Eye Movement Control During Reading: A Simulation of Some Word-Targeting Strategies

Ronan G. Reilly
Department of Computer Science
University College Dublin

J. Kevin O'Regan
Groupe Regard, CNRS
University of Paris V

Abstract

McConkie, Kerr, Reddix, and Zola (1988) demonstrated that the distributions of landing sites on a word tended to be gaussian in shape. They provided a detailed account of the behaviour of the eye once a target had been selected and a saccade initiated, but said little about the process of target selection itself. The purpose of this study was to take as a starting point the landing site distributions of McConkie et al., in particular the residuals derived from fitting the gaussians to the empirical data, and to explore by computer simulation a number of saccade targeting strategies in order to discover candidates that best accounted for the residual data. Our results indicate that the strategy that gives the best fit involves targeting the longest word in a right parafoveal window extending 20 characters to the right of the currently fixated word. The implications of this finding for models of reading are discussed.

Eye Movement Control During Reading: A Simulation of Some Word-Targeting Strategies

Ronan G. Reilly
Department of Computer Science
University College Dublin

J. Kevin O'Regan
Groupe Regard, CNRS
University of Paris V

Introduction

During reading, the eye moves along the line of print in a sequence of fixations separated by saccades. Much work has been done since the beginning of the century to understand what determines where fixations fall, and what will be their durations (for example see O'Regan, 1990; Rayner & Pollatsek, 1989). Whereas work in the past used restricted reading situations and small data corpuses, in recent years sophisticated and convenient measuring devices have made it possible to gather large corpuses of eye movement data from people reading under fairly "normal" conditions. Of particular interest to this paper is the study by McConkie, Reddix, and Zola (1985), who gathered a substantial corpus of eye movement data from 66 college students reading the first two chapters of a popular novel. A portion of these were then used by McConkie, Kerr, Reddix, and Zola (1988) in an analysis of the positions where the eyes tend to land in words.

McConkie et al. (1988) demonstrated that the distributions of landing sites on a word tended to be gaussian in shape. The centre of these distributions and their standard deviations appeared to be determined primarily by oculomotor factors. Several of the distributions are represented graphically in Figure 1. The figure shows the raw data fitted with gaussian curves. Examining the panels of the figure from left-to-right, it can be seen that there is a general tendency for the eye to land around the centre of the word, as indicated by the rightward shift of distribution means with the increase in word length. Looking at the panels from top-to-bottom, there is a clear leftward shift of distribution means with the increase in launch distance. In addition there is an increase in the spread of landing site distributions, most clearly seen in the bottom panels, where the saccade lengths are longest.

McConkie et al. (1988) argued that the above pattern of results could be accounted for by five principles: (1) The centre of the word is the functional target of a saccade; (2) a systematic range error causes the eye to be increasingly deviated from this target as a linear function of distance from the launch site; (3) this range error is somewhat less, the longer the eye spends at the launch site; (4) there is a random, gaussian-shaped distribution of landing sites around the target location; and (5) the spread of this distribution increases as a function of launch distance. These five principles can be summarised in three equations. The first is a linear equation (see Equation 1) describing how the mean landing site (*m*) on a word deviates as a function of launch distance (*d*). Note that both *m* and *d* are defined to be zero at the centre of the targetted word. In the case of a four-letter word, this would be half way between the second and third letter positions.

$$m = 3.3 + 0.49\, d \qquad (1)$$

The second is a cubic equation (see 2) describing the spread of landing positions around *m*.

$$sd = 1.318 + 0.000518\, d^3 \qquad (2)$$

The third is a gaussian equation (see 3) accounting for the random distribution of landing sites, and for which *m* and *sd* are the parameters..

$$f(x;m,sd) = \frac{1}{\sqrt{2p\,.sd}}\, e^{-\frac{(x-m)^2}{2.sd^2}} \qquad (3)$$

<Insert Figure 1 about here>

Now consider what happens if we assume that these equations really do describe the eyes' behaviour. Because the assumed gaussian landing-site distributions have tails that go beyond the particular word that is being aimed for, these tails will "fatten" the distributions corresponding to the preceding and following words. If we were to accumulate landing site data over a large corpus, we would thus expect to find deviations from pure gaussian landing site distributions. The particular aiming strategy being used by the eye (e.g. "jump to each successive word", "skip short words", or "skip high frequency words") will influence the way the fattening of the distributions occurs. In McConkie et al.'s data, the fattening is most noticeable for four-letter words, and takes place near their beginnings and ends, as one might expect. To see this more clearly, Figure 2 plots residuals defined as the difference between the actual landing site distributions and the fitted gaussian curves that McConkie et al. assumed best characterised the empirical data. The residuals are most noticeable for four letter words.

It could be argued that instead of the gaussian curves that McConkie et al. fitted to the landing site distributions, an alternative night be some form of clipped gaussian, where the tails of the distribution do not overlap neighbouring words. But there seems to be no compelling evidence for this somewhat unparsimonious assumption. Moreover, there is evidence from non-reading tasks of full gaussian landing-site distributions, with similar launch-site dependent variability to that found by McConkie et al. (Kapoula, 1985). This suggests that we are dealing with a general oculomotor aiming error underlying eye movement behaviour in a variety of visual tasks.

Another objection to our approach might be that the residuals in Figure 2 are so small as to make them unworthy of study, and furthermore that there is no reason to expect the pattern to vary significantly from strategy to strategy. It turns out, however, that the pattern of residuals is quite sensitive to changes in targeting strategy. The purpose of this study, therefore, is to explore the impact of different word-targeting strategies on the pattern of landing site distributions by means of computer simulation. Prior to describing the simulations, the next section will provide some background to the selection of possible targetting strategies.

<Insert Figure 2 about here>

Targetting strategies

One can identify three main theoretical positions with respect to the factors that influence the choice of landing site in reading. These are (1) that the factors are primarily oculomotor; (2) that they are primarily linguistic, or (3) that they involve some mixture of 1 and 2.

Oculomotor strategies

Probably the simplest oculomotor strategy is that of moving the eye forward by a constant amount at each saccade, with noise in the oculomotor system giving rise to the variations in landing position that are found empirically. However McConkie et al. considered this possibility and concluded that a simple constant-saccade strategy could not account for their landing site data, and suggested that some sort of word-targeting strategy was being used. In the following sections we will briefly consider a number of such possibilities.

WORD BY WORD (WBW)

The simplest word-targeting strategy is to target each word in strict succession. There would be no influence of lexical or attentional processing on the process. The idea that such a strategy might explain a large part of eye behaviour in reading has been favoured at one time or another by McConkie et al., (1988) and by O'Regan (1990).

Because oculomotor error gives rise to a distribution of landing sites for the targetted word, we expect some degree of overshooting of targetted words, particularly when the intended target is a short word. In other cases, when the eye is coming from a long way away, the range error described above might cause the saccade to undershoot the target and land at the end of the preceding word. Both of these events result from the assumption of an underlying gaussian landing-site distribution. In either of these cases we should be able to discern a distinct pattern of under- and overshooting of target words as a result of employing this strategy. What is in question, is whether the pattern agrees with that found by McConkie et al. and shown in the residual plots in Figure 2.

TARGET LONG WORDS (TLW)

In this case it is assumed that the saccade control mechanism locks onto the longest word among the next few words in the right parafovea. In effect, the eye is drawn to the visually most salient word in the right parafovea. In defining this strategy one needs to specify the size of the region from which the longest word is selected. In our simulations we will investigate three different values for this parameter.

SKIP SHORT WORDS (SSW)

This strategy is the complement of the TLW strategy. Instead of targeting the next longest word in the right parafovea, short words are skipped over. Parameters that need to be specified for this strategy are the size of the region in the right parafovea in which the calculation of word length takes place, and the criterion for classifying a word as short. In the simulations we will use a default strategy of targetting the next word to the right, which gets overridden if that word is below a chosen length threshold. A number of such thresholds will be examined.

Linguistic strategies

There is a consensus that if linguistic factors are involved in the moment-to-moment control of eye-movements these most likely operate at the lexical level, involving such factors as word frequency, rather than higher-level syntactic or semantic properties of the text (Rayner & Pollatsek, 1989). Of course, the latter do play a role in eye movement control, but their effects tend to be lagged rather than immediate.

SKIP HIGH-FREQUENCY WORDS (SHFW)

One simple targeting possibility that takes lexical factors into account involves the reader recognising high-frequency words and skipping over them with a probability proportional to their frequency. Something like this has been proposed by a number of theorists (Rayner & Pollatsek, 1989; McConkie, 1983). The rationale is that high-frequency words, because of their frequency, their length, or perhaps their linguistic predictability, get identified in the periphery while the eye is still fixated on the preceding word, and consequently are not targeted by the next saccade, but skipped over. The corollory of this, of course, is that if the next word in the right parafovea is not high-frequency it is targetted and is not deliberately skipped.[1]

ATTENTION SHIFT (AS) STRATEGY

A more elaborate version of the preceding strategy has been proposed by Morrison (1984) and extended by Rayner and Pollatsek (1989). We will refer to this as the Attention Shift (AS) model, and it can be sketched out broadly as follows: Assume that the word currently fixated is word *n*. In the normal course of events this word will be correctly

[1] A low frequency word may, however, be skipped because of an overshoot.

identified and attention will shift to word *n+1*. Note that foveation and allocation of visual attention are assumed to be decoupled. The process of shifting attention to the next word automatically results in the programming of a new saccade. In most cases, this program is executed. However, if the shift in attention takes place early enough to allow the identification of word *n+1* without the need to foveate it, three possibilities arise: (1) word identification takes place, the programmed saccade is cancelled, and attention shifts to word *n+2*. A new saccade is then programmed and subsequently executed; (2) identification occurs too late to delay the execution of the saccade to word *n+1*. In this case, a saccade to word *n+1* is rapidly followed by a saccade to word *n+2*; and (3) the saccadic program is modified, so that the resulting saccade causes the eye to land somewhere between word *n+1* and word *n+2*. Within this framework, one can account for the skipping of high-frequency words (i.e., readily identifiable words), saccades that land between words, and the occasional very brief fixation. The attentional shift mechanism is also a way of explaining preview effects. These occur when the encoding of a word in the current fixation benefits from it having been attended on the preceding fixation. There is a considerable amount of evidence supporting the integration of some form of information across saccades which facilitates the encoding of the subsequently fixated word in both reading and non-reading tasks (Rayner & Pollatsek, 1989).

There is a variation on the Rayner and Pollatsek (1989) version of the attentional shift model due to Henderson and Ferreira (1990). They found that the amount of parafoveal preview benefit varied as a function of foveal processing difficulty; the more difficult the foveated word (either lexically or syntactically), the less parafoveal preview benefit there was. This is a challenge to the standard AS model, since the latter would predict no difference in benefit because parafoveal processing only starts when foveal processing has finished. Thus the *duration* of parafoveal processing cannot depend on foveal processing load. In order to accommodate these new results within the AS framework, Henderson and Ferreira suggested an attentional time limit, which, if exceeded, causes attention automatically to shift to the next word in right parafovea, rather than waiting for the processing of the currently fixated word to complete. Note that this feature was not implemented in the simulations presented here.

Mixed strategies

There is also the possibility that a mixture of strategies may be involved in word-targetting. For example, there might be a default strategy of word-by-word reading, used during periods of high processing demand , but which would be overridden at certain points where the text is highly predictable and/or where the processing load is light. This model differs in a subtle way from the word-by-word model in that it permits text-level characteristics to have some influence on targeting. An empirical prediction from this proposal, is that one should find less skipping of high-frequency words in demanding texts.

At the risk of limiting the generality of our experiments, we will not be exploring the mixed-strategy option. The empirical data of McConkie et al. (1988) which serves as the benchmark for our simulation experiments was collected from 66 college students reading an undemanding contemporary novel. Consequently, because of the nature of the text and the number of subjects involved we have assumed that a single uniform strategy was used throughout. In reality, we may simply be modelling the dominant strategy of several used by this type of subject reading this type of text. However, short of attempting to model individual differences, we have no way of determining whether or not this is the case. Therefore, the reader should bear these qualifications in mind when evaluating the simulation results.

Computer simulation

The problem of refixations

Before constructing a computer simulation of the different targeting strategies discussed above, it is necessary to discuss the problem of refixations. Although the McConkie et al. (1988) data that we are aiming to model do not include refixations, their occurrence in a word will influence the launch site of a subsequent non-refixation. This might have an indirect effect on the nature of the landing site distibutions. Therefore, a mechanism that produces refixations should be incorporated into the simulation.

One needs to make a distinction between *deliberate* refixations, and *non-deliberate* ones that can as a result of undershooting the next word to the right. In the emprical data, it is impossible to distinguish between these two classes of refixation. On the other hand, in the simulation, theyare two quite distinct events.

The probability of making a refixation has been studied by McConkie, Kerr, Reddix, Zola & Jacobs (1989). The authors analysed a database of 40,000 eye fixations during normal reading and showed that the eye is least likely to make refixations when it lands at an "optimal" position just left of the middle of a word. They showed that refixation probabilities could be approximated by a curve of the form:

$$y = a + .03x^2 \qquad (4)$$

where x is the deviation in character positions from the word's optimal viewing position. The lowest point of the curve, a, is dependent on word length, as follows:

$$a = 0.15l - 0.0034 \qquad (5)$$

where l is the word length in letters.

A set of curves for different word lengths based on these equations are graphed in Figure 3. Note that while x in (4) is measured with zero as the centre of the word, in subsequent equations the zero*th* position will be assumed to be the first space to left of the word. These are the equations that will be used in the computer simulations in order to program *deliberate* refixations. Note that, as a first approximation, and following arguments made by O'Regan (1990; 1992), lexical processing is assumed not to affect the likelihood of refixating. Rather it is simply the eye's landing position which, when it deviates from the "optimal" position, makes a refixation more likely. It should also be noted that these equations are based on data that combine deliberate and non-deliberate refixations, but are being used here to drive deliberate refixations. This is not considered an important confound for the purposes of this study since,as mentioned at the beginning of this section, we are only concerned with the indirect effects of refixations.

<Insert Figure 3 about here>

Triggering of refixations

The more off-centre the fixation, the shorter the refixation latency. Furthermore, as is usually the case with reaction times (cf. Luce, 1986), the variablility of the latency also decreases when the latency decrease. In order to take these factors into account, a separate distribution will be used in the simulation for each landing position in a word, reflecting a different mean latency and standard deviation. Mean latencies will be calculated as follows:

$$m = base + range\left[1 - \frac{|loc - middle|}{(length/2)}\right] \qquad (6)$$

where *base* is the minimum delay in milliseconds before a refixation is triggered, *range* is the range of values in milliseconds between the minimum and maximum refixation latencies, *loc* is the fixation location in the word with zero at the space to the left of the word, *middle* is the middle[2] character position of the word , and *length* is the word length including the first space to its left. This equation describes a Λ-shaped function centred over the word, which approximates the pattern found by O'Regan (1990). For the purposes of the simulation, *base* has been assumed to be 80 ms, and *range* 150 ms (Rayner & Pollatsek, 1989, p. 176).

The standard deviation will be taken to be a function of *m*:

$$sd = 0.1m \qquad (7)$$

after Luce's (1986) observation that the standard deviation of a reaction time distribution varies as a fixed proportion of the mean.

Lexical identification

A further temporal ingredient in the simulation is necessary to be able to simulate the Attention Shift (AS) model. In the simulation we have adopted the following simple

[2] Fractional character positions are used here. So the midpoint of a four letter word is position 2.5, while that of a five-letter word is position 3.

principle: the average time for lexical identification is assumed to be a function of the location of the letter fixated, the length of the word, and the log cultural frequency of the word. The parameters of the probability distribution for the lexical identification process are calculated in two stages. First, the average lexical identification time for a word, centrally fixated, of a given length and $\log_{10}$ frequency is calculated using the following equation:

$$m' = base5 + 15(length - 5) + 40(1 - freq) \qquad (8)$$

where *base5* (assumed to be 150 ms) is the average time taken to identify a centrally fixated five-letter word of log frequency 1.0, *length* is the length of the word, and *freq* is its $\log_{10}$ frequency. The assumption underlying (8) is that the average fixation duration of a centrally fixated word varies linearly as a function of its length and frequency. Every extra letter in a word adds an additional 15 ms onto the base recognition time. Thus, a six-letter word will incur a 15 ms recognition penalty, while a four-letter word will have a 15 ms advantage. This figure is derived from the work of O'Regan and Jacobs (1992; p187) who made an extensive study of lexical decision times and naming latencies for isolated words of different lengths and frequencies. They found the length penalty to be around 15-19 ms. We have used the lower bound of 15 ms because of the tendency for effects obtained in isolated word experiments to be diminshed in real reading situations (Vitu, 1991). For every additional unit of $\log_{10}$ frequency there is a multiplier of 30 ms. As can be seen from (8), for lower frequency words this represents a recognition penalty, and for higher frequency words it is a recognition advantage. Again, as with the length effect, the figure of 40 ms for a frequency effect is the lower bound for that effect found by O'Regan and Jacobs (1992) in their experiments with isolated words.

Since equation (8) only gives recognition times for centrally fixated words, we now need to generalise these values for different fixation locations on the same word using the following equations:

$$m = m' + \frac{range}{(length/2)}\left|loc - middle\right| \qquad (9)$$

where *loc* is the fixation location with the space to the left of the word as zero, *length* is the length of the word including the leading space, *middle* is the middle character position of the word, and *range* defines the gap in milliseconds between the minimum and maximum recognition latencies. While the minimum will vary depending on the value of *m'*, the gap between minimum and maximum is a constant. If we graph (9) with *m* on the y-axis and *loc* on the x-axis for words of the same length, we obtain a set of V-shaped functions, centred on the middle of the word, and displaced vertically as a function of frequency. The V has a minimum at *m'*, its left arm has a slope of $-\frac{range}{(length/2)}$, and its right arm a slope of $+\frac{range}{(length/2)}$. By way of illustration, the means for lexical identification distributions for four- and eight-letter words of varying frequency are graphed in Figure 4. For all of the simulations described here, *range* had a value of 100 ms.

The resulting *m* value, along with the *sd* value derived from (7), are then used as parameters for a probability distribution. If the value generated from the distribution is less than 80 ms, it is set to 80 ms. This is is to enforce a lower bound on the time taken for lexical identification.

<Insert Figure 4 about here>

Saccade triggering

In the simulation, we assume that (unless a refixation is called for) a saccade leading out of the word is programmed immediately after lexical identification. In the program, a distinction is made between programming a saccade and executing it. Within this framework a saccade cannot occur immediately: we assume that the time between programming the saccade and it actually occurring is a purely oculomotor delay governed by a probability distribution whose mean and standard deviation are assumed to be constant, irrespective of lexical considerations, such as word frequency, word length, or location fixated. Although there is some evidence that saccade latencies may vary inversely with the size of the saccade (Kowler & Anton, 1987), we have assumed for simplicity that for all simulations the average saccade latency is 150 ms, with a standard deviation of 50 ms.

<Insert Table 1 about here>

Program overview

The simulation program takes as input "text" comprising word-length and word-frequency information derived from an actual text. The program then reads this text in a loop, checking first to see if the eye has landed near the preferred viewing position of the current word and whether a refixation should be programmed. There then follows a competition between possibly multiple saccadic programs (refixation and/or progressive) that have been programmed but not yet triggered. A saccade is triggered on the basis of a cumulative probability function (CPF), and the whole process starts over again. The algorithm is summarised in pseudo-code in Table 1.

CPFs play a significant role in the operation of the simulation. They describe the probability of an event occurring as a monotonically increasing function of time. Depending on the spread of the underlying distribution, the CPF will vary in steepness; a narrowly distributed distribution will give a steep rise in probabilities, a wide distribution will give a more gentle rise.

A variety of targeting strategies are built into the program. These strategies vary in which target word they select for the next fixation. The ultimate landing position on the selected target is determined probabilistically, where the probability is given by the systematically varying gaussians described by McConkie et al. (1988). The parameters of the gaussians (i.e., their mean and standard deviation) are calculated in one of two ways: (1) from the empirical data for word lengths and launch distances provided in Table 1 of McConkie et al. (1988), and (2) from Equations 1 and 2 for those word length and launch distance combinations not included in Table 1.

The features of the simulation program common to all targetting strategies are:

1. The programming of a *refixation* is determined independently of any lexical processes and depends only on the initial landing position of the eye in a word;

2. A distinction is made between programming a saccade and executing it. In the simulation, a saccade is programmed either when a refixation has been decided upon or when a word has been identified. Each of these events is based on a probability derived from two separate CPFs. Once a saccade has been programmed, it is triggered with a probability determined by a third CPF;

3. More than one saccadic program can await triggering, but only one will be triggered. There is provision in the simulation program used for the interaction of temporally adjacent saccadic programs, as proposed by Morrison (1984), but this has not been exploited in the current study.

Simulation Results

A number of simulations was run, each of which implemented a different targeting strategy. The text used was the same two chapters from the popular novel used in the McConkie et al. (1985) study. Note that only word length and word frequency information were used in the simulation. In all of the analyses that follow, each model's landing site distributions were subtracted from the underlying distributions used to generate the landing positions, giving a set of residuals. In all cases, the underlying distribution is a gaussian, the mean and standard deviation of which were either provided by the data from McConkie et al's (1988) Table 1 or by Equations 1 and 2 above.

What we are looking for is a pattern of residuals that provide the closest match to the pattern found in the McConkie et al. (1988) study, as illustrated in Figure 2. In all cases, the residual plots represent the average of 20 separate runs of the simulation, each using a different initial random seed. In addition, while refixations were permitted, only first fixations on a word are included in the analyses, as was the case in the McConkie et al. (1988) study. For strategies involving a number of different parameter settings such as SHFW and TLW, a table of correlations is used to help compare the empirical residuals with those derived from the simulations. The correlation coefficient used is a concordance measure, r_c (Lin, 1989), which measures the agreement of data which are measured on a continuous scale. The concordance coefficient ranges from -1 to +1.

Word-by-Word Strategy

Let us look, first, at the kind of residual effects obtained when a simple word-by-word strategy is adopted. The residuals are plotted in Figure 5. The most striking differences between these and the residuals in Figure 2 is the poor fit for landing positions at the beginning of all word lengths, and for landing positions at the end of four-letter words. There is, however, a slight trend from positive to negative residuals as word length increases, which agrees with the pattern found in McConkie et al.'s (1988) analysis.

<Insert Figure 5 about here>

It is clear, however, that the word-by-word strategy is not a realistic candidate for target selection, given the dramatically elevated word-initial landing positions. These arise from an excess of target overshoots, causing the landing positions to shift to the right adjacent word. We do not find a corresponding elevation at word-ends because the the majority of saccades in a word-by-word strategy will of necessity be short and will tend to overshoot their target (see Equation 1). There would be fewer overshoots if words further out into the right parafovea were targetted. This is, in effect, what occurs in the strategies described below.

<Insert Table 2 about here>

Skip short words

The main parameters of this strategy are (1) the criterion for classifying a word as short, and (2) the size of the region in the right parafovea within which words are considered for skipping. Four sets of simulations were run involving the pairing of window size (10 and 15 characters) and length (less than four character, and less than five characters). The table of correlations for the SSW strategy are given in Table 2. Although the general pattern of correlations for word-lengths six and eight is comparable, as we shall see, to other strategies, the correlations for the word-length four residuals are very low. This again makes the SSW strategy an unlikely candidate for the one underlying the empirical data.

Target the Longest Word

This strategy involves selecting the longest word within a predefined window to the right of the currently fixated word. It is assumed that the target word is selected on the basis of its visual weight. There are a number of parameters to be defined in this strategy. The first is the size of the window in which the target selection takes place. For the simulations described here, three window sizes of 10, 15 and 20 characters were used, where the start of the wondow was calculated from the space to the left of the next word. If the window boundary straddled a word, that word could still be choosen as a target. In addition, the probability with which a jump was made to the selected target was varied. The probability was a function of the current fixation duration. An arbitrary set of probabilities could have been chosen, but we decided that it was more realistic to tie the probability to a feature of the reading process. The assumption underlying the choice of fixation duration was that if the current word is fixated for sufficiently long enough, there is a greater likelihood of selecting the next visually most interesting word (i.e., the longest) for a saccade. The following function was used:

$$p = \mathbf{min}\{slope \times fixdur, 1\} \qquad (10)$$

where *slope* could have one of three values: .002, .004, and .008. For an average fixation duration of 250 ms, this meant that the target would be selected with a probability of .5, 1.0, and 1.0 respectively. The slope of .008 was designed to ensure that there was a value for which the target was selected for even very brief fixations. The function **min** returns the smallest of its arguments. When the target was not selected, as could be the case with a slope of 0.002, the next word to the right was the target.

<Insert Table 3 about here>

Table 3 gives the results of correlating the residual pattern obtained for each of the window/slope combinations with the residual pattern obtained by McConkie et al. (1988). Correlations were calculated separately for different word lengths, and involved combining results from the four launch sites (-1, -3, -5, and -7) and all landing positions. The highest average correlation for all word lengths (.33) was obtained for the simulation involving a 20 character window, and a slope of .004. The best single correlation (.5) was obtained for four-letter words with a window of 20 characters and a slope of .008.

<Insert Figure 6 about here>

The residual plots for the highest correlating version of the TLW strategy are given in Figure 6. In general, the graphs of the residuals are a lot smoother than those obtained empirically, particularly for landing sites in the middle of words. The critical aspects of the residual patterns are, however, the beginnings and ends of words, and here the fit is quite good particularly for words of length six and eight. For four letter words, although the overall correlation is high, the simulation residuals at word beginnings are lower than the empirical data, and are higher than the empirical data at word endings.

Skip High-Frequency Words

This strategy was implemented by getting the model to skip over high-frequency words in the right parafovea with a probability that was a linear function of the word's log cultural frequency:

$$p = \mathbf{min}\{(0.1 + slope \times freq), 1\} \qquad (11)$$

where p is the probability of skipping, *slope* could either be .1, .2, or .3, and *freq* was the log frequency. The slope value of .2 ensured that only words with a log frequency of 4.5 or greater were certain to be skipped (i.e., the articles *a* and *the*). The other slope values could either increase or decrease the probability of skipping. Another factor in this strategy was the number of words in the parafovea to assess for skipping. Three window sizes were used: 5, 10, and 15 characters. As with the TLW strategies, the size of the window was measured from the space to the right of the next word.

<Insert Table 4 about here>

Table 4 gives a breakdown of correlations between the residual patterns from the McConkie et al. data and those generated by the simulation at the various parameter settings. The best average correlation (.31) for all landing sites is given by the parameter combination of a 15 character window and a slope of .2. In the 10 and 15 character window sizes, the correlations for four-letter words are reasonably respectable, but tend to be lower for the longer word lengths. The reason for this can be seen in the Figure 7, where there is a definite curvilinear trend apparent, particularly for word-length 8. This trend is not present in the empirical data.

<Insert Figure 7 about here>

Attention Shift

The attentional shift strategy was implemented in the following way: Once a given word was recognised, and prior to a saccade being executed, attention was shifted to the next word in the right parafovea. This is operationalised in the model as an attempt to recognise a non-centrally fixated word. The lexical identification CPF is simply a generalisation of the ones displayed in Figure 4. For example, to derive the mean recognition time for a high-frequency four-letter word being attended to from the last character of the preceding word (i.e., landing site -1), we simply extend the lowermost V for the four-letter words in Figure 4 one position to the left. In addition, a component of the time accounted for by the afferent lag (assumed to be 50 ms) is subtracted from the mean identification latencies for the second and subsequent words. The afferent lag refers to the time it takes for information to reach the visual cortex from the retina. The motivation for subtracting this value is that the attentional shift mechanism is operating on some form of internally stored representation of the visual input, rather than having to await its processing through the lower visual pathways.

<Insert Table 5 about here>

The results indicate that the residual pattern found for the attentional shift (AS) strategy is almost identical to that of the word-by-word strategy (see Figure 8). The reason for this becomes clear when we look at the probabilities of skipping, say, four letter words from near launches (i.e., the word to the right), and compare them to the same data for the WBW strategy (see Table 5). As can be seen, there is almost no difference between the two sets of probabilities. This indicates that the AS and WBW stategies are behaving almost identically, implying that the only word skipping going on in AS is based on overshooting rather than successful parafoveal identification.

The lexical processing time estimates used in the AS model are based on those of Rayner and Pollatsek (1989; p. 176), who have provided the clearest articulation of the model. What the simulation results show is that there just is not enough time within the constraints of these times to identify more than a very few words using the attentional shift mechanism. Moreover, the model has not assumed any time penalty associated with shifting attention, something for which there is evidence (Posner, 1980, p. 16).

<Insert Figure 8 about here>

In order to see if the rate of probability of word-skipping could be increased, one of the parameters involved in lexical identification was adjusted. The average time taken to identify a centrally-fixated five letter word (*base5* in equation 8) was assumed to be 100 ms as opposed to 150. This only increased the probabilities marginally.

How critical are the TLW and SHFW parameter values?

In the preceding section we focused on the choice of parameter values for the AS strategy. The question also arises as to how critical is the selection of parameter values when comparing the two most successful targeting strategies, TLW and SHFW. For example, might there be a set of values that would make the SHFW strategy better than TLW?

Examining the TLW strategy first, we can see that the window size parameter has the biggest affect on the pattern of correlations in Table 3. The range of values selected for this parameter is obviously limited by what we know of the acuity of the visual system, so even a value of 20 is starting to get psychophysically unrealistic. For example, McConkie and Rayner (1975) showed that there was no significant effect on readers' eye movement behaviour when letter-space information was removed from the text further than 15 spaces to the right of the current fixation. As regards the value of the slope parameter, it seems that a value of.004 gives the best performance.

In the case of the SHFW strategy, again the window parameter has the biggest effect on correlation values, with the best performance obtained from a window size of 15 character spaces (see Table 4). However, one must keep in mind that given the fall-off in acuity into the parafovea, the likelihood of accurate identification of even a high-frequency word further than 10 characters beyond the end of the current word is remote. Therefore, extending the window size beyond the 15 characters used in the simulations would be quite unrealistic. Taking this argument further, possibly a more psychophysically realistic choice of SHFW strategy would be the one involving a 10 character window rather than the 15 character one we have chosen, despite the latter giving a better fit. If we were to do this, TLW would look an even stronger candidate for best targeting strategy.

Finally, the slope parameter for the SHFW strategy did not give rise to much variation in correlation values, with the intermediate value of .2 generally giving the best fit, suggesting that the optimal value was in the range .1 to .3, with .2 a good estimate.

In general, the choice of parameter values plays an important role in the performance of the implemented strategies, but the parameter values cannot be sampled from an infinitely large space if the modelling is to conform to the assumed constraints of visual processing. We have argued above that the choice of parameters in the most promising strategies is nearly optimal and should not affect the outcome of strategy comparisons. Furthermore, any change to the most important of these, window size, would increase the relative advantage of the TLW strategy.

Implications for a model of reading

We have seen that the simplest strategy of just moving forward word by word (WBW) does not provide a good fit to the McConkie et al. (1988) data. The best account is provided by a strategy which targets the longest word in the right parafovea most, but not all, of the time. The longest word is not targeted all of the time, because the size of the slope of the most successful TLW targeting function (.004) means that when the fixation duration on the current word is significantly less than average, a saccade is made to the next word rather than to the next longest word.

The "cleverer" strategy of skipping high frequency words (SHFW), does not give quite as good a fit. Worse still is the attention shift (AS) model of Rayner and Pollatsek (1989). As we have shown, current time estimates for the components of the lexical identification process in the AS model permit very few multiple word identifications on a single fixation. This suggests two possibilities: (1) that the time estimates are incorrect and that word identification takes considerably less time than has heretofore been assumed; or (2) that the model is incorrectly formulated. Since the time estimates seem quite reliable, and find suport form a number of sources, we feel that the details of the AS model may need some revision.

Conclusions

This paper started off with an analysis of the somewhat less than perfect fit that McConkie et al. (1988) found for their gaussian model of landing site distributions for some word lengths and landing positions. We assumed that the main source of this lack of fit was over- and undershoots from attempted landings on neighbouring words. These in turn were determined by which words were the functional target for a particular saccade. A set of targeting strategies were proposed and computationally simulated. The "word by word", the "target long words" and the "skip short words" strategies made no use of lexical processing for their execution, whereas the "skip high frequency words" and the "attention shift" strategies required that enough time be available for lexical processing of the currently fixated word to influence the eye's immediate behaviour. The strategy that gave the best fit to the data involved targeting the longest word in a right parafoveal window that extended 20 characters to the right of the currently fixated word. The two strategies requiring lexical processing were distinctly less satisfactory in accounting for the data. We argued, furthermore, that adjustments of the parameters of our simulation could probably not improve the fit of such strategies.

The results therefore suggest that the eye movement guidance system does not generally use linguistic information, but exploits word-length information in the right parafovea to target the next saccade. Recent work by Legge, Klitz, and Tjan (in press) also supports this view. Of course, adopting the hypothesis of a word-length based mechanism raises the issue of what happens to the skipped words, if their being skipped is not contingent on their immediate identification. Rather than speculate, we suggest that this is a question for further research.

Another issue which needs consideration is how might the TLW strategy deal with text in which the spaces have been removed, or with languages which have large compound words, such as German. As implemented in its present form, the simple answer is that the TLW strategy could not cope. However, it could be formulated more generally to deal with a variety of boundary types, such as spaces, and letter transitions indicating possible intra-compound boundaries. In this case, our model would predict that the longer component of a compound would be the effective target for a saccade. Another very likely possibility is of course that in German, and in reading text without spaces, readers adopt other strategies than the ones investigated here.

Finally, we think a noteworthy aspect of our results is the fact that simulations of even subtly different targeting strategies have yielded substantial differences in predicted landing site distributions, and that by comparing these to the available empirical data we have been able to narrow down effectively the field of possible explanations for eye movement control in reading.

Acknowledgements

This paper was completed while the first author was Fellow at the Netherlands Institute for Advanced Study, Wassenaar. The research was supported in part by an EOLAS/CNRS travel grant to the first author. Some of the simulations were carried out using the resources of the Centre for High Performance Computing Applications at UCD.

References

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

Coëffé, C., & O'Regan, J. K. (1987). Reducing the influence of non-target stimuli in saccade accuracy: Predictability and latency effects. *Vision Research, 27,* 227-240.

Deubel, H., Wolf, W., & Hauske, G. (1984). The evaluation of the oculomotor error signal. In A. G. Gale & E. Johnson (Eds.) *Theoretical and applied aspects of eye movement research.* North-Holland: Amsterdam, pp 55-62.

Eriksen, C. W., & St James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics, 40,* 225-240.

Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research, 21,* 347-354.

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye-movement control. *Journal of Experiemental Psychology: Learning, Memory, and Cognition, 16,* 417-429.

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.

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

Legge, G.E., Klitz T. S., & Tjan; B.S. (In press). Mr. Chips: An ideal-observer model of reading. *Psychological Review*.

Levi, D.M., Klein, S.A., & Aitsebaomo, A.P. (1985). Vernier acuity, crowding, and cortical magnification. *Vision Research, 25,* 963-977.

Lin, L.,I.-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics, 45,* 255-268.

Luce, R.D. (1986). *Response times*. New York: Oxford University Press Inc.

McConkie, G. W. (1983). Eye movements and perception during reading. In K. Rayner (Ed.), *Eye movements in reading. Perceptual and language processes.* Academic Press: New York, pp. 65-96.

McConkie, G. W., Kerr, P. W., Reddix, M. D., and Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28,* 1107-1118.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception and Psychophysics, 46,* 245-253.

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

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

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

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

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

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

Posner, M. I. (1980). The orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiquity. *Memory and Cognition, 14,* 191-201.

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

Vitu, F. (1991). The existence of a centre of gravity effect during reading. *Vision Research, 31,* 1289-1213.

Zola, D. (1984). Redundancy and word perception during reading. *Perception and Psychophysics, 36,* 277-284.

Table captions

Table 1

Psuedo code for the simulation program.

Table 2

Table of correlations for various parameter settings of the Skip Short Words (SSW) strategy. The correlations are between the residual for each landing position from the SSW strategy and the empirically derived residuals from the study by McConkie et al. (1988), and are broken down by word length.

Table 3

Table of correlations for various parameter settings of the Target Longest Word (TLW) strategy. The correlations are between the residuals for each landing position from the TLW strategy and the empirically derived residuals from the study by McConkie et al. (1988). Correlations in bold face indicate statistical significance at less than the 0.05 level.

Table 4

Table of correlations for various parameter settings of the Skip High Frequency Words (SHFW) strategy. The correlations are between the residuals for each landing position from the SHFW strategy and the empirically derived residuals from the study by McConkie et al. (1988). Correlations in bold face indicate statistical significance at less than the 0.05 level.

Table 5

Comparison of word skipping probabilities between the attention shift (AS) and word-by-word (WBW) strategies. The probabilities are broken down by log frequency of the skipped word (rounded to the nearest whole number), and the launch site of the saccade in terms of characters from the space prior to the (possibly) skipped word.

Figure captions

Figure 1

The raw data from McConkie, Kerr, Reddix, & Zola (1988) fitted with Gaussian curves (their Figure 2). There is a general tendency for the eye to land around the centre of the word, as indicated by the rightward shift of distribution means with an increase in word length. There is also a clear leftward shift of distribution means with an increase in launch distance. Finally, there is an increase in spread of landing site distributions, most clearly seen in the bottom panels, where the saccade lengths are longest.

Figure 2

The pattern of residuals found by McConkie, Kerr, Reddix, & Zola (1988) when they fitted gaussian curves to their landing site distribution. Note the preponderance of positive residuals in the four-letter word case, and the tendency towards negative residuals with increasing target word length.

Figure 3

The probability of refixating words of length 4-8 as a function of fixation location. Each of the curves can be described by the equation: $y = a + 0.3x^2$ where the offset a is a linear function of word length (l): $a = 0.15l - 0.0034$. Note that x is measured as characters from the optimal viewing position of the word (about 0.5 of a a character space to the left of the word's centre)

Figure 4

Plots of mean word recognition times as a function of word of word frequency and landing site. The means are determined by the set of cumulative probability functions described in the text.

Figure 5

Residual landing-site plots for the word-by-word (WBW) strategy for words of four, six, and eight letter words.

Figure 6

Residual landing-site plots for the target-longest-word (TLW) strategy for words of four, six, and eight letter words.

Figure 7

Residual landing-site plots for the skip-high-frequency-word (SHFW) strategy for words of four, six, and eight letter words.

Figure 8

Residual landing-site plots for the attention-shift (AS) strategy for words of four, six, and eight letter words. These results are from the simulation in which parameters of the lexical idenification functions were changed in order to maximise speed of identification.

Table 1

```
while there are words to read do
      start clock
      while saccade program has not been executed do
            program one refixation with a
                  probability based on the
                  refixation cumulative probability
                  function (CPF)
            program forward saccade(s) with
                  probability based on the lexical
                  identification CPF
            if there are any saccades programmed then
                  execute the associated saccade with a probability
                  based on the saccade execution CPF.
            endif
            increment clock
      endwhile
      extract word from current fixation
endwhile
```

Table 2

Parameter settings		Word lengths			Average r_c
Window (chars)	Skip words less than (chars)	4 letter	6 letters	8 letters	
10	4	-0.104	0.197	0.242	0.112
10	5	-0.121	0.206	0.217	0.101
15	4	-0.121	0.207	0.222	0.103
15	5	0.017	0.124	0.102	0.081

Table 3

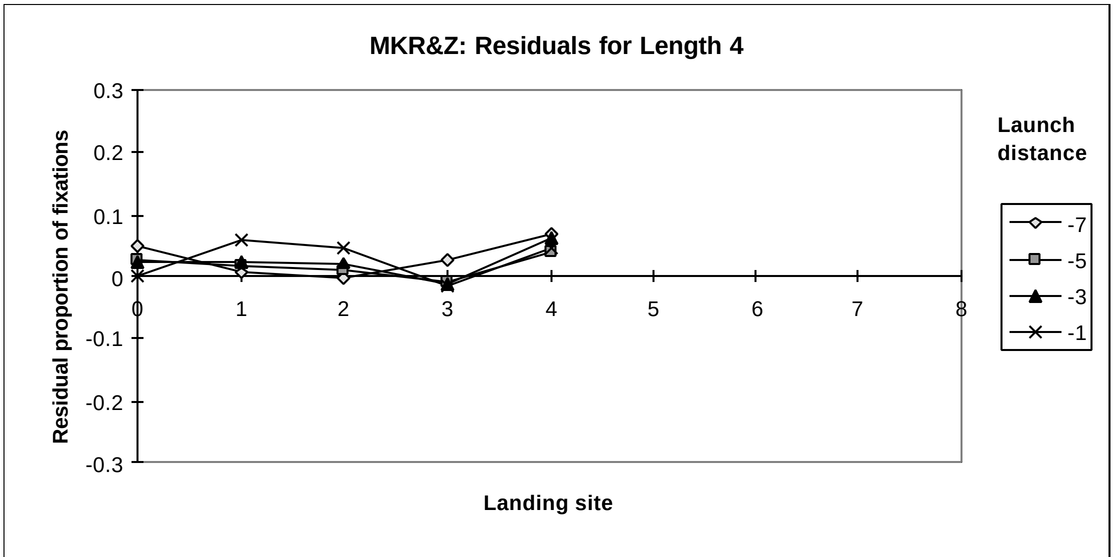
Parameter settings		Word lengths			Average r_c
Window (chars)	**Slope of probability function**	**4 letter**	**6 letters**	**8 letters**	
10	.002	0.219	0.156	0.232	0.202
10	.004	0.106	0.084	0.175	0.121
10	.008	0.318	0.158	0.120	0.199
15	.002	0.320	0.158	0.259	0.245
15	.004	**0.432**	0.221	0.154	0.269
15	.008	0.416	0.295	0.171	0.294
20	.002	0.366	0.197	0.255	0.273
20	.004	**0.470**	0.327	0.205	0.334
20	.008	**0.496**	0.287	0.162	0.315

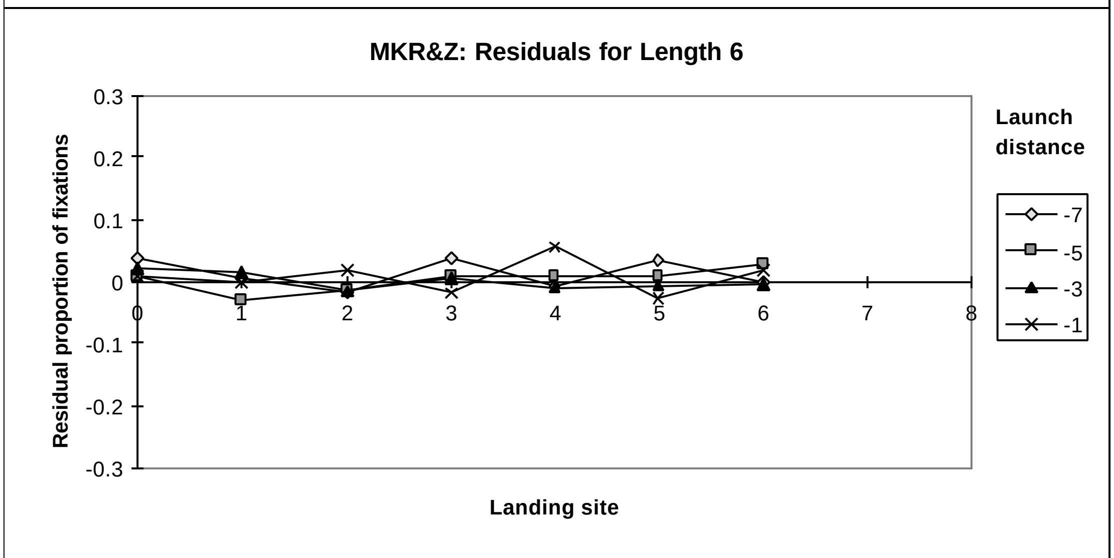
Table 4

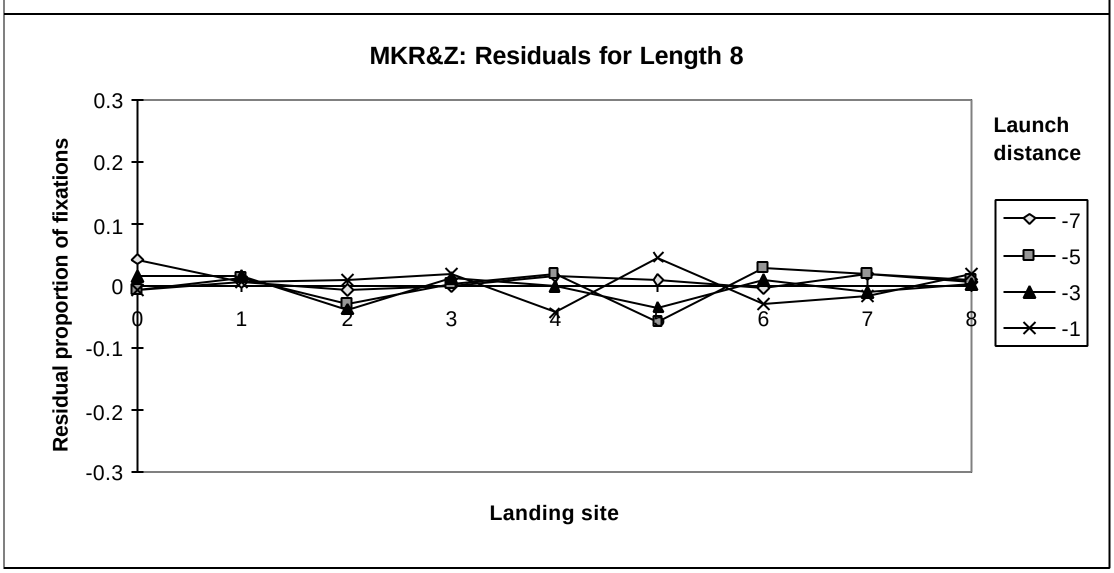
Parameter settings		Word lengths			
Window (chars)	**Slope of probability function**	**4 letter**	**6 letters**	**8 letters**	**Average r_c**
5	.1	0.091	0.143	0.209	0.148
5	.2	-0.025	0.204	0.248	0.143
5	.3	0.091	0.211	0.260	0.188
10	.1	0.304	0.176	0.232	0.238
10	.2	**0.436**	0.183	0.180	0.266
10	.3	0.294	0.095	0.197	0.196
15	.1	0.304	0.214	0.237	0.252
15	.2	**0.416**	0.314	0.191	0.307
15	.3	0.317	0.131	0.143	0.197

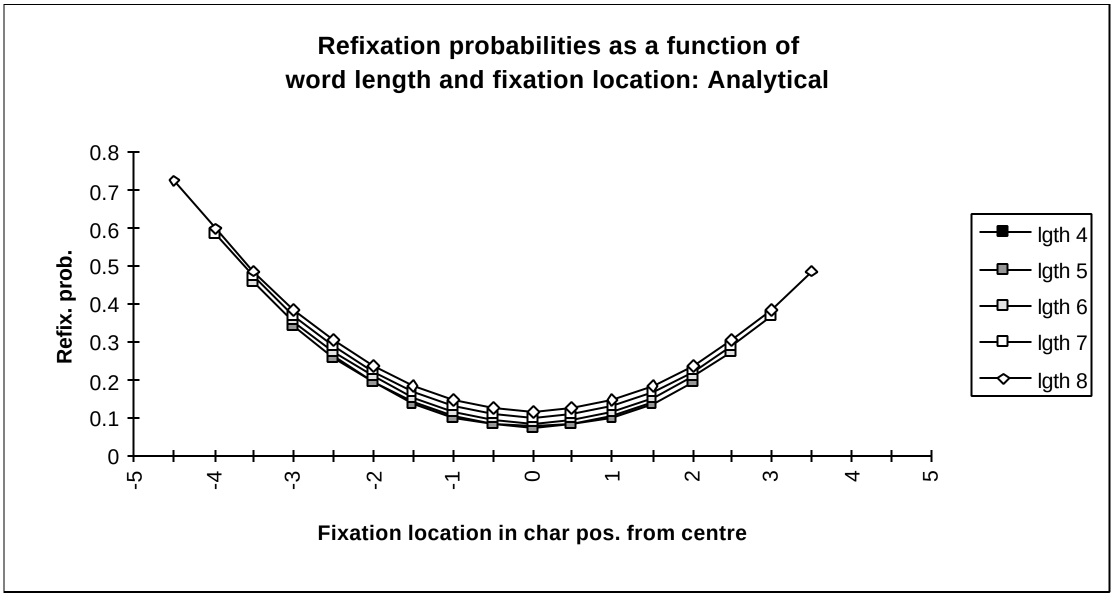
Table 5

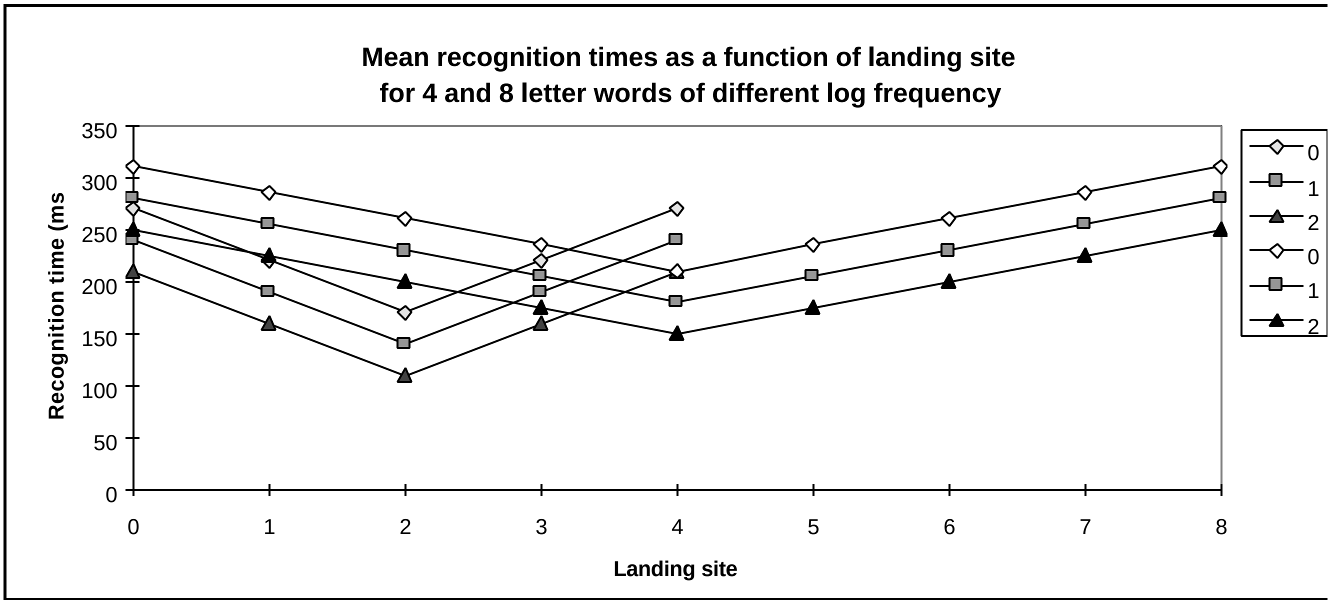
		Launch site			
Strategy	**Freq.**	**-4**	**-3**	**-2**	**-1**
attention shift	0	0.18	0.22	0.36	0.44
	1	0.20	0.28	0.34	0.46
	2	0.20	0.25	0.31	0.41
	3	0.18	0.26	0.34	0.44
	4	0.18	0.25	0.34	0.44
word-by-word	0	0.22	0.27	0.33	0.44
	1	0.20	0.26	0.36	0.42
	2	0.20	0.25	0.32	0.42
	3	0.18	0.25	0.34	0.44
	4	0.18	0.25	0.34	0.44


MKR&Z: Residuals for Length 4
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
Launch distance
-7
-5
-3
-1

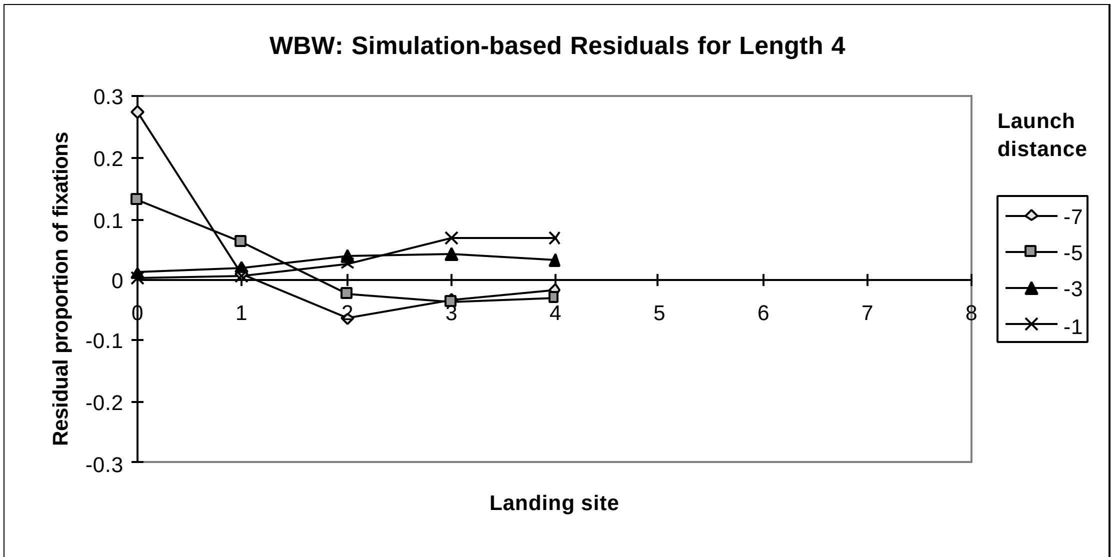

MKR&Z: Residuals for Length 6
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
Launch distance
-7
-5
-3
-1

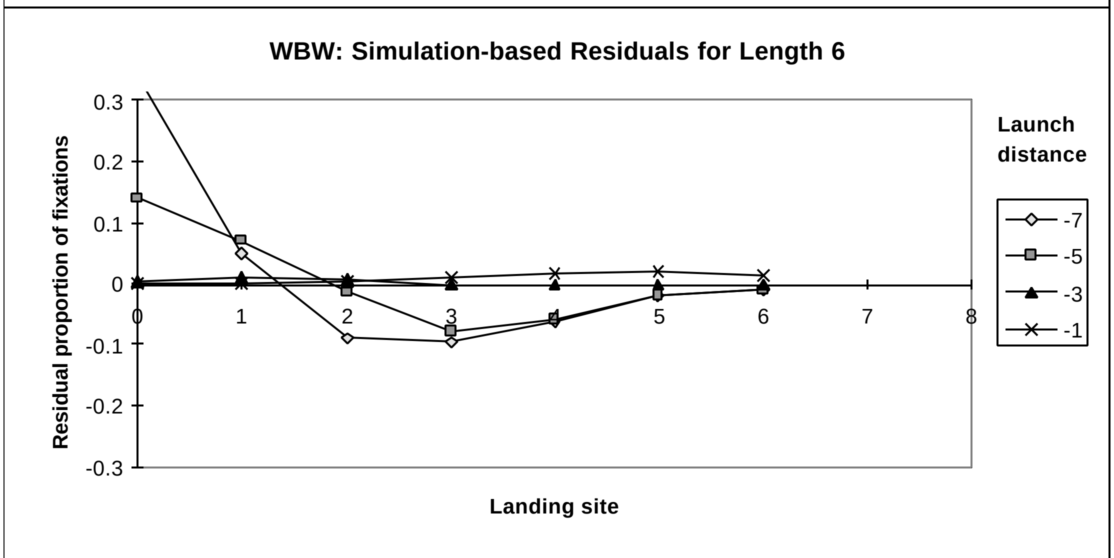

MKR&Z: Residuals for Length 8
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
Launch distance
-7
-5
-3
-1

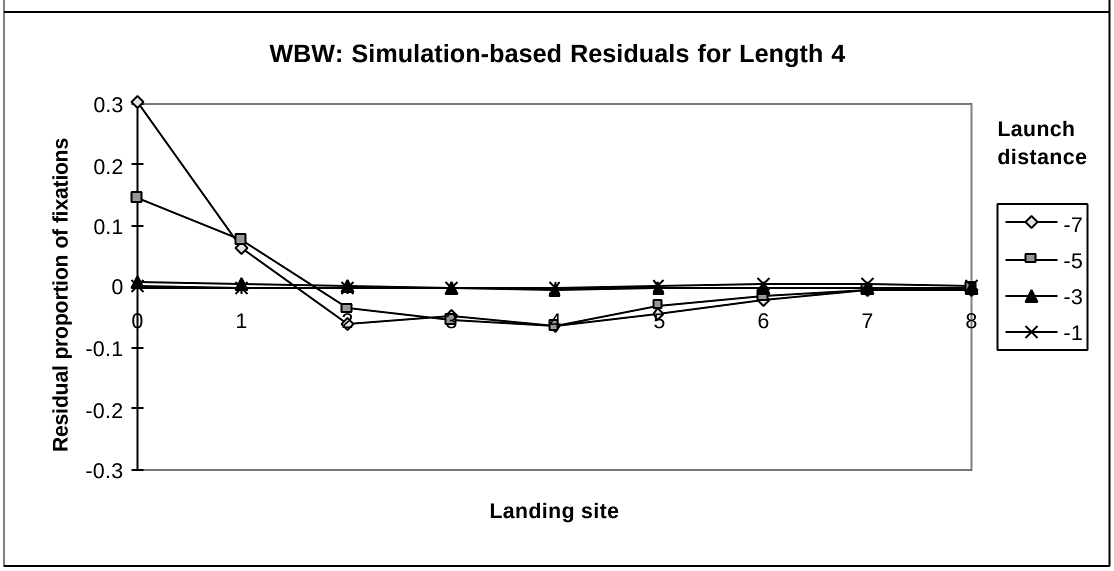

Refixation probabilities as a function of
word length and fixation location: Analytical
Refix. prob.
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-5
-4
-3
-2
-1
0
1
2
3
4
5
Fixation location in char pos. from centre
lgth 4
lgth 5
lgth 6
lgth 7
lgth 8

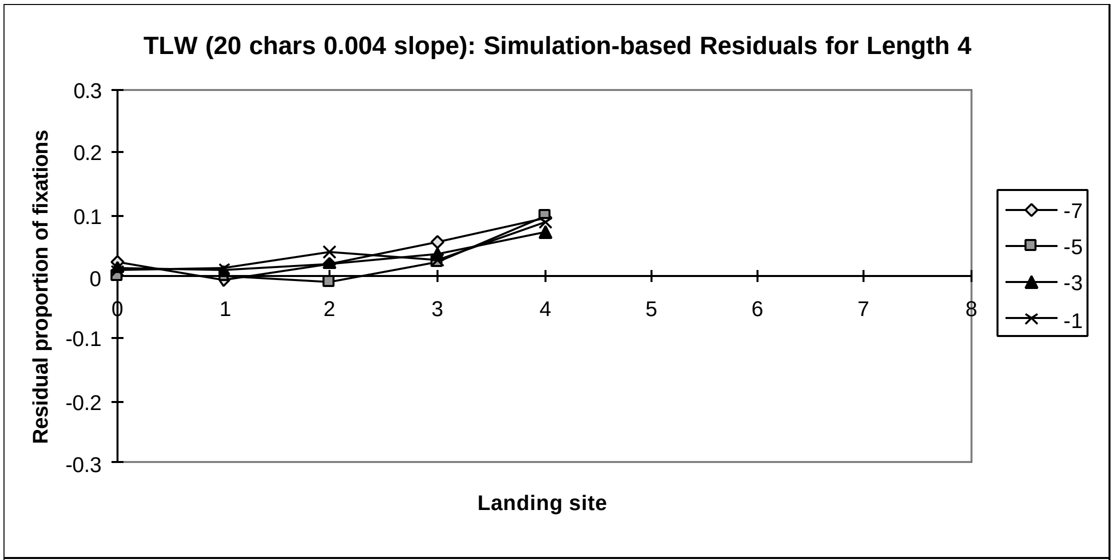


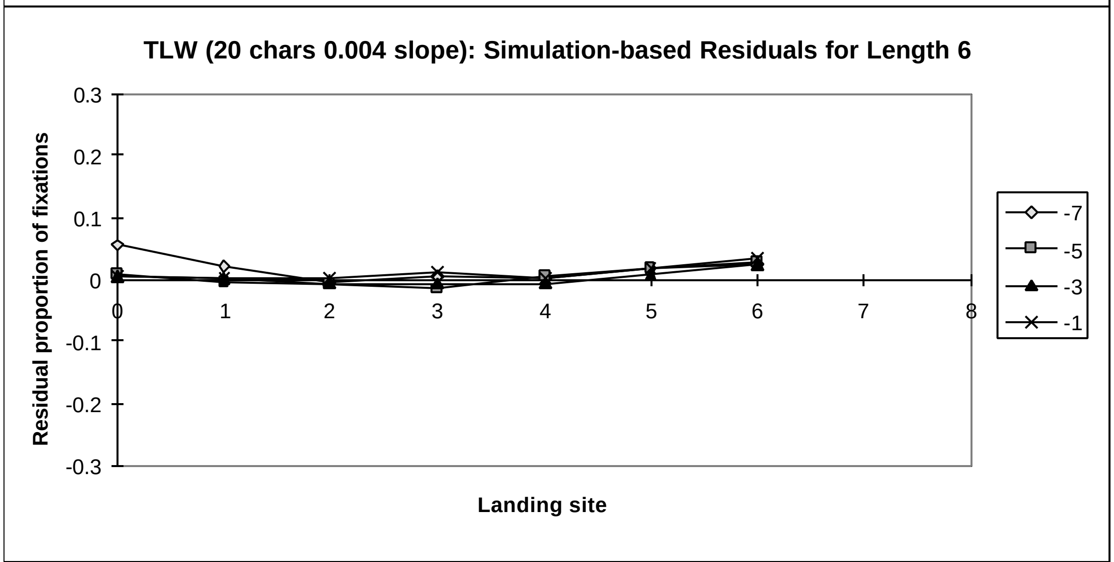
Mean recognition times as a function of landing site
for 4 and 8 letter words of different log frequency
Recognition time (ms
Landing site
0
50
100
150
200
250
300
350
0
1
2
3
4
5
6
7
8
0
1
2
0
1
2

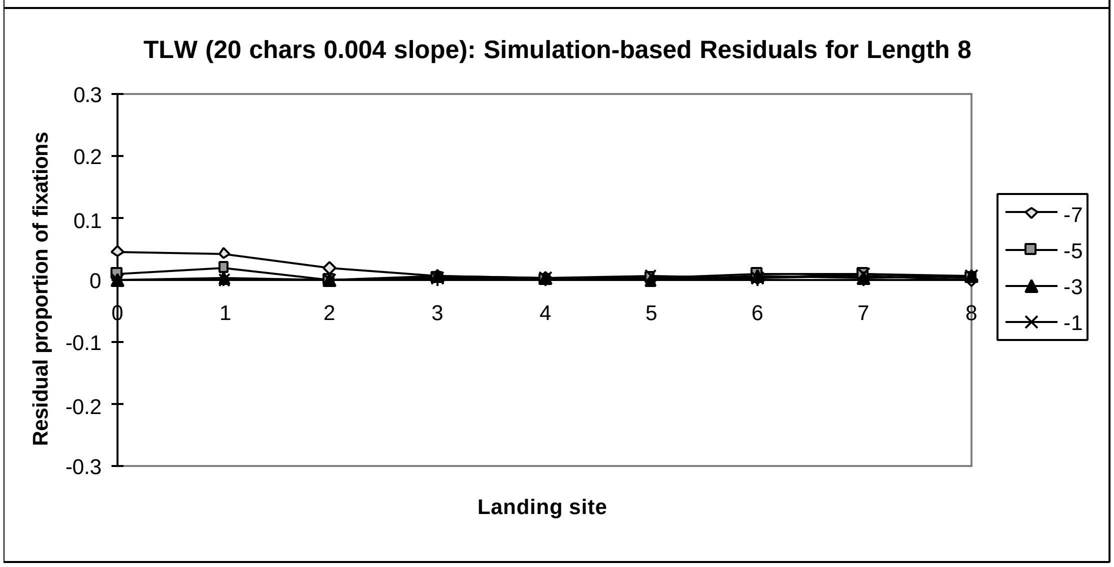

WBW: Simulation-based Residuals for Length 4
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
Launch distance
-7
-5
-3
-1

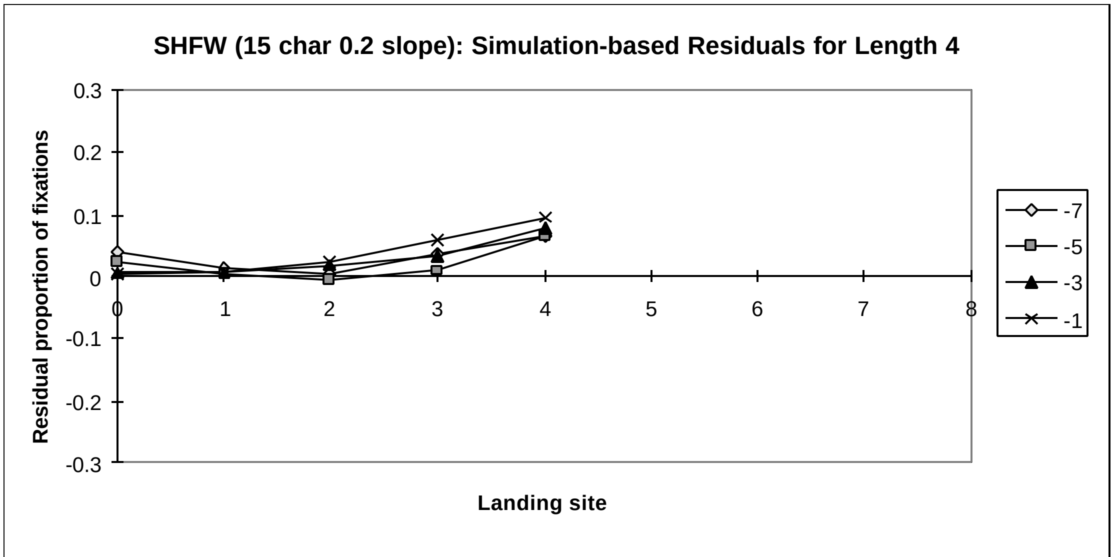

WBW: Simulation-based Residuals for Length 6
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
Launch distance
-7
-5
-3
-1

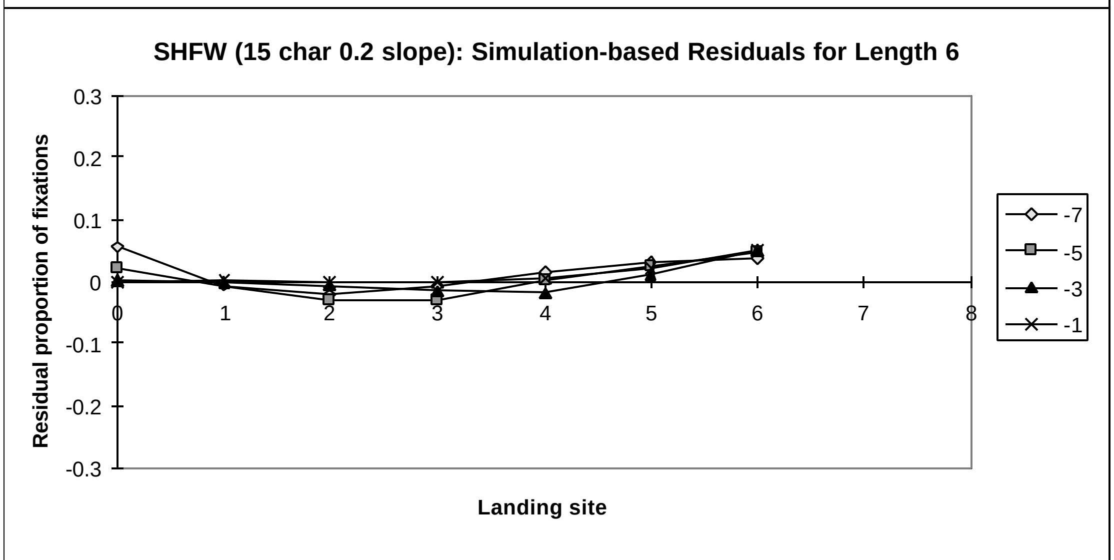

WBW: Simulation-based Residuals for Length 4
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
Launch distance
-7
-5
-3
-1

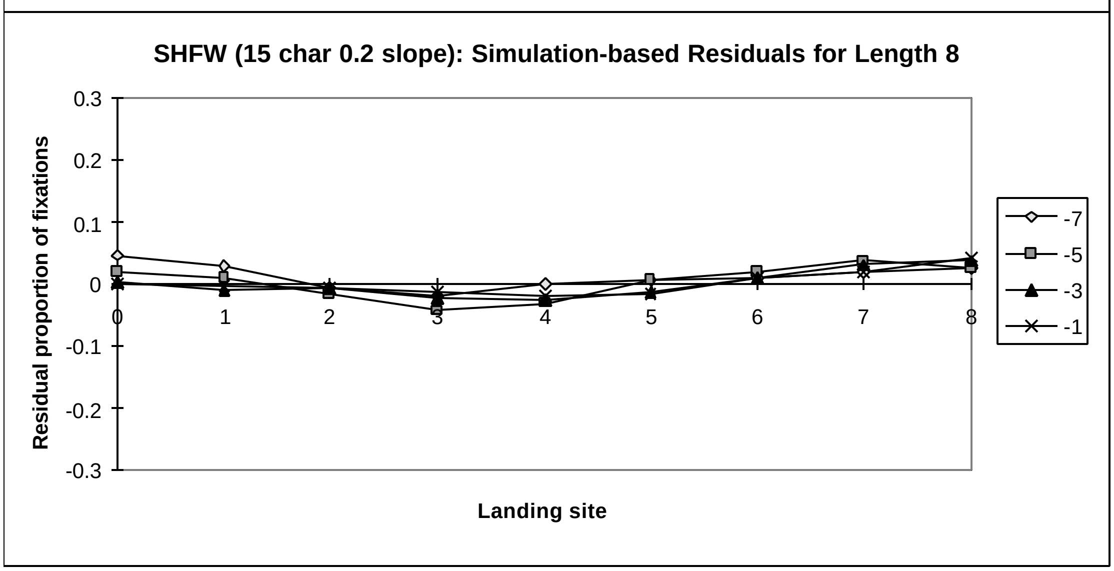

TLW (20 chars 0.004 slope): Simulation-based Residuals for Length 4
Residual proportion of fixations
Landing site
-7
-5
-3
-1

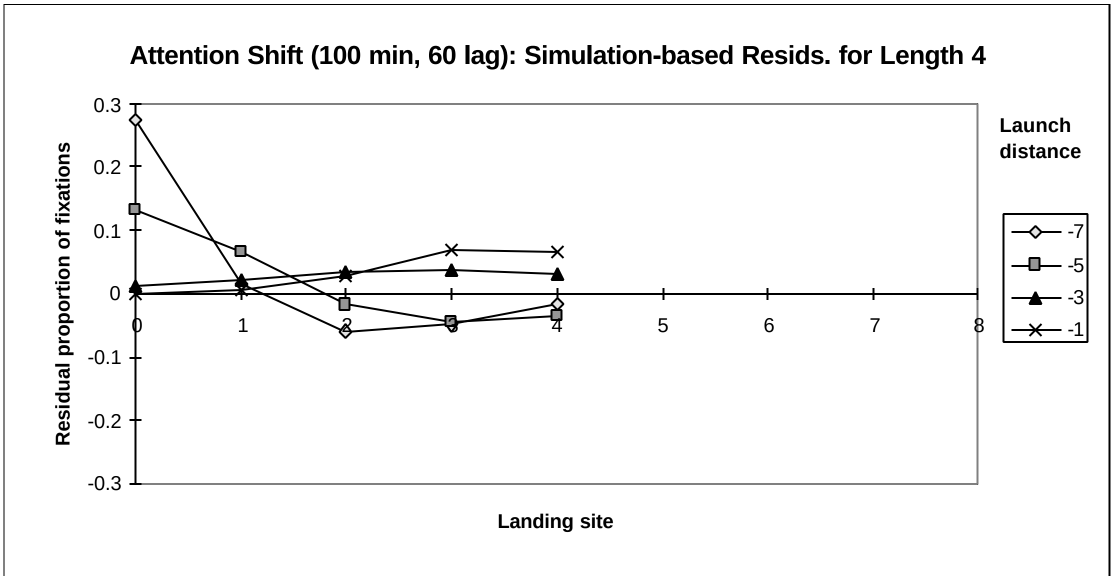

TLW (20 chars 0.004 slope): Simulation-based Residuals for Length 6
Residual proportion of fixations
Landing site
-7
-5
-3
-1

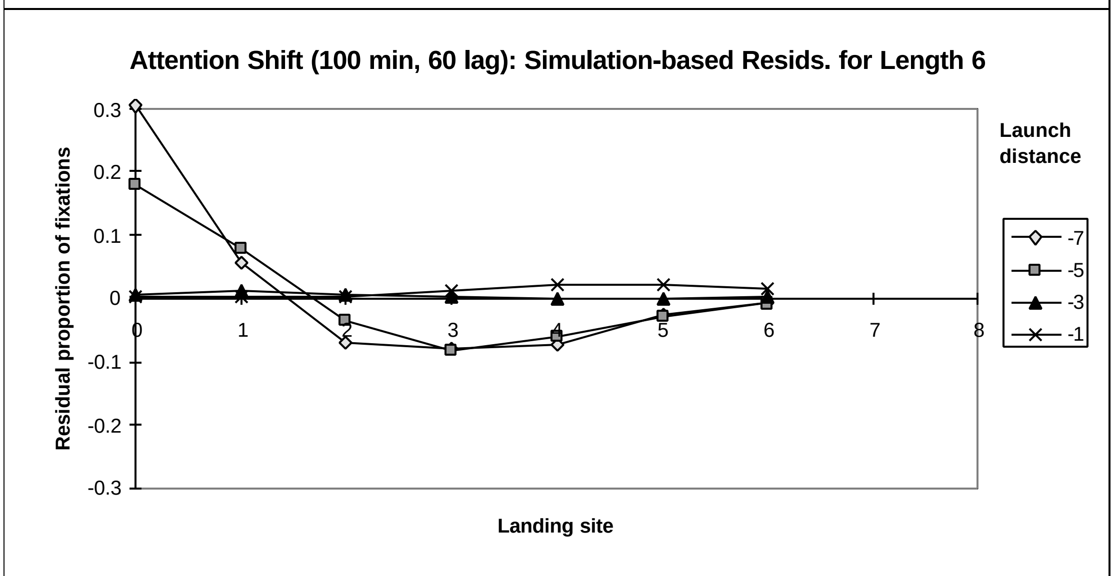

TLW (20 chars 0.004 slope): Simulation-based Residuals for Length 8
Residual proportion of fixations
Landing site
-7
-5
-3
-1

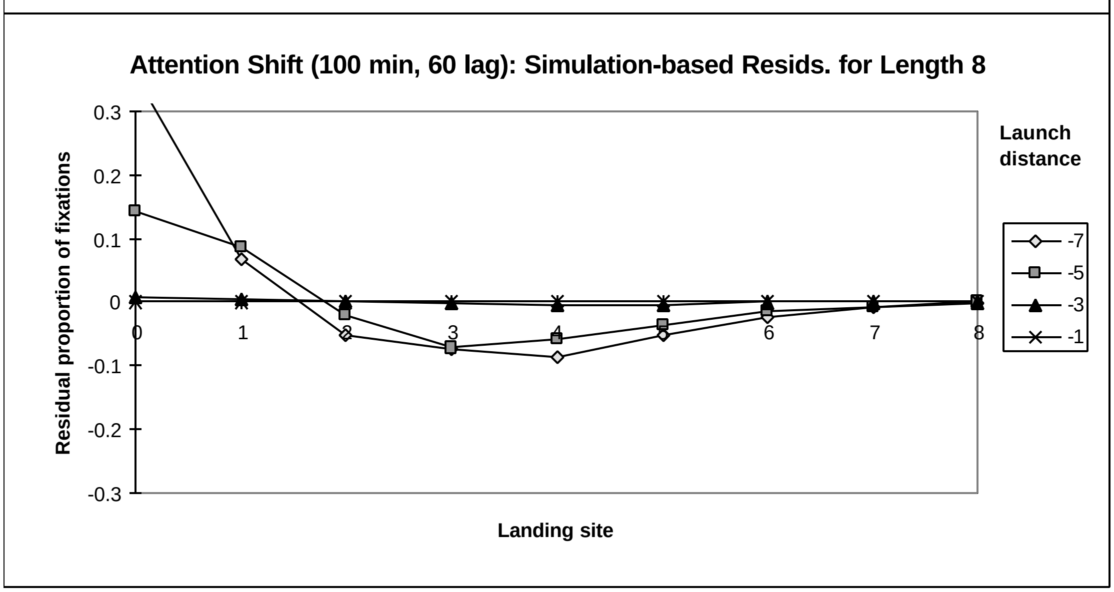

SHFW (15 char 0.2 slope): Simulation-based Residuals for Length 4
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
-7
-5
-3
-1


SHFW (15 char 0.2 slope): Simulation-based Residuals for Length 6
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
-7
-5
-3
-1


SHFW (15 char 0.2 slope): Simulation-based Residuals for Length 8
Residual proportion of fixations
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
0 1 2 3 4 5 6 7 8
Landing site
-7
-5
-3
-1


Attention Shift (100 min, 60 lag): Simulation-based Resids. for Length 4
Residual proportion of fixations
Landing site
Launch distance
-7
-5
-3
-1


Attention Shift (100 min, 60 lag): Simulation-based Resids. for Length 6
Residual proportion of fixations
Landing site
Launch distance
-7
-5
-3
-1


Attention Shift (100 min, 60 lag): Simulation-based Resids. for Length 8
Residual proportion of fixations
Landing site
Launch distance
-7
-5
-3
-1