Abstract. Four experiments in which subjects searched for the letter n yielded many detection errors on the suffix morpheme -ing. The subjects made more errors on word endings than in the middle of the words and more errors on -ing than on -en, -ion, or -ment endings. More errors were made when the -ing ending was a suffix morpheme than when it formed part of the word stem. These results suggest that the combination of the high orthographic frequency of the letter sequence -ing, its spatial predictability, and its function as a suffix morpheme lead it to serve as a single unit in reading.

In previous studies we have shown that subjects searching for instances of a given target letter in printed text make most detection errors on the very common function words the and and (Drewnowski & Healy, 1977; Healy, 1976). The subjects' tendency to make errors on the word the seems to be a function of reading skill, since the pattern of errors made by school children was directly related to their reading ability (Drewnowski, 1978). The present letter detection paradigm (Corcoran, 1966; Healy, 1976) may therefore reflect reading processes better than detection tasks used by other investigators (e.g., Johnson, 1975; Wheeler, 1970), which do not involve searching for targets in paragraphs of text. We now use the letter detection paradigm to examine the processes involved in reading multisyllabic words. Specifically, we focus on the pattern of errors that are made on frequent word endings. It has been suggested that frequent word endings and common function words can both serve as word boundary morphemes (Kean, 1977). Our hypothesis is that this linguistic similarity will be reflected in the reading task.

We have previously interpreted the disproportionate number of letter-detection errors on the most common words in terms of reading units of different sizes. Healy (1976, 1980) demonstrated more errors on common than on rare nouns and therefore proposed that the "missing letter effect" was caused by the subjects' tendency to process common words in units larger than the letter. Furthermore, the finding of more errors on function words

*Also in Journal of Verbal Learning and Verbal Behavior, 1980, 19, 247-262.
+University of Toronto
++Also Yale University

Acknowledgment. This research was supported by Grant APA 146 from the National Research Council of Canada, NICHD Grant HD01994, BRS Grant RR05596, NSF Grant BNS77-00077, and a grant from the Spencer Foundation. We are indebted to J. Chung for help with the data analyses of Experiment 4 and to W. K. Estes, T. K. Landauer, and A. G. Levitt for helpful discussions about this research.
occurring in familiar contexts than in inappropriate contexts may indicate
that subjects process very familiar phrases in units that include more than
one word (Drewnowski & Healy, 1977). The postulated reading units may be
either perceptual (visual) or response (phonetic) units. The available data
do not allow us to choose between these two possibilities. However, the data
do provide evidence against alternative explanations of these results (see
e.g., Corcoran, 1966), which refer to the syntactic and semantic redundancy of
the words containing the target letters or the atypical pronunciation of the
target letters (see Healy, 1976, 1980).

We now investigate the formation of reading units larger than the letter
but smaller than the word, and we examine the conditions under which a given
sub-word sequence of letters can serve as a single unit in the course of
fluent reading. Whereas previously we were mainly concerned with whole-word
frequencies (Thorndike & Lorge, 1944), we now distinguish between two types of
sub-word frequency that may contribute to the formation of reading units. The
first type, which we shall call orthographic frequency, is suggested by the
work of Gibson (1965) showing that letter sequences forming frequent spelling
patterns may in themselves function as higher order reading units. Central to
this notion is the assumption that through repeated exposure to printed words,
readers acquire in memory a representation of the type of letter sequences
that can be expected to occur in English texts. It seems reasonable to assume
that the most common letter sequences in English words will be most likely to
be unitized and will thus most successfully "conceal" the identity of their
constituent letters. As a result, we might expect most letter detection
errors to occur on letter sequences that are orthographically most frequent.

The second type of frequency, which we shall call positional frequency,
is suggested by the work of Mason (1975) and Katz (1977) showing that skilled
readers are sensitive to the positional frequencies of single letters within
test words. Extending this notion from single letters to letter sequences, we
might expect that most letter detection errors will occur in letter sequences
appearing in their most spatially redundant locations within the word.

EXPERIMENT 1

In order to discriminate between the effects of orthographic and posi­tional
frequency, we examined the incidence of detection errors on the target
letter n embedded in two different word endings: -ing and -ent. Whole-word
frequency, as established by Kučera and Francis (1967), was kept approximately
constant.

The two letter sequences -ing and -ent were selected following a
comparison of the Underwood and Schulz (1960) and the Mayzner, Tresselt, and
Wolin (1965) frequency tables. Underwood and Schulz (1960) tabulated digram
and trigram counts based on a random sample from the Thorndike and Lorge
(1944) word list and on two additional samples of running text. According to
their data, the embedded frequency of the trigram -ent (1778) exceeds that of
the trigram -ing (1673).

Although this analysis provides a means of comparing the orthographic
frequency of letter sequences, it does not take into account the position of
the letter sequences within the word. Such data are supplied by Mayzner et al. (1965), who tabulated trigram counts by word position and by word length for words from 3 to 7 characters in length. The trigram -ing was found to occur almost exclusively at the end of words, whereas the trigram -ent was less spatially predictable, occurring at a number of locations within the word.

The trigrams -ing and -ent thus differ in their orthographic and positional frequencies. If frequent letter sequences are perceived as higher order units (Gibson, 1965), then we might expect more letter detection errors to occur on words with -ent endings than with -ing endings. On the other hand, if positional frequency (Mason, 1975) is the critical factor in unit formation, then we might expect more errors to occur on the more spatially predictable -ing endings than on the -ent endings. The present design rules out single letter location or the word length in syllables as potential contributing variables, because the target letter n always occurred in the penultimate position within the word, and the length in syllables of test words was equated across both types of endings.

We further attempted to influence the number of letter detection errors by altering the nature of the search passage. In addition to a standard case prose passage, we employed a mixed case prose passage, in which alternating letters of the prose text were typed in upper and lower case. We also employed a meaningless passage of scrambled text. Such manipulations, designed to impede the formation of reading units larger than the letter or larger than the word, reduced the percentage of letter detection errors on the function words the and in our previous studies (Drewnowski & Healy, 1977; Healy, 1980). We now wish to determine whether the unitization of sub-word letter sequences would also be affected by such procedures.

Method

Subjects. Fifty-four male and female students at the University of Toronto served as volunteer subjects in a group experiment conducted in the classroom.

Design and materials. Three 175-word passages, typed on separate sheets of paper, were constructed for the present experiment. Each passage included 42 "test" words (words that contain the target letter) and 133 "filler" words (words that do not contain any targets). The first passage, hereafter referred to as the "prose standard case passage," contained 18 test content words ending in -ing, 18 test content words ending in -ent, and 3 instances each of the function words in and and. The nature of the test content words was manipulated further by using 12 mono-, 12 di-, and 12 trisyllabic words, with half the words in each subgroup ending in -ing, and half ending in -ent. The target letter n thus always occurred in a function word or as part of a word ending.

In the case of monosyllabic words, both -ing and -ent trigrams necessarily formed part of the word stem (e.g., wing, went). In contrast, in trisyllabic words, -ing trigrams were always inflectional suffixes (e.g., applying) and -ent trigrams were most often derivational suffixes (e.g., persistent). The difference in nature of the test trigrams (word stem or
suffix) was systematically manipulated for disyllabic test words ending in -ing, with three words including -ing as part of the word stem (during, something, sterling) and three including -ing as a suffix (having, being, going). In addition, word stress was systematically manipulated for disyllabic words ending in -ent: Three of these took primary stress on the first syllable (recent, present, payment), and three took stress on the second syllable (resent, percent, torment). Since whole-word frequency was found to affect the incidence of letter detection errors (Healy, 1976, Experiment 4), we matched word frequencies (Kucera & Francis, 1967) across conditions wherever possible: The mean frequency of -ing test words was higher (181.1) than the mean frequency of -ent test words (122.7). However, trisyllabic words ending in -ing were specifically selected so that their mean frequency of usage (30.0) would be lower than the mean frequency of usage of trisyllabic words ending in -ent (93.3).

The second passage, hereafter referred to as the "prose mixed case passage," was identical to the prose standard case passage, except that alternating letters were typed in upper and lower case. Only one version of the prose mixed case passage was prepared, so the incidence of upper and lower case ns was randomized across targets.

The third passage, hereafter referred to as the "scrambled-word passage," was derived from the prose passages. The 36 test words and the three instances each of the function words in and and remained in the same locations as in the prose passages, but the order of the 133 filler words was randomized. No particular effort was made to place every test word in an inappropriate syntactic context: In fact, wherever two test words occurred in adjacent positions in the prose passages (e.g., "resent having" or "and rent"), their locations were preserved in the scrambled-word passage.2

Each passage was typed on a separate sheet of paper. The three passages arranged in all six possible sequences and preceded by a page of instructions were stapled together into booklets. The booklets were distributed according to a fixed rotation so that passage order was approximately counterbalanced across subjects.

Procedure. The subjects were instructed to read each passage at their normal reading speed, and to circle each instance of the target letter n. The subjects were told that if they ever realized that they had missed a target, they should not retrace their steps to encircle it. They were also told that they were not expected to detect all the ns so they should not slow down their reading speed in order to be overcautious about encircling the ns. The subjects were told to read the passages in the order in which they were stapled together and to go on to the next passage as soon as they had finished the preceding one.

Results

The results are summarized in Table 1, which includes for the three passages means and standard errors of the means for error percentages on the function words in and and and error percentages on the content words ending in -ing or -ent.

130
Table 1

Means (and Standard Errors) for Error Percentages in Experiment 1
as a Function of Passage Type and Word Type

<table>
<thead>
<tr>
<th>Passage type</th>
<th>Function</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>in</td>
<td>and</td>
</tr>
<tr>
<td>Prose standard case</td>
<td>77.0 (4.3)</td>
<td>80.7 (3.7)</td>
</tr>
<tr>
<td>Prose mixed case</td>
<td>78.3 (2.5)</td>
<td>53.7 (5.0)</td>
</tr>
<tr>
<td>Scrambled-word</td>
<td>72.0 (4.3)</td>
<td>75.6 (4.3)</td>
</tr>
</tbody>
</table>

Function words. Higher error percentages were obtained for the two function words in and and than for content words ending in -ing or -ent. The present data are thus consistent with our earlier reports (Healy, 1976, 1980; Drewnowski & Healy, 1977) that a disproportionate number of letter-detection errors occurs on frequent function words. In addition, the data preclude a purely phonological interpretation of the phenomenon, such as has been suggested by Corcoran (1966) and Locke (1978): Let us suppose that subjects are more likely to detect the letter n in words that contain the sound /n/, since this is the most common pronunciation of the letter n. One might then expect that an n in the sequence -ing (pronounced as /n/) would be more difficult to detect than an n in a sequence that contains the conventional /n/ sound, such as in and and. However, considerably more errors were made on in and and than on the words containing -ing.

There were no differences in the percentage of errors on function words between the prose standard case and the scrambled-word passages [t(106) = 1.01, p > .10], possibly because the appropriateness of between-word context was not strictly controlled in the scrambled-word passage. Only one version of the prose mixed case passage was used such that all three instances of and appeared as and and all three instances of in appeared as In. Only the error frequency on the word and [t(106) = 4.35, p < .01] was reduced for the prose mixed case passage relative to the prose standard case passage. We interpret this finding to indicate that the word In was still perceived as a single unit by the readers, whereas the word and no longer was.

Content words. Table 1 also shows that more errors were made on words ending in -ing than on words ending in -ent, and that more errors were made on the prose standard case and scrambled-word passages than on the prose mixed case passage. An analysis of variance of error scores on content words showed
Figure 1. Error percentages for words ending in -ing or -ent in Experiment 1 as a function of the passage type and the length in syllables of the test words.
significant main effects of the type of word ending \( F(1,53) = 50.76, p < .01 \), and the type of search passage \( F(2,106) = 14.77, p < .01 \). The subjects' performance was similar on prose standard case and scrambled-word passages and was uniformly more accurate on the prose mixed case passage, particularly for -ent endings. The interaction of word ending by passage type \( F(2,106) = 4.72, p < .05 \) was significant.

A more detailed analysis of the data is presented in Figure 1, which shows error percentages for words ending in -ing or -ent at each syllabic length and with passage type as parameter. Error frequencies increased together with the length in syllables of content words ending in -ing, but no such effect was obtained for words ending in -ent. The main effect of the length in syllables \( F(2,106) = 28.22, p < .01 \) and the interaction between the type of ending and syllabic length \( F(2,106) = 22.56, p < .01 \) were both significant. The effect of syllabic length was greatest between mono- and disyllabic words, suggesting that the effects of syllabic length might reflect differences in the role of the test trigram. Whereas -ing (and -ent) trigrams in monosyllabic words necessarily formed part of the word stem, in di- and trisyllabic words, they mostly functioned as suffixes.

The incidence of errors on disyllabic test words ending in -ing was found to depend on whether the -ing trigram formed part of the word stem (e.g., during) or was a suffix morpheme (e.g., having). Data in Table 2 indicate that more letter detection errors were made when the -ing trigram was a suffix than when it was part of the word stem for the prose standard case \( t(53) = 2.54, p < .05 \) and prose mixed case passages \( t(53) = 2.13, p < .05 \). This difference was in the same direction but was not significant for the scrambled-word passage \( t(53) = 0.11, p > .10 \).

### Table 2

Means (and Standard Errors) for Error Percentages in Disyllabic -ing Content Words in Experiment 1 as a Function of Passage Type and Word Ending

<table>
<thead>
<tr>
<th>Word ending</th>
<th>Passage type</th>
<th>stem</th>
<th>suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prose standard case</td>
<td>32.6(5.0)</td>
<td>45.7(4.7)</td>
</tr>
<tr>
<td></td>
<td>Prose mixed case</td>
<td>21.3(4.0)</td>
<td>30.0(4.0)</td>
</tr>
<tr>
<td></td>
<td>Scrambled-word</td>
<td>43.7(5.0)</td>
<td>44.3(5.0)</td>
</tr>
</tbody>
</table>

The location of stress did not affect the incidence of errors on words ending in -ent. Instances of the target n in disyllabic words carrying stress on the first syllable (e.g., resent) were missed at the same rate as letters in words carrying stress on the second syllable (e.g., resent) for the prose
standard case \[ t(53) = 0.98, p > .10 \], prose mixed case \[ t(53) = 1.52, p > .10 \], and scrambled-word \[ t(53) = -0.77, p > .10 \] passages. The high incidence of errors on the -\textit{ing} endings thus cannot be attributed to the fact that they never receive the primary word stress in English, since the incidence of errors on -\textit{ent} endings, stressed or unstressed, was still appreciably lower than that on -\textit{ing} endings.

Discussion

Substantially more letter detection errors were made on content words with -\textit{ing} endings than with -\textit{ent} endings. These data fail to support the notion that the high percentage of errors on -\textit{ing} endings is solely due to the high orthographic frequency of the -\textit{ing} sequence, since the trigram -\textit{ent} is reported to have an even higher embedded frequency in English words than the trigram -\textit{ing} (Underwood & Schulz, 1960). The data also fail to support the notion that the observed effects are solely due to the positional frequency of the target letter within the word, since the letter n always occurred in the penultimate location in both types of test words. Instead, the data suggest that the more important factor may be the positional frequency in the language of the entire trigram. Whereas the letter sequence -\textit{ent} can occur in English at a number of locations within the word, the sequence -\textit{ing} is found most often at word endings. The data are thus consistent with the hypothesis that the positional frequency of the letter sequence may be more critical than its absolute orthographic frequency.

Another important factor seems to be the linguistic function of the letter sequence: An analysis of errors made on disyllabic words showed that more errors were made when -\textit{ing} endings were morpheme suffixes than when they formed part of the word stem. This finding also argues against a purely phonological interpretation of our results, since the n in the letter sequence -\textit{ing} is pronounced in an atypical manner in both cases.

The subjects made fewer errors on the prose mixed case than on the prose standard case passage. Case mixing reduced but did not entirely eliminate the difference in errors between -\textit{ing} and -\textit{ent} endings and between suffix and word-stem groups. These data suggest that although purely visual manipulations influence error rates, visual factors alone cannot account for the observed pattern of results. Although error rates for -\textit{inG} and -\textit{InG} were reduced relative to -\textit{ing}, they were still above those for -\textit{ent}.

There were no differences in error percentages on content words between the prose standard case and scrambled-word passages. This finding, which is consistent with our earlier reports (Healy, 1976; Drewnowski & Healy, 1977), suggests that phrases or word frames may be the largest units employed in fluent reading and that the presence of clauses or sentences has no further effect on the letter detection task. Since short-range syntactic/semantic constraints of text in the scrambled-word passage made it equivalent to the prose standard case passage in terms of the subjects' performance, we limited ourselves to scrambled-word passages in the subsequent experiments of this report.
EXPERIMENT 2

Data from Experiment 1 suggest that the positional frequency of the test trigram is critical in determining the percentage of letter detection errors. To test this hypothesis we studied variations in the location of the trigram within the word: Instances of n embedded in -ing or -ent trigrams now occurred in the middle or at the end of di- and trisyllabic content words. If the positional frequency of the trigram is the critical variable, then we would expect a large difference in error percentages with location for the more spatially predictable -ing trigrams and a smaller difference for -ent trigrams.

We also examined the possible effects of the syntactic/semantic structure of the search passage on the letter detection task. Because the context of the test words was not systematically varied in Experiment 1, the data did not rule out the possibility that the occurrence of -ing suffixes could be predicted on the basis of the words surrounding the test content words. It may be argued that the combination of surrounding context and test word stem allows the subject to anticipate and therefore skip over the word ending. We therefore presented the subjects with two passages of nonsense text differing in the appropriateness of the context surrounding the test content words (see Drewnowski & Healy, 1977, Experiment 2, for a similar manipulation using function words). Each passage was typed either in standard typecase or with alternating letters typed in upper and lower case. As in Experiment 1, this last manipulation was intended to impede the formation of reading units larger than letters using purely perceptual means.

Method

Subjects. Ninety-six male and female students at the University of Toronto served as volunteer subjects in a classroom setting. There were 53 subjects in the Standard Case condition and 43 different subjects in the Mixed Case condition. The two conditions were conducted in separate sessions.

Design and materials. Two 272-word passages of nonsense text were constructed. Each passage included 12 di- and 12 trisyllabic test words containing the trigram -ing, and 12 di- and 12 trisyllabic test words containing the trigram -ent. The location of the test trigram within the word was also systematically varied. In one half of the test words of each type and length, the -ing and -ent trigrams occurred at the end of the word, whereas in the remaining half of the test words, the -ing and -ent trigrams occurred in the middle of the word, either in the first or in the second syllable.

For di- and trisyllabic words ending in -ing, the trigram -ing was always an inflectional suffix (working, spelling, walking, hearing, falling, reading, performing, competing, amusing, exciting, amazing, visiting). In contrast, the trigram ending -ent either formed part of the word stem or occurred as a derivational suffix (recent, current, frequent, garment, absent, basement competent, evident, vehement, excellent, resident, different). Selection of the test words was further constrained in such a way that the mean frequency of usage (Kucera & Francis, 1967) of test words containing -ing trigrams (28.0) was lower than that of words containing -ent trigrams (53.8).
The nature and order of filler words surrounding each test word were also systematically varied. Each test word was the central word of a three-word meaningful phrase (e.g., late visiting hours, the garment district). In addition, 24 filler phrases that were outwardly similar in construction to the test phrases (e.g., pure carrot juice, this chocolate cake) were randomly interspersed throughout the text. Each 272-word passage thus contained 48 three-word phrases containing test words, 24 three-word filler phrases, three instances each of the function words and and in, as well as 50 of the most frequent function words selected from the Kučera and Francis list with the restriction that no word include the letter n. The letter n thus always occurred in each passage in a function word in or and, or as part of an -ing or -ent string.

In the "local context passage" (cf. Drewnowski & Healy, 1977, Experiment 2), the words in each of the test phrases occurred in their correct order (late visiting hours), but the words in each of the filler phrases were reversed (juice carrot pure). Punctuation marks (periods) were inserted to make the text appear more like prose.

In the "no context passage," the words in each of the test phrases were reversed (hours visiting late), whereas the words in the filler phrases occurred in their correct order (pure carrot juice). This manipulation of the order of words within filler phrases was intended to bring both passages to roughly the same order of approximation to English. Punctuation and the location of the test words on the printed page remained the same for both passages. The context of the function words in and and within each passage was not systematically varied.

The two passages, local context and no context, were typed either in the standard fashion (Standard Case condition), or with alternating letters typed in upper and lower case (Mixed Case condition). There were two versions of each passage in the Mixed Case condition: Version A had all even letters in upper case and version B had all odd letters in upper case, so that across versions all instances of the target letter n occurred an equal number of times in lower and upper cases.

The two passages, presented in counterbalanced order and preceded by a page of instructions to subjects, were stapled together as a booklet. Approximately half the booklets in the Mixed Case condition contained version A of the two passages and the remaining booklets contained version B.

Procedure. Instructions to subjects and details of experimental procedure were the same as those described for Experiment 1.

Results

The data are summarized in Table 3, which includes means and standard errors of the means for error percentages for each type of test word and under each set of experimental conditions.
Table 3

Means (and Standard Errors) for Error Percentages in Experiment 2 as a Function of Condition, Passage Type, and Word Type

<table>
<thead>
<tr>
<th>Condition</th>
<th>Word type</th>
<th>Function</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>in</td>
<td>and</td>
<td>ing</td>
</tr>
<tr>
<td>Standard Case</td>
<td>Local context</td>
<td>64.7 (3.0)</td>
<td>59.0 (5.0)</td>
</tr>
<tr>
<td></td>
<td>No context</td>
<td>69.7 (4.3)</td>
<td>72.3 (3.7)</td>
</tr>
<tr>
<td>Mixed Case</td>
<td>Local context</td>
<td>54.0 (4.7)</td>
<td>26.3 (2.7)</td>
</tr>
<tr>
<td></td>
<td>No context</td>
<td>56.3 (5.3)</td>
<td>35.0 (4.0)</td>
</tr>
</tbody>
</table>

Function words. More errors were made on function words in and and than on content words, in agreement with previous results (e.g., Drewnowski & Healy, 1977). Both function words were more likely to be detected when in the Mixed Case than in the Standard Case condition, and this improvement was much greater for the word and than for the word in. The main effect of test word type (in or and) was significant in the Mixed Case [F(1,42) = 48.35, p < .01], but not in the Standard Case [F(1,52) < 1] condition. A more detailed treatment of this phenomenon is presented in Experiment 4.

Content words. More errors were made on test words containing -ing trigrams than on words containing -ent trigrams both in the Standard Case [F(1,52) = 60.05, p < .01] and in the Mixed Case [F(1,42) = 48.71, p < .01] conditions. In contrast, and contrary to our hypothesis about the possible effects of syntactic/semantic contexts on the detection tasks, the effects of passage context [Standard Case: F(1,52) < 1; Mixed Case: F(1,42) = 2.52, p > .10] were not significant. Furthermore, the effect of the length of the test words in syllables [Standard Case: F(1,52) = 1.97, p > .10; Mixed Case: F(1,42) = 3.84, p > .05] was not significant. More errors were made when the two trigrams occurred at the end rather than in the middle of the test words. The main effect of target location was significant for both the Standard Case [F(1,52) = 97.13, p < .01] and the Mixed Case [F(1,42) = 42.55, p < .01] conditions.

The data, averaged over passage context and test word length in syllables, are shown in Table 4 to illustrate the nature of the interaction of trigram type by location, which was significant both for the Standard Case [F(1,52) = 61.70, p < .01] and for the Mixed Case [F(1,42) = 28.18, p < .01] conditions. Whereas the percentages of letter detection errors on -ing and on -ent trigrams were approximately equal when the trigrams occurred in the
middle of the test word, significantly more errors were made on -ing trigrams than on -ent trigrams when the trigrams occurred as word endings. The same interaction held for both the Standard Case and Mixed Case conditions, although overall error percentages in the Mixed Case condition were greatly reduced.

<table>
<thead>
<tr>
<th>Test trigram</th>
<th>Condition</th>
<th>Middle</th>
<th>End</th>
<th>Middle</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard Case</td>
<td>15.9</td>
<td>50.7</td>
<td>10.4</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>Mixed Case</td>
<td>11.4</td>
<td>30.2</td>
<td>6.1</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Discussion

The present finding that more detection errors were made in later rather than in earlier target locations within the word replicates the location effects observed by Corcoran (1966). This effect was small when the target letter was part of the letter sequence -ent and much greater when it was part of the sequence -ing. This finding of a strong interaction between the type of the test trigram and its location within the word indicates that the orthographic, or linguistic, context of the target letter must also be a contributing factor. Because the location of the target letter within the word was controlled across both types of test trigrams, we are able to rule out the possibility that single-letter positional frequency is the basis for the subjects' responses (Katz, 1977; Mason, 1975). Rather, our results point to another kind of frequency mediation—based not so much on orthographic frequency of letter sequences, but instead on their positional frequency. In other words, our results suggest that frequent letter sequences may act as reading units only if they occur in their most common locations within the word.

Although the within-word context of the target letter was found to be of importance, the present data show no trace of between-word context effects as determined by the syntactic/semantic structure of the search passage. These data qualify the notion that the combination of surrounding context and test word stem allows the subject to pass over the word ending. Because manipulations of passage context in our previous studies influenced the pattern of
errors on the function words the and and (Drewnowski & Healy, 1977), we were led to postulate that very frequent function words are processed in terms of reading units larger than the word itself. The present data do not permit a similar conclusion for the less frequent content words, since equivalent results were obtained for the local context and no context passages. Instead, it may be that in the case of multisyllabic test words, the reading unit is larger than the letter but smaller than the word. Consequently, we need to examine more closely the possibility that the -ing suffix morpheme itself serves as a processing unit in reading multisyllabic words.

EXPERIMENT 3

In Experiment 3 we asked one group of subjects to search for the letter n and another group of subjects to search for the entire -ing trigram embedded in test words. We wished to examine whether word endings per se are harder to identify than other parts of words, or whether component letters in word endings are "hidden" when the endings are read in terms of units larger than the letter. If word endings are harder to identify, then we should expect more errors on word endings than on mid-word locations for both types of search targets. In contrast, if the suffix morpheme serves as a reading unit, then we should expect more errors on word endings than on mid-word locations only for the search target n, not for the search target -ing. Only in the former case would the target be smaller than the reading unit employed; in the latter case, the reading unit and the to-be-detected target would be identical.

Method

Subjects. One hundred and sixty students at the University of Toronto served as volunteer subjects in the present experiment, which was conducted in the classroom. There were 80 subjects in the Letter Detection condition and 80 subjects in the Trigram Detection condition.

Design and materials. One 272-word passage of nonsense text, derived from the local context passage of Experiment 2, was used. The three instances each of function words in and and as well as the 24 -ent test words had been removed and replaced by words of equal length and, where possible, of equivalent frequency that did not contain the letter n. The punctuation and the location of the test words on the page remained the same as in Experiment 2, though the overall target density (24 targets per 272 words) was necessarily reduced. All the remaining instances of the letter n were part of the -ing trigram.

The present passage was then typed either in the standard fashion (standard case passage) or with alternating letters typed in upper and lower case (mixed case passage) in the same manner as the passages in Experiment 2.

Procedure. Each subject received both the standard and mixed case passages presented in a counterbalanced order. Details of experimental procedure otherwise resembled those of Experiment 2, except that half the subjects were instructed to search for the letter n and half the subjects were instructed to search for the trigram -ing.
Results

The data are summarized in Table 5, which shows error percentages for each detection task at each target location within the word and with passage type as parameter. These data have been pooled over the syllabic length of the test words.

Table 5

Means (and Standard Errors) for Error Percentages in Experiment 3 as a Function of Search Task, Passage Type, and Target Location

<table>
<thead>
<tr>
<th>Search task</th>
<th>Middle</th>
<th>End</th>
<th>Middle</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passage type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard case</td>
<td>16.8 (2.0)</td>
<td>53.5 (3.8)</td>
<td>12.9 (2.0)</td>
<td>9.8 (1.1)</td>
</tr>
<tr>
<td>Mixed case</td>
<td>14.5 (1.8)</td>
<td>41.5 (3.4)</td>
<td>11.5 (2.2)</td>
<td>7.6 (1.1)</td>
</tr>
</tbody>
</table>

The subjects made significantly more errors on the letter detection than on the trigram detection task \([F(1,158) = 69.19, p < .01]\). Subjects searching for the letter \(n\) made significantly more errors at end locations than at middle locations (in agreement with the data of Experiment 2), whereas subjects searching for -ing trigrams made somewhat fewer errors at end locations. Both the main effect of location \([F(1,158) = 73.53, p < .01]\) and the critical interaction between search task and location \([F(1,158) = 114.13, p < .01]\) were significant.

The subjects made more errors on standard case than on mixed case passages \([F(1,158) = 14.88, p < .01]\). This difference was greater in the letter than in the trigram detection task, particularly for end targets. The interaction between typeface and location \([F(1,158) = 9.39, p < .01]\) as well as the second order interaction between the search task, typeface, and location \([F(1,158) = 6.97, p < .01]\) were significant.

Discussion

These data show that most letter detection errors were made on normally typed -ing endings by subjects searching for the target letter \(n\). In contrast, subjects instructed to search for the entire -ing trigram were much more accurate in their performance and were influenced neither by target location within the word nor by the passage typeface. This pattern of results differs from that obtained previously with common function words (Drewnowski &
Healy, 1977), in which subjects made a disproportionate number of errors on the word and whether they were instructed to search for the letter n or the entire word and.

These results point to an important difference in processing content and function words. We have interpreted our previous data on function words by proposing that detection errors resulted whenever the size of the target was smaller than the size of the postulated reading unit. Thus a disproportionate number of letter detection errors on the word and implied that it was processed in terms of units larger than the letter. Similarly, the high percentage of detection errors when the entire word and was the target implicated the use of reading units larger than the word itself (Drewnowski & Healy, 1977). Following this reasoning, we might expect a sharp drop in the percentage of detection errors whenever the size of the target is equal to or is larger than the size of the postulated reading unit.

We therefore interpret the present data by proposing that -ing endings can serve as single units in the course of fluent reading. Most errors were made when the size of the postulated reading unit exceeded that of the to-be-detected target. This occurred for letter targets when the -ing trigram was in its most frequent location within the word. Fewer errors were made on mixed case passages in which the formation of -ing reading units was to some extent impeded. Few errors were also made for the target -ing, which either corresponded in size to our postulated reading unit, or exceeded the size of the available reading unit, as it did in mixed case passages.

An additional factor that may play a role when the subjects search for -ing targets is that subjects may scan word endings more thoroughly, since -ing occurs predominantly at the end of words in English. Whereas Experiment 2 suggested that subjects are, if anything, less likely to process the final -ing letter by letter when searching for ns, they might look more carefully at the word endings when searching for -ing, since -ing is a unit expected to occur at the end of words. This hypothesis implies that the subjects are sensitive to the position of letter sequences within words, in the manner postulated for single letters by Mason (1975) and explored for letter trigrams in Experiments 1 and 2.

**EXPERIMENT 4**

We have now shown that subjects make many detection errors on the letter n when it occurs as part of an -ing ending. Because such a pattern of errors is not found for -ent endings, we proposed that the critical factor determining the subjects' performance is the positional frequency of the letter sequence. However, there may be an alternative linguistic explanation for this result. First, the test words in Experiments 1 and 2 were such that -ing endings generally served as suffix morphemes, whereas this was not always the case with -ent endings. Second, -ing suffixes were always inflectional, whereas -ent suffixes were derivational (Jarvella & Snodgrass, 1974). Thus, -ing suffix test trigrams occurred in inflected forms of verbs, whereas -ent suffix test trigrams occurred in derived adjectives and nouns. Third, the inflectional and derivational suffixes were paired with different word stems.
We now wish to examine the importance of each of these factors. We therefore consider several suffix morphemes other than -ing to determine whether they also can function as perceptual units in reading. Specifically, we compare inflectional and derivational suffixes attached to the same word stems.

Work in linguistics has suggested that bound morphemes vary in their degree of attachment to the word stem and that derivational suffixes are more tightly bound than inflectional suffixes (Kean, 1977). It seems reasonable to postulate that inflectional suffix morphemes are more likely to function as single linguistic entities and hence are more likely to be unitized during reading. We therefore compare error percentages on inflectional suffixes, such as -ing and -en, to those on derivational suffixes, such as -ment and -ion. In each case, test words ending in -ing contained the same word stems as test words with each of the other three endings.

Method

Subjects. One hundred and sixteen male and female Yale undergraduates participated as volunteer subjects in a group experiment conducted in the classroom.

Design and materials. Two 100-word scrambled-word passages were constructed. The first passage, which will hereafter be referred to as the "standard case passage," included six words ending in -ment (e.g., development) and the corresponding six words (the six words with the same stem) ending in -ing (e.g., developing), six words ending in -en (e.g., taken) and the corresponding six words ending in -ing (e.g., taking), six words ending in -ion (e.g., suggestion) and the corresponding six words ending in -ing (e.g., suggesting). Each of the two words with the same stem differed only in the suffix. We therefore have three word-stem groups: -ment, -en, and -ion. Note that two of the endings employed (-ment and -ion) were derivational endings and two (-ing and -en) were inflectional endings. Also note that for two of the endings (-en and -ion), the target letter n was the final letter and for the other two endings (-ing and -ment), the target letter n was the penultimate letter. In addition to these 36 words containing target letters, each passage contained three instances of the word and three instances of the word in.

For each of the 42 test words, a word identical in length and comparable in frequency, according to Kučera and Francis (1967), but without the letter n was selected for inclusion in the passage as a filler. The mean frequency of the six test words containing a suffix of a given type was matched exactly to the mean frequency of the six analogous filler words. Similarly, there were three instances of the filler function word was and three of the filler function word of, which were matched to the instances of the test function words and and in, respectively. These constraints were instituted to insure that subjects could not determine whether a word contained a target letter on the basis of its length, its frequency in the language, or its frequency in the passage. An additional 16 function words were selected as fillers to make the passage closer to natural prose. The order of the words in the passage was pseudorandom with the constraint that the first two and last two words were selected from the 16 function-word fillers and each remaining block of 16
successive words included one of each of the six types of test word with suffix and the matched filler words, one instance of either and or in and the matched filler word, and two of the additional function-word fillers. The order of words within a block of sixteen was random.

The second passage, hereafter referred to as the "mixed case passage," was identical to the standard case passage except that every other letter was typed in capitals. There were two versions of this passage. In version A even letters were capitalized, and in version B odd letters were capitalized.

Each passage was typed on a separate sheet of paper, and each subject was shown a copy of the standard case passage and a copy of one of the two versions of the mixed case passage, presented in a counterbalanced order and preceded by a sheet of instructions. The two passages and the instruction sheet were stapled together. Half of the subjects were shown version A of the mixed case passage and half were shown version B.

Procedure. The instructions and procedure were essentially the same as in the preceding experiments.

Results

The results are summarized in Table 6, which shows the mean and the standard error of the mean error percentages as a function of passage type and test word category.

Table 6

Means (and Standard Errors) for Error Percentages in Experiment 4 as a Function of Passage Type and Word Category

<table>
<thead>
<tr>
<th>Word-Stem Group</th>
<th>Function</th>
<th>ment</th>
<th>en</th>
<th>ion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passage type</td>
<td>and</td>
<td>in</td>
<td>ing</td>
<td>ment</td>
</tr>
<tr>
<td>Standard case</td>
<td>41.7</td>
<td>51.0</td>
<td>27.6</td>
<td>13.1</td>
</tr>
<tr>
<td>(3.7) (3.3)</td>
<td>(2.5) (2.0)</td>
<td>(2.5) (1.1)</td>
<td>(2.5) (1.2)</td>
<td></td>
</tr>
<tr>
<td>Mixed case</td>
<td>14.0</td>
<td>38.7</td>
<td>9.3</td>
<td>4.9</td>
</tr>
<tr>
<td>(2.7) (3.3)</td>
<td>(1.6) (1.2)</td>
<td>(1.8) (1.0)</td>
<td>(1.6) (1.1)</td>
<td></td>
</tr>
</tbody>
</table>

Function words. As in the previous experiments, subjects made more errors on the common function words and and in than on the other test words. Also, as noted previously, the percentage of errors on the function words was greater in the standard case passage than in the mixed case passage [F(1,115)
The difference between the two passages was much larger for the word and than for the word in \( F(1,115) = 13.91, p < .01 \). A more detailed analysis of the data provides an explanation for this difference between and and in. In version A of the mixed case passage, each of the three instances of the word in was typed as in, which is a regular version of the word, whereas in version B, each of the three instances of the word in was typed as inN, which is not a regular version of the word. The mean error percentage on the word in was much larger in version A of the mixed case passage (mean = 52.3) than in version B (mean = 24.7) \( t(114) = 4.55, p < .01 \). This analysis suggests that only when the word in is typed in an irregular fashion will the percentage of letter-detection errors on it be depressed, which is consistent with the hypothesis that subjects make many letter-detection errors on the word in only when they are able to read it in units larger than the letter.

Content words. For each word-stem group, test words ending in -ing yielded more errors than the corresponding words ending in the alternative suffixes \( F(1,115) = 61.05, p < .01 \). There was also an effect of word-stem group: More errors were made on the words in the -ment group than in the other two groups \( F(2,230) = 12.59, p < .01 \), and the group factor did not interact with the factor of suffix type (-ing vs. the alternative suffixes) \( F(2,230) < 1 \). This result suggests that the different error frequencies among the three word-stem groups were not due to differences in word endings but rather to differences in word stems.

As observed in the earlier experiments, subjects made more errors on the standard case passage than on the mixed case passage \( F(1,115) = 49.2, p < .01 \). The difference between passages was greater for the words ending in -ing than for the corresponding words with other endings, and was somewhat greater for the -ment group than for the other two groups. The interaction between passage type and suffix type (-ing vs. the alternative suffixes) was significant \( F(1,115) = 43.16, p < .01 \), as was the interaction between passage type and word stem group \( F(2,230) = 8.30, p < .01 \). Hence, the difference between passages was greatest for those conditions with most errors.

Discussion

Subjects made more errors on words with -ing endings than on words with any other ending. Since the positional frequency of -ing at the end of a word is considerably greater than that of any of the other letter sequences used, the present results provide further support for unitization of letter groups on the basis of their positional frequency. Spatial predictability on its own cannot account for the present pattern of results because -ion, like -ing, occurs almost exclusively at the end of words (Mayzner, Tresselt, & Wolin, 1965). However, the letter sequence -ing is more frequent than the letter sequence -ion (Underwood & Schulz, 1960; -ing: 1673, -ion: 1370). Hence, orthographic frequency would seem to be a prerequisite, though not a sufficient condition, for the unitization of letter groups.

Although the results of Experiments 1 and 2 show that more errors were made on word suffixes than on word stems, the present task does not discriminate between different classes of suffixes. Specifically, we found no
consistent differences in error percentages between inflectional and derivational suffixes. Since the words with the inflectional ending -en yielded a similar number of errors as the words with the derivational ending -ion and fewer errors than the words with the derivational ending -ment, the preponderance of errors on the words ending in -ing cannot be attributed solely to the fact that -ing is an inflectional, as opposed to a derivational, ending. Furthermore, since the same word stems were used with -ing endings as with the other endings, the differences found between errors on the words ending in -ing and errors on the words with other endings cannot be attributed to differences in the word stems. Finally, since not all suffixes yielded many detection errors, these data argue against the existence of a suffix as the sole explanation of the observed effects; other factors must be involved as well.

The fact that the errors on the words ending in -ing were particularly disrupted by typing the words with every other letter in capitals further supports the hypothesis that the errors were caused by the subjects' reading the suffix -ing as a perceptual unit.

SUMMARY AND CONCLUSIONS

Subjects searching for the target letter n made many letter detection errors on content words ending with the suffix morpheme -ing. In this regard the suffix morpheme -ing resembled the frequent function words in and and, although the percentage of errors was smaller for -ing than for either of the two function words. We focused on three factors that may contribute to the missing letter effect: (1) the orthographic frequency of the letter sequence -ing, (2) its spatial predictability within the word, and (3) its linguistic function as a word boundary morpheme. Our data suggest that although each of these factors may be important for the missing letter effect, no factor alone is sufficient. Whereas whole word frequency has been shown to be critical to the missing letter effect (Healy, 1976), the orthographic frequency of the test trigram cannot be the sole contributing factor, since considerably fewer errors were made on words ending with the letter sequence -ent than with the sequence -ing, which is reported to be less frequent (Experiments 1 and 2). Furthermore, fewer errors were made on -ing when it occurred in the middle of the word than when it was the word ending (Experiment 2). This result agrees with our earlier studies in which considerably fewer errors were made on letter sequences the and and when they were embedded in other words (e.g., thesis, handle) than when they occurred as separate function words. The high orthographic frequency of the letter sequence in which the target is embedded may be a necessary condition for the appearance of the missing letter effect but is in itself not sufficient.

Similarly, although the results of Experiment 1 suggest that the spatial predictability of the target trigram may be important, it cannot be the sole contributing factor to the missing letter effect for two reasons. First, fewer errors were made on words ending in -ion than on words ending in -ing, despite the fact that -ion, like -ing, occurs almost exclusively at the end of English words (Experiment 4). Second, fewer errors were made on the -ing ending itself when it occurred as part of a word stem than when it served as a suffix morpheme (Experiment 1). However, the fact that the letter sequence
was a suffix morpheme could not be the only critical factor, since considerably fewer errors were made on the suffix morphemes -ment, -en, and -ion than on -ing (Experiment 4). We tentatively conclude that it is the combination of orthographic frequency, spatial predictability, and linguistic role as a word boundary morpheme that may be necessary for full expression of the missing letter effect.

In addition, we have identified a number of other factors that may affect the rate of detection errors in this task but cannot account for the specific pattern of errors in the present study: (1) the single letter spatial predictability, (2) the nature of the attached word stem, (3) the frequency of the test word, (4) the syntactic and semantic context of the test word, and (5) the class of the suffix morpheme (inflectional vs. derivational). First, subjects in Experiment 1 made many more errors on words ending in -ing than in -ent, despite the fact that n was the penultimate target letter in both cases. Second, subjects in Experiment 4 made more errors on words ending in -ing than on words ending in -ment, -en, or -ion, despite the fact that the same word stem was used for -ing as for the other endings. Third, subjects in Experiment 2 made more errors on words containing -ing than -ent, despite the fact that the mean frequency of words containing -ing was lower than that of words containing -ent. Fourth, manipulations of test word context in Experiment 2 did not affect the number of letter detection errors. Finally, manipulating the class of the suffix morphemes in Experiment 4 did not yield consistent differences between inflectional and derivational suffixes.

We have thus isolated some of the conditions responsible for the large number of letter-detection errors on the sequence -ing. We propose that these errors occurred because under these conditions -ing tends to be unitized during reading and its constituent letters are thereby "concealed." The most direct support for the hypothesis that the letter sequence -ing may serve as a single unit during reading is provided by the data of Experiment 3. Subjects searching for the letter sequence -ing made fewer detection errors than subjects searching for the letter n in the same passage. Such a pattern of results would be expected to occur if the n target but not the -ing target was smaller than the reading unit employed. Furthermore, fewer errors were made by subjects searching for the letter n in passages of mixed typecase than in normally typed passages, particularly for words with -ing endings. This finding would be expected to occur if the formation of the reading units is disturbed by the use of mixed typecase. The fact that the differences in error rates between different word endings are reduced but not eliminated by the use of mixed typecase suggests that the postulated reading units may be response units, rather than perceptual, or purely visual, ones.

The present data may be limited in their generality, because few if any suffix morphemes other than -ing may meet the constraints we have specified above for the occurrence of the missing letter effect. However, these data taken in conjunction with the results previously obtained for function words (Drewnowski & Healy, 1977) and for common nouns (Healy, 1976, 1980) provide a coherent and general framework in which to understand the nature of the reading process. In particular, this series of studies provides evidence that subjects employ units of varying sizes in reading passages of text. Earlier we showed that common words are more likely to be read in units larger than the letter than are rare words and that frequent function words in appropriate
linguistic context may be processed in units that include more than one word. We have now shown that frequent morpheme suffixes may also be read as single units. In addition, less frequent word stems may be read in terms of letter units. We therefore propose that in reading a given passage, the subject can employ units of different sizes, not only within the passage, but also within a single word.

REFERENCES


FOOTNOTES

1 The word during may seem questionable as an example of a word stem; however, this should only minimize any differences between the word stem and the morpheme suffix groups of words.

2 The present scrambled-word passage is comparable to the scrambled-word passages of Healy (1976) and Drewnowski and Healy (1977, Experiment 1) and not to the scrambled-word passage of Drewnowski (1978), in which every instance of the test word the occurred in an inappropriate context.