -1 so that, like the estimation ratings, the higher numbers would indicate greater importance. Agreement between these two methods of rating the 218 different target-containing words was fairly high [$r(216) = .78, p < .001$].

As can be seen from the correlation coefficients in Table 1, the less the rated importance of a word, the more PSEs were made on target letters occurring in that word. In other words, it is possible that the prose context increased the number of omission errors only for the less important words. However, as can

also be seen in Table 1, word frequency (from Kučera & Francis, 1967)³ and word length are also correlated with PSEs, and could plausibly be a factor in causing PSEs. Further, the two importance ratings, word frequency, and word length are all highly correlated with each other, making it extremely difficult to sort out the underlying importance of each variable.⁴

It has often been hypothesized, especially by researchers interested in education (e.g., Fries, 1952; Robeck & Wilson, 1974) and linguistics (e.g., Brown & Bellugi, 1964; Ervin-Tripp, 1970; Hormann, 1971) that the relatively small set of short, frequently occurring words that have little lexical meaning out of context seems to play a special role in our language. Although exactly how the role of these "function words" differs from that of "content words" has not been made clear, it does seem as though an individual function word makes a smaller contribution to the meaning of a prose passage than a content word. In fact, language without function words has been called "telegraphic" (Brown & Bellugi, 1964) because of its similarity to language with the least important words omitted.

Although it must be kept in mind that the distinction between function words and content words is itself a hypothesis, the distinction seems to be useful for analyzing the present data. For example, the term "function word" is more succinct and provocative than "short, high-frequency low-importance words." Moreover, the distinction provides a nonarbitrary way of dichotomizing the words in the present sample. The alternative would be to group the words arbitrarily or to limit analyses to correlational techniques.

Thus, each target-containing word was classified as a content word or a function word. The definition used to define function words agrees in general with descriptions usually given, but was simplified so it could be made as explicit, and thus as testable, as possible. The following six kinds of words were considered function words: (1) all prepositions, (2) all conjunctions, (3) all auxiliary verbs, (4) all forms of the verb "to be" (and "seems," "appears," etc., when used synonymously with forms of "to be"), (5) all pronouns, and (6) all possessive adjectives; "a," "an," and "the"; and "this," "that," "these," "those," "what," "which," and "such" when used as adjectives. All words not classified as function words were considered to be content words. *Webster's New Collegiate Dictionary* (1973) was used as the authority on part of speech. Using this method, 76 of the 218 target-containing words were classified as function words. Although the function words were defined by linguistic category, they, on the average, were shorter than the content words (3.6 vs. 6.7 letters), occurred more frequently in printed English (16,000 vs. 136 occurrences/million words), and were less important (2.71 vs. 3.48 by the estimation method; -20.7 vs. -8.4 by the telegraph method).

Each subject's error rates were computed separately

Table 1
Correlations Between PSEs, Word Importance, Log Word Frequency, and Word Length (All *dfs* = 258)

	ER	TR	LWF	WL
PSE	-.23	-.15*	.23	-.23
Estimation Rating		.78	-.75	.58
Telegraph Rating			-.46	.31
Log Word Frequency				-.76

Note—ER = estimation rating; TR = telegraph rating; LWF = log word frequency; WL = word length.

* $p < .01$; all other $ps < .001$.

Table 2
Mean Error Rates for First Target Occurrences in Content Words and Function Words and For All Target Occurrences in Prose and Scrambled Prose Contexts

	Content Words	Function Words	All Target Occurrences
Prose	2.83	13.65	7.14
Scrambled Prose	3.41	11.07	6.54
Difference (Mean PSEs)	-.58	2.58	.60

for first target occurrences⁵ in content words and function words, and the means are displayed in Table 2. An analysis of variance indicated that the error rate was higher for target letters in function words than for target letters in content words [$F(1,199) = 304.79, p < .001$], and, for first target occurrences, there was a higher error rate for prose than for scrambled prose [$F(1,199) = 5.25, p < .03$]. However, there was also a highly significant interaction between these two variables [$F(1,199) = 15.27, p < .001$]. The error rate for target letters in content words was lower when the content words were embedded in prose than when they were embedded in scrambled prose [$t(199) = 2.22, p < .03$]. This was also true when target occurrences, rather than subjects, were considered the random effects variable [$t(141) = 2.58, p < .02$]. By contrast, the error rate for target letters in function words was higher when the function words were in prose than when they were in scrambled prose [$t(199) = 3.22, p < .005$]. This difference was also significant when target occurrences were considered the random effects variable [$t(75) = 2.69, p < .01$].

To determine if the search instructions were effective in causing the subjects to make fewer prose-specific errors, a three-way analysis of variance was performed on the error rate data separated by instruction (read vs. search), meaningfulness (prose vs. scrambled prose), and target type (content word, function word, and second target occurrence). Neither the main effect of instructions, nor any interaction involving the instruction variable even approached significance (all $ps > .20$).

To further explore the effect of the instruction manipulation on search performance, each subject's word recognition test scores for prose and scrambled prose passages were computed, and the means are displayed in Table 3. An analysis of variance indicated

Table 3
Mean Percent Correct Word Recognitions for Prose and Scrambled Prose Passages and for Read and Search Instructions

	Instructions	
	Read	Search
Prose	67	63
Scrambled Prose	58	58

that recognition scores were higher in prose passages [$F(1,198) = 20.86, p < .001$], but that there was no significant effect of instructions on recognition scores [$F(1,198) < 1$] and no significant interaction between passage meaningfulness and instructions [$F(1,198) = 2.07, p < .20$]. This lack of effect of instructions on both search errors and recognition scores suggests that either (1) the read or search instructions were voluntarily ignored by most subjects, or (2) both the pattern of visual search errors and the number of words read were caused by factors not completely under voluntary control.

To determine if the prose context affected how errors were distributed within single words, the 40 words that contained at least two occurrences of a target letter were subjected to further analyses (35 of these were classified as content words, 5 as function words). Each subject's error rates on first target occurrences in prose and scrambled prose and on second target occurrences in prose and scrambled prose were computed, and the means are shown in Table 4. An analysis of variance indicated that there was no significant difference between the prose and scrambled prose error rates for these 40 words ($F < 1$). However, subjects made more errors on second target occurrences than first target occurrences [$F(1,199) = 230.23, p < .001$], and this effect interacted with passage meaningfulness [$F(1,199) = 8.15, p < .01$]. The latter result suggests that prose context did have some effect on how the errors were distributed within single words.

Since the second target occurrences appeared in later positions in the words, and since Corcoran (1966) reported more search omission errors for targets occurring in later letter positions, it is of interest whether the higher number of errors and PSEs in second target occurrences could be due to the letter position variable. The partial correlation of the position in the word of all first target occurrences with the number of errors made (controlling for word importance, word frequency, and word length) was significant [$r(212) = .15, p < .03$], indicating a tendency for there to be more errors in later letter positions. This relationship confirms Corcoran's finding, but is probably too weak to completely account for the large number of errors in second target occurrences. The finding of no similar relationship between letter position of first target occurrences and PSEs [$r(212) = -.04, p > .5$] also suggests that factors other than letter position are responsible for the high number of errors and PSEs in second target occurrences.

It is possible that the high number of errors in second target occurrences was due to subjects assuming (at least for a while) that there would be only one target occurrence per word and stopping their search of a word when the first target occurrence was found. Another possibility is that circling the first target occurrence disrupted their search briefly, thus increasing the

Table 4
Mean Error Rates for First and Second Target Occurrences for
the 40 Two-Target-Containing Words in Prose
and Scrambled Prose Contexts

	Target Occurrences	
	First	Second
Prose	2.71	10.54
Scrambled Prose	3.73	9.14
Difference (Mean PSEs)	-1.02	1.40

probability of their missing a target occurring among the next few letters. However, such explanations would leave unclear why prose context tended to increase the proportion of errors made on second target occurrences.

Correlation coefficients were computed between the subjects' responses on each of the questionnaire items and the mean number of prose-specific errors. Whether or not the PSEs of the 200 subjects were separated by target type, the obtained correlations were extremely small and none reached statistical significance.

DISCUSSION

The present results make it clear that omission errors made during visual search of prose are not due simply to random lapses of attention. Target letters that occur in function words (or short, unimportant, high-frequency words) are far more likely to be missed than target letters that occur in content words (or long, important, low-frequency words).

This conclusion is supported by previous research. Corcoran (1966) found that when subjects searched prose for the letter *e*, they were much more likely to miss the *e* in "the" than the *e*s in other words. Hatch, Polin, and Part (1974) reported that subjects made more search omission errors for target letters in function words than for target letters in content words whether the vocabulary and syntax of the passages were easy or complex. Moreover, they found that proficient readers of English who had learned it as a second language were also more likely to miss target letters that occurred in function words.

The present results also demonstrate that the distribution of search errors in prose is due to more than properties of letters and single words. The presence of a prose context increases the search accuracy of content words and decreases the search accuracy of function words. This finding suggests that the content word/function word distinction might provide a useful way of viewing previous studies of the effects of prose context on visual search for letters. While Schindler and Jacobs (Note 1) found more visual search errors for prose than for scrambled prose, they did so only for targets occurring in the function words "in," "on," and "of." Krueger's (1970) result of fewer search errors in prose than in scrambled prose fits well with the fact

that 19 out of his 24 target-containing words were content words (Krueger, Note 2). Healy (1976, Note 3) found that the *t* in the function word "the" was missed more often when it was embedded in prose than when it was embedded in scrambled prose.

However, Healy's (1976) subjects also made more errors in general on words embedded in prose context, thus causing the percentage of errors that were omissions of the *t* in "the" to remain the same. This overall effect of prose context on error rate conflicts with the present finding of no overall difference between the error rates in prose and scrambled prose passages. The overall difference Healy observed may have been due to her subjects' making more errors in the prose passages because they were able to search those passages more rapidly. Following this explanation, no general effect of prose context on error rate may have been found in the present study because subjects were given an equal amount of time to search prose and scrambled prose passages and were encouraged to use the entire time interval.⁶

If one considers that the use of a "reading strategy" is responsible for the tendency to make more visual search errors in function words than in content words when searching prose, the difference in the distribution of errors in the prose and scrambled prose passages could be due to an increased use of the reading strategy when searching prose. The higher word recognition scores for prose than for scrambled prose passages would support this interpretation, since they suggest that words were more often read in the prose passages. However, even in the scrambled prose passages, the tendency of search errors to occur in the function words was quite pronounced. If this tendency is due to the use of a reading strategy, why is such a strategy elicited at all by arrays of words that contain many syntactic and semantic illegalities?

One potential explanation for this involves the possible automaticity of the reading strategy. Although it is possible that the experimental instructions were simply ignored, it is also possible that their failure to affect the subjects' search errors and recognition scores could indicate that the reading strategy may be, to a large extent, an involuntary response to the visual features of printed English. Thus, since the scrambled prose passages contained the same words, paragraph format, punctuation, capitalization, and mixture of long and short words as the prose passages, the reading strategy may have been elicited fairly often by the scrambled prose passages even though they made little sense as language. In other words, it is possible that years of reading experience cause the reading strategy to become a conditioned response to printed English that can, to a degree, generalize to similar stimulus situations.

By what mechanisms might this reading strategy improve the perception of letters in content words and impair the perception of letters in function words? One possibility is that the reading strategy involves

forming an acoustic representation of the words, and that target letters in function words are less salient aspects of this representation. However, Healy (1976) has shown that it is not the pronunciation of the t in "the" that is responsible for its high probability of being missed, and Chen (1976) has found that even congenitally deaf subjects are more likely to miss the e in "the" than es in other words. A second possibility is that the reading strategy involves habitual patterns of eye movements. For example, it has recently been found (Rayner, 1977) that the word "the" receives fewer and generally shorter fixations than less frequent words. A third possibility is that the reading strategy consists of rules for the allocation of attention which cause very little visual attention to be given to words that are likely to be unimportant. Further, it may be that such rules become automatized over the course of years of reading experience (see LaBerge & Samuels, 1974; Miller, Galanter, & Pribram, 1960).

These possible mechanisms of the reading strategy are by no means mutually exclusive and there are no doubt other possibilities. A promising direction to take in following up this research would be to look for changes in the pattern of visual search errors caused by the scrambling of other familiar stimuli. For example, although many researchers have shown that, in general, letters can be perceived more accurately in words than in nonwords, it may be that there are unimportant letters or subletter features that are more difficult to perceive when in a word context. Or, while it may be easier to find a discrete object in a coherent scene than in a scrambled scene, the scrambled scene may have the benefit of making "embedded figures" easier to find. It may, in fact, turn out that the development of perception in general involves the making of useful tradeoffs rather than an overall increase in acuity.

REFERENCE NOTES

1. Schindler, R. M., & Jacobs, P. I. *What do we see when we read?* Paper presented at the annual meeting of the Eastern Psychological Association, New York, April 1976.
2. Krueger, L. E. Personal communication, March 23, 1976.
3. Healy, A. F. Personal communication, April 16, 1976.

REFERENCES

- BROWN, R., & BELLUGI, U. Three processes in the child's acquisition of syntax. *Harvard Education Review*, 1964, 34, 133-151.
- CHEN, K. Acoustic image in visual detection for deaf and hearing college students. *Journal of General Psychology*, 1976, 94, 243-246.
- CORCORAN, D. W. J. An acoustic factor in letter cancellation. *Nature*, 1966, 210, 658.
- CRAIK, F. I. M., & LOCKHART, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.

- CROSLAND, H. R. An investigation of proofreader's illusions. *University of Oregon Publications*, 1924, 2(6).
- ERVIN-TRIPP, S. M. Substitution, context, and association. In L. Postman & G. Keppel (Eds.), *Norms of word association*. New York: Academic Press, 1970. Pp. 383-395.
- FRIES, C. C. *The structure of English*. New York: Harcourt, Brace, & World, 1952.
- HATCH, E., POLIN, P., & PART, S. Acoustic scanning and syntactic processing: Three reading experiments—first and second language learners. *Journal of Reading Behavior*, 1974, 6, 275-285.
- HEALY, A. F. Detection errors in the word the: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, 2, 235-242.
- HORMAN, H. *Psycholinguistics*. New York: Springer-Verlag, 1971.
- HOWES, D. H., & SOLOMON, R. L. Visual duration threshold as a function of word probability. *Journal of Experimental Psychology*, 1951, 41, 401-410.
- KRUEGER, L. E. Search time in a redundant visual display. *Journal of Experimental Psychology*, 1970, 83, 391-399.
- KUČERA, H., & FRANCIS, W. N. *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press, 1967.
- LABERGE, D., & SAMUELS, S. J. Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 1974, 6, 293-323.
- LEE, M. *Bookmaking: The illustrated guide to design and production*. New York: R. R. Bowker, 1965.
- MILLER, G. A., GALANTER, E., & PRIBRAM, K. *Plans and the structure of behavior*. New York: Holt, Rinehart, & Winston, 1960.
- RAYNER, K. Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, 1977, 5, 443-448.
- ROBECK, M. C., & WILSON, J. A. R. *Psychology of reading: Foundations of instruction*. New York: Wiley, 1974.
- SCHINDLER, R. M. Visual search for letters through prose and scrambled prose. Unpublished doctoral dissertation, University of Massachusetts, 1976.
- SMITH, A. M. *Proof-reading and punctuation*. Philadelphia: Drexel Institute, 1903.
- Webster's New Collegiate Dictionary*. Springfield, Mass: G. & C. Merriam, 1973.

NOTES

1. Further details are given by Schindler (1976).
2. In this paper, all t and F tests on percentage data were performed on the arc-sine transformation of the data in order to correct for deviations from normality.
3. The log of the word frequency was used to compute the correlations since log word frequency was more highly correlated with PSEs than was simple word frequency, and since it has been previously reported (e.g., Howes & Solomon, 1951) that the effects of word frequency on visual processing are linearly related to the log frequency variable. However, this higher correlation of PSE with log word frequency might be interesting in itself, since log word frequency is, at least in the formal sense, more closely related to the informativeness of the words than is simple word frequency.
4. Examination of all first-order partial correlations between the variables in the four columns of Table 1 and PSEs failed to suggest any causal priority between the importance, frequency, and length factors. However, the partial correlations did indicate that telegraph ratings are superfluous given the estimation ratings.

5. Of the 218 target-containing words, 40 contained more than one occurrence of a target letter. In order to simplify the comparison between words of different lengths, only the first target occurrences in these words were used for the content word vs. function word analyses. The errors in the 40 second target occurrences are discussed later in the paper.

6. Healy (1976, Experiment 4) also found that more errors were made on high-frequency nouns than on low-frequency nouns in scrambled prose passages. Although in the word sample of the present study, there was no significant correlation

[$r(61) = .02, p > .5$] between word frequency and number of errors in nouns in the scrambled prose passages (frequency range = 0-949 occurrences/million words), the present results in no way rule out the possibility that factors in addition to those represented by the content word/function word distinction play a role in causing visual search errors.

(Received for publication September 27, 1977;
accepted December 5, 1977.)