Technical Report No. 406

EYE MOVEMENT CONTROL DURING READING: I. THE LOCATION OF INITIAL EYE FIXATIONS ON WORDS

G. W. McConkie, P. W. Kerr, M. D. Reddix, & D. Zola

University of Illinois at Urbana-Champaign

February 1987

Center for the Study of Reading

TECHNICAL REPORTS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

CENTER FOR THE STUDY OF READING

Technical Report No. 406

EYE MOVEMENT CONTROL DURING READING:
I. THE LOCATION OF INITIAL
EYE FIXATIONS ON WORDS

G. W. McConkie, P. W. Kerr, M. D. Reddix, & D. Zola

University of Illinois at Urbana-Champaign

February 1987

University of Illinois at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

The work upon which this publication is based was performed pursuant to Grant No. HD18116 of the National Institute of Child Health and Human Development. It does not, however, necessarily reflect the views of this agency. The authors wish to thank Gary Wolverton for preparing computer programs to collect and reduce the data from this study, and Dr. Arthur Jacobs for motivating and helping to create the data matrix used in the analyses. Nancy Bryant and John Grimes provided helpful suggestions in preparing this manuscript.

Abstract

Sixty-six college students read two chapters from a contemporary novel while their eye movements were monitored. The eye movement data were analyzed to identify factors that influence the location of a reader's initial eye fixation on a word. When the data were partitioned according to the location of the prior fixation (i.e., launch site), the distribution of fixation locations on the word (i.e., landing site distribution) was highly constrained, normal in shape, and not influenced by word length. The locations of initial fixations on words can be accounted for on the basis of five principles of oculomotor control: A word-object has a specific functional target location, a saccadic range error occurs that produces a systematic deviation of landing sites from the functional target location, the saccadic range error is reduced somewhat for saccades that follow longer eye fixations, there exists oculomotor variability that is a second, non-systematic source of error in landing sites, and the oculomotor variability increases with distance of the launch site from the target.

EYE MOVEMENT CONTROL DURING READING: I. THE LOCATION OF INITIAL EYE FIXATIONS ON WORDS

During reading, people tend to direct their gaze to a point just left of the center of a word more frequently than to other locations. Rayner (1979) referred to this as the *preferred viewing location*. O'Regan (1981) elaborated on this preference in his *convenient viewing position* hypothesis, suggesting that each word has a location from which it is most likely to be identified. This location depends on the distribution of information across the word, but over a group of words, the average will be near their center. Readers learn to adjust their gaze to place the eyes at that "convenient" position, in essence tuning their oculomotor behavior for maximum performance under the constraints of the visual system and characteristics of the text stimuli.

McConkie and Zola (1984) argued on the basis of this tendency to send the eyes to the centers of words, that the eye movement guidance system directs the eyes in terms of word units during reading. These investigators identified eye fixations located at different distances to the left of a six-letter word, and examined the lengths of rightward (forward) saccades following these fixations. The frequency distributions of rightward (forward) saccades were sometimes uni-modal and sometimes bi-modal, depending on where the eyes were centered with respect to the word, further demonstrating the word-based nature of eye guidance. McConkie and Zola (in press) suggested that early visual processes parse the text stimulus array into a hierarchical object structure, and that skilled readers typically attend to, and move their eyes with respect to, objects at the word level.

The study reported here further investigated the tendency for readers to send their eyes to the centers of words. A large set of readers' eye movement data, consisting of over 40,000 eye fixations, was analyzed to observe the effects of three variables on the distribution of initial fixation locations in a word: length of the word, distance of the eyes from the word during the prior eye fixation, and the duration of the prior fixation.

Increased word length could be expected to broaden the distribution of initial fixation locations on a word, which we refer to as *landing sites*, for any of several reasons. First, if word-level objects serve as targets for eye guidance, longer words may place less constraint on appropriate landing sites, so less precision is exercised in programming the saccade. Second, if the eyes are sent to specific locations in the word depending on where needed information is expected or peripherally localized, longer words provide a wider range of potentially informative locations. Finally, if readers vary in their preferred fixation locations in words, longer words provide increased opportunity for this variance to occur, again broadening the distribution of landing sites in data combined across subjects.

The location of the eye fixation prior to a saccade to a word, which we refer to as the *launch site* for that saccade, could be expected to influence both the most frequent landing site on a word, and the breadth of the landing site distribution. Since saccades tend to undershoot their target and this tendency increases with saccade distance (Henson, 1979), it is expected that with more distant launch sites the distribution of landing sites would shift leftward in the word. Furthermore, with greater distance, it is likely that the reliability of eye placement will be reduced (de Bie, van den Brink, & van Sonderen, 1987), thus leading to a broader landing site distribution.

Recent research indicates that longer oculomotor response latencies are associated with more accurate saccades (de Bie, van den Brink, & van Sonderen, 1987; Coeffe, 1987). Thus, it is possible that saccades launched following shorter fixations during reading will show a broader distribution of landing sites within the target word. On the other hand, since the text pattern was visible during prior fixations, it is possible that the accuracy of locating a target word and programming the saccade to it does not depend on the duration of the eye fixation immediately preceding that saccade. In this case, fixation duration would have no influence on the distribution of landing sites in the word fixated next. The analyses sought to distinguish between these alternatives, as well as to test the hypotheses stated above.

Methods

Data were taken from a series of studies conducted by McConkie, Reddix, and Zola (1985). Sixty-six college students read the first two chapters of a contemporary novel, displayed one line at a time on a computer screen. A reader could advance to the next line of text by pressing a button. The text was 994 lines long with a maximum of 73 characters per line. Four characters subtended approximately one degree of visual angle. Eye movement data were collected during the reading of 302 of these lines, each of which contained a word location of interest in the studies. At this word location the original word could be replaced by one of five alternative strings which violated constraints at the orthographic, lexical, syntactic, semantic, or discourse level. A sixth condition was a control condition in which the original word remained in the text. For the current investigation, only data from the control condition lines were used, and this was accepted as a sample of normal reading data for the subjects. Depending on the study, this amounted to either 50 or 100 lines of text per subject, yielding a total data matrix containing 43,668 eye fixations.

Subjects were told that the text contained errors, but their task was simply to read the novel as normally as possible. Fifteen times during reading they were asked questions about the content in order to make sure that they were comprehending what they read. These questions never dealt with the errors in the text.

Eye movements were monitored using a Dual Purkinje Image Eyetracker (Cornsweet & Crane, 1973) sampling eye position every ms. Subjects required one or two two-hour sessions to complete the study, depending on their reading speed and ease of eyetracking.

Results

Preferred Viewing Position

Figure 1 presents fixation location data for the first fixation made on words of different lengths. The number of eye fixations on which these percentages are based ranges from 1878 to 7089 for the different word lengths. For words with four or more letters, there are more fixations near the center than at either end. In most cases, in agreement with previous observations (O'Regan, 1981; Rayner, 1979; Zola, in press), the most frequently fixated location is at, or just left of, the center of the word.

[Insert Figure 1 about here.]

Influence of Launch Site

Where the eyes land in a word (i.e., the landing site) depends strongly on the location of the prior fixation (i.e., the launch site). This is illustrated in Figure 2, which shows the distribution of landing sites on four-, six- and eight-letter words following launch sites at different distances to the left of the word. Launch sites and landing sites are indicated by designating a particular letter position as zero and indicating character positions to the left of that location with negative numbers, and character positions to the right with positive numbers. In Figure 2, the space immediately to the left of the word is indicated by zero. Hence, a launch site of -1 is actually two letter positions to the left of the first letter of the word.

Each curve in Figure 2 shows a strong tendency for the eyes to go to a particular part of the word, but the most frequently fixated location varies with the launch site. When the launch site lies further to the left of the word, the most preferred landing site in the word also shifts to the left.

[Insert Figure 2 about here.]

The landing sites in a word, given a constant launch site and word length, appear to be normally distributed. Therefore, the normal curve was fit to 35 landing site distributions, representing four- to eight-letter words, with launch sites of -1 to -7 with respect to the space in front of the word (.5 to 2 degrees of visual angle to the left of the word). This was accomplished by using the PAR program from the BMDP statistical package (Dixon, 1981). In each case, the dependent variable was the proportion of initial fixations on the word that were centered at each of the letter positions in that word, including the space immediately to the left of it.

Table 1 presents the results of fitting curves to those distributions for which this is appropriate. Other distributions have too few fixations, or do not have the type of relative maximum that is necessary for fitting a normal curve. Table 1 includes the mean, standard deviation, and average residual (i.e., mean of the absolute values of the differences between the best fit curve and each actual data value), for each of the distributions where estimation is possible, plus the total number of fixations in the distribution. Figure 2 presents the fitted curves for a sampling of these distributions.

[Insert Table 1 about here.]

Five of the cases included in Table 1 have landing site distributions with a mode on the final letter of the word. In these cases, without a clear relative maximum, a true fit of the normal curve was not possible, and the obtained parameters are questionable. Therefore, these instances are not used in further data presentation. In other cases, as seen in Figure 2, the fit was consistently quite good, with an average residual of .018. We conclude that the obtained parameters for mean and standard deviation can be accepted as characterizing the distributions, and can be used to test hypotheses about changes in the central tendency and variation of landing site distributions as a function of the variables being investigated. There is some tendency for obtained data values at the ends of the words to be higher than expected, and a rationale for this departure from normality is presented later.

Effect on mean landing site. Estimated mean landing site was plotted against launch site for four- to eight-letter words. For each word length, the relationship is linear with very similar slopes, ranging from -0.45 to -0.53 and averaging -0.49 and with y-intercepts of these regression lines that increase with word length. However, when the estimated mean landing sites were plotted against launch site measured from the center of the word, rather than from its beginning, the regression lines came together, as shown in Figure 3. From the lines fit to the means for each word length, a predicted mean landing site was calculated for a launch site eight character positions to the left of the center of the word (the middle of the launch site range studied). These predicted means were within three-tenths of a character position from each other, ranging from -0.6 to -0.9 character positions from the center of the word, and averaging -0.7.

[Insert Figure 3 about here.]

From these results, it is concluded that word length has very little effect on the means of the landing site distributions. Rather, the means of these distributions are simply a linear function of the distance of the launch site from the center of the word. Furthermore, for each character position further to the left that the launch site lies, the mean of the landing site distribution in a word is shifted 1/2 character position to the left. Finally, the mean of the landing site distribution is at the center of the word when the launch site is between six and seven character positions (1.50 to 1.75 degrees) to the left of the center of the word for all word lengths studied. When the launch site is closer than this, the mean is to the right of the center of the word; when it is farther away, the mean is to the left of the center.

Effect on landing site variability. Figure 4 presents a plot of the estimated standard deviations of the landing site distributions as a function of the distance of the launch site from the center of the word.

There is a clear trend for landing site distributions to show greater variability when the launch site is more distant.

[Insert Figure 4 about here.]

Influence of Word Length

Effect on mean landing site. As shown in Figure 3, word length appears to have little influence on the mean of the distribution of landing sites on words, once distance of the launch site from the middle of the word is taken into account. The data for eight-letter words is slightly below that for shorter words. This finding receives further comment in the discussion section.

Effect on landing site variability. In order to determine whether there is a relationship between the lengths of words and the estimated standard deviation of the distribution of landing sites on them, it was necessary to hold launch site constant. This was accomplished by grouping the standard deviations presented in Table 1 by launch site as measured from the center of the word. This procedure yielded groups of standard deviations for words of different lengths from launch sites ranging from -4.5 to -10.5. For each pair of word lengths in each group in which word length differed by two letter positions, the slope and y-intercept of the line joining them was calculated. Thus, these slopes indicated the change in standard deviation that resulted from a 33% to 50% increase in word length. The median value of the 14 resulting slopes was .005, indicating little, if any, change with word length. Thus, with launch site held constant, there is no evidence that the variance of landing site distributions increases for longer words.

Influence of Fixation Duration

The data were examined to determine whether there is greater variability in landing sites following shorter eye fixations. This was accomplished by partitioning the data on the basis of prior eye fixation duration so that the means and standard deviations of the landing site distributions could be compared as a function of prior fixation time. The four categories used were 20-119, 120-219, 220-319, and 320-519 ms. In order to permit this breakdown, data were collapsed across pairs of word lengths and pairs of launch sites. A landing site distribution was obtained for each prior fixation duration category, for words of length 5-6 and 7-8, and for launch sites from -5 to -6, -7 to -8 and -9 to -10, as measured from the center of the word. This selection excluded conditions with very small sample sizes, and without a relative maximum in their landing site distributions.

Examination of the data indicated that there were still three conditions that could not be fit appropriately with a normal curve. The estimated means from the remaining distributions are presented in Figure 5. Each point represents an estimated mean, based on a sample size between 35 and 1147, with the largest sample sizes appearing in the two middle fixation duration categories. Each line represents the means of the landing site distributions following fixations of different durations for a given word length-launch site condition.

[Insert Figure 5 about here.]

The lines in Figure 5 are of different heights due to the previously discussed effect of launch site. However, the most interesting aspect of Figure 5 is the interaction between condition and fixation duration. The means for the different conditions are furthest apart following short fixations, and show a definite tendency toward convergence following longer fixations.

In order to determine whether this convergence was due in some way to systematic differences in the variability of the distributions, estimates of the standard deviations of each distribution were examined. Several of the distributions for the extreme fixation duration categories had small sample sizes, and the standard deviations of these varied greatly, ranging from 1.17 to 2.52. However, estimated standard deviations for 13 distributions having sample sizes greater than 100, including one for the shortest fixations and two for the longest, ranged between 1.33 and 1.62 with only one exception, and showed very little relationship with prior fixation duration. While there is a need for further investigation of saccades following very short eye fixations, the current data did not support the hypothesis that variance in landing site distributions is reduced following longer eye fixations.

These data indicate that mean landing site does vary somewhat with prior fixation duration. The consistency of landing sites, as indicated by the standard deviation of landing site distributions, does not appear to increase with fixation duration, but the target location within the word, as indicated by the mean of the distribution, does change, apparently converging toward a point slightly to the left of the center of the word

Discussion

Relation of Landing Site Distribution to Preferred Viewing Position

The results of this study have confirmed the observation that led to the Preferred Viewing Position (Rayner, 1979) and Convenient Viewing Location (O'Regan, 1981) hypotheses, namely, that during reading the eyes go most frequently to a location slightly left of center in a word. However, the study has shown that this tendency is not a fundamental phenomenon itself, but instead is the result of more basic phenomena. The Preferred Viewing Position was simply the maximum point in a distribution of all fixations on the word, which is referred to as the *composite distribution*. This composite distribution is actually a summation of many distributions of landing sites on the word, each contingent on the site from which the saccades were launched. These more fundamental distributions are referred to as *Launch Site Contingent Landing Site Distributions*, or simply as *landing site distributions*. Landing site distributions approximate the normal curve, though with some systematic deviation that will be discussed later. We refer to the estimated mean of a landing site distribution as the *Estimated Mean Landing Site*, or simply as the *mean landing site*, for saccades launched from a particular location. This is not the same as the mean eye position of all fixations on a word following saccades from a given location. Rather, it is a parameter of a distribution fit to those data.

The composite distribution, investigated in earlier studies, is complex in two ways. First, it includes both initial and later fixations on the word. The present paper has not attempted an analysis of the locations of refixations of a word. Second, for initial fixations on the word, which predominate in the data from skilled readers, the composite distribution includes a number of landing site distributions, each resulting from saccades coming from a different launch site. Thus, the particular shape of a given composite distribution, and hence the Preferred Viewing Position, is largely determined by the proportion of initial fixations on the word, and the frequency with which these fixations were preceded by fixations at different launch sites.

Relation of Landing Site Variance to Accuracy of Fixating Small Targets

The variance of the landing site distributions is much smaller than that of the composite distribution, as can be seen by comparing the distributions in Figures 1 and 2. The fact that landing site variance shows little relation to word length within the range of lengths studied indicates that the oculomotor system is not simply avoiding blank spaces in placing the eyes in a word (Abrams & Zuber, 1972-73). If that were the case, the standard deviation would be much larger for longer words, a relationship not observed in the data. Apparently, then, the visual pattern of a word provides the oculomotor system with a specific target site that is much more constrained than simply the space occupied by the word. Rather, it is as if the center of the word, or perhaps some "center of gravity" within the word (Kaufman & Richards, 1969), was the target used in computing the length of the saccade taking the eyes to that word.

If the target site for the saccade is a specific location within the word, then this fact raises the question of whether the variability of landing sites in a word, given a particular launch site, is similar to the variability obtained with smaller targets that are used in studies of oculomotor control. That is, does a word that is 1-2 degrees in length provide just as constrained a target location for the next saccade as does a .1 to .5 degree target to which subjects attempt to move their eyes? This comparison is made difficult by the fact that most studies of oculomotor control have used much longer saccades than those typically observed in reading. However, Kapoula (1985) examined the locations of fixations on a .5 degree square target following saccades of about 2.7 degrees, similar to the 2.75 degrees distance in the -11 launch site condition of the present study. She obtained standard deviations of fixation location distributions ranging from .28 to .36 degrees for individual subjects, with an average of .34 degrees, smaller than the .48 degrees standard deviation obtained in the present study following saccades of comparable length. However, some of this difference could be the result of variability among our readers in the precise location within a word that serves as the target location, since, unlike Kapoula, our data are combined across many subjects. A future study should obtain enough reading eye movement data from individuals to examine the characteristics of their individual landing site distributions, and to relate these to distributions of fixation locations when the same subjects send their eyes to small targets at distances similar to those of saccades typically made during reading.

Kapoula (1985) observed an increase in variability of fixation locations following saccades of different lengths (see also de Bie, van den Brink and van Sonderen, 1987). As in our data, the further away the target, the greater the variability in landing sites when trying to go to it, a phenomenon typical of motor control (Poulton, 1981). The reading data showed a strong nonlinear relationship, but the linear component had a slope of .097, slightly greater than the .064 reported by Kapoula, which may reflect differences between the tasks in the two studies.

The Basis for Launch Site Influences on Landing Sites

Why does the mean landing site vary with launch site? Consider two extreme data patterns that could have been obtained in this study. First, it would have been possible to find that the eyes always go to the centers of words, with some random error in achieving that target. In the extreme case, we would have obtained no relation between launch site and mean landing site (i.e., slope = 0), indicating a totally word-based determination of fixation location. Second, it would have been possible to find that the eyes are simply sent a particular distance, regardless of word locations, again with some variability. In this case, there would have been a perfect relationship between the mean launch site and the mean landing site (i.e., slope = 1), with words having no influence on fixation location. It is interesting, then, that the actual data yielded a slope that is just halfway between these extremes (i.e., slope = .49). There are three possible explanations that can account for this result.

First, it would be possible to obtain a slope of .49 from distributions containing a mixture of two types of saccades, some of which are directed to the centers of words, and some of which are simply cast a certain average distance. If this were the case, each landing site distribution would actually be a composite of two distributions of saccades. However, this would have produced much greater differences in standard deviations of landing site distributions with different launch sites than were observed in the data. Also, estimated standard deviations of landing site distributions should have increased with word length, and on longer words there should have been conditions where these distributions became bimodal in shape. None of these predictions was confirmed.

Second, changes in mean landing site with different launch sites could also be predicted from a "peripheral preview" hypothesis, in which partial information is assumed to be acquired from a word during the fixation prior to that on which it is directly fixated (Rayner, Well, Pollatsek, & Bertera, 1982). The closer the eyes are to a word, the more information might be obtained from the early parts of that word. For example, initial letters in the word might be identified during the fixation prior to going to that word (Rayner, McConkie, & Zola, 1980). Furthermore, with greater identification of the first part of the word, the eyes could be sent further into the word for the next fixation. Thus, the fixation location within the word that would be most effective for further processing would shift rightward as the prior fixation comes closer to the word, leading to the shift in mean landing site observed in the study. However, such an hypothesis would predict a non-linear relationship between launch site and mean landing site. Little or no letter-level information is acquired during reading more than about seven or eight letter positions to the right of the currently fixated letter (Underwood & McConkie, 1985). Thus, this hypothesis could not account for a continued linear relationship, such as was found in the current investigation, but would predict a negatively-accelerated curve that asymptotes at a launch site no more than seven letter positions to the left of the word.

A third hypothesis, and the one that is most consistent with the current data, is that the initial placement of the eyes within a word is determined strictly by principles of oculomotor control. Early visual processes parse the text stimulus pattern into objects, perhaps organized hierarchically (Duncan, 1984), and the reader selects a word-object as the target for a given saccade (McConkie & Zola, 1987). Of course, it is possible that this selection might be made for a saccade beyond the immediately following one. There is a specific location within words of a given length that is the true target for the saccade. We refer to this as the *functional target location*. Within the range of word lengths studied, it is appropriate to represent the functional target location as the center of the word, though in reality it is unclear just how this location can be positively identified. For eight-letter words the functional target location may have been slightly further to the left than it was for other word lengths, as indicated by the fact that the line relating mean landing site to launch site fell slightly below the lines for other word lengths in Figure 3. Thus, the functional target location may actually be some sort of "center of gravity" within the word (Kaufman & Richards, 1969) with less weight being given to more distant regions of longer words. This possibility requires further study.

As with all muscle control, however, there is some error in directing the body part quickly to the functional target location. This error has two components, systematic and random. First, there is a systematic error related to the length of the movement, which Poulton (1981) refers to as the constant error (CE). This has appeared in oculomotor studies as a tendency to undershoot the target (Henson, 1979) typically by about 10% of the total distance (Becker, 1972). However, Kapoula (1985; Kapoula & Robinson, 1986) has recently demonstrated that, at least under some circumstances, this is actually an example of a range effect in muscle control that has been described by Poulton (1981). In this range effect, when the eyes are close to a target, thus requiring short saccades, the eyes tend to overshoot their target, whereas when the eyes are further away, thus requiring longer movements, they tend to undershoot their target. Kapoula and Robinson demonstrated that the point at which an accurate mean saccade length is obtained (i.e., the point at which overshoots and undershoots are equal) varies with the lengths of saccades being executed within the task. Thus, it is the range of saccade lengths required in the task that determines the point of equality, with this point tending to be toward the shorter end of the continuum of lengths of saccades being executed. As Poulton (1981) points out, this is actually a central tendency bias, or contraction bias, since the range of saccade lengths is smaller than the range of target distances to which the eyes are sent in the task. We refer to this source of constant error as the *saccadic range error.*

The shift of mean landing site as launch site varied in the current study can therefore be understood as the manifestation of a basic principle of controlled muscle movement, realized in the form of the saccadic range error. In order to measure the actual amount of error, the deviation of the mean from the target, it would be necessary to know where the functional target is within the word. This is an empirical issue that cannot be firmly established from the current data. However, if we assume that the functional target is the center of the word, then the mean of the landing site distribution is accurate when the launch site is between six and seven character positions to the left of that target. For saccades coming from this region, undershoots and overshoots are balanced. This is slightly below the median saccade length for the data set, which was 7.0 character positions.

The second type of error is random variation of landing sites around the mean, which is quite normally distributed and increases with distance to the target. This source of error is referred to as *oculomotor variability*. The size of this error is indicated by the standard deviations of the landing site distributions, though these standard deviations are undoubtedly inflated somewhat due to inaccuracy in the eye position signal, and to the fact that each launch site was a zone, rather than a point. If there are individual differences in the functional target within words, then they also inflate the standard deviations.

Finally, the degree to which the final location of the eyes is displaced from the functional target location by the saccadic range influences is determined by the duration of the preceding eye fixation. When that fixation is relatively short, the saccadic range influences are greater, thus spreading out the mean landing sites for saccades coming from different launch sites, whereas when the prior fixation is longer, the range error is reduced and the mean landing sites converge somewhat within the word. At the same time, the oculomotor variability is affected very little by the duration of the prior fixation, though more data are needed in the case of very short fixations (i.e., under 120 ms). Thus, the location of the target word has a strongly constraining influence on the lengths of saccades being sent to that word, even following relatively short eye fixations. Also, there is a large saccadic range error even following rather long eye fixations. Since the increase in accuracy that results from longer fixation time is specific to the constant error, apparently without affecting the oculomotor variability, it is likely that this results from more accurately locating the functional target in space, rather than from increasing accuracy in the oculomotor programming and control itself. Particularly in the case of short fixations, it is possible that the programming of the saccade is actually based on information concerning the location and length of the targeted word that was acquired during a prior fixation. This remains to be studied using eye movement contingent display manipulation techniques, by which the location and length of a word can be changed between fixations (Reder, 1973).

In summary, it appears that the current data can be well accounted for by a strict oculomotor control theory of the location of initial fixations in words during skilled reading. Once a word-object has been selected as the target for a saccade during reading, the location of the fixation on that word is determined by five principles: (1) a stimulus object has a functional target location, (2) a saccadic range error occurs which produces an error in the mean landing site that is a linear function of the distance of the launch site from the functional target location, (3) the influence of the saccadic range error is lessened following longer eye fixations, (4) oculomotor variability occurs in the form of random placement error, and (5) such oculomotor variability increases with the distance of the launch site from the target.

Implications for the Use of Eye Movements in Studying Cognition

There are two implications of these results that are important to consider when using eye movement data to study cognition. First, it is important to note that the combination of systematic and random error in landing site distributions on a word can lead to eye fixations that were destined for one word but are actually centered on another. This is particularly the case for short words. In the distributions of landing sites (see Figure 2), there are cases where the mean of the distribution appears to have actually been at the edge of the word or even on the adjacent word. There are other cases where the peak of the distribution is well within the word, but the distribution appears to extend beyond the word boundaries into an adjacent word. Thus, the set of initial eye fixations on a word consist of a combination of some fixations intended for that word, and others that were actually intended for an adjacent word. Of course, these mislocated fixations would occur most frequently at the beginning and ends of words. This is probably the reason that obtained landing site frequencies at the beginnings and ends of words in the current study tended to be higher than the frequencies expected from fitting the normal curve to the data. Furthermore, since there is a bias toward undershooting a target location, and since the landing site distribution appears quite symmetric, there are probably more mislocated fixations at the ends of words (i.e., undershooting the following word) than at the beginnings of words (i.e., overshooting the prior word). Because of these factors, more accurate estimates of the landing site distributions will require simultaneous fitting of curves to data for two or more adjacent words, allowing the overlap of the distributions for different words to be taken into account. Fitting multiple adjacent distributions will also yield estimates of the frequencies of these "mislocated" fixations. Mislocated eye fixations serve as a source of error for forms of data analysis that assume that fixation time on a word is a measure of processing time for the fixated word (Just & Carpenter, 1980).

A second implication of the current results concerns the relation between the oculomotor system and higher cognitive processes during reading. Although eye movements must be controlled in such a way as to meet the needs of language processing, the current study indicates that at least one aspect of eye guidance is largely free from cognitive control, namely, where the initial eye fixation on a word is located within that word.

References

Abrams, S. G., & Zuber, B. L. (1972-73). Some temporal characteristics of information processing during reading. *Reading Research Quarterly, 8*, 42-51.

Becker, W. (1972). The control of eye movements in the saccadic system. *Bibliotheca Ophthalmologica, 82*, 233-243.

Bie, J. de, Brink, G., van den, & Sonderen, J. F. van (1987). The systematic undershoot of saccades: A localization or an oculomotor phenomenon? In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements from physiology to cognition* (pp. 85-95). Amsterdam: Elsevier North-Holland.

Coeffe, C. (1987). Two ways of improving saccade accuracy. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements from physiology to cognition* (pp. 105-115). Amsterdam: Elsevier North-Holland.

Cornsweet, T. N., & Crane, H. D. (1973). Accurate two-dimensional eye tracker using first and fourth Purkinje images. *Journal of the Optical Society of America, 63*, 6-13.

Dixon, W. J. (1981). *BMDP Statistical Software*. Berkeley: University of California Press.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501-517.

Henson, D. B. (1979). Investigation into corrective saccadic eye movements for refixation amplitudes of 10 degrees and below. *Vision Research, 19*, 57-61.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

Kapoula, Z. (1985). Evidence for a range effect in the saccadic system. *Vision Research, 25*, 1155-1157.

Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: Saccades can be accurate. *Vision Research, 26*(5), 735-743.

Kaufman, L., & Richards, W. (1969). Spontaneous fixation tendencies for visual forms. *Perception and Psychophysics, 5*, 85-88.

McConkie, G. W., Reddix, M., & Zola, D. (1985, November). *Chronometric analysis of language processing during eye fixations in reading.* Paper presented at the Annual Meeting of the Psychonomic Society, Boston, MA.

McConkie, G. W., & Zola, D. (1984). Eye movement control during reading: The effects of word units. In W. Prinz & A. F. Sanders (Eds.), *Cognition and Motor Processes.* Berlin Heidelberg: Springer-Verlag.

McConkie, G. W., & Zola, D. (1987). Visual attention during eye fixations while reading. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 385-401). Hillsdale, NJ: Erlbaum.

O'Regan, K. (1981). The "convenient viewing position" hypothesis. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 289-298). Hillsdale, NJ: Erlbaum.

Poulton, E. C. (1981). Human manual control. In V. B. Brooks (Ed.), *Handbook of physiology (Section 1, Volume II, Part 2)* (pp. 1337-1389). Bethesda, MD: American Physiology Society.

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

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12*, 206-226.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31*, 537-550.

Reder, S. M. (1973). On-line monitoring of eye position signals in contingent and noncontingent paradigms. *Behavior Research Methods and Instrumentation, 5*, 218-228.

Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly, 20*, 153-162.

Zola, D. (in press). Redundancy and word perception during reading. *Perception and Psychophysics*.

Table 1

Estimated Mean and SD of Landing Site Distributions

Launch	4-letter Words				5-letter Words				6-letter Words				7-letter Words				8-letter Words			
Site***	Mean	SD	Res**	N**	Mean	SD	Res	N	Mean	SD	Res	N	Mean	SD	Res	N	Mean	SD	Res	N
-1	3.50*	1.03	.035	268	4.35*	1.32	.038	294	4.51	1.35	.023	329	4.82	1.53	.006	229	4.61	1.67	.021	113
-2	3.49*	1.00	.032	467	3.85	1.24	.017	452	4.16	1.44	.009	454	4.38	1.55	.014	308	4.37	1.34	.026	153
-3	3.04*	1.36	.034	630	3.61	1.51	.019	585	3.70	1.39	.010	453	3.86	1.49	.016	323	3.73	1.57	.018	192
-4	2.41	1.47	.023	775	2.91	1.45	.015	589	3.07	1.54	.005	486	3.50	1.47	.016	351	3.39	1.44	.021	175
-5	2.09	1.53	.021	677	2.51	1.57	.014	595	2.51	1.54	.009	630	2.84	1.75	.007	344	2.95	1.61	.021	194
-6	1.53	1.61	.026	689	2.08	1.75	.018	630	2.26	1.89	.025	441	2.56	2.00	.029	342	2.35	2.30	.030	143
-7	0.96	1.45	.033	556	1.13*	1.66	.032	526	1.70	1.73	.018	398	2.00	2.00	.016	247	2.05	2.14	.013	156

* indicates a distribution in which the mode is not at a relative maxima, but is at the extreme left or right. In this case, the curve-fitting algorithm did not fit the data well, and the estimated parameters are questionable.

** Each value in the Res column is the average of the absolute values of the residuals for the data points in the distribution. Each value in the N column is the number of data values (observations) in the distribution, and hence is the N on which individual proportions were based.

*** Launch site is measured in character positions relative to the space immediately to the left of the word, designated character position zero. Negative numbers indicate positions to the left of that space.

Figure Captions

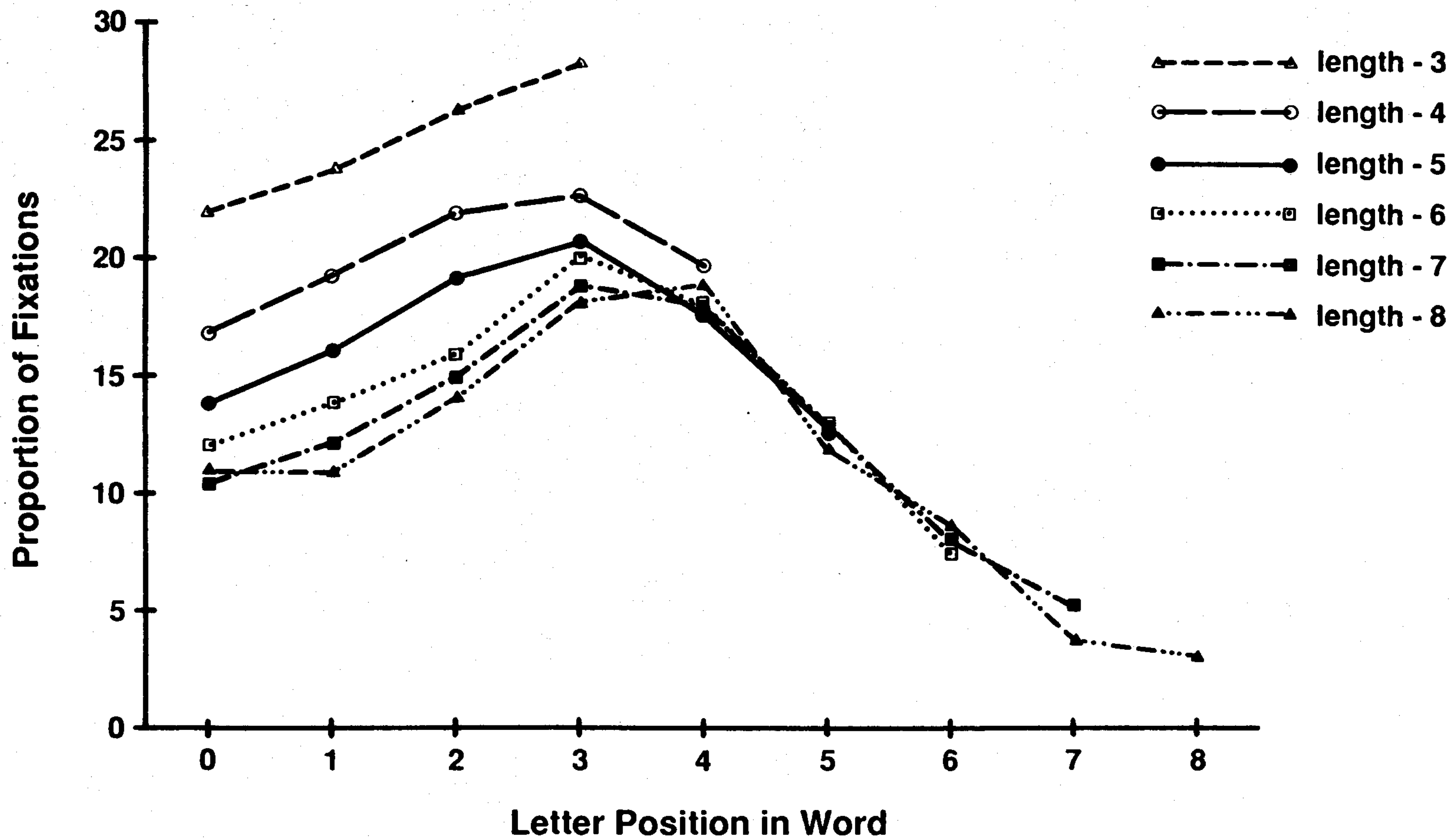
Figure 1. Proportion of eye fixations that are centered at different letter positions within words of different lengths. Letter position 0 is the space to the left of the word. Proportions are based on sample sizes of 1878 to 7089 eye fixations. For words of four or more letters, fixations are more frequent in the center than toward the ends.

Figure 2. Landing site distributions for different launch sites and word lengths. Letter locations are measured with respect to the space to the left of the word, which is designated character position 0. Locations to the left are designated by negative numbers; location to the right by positive numbers. Also presented is the best fit normal curve for each distribution. Parameters and sample sizes are presented in Table 1.

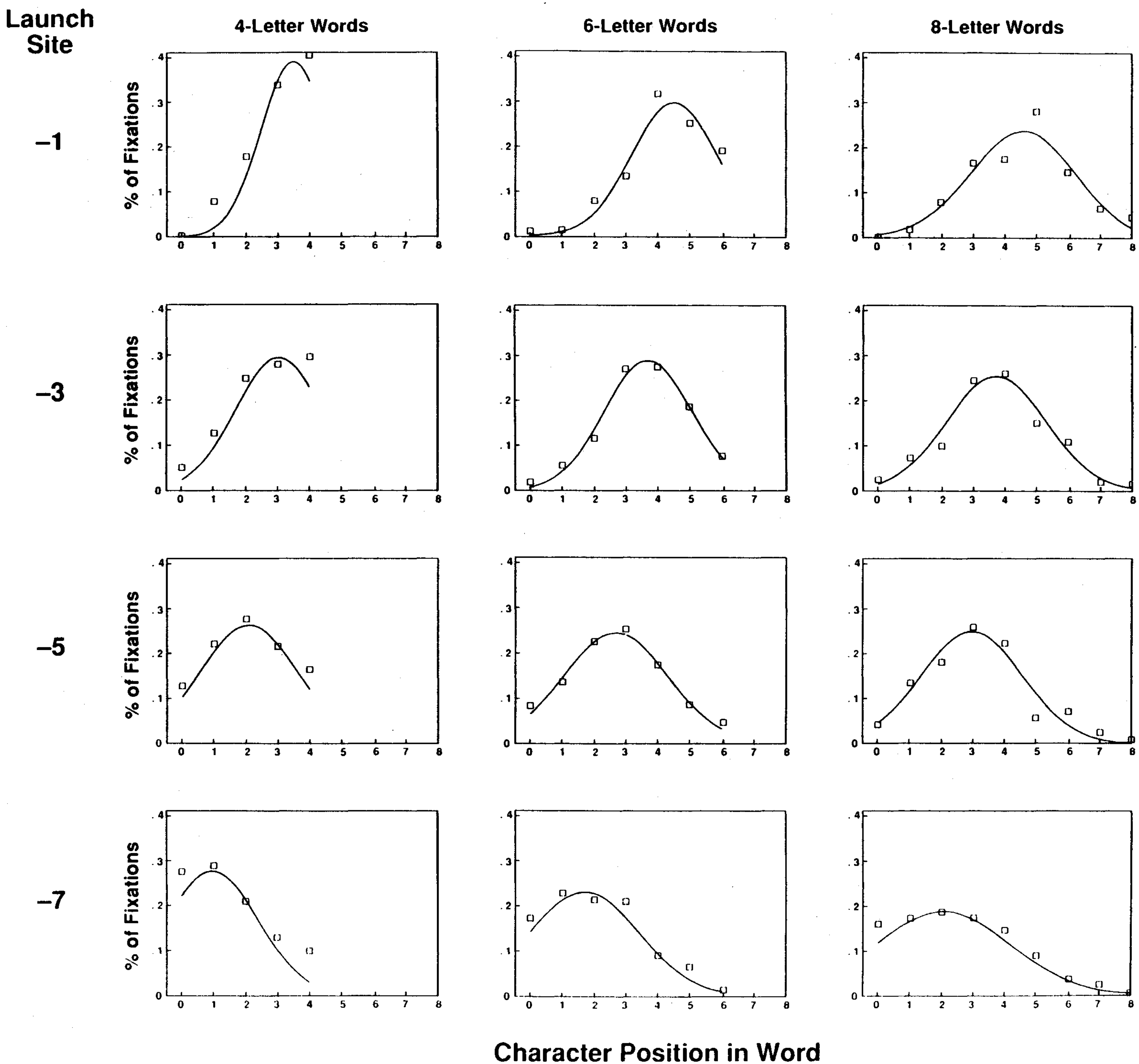
Figure 3. Estimated means of landing site distributions as a function of launch site for words of different lengths. Distances are from the center of the word. Best fit linear curves are also presented for each word length. Average slope is 0.49 character position.

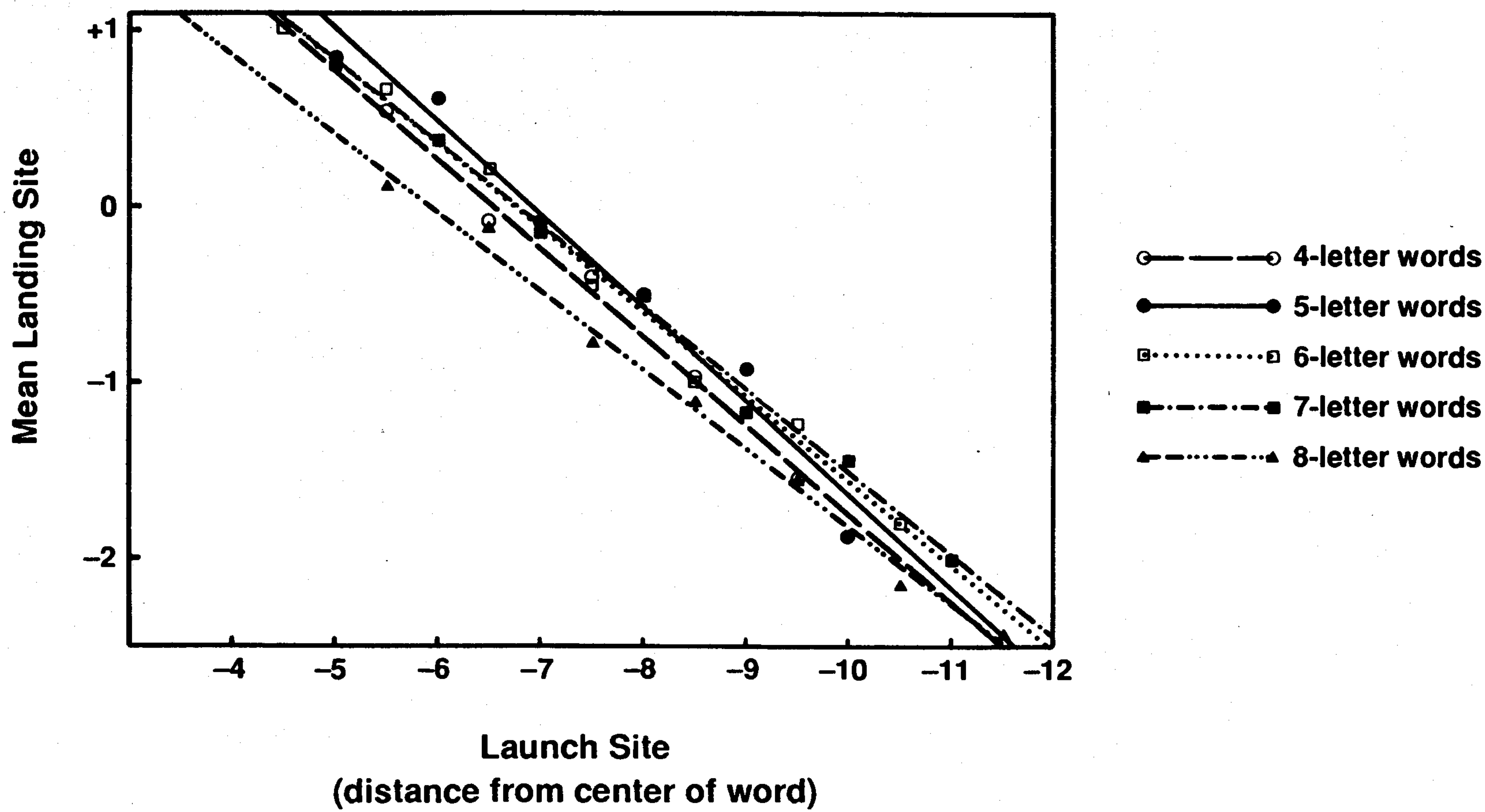
Figure 4. Estimated standard deviations of landing site distributions for different launch sites and word lengths, measured as distance from the center of the word. The data are fit by a curve of the form $Y = A \& BX^3$, where $A = 1.318$ and $B = .000518$.

Figure 5. Estimated means of the landing site distributions following fixations of different durations for different word length-launch site combinations.

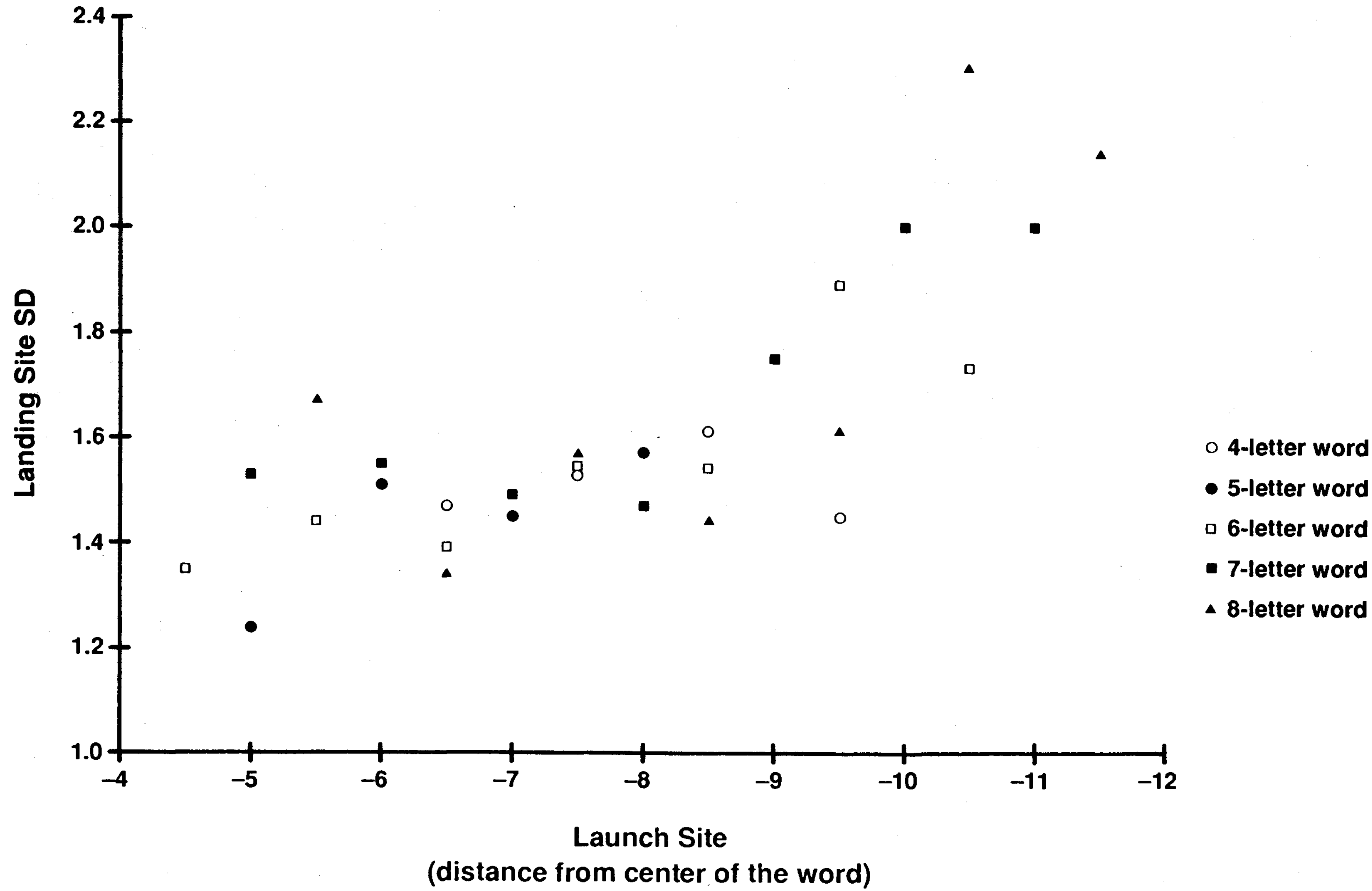


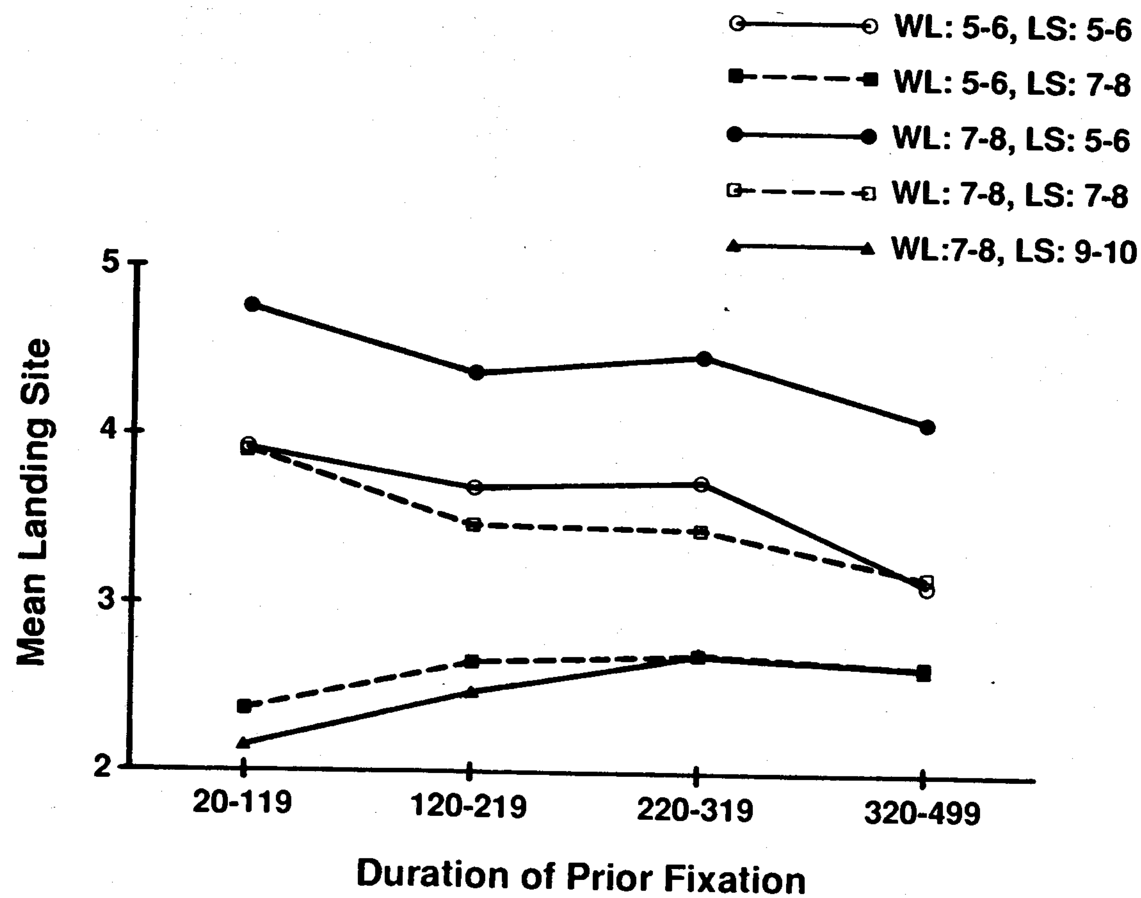
Proportion of Fixations
30
25
20
15
10
5
0
0
1
2
3
4
5
6
7
8
Letter Position in Word
length - 3
length - 4
length - 5
length - 6
length - 7
length - 8


Launch Site
4-Letter Words
6-Letter Words
8-Letter Words
-1
-3
-5
-7
% of Fixations
Character Position in Word



Mean Landing Site
+1
0
−1
−2
−4
−5
−6
−7
−8
−9
−10
−11
−12
Launch Site
(distance from center of word)
4-letter words
5-letter words
6-letter words
7-letter words
8-letter words


2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
Landing Site SD
–4
–5
–6
–7
–8
–9
–10
–11
–12
Launch Site
(distance from center of the word)
4-letter word
5-letter word
6-letter word
7-letter word
8-letter word



WL: 5-6, LS: 5-6
WL: 5-6, LS: 7-8
WL: 7-8, LS: 5-6
WL: 7-8, LS: 7-8
WL:7-8, LS: 9-10
Mean Landing Site
5
4
3
2
20-119
120-219
220-319
320-499
Duration of Prior Fixation