Perception & Psychophysics
1979, Vol. 25 (6), 501-509

Saccade size control in reading: Evidence for the linguistic control hypothesis

KEVIN O'REGAN
Centre d'Étude des Processus Cognitifs et du Langage
Centre National de la Recherche Scientifique, Paris, France

In Experiment 1, it is shown that during reading the eye makes larger saccades near long words than near short words. The effects are reduced when the subject's peripheral vision is diminished by the use of a moving "window" centered on the subject's fixation point, outside of which letters are replaced by Xs. In Experiment 2, it is shown that even if linguistic predictions are kept constant, the eye tends to make longer jumps when approaching THE than when approaching a three-letter verb. This "THE-skipping" effect is weaker if THE is compared with an auxiliary (HAD, WAS, or ARE) than if it is compared with a less frequently occurring verb (ATE, RAN, MET). It follows that knowledge of the lexicon can combine with information from peripheral vision fast enough to influence saccade size from moment to moment.

Until recently, many authors have supposed that during reading, the eye scans the text in a fairly rhythmical way. Under this "rhythmical scan" hypothesis, fixation durations are assumed to be approximately constant and saccade sizes are assumed to adjust only slowly to overall difficulty. It is assumed that these adjustments are made so that, on average, the rate of intake of information matches the rate at which it is processed (Bouma & de Voogd, 1974; Kolers, 1976; Morton, 1964; Woodworth & Schlossberg, 1954). Rayner (1975) has shown that the rhythmical scan hypothesis must be rejected as concerns fixation durations, since these vary rapidly from moment to moment as a function of the kind of processing going on during reading. However, saccade size variations have not received direct attention. Abrams and Zuber (1972) suggested that the eye skips abnormally large blanks placed at random positions in a text, but they did not evaluate this observation statistically. Some authors have found that the number of fixations falling on a word depends on its length (Rayner & McConkie, 1976) or on its syntactic category or position in the sentence (Communale, 1973; Klein & Kurkowski, 1974; Mehler, Bever, & Carey, 1967; Rayner, 1975; Wanat, 1976). However, these studies consider only the overall number of fixations falling on different parts of a sentence. One cannot conclude from them that the eye is adapting rapidly to the text, since the differences found may only be generated after an initial rhythmical scan, when the eye comes back to regions of particular difficulty.

The potential use of instantaneous saccade size changes as a clue to ongoing linguistic processing in reading makes it interesting to try to find instances were saccade sizes do in fact adjust from moment to moment. As a first step in this direction, the present paper presents evidence that saccade sizes vary systematically as a function of word length. It also attempts to show that these variations are probably caused by ongoing linguistic processing ("linguistic control"), and not by an eye-movement strategy such as "fixate each word once" which would require no linguistic processing to be done ("length control"). As pointed out by Rayner (1979), eye-movement control hypotheses have never been very precisely formulated. It is therefore worthwhile to try to distinguish these two kinds of control better. However, before this is done, it is necessary to consider what information is available at each fixation in reading.

Visibility and Perceptibility

Central vision permits unambiguous identification of two to three letters to the right and left of the fixation point. This is true more or less independently of viewing distance.[1] From the periphery, fairly good information about word length is available up to 21 letters on either side of the fixation point (Schiepers, 1976). While some information is also available concerning letters flanked by spaces, for reasons connected with lateral masking (Bouma, 1973) or acuity (Banks, Bachrach, & Larson, 1977), only crude clues are available concerning letters flanked by other letters. The quality of these clues also gets worse as we move further into the periphery.

I would like to thank Don Norman and the members of the LNR group (in particular, J. McClelland, D. Rumelhart, and A. Stevens) for their hospitality and the use of their equipment during my stay at the University of California at San Diego, where Experiment 2 was done in 1978. I would also like to thank J. Morton for help with Experiment 1, and Martha Teghtsoonian for her valuable editorial comments after submission of the manuscript. Queries and reprint requests should be sent to me at Laboratoire de Psychologie, 54 Boulevard Raspail, 75006 Paris, France.

0031-5117/79/060501-09$01.15/0

Define the "visibility" of a given letter in a given position of the line as the probability of reporting that letter without making use of contextual constraints (that is, as if it were inserted in a random letter sequence instead of being a word). In general, the visibility will then take values near 1.0 in the 6-7-letter region surrounding the fixation point. Visibility will drop as letters move further into peripheral vision. Superimposed on this gradual decrease will be bumps of high visibility corresponding to the first and last letter of each word and to the more visible inside letters.

"Perceptibility" may be defined as the probability of correctly reporting a given letter of a word when, in addition to visual information, all possible lexical, syntactic, and semantic constraints are taken into account. Many different types of knowledge will contribute to perceptibility. However, O'Regan (1975) has argued that the strongest constraints available come from the reader's knowledge of the lexicon. From counts in the Kučera and Francis (1970) corpus of English, O'Regan estimated that for a word of any length, knowledge of its length, its first six or seven letters, and its last letter generally suffice to determine its identity uniquely. It follows that for a fixation point near the second or third letter of a word of any length, the perceptibility will be near 1.0 for the whole word, even if it is long.

Looking now at information in peripheral vision, if we assume that first and last letter and length of the word beyond the one being fixated are known, then here also lexical constraints decrease the number of possible choices. Exhaustive counts done by the author an all words of length 4-13 in the Kučera and Francis corpus showed that, at worst, knowledge of a word's first and last letters and its length decreases the number of possible choices to less than 19 words 75% of the time. If grammatical constraints and crude information about inside letters further diminish the number of choices, then perceptibility may approach 1.0 for many of the letters in the word.[2]

Linguistic Control

The linguistic control hypothesis was originally detailed by O'Regan (1975). Similar ideas have been proposed by Haber (1976), Hochberg (1970), and Rayner and McConkie (1976). Two things are supposed: first, that lexical, syntactic, and semantic constraints can be used to disambiguate poor-quality visual information so that perceptibility is greater than visibility; second, that this process can occur fast enough for eye movements to be influenced in useful time. In particular, the eye should skip regions about which good predictions can be made and tend to go to a region in the periphery about which there are no confident hypotheses. Since immediate use is made of all possible constraints in guiding the eye, linguistic control is the most efficient way eye movements could be controlled in reading.

Length Control

If time is not available for linguistic constraints to be accessed and to interact with visual information, and for the result of this interaction to affect eye movements, then a workable eye-movement strategy should ensure that, on average, the eye behaves the same way it would under linguistic control, but without the necessity for ongoing linguistic analysis. Thus, words should in general be fixated once, about three or four letters from the beginning, to make sure that the most constraining visual information is available. Two suitable rules are (a) "From any position, jump three letters in the next word," and (b) "From any position, jump just left of the middle of the next word."

Predictions from Linguistic and Length Control

Both linguistic and length control hypotheses predict systematic changes in saccade size from moment to moment during reading. The eye should tend to jump completely out of a word after it has fixated once near the beginning. A measurable consequence of this is that saccades leaving a long word should be longer than saccades leaving a short word.

Under linguistic control, the eye may make longer saccades going towards long words than going towards short words. This is because perceptibility drops considerably at the end of a word, and this is further away in the case of long words. Under the length control strategy, "Fixate three letters into the next word," no such effect should be found. But a strategy like "Fixate just left of the middle of the next word" would have the effect, since the middles of long words are further away than the middles of short words.

A prediction made by a linguistic control hypothesis, but not by a length control hypothesis, is that the next word should tend to be skipped if it is short and highly predictable. An example would be the word THE following the main verb of a sentence.

A further test enabling linguistic and length control to be distinguished will be used in the experiment to be described. The technique makes use of a "window condition," in which the text display changes as the subject moves his eyes. Wherever the subject is fixating, all the letters further than 7 character positions away are automatically replaced by Xs, blanks being left blank. The subject is therefore effectively reading through a "window" 15 spaces wide (7 on either side of where he is fixating). Outside the window, only Xs and spaces are visible. Wherever his eye moves, the window also moves. Pilot experiments showed that subjects do not notice the window at all when it is 15 spaces wide and the text is in capital letters. McConkie and Rayner (1975) used similar window conditions, and their work confirms that when display changes occur sufficiently rapidly in peripheral vision, subjects need not be aware of them.

What predictions can be made about eye-movement strategies in the window condition? With length control, there should be no change at all in saccade lengths. Length information is not interfered with by the Xs, since the spaces in the text are not removed. Under linguistic control, however, the information available from the periphery is reduced. Information about letters is unavailable further than 7 letters from the fixation point. The predictability of the farther regions of the visual field will be reduced and the longer saccades will tend to be eliminated.

EXPERIMENT 1

The first purpose of Experiment 1 was to try to reject the rhythmical scan hypothesis by demonstrating that there exist systematic changes in saccade sizes as a function of word length. The second purpose was to use the window condition to determine if these changes are more compatible with length control or linguistic control.

Method

Materials. The following sentence structure was used: THE (noun) (verb) THE (prepositional phrase). In each noun or verb position was put either a SHORT 3-letter word or a LONG 10-14 letter word. There were eight sentences in all, corresponding to the eight possible combinations of SHORT and LONG words. They were:

SHORT SHORT THE SHORT:
The CAT SAW THE RAT in the barn

SHORT SHORT THE LONG:
The MAN RAN THE ADMINISTRATION of the firm

SHORT LONG THE SHORT:
The MAN CONGRATULATED THE BOY for his courage

SHORT LONG THE LONG:
The MAN ACKNOWLEDGED THE INTELLIGENCE of his boss

LONG SHORT THE SHORT:
The HIPPOPOTAMUS SAW THE MAN in the swamp

LONG SHORT THE LONG:
The PROFESSIONAL WON THE COMPETITION at the show

LONG LONG THE SHORT:
The COMMISSION REPRESENTED THE MAN at the trial

LONG LONG THE LONG:
The PERFORMANCE FASCINATED THE SPECTATORS that day

These sentences provide eight types of transitions. There are four subject transitions, starting from the subject of the sentence: SHORT—SHORT, from a SHORT subject to a SHORT verb; SHORT-LONG; LONG-SHORT; and LONG-LONG. There are also four object transitions, starting from the verb of the sentence: SHORT-THE SHORT starting from a SHORT verb and going towards an object of type THE SHORT; SHORT-THE LONG; LONG-THE SHORT; and LONG-THE LONG. For each transition type, there are two test sentences. The test sentences were randomly inserted in a list of 36 other unrelated sentences of diverse length and structure so that no set for the special construction used would be established.

Subjects. A total of 40 paid American or English tourists in Paris (mostly college students) took part in the experiment: 20 did the window condition and 20 did the normal condition. As none had had any experience with eye-movement measurement before, they had some difficulty relaxing, and calibration was not always maintained. This meant that a different number of subjects contributed data to each sentence, the smallest number being 10 subjects for sentence SHORT LONG THE LONG in the window condition, and the largest number being 18 sujbects for sentence LONG LONG THE LONG in the normal condition. It should also be noted that, through a programming fault, the scores for sentence LONG LONG THE SHORT were lost.

Procedure. The subject reclined in an armchair with his head supported by cushions. He clasped a bite-board in his teeth so as to minimize head movements. The experimenter adjusted the eye-movement recording glasses (photoelectric detection of scleral boundary),[3] and the subject executed a "smooth pursuit" calibration, which was verified by the "zero drift" method (cf. O'Regan, 1978) to be accurate to a distance corresponding to two letter spaces on the screen. When the calibration had been done satisfactorily, the subject fixated a "control point" on the left of the screen. If there was no calibration error, the subjects's fixating this point caused it to disappear and a sentence to appear. The position of the point coincided with the space following the word THE at the beginning of each sentence. When the subject finished reading the sentence and brought his eye backwards to the left edge of the screen, the sentence automatically disappeared and the control point reappeared. The cycle then repeated, unless, through head movement, change in average lower eyelid position, or amplifier drift, an error was present when the subject fixated the control point. If the error was greater than 3 letter spaces, the experiment was interrupted and a new calibration executed. If it was less than 3 letter spaces, the calibration was corrected by this amount and the experiment continued.

The sentences were displayed one at a time in upper case (6 by 4 matrix of dots) in the center of a Digital Equipment Corporation VR 12 screen controlled by a PDP-12 computer. Continuous 20-msec sampling of eye position and the modification of the display necessary in the window condition were done by the computer, as described in O'Regan (1975).

The subject's task was to read silently the sentences which appeared on the screen. In order that the subject should read attentively, he was told that, since after the experiment he would be asked questions about them, he should try to remember the sentences as accurately as possible. Including the calibration, the experiment lasted 20 min.

Results

The raw data consisted of 20-msec samples of the eye's position (measured in units of ¼ of a letter) during the reading of each sentence. A saccade was defined as a movement in which two consecutive 20-msec samples were in the same direction and together covered a distance of more than 1 letter. A

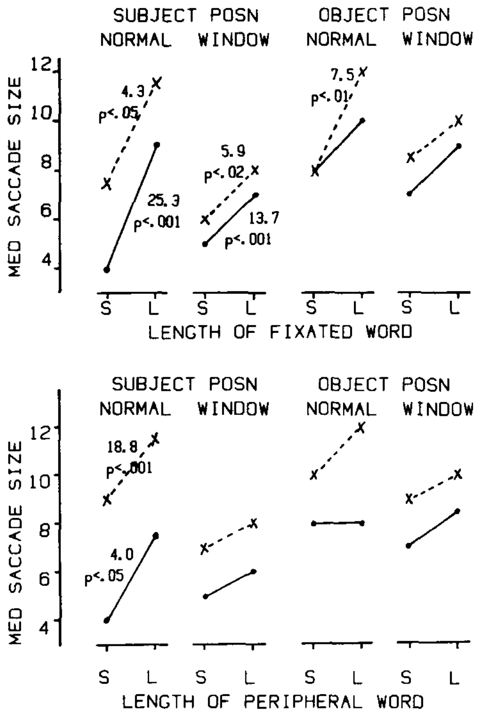

Figure 1. (a) The effect of length of fixated word on median size, measured in letters, of next saccade made. The solid lines indicate a SHORT word in peripheral vision; the dotted lines, a LONG word. (b) The effect of length of word in peripheral vision on the median size of the next saccade. The solid lines indicate fixation on a SHORT word; the dotted lines indicate fixation on a LONG word. In both (a) and (b), chi square and significance of the effect are indicated when $p < .05$ by the median test (Winer, 1971).

fixation was defined as the interval between the beginnings of successive saccades. The position of a fixation was taken as the last position occupied before the saccade. Saccade size was defined as the distance between two fixations.

Median saccade sizes (measured in letters) made by the eye starting from the first word of each type of transition were calculated and plotted in Figure 1. For clarity, the same data are plotted in two different ways in the figure: in the upper portion, to show the effect on saccade size of the length of the word the eye is fixating, and in the lower portion, to show the effect on saccade size of the length of the word in peripheral vision. Each plotted value is a median taken over data for 20 subjects reading the two sentences that contain a given type of transition. Only those saccades were taken into account which brought the eye all the way out of the first word of the transition. Saccades preceded by regressions coming from regions beyond the word were excluded, since we were interested only in what the eye does if it has not already scanned a word. Because of such rejected records, a sentence sometimes did not receive a score from each subject. Another reason scores were missing was that sometimes a subject did not fixate the first member of a transition at all. Finally, transition LONG-LONG and LONG-THE SHORT received only half their full complement of data because sentence LONG LONG THE SHORT was missing from the analysis. For all these reasons, an analysis of variance or related samples tests could not be performed, and fairly conservative median tests (cf. Winer, 1971) were used to compare the distributions of interest. Chi-square values and significance levels obtained are indicated in Figure 1.

Saccade sizes and the word the eye is fixating. The upper portion of Figure 1 shows the effect on saccade size of the length of the word being fixated. The essential aspect of the data is clear: The eye makes a larger saccade when it has been fixating a LONG word than when it has been fixating a SHORT word. The effect is strong in Subject Position: the two chi-square tests for the normal condition and the two chi-square tests for the window condition are all statistically significant (values given in Figure 1). While the same trend exists in Object Condition, only one of the comparisons is statistically significant. The weakness of the effect in Object Position is probably due to the intervening THE: as will be discussed later, saccade sizes were generally larger when the eye was heading for THE LONG or THE SHORT than when there was no intervening THE. Any additional lengthening of saccades due to the length of the word after the THE may have been hidden.

As seen from Figure 1, the effect on saccade size of the length of the word being fixated was reduced but not completely eliminated in the window condition. This is to be expected, because the size of the window (seven letters on either side of the fixation point) would generally allow the whole word to be seen from a fixation point within it.

Saccade sizes and the word in peripheral vision. The influence of the length of the word in peripheral vision on saccade sizes is shown in the lower portion of Figure 1. There is a trend for the eye to make larger saccades when a LONG word is in the periphery than when a SHORT word is in the periphery. However, the differences are only statistically significant in Subject Position for the normal condition (values of chi-square are given in Figure 1). This pattern is to be expected. In Object Position, the presence of the intervening THE may weaken the influence of the length of the word beyond it. The absence of a significant effect in the window condition shows that something other than just word length information is being used in the control of eye movements, since in the window condition length information still is available. The linguistic control hypothesis therefore is supported by the results.

Landing positions and the word in peripheral vision. Even if saccade sizes vary with the length of the word visible in the periphery, this does not imply

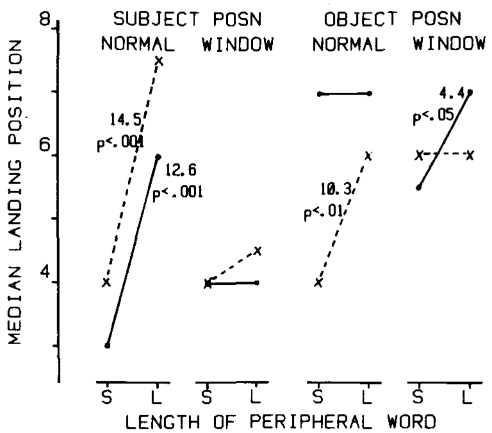

Figure 2. The eye's median landing position, measured in letters from the end of the 7-letter departure zone preceding the second word of a transition, as a function of the length of the second word. The solid lines are for cases where the departure zone contained a SHORT word; the dotted lines, a LONG word. Chi square is shown only when p < .05.

anything about where the eye actually lands: for example, the fact that saccades are larger when the eye jumps into a long word does not mean that the eye necessarily lands further into the word, since it may have started off earlier in the preceding one.

In plotting the positions where the eye lands in the second word of a transition, it is important to consider only certain saccades. First, regressions coming from regions beyond the word must be excluded as before. Second, saccades starting from too far to the left of the second word are unlikely to vary with its characteristics because it will have been too far in peripheral vision for useful data to be extracted. A "departure zone" consisting of the 7 letter spaces preceding the second word was defined. In the case of LONG first words, this departure zone contained the last 6 letters of the word and the space following it. In the case of SHORT first words, the departure zone contained the whole SHORT word, the space following it, and the 3 letter spaces preceding it. Figure 2 shows the median landing positions in the second word of a transition for saccades starting in these departure zones. The data are similar to the saccade size data of the preceding section.

In Subject Position, a strong effect is found in the normal condition: the eye lands further to the right when there is a LONG word in peripheral vision than when there is a SHORT one in peripheral vision (both chi-square tests significant at $p < .001$). The effect disappears in the window condition, as expected from the linguistic control hypothesis. The linguistic control hypothesis is therefore strongly supported.

In Object Position, for the normal condition, the effect of peripheral word length on landing position is weaker because of the intervening THE: Only one of the two tests is significant (chi square = 10.3, $p < .01$). For the window condition, one of the comparisons is just significant (chi square = 4.39, $p < .05$). This puzzling aspect of the data is probably an artifact of the missing data for the transition LONG THE SHORT.

THE-skipping. As noted in the previous sections, the data for Subject Position are generally quite different from those for Object Position. Saccade sizes are larger and landing positions are further to the right. These differences may be related to the presence of THE in Object Position. To show this, we can compare where the eye goes when approaching THE with where it goes when approaching a different 3-letter word. Table 1 compares (a) landing positions for saccades leaving the 7-letter departure zone preceding a 3-letter verb, with (b) landing positions for saccades leaving the 7-letter departure zones preceding THE.

In the normal condition, the eye jumps from the 7-letter departure zones THE CAT and THE MAN to a median landing position 3 letters to the right, i.e., to the middle letter of the 3-letter verbs SAW

Table 1
Median Landing Position for Transitions Classified by Length of Word at End of Departure Zone, by Word at Beginning of Landing Zone, and by Condition; Chi Square and p for Comparisons of Landing Zone-SHORT with Landing Zone-THE SHORT and -THE LONG

		Condition					
		Normal			Window		
Departure Zone	Landing Zone	Median	Chi Square	p	Median	Chi Square	p
SHORT	SHORT	3			4		
	THE SHORT	7	19.7	<.001	5.5	3.84	<.05
	THE LONG	7	16.7	<.001	7	14.0	<.001
LONG	SHORT	4			4		
	THE SHORT	4	.18	n.s.	6	3.97	<.05
	THE LONG	6	15.9	<.001	6	2.97	n.s.

Note—Landing positions are measured in letters from the end of the 7-letter departure zone. All medians are based on two transitions except for LONG-THE SHORT, which is based on one.

and RAN. But when the word THE is in peripheral vision, the eye jumps significantly further. From the 7-letter departure zones. . .MUS SAW and CAT SAW, the eye jumps to a median landing position 7 letters to the right, into the word following THE. The same is true when the eye is leaving . .NAL WON and MAN RAN (all comparisons significant at $p < .001$, cf. Table 1).

There are two comparisons for cases where the eye is leaving the 7-letter departure zone at the end of a long word and going towards either a 3-letter verb or the article THE. One is strongly significant (chi square = 15.9, $p < .001$). The other gives no effect (cf. Table 1); however, this is one of the cases where there were missing data.

The results of this section, showing that THE is skipped while a 3-letter verb is not, strongly argue that it is not simply length control that determines the size of saccades, since if that were so, eye movements near THE would be the same as eye movements near other 3-letter words.

The THE-skipping result is, however, consistent with the linguistic control hypothesis, under which linguistic constraints combine with visual information available during fixation of the preceding word to influence the size of the next saccade. Since THE is both highly predictable from the preceding context and occurs extremely frequently in English, it might be recognizable from further into peripheral vision than a 3-letter verb, and so would tend to be skipped more often.

More support for the THE-skipping phenomenon comes from the window condition data of Table 1. The same pattern of results is found as in the normal condition. The landing positions are always further to the right when there is a THE in peripheral vision (landing positions at the 6th or 7th letter beyond the departure zone) than when there is a 3-letter verb in peripheral vision (landing position at the 4th letter beyond the departure zone). The comparisons are statistically significant at $p < .05$ or better in three cases out of four[4] (see Table 1).

EXPERIMENT 2

Within the linguistic control hypothesis, there are two mechanisms that might be put forward to account for the THE-skipping found in Experiment 1.

The first concerns grammatical predictions. It may be that, through grammatical constraints imposed by the preceding text, the predictability of the main verb in the experiment was much smaller than the predictability of the THE that followed it. Saccades towards the verb would therefore be smaller than saccades towards THE. This mechanism makes few demands on the oculomotor system, because it is not necessary for poor-quality visual information from the periphery to be processed and immediately to affect the size of the next saccade. Instead, predictions from the preceding text are used.

A second mechanism that makes greater demands on the oculomotor system may be postulated. It is possible that THE-skipping can occur even without the help of grammatical predictions. Poor-quality information from peripheral vision may suffice to enable the lexical category of a word to be determined, and oculomotor programming may be fast enough to make use of this information. If this were true, then, even when grammatical constraints are held constant, a frequently occurring word like THE would be more likely to be skipped than a 3-letter verb. Furthermore, the difference should be most pronounced when comparatively rare verbs are compared to THE. Keeping the context preceding a THE and a 3-letter verb constant, the present experiment compared saccade sizes approaching THE (frequency 69,971 in the Kučera and Francis corpus) to saccade sizes approaching the relatively frequent 3-letter auxiliaries ARE, HAD, and WAS (frequencies 4,393, 5,133, and 9,825, respectively), and also approaching the rarer verbs ATE, MET, and RAN (frequencies 16, 132, and 134, respectively).

Method

Materials. Sentences having two possible endings were constructed as follows. The two versions of each sentence began in the same way. After a critical position, the sentence continued either with the word THE and some other words or with a 3-letter verb and some other words. To emphasize the fact that the verb has 3 letters, it will be referred to as VRB. In both versions of a sentence, the word following the critical THE or VRB had the same number of letters and began with the same letter. For each test sentence, a second, matched, sentence was constructed, also with two versions. This matched sentence had approximately the same surface syntactic structure, its beginning had words of the same length, the same VRB was used, the words preceding the critical THE/VRB began and ended with the same letters, and the word following the critical THE/VRB began with the same letter and had the same length. An example is given in Table 2. There were two experimental sessions separated by at least 3 days. If a subject read sentence 1-THE and 2-VRB in Session 1, he read sentence 1-VRB and 2-THE in Session 2. For each of the six verbs used in the experiment, there were four matched pairs of sentences: in each session, the subject read four VRB and four THE sentences for each verb that was used. The six verbs used were divided into a high-frequency class, consisting of the frequent auxiliaries HAD, WAS, ARE, and a low-frequency class, consisting of the rarer verbs RAN, ATE, MET. The experiment was thus divided into two conditions, one with 12 high-frequency VRB and 12 THE sentences for each session, and one with 12 low-frequency VRB and 12 THE sentences for each session. All 48 experimental sentences for each session were randomly intermixed with 96 filler sentences of diverse structure.

Subjects and Procedure. Seven students of psychology at the University of California at San Diego served as subjects. In most respects, the procedure was identical to that for Experiment 1. Differences were as follows. Sentences were displayed in capital letters (5 by 7 matrix) on a Tektronix Type 601 display unit, using storage mode. The recording apparatus was constructed at the LNR laboratory of the Department of Psychology of the University of California at San Diego. It was equivalent to the author's

Table 2
Example of Matched Pairs of Sentences Used in Experiment 2

	THE OTHER DAY WAS CAUGHT	1-THE
THE BEAR THAT JOE WAS HUNTING		
	HAD OFTEN BEEN SEEN	1-VRB
	THE OTHER DAY WAS THICK	2-THE
THE ROPE THAT JIM WAS HOLDING		
	HAD OFTEN BEEN USEFUL	2-VRB

Note–Vertical lines show where the same letters were used across sentences of any matched pair.

apparatus except that the amplifier had a rise time of 20 msec. Subjects' heads were secured by a moulded dental impression, and they sat upright instead of recumbent. The computer, which was a PDP-9 sampling at 10-msec instead of 20-msec intervals, did not adjust the calibration when the initial fixation point was slightly wrongly fixated. Instead, an accuracy of ½ letter was required before each new sentence appeared, and a new calibration was performed when this accuracy was not attained. After every five or six sentences, a question sentence appeared. The subject had to decide whether this question sentence had appeared in the preceding group of five or six sentences. The task was very easy.

Results

Figure 3 shows, for high- and low-frequency verb types, the mean size of the saccades made from each sentence position. The solid line corresponds to the VRB, the dotted line to the THE sentence of each pair. The values plotted were calculated as follows. For each sentence, a 30-letter region was defined consisting of the 19 spaces to the left of the THE/VRB, the THE/VRB itself, and the 8 spaces to the right of it. Combining data for all 12 sentences of a given type in both experimental sessions (24 sentences in all), the mean size of the saccades made from every position in the 30-letter region was calculated for each subject. Note that since a subject fixates only about every 7th letter in a sentence, the amount of data varies from position to position in the sentence. Means and standard errors were then taken across the seven subjects, and are shown in Figure 3.

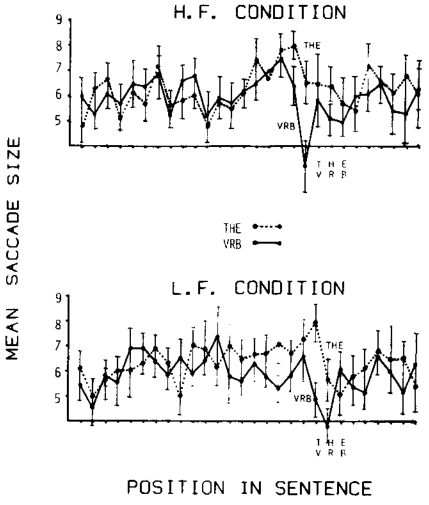

Figure 3. The mean saccade sizes (measured in letter spaces) made by the eye when leaving each of 30 positions of the sentences, for high- and low-frequency types of critical THE/VRBs. The positions where the critical THE or VRB occurred are shown on the abscissa. The curves for VRB sentences are solid lines; for THE sentences, they are dotted lines. The vertical bars show one standard error on either side of the mean.

The (dotted) THE and (solid) VRB curves follow each other closely in the region far to the left of the critical THE/VRB. This is to be expected, since THE and VRB sentences had the same beginnings. Indeed, the degree of compatibility of the two curves in this region gives an estimate of the data's variability which can be used in considering the differences in the curves found near the THE/VRB region.

As can be seen, such differences start appearing after a point 2 letters to the left of the THE/VRB region in the upper half of Figure 3 (high-frequency condition) and after a position 7 letters to the left of the THE/VRB region in the lower half (low-frequency condition): to the right of these points, the THE and VRB curves remain consistently separate for several sentence positions, showing that the effect is not a random variation. Apart from observing the consistent differences in the curves for these regions, it is difficult to estimate more precisely the statistical significance of the effect, since different numbers of subjects and different numbers of sentences contribute to each point. The above observations will therefore have to suffice in confirmation of what is directly visible from the curves: first, the eye makes larger saccades when approaching THE than when approaching a 3-letter VRB; second, the effect appears earlier when the VRB is a (frequently occurring) auxiliary (WAS, ARE, HAD) than when it is a rarer verb (MET, ARE, RAN).

Landing positions. The fact that longer saccades are made approaching THE than approaching VRB does not mean that the eye actually fixates THE less than VRB, since it may tend to start off from a different position. To determine where the eye lands, a 7-letter departure zone preceding the THE/VRB region was defined, consisting of the 7th to 1st letters preceding the THE/VRB. Cumulating across sentences of a given type, the mean position where each subject's eye landed when it started from anywhere within this departure zone was calculated. As for Experiment 1, only forward-going saccades when the eye had not already been beyond the departure zone were taken into account. Figure 4 plots the means across subjects of these means. As can be seen from the figure, the eye tends to land further to the right when THE is in peripheral vision than when VRB is is in peripheral vision.

An analysis of variance done on the data confirms that the main THE/VRB difference in landing position is significant at $p < .001$ [$F(1,6) = 26.04$]. The only other significant effect is the interaction of this main effect with the high/low frequency factor [$F(1,6) = 8.002$, $p < .05$]. As can be seen from the figure, this interaction stems from the fact that, as expected, the THE-skipping effect is stronger when THE is compared to a low-frequency VRB than when it is compared to a high-frequency auxiliary.

The landing-position data are therefore consistent with the saccade size data. When starting in the departure zone consisting of the 7 letters preceding THE/VRB, the eye lands further to the right when there is a THE in peripheral vision than when there is a VRB. The effect is stronger when THE is compared to a fairly rare verb than when it is compared to a more frequent auxiliary. For simplicity, we will call this effect the THE-skipping effect, even though, in fact, the eye skips both THE and VRB: the mean landing positions are in the 2nd letter beyond THE and the 1st letter beyond VRB.

The existence of the effect shows that even when syntactic and semantic predictions are held constant, peripheral visual information and lexical processing can combine to influence the size of the next saccade that will be made. This hypothesis is further supported by the difference in strength of the effect that was observed as a function of the type of the VRB used. The higher frequency auxiliaries (ARE, WAS, HAD), being perhaps easier to see or to process, induce less of a difference with THE than the lower frequency verbs ATE, RAN, MET.

Finally, it is interesting to note that, though statistically significant, the THE-skipping effect found in Experiment 2 is small (1 or 2 letters' difference between THE and VRB). When, as in Experiment 1, the possibility of making syntactic and semantic predictions is added to lexical differences, the size of the differences in landing position are larger, of the order of 4 letters.

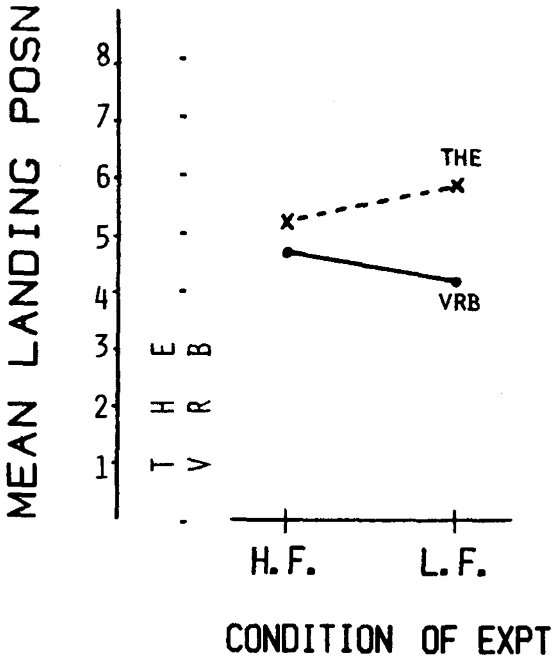

Figure 4. Mean over subjects of the positions on which the eye lands when leaving the 7-letter departure zone preceding THE (dotted line) or preceding VRB (solid line). Positions were measured in letter spaces from the end of the departure zone.

DISCUSSION

The most important aspect of the results of Experiment 1 is that they show a systematic dependence of eye movements on word length during reading. First, the eye makes larger saccades leaving LONG words than leaving SHORT words. Second, the eye lands further into LONG words than into SHORT words. These two effects on saccade size, of the length of the word currently fixated and of the word in peripheral vision, show that saccades in reading are not constant or randomly distributed about some mean. A "rhythmical scan" strategy in which saccade sizes adjust only slowly to overall aspects of the text being read would also not produce these results, since word lengths varied randomly from place to place in the sentences. Some degree of moment-to-moment control of saccades must be postulated.

The THE-skipping and window condition data available from Experiment 1 speak to the issue of whether this moment-to-moment control occurs as a consequence of a simple length control strategy or a linguistic control strategy. The effects of word length were weaker in the window condition. Saccades leaving LONG words differed less from saccades leaving SHORT words in the window condition than in the normal condition. Saccade length and landing positions were less dependent on the length of the word into which the eye was jumping. These findings are inexplicable in terms of a length control strategy, since word-length information was preserved in the window condition. On the contrary, the findings strongly support the idea of a linguistic control strategy. In the window condition, the presence of Xs beyond 7 letters in peripheral vision prevents processing from going as far as when these letters are not obscured. Therefore, all the effects that were found in the normal condition can, in the window condition, only act within a region extending 7 letters left and right of the fixation point. The longer saccades of the normal condition are eliminated in the window condition.

The THE-skipping data of Experiment 1 further confirm the linguistic control strategy hypothesis. The eye jumps further when approaching the word THE following the main verb of the sentences used in the experiment than when approaching 3-letter verbs. Under length control, the eye should behave the same way when approaching any 3-letter word, independently of its function in the sentence.

The question was left open, by the THE-skipping result of Experiment 1, as to whether THE was identified using peripheral visual information or whether syntactic and semantic predictions from the preceding text were used to program a larger saccade approaching THE. Evidence that poor-quality information concerning the word THE in peripheral vision can influence saccade sizes without the aid of syntactic and semantic predictions came from Experiment 2. There, the same sentence beginnings were followed with an ending that started either with THE or with a 3-letter verb. A difference in saccade size was again found: larger saccades are made when the eye approaches THE than when it approaches a 3-letter verb, the difference being smaller for auxiliaries than for rarer verbs. These data show that lexical information and poor-quality visual information suffice for a larger saccade to be immediately programmed. For this to be possible, perceptual processing and oculomotor programming must take place within the approximately 250-msec duration of the preceding fixation, that is, surprisingly quickly.

REFERENCES

Abrams, S. G., & Zuber, B. L. Some characteristics of information processing during reading. *Reading Research Quarterly*, 1972, **8**, 40-51.

Banks, W. P., Bachrach, K. M., & Larson, D. W. The asymmetry of lateral interference in visual letter identification. *Perception & Psychophysics*, 1977, **22**, 234-240.

Bouma, H. Visual interference in the parfoveal recognition of initial and final letters of words. *Vision Research*, 1973, **13**, 767-782.

Bouma, H., & de Voogd, A. J. On the control of eye saccades in reading. *Vision Research*, 1974, **14**, 273-284.

Communale, A. S. Visual selectivity in reading: A study of the relationship between eye movements and linguistic structure. *Dissertation Abstracts International*, 1973, **34**, 1692-1693.

Haber, R. N. Control of eye movements during reading. In P. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes*. Hillsdale, N.J: Erlbaum, 1976.

Hochberg, J. Components of literacy: Speculations and exploration research. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, 1970.

Huey, E. B. On the psychology and physiology of reading. I. *American Journal of Psychology*, 1900, **11**, 283-302.

Klein, G. A., & Kurkowski, F. Effect of task demands on relationship between eye movements and sentence complexity. *Perceptual and Motor Skills*, 1974, **39**, 463-466.

Kolers, P. A. Buswell's discoveries. In P. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes*. Hillsdale, N.J: Erlbaum, 1976.

Kucera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, R.I: Brown University Press, 1970.

McConkie, G. W., & Rayner, K. The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 1975, **17**, 578-586.

Mehler, J., Bever, T. G., & Carey, P. What we look at when we read. *Perception & Psychophysics*, 1967, **2**, 213-218.

Morton, J. The effects of context upon speed of reading, eye movements and eye voice span. *Quarterly Journal of Experimental Psychology*, 1964, **16**, 340-354.

O'Regan, J. K. *Structural and contextual constraints on eye movements in reading*. Unpublished doctoral dissertation. University of Cambridge, 1975.

O'Regan, J. K. A new horizontal eye movement calibration method: Subject controlled "smooth pursuit" and "zero drift." *Behavior Research Methods & Instrumentation*, 1978, **10**, 393-397.

Rayner, K. Parafoveal identification during a fixation in reading. *Acta Psychologica*, 1975, **39**, 271-282.

Rayner, K. Eye movements in reading: Eye guidance and integration. In P. A. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing of visible language 1*. New York: Plenum Press, 1979.

Rayner, K., & McConkie, G. W. What guides a reader's eye movements? *Vision Research*, 1976, **16**, 829-837.

Schiepers, C. W. J. Global attributes in visual word recognition. Part 1. Length perception of letter strings. *Vision Research*, 1976, **16**, 1343-1349.

Wanat, W. F. Language behind the eye: Some findings, speculations, and research strategies. In P. A. Monty & J. W. Senders (Eds.), *Eye movements and psychological processes*. Hillsdale, N.J: Erlbaum, 1976.

Winer, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.

Woodworth, R. S., & Schlossberg, H. *Experimental psychology*. New York: Holt, 1954.

NOTES

1. This is the reason why, as suggested by Huey (1900), for studies in reading, the extent of central vision should be measured in letters and not in degrees.

2. For example, if there happen to be only two words in the lexicon that are compatible with all the constraints, and if these words differ by only 1 letter (e.g., CONSTRUCT and CONSTRICT), then perceptibility will be 1.0 for all positions in the word except the one corresponding to the uncertain letter, where it will drop to .5.

3. The glasses and electronics were constructed by the author (cf. O'Regan, 1975). The amplifier has a .5-msec rise time.

4. The reason the window condition did not destroy the THE-skipping effect in the same way it destroyed the effect of word length on saccades is that the window extended 7 letters right and left of where the eye was fixating. From a fixating point in the 7-letter region preceding THE, the eye will often be able to see THE, since it is only 3 letters to the right of the 7-letter region.

(Received for publication September 5, 1978;
revision accepted December 12, 1978.)