Letter detection in very familiar texts

SETH N. GREENBERG and JOANNA TAI
Union College, Schenectady, New York

In the present study, we investigated whether patterns of letter detection for function and content words in texts are affected by the familiarity of the material being read. In Experiment 1, subjects searched for target letters in sentences that had been rehearsed prior to performing the letter detection on them as well as on unfamiliar sentences. In Experiment 2, subjects searched for target letters in highly familiar verses (e.g., nursery rhymes) and in unfamiliar sentences that were matched to the familiar verses. A disadvantage in letter detection for function as compared with content words consistently found with unfamiliar passages was reduced significantly with the familiar material in both experiments. Specifically, letter detection for content words grew worse in familiar text, but letter detection for function words showed a contrasting modest, though nonsignificant, improvement. The results are consistent with the proposition that in very familiar texts, parafoveal analysis permits the identification of generally less familiar content words. Simultaneously, the normal pattern of weighing the structure and content elements of text changes so that more fixations on function words occur than when one is reading unfamiliar texts.

The impetus for the present study came from nearly two decades of work on what Healy (1980) dubbed the "missing-letter effect" (MLE). The effect describes a pattern in which readers miss target letters appearing in frequently occurring words more than they miss target letters appearing in less frequently occurring words (Healy, 1976, 1994; Healy & Drewnowski, 1983; and Healy, Oliver, & McNamara, 1987). The most robust and consistent disadvantage is observed when target letters are embedded in highly familiar function words such as *the*, *and*, and *for*. If one assumes that the letter detection task is a relatively unobtrusive "online" procedure for capturing processes that mediate fluent reading, then the function word disadvantage indicates that different classes of words are processed somewhat differently during normal reading (see, e.g., Haber & Schindler, 1981; Healy, 1976, 1980, 1994; Koriat & Greenberg, 1994). The question raised in the present study was whether the reader's fluency with the text attenuates the distinctive processing of these two classes of words. Although we know that both reading speed and comprehension improve when texts are reread (Levy & Begin, 1984; Levy, Di Persio, & Hollingshead, 1992; Raney, 1993), the effect of rereading on the processing of words belonging to these different classes has not been systematically investigated (but see Raney & Rayner, 1995). Presumably, if the pattern of letter detection undergoes a change in a familiar passage, it would signal that the pattern of encoding and/or representation of these two classes of linguistic units is also al-

tered. In the present project, we used procedures and stimuli that maximized reader familiarity with texts, in order to test how text familiarity would affect the processing of function and content words.

It was unclear whether rereading familiar text would maintain or alter the magnitude of letter detection for function and content words. Presumably, if rereading affects processes that operate independent of those properties that distinguish function words from content words, then the impact of rereading on the MLE should be negligible inasmuch as error rates for function and content words should be uniformly affected (see Scarborough, Cortese, & Scarborough, 1977). However, if rereading alters text processing to either reduce or accentuate the properties that distinguish function from content words, then the size of the MLE should shrink or increase with text familiarization.

Levy and her colleagues (e.g., Levy, 1983; Levy & Begin, 1984; and Levy, Newell, Snyder, & Timmins, 1986) found that the ability to spot misspellings in function and content words is not differentially affected by the familiarity of a text. For both unfamiliar and familiar texts, letter detection in misspelled function words was slightly poorer than in misspelled content words. Detection for both types of misspelled words improved uniformly in familiar passages even as these passages were read more quickly. Levy concluded that the effects of sentential role and passage familiarity are orthogonal.

Although these results are suggestive, there is reason to question whether similar effects of text familiarity would be observed when readers detect letters in function and content words. Levy's data captured the reader's ability to detect misspellings in what she aptly termed function and content *nonwords*. Healy (1980) indicated that proofreading of nonwords denies the reader direct access to word codes,

thereby compromising possible processing differences that might normally arise from the processing of frequent function and less frequent content *words*. Indeed, Healy and Drewnowski (1983) found that proofreading yielded little difference in letter detection performance in misspelled function and content words, and Healy (1980) found that the misspelling of the function word *the* actually led to an advantage in letter detection relative to misspelling of content words. One can contrast those findings with the robust and consistent function word disadvantage in letter detection tasks in which target words are spelled correctly (e.g., Greenberg & Koriat, 1991; Healy, 1976, Healy & Drewnowski, 1983; Koriat & Greenberg, 1991; Moravcsik & Healy, 1995).

At this point, we know of only one experiment in which letter detection performance for correctly spelled function words was compared with that for content words after readers had first been familiarized with the text (Experiment 2, Healy et al., 1987). Sequences of four and one word displays forming a connected passage were presented on a screen to the reader. Each sequence of displays was presented three times. Subjects detected the same letter for every appearance of a particular display. The results yielded a reduction in letter detection errors with each additional exposure, but with a greater reduction for function than for content words. Thus, these results indicated an interaction between text familiarity and factors that distinguish function from content words.

Unfortunately the presentation condition of the aforementioned study clouded the impact of text familiarity. Familiarity with the passage was clearly confounded with the practice of detecting letters for a particular display. Proctor and Healy (1995) have demonstrated that specific training in letter detection allows readers better access to letter-level information, even in the context of reading, and that such training can greatly reduce or even eliminate the frequent function word disadvantage. Thus, the benefit provided function words in their experiment might be attributed more to the practice received on letter detection than to the repeated encounters with the text (see Proctor and Healy's discussion).

Although the proofreading task differs from the standard letter detection task, Levy and Begin's (1984) account of proofreading error patterns does provide a basis for making predictions with regard to letter detection in intact words. Levy and Begin proposed that familiarization with a message allows for more efficient reprocessing of the syntactic and semantic information, freeing additional resources for a more detailed analysis of visual patterns. Presumably, the availability of additional processing resources should improve letter detection in correctly spelled function and content words as well, thereby leaving the size of the MLE relatively unchanged.

The predominant explanations of the MLE, though— that is, the unitization position of Healy (1994) and the structural position of Koriat and Greenberg (1994)— suggest different outcomes. The unitization model's understanding of the MLE (Healy, 1976, 1994) assumes that familiarity with specific linguistic patterns permits a reader to rapidly process such text segments in higher level word and phrase units. Once a large unit is identified, readers quickly proceed to the next text pattern before a reader can complete identification of the embedded letters. From this perspective, increased familiarity with a text ought to expedite higher level unit identification, thereby inhibiting identification of lower level constituent letters. However, if readers are familiarized with a text by the typical procedure of repeated readings (as was done in Experiment 1), letter detection in content words ought to be adversely affected, whereas detection in function words might remain stable or deteriorate only slightly. The basis for this prediction is that repeated readings prior to performing the detection task should enhance the familiarity of lower frequency content words. In contrast, owing to their already long-standing familiarity, additional encounters with function words ought to have little impact on their unitization and hence detection of their constituent letters.

The structural position attributes the MLE to a postlexical period when identified function words are rapidly pushed to the background of the linguistic representation of the phrase and sentence (see Koriat & Greenberg, 1994). Function words are identified early on in an effort to prepare the structural framework for the meaning-laden content units (Bock & Loebell, 1990). However, the structural units are less available than the content words immediately after the structure is established. The MLE is thus seen as a reflection of the diminished availability of the function word in the maturing cognitive representation. Presumably, then a familiar passage whose structure is completely known in advance could even more completely and quickly inhibit the processing of structural units, further compromising letter detection in such items. With regard to content words, the structural model would predict little change in processing, since it is presumed that content units are maintained in the maturing representation.

EXPERIMENT 1

Method

Subjects. Twenty-four undergraduate students at Union College were either paid $4 or given credit toward fulfilling a course option for an out-of-class activity. The students had no previous experience with the letter detection task. They were assigned to one of four presentation conditions.

Design and Materials. Two booklets were constructed for each subject. A first "familiarization" booklet familiarized each subject with a subset of critical sentences that contained the target function and target content words. A second "detection" booklet included the familiar sentences from the familiarization booklet as well as unfamiliar target sentences not from that booklet.

We prepared 40 short sentences, 30 of which were deemed critical and were included in the data analysis. Half of these sentences contained target *t* words and the other half contained target *d* words. The *t* and *d* target words always appeared in the middle of the sentences and were never preceded or followed by another word with the target letter. For the *t* target condition, the target function word was *the*, and the target content words were familiar three-letter

words beginning with *t*: *tip*, *tax*, *two*, *tie*, *tea*, *top*, *try*, and *ten*. For the *d* target sentences, the target function word was *and*, and the critical content words were familiar three-letter words ending in *d*: *kid*, *led*, *old*, *end*, *bad*, *aid*, *bed*, and *red*. In each of the 15 critical sentences for each target letter, the assigned function word appeared once along with one of the appropriate content words. Across the 15 sentences for each target letter, seven of the target content words for each set appeared in 2 sentences, whereas the words *tip* and *kid* appeared just once.

In addition to these critical sentences, five "distractor" sentences were constructed for each of the target letters *r* and *s*. These sentences were not used for the data analysis but were incorporated into the materials simply to discourage subjects from trying to focus on the appearances of *t* and *d*.

From the collection of sentences described above, four versions of familiarization booklets were constructed. Of the 15 *t* and 15 *d* critical sentences, 5 were chosen for the two-repetition familiarization condition, and 5 were chosen for the five-repetition condition. The 5 remaining critical *t* and *d* sentences were reserved for the nonrepetition sentences, and so they did not appear in these booklets. In the two-repetition condition, a target sentence appeared by itself twice on a page, whereas in the five-repetition condition, a sentence appeared by itself five times on a page. Thus, the familiarization booklet contained five pages each of *t* and *d* sentences for each of the repetition conditions. All eight target content words appeared at least once in the two- and five-repetition sentences, and five of the target content words also appeared in the nonrepetition sentences. Ten other pages in the booklet presented the distraction *r* and *s* target sentences, with 2 *r* and 3 *s* sentences assigned to the two-repetition condition, and 3 *r* and 2 *s* sentences assigned to the five-repetition condition. The first two pages and the last page of each familiarization booklet contained distraction sentences. Otherwise, the ordering of the combination of repetitions for each target letter was randomized across the booklet. A second familiarization booklet reversed the page order of the first booklet. Finally, two other familiarization booklets were generated by reassigning sentences that had been first assigned to the two-repetition condition to the five-repetition condition and vice versa.

Each page of the detection booklets contained four sentences, including one that matched the sentence on the corresponding page of its partner familiarization booklet. When the matched sentence was a *t* or *d* critical sentence from the familiarization booklet, it appeared as either the second or the third sentence on the detection page. When, the matched sentence was for the *r* or *s* targets, the sentence was either the first or the last sentence on the page. The location of the matching sentence was varied so that subjects could not anticipate the location of the repeated sentences on a page. In addition to the matching sentences, on half the *t* and *d* target pages there was a second critical sentence. The second critical sentence did not appear in the familiarization booklet. These nonrepetition (unfamiliar) *t* and *d* critical sentences also appeared in either the second or the third sentence location, wherever the repeated (familiar) sentence did not appear. On all pages, words containing *t*s or *d*s appeared in noncritical sentences as well. The sentences were printed on single lines rather than in a paragraph format, although the four sentences actually formed reasonably cohesive passages. Since the detection booklet's pages corresponded to those of its companion familiarization booklet, the order of its critical and distraction sentences was similarly randomized.

Procedure. Sessions lasted 20–25 min. The subjects were given a familiarization booklet and the companion letter detection booklet. They read a page from the familiarization booklet aloud into a tape recorder, while looking at the sentence on the page in the familiarization booklet. After reading a page of repeated sentences from the familiarization booklet, the subjects next read the corresponding page from the detection booklet silently. Detection pages were not revealed until after a page of sentences was read from the

familiarization booklet so as not to indicate the target letter prior to the performance of the detection task. When they were performing the letter detection task, the subjects were told to read at a normal pace while crossing out the target letter indicated at the top of the detection page. They were further instructed not to go back to cross out letters that they had missed. They were timed to encourage their reading at an even pace. Finally, after every 10 pages, following the letter detection, the subjects were stopped and asked simple questions about the content of the last passage read, thereby encouraging reading for content. They were immediately informed of the accuracy of their responses.

Results

The central concern was how prior familiarization with a text would affect the pattern of letter detection in the structure-supporting function words relative to the meaning-laden content words. All analyses used an alpha level of $p = .05$, unless otherwise indicated. Prior to looking at the detection data, we compared the average reading times on first and last readings of repeated sentences in the familiarization booklet for the repeated conditions (two and five). There was a slight saving in reading time ($M = 3.06$ sec on the first reading vs. 2.92 sec on the final reading) that proved significant [$F(1,19) = 16.39, MS_e = 0.04$]. Reading times were not recorded for the first 5 subjects. Although these measures were not taken as subjects performed a detection task, the savings in reading time across readings is consistent with the "reprocessing" advantage disclosed by Levy (1983; Levy & Begin, 1984; and Levy et al., 1986). The results demonstrated an effect of prior familiarization immediately before the subjects performed the letter detection task.

The letter detection data are reported in Table 1 and were initially analyzed in an analysis of variance (ANOVA) that included all three levels of repetition (five, two, and zero), the two target letters (*t* and *d*), and the two word roles (function and content). Consistent with past research on the MLE, there were more errors for function words (73.55%) than for content words (13.95%) [$F(1,23) = 349.5, MS_e = 1.83$]. In addition, a main effect of target letter was found, indicating that the detection of *d*s in the final position of a word was more difficult than the detection of *t*s in the initial position of a word [$F(1,23) = 140.1, MS_e = 0.92$]. As can be seen in Table 1, in particular, the letter *d* in *and* was more difficult to detect than the *t* in *the*

Table 1
Mean Percentages of Letter Detection Errors for Target Function and Content Words in Repeated and Nonrepeated Sentences

| Word Class | Letter | No. of Repetitions | | | |
		0	2	5	*M*
Function	T	57.4	51.6	53.4	54.1
	D	95.5	93.3	90.1	93.0
	M	76.5	72.5	71.8	73.6
Content	T	4.1	5.6	7.5	5.7
	D	15.0	29.0	22.5	22.2
	M	9.6	17.3	15.0	13.9

$[F(1,23) = 66.5, MS_e = 1.34]$, and it was more difficult to detect in content words as well $[F(1,23) = 33.6, MS_e = 0.40]$.

The primary focus, though, was the error patterns for function and content words across the various repetition conditions. In the repetition conditions, letter detection in function words showed an approximate 4.35% improvement in comparison with letter detection in the nonrepetition conditions, whereas content words suffered about a 6.6% increase in letter detection errors in the repetition conditions, yielding a role × repetition (familiarity) interaction $[F(2,23) = 4.27, MS_e = 0.54]$. The two and five repetitions produced similar effects. An analysis comparing both repetition conditions (two and five) indicated no main effect of repetition or interaction involving it.

To further understand the nature of the role × repetition interaction, we combined the two repetition conditions and compared the collapsed condition against the nonrepetition condition. The ANOVA demonstrated that the key interaction between role and repetition remained reliable $[F(1,23) = 11.2, MS_e = 0.30]$. We also performed a more restricted version of this analysis. In this comparison, only repeated sentences that contained the five content words that appeared in the nonrepetition condition were included in the computation of error rates for the collapsed repetition condition. Although this procedure resulted in a more limited number of trials, it confirmed the role × repetition interaction $[F(1,23) = 5.96, MS_e = 0.76]$ with error rates for function and content words of 76% and 9%, and 68% and 15% for the nonrepetition and repetition conditions, respectively. Finally, separate analyses of function words and content words showed that repetition marginally reduced errors for function words $[F(1,23) = 2.89, MS_e = 0.33, p < .10]$ and increased errors for content words $[F(1,23) = 9.3, MS_e = 0.29]$.

Discussion

The observed interaction was a dramatic change from what Levy and her colleagues (Levy, 1983; Levy & Begin, 1984; Levy et al., 1986) found when they investigated the effect of familiarity on proofreading. Given past inconsistencies between proofreading and letter detection tasks in connected text (Healy, 1980), this inconsistency was not unexpected. The results make clear the importance of lexical status when one is assessing the effects of text familiarity during reading. Importantly, they argue against a processing benefit from such familiarization that spills over to letter identification. Freeing of resources as a result of prior readings apparently does not facilitate letter detection in content words. Although proofreading text for misspelled words simply depends on the available processing resources, detecting letters in intact words is more complexly related to text familiarity.

The interaction between role and repetition was not entirely consistent with the predictions of the structural and unitization models either. According to the structural model, one would expect a further deterioration in letter detection for function words, and this did not occur. The uni-

tization model performed somewhat better, in that it predicted a detriment for content word letter detection. However, it too predicted that if there were to be an effect of familiarity on function word letter detection, this would also be to further inhibit detection. Although the effect of text familiarization did not quite produce a significant positive effect on letter detection in function words, the interaction and direction of the effect on function words was not what the unitization model anticipated. The interaction, of course, indicated that text familiarity has an impact on at least one feature that helps to differentiate the processing of function and content words. One curious finding was that five repetitions produced no greater effect than did two repetitions. Perhaps reading the text aloud puts so much focus on each individual word that the effects reach some asymptote with only two repetitions (see, e.g., Carlson, Alejano, & Carr, 1991, regarding effects of reading words aloud).

EXPERIMENT 2

The strategy of Experiment 1 was to familiarize readers with certain sentences immediately before they performed the letter detection task. The results of prior readings of those sentences likely induced a temporary activation of target word lexical and perceptual codes due to implicit priming of the target words. A variety of implicit memory priming studies have revealed that prior exposure to a word often enhances its subsequent perceptual identification (e.g., Jacoby, 1983).

Since the unitization model's predictions matched the outcome of text familiarity observed in Experiment 1 best, it was considered useful to test that model's ability to predict the effect of text familiarity under a related though somewhat different manipulation of this variable. The elevated error rates for content words observed in Experiment 1 were predicted by the unitization model because of the temporary increase in the familiarization of the target content words due to the repeated readings of sentences. As has been noted, though, the model did less well with regard to function words. In any case, in Experiment 2, text familiarity was manipulated without encouraging a *temporary* activation of content word familiarity. Instead, subjects read verses with which they were already quite familiar. Presumably, the unitization model would predict that when readers tackled *well-known verses*, the phrases within would likely be unitized, leading to a uniform increase in the concealment of letters in all target words, as opposed to when the same words appeared in matched unfamiliar verses. Drewnowski and Healy (1977) have in fact suggested, though, that in highly familiar phrases, function words such as *the* are somewhat more apt to suffer poorer letter detection than are constituent content words. In sum, according to the unitization position, this present manipulation of familiarity should lead to an increase in letter detection errors for both classes of words in familiar texts, although the impact on function words might be slightly greater.

Method

Subjects. Thirty adult subjects who were familiar with the experiment's set of familiar reading material participated in the study. The subjects were awarded $3 for their participation. [1]

Design and Materials. Familiar texts were taken from popular children's books and movies. The material consisted of familiar sentences from highly recognizable nursery rhymes, portions of children's song verses, and jingles from childhood movies. Twenty familiar verses were collected in all. From these 20 *familiar* verses, 20 more matching *unfamiliar* verses were derived. Sixteen verses from each set (familiar and unfamiliar) were included in the experimental portion of the study. Two verses from each set served as practice, and 2 more were used as the initial warm-up verses in the experimental portion but were not scored.

The 16 experimental verses for each condition contained 12 target content and 12 target function words. Four verses per condition contained 1 target function word, whereas another 4 per condition contained 1 target content word. Finally, 8 verses per condition contained 1 of each type of target word, function and content. The target words were neither preceded nor followed by another word containing the target letter, and the target words never appeared as the first word in a sentence. The target function word was always *the*, whereas the target content words were short words with a single *t* in either the first or the last position. Other appearances of *the* and content words with *t* (nontarget words) in the familiar and unfamiliar verses were not included in the data calculation.

Unfamiliar verses derived from the familiar verses were matched for word length and contained the same target words in the same locations as those for the familiar verses with which they were matched. Furthermore, the grammatical structure surrounding the target words was held constant for the familiar verse and its derived unfamiliar counterpart. Each verse contained from 11 to 48 words, with an average of 23 words per verse appearing in one to four sentences. No more than 11 appearances of the target *t* in target and nontarget words occurred per verse. Finally, the number of appearances of *t* (for nontarget as well as target words) was matched for the familiar and derived unfamiliar verses. Below are two examples of familiar and derived unfamiliar verses. Target words are italicized.

Familiar sets:

A. Hickory, dickory, dock! The mouse ran up *the* clock. The clock struck one, and down he run, hickory dickory dock!

B. Jack and Jill *went* up *the* hill to fetch a pail of water.

Derived unfamiliar sets:

A. Fee, fie, foe! The boy climbed up *the* pole. The ogre struck once, and made him fall, fee fie foe!

B. Amy and John *went* by *the* lake to catch a fish for later.

Experimental booklets consisted of a front page with instructions, followed by four pages of reading material. The first page of reading material (second page of booklet) contained the four familiar and derived verses chosen for practice. The three succeeding pages contained alternating familiar and unfamiliar verses, with the four initial verses (two familiar and two unfamiliar) not being included in the data. Verses were randomly distributed over the three pages, with the stipulation that a familiar verse and its derived unfamiliar verse not appear within 10 verses of one another. Thus, though the types of verses were alternated, placement of related verses was such that readers would not come upon two related verses for some time. To control for order effects, two booklets were constructed. In one booklet, a familiar verse began the alternating sequence, but in the second booklet an unfamiliar verse began the sequence. The booklets were randomly assigned to the subjects, with half the subjects receiving each of the two orders.

Procedure. Each subject was instructed to read each verse normally, and while doing so to cross out a *t* wherever it appeared in the text. They were issued the standard letter detection warning about not going back to mark target letters already skipped. They were specifically warned not to simply scan the text for *t*s.

Results

Letter detection errors were recorded for the target function and content words in the familiar and their structurally equivalent unfamiliar sets of verses. Table 2 indicates that the difference in detection errors between the function target words and content target words was considerably smaller for the familiar than for the unfamiliar verses (approximately a 16% function word disadvantage in the unfamiliar condition, but only 6.4% function word disadvantage in the familiar condition). Again reported significance is at $p = .05$, except when otherwise indicated. A two-way ANOVA to evaluate the effects of familiarity (familiar vs. unfamiliar) and word role (function vs. content) yielded a main effect of word role [$F(1,29) = 7.50, MS_e = 7.10$] and an interaction [$F(1, 29) = 6.67, MS_e = 1.44$]. A further analysis was run to compare letter detection in *the* and only content words beginning with a *t*. Experiment 1 suggested that the location of a letter in a target word could be of some importance. Unfortunately, highly familiar verses do not lend themselves to a symmetrical representation of short content words for beginning and final positions. Thus, only three target content words were included in this analysis. Despite the limited number of target content words, the interaction of role × familiarity was still significant [$F(1,29) = 4.41, MS_e = 1.11$], with mean error rates for these content words of 6.6% and 13.2% for unfamiliar and familiar texts, respectively (the function data are the same as in Table 2). Follow-up comparisons demonstrated that the effect of word role was highly significant in the unfamiliar condition [$F(1,29) = 11.86, MS_e = 4.56$], but not in the familiar condition [$F(1,29) = 2.21, MS_e = 3.99, .10 < p < .15$]. The difference in letter detection for function words across the two familiarity conditions only approached significance [$F(1,29) = 3.06, MS_e = 1.57, p < .10$], although the difference in error rates for content words did reach significance [$F(1,29) = 6.44, MS_e = .75$].

Discussion

The results of this experiment converged with the key findings of Experiment 1, indicating that the MLE is more pronounced when readers are required to detect letters in unfamiliar texts. The only difference of substance was that the size of the MLE was not as large as that found in Experiment 1. It is difficult, though, to compare the relative sizes of these effects, given that we have no method of de-

Table 2
Mean Percentages of Letter Detection Errors
for Content and Function Words

| Word Class | Familiarity | | *M* |
	Unfamiliar	Familiar	
Function	34.7	30.0	32.4
Content	18.9	23.6	21.3
M	26.8	26.8	26.8

termining the relative impacts of short-term repetition of unfamiliar sentences versus long-term retention of classic rhymes on the attention given to individual words. However, others have speculated that reading aloud tends to draw attention to the lexical level of processing (see, e.g., Carlson et al., 1991), perhaps thereby encouraging unitization of individual words that in turn might enhance the MLE.

Of greater interest, though, is that once again the results challenge the predictions of existing models regarding the MLE in familiar texts. The unitization model, which performed best with Experiment 1, posits that letter detection for both classes of words should be uniformly and adversely affected by familiarization achieved through the presentation of very well known verses, rather than through repetition of text. Unlike in Experiment 1, under these test conditions, the less familiar content words should not gain a short-term elevation in familiarity that would lead to some enhanced concealment of their constituent letters. However, a different set of propositions generated by the unitization model could handle the increased difficulty in letter detection for content words in the present experiment. Specifically, Drewnowski and Healy (1977) explained that unitization might encompass whole phrases when phrases are highly familiar. Nursery rhymes certainly contain such multiword phrase units. The model suggests, however, that when such phrases appear, function words are at least as likely as content words to be included and disguised by such comprehensive units. According to the model, this extra shell of unitization further conceals a letter from detection, and thus letter detection in function (and other constituent) words should only get worse. Alternatively, Moravcsik and Healy (1995) attributed letter detection patterns to the time spent processing a word. Presumably, in well-known verses less time is spent in processing each word. Yet this too leads to a prediction of a fairly uniform increase in errors across all words (at the very least some increase in errors for function words). Thus, for these propositions it is problematic to find even a null effect for function words when content words show a decrease in letter detection.

The structural model would anticipate an increase in errors for function words. As has been indicated, the structural model holds that letter detection patterns mirror the prominence of words in an evolving cognitive representation. With highly familiar texts, both content and function words should be less prominent, and hence letter detection should have shown a general decline. Thus, it appears that neither model handles the comprehensive effect of text familiarity. In the General Discussion, we offer an alternative explanation for how these classes of words are affected by familiar text.

GENERAL DISCUSSION

The MLE diminished in both experiments primarily because letter detection in content words suffered substantially as familiarity increased. In addition, a modest increase

occurred for letter detection in function words, replicating an earlier finding of Healy et al. (1987). The key point, though, is that letter detection did not follow the same pattern for both classes of words.

In order to understand these results, it is helpful to consider some evidence and understandings provided by eye movement research. A particularly pertinent observation is that function words are apparently skipped more often than content words. Carpenter and Just (1983) found that readers fixated 83% on content words and only 38% on function words. The pattern, in part, reflects that function words are generally more predictable and frequent than content words. Investigations of eye movement routinely demonstrate that predictability enhances word skipping (Balota, Pollatsek, & Rayner, 1985; Rayner & Well, 1996), as does frequency (Reichle, Pollatsek, Fisher, & Rayner, 1998). Hadley and Healy (1991) posited that function words, because of their high familiarity, are identified when they appear in the parafoveal region and consequently rarely draw foveal attention (see also Morrison, 1984). Moreover, Hadley and Healy assumed that in the parafoveal region acuity is reduced to a point where the identification of a word's details, including its constituent letters, is compromised. According to Hadley and Healy, foveal analysis is generally required for constituent letter identification. Indeed, Saint-Aubin and Klein (2001) have recently found that letter detection is worse for both skipped function and content words than for fixated words.

If one assumes that prior familiarity with a text can increase a word's predictability, or momentarily enhance its familiarity, then the identification of low-frequency, less predictable content words would stand to benefit from parafoveal analysis in very familiar text. Balota et al. (1985) determined that in highly constrained predictive contexts, predictable words received considerably more parafoveal analysis than did less predictable words.

Raney and Rayner (1995) examined how eye movements in text were altered when a text was read more than once. Importantly, they considered the impact of text familiarization on the processing of high- and low-frequency content words. Target content target words were read faster during the second reading (13–18 msec overall), but the decrease in time was much greater than the average decrease in fixation duration (only 4 msec). They observed that some "most likely function words" (p. 164) showed no repetition benefit at all. Raney and Rayner stressed that even in initial readings, parafoveal analysis of function words permitted them to be skipped or led to their rapid foveal identification. Moreover, they obtained a marginally significant interaction of word frequency × text repetition with regard to the probability of fixation. Fixating target words was less for high-frequency words than for low-frequency words on the first reading, but not on the second. However, it is also worth noting that on a variety of other measures of lexical identification (e.g., first fixation or gaze time), improvement was consistent across high- and low-frequency words. Nevertheless, the fixation measure appeared to be consistent with Hadley

and Healy's (1991) proposition. Of course, it follows that if text familiarity increases parafoveal analysis of content words, but not function words, then letter detection should deteriorate for only content words. It also is of interest that in Experiment 1 the effect of repeated readings was strongest for the target at the end of target word (i.e., d detection in *and* plus its content word matches). Perhaps, the first position target letters were less subject to processing changes due to repetition since they might already have experienced parafoveal analysis in a single reading (see Rayner & Pollatsek, 1989).

What remains a mystery, though, is why letter detection in function words did not follow the same pattern, and indeed gave some hint of improvement. We submit that in the present study, the effect of familiarization was more substantial than that investigated by Raney and Rayner (1995). In Experiment 1, subjects read relatively short passages (single sentences) several times. Presumably, readers acquired a strong—perhaps even "verbatim"—representation for such sentences, one much stronger than that obtained in reading a long paragraph twice as in Raney and Rayner. In Experiment 2, subjects read material so familiar that they could probably have recited much of each verse word for word.

In line with the structural model's reasoning, under these reading conditions, eye movements would be less apt to be guided by an early pursuit of structural units used to lay the framework for meaning (see Dell, 1990; Koriat & Greenberg, 1994). According to the structural model, readers take advantage of the familiarity (frequency) of function words to process them parafoveally for structural definition, and they generally direct their foveal attention toward the content units that these words introduce (see also Saint-Aubin & Klein, 2001). Evidence both from letter detection study (Koriat & Greenberg, 1996) and eye movement work (Schmauder, Morris, & Poynor, 2000) has indicated that in unfamiliar texts, processing and/or landings on content words following function words is higher than that on content words following other content words. In already very familiar texts, though, foveal landings are apt to be less selective. As is suggested by fuzzy-trace theory (Brainerd & Reyna, 1990), when a text is already known, readers may be approaching that text with its gist already stored in working memory, and this may derail the normal pattern of verbatim processing. Consistent with this proposition, a variety of studies have demonstrated that rereading benefits the processing of paraphrases of an unfamiliar text along with the text itself (Levy, Barnes, & Martin, 1993; Raney & Rayner, 1995; and Tardif & Craik, 1989). Presumably, then, the reader's eyes move with less regard for the informative nature of an upcoming word, thereby setting the stage for more haphazard landings and thus greater possibility of fixating function words. After all, if most words are suitable for parafoveal identification due to a stored gist, the chance that any content word will be skipped increases. Given that readers are unlikely to inspect words well beyond the first word in the parafoveal region (Rayner & Pollatsek,

1989), the increase in skipped content words could also lead to an increase in arbitrary landings function words. In accord with Hadley and Healy's (1991) position, such occurrences should result in some improvement in function word letter detection.

Our position is that the characteristics that differentiate function and content words, such as frequency and predictability, undergo temporary adjustments in highly familiar texts that lead to a changing pattern of eye movements. Although for the present study it may not be necessary to consider the sentential roles of words as well, in order to account for the impact of text familiarity, other evidence suggests that this factor should not be discounted. Greenberg, Inhoff, Koriat, and Seely (2001) have found that skipping rates for very high and low frequency content words are not affected by preview. However, skipping of function words increases appreciably under preview conditions. Saint-Aubin and Klein (2001) have determined that for both skipped as well as fixated words letter detection is still worse for function than for content words. Finally, Schmauder et al. (2000) have discovered that for measures of eye movement that reflect text integration processes, function words but not content words require additional processing. They concluded, in accord with the structural model, that in the initial representation of a text, function words are not as prominent as content words, and so when passages require more integration, structural cues are revisited. Of course, letter detection studies from a variety of languages (e.g., Koriat & Greenberg, 1991, in Hebrew; Greenberg & Koriat, 1991, and Greenberg, Koriat, & Shapiro, 1992, in English; Saint-Aubin & Poirier, 1997, in French) have also pointed to the importance of sentential role. Thus, it is premature to suggest that the hypothesized change in discriminative parafoveal processing results only from momentary adjustments to a word's frequency or predictability. What the present results appear to signal is that models that provide reasonable insight into the processing of classes of words in unfamiliar texts (the vast majority of reading) are less able to account for performance in very familiar texts, or at least require some modification.

In sum, although the exact mechanism and variables that contribute to a shift in the processing of individual words and word classes is unclear, the present results leave little doubt that these two classes of words are handled quite differently. Moreover, the disparate effects of familiarity on function and content words suggests that in an environment in which parafoveal identification can play a more dominant role (i.e., in very familiar text), the pattern of parafoveal analysis is modified considerably. Clearly these results demonstrate that the reading of very familiar texts engages a different pattern of processing than that witnessed for the reading of unfamiliar texts.

REFERENCES

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, **17**, 364-390.

BOCK, K., & LOEBELL, H. (1990). Framing sentences. *Cognition*, **35**, 1-39.

BRAINERD, C. J., & REYNA, V. F. (1990). Gist is the grist: Fuzzy-trace theory and the new intuitionism. *Developmental Review*, **10**, 3-47.

CARLSON, L. A., ALEJANO, A. R., & CARR, T. (1991). The level-of-focal-attention hypothesis in oral reading: Influence of strategies on the context specificity of lexical repetition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 924-931.

CARPENTER, P. A., & JUST, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275-307). New York: Academic Press.

DELL, G. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language & Cognitive Processes*, **5**, 313-349.

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.

GREENBERG, S. N., INHOFF, A., KORIAT, A., & SEELY, R. (2001). *Function and frequency interact with preview*. Manuscript in preparation.

GREENBERG, S. N., & KORIAT, A. (1991). The missing-letter effect for common function words depends on their linguistic function in the phrase. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1051-1061.

GREENBERG, S. N., KORIAT, A., & SHAPIRO, A. (1992). The effects of syntactic structure on letter detection in adjacent function words. *Memory & Cognition*, **20**, 663-670.

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

HADLEY, J. A., & HEALY, A. F. (1991). When are reading units larger than the letter? Refinement of the unitization reading model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1062-1073.

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

HEALY, A. F. (1980). Proofreading errors on the word *the*: New evidence of reading units larger than letters. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 45-57.

HEALY, A. F. (1994). Letter detection: A window to unitization and other cognitive processes in reading text. *Psychonomic Bulletin & Review*, **1**, 333-344.

HEALY, A. F., & DREWNOWSKI, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception & 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 & Performance*, **13**, 279-290.

JACOBY, L. L. (1983). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **9**, 21-38.

KORIAT, A., & GREENBERG, S. N. (1991). Syntactic control of letter detection: Evidence from English and Hebrew nonwords. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1035-1050.

KORIAT, A., & GREENBERG, S. N. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin & Review*, **1**, 345-356.

KORIAT, A., & GREENBERG, S. N. (1996). The enhancement effect in letter detection: Further evidence for the structural model of reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1184-1195.

LEVY, B. A. (1983). Proofreading familiar text: Constraints on visual processing. *Memory & Cognition*, **11**, 1-12.

LEVY, B. A., BARNES, L., & MARTIN, L. (1993). Transfer of fluency across repetitions and across texts. *Canadian Journal of Psychology*, **44**, 465-482.

LEVY, B. A., & BEGIN, J. (1984). Proofreading familiar text: Allocating resources to perceptual and conceptual processes. *Memory & Cognition*, **12**, 621-632.

LEVY, B. A., DI PERSIO, R., & HOLLINGSHEAD, A. (1992). Fluent rereading: Repetition, automaticity, and discrepancy. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 957-971.

LEVY, B. A., NEWELL, S., SNYDER, J., & TIMMINS, K. (1986). Processing changes across reading encounters. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 467-478.

MORAVCSIK, J. E., & HEALY, A. F. (1995). The effect of meaning on letter detection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 82-95.

MORRISON, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 667-682.

PROCTOR, J. D., & HEALY, A. F. (1995). Acquisition and retention of skilled letter detection. In A. F. Healy & L. Bourne (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 282-299). Thousand Oaks, CA: Sage.

RANEY, G. E. (1993). Monitoring changes in cognitive loading during reading: An event-related brain potential and reaction time analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 51-69.

RANEY, G. E., & RAYNER, K. (1995). Word frequency effects and eye movements during two readings of text. *Canadian Journal of Experimental Psychology*, **49**, 151-172.

RAYNER, K., & POLLATSEK, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

RAYNER, K., & WELL, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, **3**, 504-509.

REICHLE, E. D., POLLATSEK, A., FISHER, D. M., & RAYNER, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, **105**, 125-157.

SAINT-AUBIN, J., & KLEIN, R. M. (2001). Influence of parafoveal processing on the missing-letter effect. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 318-334.

SAINT-AUBIN, J., & POIRIER, M. (1997). The influence of word function in the missing-letter effect: Further evidence from French. *Memory & Cognition*, **25**, 666-676.

SCARBOROUGH, D. L., CORTESE, C., & SCARBOROUGH, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 1-17.

SCHMAUDER, A. R., MORRIS, R. K., & POYNOR, D. V. (2000). Lexical processing and text integration of function and content words: Evidence from priming and eye fixations. *Memory & Cognition*, **28**, 1098-1108.

TARDIF, T., & CRAIK, F. I. M. (1989). Rereading a week later: Perceptual and conceptual factors. *Journal of Memory & Language*, **28**, 107-125.

NOTE

1. Three subjects who participated in Experiment 2 were not included among the 30 subjects counted in the data, since they had indicated during postexperiment questioning that they were unfamiliar with most of the "familiar" verses. Checks of the other subjects indicated that they were very familiar with these verses.