Pergamon

PII: S0042-6989(97)00095-3

Vision Res., Vol. 37, No. 20, pp. 2899–2914, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Fillers and Spaces in Text: The Importance of Word Recognition During Reading

JULIE EPELBOIM,*§ JAMES R. BOOTH,† REBECCA ASHKENAZY,‡ ARASH TALEGHANI,‡ ROBERT M. STEINMAN‡

Received 13 March 1996; in revised form 22 October 1996; in final form 24 March 1997

Current theories of reading eye movements claim that reading saccades are programmed primarily on the basis of information about the length of the upcoming word, determined by low-level visual processes that detect spaces to the right of fixation. Many studies attempted to test this claim by filling spaces between words with various non-space symbols (fillers). This manipulation, however, confounds the effect of inserting extraneous characters into text with the effect of obscuring word boundaries by filling spaces. We performed the control conditions necessary to unconfound these effects. Skilled readers read continuous stories aloud and silently. Three factors were varied: (i) position of the fillers in the text (at the beginning, the end, or surrounding each word); (ii) the presence or absence of spaces in the text; and (iii) the effect of the type of filler on word recognition (from greatest effect to least effect: Latin letters, Greek letters, digits and shaded boxes). The effect of fillers on reading depended more on the type of filler than on the presence of spaces. The greater effect the fillers had on word recognition, the more they slowed reading. Surrounding each word with digits or Greek letters slowed reading as much as filling spaces with these symbols. Surrounding each word with randomly chosen letters, while preserving spaces, slowed reading by 44–75%—as much as, or more than, removing spaces from normal text. Removing spaces from text with Latin-letter fillers slowed reading by only 10–20% more. We conclude that fillers in text disrupt reading by affecting word recognition directly, without necessarily affecting the eye movement pattern. © 1997 Elsevier Science Ltd

Reading Word recognition Saccadic programming Optimal viewing position

INTRODUCTION

A popular theory of reading eye movements asserts that reading saccades are programmed on the basis of information about the length of the upcoming word. Word-length information is obtained by low-level visual processes that segment the text immediately to the right of each fixation into words by detecting the highly visible spaces that separate them. This information is then used to program efficient reading saccades that land the line-of-sight near the "optimal viewing position" (OVP), located near the center of each word. Missing the OVP is thought to be detrimental to reading speed (see O'Regan, 1990; Rayner, 1993 for reviews).

Although this scheme sounds reasonable, it is specific only to some modern Western languages. Most ancient languages (Latin, Greek and Hebrew, for example) did not separate words with spaces. Some modern Asian languages (such as Thai or Japanese) still write without spaces. Even among the modern Western languages there are some that use spaces much more sparingly than the languages for which the popular space-based theory was developed (i.e., English and French). German and Dutch, for example, have a practice of stringing compound nouns together into single, unspaced words, often over 15 letters long. The readers of these unspaced and sparingly spaced languages must use something other than spaces to program reading saccades.

In addition, there is evidence that some readers of English can read English text from which spaces were removed, as quickly as they read ordinary text, without any practice with unspaced text. Those who do slow down when spaces are removed use the same eye movement strategy with unspaced text as they use with ordinary text (Epelboim *et al.*, 1994). These observations suggest that the saccadic strategy used when ordinary texts are read can be adapted to unspaced text, which means that spaces are not as important to programming reading saccades as currently popular OVP-based theories propose.

The importance of spaces in reading has often been studied by filling spaces between words with various

*Center for the Study of Language and Information, Stanford University, Stanford, CA 94305-4115, U.S.A.
†Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213-3890, U.S.A.
‡Department of Psychology, University of Maryland, College Park, MD 20742-4411, U.S.A.
§To whom all correspondence should be addressed [*Email* yulya@brissun.umd.edu].

extraneous characters (fillers). For example, Morris *et al.* (1990) used a text in which spaces had been filled with "x"s and concluded that "the information about the location of the first and second space both have an effect on saccadic behavior independent of letter masking" (p. 279). Epelboim *et al.* (1996a) pointed out, however, that experiments like this, which used the "filler technique", lacked crucial control conditions in which both fillers and spaces were present. Such controls unconfound the effect of removing spaces and obscuring word boundaries from the effect of inserting irrelevant information into the text. Here, we report the effects that these neglected controls have on reading speed.

Three factors were manipulated in our study: (i) the position of the fillers with respect to words (at the beginning of each word, at the end of each word, or surrounding each word); (ii) the presence or absence of spaces; and (iii) the type of filler. Filler types were: Latin letters, Greek letters, digits and shaded boxes (listed in the order of the magnitude of their effect on the speed of recognition of individual words—from the greatest effect to the smallest effect).

It is known that reading speed is limited by the speed required to program saccades when ordinary text is read. Reading speeds increase three-fold for reading aloud and double for reading silently, when words are presented one at a time—a technique that eliminates the need to use eye movements for processing text (Rubin & Turano, 1992). This means that a manipulation that slows saccadic programming should slow reading. So, if saccades are programmed by detecting spaces to the right of fixation, reading text that contains fillers should be slower when these fillers occlude spaces than when spaces are preserved. This was not the case.

We found that surrounding each word with digits or Greek letters, while preserving spaces, slowed reading by the same amount as filling spaces with these characters. This pattern was observed for both silent reading and in reading aloud. The more fillers affected word recognition, the more they slowed reading. In fact, placing randomly chosen Latin letters at the beginning and the end of each word while preserving spaces, slowed reading more than removing spaces from ordinary text. Although texts in which spaces were filled with Latin letters were read more slowly than texts in which words were surrounded by Latin letters and spaces were preserved, the effect of introducing extra letters was much greater (44–75%) than the additional effect of removing spaces (10–20%). We will argue that this relatively small additional effect of removing spaces was caused by the formation of incorrect groupings of letters around word boundaries, rather than by the disruption of saccadic programming.

We conclude that (i) prior studies of reading that used the filler technique must be re-evaluated because they lacked essential controls in which both fillers and spaces are present; (ii) word recognition plays a more significant role in determining reading speed than local, physical features of the text; and (iii) saccadic programming during reading is more sophisticated than suggested by theories emphasizing spaces between words as markers for guiding the programming of reading saccades.

METHOD

Subjects

Performance of each subject will be reported separately, continuing the approach taken in our prior work (Epelboim *et al.*, 1994; Booth *et al.*, 1995). This is desirable because large individual differences in reading performance are well known and, therefore, examination and interpretation of averaged data is useful only after the data of individual subjects has been analyzed and reported.

Subjects were selected on the basis of their ability to read fluently and to understand what they were reading. A total of 62 subjects were screened, 48 proving suitable for these experiments. The following criteria were used to screen subjects:

1. *Fluent reading of normal text.* Reading fluency was judged while each subject read a paragraph of ordinary text during a brief practice session before the experiment began (see Procedure). Subjects having obvious difficulty reading, speech problems or difficulty seeing the text clearly, were thanked and dismissed at the end of this practice session. They left under the impression that they had completed our experiment and had no reason, whatsoever, to assume that they were unsuitable for additional participation. Three subjects were excluded for this reason.
2. *Comprehension of material read.* Each subject had to score at least 75% correct on the comprehension test. Most subjects answered nearly all of the questions correctly. Most of the 11 subjects who did not, missed more than 30% of the questions. The data of the latter group were not used because their inability to answer simple questions about the text suggested that they did not follow the instruction to read the text for meaning as well as for speed. The subjects whose data were used in the analyses had a mean score of 89.4% correct on the comprehension questions. The subjects whose data were not analyzed had a mean score of 62.1% correct. It was fortunate that our population fell into two distinct groups because recent studies have shown that a *post-hoc* comprehension test, such as ours, can be a poor indicator of reading comprehension (Katz *et al.*, 1990).

Allocation of subjects to conditions

In general, different groups of subjects were used for each filler-type condition in order to avoid interference and practice effects. There were the following exceptions: (1) the same six subjects were used to read text with digit fillers and text with Latin-letter fillers in silent reading conditions. (2) Four subjects had previously

TABLE 1. Placements of fillers tested with different types of fillers and reading conditions

Filler	Reading	n	Filler placement						
			Normal	Begin	End	Surround	Filled	Fill-2	Unspaced
Shaded boxes	Aloud	8	✓			✓		✓	✓
Digits	Aloud	17	✓	✓	✓	✓	✓		✓
Digits	Silent	6	✓	✓	✓	✓	✓		✓
Greek letters	Aloud	9	✓			✓	✓	✓	✓
Latin letters	Aloud	8	✓	✓	✓	✓		✓	✓
Latin letters	Silent	6	✓			✓	✓		✓

participated in experiments that required them to read unspaced text. They had no prior experience of reading texts with fillers. (3) Two subjects participated in two different filler-type conditions, reading aloud. The subjects who participated in more than one session were given a new text each time they participated. Practice or interference effects were not observed in the patterns of results of the subjects participating in more than one experiment, nor was there any appreciable difference in the pattern of results shown by these subjects as compared with those subjects who served in only one session.

Reading aloud and silently

Most subjects read aloud. Their speech was recorded and scored for reading errors (see below). We asked subjects to read aloud because it provides an unambiguous, ongoing measure of reading accuracy and comprehension. Reading aloud also forces the subject to read every word, rather than to skim the text, hoping to pick up enough information to answer *post-hoc* comprehension questions correctly. Such a "skimming strategy", possible only in silent reading, could be particularly useful to a "lazy" subject reading unusual texts, such as texts with fillers. A "skimming strategy" allows a lazy subject to "read" a text much faster than it could be read when each word is recognized and processed as it must be when reading aloud. The reverse is also possible: namely, a subject might get away with skimming ordinary text, but be forced to read more carefully and, therefore, more slowly when text is altered. Interactions between reading strategy and the type of text being read make the results of a subject's silent reading more difficult to interpret than results obtained when a subject reads aloud.

Most reading, however, is silent, both in the laboratory and in everyday life. This fact, in itself, does not force one to confine studies to silent reading in the laboratory, because there is evidence that reading aloud, although somewhat slower, is not fundamentally different from reading silently. For example, Legge *et al.* (1985) and Legge & Rubin (1986) showed that reading speeds vary with character size and wavelength in the same fashion when reading silently and aloud. Rubin & Turano (1992) showed that reading speeds increase 2–3-times when the need for eye movements is eliminated for both types of reading. Epelboim *et al.* (1994) showed that eye movement patterns are similar and that global and local eye movement features are scaled in the same way for both types of reading when spaces between words were removed. Such observed similarities between silent reading and reading aloud, combined with the benefits (described above), make reading aloud a better choice than silent reading for assessing the role of spaces and fillers in text. Key conditions, shown in Table 1, were replicated with silent reading to encourage researchers accustomed to studying only silent reading to accept our results as representative of reading in general. As expected from the prior research just cited, the pattern of results observed was the same for both kinds of reading.

Scoring pronunciation

Analyses of pronunciation errors with texts read aloud allowed measurement of and correction for effects of any speed/accuracy trade-off induced by asking a subject to read both quickly *and* accurately. Each subject's speech was scored for the following reading errors: (1) corrections—words, initially read incorrectly, but subsequently corrected; (2) insertions—words said but not present in the text; (3) omissions—words present in the text but not read; and (4) substitutions—incorrect words substituted and not subsequently corrected.

"Corrected reading time" was calculated as follows: time to read each paragraph (to the nearest 10 msec) was divided by the number of words read correctly (total number of words in the paragraph minus the number of omissions and substitutions). Insertions were rare (mean = 0.16 insertions/paragraph) and their frequency did not depend on text-condition. Corrections were more common (mean = 2.7 corrections/paragraph), and were affected by text-condition with more corrections in conditions that produced slower reading speeds ($P < 0.05$). Insertions and corrections were not used in calculating corrected reading time because (i) corrections were not really errors; and (ii) neither corrections nor insertions resulted from sacrificing accuracy to gain speed, i.e., they increased, rather than decreased reading time. Our corrected reading time is similar to the measure of reading speed used by Legge and collaborators in experiments in which subjects read sentences aloud as they drifted across a screen (e.g., Legge *et al.*, 1985, 1987).

Mean corrected reading time for a given subject and condition was calculated as the sum of the corrected

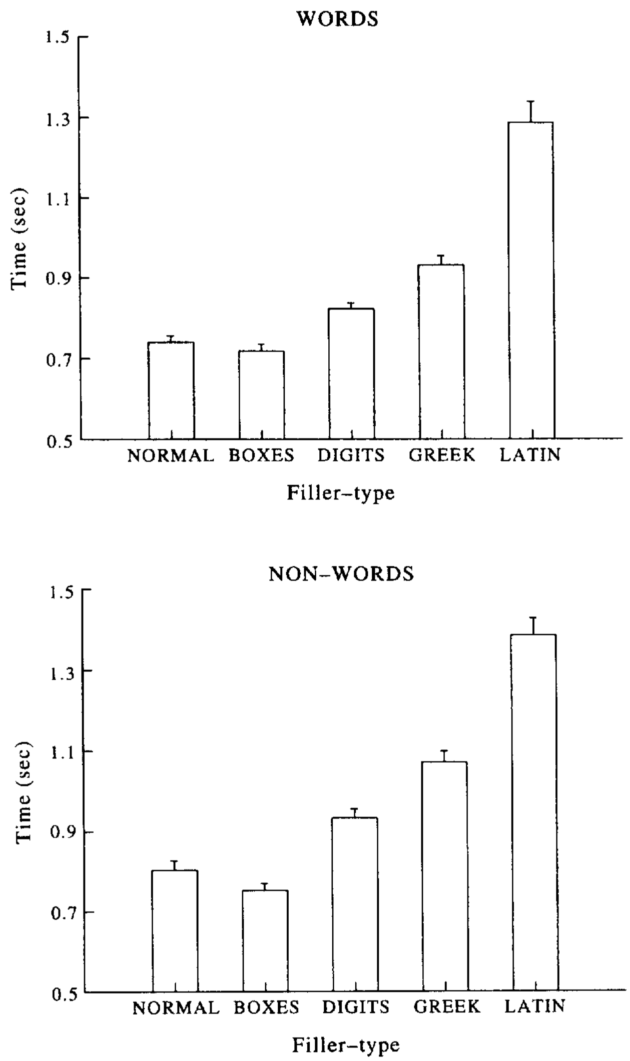

FIGURE 1. Mean reaction times for deciding wheter a six-letter string surrounded by fillers is an English word or a non-word. Performance of a typical subject is shown. The other two subjects had the same pattern of results. Each bar is based on 91–97 trials. Error bars show 1 SE.

reading times for all paragraphs read by that subject in that condition divided by the number of paragraphs.

In general, subjects, selected for reading fluency and comprehension (discussed above), made very few errors of any type. The mean difference between corrected reading time and raw reading time (time/number of words in the paragraph) tended to be slightly larger for more difficult conditions, but the difference did not exceed 10% for any subject or condition, and was usually much smaller. All analyses of reading aloud used corrected reading time, calculated as described.

Scoring reading accuracy was impossible during silent reading, so raw reading time was used (time to read a paragraph/total number of words in that paragraph). Comprehension here was assessed only by asking the subject to retell the story read and asking questions about it at the end of the experiment. All six subjects who read silently answered over 90% of these questions correctly.

Types of fillers

Four types of fillers were used in this study:

1. *Shaded boxes*: ASCII character 176 was used as the filler. This character appeared similar to the "grating" shown in Fig. 4 of Pollatsek & Rayner (1982).
2. *Digits*: Eight single digits from 2 to 9 were used as fillers. Digits "1" and "0" were not used because they resemble letters "l" and "O". Digits at each filler location were selected randomly by the computer before the entire paragraph was displayed.
3. *Greek letters*: Eight lower-case Greek letters were also used as fillers. The following Greek letters were used: α, π, τ, δ, θ, ϕ, μ and σ. Their corresponding ASCII numbers were: 224, 227, 231, 235, 233, 230, 237 and 229. Only eight letters were used so as to match the Digit-filler condition in which eight different digits were used. A Greek letter was selected randomly from this set and assigned to each filler position before the entire paragraph was displayed. Greek-letter fillers had not been used as fillers in prior experiments to our knowledge. However, they are similar to "letter-like" symbols that had been used as fillers (e.g., Malt & Seamon, 1978).
4. *Latin letters*: A set of eight lower-case Latin letters was selected randomly for each subject. Eight, rather than 26, letters were used as fillers also to match the Digit-filler condition in which only eight digits were used. The letter "a" was not used because it is a common word in itself. A random letter was selected from the set of eight at each position where a filler was required.

The rationale for selecting these types of fillers was two-fold: (i) these or similar types fillers had been used in prior experiments (e.g., Pollatsek & Rayner, 1982; Malt & Seamon, 1978); and (ii) each filler-type slowed word recognition to a different degree, as described next.

Effect of different filler-types on word recognition

The following experiment was performed to determine how filler-type affected word recognition: a six-letter string of letters, flanked by a filler on each side (e.g., 1shower3) appeared in the center of a computer screen. The six-letter string was either an English word or an orthographically legal anagram (e.g., 1rowsher3). Subjects judged whether the letter-string between fillers was a word or not a word and their reaction times were recorded. Three subjects participated. Each ran 1000 trials in a single session. Trials were randomized in blocks of 200 (100 words and 100 non-words). Each block used one of the following filler-types: no fillers (just a word), shaded boxes, digits, Greek letters and Latin letters (taken from a randomly selected set of eight). The same font and character-size was used in the word recognition experiment as in the main reading experiment. Subjects made few errors (<6%) with most of the errors occurring with Latin-letter fillers. Only correct responses were included in the analyses. Mean reaction times for a representative subject are shown in Fig. 1. All subjects gave similar results: namely, shaded boxes did not have a reliable effect on word recognition time ($P > 0.3$). All other fillers did ($P < 0.05$). Latin letters slowed word recognition the most, digits the least, and Greek-letter fillers slowed word recognition more than digits, but less than Latin letters.

Placement of fillers

The following filler-placements were used (examples with digit fillers are shown in parentheses):

1. Normal—Normal text (this is an example);
2. Begin—A filler at the beginning of each word, spaces preserved (1this 3is 7an 2example);
3. End—A filler after the end of each word, spaces preserved (this1 is3 an7 example2);
4. Surround—Fillers surrounding each word, spaces preserved (9this1 4is3 6an7 8example2);
5. Filled—A filler filling each space (9this2is5an8example4);
6. Fill-2—Two fillers filling each space (42this54is89an72example39);
7. Unspaced—Spaces removed, no fillers (thisisanexample).

Spaces were removed and fillers were added without adjusting the length of the line. Thus, lines of unspaced text were shorter than lines of normal text, and lines with both spaces and fillers were longer than lines of normal text. When a word was followed by a punctuation mark, the punctuation mark was considered part of the word, and the filler was inserted after it.

Texts

Subjects read stories presented in paragraphs of 9–11 lines of text. The mean number of words/paragraph was 89. Each paragraph was presented in its entirety on the computer screen before the subject started reading, and remained there until the subject finished reading. Different filler-placement conditions were presented in blocks of five paragraphs. The order of filler-placement conditions was randomized, with the constraints that the same condition could not appear in two consecutive blocks, and that blocks in all conditions were read once, before the next full set of conditions started. Only one type of filler was used in each experimental session.

The text used for sessions in which digits were used as fillers and reading was aloud was taken from the Scholastic Aptitude Test (SAT) preparatory books. Twelve five-paragraph stories were used. Mean number of words/story was 393 (SD = 33). Mean word length was five letters (SD = 0.5).

Assignment of stories to conditions was counter-balanced across subjects so that each story was used in each condition at least twice. *Post-hoc* readability analysis of the stories using the RightWriter 4.0 program (1990, MacMillan by Richard Schreinert) showed that the stories ranged in readability levels from 7 to 17 school-years. RightWriter calculated readability levels on the basis of number of words/sentence, number of syllables/word and word frequency. There was a small, but significant interaction between text condition and story readability level. Reading speed in more difficult conditions, such as unspaced text and text with fillers, was affected by story readability level to a larger extent than reading speed for normal text ($F(11,5) = 8.1$, $P < 0.001$). This interaction did not affect the pattern of results because (i) the effects were quite small (<10%); and (ii) the stories and conditions were counterbalanced among subjects.

For all subsequent sessions the issue of readability was avoided by using text from a single, fairly difficult novel: the "War of the Worlds" by H. G. Wells. Its mean word length was 4.6 letters (SD = 0.4). Although mean word length was lower than for the SAT stories, the novel had more complex syntax and was more advanced conceptually than most of the stories.

Fairly difficult text was used to avoid a ceiling effect introduced by irreducible time required to pronounce each word clearly. Furthermore, increasing text difficulty is known to slow reading of unspaced, filled and flanked text more than it slows the reading of ordinary text (see above), so the effects of these manipulations should be more pronounced than they would have been had easier text been used.

The text was displayed on the computer screen in white letters on blue background, using the default MS-DOS fixed-width Courier font (in text mode). Subjects sat at a "comfortable reading distance", so the angular size of letters varied from ≈18 to 25 minarc.

Procedure

Each session lasted from 1 to 1.5 hr. The session started with a practice story that illustrated the text conditions in that particular session (one condition per paragraph). Five multiple choice questions on the practice story followed. The experimenter then determined whether the subject read fluently on the basis of how well she read the paragraph of ordinary text. If the subject was judged to be a fluent reader, she continued in the experiment. If not, the subject was told that the experiment was completed, thanked and dismissed (see above).

Each trial began with the introductory sentence "Look at the asterisk and press space bar when ready" that appeared at the top of the screen. The message illustrated the filler-placement condition of the upcoming paragraph. For example, if the next paragraph was going to be unspaced, the introductory sentence "Lookatthe asteriskandpressspacebarwhenready" appeared. The asterisk mentioned in this sentence was located directly below its first letter. Subjects were instructed to read the introductory sentence silently, making sure that they could see all letters clearly, then to fixate the asterisk, and finally to press the space bar which started the trial. Immediately thereafter, the entire next paragraph appeared on the screen, its first letter positioned where the asterisk had been located. The subject read the

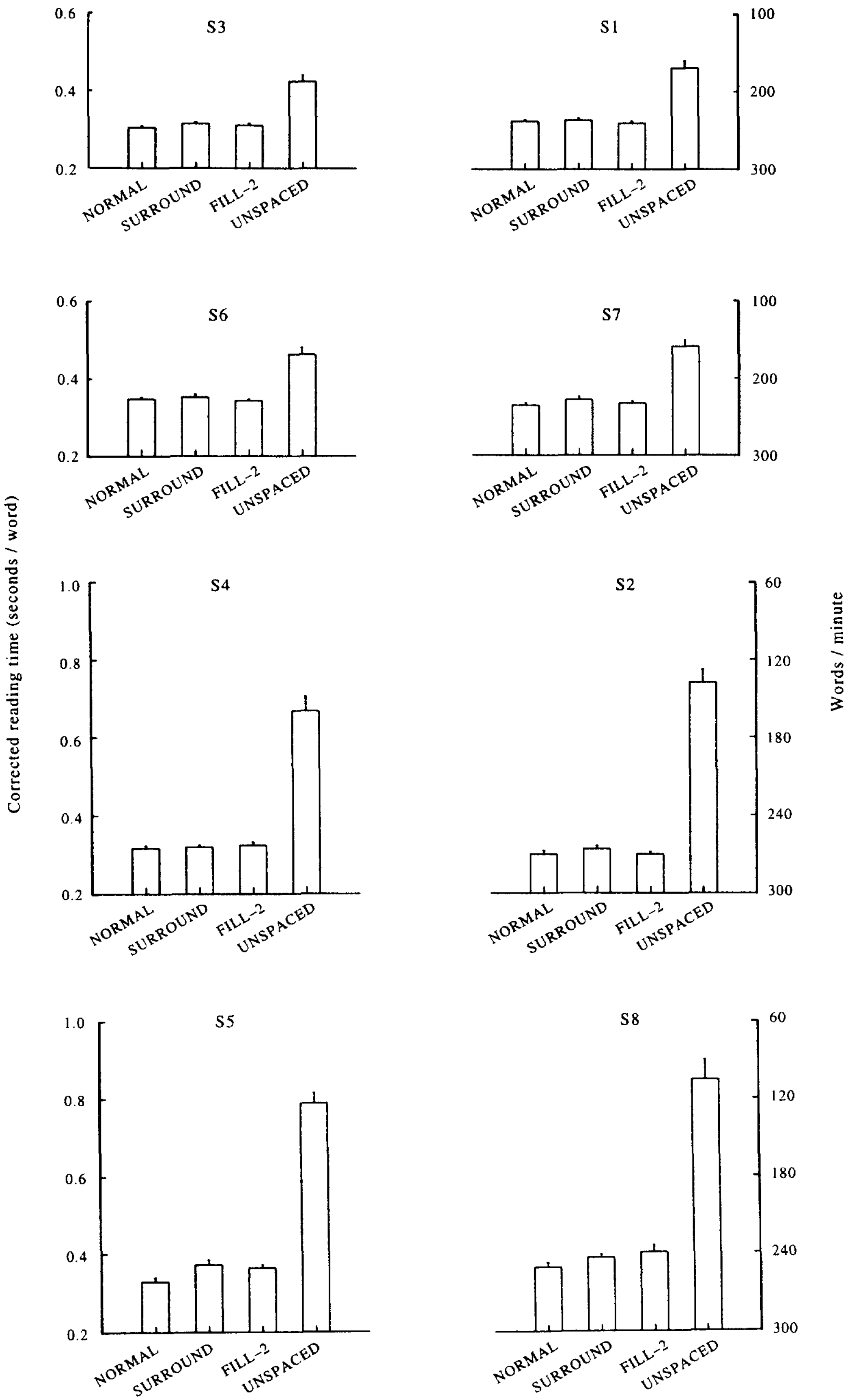

FIGURE 2. Mean corrected reading times (seconds/word on the left ordinate, words/minute on the right ordinate) for the four different text conditions in the shaded-box filler condition. Each graph shows data for an individual subject. Subjects are ordered from left to right and top to bottom, according to how fast they read unspaced text—from the fastest to the slowest. Each bar is based on 20 paragraphs. Error bars show 1 SE.

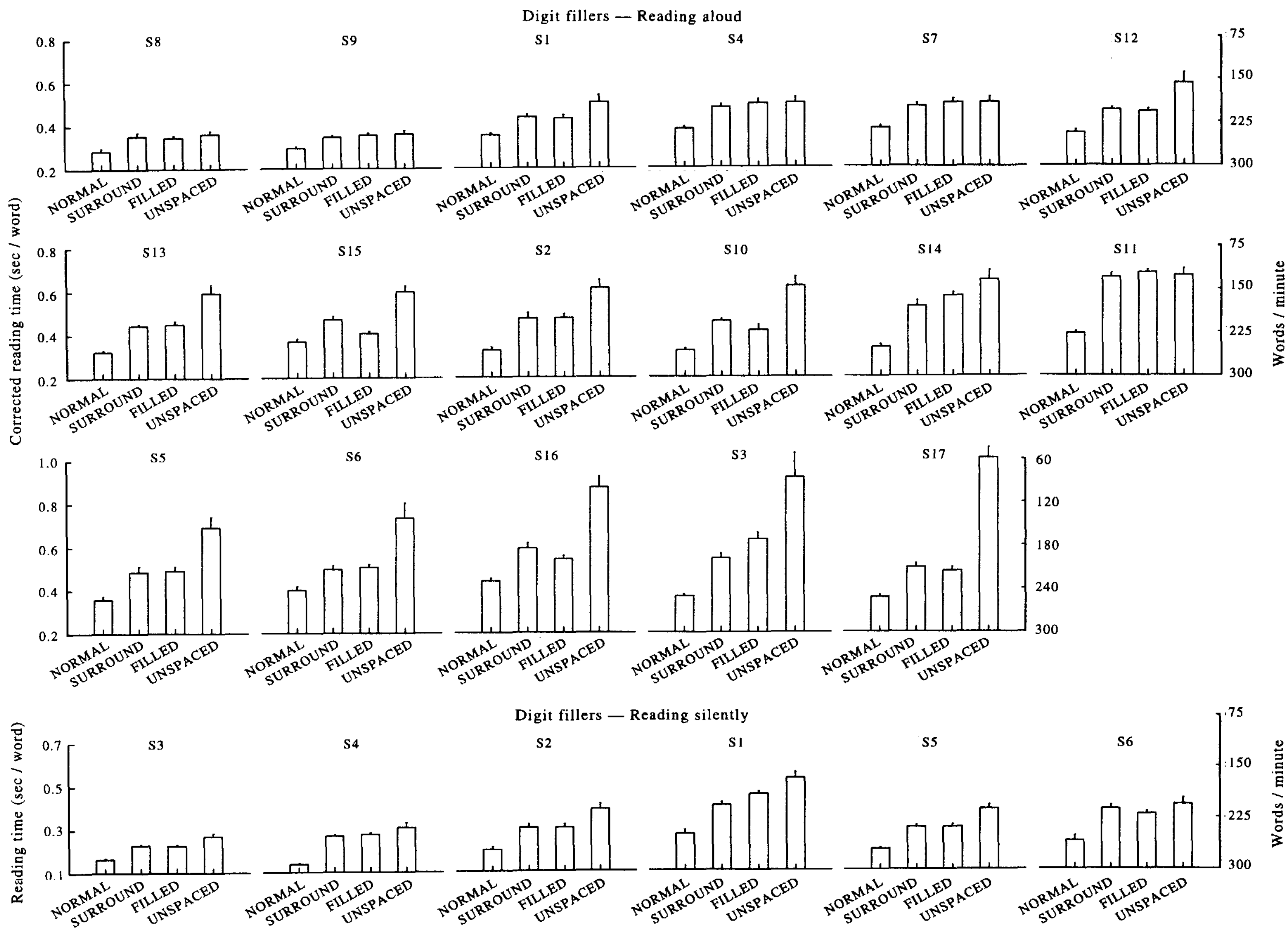

FIGURE 3. Mean corrected reading times (seconds/word on the left ordinate, words/minute on the right ordinate) for the six different text conditions in the digit filler condition. Each graph shows data for an individual subject. Subjects are ordered from left to right and top to bottom, according to how fast they read unspaced text—from the fastest to the slowest. Each bar is based on 10 paragraphs. Error bars show 1 SE.

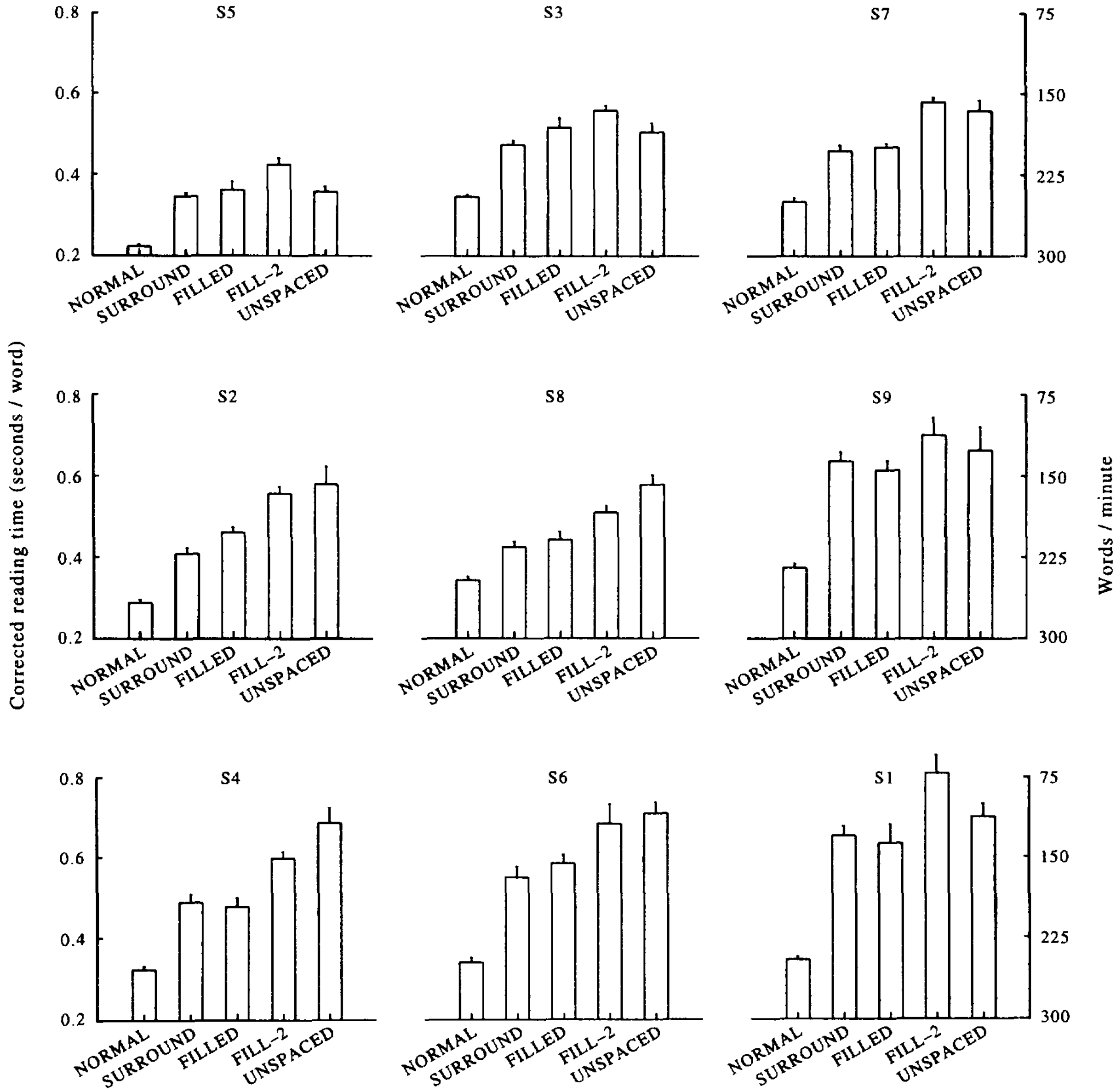

FIGURE 4. Mean corrected reading times (seconds/word on the left ordinate, words/minute on the right ordinate) for the five different filler placements in the Greek-letter filler condition. Each graph shows data for an individual subject. Subjects are ordered from left to right and top to bottom, according to how fast they read unspaced text—from the fastest to the slowest. Each bar is based on 10 paragraphs. Error bars show 1 SE.

paragraph pressing the space bar when finished. The time between the two presses of the space bar was recorded. When subjects read aloud, their speech was recorded and later scored for pronunciation errors. Subjects were instructed to read quickly, with expression and for comprehension when they read aloud, and to read quickly for comprehension when they read silently. Multiple-choice comprehension questions were presented either at the end of each five-paragraph story (Digit fillers/reading aloud), or at the end of the session.

RESULTS

The results for individual subjects in the four filler-type conditions are presented in Figs 2–5. Figure 6 presents group data by showing mean differences (in sec/word) in reading times between reading normal text and reading texts in the main experimental conditions, averaged across subjects. Percent differences averaged across subjects are summarized in Table 2.

The following effects were observed.

Normal vs unspaced text

Only 1 of 48 subjects read unspaced text as quickly as he read normal text. All other subjects read unspaced text more slowly ($P < 0.05$). The % difference in mean reading times between the unspaced and normal conditions ranged widely, from 13 to 64%. Mean percent difference between normal and unspaced reading times was 44% for reading aloud and 48% for reading silently.

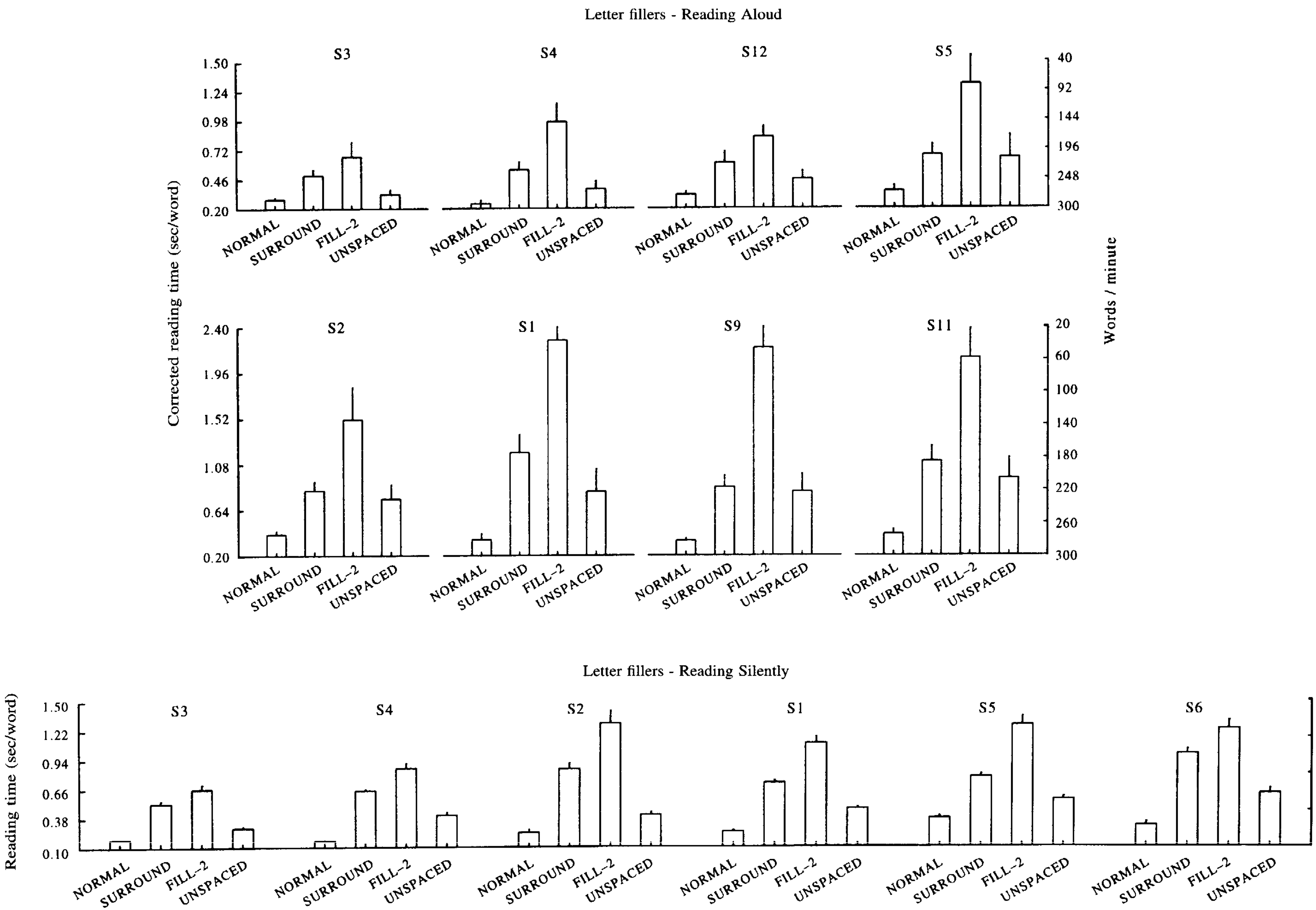

FIGURE 5. Mean corrected reading times (seconds/word on the left ordinate, words/minute on the right ordinate) for the six different placements of fillers in the Latin-letter filler condition. Each graph shows data for an individual subject. Subjects are ordered from left to right and top to bottom, according to how fast they read unspaced text—from the fastest to the slowest. Each bar is based on 10 paragraphs. Error bars show 1 SE.

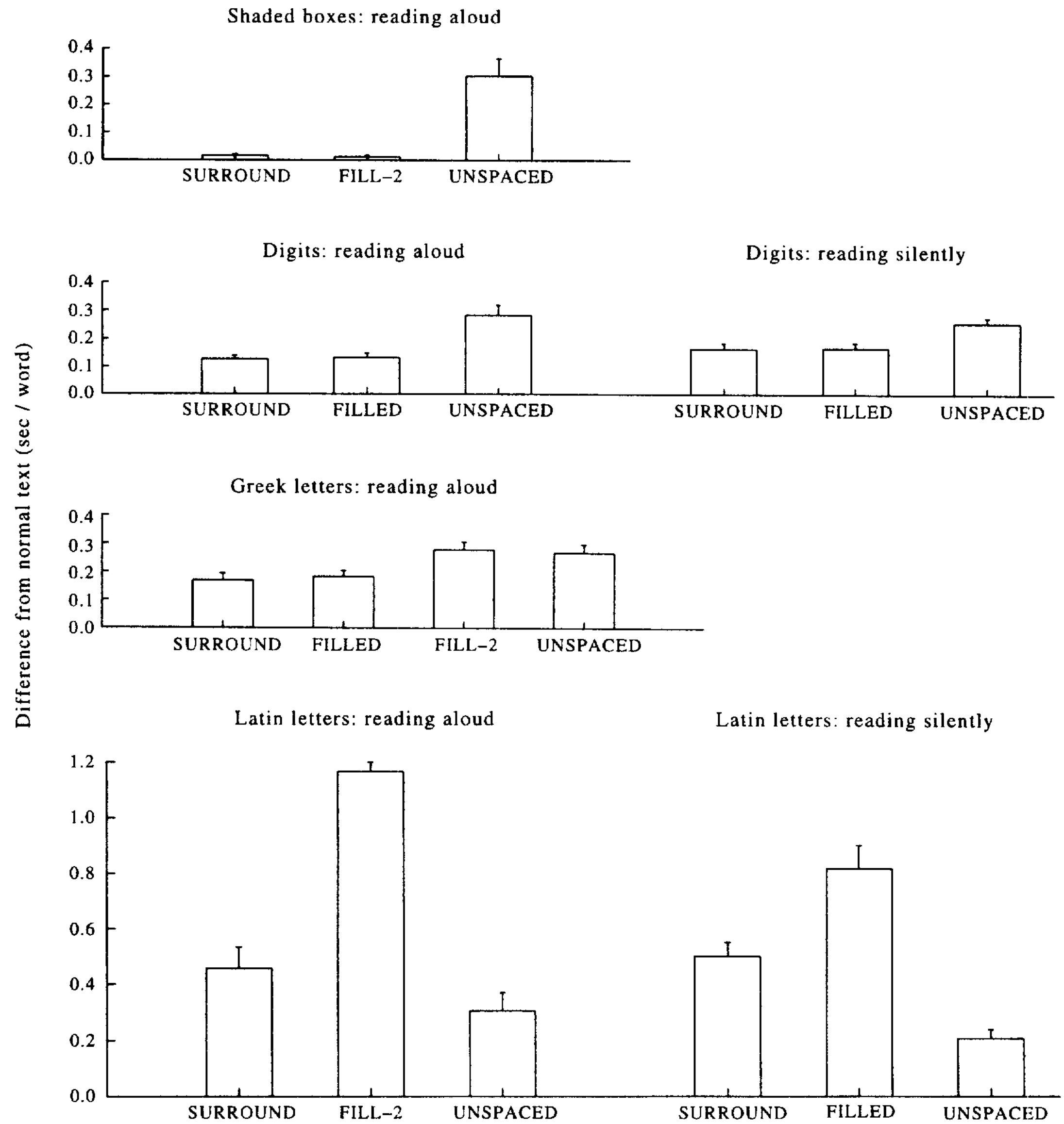

FIGURE 6. Mean difference in reading time (sec/word) between normal text and other conditions. The data are averaged across all subjects. Each bar is based on means computed for 6–17 subjects. Error bars show 1 SE.

Shaded boxes

Figure 2 shows the "corrected reading times" with shaded boxes as fillers. A Tukey HSD multiple comparisons t-test, performed for each subject, showed that the corrected reading times in the surround and fill-2 conditions did not differ significantly from corrected reading times for normal text ($P > 0.3$) or from each other ($P > 0.9$).

This result differs from the result reported by Pollatsek & Rayner (1982). They found that filling spaces with

TABLE 2. Mean percent differences in corrected reading times

Filler type	n	Surround	Filled	Fill-2	Unspaced
Shaded boxes (aloud)	8	$\bar{x} = 4\%^{ns}$ (1–12)	———	$\bar{x} = 3\%^{ns}$ (−2–9)	$\bar{x}$ = 44% (25–60)
Digits (aloud)	17	$\bar{x}$ = 26% (15–40)	$\bar{x}$ = 26% (10–42)	———	$\bar{x}$ = 42% (19–64)
Digits (silently)	6	$\bar{x}$ = 40% (32–55)	$\bar{x}$ = 41% (32–58)	———	$\bar{x}$ = 52% (45–63)
Greek letters (aloud)	9	$\bar{x}$ = 33% (20–47)	$\bar{x}$ = 36% (23–45)	$\bar{x}$ = 46% (33–57)	$\bar{x}$ = 44% (32–53)
Latin letters (aloud)	8	$\bar{x}$ = 55% (44–70)	———	$\bar{x}$ = 74% (58–84)	$\bar{x}$ = 43% (13–58)
Latin letters (silently)	6	$\bar{x}$ = 66% (51–75)	$\bar{x}$ = 75% (68–81)	———	$\bar{x}$ = 45% (33–60)

Each $\bar{x}$ shows the mean of the individual subjects' mean percent differences in reading times between the particular text condition (column) and "normal" text. Each row show the results for a particular type of filler. The intersubject range of mean percent differences in reading times are given in parentheses. See text for a complete explanation. ^{ns} indicates that the differences in corrected reading time between the particular text condition and normal text was not statistically significant for any of the subjects.

gratings (similar to our shaded boxes according to their Fig. 4), slowed reading by about 30%—almost by the same amount as filling spaces with random digits slowed reading in their experiment. We did not observe any effect when spaces were filled with shaded boxes, but did observe a statistically significant increase in reading time (10–41%, mean = 26%) when spaces were filled with random digits (see below). Our result with shaded boxes is similar to the result reported by Malt & Seamon (1978), who found that using red boxes to fill spaces had no effect on reading speed. Our shaded boxes might have been more similar to the red boxes used by Malt & Seamon (1978) than to the gratings used by Pollatsek & Rayner (1982). It is not possible to say more about this because Pollatsek & Rayner (1982) did not define their "gratings". Ours were ASCII character 176.

Another similarity between the present and the Malt & Seamon (1978) experiments is that both used paragraphs of coherent text that were not perturbed while subjects read. In contrast, Pollatsek & Rayner (1982) used single sentences and introduced fillers, contingent of fixation position, while the subject was reading. It is likely that reading a continuous, unperturbed story, rich in meaning, would facilitate reading speed. It also seems likely that highly discriminable additions to the text like shaded boxes and red squares would be less disturbing when an unchanging text is read than when these characters appear during reading, the technique used by Pollatsek & Rayner (1982). This perturbation technique has been criticized by O'Regan (1990) and Epelboim *et al.* (1996).

Digits

Figure 3 shows mean reading times for individual subjects in the four most interesting text conditions. Performance varied considerably among subjects, but the pattern of differences among conditions was similar for all subjects and for reading aloud and reading silently (bottom row). A Tukey HSD multiple comparisons *t*-test performed separately on each subject's data compared pairs of conditions with the following results.

Only 4 of the 17 subjects slowed down significantly in the "begin" condition when compared with the "normal" condition ($P < 0.05$). Two of these subjects also slowed down significantly in the "end" condition when compared with the "normal" condition ($P < 0.05$). The percent differences in mean corrected reading times between the "begin" and "end" condition and the "normal" ranged from 0 to 27%. Inserting digits at the beginnings of words had the same effect on reading time as inserting digits at the ends of words. None of the 17 subjects showed a significant difference in corrected reading times between the "begin" and the "end" conditions ($P > 0.7$).

Inserting digits at *both* ends of each word had a greater effect on reading time than inserting a digit at only one end of each word. Fifteen of 17 subjects who read aloud and all six subjects who read silently slowed down significantly in the surround condition when compared with the "normal" condition ($P < 0.05$). The percent difference in mean corrected reading times between the "surround" and "normal" conditions ranged from 15 to 55% (mean aloud = 26%; mean silently = 40%).

Twenty of 23 subjects slowed down significantly ($P < 0.05$) when spaces were filled with digits ("filled"). The % differences in mean corrected reading times between the "filled" and the "normal" conditions ranged from 10 to 58% among the subjects (mean aloud = 26%; mean silently = 41%).

Not one of the 23 subjects reading aloud and silently showed a statistically significant difference in corrected reading time between the "surround" condition and the "filled" condition. Only one subject (S15) even came close to having a significant difference in corrected reading time between these two conditions ($P = 0.079$). This subject's reading time was 13% *faster* in the "filled" condition than in the "surround" condition. The differences in corrected reading times between these two conditions were not significant for the other 22 subjects (all $P > 0.2$).

Greek letters

Figure 4 shows mean corrected reading times with Greek-letter fillers. Subjects' performances varied, but the pattern of results was similar for all nine. A Tukey HSD Multiple comparisons *t*-test performed separately on each subject's data, compared pairs of text conditions with the following results.

All nine subjects slowed down significantly ($P < 0.05$) when each word was surrounded by Greek letters with the spaces preserved ("surround"). The difference in corrected reading time between the "surround" and "normal" conditions ranged from 20 to 47% (mean = 33%). This effect was somewhat greater than the effect observed in the "surround" condition with digit fillers (15–40%, mean = 26%), but smaller than the effect observed in the "surround" condition with Latin-letter fillers (44–70%, mean = 55%).

All nine subjects slowed down significantly ($P < 0.05$) when each space was filled with a single Greek letter ("filled"). The percent differences in mean corrected reading times between the "filled" and "normal" conditions ranged from 23 to 45% (mean = 36%). This effect was somewhat greater than the effect observed in the "filled" condition with digit fillers (10–42%, mean = 26%).

All nine subjects slowed down significantly ($P < 0.001$) when each space was filled with two *different* Greek letters ("fill-2"). The percent differences in mean corrected reading times between the "fill-2" and "normal" conditions ranged from 33 to 57% (mean = 46%). Reading was significantly slower in the "fill-2" condition than in the "filled" condition for six of the nine subjects ($P < 0.05$). The percent differences ranged from 9 to 23%.

Not one of the nine subjects showed a statistically significant difference in corrected reading time between the "surround" condition and the "filled" condition (all $P > 0.3$), as was the case with digit fillers. Two Greek-letter fillers had a greater effect on reading time. Eight of the subjects read significantly more slowly in the "fill-2" condition than in the "surround" condition ($P < 0.05$), but percent differences in corrected reading time between these conditions were relatively modest (13–27%, mean = 18%).

Latin letters

Figure 5 shows mean corrected reading times for the four most interesting placements of fillers. Intersubject variability was considerable, but the pattern of results was similar for all eight subjects. A Tukey HSD multiple comparisons *t*-test was performed separately on each subject's data with the following results.

Subjects slowed down by 14–75% (mean = 40%) from their normal corrected reading times in the "begin" condition. For three of the eight subjects, the difference between corrected reading times in the "normal" and "begin" conditions was statistically significant ($P < 0.05$). The corrected reading times were significantly slower in the "end" condition than with "normal" text for all eight subjects ($P < 0.05$). This difference ranged from 22 to 49% (mean = 38%). Although none of the subjects showed a significant difference in corrected reading time between the "begin" and "end" conditions ($P > 0.09$), all but one subject (S1) tended to read more slowly when letters were inserted at the ends of words than when letters were inserted at the beginnings of words.

Inserting randomly chosen letters at both ends of each word ("surround") slowed reading much more than inserting a letter at either end of each word ($P < 0.05$). All subjects read significantly more slowly in the "surround" condition than in the "normal" condition ($P < 0.05$). The percent differences in mean corrected reading time between these two conditions ranged from 44 to 75% (mean aloud = 55%; mean silently = 66%). Surrounding each word with randomly chosen letters, while preserving spaces, had the same effect on corrected reading times as removing spaces from normal text ("unspaced") for four of the eight subjects reading aloud ($P > 0.1$). The remaining four subjects reading aloud, as well as all six subjects reading silently, slowed down more in the "surround" condition than in the "unspaced" condition (all $P < 0.05$).

The greatest challenge to reading for all subjects was observed when spaces were filled with either one or two Latin letters ("filled" and "fill-2" conditions). Reading aloud in the "fill-2" condition was 58–84% (mean = 74%) slower than reading "normal" text, 26–61% (mean = 45%) slower than reading text in the "surround" condition, and 46–64% (mean = 55%) slower than reading "unspaced" text ($P < 0.001$).

Reading silently in the "filled" condition was: (i) 68–81% (mean = 75%) slower than reading "normal" text; (ii) 21–41% (mean = 29%) slower than reading text in the "surround" condition; and (iii) 49–68% (mean = 56%) slower than reading "unspaced" text ($P < 0.001$).

The main results, averaged over all subjects, are summarized in Fig. 6. These summaries are fairly representative of the individual performance patterns observed.

DISCUSSION

Table 2 summarizes the results by listing mean and range in percent differences in reading times for different filler-placements and filler-types. It is clear that the size of the effect of fillers on reading time depended on the effect of the particular filler-type on word recognition, determined empirically for individual words in a word recognition experiment (see Method and Fig. 1).

Shaded-box fillers did not slow recognition of individual words. They also had no statistically reliable effect on reading time. Digits had a reliable effect on reading time, but this effect was smaller than the effect of Greek letters, which had a greater effect on recognition of individual words than digits. As would be expected from the results of our word recognition experiment, (as well as intuitively), the largest effect of fillers on reading time was observed when Latin letters were used.

The remainder of this section will consider how fillers could have slowed reading either by obscuring word-length information and thereby disturbing the eye movement pattern, or by slowing word recognition directly. It will be argued that the primary effect of fillers in text is to disrupt word recognition, rather than to disrupt low-level visual processes postulated as being used for programming reading saccades.

Reading eye movement pattern

Likely effects of our conditions on the reading eye movement pattern will be discussed notwithstanding the fact that eye movements were not recorded. This is reasonable because much is already known about the reading eye movement pattern (see O'Regan, 1990, for a review). It is known that during reading the first fixation within each word tends to fall on what has been called the "preferred landing position" (PLP, Rayner, 1979), which is located somewhat left of the center of each word. The PLP is believed to coincide with what is called the "optimal viewing position" (OVP). The OVP is the position within a word that permits the fastest recognition of that word. Fixating a position other than the OVP makes it take longer to recognize the word. This effect of landing position on the latency of word recognition has been shown to be important for isolated words (O'Regan *et al.*, 1984; O'Regan & Jacobs, 1992), and possibly present in a much weaker form when words within a normal, continuous text are read (Vitu *et al.*, 1990). Despite the weakness of the OVP effect in continuous text, O'Regan (1990) proposed that the goal of saccadic programming during reading is to aim each saccade into an upcoming word so as to land near its OVP. The relative success of this strategy is an important determinant of reading speed. O'Regan's (1990) theory asserts that spaces between words are used to obtain information about the length of the next word in the text. This information is used to guide each reading saccade to its OVP. If a saccade fails to fall on the OVP, word

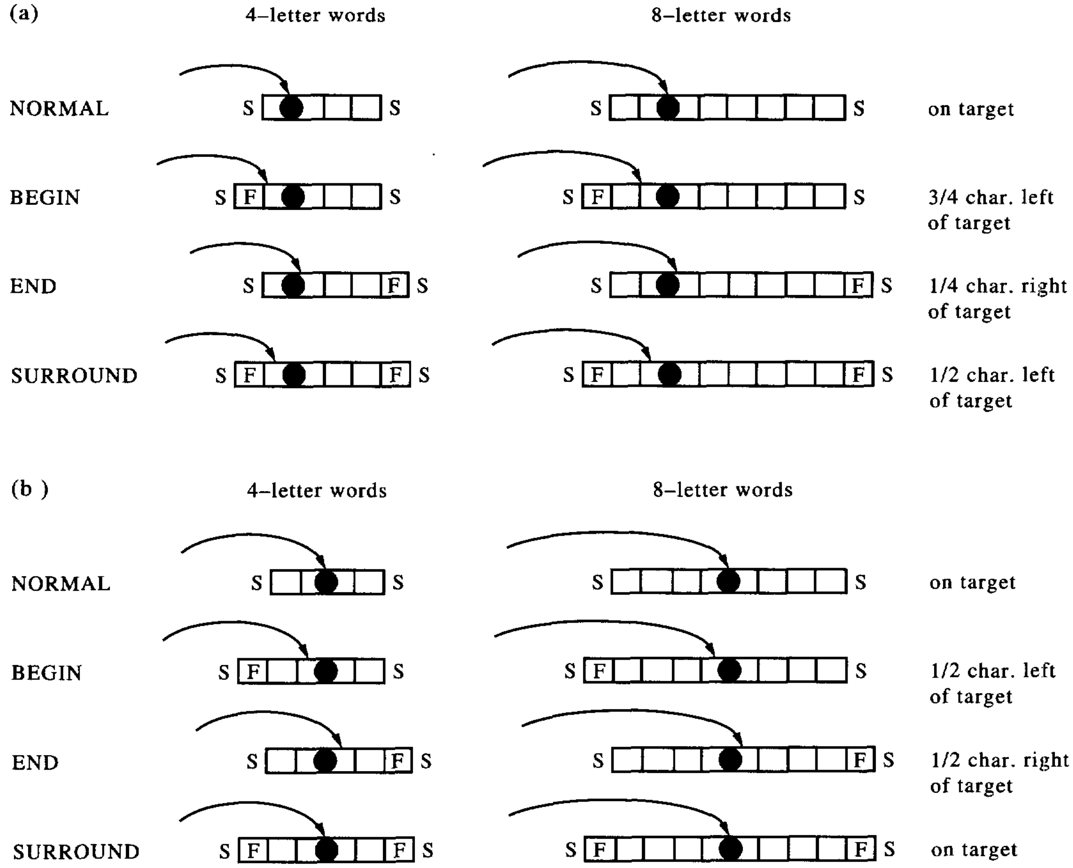

FIGURE 7. Possible effect of fillers (F) on the landing positions of saccades for four-letter words (left) and eight-letter words (right). The diagram is based on the assumption that the saccades use spaces (S) to detect the length of the upcoming word and aim at the OVP, which is assumed to be located halfway between the word's beginning and its center (a) or at the word's center (b). The assumed position of the OVP within the word is indicated by filled circles.

recognition is slowed, or, a refixation to compensate for this error is made. Both would slow reading.

Figure 7 illustrates where saccades, programmed according to an OVP-based theory, will land when fillers (F) are inserted into the various locations used in our experiment. Figure 7(a) shows that if the OVP is assumed to be located halfway between the word's center and its beginning, saccades miss the OVP by 3/4 character to the left in the "begin" condition, 1/4 character to the right in the "end" condition and 1/2 character to the left in the "surround" condition. Figure 7(b) shows that if the OVP is assumed to be located at the word's center, saccades would miss the OVP by 1/2 character to the left in the "begin" and "end" conditions. Saccades would be accurate in the "surround" condition.

Regardless of the assumed location of the OVP, if saccades are aimed on the basis of spaces between words, they would frequently miss the OVP by an unpredictable number of characters in all conditions where words are not separated by spaces ("filled", "fill-2" and "unspaced").

The relationships between reading times for our different text conditions can be predicted from these considerations by assuming that reading speed is determined, at least in part, by the absolute distance between the OVPs of the words and the actual landing locations of the reading saccades. Specifically, reading text without spaces should be slower than reading text with spaces because most saccades would miss the OVP by unpredictable distances. The OVP would also be missed in conditions in which both fillers and spaces were present, but these errors would be quite small (less than 1 character, see Fig. 7) and completely predictable. Furthermore, in these conditions, saccades could adapt easily to the new pattern because adaptation would be the same for each word. So, according to the OVP-based theory, reading text with spaces, even when fillers are present, should be easier than reading text in which spaces are either filled or removed.

Note: it is possible that any manipulation that makes reading very difficult could overshadow the effects of missing the OVP. For example, lowering the contrast so as to make the text barely legible, or, inserting randomly selected Latin letters next to words, could cause so great a detriment to reading that the OVP would become irrelevant. In other words, if the subject must struggle to make out the text, one letter at a time, it does not really matter where the line of sight initially falls within each word. This type of letter-by-letter reading may have been used for particularly difficult passages by a few of our subjects in our Latin-letter filler conditions. In texts in which spaces were filled with digits or with Greek letters, subjects did not have to make out each word letter-by-letter. Most subjects read such texts fluently, pronouncing whole words or phrases at a time, albeit slower than when they read ordinary text. This fact suggests that the role of the OVP in texts with digit or Greek-letter fillers should be the same as it is in ordinary text.

The pattern of results observed does not support predictions of the OVP-based theory. The effects of different text conditions on reading depended more on the particular type of filler used than on the presence or absence of spaces.

Shaded-box fillers, which had no effect on word recognition, also had no effect on reading speeds in any condition, despite the fact that they altered word spacing, which should have affected reading saccades (see Fig. 7). However, shaded boxes look more like spaces than letters, so it is possible to argue that word-length information could have been obtained by low-level visual processes, even when spaces were filled with shaded boxes.

Digits and Greek letters at both ends of words slowed reading by the same amount as when these symbols filled the spaces between the words. Although one might still argue that these symbols could function as spaces, and be detected by low-level visual processes, this argument is less than convincing, especially in the case of Greek letters, which share many features with letters making up ordinary text. Furthermore, one must explain why two Greek letters filling a space were less likely to be detected than just one Greek letter. If words are segmented using low-level visual features, such as spatial frequency, or texture gradient, two Greek letters should be at least as detectable as one. Actually, a pair of Greek letters between words should be better for programming saccades because only one of the letters would have to be detected in order to segment a word, and the visual system has two chances of detecting one. However, texts in which spaces were filled with two Greek letters were read *more slowly* than texts in which spaces were filled with just one Greek letter, a result clearly counter to this prediction.

Effects of fillers on reading times were greatest when randomly chosen Latin letters were used, but, again, the pattern of effects was not consistent with the pattern predicted by the OVP-based theory. Inserting extra letters, either at the beginning or at the end of each word, had the smallest effects on reading times. The greatest effect of fillers on reading times was in conditions in which spaces were filled with one or two randomly chosen Latin letters. These unspaced conditions slowed reading most, but it is unlikely that this slowing was caused by the lack of spaces *per se*, because the second slowest reading was observed when spaces were present but each word was flanked on both sides by a randomly chosen letter. This manipulation slowed reading as much as, or more than, removing spaces from normal text. Furthermore, surrounding words with Latin letters while preserving spaces, slowed reading by 55% (reading aloud) and by 66% (reading silently). Removing spaces from this condition slowed reading by only 10–20% more.

This pattern of results suggests that the biggest difficulty did not lie in estimating word lengths, but, rather, in separating words from the characters surrounding them. If spaces were used to program saccades, reading should be easier in the "surround" condition because strings of letters separated by spaces present clear targets for reading saccades. Once each saccade landed, all the reader had to do was to ignore the first and the last character (be it Latin or Greek letter or a digit) of each string of characters. In the conditions without spaces (with or without fillers), however, the words were not delimited, so saccades would be likely to miss their intended targets, and each word would still have to be parsed from the letters surrounding it. Reading text with words surrounded by extra letters was 10–20% slower than reading text without spaces and without fillers, and only 10–20% faster than reading texts in which spaces were filled with Latin-letter fillers. This pattern shows that having clear saccadic targets affords only a small advantage (at most 20%) whereas extraneous letters flanking words are very detrimental (at least 44%). Furthermore, when flanking symbols are digits or Greek letters, having spaces between words affords no advantage whatsoever, but flanking words with these symbols still causes a detriment.

These results, taken together, suggest that fillers surrounding words affect word recognition directly, without necessarily having any effect on the eye movement pattern. Possible effects of fillers on word recognition will be discussed next.

Word recognition

One plausible mechanism that could account for the detrimental effect fillers have on reading speed, is the kind of lateral interaction (masking) reported by Bouma (1970), who showed that letters, presented some distance away from central fixation, took longer to recognize when they were flanked by other letters, than when they were presented alone, especially when the flanking letters shared features with the target letter. The fact that the degree of similarity between adjacent letters increased the size of this interaction suggests that a similar interaction may have been, at least in part, responsible for the pattern of our results. It seems reasonable to propose that the effects of our fillers (flanking symbols) increased as the similarity between filler symbols and letters of the text increased. Shaded boxes shared no features with letters, and had no effect on reading times, or on the recognition of individual words. The digits used shared more features with letters, and therefore slowed reading more (1 and 0 were excluded for just this reason). Greek letters are still more similar to Latin letters, and predictably, had still greater effects. The greatest effect was observed with randomly chosen Latin letters, often identical to and always very similar to the adjacent letters making up the words of the text. Bouma (1970) studied these masking effects with only single letters, but such lateral interactions may have operated in our experiments, where they would slow word recognition by delaying the recognition of letters adjacent to fillers, differentially on the basis of the similarity between the filler and the letter.

Automatic word recognition processes, like those observed in the Stroop effect (Stroop, 1935) provide another plausible hypothesis of how fillers can affect word recognition, especially when randomly chosen Latin letters are used. When a word made up of Latin letters is surrounded by extra Latin letters, the resulting string of letters is read "as is", automatically. The extra letters cannot be ignored easily, just as the word naming the inappropriate color cannot be ignored in the classic Stroop task. A similar phenomenon was observed by O'Hara (1980), who asked subjects to match two letters flanking either a word or an anagram of a word ("a word b" or "a dwor b"). The subjects performed the matching task faster when the string between the letters, which had nothing to do with the task, was a word, than when it was an anagram, unless the word was incompatible with the task, because it sounded like a name of a letter (e.g., "b tea B"). Subjects were not able to ignore the word in the middle, despite instructions to do so. Such automatic word recognition processes probably hindered reading when fillers flanked words in our experiments, even when the flanking symbols were not actual Latin letters. Similarity between filler-symbols and letters could be sufficient to activate such automatic word recognition processes.

In unspaced text, the same kinds of automatic word recognition processes may actually have helped the reader. Familiar, highly practiced words and words predictable through meaning conveyed by context may have "jumped out" at the reader, despite the absence of spaces that delimit words in ordinary text. Of course, occasionally, letters from around word boundaries would be grouped incorrectly in unspaced text, and form words that could jump out and be read automatically. Such words, however, would not fit into the semantic and syntactic structure of the text being read. The reader would detect such inconsistencies and be forced to spend time figuring out how to correct the now obvious error by trying other grouping of letters. Incorrect grouping of letters could also account for the fact that reading was slowest when spaces were filled with two different letters, either Latin or Greek. Reading was slower with two fillers than with just one because two extraneous Latin letters, or two Greek letters mistaken for Latin letters, provide many more possible incorrect groupings simply because the larger the total number of letters, the more different ways they can be grouped. The detrimental effect of changing familiar letter groupings on reading has been studied (see Kolers, 1968; Kowler & Anton, 1987).

In summary

The pattern of results observed suggests that fillers slow reading because they affect the recognition of words directly, not because they obscure word-length information, presumed to be necessary for an effective reading eye movement pattern. The effect of fillers on reading eye movements are now under study. Preliminary analyses of eye movement data obtained in experiments with filler conditions like those reported here tend to support the notion that the presence of fillers does not have any first-order effects on the local properties of the reading eye movement pattern (Epelboim *et al.*, 1996b).

We suggest that word-length obtained by a low-level visual process that detects spaces, if used at all, is only one of many sources of information used to program efficient reading saccades. Other sources of information may include (i) global saccadic patterns adjusted on the basis of global text properties (such as conceptual or visual difficulty); and (ii) information about words obtained by using fast word recognition processes, or by predicting them on the basis of the context of the text. Once it is allowed that many sources of information can be used for programming reading saccades under unusual circumstances or in languages that do not use spaces, it becomes simpler to assume that this kind of information is also used when ordinary, spaced text is read. Assuming different saccadic programming strategies are needed to read spaced and unspaced text seems unnecessary.

REFERENCES

Booth, J. R., Epelboim, J. & Steinman, R. M. (1995). The relative importance of spaces and meaning in reading. *Proceedings of the Cognitive Science Society, 17*, 533–538.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177–178.

Epelboim, J., Booth, J. R., Ashkenazy, R., Taleghani, A. & Steinman, R. M. (1995). A comparison of fillers and spaces in text: evidence for the importance of word recognition. *Investigative Ophthalmology and Visual Science (Suppl.), 35*, B708.

Epelboim, J., Booth, J. R. & Steinman, R. M. (1994). Reading unspaced text: implications for theories of reading eye movements. *Vision Research, 34*, 1735–1766.

Epelboim, J., Booth, J. R. & Steinman, R. M. (1996a). Much ado about nothing: the place of space in text. *Vision Research, 36*, 465–470.

Epelboim, J., Booth, J. R., Taleghani, A., Ashkenazy, R. & Steinman, R. M. (1996b). Fillers and spaces in text: implications for the relative importance of word recognition and physical features of the text during reading. *Perception, 25 suppl.*, 12–13.

Katz, S., Lautschlager, G. J. & Blackburn, A. B. (1990). Answering reading comprehension items without passages on the SAT. *Psychological Science, 1*, 122–127.

Kolers, P. A. (1968). The recognition of geometrically transformed text. *Perception and Psychophysics, 3*, 57–64.

Kowler, E. & Anton, S. (1987). Reading twisted text: implications for the role of saccades. *Vision Research, 27*, 45–60.

Legge, G. E., Pelli, D. G., Rubin, G. S. & Schleske, M. M. (1985). Psychophysics of reading—I. Normal vision. *Vision Research, 25*, 239–252.

Legge, G. E. & Rubin, G. S. (1986). Psychophysics of reading: IV. Wavelength effects in normal and low vision. *Journal of the Optical Society of America A, 3*, 40–51.

Legge, G. E., Rubin, G. S. & Luebker, A. (1987). Psychophysics of reading—V. The role of contrast in normal vision. *Vision Research, 27*, 1165–1177.

Malt, B. C. & Seamon, J. G. (1978). Peripheral and cognitive components of eye guidance in filled-spaced reading. *Perception and Psychophysics, 23*, 399–402.

Morris, R. K., Rayner, K. & Pollatsek, A. (1990). Eye guidance in reading: the role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 268–281.

O'Hara, W. P. (1980). Evidence in support of word unitization. *Perception and Psychophysics, 27*, 390–402.

O'Regan, J. K. (1990). Eye movements and reading. In Kowler, E. (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 395–453). Amsterdam: Elsevier.

O'Regan, J. K. & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: a challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 185–197.

O'Regan, J. K., Lévy-Schoen, A., Pynte, J. & Brugaillere, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 250–257.

Pollatsek, A. & Rayner, K. (1982). Eye movement control in reading: the role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 817–833.

Rayner, K. (1979). Eye guidance in reading: fixation locations within words. *Perception, 8*, 21–30.

Rayner, K. (1993). Eye movements in reading: recent developments. *Current directions in psychological science, 2*, 81–85.

Rubin, G. S. & Turano, K. (1992). Reading without saccadic eye movements. *Vision Research, 32*, 895–902.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Vitu, F., O'Regan, J. K. & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception and Psychophysics, 47*, 583–600.

Acknowledgements—Some of the results from this study were presented at the annual meeting of the Association for Research in Vision and Ophthalmology (Epelboim *et al.*, 1995). This research was supported by NIMH grants 5-F32-MH11282-02 and 5T-32-MH191-02 and by AFOSR grant F49620-94-1-0333.