PROOFREADING ERRORS ON THE WORD THE: NEW EVIDENCE ON READING UNITS

Alice F. Healy

Abstract. In three experiments, subjects read passages and circled misspellings in them. In Experiment 1, misspellings were introduced by transposing two adjacent letters in a word. Subjects made a disproportionately small number of errors on the word the in the transposition-proofreading task. In Experiments 2 (prose passages) and 3 (scrambled nouns), misspellings were introduced by replacing instances of the letter t with the letter z. Letter-detection tasks in which subjects searched for instances of t in passages without misspellings were compared to the substitution-proofreading tasks in which the subjects, in effect, searched for z. Subjects in Experiment 2 made a disproportionately large number of errors on the in letter detection but not in proofreading. In Experiment 3, subjects made more errors on common than on rare nouns in letter detection but not in the proofreading. The results provide evidence that common words are normally read in units larger than letters but are read in letter units when they are misspelled.

INTRODUCTION

Letter-detection tasks have been used to obtain evidence for the size of the units used in reading (Healy, 1976; Drewnowski & Healy, 1977). The present study sought to obtain evidence on the same issue by employing various proofreading tasks and by comparing proofreading and detection tasks.

Subjects have been found to be especially likely to make errors on the word the in letter-detection tasks (Corcoran, 1966; Healy, 1976). By ruling out hypotheses concerning the pronunciation and location of the target letters and the semantic and syntactic redundancy of the, Healy (1976) argued that the preponderance of letter-detection errors on the was due to its high frequency, which made it especially likely to be read as a unit, or chunk, rather than in

*This paper is to appear in the Journal of Experimental Psychology: Human Perception and Performance, 6.

†Also Yale University.

Acknowledgment: This research was supported in part by FHS Grants M926573 and RR07015, NSF Grant BNS 77-00077, and a grant from The Spencer Foundation to Yale University and NICHD Grant HD01994 and BRS Grant RR05596 to Haskins Laboratories. The author was supported by a Junior Faculty Fellowship from Yale University during the preparation of this manuscript. The author is indebted to C. Paul for help with the design and conduct of the experiments, to J. Gottlieb and J. Chung for help with the data analyses, and to R. G. Crowder, A. Drewnowski, W. K. Estes, M. Kubovy, A. G. Levitt, and B. H. Repp for helpful discussions about this research.

[HASKINS LABORATORIES: Status Report on Speech Research SR-58 (1979)]
terms of its component letters. As further support for this argument, Healy (1976) demonstrated that in a passage of scrambled nouns, subjects were more likely to make letter-detection errors on common nouns than on rare nouns. These findings were extended by Drewowski and Healy (1977) who employed the trigram the as well as the letter t as a detection target. A preponderance of detection errors was found on the word the, rather than on words with embedded the trigrams (such as bathed), for both targets. In addition, more detection errors were found on the word the when it occurred in an appropriate syntactic context than when it did not. Finally, more errors occurred on the word the in a passage typed in standard paragraph format than in passages in which word-group identification was disturbed by the use of mixed typecases or a list format. On the basis of these findings, Drewowski and Healy (1977) concluded that familiar word sequences may be read in units larger than the word, probably short syntactic phrases or word frames, such as "on the ---."

A set of five hypotheses about the reading process that are consistent with the findings from detection tasks was proposed by Drewowski and Healy (1977). These hypotheses will be referred to here as the "unitization hypotheses." Specifically, (1) a hierarchy of processing levels was identified and was defined in terms of the units available at each level, such as letters, words, and phrases. (2) It was proposed that the completion of processing at a given level is tantamount to the identification of the unit at that level. In accordance with this hypothesis, detection tasks that require subjects to identify targets at a given level allow one to monitor the completion of processing at that level. For example, letter-detection tasks allow one to monitor the completion of processing at the letter level. It was further postulated (3) that subjects process text in parallel at the various levels available to them and (4) that once a unit has been identified at a given level, the subjects will proceed to the next unit at that level without necessarily completing the processing of units at the lower levels in the hierarchy. For example, once the subjects have identified a word, they may move on to the next word in the text, which they will process at all levels in parallel, without necessarily completing the processing of all the letters within the word just identified. (5) Familiarity with a unit at a given level should facilitate processing of it. For example, common words should be processed at the word level more easily than rare words. In particular, the, which is the most common word in English, should be processed at the word level more easily than other words.

It should be noted that the unit processing depicted by these hypotheses is what has been described elsewhere as "automatic processing" (see, for example, LaBerge & Samuels, 1974; Shiffrin & Schneider, 1977). Hence, it is assumed that when a given unit is processed, the constituent elements of that unit may be hidden from conscious perception. For example, when a word unit is processed, the constituent letters of that word may be hidden from conscious perception, even when the word itself is consciously perceived. The hypothesis that familiarity of a given unit facilitates processing is compatible with the extensive evidence that automatic processes require considerable training to develop. Despite the fact that these hypotheses describe automatic processes, rather than the slower controlled processes, the hypotheses are certainly compatible with the possibility that controlled processes may be occurring in the reading task as well as automatic processes. For example, in the case when all the letters in a given word are identified before the word itself, the subjects may use controlled processing to identify
the word before moving on to the next string of letters. In fact, such a strategy seems reasonable for subjects who are reading for meaning. Although it seems reasonable to postulate controlled processing of a higher-level unit after the automatic processing of the constituent lower-level units is completed, it seems less reasonable to postulate controlled processing of the constituent lower-level units after the automatic processing of the higher-level unit is completed. Indeed such an asymmetry between the levels of processing is implied by the fourth hypothesis.

On the basis of the unitization hypotheses, one would expect that subjects would be likely to make many letter-detection errors on the word **the** when reading standard text, since they would tend to complete processing at the word or phrase level before the letter level. More generally, these hypotheses lead one to expect more letter-detection errors on common words than on rare words, since the probability of faster processing at the word level than at the letter level would be greater for common words than for rare words. Another prediction from these hypotheses is that the tendency to make letter-detection errors on **the** would be greatly reduced if the processing of units larger than the letter were disturbed by typing every other letter in capitals. These predictions have all been confirmed in the previous studies by Healy (1976) and Drewnowski and Healy (1977) and were examined again in the present study, along with an examination of proofreading errors.

What do the unitization hypotheses lead one to expect about errors subjects make in proofreading? In order to answer this question, we must be more explicit about the nature of the proofreading task. In the proofreading task used in the first experiment of the present study, which we call a "transposition-proofreading" task, subjects were told to encircle every instance of a misspelling, and misspellings were introduced into a prose passage by transposing two adjacent letters within a word. Clearly, in such a task the processing of the misspelled words should be minimally disturbed at the letter level, since they contain the same letters as their correct versions. Effects of letter transpositions would be expected at the letter level only to the extent that there are sequential dependencies in letter recognition. (For example, processing of the letter **y** may be facilitated if it occurs after the letter **g**.) However, automatic processing at the word level and above should be greatly disturbed, if not prohibited, for the misspelled words, since no word-level unit would correspond to the anomalous sequence of letters. Subjects would be limited to the slower controlled processing at the word level and above. On the basis of these considerations alone, one would expect no difference between **the** and other words. However, an additional factor is also relevant to this task—namely the subjects' ability to identify a letter string as a misspelling. Subjects should find it easier to identify a misspelling caused by a single transposition in a very short word such as **the** than in a longer word, because in a short word the misspelled letter string would have less in common with the correctly spelled word. (See Holbrook, 1978a, for a similar argument and a demonstration that the perceived similarity between a word in its misspelled and correct forms depends on word length.) Thus, according to these hypotheses, in this transposition-proofreading task subjects should make relatively few errors on **the**, as opposed to the preponderance of errors on **the** in letter-detection tasks.
Although the unitization hypotheses lead one to expect large differences in the patterns of errors in the transposition-proofreading task and in a comparable letter-detection task, other reasonable views would lead one to expect similar patterns in the two tasks. For example, Corcoran (1966) explained the preponderance of letter-detection errors on the word the in terms of the redundancy of the. Healy (1976) termed this the "redundancy hypothesis." Although there are several different kinds of redundancy (Smith, 1971), only semantic and syntactic redundancy is referred to by this hypothesis. Specifically, Corcoran (1966) suggested that the may be "'taken for granted' and thus not scanned" (p. 658). (See Hatch, Polin, & Part, 1974, for an expanded discussion of the importance of predictability, or syntactic and semantic redundancy, in this task.) According to this hypothesis, a preponderance of errors would be expected on the in the transposition-proofreading task as well as in a comparable letter-detection task, since the would be syntactically and semantically redundant in both situations. If subjects fail to scan the because they take it for granted, they would not be able to detect misspellings of it.

Similarly, Schindler (1978) proposed an explanation in terms of eye-movement patterns for his finding that subjects made more letter-detection errors on function words than on content words. He described results from an experiment by Rayner (1977) demonstrating that the received fewer and usually shorter fixations than other words. Although Schindler did not propose an explanation for the eye-movement patterns, it is likely that they would be guided by the subjects' expectations based on prior word context. Words expected to be of little informational content would be likely to be skipped. Schindler (1978) also considered the possibility that subjects give very little visual attention to words that are likely to be unimportant. Both of these hypotheses by Schindler are consistent with, if not merely restatements of, the redundancy hypothesis.

A recent proofreading experiment by Holbrook (1978b) provides some direct support for the redundancy hypothesis. Holbrook found a significant positive correlation between the subjective verbal uncertainty of words, as measured by a Cloze test (Taylor, 1953), and the detection of typographical errors in those words.

The results of an earlier proofreading task, although inconclusive, might also lead one to expect a similar pattern of results in the transposition-proofreading and comparable letter-detection tasks. Corcoran (1967) conducted a proofreading experiment in which letters were omitted from various words in a prose passage, and subjects were asked to indicate the locations where letters were missing. As in his earlier letter-detection task (Corcoran, 1966), Corcoran found that subjects made more errors on silent es than on pronounced es in the proofreading task. Although errors were most frequent on the when the task was to detect the letter e (Corcoran, 1966), when the task was proofreading, the probability of failing to detect a missing e from the was roughly equivalent to the overall probability of failing to detect a missing letter. Corcoran (1967) does indicate, though, that subjects made significantly more errors in proofreading on the e in the than on other terminal pronounced es. The results from Corcoran's (1967) proofreading task are, therefore, somewhat ambiguous for the. It is also difficult to determine what to expect on the basis of the unitization hypotheses in proofreading with
omitted letters, because automatic processing at both the word and letter levels (in the case of the omitted letters) would be virtually prohibited. For that reason Corcoran's proofreading task was not employed in the present study.

EXPERIMENT 1

Method

Subjects. Ninety-six male and female Yale undergraduates participated as subjects. The data from five additional subjects were not analyzed because those subjects had participated previously in similar experiments.

Design and materials. A single typewritten passage was employed. The passage was based on a 321-word prose passage taken from The Social Animal by Elliot Aronson. Forty misspellings were introduced into this passage in a pseudorandom fashion so that misspellings occurred on exactly two words in every block of 16, excluding the final word of the passage. Each of these misspellings involved a transposition of two adjacent letters in a word. None of these transpositions yielded a new word except for two which yielded very infrequent words (fro from for and eh from he)\(^1\). Exactly 11 of the 40 misspellings involved the word the. (There were 38 thes in the passage as a whole.) Six of the misspellings of the were obtained by transposing the last two letters, forming a letter string which is pronounceable (teh), and five of the misspellings were obtained by transposing the first two letters, forming a letter string which is not pronounceable (hte).

Each subject was shown a mimeographed copy of the passage, preceded by a mimeographed sheet of instructions.

Procedure. The subjects were tested in a group session conducted in a classroom. The subjects were instructed to read the prose passage at their "normal reading speed," but whenever they came to a spelling error they were to encircle it with their pen or pencil. The subjects were told that if at any time they realized that they missed an error in a previous word, they should not retrace their steps to encircle it and that they should not slow down their reading speed in order to be overcautious about getting the errors.

Results and Discussion

The results are summarized in Table 1, which includes the means and the standard errors of the means for the percentage of proofreading errors made by the subjects (out of 40 possible errors), for the percentage of proofreading errors on the (out of 11 possible errors), and for the conditional percentage of proofreading errors on the given a proofreading error. All errors considered here and in the subsequent analyses in this paper were omission errors (misses). The conditional percentage of proofreading errors on the was derived for a given subject by determining the ratio of the number of proofreading errors on the to the total number of proofreading errors. By chance alone, the conditional percentage should be 27.5%, since 11 of the 40 transpositions occurred in the. Healy (1976) and Drewowski and Healy (1977) found this measure to be the most sensitive index of performance in their
detection tasks, since it is unaffected by the speed-accuracy tradeoff typically found in such tasks.

Table 1

Means and Standard Errors of Means for Percentages of Transposition-Proofreading Errors in Experiment 1

<table>
<thead>
<tr>
<th>Error percentage</th>
<th>p(error)</th>
<th>p(error on the)</th>
<th>percentage</th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>11.6</td>
<td>3.3</td>
<td>6.2</td>
<td>93</td>
</tr>
<tr>
<td>(SE_M)</td>
<td>0.8</td>
<td>0.7</td>
<td>1.4</td>
<td></td>
</tr>
</tbody>
</table>

Note. In this table and in the succeeding tables in this paper, the total number of subjects (N = 96) does not equal the number of subjects on which the mean conditional error percentage is based (N'), since not all subjects made errors on the passage.

\(^a\)The value of \(p(error on the|error)\) expected by chance alone is 27.5.

Although subjects failed to detect about 12% of all misspellings, they missed only about 3% of the misspellings that involved the. In fact, of the 93 subjects who made proofreading errors on this passage, only 26 made any error on the. The conditional percentage of proofreading errors on the given an error was significantly below chance level, \(t(92) = 21.5, p < .001\). Contrary to the redundancy hypothesis, it is clear from these results that subjects do not fail to scan the when proofreading, but rather are extremely accurate at detecting misspellings of this word. Further, the pronounceability of misspellings seems to be of little consequence, since very few errors were made on the, whether it was misspelled as teh or as hte.

In order to test directly the hypothesis that the subjects' ability to identify a letter string as a misspelling depends on word length, the percentage of proofreading errors was computed as a function of word length for all misspelled words excluding the. (See Table 2.) There were six misspelled two-letter words in the passage, four misspelled three-letter words (excluding the), six misspelled four-letter words, four misspelled five-letter words, and nine misspelled words six to ten letters long. In accordance with the hypothesis proposed above, subjects made more errors on the words five to ten letters long than on the shorter words, \(F(4,380) = 81.4, MS_e = 162, p < .001\). Furthermore, the percentage of errors on the (see Table 1) was slightly, but significantly, greater than the percentage of errors on other three-letter words (see Table 2), \(F(1,95) = 14.3, MS_e = 26, p < .001\).
Table 2

Means and Standard Errors of Means for Percentages of Transposition-Proofreading Errors Excluding those on the Word the in Experiment 1 as a Function of Word Length

<table>
<thead>
<tr>
<th>Word length</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6-10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error percentage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>7.6</td>
<td>0.5</td>
<td>6.2</td>
<td>22.9</td>
<td>27.9</td>
</tr>
<tr>
<td>SE&lt;sub&gt;M&lt;/sub&gt;</td>
<td>1.2</td>
<td>0.5</td>
<td>1.1</td>
<td>1.8</td>
<td>2.1</td>
</tr>
</tbody>
</table>

EXPERIMENT 2

Although the results of Experiment 1 are compatible with the unitization hypotheses, the additional hypothesis concerning the subjects' ability to identify a letter string as a misspelling must be added to account for performance. (On the basis of the unitization hypotheses alone, no differences between misspellings of the and of other words were expected, since no word-level units would correspond to the misspelled letter strings. However, the hypothesis that subjects would find it easier to identify a misspelling in a short word than in a longer word correctly predicted that subjects would make relatively few errors on the short word the.) For this reason, we constructed a new "substitution-proofreading" task that eliminated the need for the subjects to identify a letter string as a misspelling. The new task therefore enabled us to test more directly the unitization hypotheses as they apply to proofreading. The new task also had the advantage of permitting an elegant comparison of proofreading and letter detection.

Specifically, subjects were told to encircle every instance of a misspelling, and misspellings were introduced by replacing each instance of the letter t with the letter z. Subjects were informed of this fact as well as of the important fact that there were no other zs in the passage, so that each z represented a misspelling. Superficially, this substitution-proofreading task is strictly analogous to a detection task in which subjects search for ts in the corresponding passage without misspellings. According to the unitization hypotheses, though, the pattern of results should be very different for the analogous substitution-proofreading (z-circling) and letter-detection (t-circling) tasks. In the substitution-proofreading task, as in the transposition-proofreading task, processing at the letter level should be minimally affected in the misspelled words. Such an effect is expected only to the extent that the subjects' baseline ability to detect z is different from their baseline ability to detect t. In addition, as in the transposition-proofreading task, automatic processing of a given letter sequence at the word level and above should be greatly disturbed, if not
prohibited, by z substitutions. Subjects would be restricted to the slower controlled processing at these higher levels. The additional factor considered for the transposition-proofreading task—that the subjects' ability to identify a letter string as a misspelling depends on word length—should not be relevant to the present substitution-proofreading task, since it is made clear to the subjects that all zs represent misspellings. Considering the factors that are relevant to this task, one would not expect subjects to make either a disproportionately large or a disproportionately small number of errors on the in the substitution-proofreading task, but one would expect subjects to make a disproportionately large number of errors on the in the analogous letter-detection task.

Although a different pattern of results is expected for the present substitution-proofreading and letter-detection tasks according to the unitization hypotheses, no difference between the two tasks is expected according to the redundancy hypothesis, because the syntactic and semantic redundancy of the would not be changed by substituting z for t.

The effect of mixed typecase was also examined in this experiment. As in the study by Drewnowski and Healy (1977, Experiment 3), subjects were given passages typed in standard format and passages typed with every other letter in capitals in order to disturb the use of reading units larger than the letter. According to the unitization hypotheses, the conditional percentage of errors on the given an error should be reduced in the passages with mixed typecase relative to the passages typed in the standard fashion. In contrast, the redundancy hypothesis could not provide a simple account of any differences in conditional percentages for the two types of passages, since the syntactic and semantic redundancy of the would not change with a change in typecase.

Method

Subjects. The same subjects were employed as in Experiment 1. The subjects performed Experiment 2 immediately after completing Experiment 1.

Design and materials. Four passages were used, all based on a 100-word prose passage from Golding's Lord of the Flies. The basic passage (the "unmixed t passage") included 40 ts, 11 of them in the word the. This passage was identical to that employed by Healy (1976, Experiment 1). The second passage ("unmixed z passage") was identical to the unmixed t passage except that every t was replaced by a z. There were no other zs. None of the letter strings containing a z in the unmixed z passage were English words. The third and fourth passages ("mixed t and mixed z passages") were identical to the unmixed t and unmixed z passages, respectively, except that every other letter was typed in capitals. There were two versions of these passages. In version A the odd letters were capital, and in version B, the even letters were capital.

Each subject was shown mimeographed copies of all four passages, typed on separate sheets of paper and stapled together. For all subjects, the four passages were divided into two sets—the t passages and the z passages. Each set of passages was preceded by a mimeographed sheet of instructions. The order of the two sets was counterbalanced across subjects. A given subject
was shown only one version (A or B) of the mixed passages. The version of the mixed t passage shown to a given subject matched the version of the mixed z passage shown to him or her (e.g., subjects shown version A of the mixed t passage were shown version A of the mixed z passage). Furthermore, the order of the unmixed and mixed passages was counterbalanced across subjects, but the order of the unmixed and mixed t passages shown to a given subject corresponded to the order of the unmixed and mixed z passages shown to him or her. The three divisions among subjects (t first versus z first; version A versus version B of mixed passages, and order of unmixed and mixed passages) were orthogonal to each other, so that there were approximately equal numbers of subjects (11-13) in the eight subgroups of subjects.

Procedure. The subjects performed the experiment in two stages, one stage for each set of passages. In a given stage the subjects read the instructions for the appropriate set of passages, were given an opportunity to ask questions about the instructions, and then were allowed to perform the task for that set of passages.

The subjects were instructed to read the t passages at their "normal reading speed," but whenever they came to a letter t, they were to encircle it with their pen or pencil. In analogy with the instructions for Experiment 1, the subjects were told that if at any time they realized that they missed a t in a previous word, they should not retrace their steps to encircle it and that they should not slow down their reading speed in order to be overcautious about getting the t's. These instructions were identical to those used by Healy (1976).

The subjects were told that the z passages would each contain a number of spelling errors all of the same type: Each instance of the letter t was replaced by the letter z. They were further told that there were no other zs in the passages. The other instructions for the z passages were analogous to those for the t passages except the letter t was replaced by z.

Results and Discussion

The results are summarized in Table 3, which is analogous to Table 1 except that it includes data from two tasks (detection and substitution proofreading) for two passages (mixed and unmixed). The percentage of errors was greater in the detection (t-circling) task (mean = 15%) than in the substitution-proofreading (z-circling) task (mean = 5%), F(1,95) = 91.0, MS_e = 114, p < .001. Overall, the percentage of errors made in the unmixed passage (mean = 12%) was greater than in the mixed passage (mean = 9%), F(1,95) = 10.8, MS_e = 85, p = .002, and, whereas there was a large difference in error percentages between the two passages for t detection, the difference was smaller and in the opposite direction for proofreading, F(1,95) = 65.9, MS_e = 51, p < .001. A similar pattern of results was found when the percentage of errors on the was considered. The percentage of errors on the in t detection (mean = 25%) was greater than in proofreading (mean = 4%), F(1,95) = 103.7, MS_e = 453, p < .001. Overall the percentage of errors on the in the unmixed passage (mean = 21%) was greater than in the mixed passage (mean = 8%), F(1,95) = 81.2, MS_e = 175, p < .001, and whereas the difference between the two passages in percentages of errors on the was large for t detection, it was smaller and in the opposite direction for proofreading, F(1,95) = 103.8, MS_e = 146, p < .001.
### Table 3

Means and Standard Errors of Means (in Parentheses) for Letter-Detection (L-circling) and Substitution-Proofreading (z-circling) Error Percentages in Experiment 2 as a Function of Passage Type

<table>
<thead>
<tr>
<th>Task</th>
<th>p(error)</th>
<th>p(error on the)</th>
<th>p(error on the</th>
<th>error) percentage</th>
<th>N'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmixed</td>
<td>19.7 (1.5)</td>
<td>38.1 (3.1)</td>
<td>51.9 (2.9)</td>
<td></td>
<td>91</td>
</tr>
<tr>
<td>Mixed</td>
<td>10.7 (1.5)</td>
<td>13.4 (2.2)</td>
<td>26.2 (3.2)</td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>Proofreading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unmixed</td>
<td>3.4 (0.5)</td>
<td>3.4 (0.8)</td>
<td>23.4 (4.9)</td>
<td></td>
<td>53</td>
</tr>
<tr>
<td>Mixed</td>
<td>6.2 (0.9)</td>
<td>3.8 (1.0)</td>
<td>13.2 (3.2)</td>
<td></td>
<td>68</td>
</tr>
</tbody>
</table>

*The value of p(error on the | error) expected by chance alone is 27.5.

Most critical is the pattern of results for the conditional percentages, which, unlike the absolute percentages of errors, should not be affected by any speed-accuracy tradeoff. The conditional percentages were significantly larger for detection (mean = 39%) than for proofreading (mean = 18%), F(1,39) = 66.0, MSE = 630, p < .001, and significantly larger for the unmixed passage (mean = 38%) than for the mixed passage (mean = 20%), F(1,39) = 59.5, MSE = 520, p < .001. In addition there was a significant interaction between these two factors, F(1,39) = 11.3, MSE = 520, p = .002, reflecting the larger difference between the unmixed and mixed passages in detection than in proofreading (although the difference was in the same direction in both tasks). The only passage for which the conditional error percentage was significantly greater than the chance level (27.5%) was the unmixed passage employed in the L-detection task, t(90) = 8.2, p < .001.

It is interesting to note in Table 3 that, for the mixed z passage, the conditional error percentage was actually significantly less than chance, t(67) = 4.8, p < .001. In other words, subjects were less likely to make an error on zhe than on other letter strings containing a z. One possible explanation for this finding is based on the fact that for the mixed z passage there are two factors—mixed typecase and misspellings—each of which should disturb the use of reading units larger than letters. It is reasonable to expect that the combined effect of the two factors would be larger than the effect of each factor individually. For a pure letter scanning (z-scanning) strategy, one might expect a conditional percentage less than chance on the
basis of the finding by Healy (1976, Experiment 1) that the conditional percentage of t-detection errors in the locations was significantly less than chance in a passage of scrambled letters\textsuperscript{5}. This finding can be understood by noting that the t in the word the occurs in the first position of the word, and Corcoran (1966) and Smith and Groat (Note 1) observed that the position of the letter within a word affects letter detection and that early letters are more likely to be detected than later ones.

In any case, the major results are in line with the predictions that subjects would automatically process the word the in units larger than the letter in the unmixed passage but not in the mixed passage, and that subjects could not automatically process the letter string zhe as a single unit, as suggested by the unitization hypotheses. In contrast, these results are not consistent with the redundancy hypothesis.

**EXPERIMENT 3**

According to the unitization hypotheses, the pattern of errors on the in letter detection and substitution proofreading can be attributed to the fact that the is an extremely common word. On the basis of these hypotheses, then, one should see an analogous pattern of results when comparing common and rare words in letter-detection and substitution-proofreading tasks as seen when comparing the and other words. Specifically, subjects should make more errors on common than rare words in letter detection but not in substitution proofreading. We tested this prediction in Experiment 3 by comparing common and rare nouns. Word frequency was controlled in this experiment across the lengths of the words and the locations of the target letter. Each word occurred only once in each of the passages of Experiment 3, so that any differences between common and rare words could not be attributed to different numbers of occurrences of these words in the test passage, a factor that was not controlled in Experiments 1 and 2. Further, in Experiment 3, unlike Experiments 1 and 2, a passage of scrambled nouns was used, so that syntactic and semantic redundancy was virtually eliminated.

Experiment 3 allowed us to test another hypothesis as well, namely the standard explanation of proofreaders' errors: "the common belief that misspellings are more difficult to detect in more familiar words, presumably due to incomplete processing of, or inattention to, orthographic features" (Krueger & Weiss, 1976, p. 204). In a letter-search task with mutilated targets, not unlike the present substitution-proofreading task, Krueger and Weiss (1976) provided support for this hypothesis. In particular, they found that subjects were more likely to detect a mutilated target (created by changing E to F) when it occurred in a nonword than when it occurred in a word. On the basis of these results, we would expect subjects to make more substitution-proofreading errors on common than on rare words in Experiment 3.

**Method**

**Subjects.** The same subjects were employed as in Experiments 1 and 2. The subjects performed Experiment 3 concurrently with Experiment 2 (see below).
Design and materials. The two passages used in Experiment 3 had the same punctuation as those used in Experiment 2; only the words differed. The words employed in the first passage ("i passage") were taken from a list of nouns composed by Paivio, Yuille, and Madigan (1968). The passage included 50 common nouns (AA on the Thorndike-Lorge scale, 1944) and 50 rare nouns (5 or less on the Thorndike-Lorge scale). There were 40 i s in this passage; 20 in common words and 20 in rare words. The words were selected with the following constraint: For every common word, a rare word was chosen that was the same length. Wherever a i, if any, occurred in the common word, a i occurred in the same location in the corresponding rare word. For example, the common word fact was matched with the rare word pact. There were two versions of this passage (versions A and B), which included the same words but differed in the order of the words. Wherever a common word containing a i occurred in one version, its rare mate occurred in the other version. Thus, for example, fact in one version was replaced by pact in the other version, and pact was replaced by fact. The order of the words was otherwise random and the same for both versions. The two versions of this passage were identical to those used by Healy (1975, Experiment 4).

The second passage ("z passage") was the same as the i passage except that every i was replaced by a z. There were no other zs. No letter strings containing z in the z passage formed English words except one which formed a very infrequent word (razing from rating). There were two versions of the z passage, which corresponded to the two versions of the i passage.

Each subject was given mimeographed copies of both passages, typed on separate sheets of paper. The i passage was placed between the two i passages for Experiment 2, and the z passage was placed between the two z passages for Experiment 2. A given subject was shown only one version (A or B) of the passages. The version of the i passage shown to a given subject matched the version of the z passage shown to him or her. This division among subjects was orthogonal to the three divisions of the subjects made for Experiment 2, so that there were approximately equal numbers of subjects (5-7) in the sixteen (2X2X2X2) subgroups of subjects.

Procedure. The procedure was the same as that used in Experiment 2.

Results and Discussion

The results are summarized in Table 4, which includes for both passages the mean and standard error of the mean percentages of errors on the common and rare nouns. More errors were made on common nouns than on rare nouns for i detection, but a small difference in the opposite direction was found for proofreading, for which, overall, errors were less frequent. An analysis of variance conducted on these data revealed that the main effect of task, F(1,95) = 56.0, MS_e = 73, p < .001, the main effect of word frequency, F(1,95) = 5.2, MS_e = 30, p = .023, and the interaction of these two factors, F(1,95) = 7.0, MS_e = 37, p = .009, were all significant.
Table 4
Means and Standard Errors of Means for Letter-Detection (t-circling) and Substitution-Proofreading (g-circling) Error Percentages in Experiment 3 as a Function of Word Frequency

<table>
<thead>
<tr>
<th>Task</th>
<th>M</th>
<th>SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common</td>
<td>14.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Rare</td>
<td>11.3</td>
<td>1.1</td>
</tr>
<tr>
<td>Proofreading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common</td>
<td>6.0</td>
<td>0.8</td>
</tr>
<tr>
<td>Rare</td>
<td>6.4</td>
<td>0.6</td>
</tr>
</tbody>
</table>

This pattern of results is consistent with the unitization hypotheses but cannot be explained by the redundancy hypothesis. Finding fewer errors in proofreading on common than on rare words is also inconsistent with the proposal by Krueger and Weiss (1976) that misspellings are more difficult to detect in more familiar words and with the demonstration by them that mutilated targets were more often missed in a letter-search task when they occurred in words than in nonwords. There were many procedural differences between the present substitution-proofreading task and the letter-search task of Krueger and Weiss (1976). Perhaps the most important difference between the two studies is that the mutilation of the target in the study by Krueger and Weiss (1976) (changing an R to an F) was much smaller than the mutilation of the target in the present study (changing a t to a g). In fact, Krueger and Weiss (1976) proposed that the level of target mutilation may determine whether the mutilated target will be assimilated into the familiar word schema, becoming more difficult to detect, or will be contrasted with the familiar word schema, becoming easier to detect.

**SUMMARY AND CONCLUSIONS**

In summary, the pattern of errors in this study for proofreading was quite different from the pattern for letter detection. Whereas subjects made an inordinate number of errors on the in letter detection, the number of errors on the was no greater than chance in proofreading, and, in fact, was significantly less than chance in Experiment 1. Likewise, whereas subjects made more errors on common than on rare words in letter detection, a small difference in the opposite direction was found in proofreading. These results provide clear evidence that subjects do not skip over or give inadequate attention to the word the in reading prose, thereby refuting the redundancy
hypothesis. In contrast, these results are consistent with the unitization hypotheses put forth by Drewnowski and Healy (1977). In particular, they support the notion that in reading normal prose subjects are able to process automatically common words, especially the most common word the, in units larger than the letter. When the formation of these larger units is disturbed, as it is when every other letter is typed in capitals or when misspellings are introduced, the subjects are more likely to complete the processing of the words at the letter level and, hence, are less likely to make letter-detection errors on the words.

Another, less attractive, explanation is available for the difference between the pattern of results in the comparable substitution-proofreading (g-circling) and letter-detection (t-circling) tasks of Experiments 2 and 3: Whereas subjects may have read the passages for meaning when performing the letter-detection task, subjects may have been able to scan the text for letters (g's), ignoring meaning altogether, when proofreading. (Although a pure letter scan may not be a reasonable strategy in many proofreading situations, it would be reasonable in the particular substitution-proofreading task used in Experiments 2 and 3, since subjects knew that all misspellings involved g.) In the case of a pure letter scan, one would expect to find no differences between common and rare words, as indeed was the case for proofreading. However, three factors argue against such an explanation. First, such a letter-scanning strategy is impossible in the transposition-proofreading task of Experiment 1, which yielded results consistent with those of Experiments 2 and 3. In the task of Experiment 1, in which misspellings consisted of transpositions, the subjects were forced to access the lexicon in order to determine whether a given letter string included a misspelling. Second, the same subjects performed the letter-detection and substitution-proofreading tasks, and the subjects did the two tasks in immediate succession. Since the tasks were superficially strictly analogous, it would seem unlikely that the subjects would employ radically different strategies in the two tasks. In support of this argument, the order in which the two tasks were performed was not found to be a factor that influenced error frequency.7 The third argument against this explanation is that the major portions of the text were identical in the passages for the two tasks. Sixty of the 100 words in the substitution-proofreading task did not contain misspellings so were identical in all respects to the analogous words in the detection task. It seems unreasonable that subjects would process these words in a different way in the two tasks.

Even if an explanation in terms of strategy differences in the two tasks could not be ruled out, the difference between the pattern of results for comparable substitution-proofreading and letter-detection tasks would be of interest. One would still be left with the interesting question: Why were the subjects able to use the more efficient (in terms of numbers of errors) letter-scanning strategy in the substitution-proofreading task but not in the comparable letter-detection task? The only difference between the two tasks (apart from the trivial difference between the identities of the target letters—t versus g) was that the target letters occurred within real words in letter detection but not in substitution proofreading. The question would then be: Why were the subjects able to use the more efficient letter-scanning strategy when the target letters did not occur within real words but not when the target letters did occur within real words? The most plausible answer to
such a question again seems to be in terms of the size of the reading unit available to the subjects. When and only when units are available at the word level may subjects fail to use a pure letter-scanning strategy. Hence, even under the assumption of a strategy difference between tasks, there is support for the central unitization hypothesis.

A possible explanation for the fact that subjects did not make many errors on the in the proofreading tasks of Experiments 1 and 2 is that a large percentage (27.5%) of the misspellings occurred on the in the passages used in these experiments. It could be argued that because of the preponderance of misspellings involving the, subjects gave more attention to that word than they would have otherwise. However, such an explanation could not account for the fact that subjects made a disproportionately large number of errors on the in the t-detection task of Experiment 2, even though an equally large percentage (27.5%) of the ts occurred in the in the passages employed for that task. In addition, this explanation cannot account for the pattern of proofreading errors in Experiment 3, since each word occurred only once in the passages for that experiment. Although a number of ad hoc explanations, like this one, could be constructed to account for a subset of the present results, it appears that only the unitization hypotheses are able to account for the full range of results presented here.

Although this study gives further support to the notion that subjects may read common words in units larger than the letter, the nature of these reading units has not been further clarified by this study. As Healy (1976) has remarked, the units may be perceptual (visual) units or response (phonetic) units. On one hand, the possibility that response units, presumably formed by phonetic recoding, are at issue, rather than visual units, is supported by Corcoran's (1966) study of e detection and subsequent follow-up studies by Mohan (1978), Chen (1976), and Locke (1978), which demonstrated more letter-detection errors on silent than on pronounced letters for normal adult subjects, suggesting that these subjects scan a phonetic representation when searching for a target letter. (However, in a letter-detection experiment by Smith and Groat, Note 1, and in an unpublished experiment by Veneky reported in Hatch, Polin, and Part, 1974, the effect of silent versus pronounced letters was not replicated.) On the other hand, the effects of typecase demonstrated in Experiment 2 of the present study and in Experiment 3 of the study by Drewnowski and Healy (1977) suggest that visual units may be at issue.

Finally, it may be argued that these letter-detection and proofreading tasks have little to do with normal reading for meaning. However, a recent developmental study by Drewnowski (1978) demonstrated that the tendency to make letter-detection errors on the word the was a function of reading level; the pattern of errors on a series of letter-detection tasks provided a good index of the subject's reading ability, presumably because it provided a good index of the size of the reading units employed by the subject. Mohan (1978) has also demonstrated an increasing tendency to make letter-detection errors on the with increases in grade level. These studies demonstrate the relevance of letter-detection tasks to normal reading. If we can understand why subjects make a preponderance of letter-detection errors on the word the, we may indeed advance our understanding of the reading process.
REFERENCES


REFERENCES


Holbrook, M. B. Effect of subjective interletter similarity, perceived word similarity, and contextual variables on the recognition of letter substitutions in a proofreading task. *Perceptual and Motor Skills, 1978, 47*, 251-258. (a)


The fact that words were created by these transpositions seems inconsequential, since none of the 96 subjects made an error on the word *fro* and only three made errors on the word *eh*.

Word length was confounded with two potentially important variables: word frequency and the location of the transposed letters in the word. The mean frequency (Kučera & Francis, 1967) of the misspelled words excluding *the* monotonically decreased as a function of word length. Mean frequency per 1,014,232 words of text was 18,972, 3,820, 1,625, 70, and 69 for the misspelled words of length two, three, four, five, and six to ten letters, respectively. However, word frequency cannot account for finding a greater mean percentage of errors on the most frequent word *the* (frequency = 69,971) than on other misspelled three-letter words or finding a greater mean percentage of errors on two-letter words than on three-letter words (*p* < .01 by a Newman-Keuls test).

The location of the transposed letters in a word may be critical, since the transpositions necessarily involved an end letter (either initial or terminal) in all the two- and three-letter words but involved only intermediate letters in three of the six four-letter words, two of the four five-letter words, and seven of the nine words six to ten letters long. For the four- and five-letter words, more errors were made on the five words with transpositions involving only intermediate letters (mean = 21.5%) than on the five words with transpositions of an end letter (mean = 4.6%), *F*(1,95) = 97.4, *MSe* = 140, *p* < .001. However, other factors must also be critical, since the percentage of errors on the long word *separate* when misspelled as *esparate* was quite high (21.9%), although it involved a transposition of the initial letter.

The difference between errors percentages on the word *the* and other three-letter words is due in part to the relatively large percentage of errors (14.6%) made on a single instance of the word *the* (the only instance involving a capital letter: *The* was misspelled as *Hte*). Excluding that instance, the mean percentage of errors on *the* was greatly reduced (mean = 2.2%). The especially low percentage of errors on other three-letter words (only two errors out of 384 opportunities across subjects) may be due to peculiar aspects of the particular misspellings employed: *for* was misspelled as *fro*, *let* as *lte*, *low* as *olw*, and *two* as *wto*. However, with the exclusion of the one instance of *the* described above, the difference between errors made on the misspelled as *hte* (mean = 2.3%) and *teh* (mean = 2.1%) was not significant, *F*(1,95) < 1.

In the analysis of variance, there were missing cells for the cases when a subject made no errors in a given task on a given passage so that a conditional percentage could not be computed. Each of these missing cells was replaced by the appropriate mean conditional percentage. A conservative estimate of the degrees of freedom was computed by subtracting the number of
subjects (56) who contributed one or more missing cells.

In Healy's (1976) experiment the scrambled letter passage was derived from a prose passage by retaining the punctuation, word boundaries, and locations of the 's in the prose passage but scrambling all the remaining letters. A t in a the location in the scrambled letter passage was a t in a location where the word the occurred in the corresponding prose passage.

The fact that a word was created in this case seems relatively inconsequential, since only 6 of the 96 subjects made an error on the word razing.

Unweighted analyses of variance for unbalanced designs (unequal cell frequencies) were performed on the total error scores in Experiments 2 and 3. The factor of test order (proofreading first versus letter detection first) was not found to be a significant main effect or to enter into any significant interactions in these analyses (except for a significant, p = .040, five-way interaction in Experiment 3, involving the four between-subjects factors and the factor of task.)