Journal of Experimental Psychology:
Learning, Memory, and Cognition
1991, Vol. 17, No. 1, 66–80

0278-7393/81/$3.00

The Missing-Letter Effect in Hebrew: Word Frequency or Word Function?

Asher Koriat
University of Haifa

Seth N. Greenberg
Union College

Yona Goldshmid
University of Haifa

Letters are often missed in processing highly common words, such as *THE* and *AND* (the missing-letter effect). According to the unitization model, this occurs because familiar words access their whole-word representations, preempting processing of their letters. In contrast, we attribute the missing-letter effect to the linguistic function of these words. In Hebrew, some function morphemes appear as single letters prefixed to content words. Although prefix words are not more frequent than matched-content words, they engendered a missing-letter effect that was confined to the function morpheme letter. Interpretation of theses results in terms of a greater unitization of the phrase as a whole was not supported. Also ambiguous Hebrew words produced more omissions when interpreted as prefix plus stem combinations than as unprefixed content words. The missing-letter effect reflects postlexical processing, where function morphemes are lost in the transition from structure to meaning.

A major issue in reading research concerns the units that govern the reading of text (see Henderson, 1982). Healy and her associates (e.g., Healy, 1976; Healy & Drewnowski, 1983) posited a hierarchy of processing levels, and assumed that readers process a text in parallel at the various levels of analysis available to them. They further postulated that familiarity with a unit at a given level facilitates its processing by permitting access to higher order unitized representations. Thus, highly familiar words are encoded more easily than rare words, because they activate their unitized representations at the whole-word level. Moreover, once a unit at a given level is identified, subjects proceed to the next segment of text without completing the processing of units at lower levels in the hierarchy (e.g., constituent letters). Presumably, it should be more difficult to detect a target letter in a familiar unitized word than in a less familiar word that engenders slower access to whole-word representations.

Much of the impetus for the unitization position comes from a phenomenon referred to by Healy and others as the missing-letter effect, where, under a variety of conditions, it has been consistently shown that high-frequency words tend to conceal their component letters (e.g., Drewnowski & Healy, 1977; Healy, 1976; Healy, Oliver, & McNamara, 1987; Proctor & Healy, 1985). Specifically, readers find it more difficult to detect a target letter *T* embedded in a highly familiar words (e.g., *THE*) than in less familiar words (e.g., *WEATHER*) when searching connected text.

Alternatively, the missing-letter effect has been attributed to attentional processes (e.g., Corcoran, 1966; Krueger, 1989; Schindler, 1978). This position stresses the role of semantic and syntactic factors in reading. Because readers can secure the meaning of a sentence without identifying each of the component words, they may invest little attention in the familiar and redundant function words (e.g., *THE* and *AND*). This idea is supported by research on eye movement during reading, indicating that readers make fewer and shorter fixations on short, predictable function words (Carpenter & Just, 1983; Rayner, 1977). O'Reagan (1979) also found that longer eye movements were made when the eye approached the word *THE* than when it approached a three-letter verb. Further, Haber and Schindler (1981) discovered that spelling errors were missed more frequently in function than in content words when the two classes of words were matched for word length and whole-word shape (also see Schindler, 1978).

One study that directly contrasted the unitization and redundance–attentional explanations appeared to favor the unitization position (Proctor & Healy, 1985). Subjects performed one of two secondary tasks while reading for comprehension: detection of the letter *T*, or detection of a double-dot sign that appeared beneath one of the letters. Whereas letter detection was poorer for high-frequency than for low-frequency words, dot detection did not differ. Presumably, if less attention had been paid to the critical function words, dot detection should have also suffered when dots appeared below the critical word *THE*. However, Krueger (1989) suggested that dot detection may depend in part on peripheral vision, and may not be

This research was supported by Grant No. 88-00395 to Asher Koriat and Seth N. Greenberg from the United States-Israel Binational Science Foundation, Jerusalem, Israel.

Experiments 1, 2, and 4 were conducted at the Institute of Information Processing and Decision Making, University of Haifa, while Experiment 3 was conducted at Union College.

We are indebted to Alice Healy, Albrecht Inhoff, Judith Kroll, and Keith Rayner for their helpful comments on an earlier version of the article. We also thank Hamutal Kreiner for her help in setting up and conducting the experiments.

Correspondence concerning this article should be addressed to Asher Koriat, Department of Psychology, University of Haifa, Haifa, Israel, or to Seth N. Greenberg, Department of Psychology, Union College, Schenectady, New York 12308.

sufficiently sensitive to capture attentional differences. Thus, the findings may not have provided a critical test of the two positions.

One question that relates to the unitization–attention controversy, and is also of interest in its own right, is whether the greater rate of omissions for function words derives from their very high frequency or from their linguistic role in text. Whereas the unitization hypothesis assumes that subjects fail to detect letters in *THE* because of its greater unitization as a high-frequency word, the attentional hypothesis appears to stress its role as a function word in a sentence. Function words are highly redundant in text, and can be readily anticipated on the basis of their surrounding context. Therefore, they may receive less attention than the more informative content words during text processing (Corcoran, 1966). Although it was observed that the missing-letter effect was obtained even where low and high-frequency words were matched for linguistic class (e.g., Healy, 1976), the most dramatic effects have been found with the most frequent function words (i.e., *THE* and *AND*).

Although Healy and her colleagues investigated the relative importance of these two factors, their results leave open the question of whether a word's function contributes to the missing-letter effect. Healy (1976) and Drewnowski and Healy (1977) demonstrated that subjects continued to make more omission errors on the word *THE* than on control words even when these were presented in a scrambled-word passage. Also alterations in typecase in a sentential context reduced misses in the word *THE*. Neither effect is consistent with the assumption that the higher error rate for the word *THE* was due to its function in a sentence. Furthermore, using a proofreading task, Healy (1980) found that subjects were more accurate in detecting misspellings of *THE* than of other words, suggesting that they did not simply give insufficient attention to *THE*. This conclusion received further support from Healy and Drewnowski's finding (1983) that omission errors were greatly reduced when *THE*, which appeared in prose, was misspelled.

However, other findings suggested that, indeed, function is critical to letter detection. Although scrambling the words in a passage had little effect on letter detection in content words (Drewnowski & Healy, 1980), it did exert a dramatic effect on function words. In particular, where the scrambling of a passage retained local context (words immediately around the critical word, as when *THE* was followed by a noun), letter detection remained difficult. In contrast, where local context was inconsistent with function (*THE* plus verb), letter detection was much better (Drewnowski & Healy, 1977; Healy, 1976). Drewnowski and Healy attributed the local context effect to unitization of familiar phrases. They proposed that, in addition to words, frequent phrases can also be unitized, and that the unitization of these phrases helps to conceal function words and their component letters. However, it was not clear from their study that the phrases themselves were indeed high-frequency word sequences. Rather consistent local context simply meant that the ordering of words in the phrases was syntactically legal. Thus, it is conceivable that local context reduced letter detection by supporting the interpretation of *THE* in its syntactic function rather than by contributing to phrase unitization. In short, aspects of Drewnowski and Healy's data can be explained by assuming that subjects allocate less attention to function words rather than by postulating unitization based on higher order units. Additional support for the proposition that function mediates letter detection is that locating misspellings in verbs appears to depend on whether the verb has a primary or auxiliary function in a phrase (Abramovici, 1983). Thus, proofreading of *was*, for example, was more accurate when it appeared in "he was big" than in "he was coming."

Moreover, a comparison of some of the results obtained by Drewnowski and Healy (1977, 1980) indicated that subjects may be processing letters in function and content words differently. Briefly, although detection of letters in frequent function words was affected by manipulation of context, detection of letters in frequent suffixes (e.g., *ING*) was not affected by the same manipulation. Further, subjects made a disproportionately high number of errors on the word *AND* whether they searched for the letter *N* or for the entire word *AND* (Drewnowski & Healy, 1977). In contrast, they were much more accurate on the *ING* ending when they searched for the entire *ING* trigram than when they searched for the letter *N*. Because both *ING* and *AND* are highly familiar patterns, perhaps, the difference between *ING* and the high-frequency function words depends upon their respective linguistic roles. Thus, further research on the contribution of linguistic function to the missing-letter effect appears warranted. The present research attempted to unconfound frequency and function by taking advantage of some interesting properties of the Hebrew language.

Experiment 1

Hebrew provides alternate forms for several of its function words. For example, the word *to* can appear either as the two-letter word *EL* or as a single letter prefix *L* (*Lamed*) to successing words. Because both forms have nearly identical functions and meanings in the Hebrew language, two equivalent phrases can be constructed, although the prefix form is somewhat more common. However, the word *EL* is a very frequent, short word, whereas *L* prefixed to particular words creates whole patterns that are not necessarily very common. If the frequency of the word as a whole is the critical factor in the missing-letter effect of Healy and her associates, then the designated target letter *L*, which appears in both forms of "*to*," should be more difficult to detect in the word *EL* than in either the prefix words or in content words that also begin with the letter *L*. The prefix words should produce no more detection errors than their companion content words. On the other hand, if word function is a crucial factor in letter detection, then both forms of *to* should produce more detection errors than the content words.

Method

Subjects. Sixty University of Haifa students (40 women, 20 men) whose native language was Hebrew participated in the experiment, 51 for course credit and the rest as volunteers.

Design. The design of the experiment was based on four function morphemes in Hebrew. Each of these can be represented by two forms, either as a separate function word, as is typical in English (function word), or as a single consonant attached as a prefix to a noun in a sentence (function prefix). The function words were *BETOCH* (*in*), *EL* (*to*), *MIN* (*from*), and *ASHER* (*that/who*), and their corresponding function prefixes where *B*, *L*, *M*, and *S*, respectively. In every case, the function word contained the single letter that defined its matching prefix. Thus, it was possible to have subjects search for the same target letter in both versions of the function form (e.g., *EL* and *L* both contained the target consonant *L*, "lamed"). To illustrate, the expression "Who finished" can be translated either as *ASR GMR*, using the function word, or as *SGMR*, using the function prefix. These two expressions are pronounced *ASHER GAMAR* and *SHEGAMAR*, respectively; but in unpointed Hebrew orthography, the vowels are normally not explicitly expressed (See Koriat, 1984; Navon & Shimron, 1984). Note that both of the aforementioned expressions incorporate *S* in the function unit. Similarly, the expression "to Dan" can be written either as *EL DN*, or as *LDN*. The orthographic forms *ASR* or *EL* are similar to the function words in English, being highly frequent and relatively short. In contrast, the orthographic patterns *SGMR* or *LDN* are less frequent and generally longer than their function-word counterparts in the study.

The design of the experiment involved the use of three types of matched sentences for each of the target letters. These are referred to as function-prefix, function-word, and content-word sentences. In the function-prefix sentences, one word (the critical word) contained the target letter as a function prefix. The matched function-word sentence was identical, except that the function prefix was replaced by the equivalent function word placed in front of the same noun as the prefix (e.g., *LDN* became *EL DN*). The content-word sentence also matched the prefix sentence, except that the critical prefix word was replaced by a content word. However, to construct a semantically and syntactically corrected content-word sentence, it was sometimes necessary to modify one or two words. The content words yoked to the prefix words also displayed the target letter in the initial position.

There is some difficulty involved in controlling for frequency in Hebrew. The frequency counts available are not only inadequate, but also omit a separate listing of the frequency of specific prefix plus stem combinations (e.g., "to Haifa," "in Haifa," and so on). Therefore, in constructing the materials for all Hebrew experiments to be reported, we had to rely on our own intuitions when controlling for frequency, and chose content and prefix words all having a moderate-to-high frequency of usage in Hebrew.

Stimulus materials. The 144 experimental sentences used contained between 12 and 24 words each, and represented 12 replications of each factorial combination of the three sentence types (function prefix, function word, and content word) and four target letters, *B* (*Beth*), *L* (*Lamed*), *M* (*Mem*), and *S* (*Shin* or *Sin*). There was only one critical word in each sentence, and it contained only one target letter. Note, however, that within an experimental sentence target letters appeared in noncritical words as well.

For each of the four target letters, then, there were 36 experimental sentences: 12 parent function-prefix sentences, 12 matched function-word sentences, and 12 matched content-word sentences. The 36 sentences belonging to a particular target letter were randomly distributed across 3 different pages in a subject booklet, with the following restrictions. Each page exhibited exactly 4 experimental sentences from each condition, and no page allowed more than one version of a particular sentence. Thus, if a sentence appeared in prefix form on a page, it was not repeated in any other form on that page. Given four target letters and three pages per letter, a booklet had 12 pages of sentences.

In addition to the experimental sentences, 36 filler sentences were constructed that did not include any of the critical function words or function prefixes. These were also distributed equally across pages, so that 3 filler sentences appeared on each page, making a total of 15 sentences per page. A page appeared as one long paragraph of continuous text composed of unrelated sentences, with a period at the end of each sentence. The order of sentences on a page was random. The filler sentences and noncritical words in experimental sentences were used so subjects would not expect targets to occur only in function morphemes in the majority of sentences.

The sentences were typed double-spaced by a 6630GP laser printer. Words appeared in Hebrew unpointed orthography, as is most commonly used in adult reading material (see Koriat, 1984). This orthography lacks much of the vowel information that is present in most alphabetic scripts. Thus, the Hebrew function words contained two to four letters each, with *ASHER* spelled as *ASR*, *MIN* as *MN*, and *BETOCH* as *BTOC*. Figure 1 illustrates the sentences used.

Procedure. Subjects were run in small groups. They received a self-contained booklet that included 1 page of instructions, 2 pages of practice, and 12 experimental pages. The instructions indicated that subjects were to read passages at their normal reading speed, but whenever they came across the target letter, designated at the top of each page, they were to circle it. They were further instructed not to slow down their reading speed to catch all target letters, and not to go back to circle a letter they had missed. The two practice pages were included to familiarize the subjects with the task of having to search for different letters on different pages. The target letters searched (*Cheth* and *Aleph*) were not among those included in the experiment proper.

The 12 experimental pages were arranged in three blocks of 4 successive pages, where each page in a block was devoted to a different target letter. For each subject, the sequence of the target-letter assignments was kept constant across all three blocks, but the ordering of letter-target presentation within blocks was roughly counterbalanced across all subjects. Subjects proceeded from one page to the other at their own pace.

It should be noted that because the letter *Mem* has two different forms in Hebrew, according to whether it occupies a terminal or a nonterminal position, subjects were explicitly instructed to circle both forms of the letter.

Results

Table 1 presents the means and standard errors of the percentage of omission errors by letter target and condition. These means were derived by calculating percentages for each subject and then averaging across subjects. For each subject, these percentages were based on a total of 12 observations per cell, except for the filler percentages, which are based on 143, 224, 126, and 153 observations for the letters *B*, *L*, *M*, and *S*, respectively. Finally, note that the filler percentage for the letter *M* included both Hebrew versions of this letter, normal and terminal.

Consider first the effect of letter. It is clear that the four letters differed markedly in their likelihood of omission, $F(3, 177) = 91.55$, $p < .0001$. The letters were ordered as follows in terms of increasing likelihood of omission: *S*, *L*, *M*, and *B*. The same order was generally maintained across the four conditions.

However, the primary concern in the present experiment was letter-detection performance in the three matched conditions: function word, function prefix, and content word. A two-way analysis of variance (ANOVA), Condition × Letter Target, yielded significant effects for letter, $F(3, 177) = 62.02$,

TARGET LETTER: מ

FUNCTION WORD

1. יהיה זה מוזר אם ראש הוועדה יעדר מן העיר ולא יהיה נוכח בישיבה.

FUNCTION PREFIX

2. יהיה זה מוזר אם ראש הוועדה יעדר מהעיר ולא יהיה נוכח בישיבה.

CONTENT WORD

3. יהיה זה מוזר אם ראש הוועדה יעדר מחר ולא יהיה נוכח בישיבה.

FUNCTION WORD

1. It would be strange if the committee chairman were to be away from town and not be present at the meeting.

FUNCTION PREFIX

2. It would be strange if the committee chairman were to be away from town and not be present at the meeting.

CONTENT WORD

3. It would be strange if the committee chairman were to be away tomorrow and not be present at the meeting.

Figure 1. The Hebrew sentences used in Experiment 1 and their English translations. (The top panel shows three matched sentences containing a critical function word, function prefix, or content word, with the critical letter marked with an arrow. The target letter is *Mem*, designating "from" when used as a function prefix. Note that Hebrew is read from right to left.)

$p < .0001$, condition, $F(2, 118) = 41.24$, $p < .0001$, and the interaction, $F(6, 354) = 18.32$. $p < .0001$. It may be seen that the Condition × Letter interaction resulted primarily from the three conditions being more clearly differentiated for the letters with higher overall likelihood of omission than for those with lower likelihood of omission.

Content versus function words. Function words induced 17.5% errors compared with 10.7% for the content words, $F(1, 59) = 27.89$, $p < .0001$. This effect was significant only for the letter *B*, $F(1, 59) = 58.31$, $p < .0001$, but was in the expected direction for the letters *L* and *M* as well. It should be noted that the letters *L* and *S* occupied the first position in the content words but the second position in the function words. Given that initial letters are somewhat more available than medial or final letters (e.g., Drewnowski & Healy, 1980), the function disadvantage for *S* and *L* function words is probably overestimated by the results presented in Table 1. The data for individual subjects indicated that 41 subjects made more errors on function than on content words compared with 16 who showed the opposite pattern 3 subjects evidenced equal rates of errors for both classes of words).

Content words versus function prefixes. Initial letters of words were also more likely to be missed when they constituted function prefixes (22.6%) than when they formed part of a content word (10.7%), $F(1, 59) = 71.50$, $p < .0001$. This difference was significant for the letter *B*, $F(1, 59) = 90.41$, $p < .0001$, the letter *M*, $F(1, 59) = 21.82$, $p < .0001$, and the letter *S*, $F(1, 59) = 10.36$, $p < .0001$, and nearly significant for the letter *L*, $F(1, 59) = 3.11$, $p < .10$. The data for individual subjects indicated that for 51 subjects error rate was higher for function prefixes than for content words compared with 4 subjects who exhibited the opposite trend (there was a tie for the remaining 5 subjects).

Function prefixes versus function words. Not only did function prefixes yield higher error rates than the content words, but they also induced more errors than the equivalent function words, $F(1, 59) = 16.92$, $p < .0001$. The differences were relatively small, but nevertheless significant for the letter

Table 1
Means and Standard Errors for Percentages of Omission Errors by Condition and Target Letter (Experiment 1)

Target letter	Function word		Function prefix		Content: Word		Filter		All	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
B	31.5	3.1	38.1	3.0	9.9	1.7	21.9	1.8	22.8	1.7
L	17.2	2.8	17.8	2.1	14.0	1.9	16.5	1.5	16.5	1.5
M	19.4	2.6	28.8	2.9	16.3	2.1	20.3	1.4	20.4	1.4
S	1.7	0.6	5.8	1.3	2.8	0.9	3.5	0.6	3.5	0.6
All	17.5	1.8	22.6	1.7	10.7	1.3	16.0	1.1	16.9	1.4

B, $F(1, 59) = 7.24$, $p < .01$, the letter *M*, $F(1, 59) = 12.05$, $p < .001$, and the letter *S*, $F(1, 59) = 13.21$, $p < .001$. The individual data indicated that 43 subjects made more errors on function prefixes than on function words compared with 14 who exhibited the opposite pattern.

Comparing subjects with different error rates. Subjects were divided into two equal groups on the basis of the number of errors committed on the filler words. A two-way, Subject Type × Condition ANOVA yielded significant effects for subject type, $F(1, 58) = 20.52$, $p < .0001$, and for condition, $F(2, 116) = 42.09$, $p < .0001$, but not for the interaction $F(2, 116) = 2.21$. Mean percentage of errors for the content-word, function-word, and function-prefix conditions averaged 6.3, 10.3, and 17.2%, respectively, for the low-erring subjects, and 15.1, 24.6, and 28.1%, respectively, for the high-erring subjects. Thus, the differences as a result of condition replicate across subjects with different rates of errors.

Analysis of first block only. Because subjects saw similar versions of the same sentence in different blocks of the experiment, we also analyzed the results for the first block only. Percent errors for this block averaged 11.3, 20.0, and 25.2%, for the content-word, function-word, and function-prefix conditions, respectively, $F(2, 59) = 23.66$, $p < .0001$. Mean error rate was significantly lower for the content words than for either the prefix words, $F(1, 59) = 48.95$, $p < .0001$, or the function words, $F(1, 59) = 17.16$, $p < .0001$. The difference between the latter two types of stimuli was also significant, $F(1, 59) = 6.50$, $p < .05$.

Discussion

The results of Experiment 1 yielded two major findings. First, the detection of target letters in running text was more difficult for letters in a high-frequency function words than in lower frequency content words. This result replicated the findings of Healy and her colleagues (e.g., Drewnowski & Healy, 1977; Healy, 1976; Proctor & Healy, 1985), but with Hebrew-printed text rather than English. Healy and her associates explained that high-frequency words are unitized and therfore tend to conceal their constituent letters.

The second finding, however, presents difficulties for the unitization account. The critical and novel condition involving prefix words produced significantly more detection errors than did the content words and, in fact, more than the high-frequency function words. It appears that the missing-letter effect cannot simply be attributed to orthographic frequency. The prefix words, as whole-word patterns, are apparently no more frequent than the content words, making it difficult to explain these results simply on the basis of frequency. Thus, if we were to generalize from these results to English, it appears that the high rate of omission errors observed for *THE* does not derive solely from its high frequency, but may have much to do with the specific syntactic role of this word in the text.

These results appear problematic for the unitization account for two reasons. First, in the unitization account, the higher rate of omission errors observed for *THE* and *AND* is attributed to the very high frequency of these words in written text, which results in their being processed as whole-word units. However, the prefix Hebrew words of the present study were not particularly high frequency so as to promote their processing as whole-word patterns, and might, in fact, be expected to be less "unitized" than the corresponding content words.

The second problem raised by these findings is illustrated by comparing our results with those of Drewnowski and Healy (1980). They found that detecting *N* in *ING* is more difficult when *ING* represents a morphological unit (e.g., *HAVING*) than when it is part of the word stem (e.g., *DURING*), suggesting that subjects process morphological components. They concluded that complex, low-frequency words may be read at a level smaller than a word, although larger than a letter (e.g., morpheme). Because the prefixes in the present study also constitute morphological units, perhaps they are less likely to be detected for the same reason that *N* is not detected in the suffix *ING*.

However, in the prefix words of the present study, the morphological unit represented by the prefix was identical to the target unit. Therefore, the prefix disadvantage observed is inconsistent with the key assumption of the unitization model that errors occur where the size of the target is smaller than the size of the postulated reading unit. Thus, in the case of the *ING* unit, for example, subjects were assumed to miss the letter *N* because the availability of the encompassing morphological unit conceals its letters. However, were that the case here, subjects should have detected the target letter in prefix words more, not less frequently, than in the content words, because the target letter coincided with the reading unit. One solution to this difficulty is provided by results suggesting that the effective reading unit for common English words includes the adjacent interword spaces (see Healy, Conboy, & Drewnowski, 1987). The effective unit for the prefix morphemes may also include the previous interword space, which would make the unit larger than the target.

We should note that, although the results on the whole suggest a role for function in the missing-letter effect, the higher error rate for the function prefixes compared with the function words indicates that yet another factor may be involved. Perhaps the function-prefix disadvantage could be attributed to the wider range of semantic usage of the function prefix compared with its function-word counterpart. Thus, the four prefixes used in the present study convey a variety of meanings in Hebrew in addition to the one captures by the corresponding function words. For example, both *B* and *BETOCH* can be used to signify "inside," but *B* can also mean "in" or "with." Alternatively, Hebrew readers may use prefix versions of expressions more often than the equivalent function word versions. This latter possibility gains support from a follow-up experiment where 16 subjects were presented with all 48 experimental sentences in both a prefix and a function-word version. Subjects judged which of these versions was more likely to occur. The prefix version was considered more likely in 84% of the comparisons across all subjects and sentences. However, although these data indicate that the function-prefix version is more frequently used, note that the prefix words themselves (e.g., *LDN*) occur much less frequently than the functions words corresponding to the prefix (e.g., *EL*). Thus, the function prefix disadvantage should not be construed as support for the unitization position.

Finally, the results of the present study can help reject another account of the missing-letter effect based on word length. As is well known, word frequency is correlated not only with linguistic class but also with word length, so that the most frequent words tend to be relatively short (Zipf, 1949). Although the very frequent function words tend to be very short in both English and Hebrew, the prefix words used in the present study were generally no shorter than their matched control words. The high rate of omission errors observed for the prefix words further supports the claim (see Healy, 1976) that word length is not responsible for the higher error rate found for the frequent function words in English or Hebrew.

Experiment 2

The high rate of omission errors associated with the not particularly frequent prefix words suggest that, independent of frequency, a word's function contributes to the missing-letter effect. This finding seems to raise problems for the unitization model stressing the higher frequency of function words.

However, the unitization position may handle these results by attributing the function-prefix disadvantage to processing at levels higher than the word (see Drewnowski & Healy, 1977). Specifically, function prefixes together with the stem to which they are attached could be assumed to form a cohesive short phrase that conceals its constituent letters. Indeed, Drewnowski and Healy suggested that "frequent function words, although separated from other words by word boundaries, may be read as prefixes or suffixes of the neighboring word" (1977, p. 646). This proposition is even more applicable to the Hebrew prefix words of Experiment 1. Note that subjects judged the prefix versions of the sentences used in Experiment 1 as being more likely to occur than their matched function-word versions. So perhaps function prefixes are particularly effective in generating well-integrated, familiar phrases.

Hence, according to a unitization account of the prefix disadvantage, when a function prefix (e.g., *L*) is appended to a content word (e.g., *LDN*, meaning "to Dan"), a more unitized phrase-level unit is created that conceals its letters. If this is correct, then letter omissions should increase not only for the prefix letter, but also for other constituent letters as well (e.g., *N*). In contrast, a function account attributes the missing-letter effect to the role of the function prefix within the phrase. Therefore, a high rate of omissions ought to be observed only for the letter serving as a function prefix, and not for the other letters. Such an outcome would imply that the prefix word is parsed into its constituent morphemes (*L* + *DN*). Once the parsed phrase is properly interpreted, it leaves behind a memory trace where the letters that support the function element serve only as a background for those letters supporting the content element, and are less likely to be available.

Method

Subjects. Sixty-four University of Haifa students (51 women, 13 men) whose native language was Hebrew participated in the experiment for course credit. None had participated in Experiment 1.

Stimulus materials. Brief sentences were used, each containing a critical word. Each critical word contained two target letters, one letter (*B*, *L*, *M*, or *S*) in the initial position, and a different letter drawn from the same set in the last (*B*, *L*, or *S*) or penultimate position (*M* only). Note, that the Hebrew equivalent of *M* changes print form in the final position, so the penultimate position was used to keep its print form constant across tested positions. Three critical prefix and three matched content words were selected for each of 12 cells. These cells reflected all possible combinations of "initial" and "other" target letters (i.e., the letter occupying either the last or next-to-last positions), under the constraint that each target letter appears only once in a target word (yielding [4 × 3 =] 12 combinations of letters across two positions). In the prefix words, the "initial" target letter formed the prefix, whereas the "other" letter appeared as part of the word stem. The content words, in contrast, displayed both their "initial" and "other" target letters as part of the word stem. The content words had the same number of letters and contained the same initial and final (or next-to-final, in the case of *M*) letters as did their matched prefix counterparts. For example, the content word *LHB* (meaning "blade") was matched with the prefix word *LRB* (meaning "to the Rabbi," with *L* serving as a prefix).

Each critical word was assigned to its own sentence. Thus, with two types of critical words (prefix and content), 12 combinations of target letters across the two critical positions (initial and other), and three exemplars for each combination of letters, there were 72 sentences; 36 prefix and 36 content. Sentences were constructed for each pairing of content and prefix word so that the number of words in these yoked sentences was the same, and the position of the critical words in the sentences was also the same. Finally, in none of the sentences did the target letter appear in the word that preceded or followed the critical word, nor did the letter appear as a function prefix except in the critical word.

Recall that each critical word housed a different target letter in the two positions. Therefore, each of the 72 sentences was used twice throughout the experiment so that on one occasion the target letter that had to be searched was in the initial position of the critical word, whereas on the other occasion it appeared in the other position. The 144 sentences were distributed among eight different pages. Each of the four letters (*B*, *L*, *M*, or *S*) served as the target letter in two pages. Across the 18 sentences on a page, the tested target letter appeared four or five times in each of the four conditions, generated by the factorial combination of critical word (content vs. prefix) and position (initial vs. other). However, over the two pages each letter was tested exactly nine times in each of the four conditions. An extra sentence was added at the beginning of a page and served as a filler (*warm-up*).

Booklets were arranged so that the test order of the four targets was the same for both the first and second set of four pages comprising a booklet. The test order for target letters was counterbalanced across subjects. It should be noted that the two appearances of each sentence occurred in different halves of the booklet. Thus, if a sentence containing the critical word *LHB* (blade) with the target letter *L* appeared in the first half of the booklet, the same sentence with *LHB* and target letter *B* would not appear again until the second half of the booklet.

Procedure. The instructions, procedure, and practice were similar to those of Experiment 1, except that when the target was the Hebrew *M* subjects were told to mark only the nonterminal version of that letter.

Results

The results for omission errors in Experiment 2 were expected to yield an interactive pattern between word type (prefix vs. content) and target position (initial vs. other). Mean percentage of errors is presented in Table 2 for the four combinations, and it can be seen that the results confirm this expectation: A two-way ANOVA on these means yielded $F(1, 63) = 57.52$, $p < .0001$, for word type, $F(1, 63) = 3.96$, $p < .06$, for target position, and $F(1, 63) = 84.93$, $p < .0001$, for the interaction. When the target letter was in the initial position of a word, rate of omission errors was higher when that letter constituted a function prefix (23.1%) than when it was part of the stem of a content word (10.3%). In contrast, when the letter was in the last or next-to-last positions, and therefore constituted part of a content morpheme in both types of words, percentage of errors was not greater for the prefix words (14%) than for the content words (14.5%). One-way ANOVAs confirmed that the difference between prefix words and content words was highly significant for the letter occupying the initial position, $F(1, 63) = 141.47$, $p < .0001$, but was not significant for the letters in the other position, $F < 1$. Also, for prefix words, omission errors were higher for the initial prefix letter than for the letter in the other position, $F(1, 63) = 33.47$, $p < .0001$, whereas for the content words, error rate was actually higher for the letters occupying the last or next-to-last position, $F(1, 63) = 10.31$, $p < .005$.

It should be noted that the interactive pattern between word type and letter position was obtained for each of the four target letters. Thus, the Word Type × Target Position interaction yielded $F(1, 63) = 49.21$, $p < .0001$, for the letter *B*, $F(1, 63) = 16.13$, $p < .0002$, for the letter *L*, $F(1, 63) = 31.73$, $p < .0001$, for the letter *M*, and $F(1, 63) = 9.01$, $p < .005$, for the letter *S*.

Because each sentence was repeated across the two blocks of the experiment, we also analyzed the results for the first block only. A two-way ANOVA, Word Type × Letter Position, yielded $F(1, 63) = 20.64$, $p < .0001$, for word type, $F < 1$ for target position, and $F(1, 63) = 61.46$, $p < .0001$, for the interaction. For the initial position, error rate was higher for the prefix (22.0%) than for the content words (9.8%), $F(1,63) = 79.44$, $p < .0001$, whereas for the other positions it was somewhat higher for the content (16.2%) than for the prefix words (13.3%), $F(1, 63) = 4.18$, $p < .05$.

Discussion

The results of Experiment 2 clearly indicated that the relatively high rate of omission errors observed for prefix words is confined to the letter serving as a function prefix, and does not extend to the other letters in the word. These results suggest that the missing-letter effect is a manifestation not of a "word frequency disadvantage," as Proctor and Healy (1985) implied, but rather of a "*morpheme* disadvantage."

This conclusion has two important implications. First, the missing-letter effect observed in the present study for Hebrew prefixes does not seem to derive from a greater unitization of the word as a whole. If this were the case, the detection of

Table 2
Means and Standard Errors for Percentages of Omission Errors by Position and Target Letter for the Prefix and Content Words in Experiment 2

	Position									
	Initial				Other					
	Content word		Prefix word		Content word		Prefix word		All	
Target letter	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
B	10.9	1.6	36.3	3.0	16.0	2.1	20.0	2.5	20.8	1.6
L	5.0	1.2	12.8	1.5	16.3	2.1	13.7	1.9	12.0	1.3
M	21.2	2.5	35.8	3.1	20.3	2.4	18.4	2.5	23.9	2.0
S	4.2	1.0	7.5	1.5	5.4	1.1	3.6	1.1	5.2	0.9
All	10.3	1.1	23.1	1.8	14.5	1.4	14.0	1.5	15.5	0.8

letters other than the prefix letter should have been impaired as well. According to the unitization account, the missing-letter effect is due to the greater unitization of frequent words. This unitization allows rapid access to the word-level code, which preempts access to the letter-level entries. However, the differential effects observed for the initial and other letters of the prefix words suggest that these words, if anything, are less unitized than their matched content words.

Second, the results place the locus of the missing-letter effect at a postlexical stage. Because impaired letter detection is confined to the letter serving as a function prefix, this impairment must occur after the word has been parsed into its constituent morphemes, and the initial letter correctly interpreted as a function prefix. This observation suggests a rather different account of the missing-letter effect. According to this account, function morphemes are initially used to establish the overall meaning frame of the phrase or the sentence. Once such a frame has been specified, the function morphemes recede into the background as attention shifts to the extraction of the phrase's substance.

Experiment 3

Experiments 1 and 2 used Hebrew, and so in Experiment 3 we examined whether some of the conclusions could be extended to English. Although Experiment 3 was based on Experiment 2 and generally had similar aims, its design had to be modified to suit the constraints of English. Specifically, Experiment 2 examined the possibility that the function-prefix disadvantage derives from unitization at the phrase level. Because in English, function morphemes are represented by independent words, Experiment 3 evaluated whether the presence of a function word increases the omission errors in neighboring content words in connected discourse. This should be the case if function words in English contribute to a greater unitization of the phrases in which they are embedded.

Drewnowski and Healy (1977) argued that letters in English function words are missed not only because these words are unitized (because of their higher frequency), but also because function words promote unitization of the entire phrase. This conclusion was based on the finding that *T* in the word *THE* was missed more often when *THE* appeared in appropriate local context than in scrambled text. Furthermore, it was also more difficult to detect whole function words, (e.g., *THE*) in intact than in scrambled phrases. If short syntactic phrases formed around function words (e.g., *BOY AND GIRL*) are highly unitized, then the error rate should also be disproportionate for the content words joined by the function word (e.g., *BOY*). This implies that where word order is scrambled (e.g., *GIRL BOY AND*) this should yield similar benefits for letter detection in both the function (*THE*) and content (*BOY*) words. Thus, the focus of Experiment 3 was whether disruption of local context in short phrases improves letter detection in content as well as in function words. To this point, a systematic comparison of letter detection in function and content words within the same familiar and scrambled phrases has not been made.

In contrast with the unitization position, we propose that local context facilitates the semantic/syntactic integration of the function word in the phrase, further closing it off from independent inspection. This effect need not extend to the content word whose semantic identity may depend less upon the embedding phrase. Thus, according to the present position the benefit derived from scrambling text should be confined to function words. Of course, such an outcome would be consistent with some of the findings already obtained with Hebrew materials.

Method

Subjects. Forty-eight students at Union College were paid $3.00 for their participation in the experiment. English was the primary language for all of the participants.

Stimulus materials. Twenty critical three-word prepositional phrases were constructed. Half the phrases began with the preposition *FOR*, whereas the other half began with the preposition *WITH*. Within each prepositional group, the second word was equally often *THE* or *A*. The third word was chosen from a set of moderate- to high-frequency content words (frequency counts ranged from 70 to 450) (Kučera & Francis, 1967) beginning with either the consonant *W* or *F*. *WITH* phrases incorporated only *F* content words (e.g., *WITH THE FAMILY*), whereas *FOR* phrases used the *W* content words (e.g., *FOR A WOMAN*). *WITH* and *FOR* phrases were matched for frequency of content words. Finally, each critical phrase contained only a single *F* and a single *W*.

Two passages of normal text containing 20 "critical" and 20 "filler" sentences were then prepared. Of the 20 critical sentences, 10 had *WITH* phrases and 10 had *FOR* phrases. The critical phrases never appeared at the beginning or end of a critical sentence or at the beginning or end of a line of text. Moreover, words with an *F* or *W* never appeared immediately before or after a sentence's critical phrase. Across the 20 filler sentences, each preposition (*FOR* and *WITH*) appeared five times, and content words beginning with *F* or *W* appeared four or five times each. In none of these sentences, however, were the critical prepositions joined with *F* and *W* content words in the same phrase. Critical and filler sentences alternated throughout a passage. In addition, critical sentences containing *WITH* phrases and those containing *FOR* phrases also alternated over a passage. Thus, the various types of sentences were evenly distributed across a passage.

Both of the normal text passages were then used to generate scrambled passages in which normal word order was altered to keep phrases to a minimum (i.e., similar to the "no local context" condition of Drewnowski & Healy, 1977). To accomplish scrambling within the critical phrase, the article (*THE* or *A*) was moved to another location in the sentence, whereas another word from the sentence replaced it (e.g., *FOR AND WOMAN*). The effect of this change was to destroy the critical phrase, while maintaining the position of the phrase's preposition and content word within the sentence. Additional care was taken so that restructuring of sentences never violated principles used to generate the original passages (e.g., words containing *F* or *W* did not appear immediately before or after the "critical phrase"). Movement of other words in a scrambled passage destroyed other normal phrasing as well. Finally, because whole sentences were not moved, the equal distribution of the types of critical target words was maintained in the scrambled passages.

Procedure. Subjects received a booklet containing two successive normal and two successive scrambled passages. The order of presentation of the scrambled passages was matched to the order of presentation of the parent normal passages. The search for each of the two target letters alternated across the four passages. Thus, each subject searched a normal passage and its associated scrambled passage for the same target letter. Order of letter searches, order of passage types, and order of passages within a type was counterbalanced across subjects. Instructions were similar to those of the earlier experiments.

Results

Mean percentage of omission errors for function and content words in the critical phrases is presented in Table 3. Clearly the results demonstrated that letter detection in content and function words differs within phrases (25.4% vs. 4.2%, respectively). Moreover, the effect of destroying phrases (i.e., scrambling passages) was much different for function words, where error rates dropped from 25.4% to 18.5%, than for content words where rates increased from 4.2% to 6.2%. A three-way ANOVA, Word Type (function vs. content) × Passage Type (normal vs. scrambled) × Letter Target (*F* vs. *W*), confirmed the significance of word type, $F(1, 47) = 95.97$, $p < .0001$, and the key interaction of word type with passage type, $F(1, 47) = 9.56$, $p < .005$, thereby supporting the contention that letter detection in function and content words is not similarily affected by the structural unity of phrases. Moreover, independent analyses of the function and content-word data revealed that the reduction in function-word errors from the normal to the scrambled passage was significant, $F(1, 47) = 5.11$, $p < .03$, whereas the increase in content errors approached significance, $F(1, 47) = 2.85$, $p < .10$.

Because all subjects saw the same passages in both their normal and scrambled versions, there was some possibility that they retained a portion of the text in memory from their first to their second encounter. Therefore, we also performed a separate analysis based only on the first two passages of each subject. A two-way Passage Type (between subjects) × Word Type (within subjects) ANOVA yielded the following results: Passage and word type were both significant, $F(1, 46) = 7.32$, $p < .01$, and $F(1, 46) = 89.03$, $p < .001$, respectively, as was the interaction $F(1, 46) = 16.46$, $p < .001$. Although error rate for function words decreased dramatically from normal to scrambled text (35% to 18.5%), the error rate for content words increased slightly (5.5% to 6.25%).

The results also indicated a main effect for letter (*F* vs. *W*), $F(1, 47) = 4.60$, $p < .05$. As seen in Table 3, the mean error rates were generally higher for *W* than for *F* (15% vs. 12% over all conditions). Importantly, though, letter detection for both *F* and *W* followed the overall patterns discussed previously. Independent analyses of *F* and *W* data revealed higher error rates for function compared with content words, $F(1, 47) = 76.50$, and $F(1, 47) = 48.90$, $p < .0001$, for *F* and *W*, respectively, and the Word Type × Passage interaction was also significant, $F(1, 47) = 3.49$, $p < .07$, and $F(1, 47) = 8.87$, $p < .005$, for *F* and *W*, respectively.

Discussion

The pattern of results in Experiment 3 is consistent with that found with Hebrew words in Experiment 2, enabling us to generalize across two very different languages. Specifically, letter detection in function words is worse than in content words even where both types of words are part of the same phrase. Furthermore, scrambling phrases had opposite effects on letter detection in function and content words. These results are not consistent with the idea that the higher error rate observed for function words derives primarily from the greater unitization of a phrase. If this were so, destroying phrase unity should have produced similar reductions in errors in content and function words.

Rather these findings reinforce the position that the role of individual words within a phrase affects the availability of their respective constituent letters. Function words are particularly important for establishing a semantic/syntactic framework for the phrase. Therefore, these items are more sensitive to the removal of local context. Content words, in contrast, tend to maintain a semantic independence in or out of context. In fact, the present results suggest that content words are more available in a phrase than out of one. Perhaps the content words shouldered more of the structural burden in the scrambled text and thereby suffered some of the same fate as the function words in normal text.

However, although scrambling words had its most profound effect on function words, particularly across the initial two passages, a strong function-word disadvantage was still observed in the scrambled condition. This residual effect is consistent with the notion that function words, in addition to contributing to the unitization of the phrase, are also unitized themselves by virtue of their high frequency (Healy, 1976). The effect is somewhat more difficult to reconcile with our

Table 3
Means and Standard Errors for Percentages of Omission Errors for Function and Content Words in Normal and Scrambled Passages in Experiment 3

	Passage									
	Normal				Scrambled					
	Function word		Content word		Function word		Content word		All	
Target letter	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
F	23.5	3.3	2.3	0.7	18.5	2.3	4.4	1.1	12.2	1.2
W	27.3	3.5	6.2	1.3	18.5	2.4	7.9	1.4	15.0	1.5
All	25.4	2.8	4.2	0.8	18.5	2.0	6.1	1.1	13.6	1.2

interpretation. If, as we proposed, the missing-letter effect derives from the role that function words play in the processing of text, why should it persist despite the removal of local context? Our answer is only speculative. Subjects reported that, in the scrambled passages, they still strived to piece together the meaning, although they did so with great difficulty. The results of Drewnowski and Healy (1977) and Healy, Oliver, and McNamara (1987) suggested that only a minimal local context, perhaps as small as a two-word phrase, is sufficient to maintain a missing-letter effect. Perhaps, then, the attempt of subjects to extract meaning from scrambled passages is responsible for part of the residual function-word disadvantage obtained here.

Experiment 4

The previous experiments suggest that the missing-letter effect for frequent function morphemes is apparently due to the linguistic role of these morphemes in connected discourse. This effect is selectively observed for the letter or letters representing the function morpheme, and does not generalize to the entire word (Experiment 2) or to the entire phrase (Experiment 3) in which the function unit is embedded. These results suggest that the locus of the missing-letter effect is not at the stage where the lexical entry is accessed in memory, but at a postlexical stage, in which the interpretation of individual morphemes is subordinated to the interpretation of the phrase or sentence.

Experiment 4 was intended to provide further support for this proposition. The design used ambiguous Hebrew words that could be interpretated either as prefix words, consisting of a function prefix plus stem combination, or as single morphemes where the initial letter is part of the stem. Hebrew offers many such examples (particularly because of its unpointed orthography) (e.g., *SMR* ["that Mister. . . ' or "kept"]).

Each ambiguous word was placed in a disambiguating context that biased its interpretation toward one of the two possible meanings. The disambiguating context either preceded the ambiguous word in the text (prior disambiguation) or followed it (following disambiguation). Subjects were asked to detect a target letter corresponding to the initial letter of the critical, ambiguous word. We expect greater difficulty in detecting the initial letter of the ambiguous word (e.g., *S* in *SMR*) when the context biases toward interpreting it as a function prefix than when it biases toward interpreting it as part of the stem. Findings consistent with this prediction will indicate that the missing-letter effect does not depend solely on the frequency of the orthographic pattern as such, as postulated by the unitization account, but may have much to do with the semantic–syntactic role of the function morpheme within the phrase.

Also the use of disambiguating context both before and after the critical units allows us to further specify the stage at which the missing-letter effect occurs. If a subsequent disambiguating context affects letter detection, this suggests that the missing-letter effect occurs even when the interpretation of a word (and its parsing) must be delayed until more information is available to allow its integration into the sentence (see Rayner & Frazier, 1989). In contrast, if the effect of context is confined to the prior-disambiguation condition, this may imply that although the missing-letter effect is presumably postlexical, it is affected only by the kind of parsing that occurs on line.

Method

Subjects. Forty-eight University of Haifa students (31 women, 17 men) whose native language was Hebrew participated in the study; 9 were involved in the study for course credit, and 39 were paid NIS 6 (about $3) for their participation in the experiment. None had participated in the previous experiments.

Design. The design of the experiment called for four conditions, defined by whether the disambiguating context occurs before or after the critical ambiguous word, and which of the two interpretations is favored by it. As in Experiment 1, four letters served as the target letters: *B*, *L*, *M*, and *S*. They always occupied the initial position of the critical, ambiguous word.

Stimulus materials. Forty Hebrew ambiguous words were chosen, 10 beginning with each of the letters *B*, *L*, *M*, and *S*. Each ambiguous word could be interpreted either as a prefix word, with the initial letter serving as a function prefix, or as an unprefixed content word. In most cases, the two interpretations also called for different pronunciations of the word (e.g., *SMR* pronounced as *shamar* or as *shemar*).

Each ambiguous word was placed in disambiguating contexts that biased its interpretation toward one of the possible meanings. The disambiguating context followed the critical word in two of every four matched sentences (following-disambiguation—prefix and following-disambiguation—stem). The beginning of these two sentences was identical, hence ambiguous. In the other two sentences (prior-disambiguation—prefix and prior-disambiguation—stem), disambiguating text preceded the critical word. The text following the critical word in these two sentences was similar, but not always identical.

The 160 experimental sentences that housed the 40 critical words contained between 10 and 20 words each. There was only 1 critical word in each sentence, and it never appeared at the beginning or at the end of the sentence. The target letter occurred only once in the critical word. It could occur in other words in the sentence, but not in those immediately preceding or following the critical word. In addition, 96 filler sentences were also formed, which did not include ambiguous phrases. These were used so subjects would not expect targets to occur only in ambiguous words.

Thus, for each target letter there were 40 experimental sentences resulting from the combination of 10 critical words, context location (2), and favored interpretation (2). These sentences were distributed across four different pages in a booklet. On each page appeared 10 experimental sentences, each housing a different critical word. Of the 10 sentences in each page, 5 were of the prior-disambiguation type, and 5 were of the following-disambiguation type; among the 5 sentences in each category, 2 to 3 sentences favored the function-prefix interpretation of the initial letter, and the rest favored the part of stem interpretation. Each page also included 6 filler sentences, 3 of which were placed at the beginning of the page, and the remaining 3 sentences were evenly distributed throughout the page.

In total, booklets contained 2 practice pages, followed by 16 experimental pages, arranged in four blocks of 4 pages each. Within a block, 1 page was devoted to each target letter. The order of the four targets was the same for each subject across the four blocks, but was counterbalanced across subjects. Sentences were so assigned to pages that the matching sentences representing the same level of location of disambiguating context or the same level of favored interpretation were never repeated within the same half (the first two blocks or the last two blocks) of a booklet. Thus, for example, if the "prior disambiguation—prefix" version was assigned to Block 1, then the "following disambiguation—stem" had to be assigned to Block 2.

The order of the experimental sentences on a page was random. A page appeared as one long paragraph, with a period at the end of each sentence. The ambiguous words never appeared at the beginning or end of a line.

Procedure. The instructions, procedure, and practice were similar to those of Experiment 2.

Results

Table 4 presents mean percent errors for prefix-favored and stem-favored interpretations when the disambiguating context preceded the critical word and when it followed the critical word.

A two-way ANOVA, Context Location (prior vs. following) × Favored Interpretation (prefix vs. stem) yielded $F(1, 47) = 27.30$, $p < .0001$, for context location, $F(1, 47) = 3.21$, $p < .10$, for favored interpretation, and $F(1, 47) = 13.31$, $p < .001$, for the interaction. As may be seen in Table 4, the prior-disambiguation condition induced a higher miss rate overall (14.8%) than the following-disambiguation condition (11.4%). The interaction reflected the fact that the expected difference between the two types of favored interpretations was found only for the prior-disambiguation conditions. When the disambiguating context preceded the critical word, error rate was higher when that context favored the prefix interpretation than when it favored the part-of-stem interpretation, $F(1, 47) = 13.57$, $p < .001$. In contrast, when the disambiguating context followed the critical word, there was little difference between the two interpretations, $F(1, 47) = 1.43$, *NS*, and if anything, percent errors were slightly smaller for the function-prefix interpretation.

The effect of favored interpretation for the prior-disambiguation condition was of a rather small magnitude, but was consistently found for each of the target letters, although it was significant only for the target *B*, $F(1, 47) = 13.88$, $p < .001$. The effect for the following-disambiguation condition was not systematic across the four target letters, and for the target letter *M*, error rate was in fact higher for the part-of-stem interpretation, $F(1, 47) = 5.39$, $p < .05$.

It may be recalled that, for each critical word, the two sentences representing the same level of either factor were repeated across the two halves of the experiment (that is, the neutral part of each sentence was repeated). Therefore, it is important to examine the results for each half separately. For the first half (Blocks 1 and 2), the results for the prior-disambiguation condition indicated a higher error rate for the prefix interpretation (18.6%) than for the part-of-stem interpretation (14.7%), $F(1, 47) = 5.31$, $p < .05$, whereas in the following-disambiguation condition it was actually lower for the prefix (9.4%) than for the part of stem interpretation (12.2%), $F(1, 47) = 3.90$, $p < .06$. In the second half, only the prior-disambiguation condition yielded a higher miss rate for the prefix (15.4%) than for the part-of-stem interpretation (11.3%), $F(1, 47) = 7.84$, $p < .01$, whereas the following-disambiguation condition yielded practically identical miss rates for both interpretations (11.6% and 11.3%), $F < 1$.

Finally, as noted previously here, error rate was higher in the prior-disambiguation condition than in the following-disambiguation condition. However, this effect was significant only for sentences favoring the prefix interpretation, $F(1, 47) = 30.21$, $p < .0001$, and not for those favoring the stem interpretation, $F(1, 47) = 1.74$. This pattern implies that the interaction between context location and favored interpretation derives mostly from the increase in error rate as a result of the presence of a prefix-favoring context occurring before the ambiguous word.

Table 4
Means and Standard Errors of Percentages of Omission Errors for Prefix-Favored and Stem-Favored Interpretations, When Disambiguating Context Occurred Before (Prior Disambiguation) and After (Following Disambiguation) the Critical, Ambiguous Word (Experiment 4)

	Prior disambiguation				Following disambiguation			
	Prefix		Stem		Prefix		Stem	
Target	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
B	20.4	2.7	11.0	1.9	12.3	2.0	10.8	2.1
L	17.1	2.2	15.2	2.0	13.1	2.4	14.0	1.9
M	24.0	2.5	21.7	2.7	13.5	2.3	18.1	2.6
S	5.4	1.7	4.0	1.1	4.2	1.3	4.8	1.3
All	16.7	1.8	13.0	1.5	10.8	1.7	11.9	1.6

Note. Prefix = prefix-favored interpretation; Stem = stem-favored interpretation.

Discussion

The results of the prior disambiguation condition indicate that error rate for the identical letter strings differs according to the interpretation instantiated by its prior context: Detection of the initial letter of ambiguous words was poorer where disambiguating context favored its interpretation as a function prefix than as part of a content word. This finding is inconsistent with a unitization account according to which the missing-letter effect depends strictly on the frequency of the orthographic pattern of the word. On the positive side, these results have two implications. First, they provide strong support for the function interpretation of the missing-letter effect, because they show that the error rate for an orthographic unit can vary with its role in a sentence or phrase. Second, these results strengthen the position that the missing-letter effect occurs relatively late in processing, some time after an ambiguous word has been parsed into appropriate morphological units, and its meaning has been accessed or selected.

Although the function disadvantage was observed when the disambiguating context preceded the ambiguous letter string, there was no evidence of the effect when the context followed. This result helps to place an upper limit on the stage at which the missing-letter effect can occur. It suggests that the effect does not occur when the interpretation of the ambiguous word must presumably be delayed until disambiguating information is encountered (see Rayner & Frazier, 1989). Thus, apparently the function disadvantage does not depend on the successful interpretation of the sentence as a whole.

Beyond the general characterization just presented, however, the results do not afford more refined conclusions regarding the exact locus of the missing-letter effect. For one thing, previous work on the reading of ambiguous words has led to conflicting views on exactly how disambiguating context works to resolve lexical (and syntactic) ambiguity (see Gorfein, 1989). According to the selective-access model, context directs access to only the appropriate meaning of the word (e.g., Tabossi, 1988), and in the absence of disambiguating information, the primary meaning is accessed. According to this view, with prior disambiguation, the missing-letter effect must occur after (or simultaneously with) the retrieval of the function-prefix interpretation. Presumably, the primary interpretation of the ambiguous words of Experiment 4 was sometimes as a prefix word, and sometimes not (i.e., a nonprefixed content word). This view should predict more omission errors in the following disambiguation conditions than in the prior-disambiguation—stem condition. This, however, was not the case.

The alternative, multiple-access model (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982) posits that all meanings of an ambiguous word are retrieved, and context influences a subsequent selection process. It follows from this view that the missing-letter effect occurs at a postaccess, postselection stage. This implies that in the case of the prior-disambiguation—prefix condition, which yielded the highest error rate, a function-prefix interpretation is selected, resulting in the subsequent loss or masking of the function morpheme. As long as no disambiguating information exists (as is the case with following disambiguation), both interpretations are entertained, and the initial letter of the critical word remains available. This implies that the missing-letter effect occurs only if a function interpretation has been instantiated immediately upon (or very shortly after) encountering the target letter. The search for disambiguating information may be responsible for the relatively low error rate found in the following-disambiguation condition. This, however, may also result from a reanalysis of the ambiguous word when subjects realize that the inappropriate meaning has been selected.

Experiments that monitored eye movements during reading suggest that the interpretation of a lexically ambiguous word generally occurs on-line immediately upon or very shortly after the word is processed (Rayner & Frazier, 1989). In contrast, when syntactic ambiguity was present, there was a tendency to delay syntactic category assignments (Frazier & Rayner, 1987). It should be noted that the ambiguous words of Experiment 4 also entailed syntactic ambiguity. Regardless, the function disadvantage was not obtained when the function interpretation was clarified by subsequent context, suggesting that it is confined to the kind of selection that occurs on line. Thus, more work is needed to relate the missing-letter effect to the processes involved in resolving lexical and syntacic ambiguity. Perhaps more insight into the dynamics of this effect could be gained by using ambiguous words that differ in the relative frequency of the prefix and stem interpretations (see Rayner & Frazier, 1989) and by monitoring letter detection on line. For the present, the critical point of Experiment 4 is that prior context apparently succeeded not only in identifying the appropriate meaning of the ambiguous word, but also in guiding its parsing into the appropriate morphological units. This effect was strong enough to result in the masking of those letters that were assigned the role of a function morpheme.

General Discussion

The missing-letter effect has been extensively replicated under a variety of conditions, and it is probably one of the most robust effects in reading research. This effect refers to the phenomenon where letters are more difficult to detect in short, frequent function words such as *THE, AND,* and *OF* (e.g., Corcoran, 1966; Healy, 1976; Healy & Drewnowski, 1983; Read, 1983). This finding, as well as several other results obtained with the letter-detection task, provided the impetus for the unitization model, first proposed by Healy (1976), and subsequently extended by Drewnowski and Healy (1980) and Healy and Drewnowski (1983). According to the model, the processing of text occurs at several levels in parallel in terms of the units available at each level. Familiarity of a unit at a given level facilitates its processing, and once such a unit has been identified, subjects proceed to the next location without completing the processing of lower level units. The tendency of the common words *THE* and *AND* to conceal their constituent letters is thus seen to ensue from their greater unitization, allowing direct recognition at the word level without complete identification of the constituent letters. It was further proposed that familiar phrases may also be processed in terms of supraword units that consist of short syntactic phrases or of "word frames" such as "on the ____."

The basic tenet of the unitization account is that the missing-letter effect is due to unit familiarity: *THE* and *AND* engender a high rate of letter omissions because of their high frequency. However, these words also have a specific role in text, and it is not clear that their disadvantage does not derive from their status as function words. Indeed, Corcoran's (1966) attentional account attributes this inferiority to the predictability and high redundancy of function words in text, which permit subjects to skip over them.

Because frequency and function are highly confounded in English (see Haber & Schindler, 1981), the first aim of the present study was to exploit some of the characteristics of Hebrew to assess the independent contributions of these two factors. Specifically, we examined the question of whether the linguistic function of a morphemic unit might contribute to the missing-letter effect when the role of orthographic frequency is minimized. Although there is evidence suggesting that frequency and orthographic familiarity do affect letter detection (e.g., Drewnowski & Healy, 1977; Healy, 1976), it is not clear that word function does not. In fact, some of the results that forced Healy and Drewnowski to interject the notion of phrase unitization (e.g., the finding that destroying local context reduced the size of the missing-letter effect; see Drewnowski & Healy, 1977) are also consistent with a redundancy account according to which sentential context makes it unnecessary to attend to the predictable function words. Also, although some of the results reported by Healy and

Drewnowski argue against the proposition that common function words are merely skipped over by virtue of their high predictability in text (e.g., Healy & Drewnowski, 1983; Proctor & Healy, 1985), these results do not deny the possibility of word function affecting letter detection through a mechanisms other than simple inattention.

The second aim of the study was to obtain preliminary clues regarding the process by which word (or morpheme) function might affect letter detection. As Proctor and Healy (1985) noted, both the attentional and unitization accounts imply that common function words receive less attention or processing than low-frequency words. The main difference between these two accounts, however, lies in that only the unitization model assumes that common words receive sufficient attention to allow processing to the point of word identification. The results of the present study, on the whole, seem to suggest a mechanism that differs from that implied by either of these accounts. Although linguistic function seems to make a contribution to the missing-letter effect over and above that of frequency, it appears clear that function morphemes are not simply skipped over, but are processed at least to the point of lexical access and parsing.

In Experiment 1, we compared detection errors in content words, function words, and function prefixes. The results replicated the common finding in English of a higher proportion of errors for the common function words than for their matched moderate-frequency content words. In addition, however, the prefix words also engendered significantly more errors than the content words, despite the fact that they were apparently no more frequent. These results are not consistent with the assumption that the missing-letter effect for *THE* and *AND* derives from their high frequency. The finding of a similar disadvantage for function prefixes suggests the possibility that the missing-letter effect occurs at a stage that is subsequent to the parsing of these words into their appropriate constituent morphemes.

Although the function-prefix disadvantage is consistent with the redundancy–attentional hypothesis (Corcoran, 1966), it could also be accommodated by a unitization account, assuming that prefix words form unitized phrases that tend to conceal their constituent morphemic units (see Drewnowski & Healy, 1977). This hypothesized phrase-level unitization, however, was not supported by the results of Experiment 2: The same letter was detected equally well at the beginning and end of a content word, but where the beginning letter was a function prefix, detectability was much worse at the beginning than at the end of the letter string. Thus, the missing-letter effect is confined to the letter serving as a function prefix, and does not generalize to the other letters in the word. This observation suggests that prefix words, if anything, are less unitized, not more unitized, than their matched content worse, and that the missing-letter effect apparently occurs at a postaccess, postparsing stage. Presumably, the prefix words are first parsed into their constituent morphemes, and then the letters supporting the function element are somehow lost or masked.

These results also imply that the missing-letter effect occurs at the level of the morpheme rather than at the level of the entire word. In this sense, the results parallel the finding that *ING* is more likely to conceal the better *N* when *ING* represents a separate morpheme than when it is part of the stem (Drewnowski & Healy, 1980). However, in the Hebrew prefix words, the target coincides with the function morpheme, and therefore the prefix disadvantage cannot be explained in terms of the tendency of reading units to conceal their constituent letters. However, perhaps the effective unit in the case of function prefixes also includes their previous interword spaces. Indeed, Healy, Conboy, and Drewnowski (1987) obtained evidence suggesting that familiar word sequences that often include the words *THE* may be read in terms of "word frames" that also contain the preceding word and space. Thus, perhaps the reading unit in the case of the Hebrew prefix words also contains the previous word and interword space, making the unit larger than the target letter. However, there is no compelling evidence that such a word frame is any more frequent and unitized than the corresponding frame containing the initial letter of the matched content words.

Nor was there any sign for English that phrases formed around function words are better unitized (Experiment 3): Scrambling text had a strong, beneficial effect on letter detection in the function portion of a phrase, but a weak, deleterious effect on letter detection in the content portion of the same phrase. These results are consistent with those obtained with Hebrew (Experiment 2) in rejecting the notion that letter detection might be due to unitization at the phrase level.

Finally, Experiment 4 suggests that the detection of letters in the same exact orthographic units differs according to the linguistic role of that unit in text, suggesting that the missing-letter effect does not depend entirely on the frequency of the perceptual unit (word) as such. There was a greater tendency to miss the first letter of an ambiguous word when previous context indicated that this letter represented a function morpheme than when it indicated that the letter was part of the stem. It appears that the impact of function rests on an appropriate parsing of the letter sequence. Clearly, from Experiments 1 and 2, the function prefix suffers because the reader discovers at some point during processing that the target letter is a function morpheme. This point is made even more dramatically in Experiment 4, where the letter sequence reveals or conceals the target letter depending on the parsing favored by the prior context.

Taken together, the results of the present study are not consistent with a simple version of the unitization account. They suggest that word function exerts a profound effect on letter detection independent of that of the frequency of the orthographic pattern as a whole. Furthermore, the function inferiority observed for Hebrew prefix words is confined to the letter representing the function morpheme, and does not generalize to the other letters in the word. These results are not consistent with an account that ascribes the missing-letter effect solely to the greater unitization of frequent perceptual units, either words or phrases. On the other hand, however, the results are also not consistent with a simple version of the attention account, which places the locus of the missing-letter effect at a prelexical stage, where sentential redundancy permits little or no attention to be allocated to highly predictable locations. Rather it appears that letters are lost at a postaccess stage, so that words containing function morphemes are processed and parsed before they conceal their function constituents.

Although these results help to place some limits on a model for the missing-letter effect, they are not sufficient to specify such a model in full. The structural account presented later represents a tentative, preliminary sketch of such a model. According to our position, the missing-letter effect occurs in the transition from structure to meaning. We propose that the reading process revolves around a part-whole polarity. On the one hand, it is a serial process involving the sequential analysis of different orthographic segments. On the other hand, comprehension requires that each segment be interpreted in terms of the largest possible unit. This calls for a continuous subordination of the representation of reading segments to the whole in which they are embedded. The missing-letter effect is assumed to derive from the process in which the already activated individual units are integrated within an overall meaning schema, and come to assume their specific linguistic roles within this schema.

Specifically, we propose that although the processing of text requires both coding of structure and coding of meaning, the coding of structure leads the way. Thus, subjects tend to rely on a variety of cues, mostly syntactic, to establish an abstract framework for the sentence as a whole. This framework then guides the extraction of semantic properties, and the integration of individual elements into an overall meaning representation (See Rayner & Frazier, 1989). Function morphemes assume an important role in the establishment of this structural framework, but tend to recede to the background as the meaning of a text segment unfolds (see Aaronson & Ferres, 1983). Thus, the missing-letter effect is assumed to derive from the shift in attention from structure to meaning. Indeed, Healy, Oliver, and McNamara (1987) demonstrated that it takes a function word plus a second related word to generate a missing-letter effect.

This view is consistent with several formulations presented by others (see Aaronson & Ferres, 1986; Forster & Ryder, 1971; Rayner, Carlson, & Frazier, 1983). Furthermore, it accords with the frame-and-slot models of sentence production (e.g., Dell, 1986). According to these models, the internal representation of utterances to be spoken are constructed by inserting linguistic items into slots in independently created structures that define the order of the slots. This assumes that the processing of higher level representations tends to lead the way for the processing of lower level representations. Perhaps, in a similar manner, sentence comprehension too is guided by an attempt to define an encompassing frame-and-slot structure that can allow the assimilation and integration of individual units. If such is the case, the letter-detection task of Healy and her associates may provide a powerful tool for exploring the dynamic interplay between structure and meaning that is assumed to occur during sentence comprehension.

References

Aaronson, D., & Ferres, S. (1983). Lexical categories and reading tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 675–699.

Aaronson, D., & Ferres, S. (1986). Reading strategies for children and adults: A quantitative model. *Psychological Review, 93,* 89–112.

Abramovici, S. (1983). Errors in proofreading: Evidence for syntactic control of letter processing? *Memory & Cognition, 11,* 258–261.

Carpenter, P., & Just, M. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading* (pp. 275–307). New York: Academic.

Corcoran, D. W. J. (1966). An acoustic factor in letter cancellation. *Nature, 210,* 658.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Drewnowski, A., & Healy, A. F. (1977). Detection errors on *the* and *and*: Evidence for reading units larger than the word. *Memory & Cognition, 5,* 636–647.

Drewnowski, A., & Healy, A. F. (1980). Missing -ing in reading: Letter detection errors on word endings. *Journal of Verbal Learning and Verbal Behavior, 19,* 247–262.

Forster, K. I., & Ryder, L. A. (1971). Perceiving the structure and meaning of sentences. *Journal of Verbal Learning and Verbal Behavior, 10,* 285–296.

Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language, 26,* 505–526.

Gorfein, D. S. (1989). (Ed.). *Resolving semantic ambiguity.* New York: Springer-Verlag.

Haber, R. N., & Schindler, R. M. (1981). Error in proofreading: Evidence of syntactic control of letter processing? *Journal of Experimental Psychology: Human Perception and Performance, 7,* 573–579.

Healy, A. F. (1976). Detection errors on the word THE: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 235–242.

Healy, A. F. (1980). Proofreading errors on the word THE—new evidence on reading units. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 45–57.

Healy, A. F., Conboy, G. L., & Drewnowski, A. (1987). Characterizing the processing units of reading: Effects of intra- and interword spaces in a letter detection task. In B. K. Britton & S. M. Glynn (Eds.), *Executive control processes in reading* (pp. 279–296). Hillsdale, NJ: Erlbaum.

Healy, A. F., & Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 413–426.

Healy, A. F., Oliver, W. L., & McNamara, T. P. (1987). Detecting letters in continuous text: Effects of display size. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 279–290.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic.

Koriat, A. (1984). Reading without vowels: Lexical access in Hebrew. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 227–242). Hillsdale, NJ: Erlbaum.

Krueger, L. E. (1989). Detection of intraword and interword letter repetition: A test of the word unitization hypothesis. *Memory & Cognition, 17,* 48–57.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Navon, D., & Shimron, J. (1984). Reading Hebrew: How necessary is the graphemic representation of vowels? In L. Henderson (Ed.). *Orthographies and reading* (pp. 91–102). Hillsdale, NJ: Erlbaum.

O'Regan, J. K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics, 17,* 578–586.

Proctor, H. D., & Healy, A. F. (1985). A secondary task analysis of a word familiarity effect. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 286–303.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition, 9,* 443–448.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior, 22,* 358–374.

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15,* 779–790.

Read, J. D. (1983). Detection of F's in a single statement: The role of phonetic recoding. *Memory & Cognition, 11,* 390–399.

Schindler, R. M. (1978). The effects of prose context on visual search for letters. *Memory & Cognition, 6,* 124–130.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meaning of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14,* 489–537.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language, 27,* 324–340.

Zipf, G. K. (1949). *Human behavior and the principle of least effort.* Cambridge, MA: Addison-Wesley.

Received March 9, 1990
Revision received June 4, 1990
Accepted June 11, 1990 ■