Evidence for cognitively controlled saccade targeting in reading

Klinton Bicknell¹ (kbicknell@ucsd.edu), Emily Higgins¹ (ehiggins@ucsd.edu), Roger Levy² (rlevy@ucsd.edu), & Keith Rayner¹ (krayner@ucsd.edu)
¹Department of Psychology, UC San Diego, La Jolla, CA, USA
²Department of Linguistics, UC San Diego, La Jolla, CA, USA

Abstract
It is generally assumed that the character position targeted within a particular word is not under direct cognitive control, but is rather determined by oculomotor processes sensitive only to word length and distance. An alternative view is that readers target more distant characters in words when they have parafoveally processed these words more. These possibilities are difficult to distinguish because the actual landing site within a word has large effects on subsequent word processing measures. In two experiments, we decoupled the targeted location from the actual landing site by shifting the text 3 characters during the saccade into a target word. Results show that subsequent word processing time given a particular landing site was lower/higher when the eyes would have landed further forward/backward in the word. This effect remains significant in some cases when controlling for saccade launch site. These data provide evidence against the oculomotor theory and support a cognitive account of saccade targeting.

Keywords: eye movements; reading; display change

Introduction
Reading is a complex process that requires the combination of language processing with visual information to make decisions about when and where to move the eyes. These decisions are made very rapidly: saccades in reading take around 150 ms to program (Rayner, Slowiaczek, Clifton, & Bertera, 1983), yet fixation durations in reading are around 200–250 ms, leaving only 50–100 ms to decide when and where to send the eyes next. Given these temporal requirements, one central question of reading research is the extent to which these decisions are made by the cognitive system – and thus are sensitive to ongoing linguistic processing – or made by faster, low level oculomotor heuristics. Much of this debate has focused on how readers decide when to make a saccade, e.g., investigating the sensitivity of the distribution of fixation durations to the linguistic properties of a fixated word such as its frequency or predictability (Staub, White, Drieghe, Holloway, & Rayner, 2010; Staub, 2011; Feng, 2009b). It is generally assumed, however, by researchers on both sides of this debate that it is via oculomotor heuristics that readers decide where within a word to target their eyes.¹ In this paper, we provide evidence against this view, suggesting that character-level saccade targeting decisions are under cognitive control, and thus supporting a view in which even the fine details of eye movements are sensitive to ongoing linguistic processing.

Character-level saccade targeting
It has been known since Rayner (1979) that the eyes’ modal landing position in (medium and long) words is slightly left of the center, and Rayner initially suggested that readers may intentionally send their eyes to this position because it is the most efficient location from which to process the word (cf. O’Regan, 1981). However, Rayner, Well, Pollatsek, and Bertera (1982) found evidence from a display change paradigm in which they controlled the amount of preview – visual information available about the next word – by replacing some letters with Xs that readers send their eyes further into the following word when they had received more preview. They suggested a cognitive account of character-level saccade targeting, in which readers target a position further into a word when they have already processed more of the word. For example, if readers are able to identify initial letters in a word, they no longer need visual information about those letters, and it is an efficient reading strategy to target the eyes at the latter, still-unidentified part of the word (Rayner, McConkie, & Zola, 1980). However, it is possible that Rayner et al.’s (1982) results do not reflect normal reading behavior, and may instead reflect an experiment-specific strategy, e.g., making shorter saccades when the next word contains more Xs.

McConkie, Kerr, Reddix, and Zola (1988) investigated this issue with an analysis of the effect of preview on saccade targeting in a corpus of naturalistic reading. To assess the effect of preview, McConkie et al. investigated the effect on landing position of launch site, the distance of the previous fixation from the beginning of the word. Because the quality of visual information rapidly decreases away from the fovea, nearer launch sites would be expected to yield more preview, and – under Rayner and colleagues’ cognitive account – landing positions further into the word. McConkie and colleagues’ results confirmed this prediction, showing that the modal landing position was more rightward for nearer launch sites. However, McConkie et al. presented analyses suggesting that this result was not best explained by the cognitive account. Specifically, they presented evidence that the relationship between launch site and modal landing position was linear, and argued that an account that explains the shift in modal landing position in terms of parafoveal preview should predict a non-linear relationship. Because readers only obtain significant information about letter identities from 7 or so characters away (Underwood & McConkie, 1985), they argued that a preview account would predict that the effect of launch site on landing site should asymptote by launch sites of 7 characters. McConkie and colleagues presented evidence that the shift in modal landing position was well modeled as a linear function of launch sites from 1 to 7 characters, with no evidence of becoming smaller near 7 characters. Neverthe-

¹This is specifically the case for decisions about where within a word to target the eyes. The control of decisions about which word to target is known to reflect cognitive processing.
less, because they did not analyze launch sites past 7 characters, this is not strong evidence against the cognitive account. Based on their evidence, however, McConkie et al. (1988) proposed an oculomotor account of character-level saccade targeting, in which the functional target of the eyes is always the center of the word, but in which systematic error biases saccade lengths toward 7 characters (and happens to do so linearly). They further suggested that this systematic error is related to range error found in other saccadic (Kapoula, 1985) and manual (Poulton, 1981) tasks, which biases saccades toward the mean saccade length. This oculomotor account of saccade targeting has since become the dominant theory, and is encoded in all major models of eye movement control in reading (e.g., Reichle, Pollatsek, Fisher, & Rayner, 1998; Engbert, Nuthmann, Richter, & Kliegl, 2005).

**Goals**

In the present work, we tease apart the cognitive and oculomotor accounts by testing their predictions for word processing that occurs after landing on a new word. This is difficult to disentangle in natural reading, because there are large effects of the actual landing site on eye movement measures that indicate word processing time, such as gaze duration and refixation rates (e.g., O’Regan, 1981; McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; Rayner, Sereno, & Raney, 1996). Here, we experimentally decouple intended landing site from actual landing site using a sentence shift paradigm, which allows us to investigate the relationship between target word processing and where a reader would have landed in the word. The cognitive account, in which readers direct their eyes to later character positions in upcoming words when they have performed more parafoveal processing of the beginnings of these words already, predicts that – when controlling for actual landing site – readers will require less time to finish processing a word when they had targeted a later character. The oculomotor account, which holds that where a reader lands in a word is purely a function of launch site, does not obviously make this prediction. However, because parafoveal preview should be larger when the eyes are closer to the word (i.e., for closer launch sites), the oculomotor account may also make this prediction, because it predicts that landing sites are correlated with launch site. Crucially, though, in the oculomotor account, all effects of original landing position (i.e., the position at which the eyes would have landed had we not shifted the sentence) must be mediated by launch site. The cognitive account by contrast, under the assumption that the amount of parafoveal processing performed is variable even for a constant launch site, predicts that word processing times will be smaller when the eyes would have landed further into the word, even when controlling for effects of launch site. (Note, however, that a large amount of the amount of parafoveal preview obtained is likely to be correlated with launch site even on this account). We test these predictions in the following two experiments.

**Experiment 1**

We use a sentence-shifting paradigm (McConkie, Zola, & Wolverton, 1980; O’Regan, 1981; Inhoff, Weger, & Radach, 2005; Nuthmann, 2006; Feng, 2009a) to tease apart effects of intended landing site from actual landing site. In our first experiment, we shift the sentence to the right during the saccade into a target word, as described below. This paradigm allows us to align actual landing sites and compare two cases: (1) when the actual site was the intended landing site (when no shift occurred), and (2) when the intended landing site was, instead, further into the target word (when the sentence shifted to the right). The cognitive account predicts that the latter case – when more distant locations were targeted – is more likely to reflect instances in which readers had parafoveally processed the word to a greater extent. Thus, the account predicts that word processing times on the target word should be reduced compared with the control, no shift condition. The oculomotor account may make this same prediction, but only as mediated by effects of launch site, since under this view landing position is strictly a function of launch site, and closer launch sites may yield more parafoveal preview. We thus seek to answer two questions: (1) whether any information about upcoming word processing can be recovered from original landing site, as measured by whether there is an effect of shift on subsequent eye movement measures aligned by actual landing site, and (2) whether this effect is completely mediated by launch site.

**Method**

**Subjects** All subjects were students at the University of California, San Diego who received course credit for participation. All were naive to the purpose of the experiment and reported that they were native speakers of English with normal or corrected-to-normal vision. Data from 40 subjects were included in our analyses. Five additional subjects participated in the experiment but were excluded from analysis for reasons discussed below.

**Apparatus** Eye movements were monitored with an SR Eyelink 2000 eye tracker (SR Research Ltd., Kanata, Ontario, Canada) sampling at 1000 Hz. The system was configured in ‘tower mode’ and equipped with a chin rest. While subjects read binocularly, only one eye (the right eye by default) was tracked. Sentences were displayed on an HP p1230 20 in. CRT monitor with refresh rate set to 150 Hz and resolution set to 1024 × 768 pixels. Viewing distance was approximately 60 cm. Approximately 2.4 characters were encompassed by 1° of visual angle. We used custom software (EyeTrack, developed at the University of Massachusetts, Amherst) to present and update the display.

**Materials** One hundred and sixty experimental sentences were included in this study. Eye movement measures were obtained from a single, pre-selected target word (always a 7-letter verb) within each sentence, which was immediately preceded by a 3- or 4-letter noun. Each sentence appeared alone
Figure 1: Example sentence. The first line depicts an experimental sentence and a boundary (invisible to subjects but shown here after the pre-target word ‘fig’) that, when crossed, will trigger a display change. The target word is ‘couples’. The second line shows the control condition in which the sentence remains in place after the boundary has been crossed. The final two lines depict rightward (Experiments 1 and 2) and leftward shifts (Experiment 2 only) respectively.

on a single line of the screen in Courier New 14 pt. font.

Procedure After giving informed consent and receiving experimental instructions, subjects placed their heads in the chin rest and performed a 3-point horizontal calibration. Subjects then read 6 practice sentences, all without display changes, before beginning the experiment. Subjects read each experimental sentence silently for comprehension. For each subject, the order of sentence presentation was randomly and independently selected. After one third of trials, a simple comprehension question was presented to encourage attentive reading. Breaks were offered approximately halfway through the experiment and were available at any other time upon request. We used the gaze-contingent boundary technique (Rayner, 1975) to update the display when subjects’ eyes crossed an invisible boundary placed after the last letter of the pre-target word. When this boundary was crossed, on half of the trials the display was re-drawn so that the entire sentence was shifted 3 characters to the right (the Right Shift condition, see Figure 1). In the remaining half of trials, the sentence was simply re-drawn in its original location (No Shift). The assignment of items to shift conditions was counterbalanced.

Analysis Data were processed using a suite of custom software developed at the University of Massachusetts, Amherst and the University of California, San Diego. Fixations shorter than 80 ms that occurred within a single character width (11 pixels) of an adjacent fixation were combined, and those that did not were removed. Trials containing a fixation longer than 1000 ms or a blink on or immediately preceding or following the target word region were also excluded. Trials were also excluded if the display change completed more than 9 ms after the beginning of the following fixation. Subjects were excluded from analysis for excessive data-loss, defined as 25% or more of trials being excluded for blinks or 50% or more of trials being excluded for late display changes. Trials were also excluded if the eyes (1) would not have landed on the target had no shift taken place, or (2) would have landed on the target under natural circumstances but were ‘thrown off’ by the shift. This requirement meant that all data from the shift condition was limited to actual landing positions 1–4. In our statistical analysis of the effect of shift, we thus compared the two shift conditions at only these four positions. Finally, in order to increase the probability that all fixations were intended for the target word, and not mislocated fixations intended for the previous word, we also excluded cases in which the previous word was skipped. Note that it is possible that two classes of unintentional fixations of the target word remain in the data: (a) fixations intended to be refixations of the previous word and (b) fixations that were intended to skip over the target word and fixate a subsequent word. However, each of these possibilities is unlikely to represent a substantial portion of the dataset, as the refixation probability for words of length 3–4 is very low (about 13% for words of length 4 and even lower for words of length 3; McConkie et al., 1989) and the probability of skipping over a 7-letter word is only about 10% (Drieghe, Brysbaert, Desmet, & De Baecke, 2004).

We analyzed two measures of word processing: (1) gaze duration, defined to be the summed duration of all fixations made on a region prior to leaving it and (2) refixation probability, defined as the probability of making more than one fixation on a region prior to leaving it. We analyzed the effect of shift on gaze duration with linear mixed-effects regression (Pinheiro & Bates, 2000) and on refixation probability with logistic mixed-effects regression (Agresti, 2002). In addition to a fixed effect of shift, all models included random intercepts and random slopes for shift for both subjects and items. As a control variable, the actual (post-shift) landing site was included as an unordered categorical fixed effect, and random slopes for landing site were included for subjects and items. In cases of nonconvergence, we iteratively removed random slopes of landing site until the model converged. We do not report control variable effects. Outlier gaze durations were excluded by removing all gaze durations more than 2 standard deviations from a subject’s mean, without respect to experimental condition.

We report two analyses to answer the two questions described above. The first analysis seeks to determine whether any information about upcoming word processing can be recovered from original landing position by testing for an ef-
Refixation rate
Gaze duration

0.0 250 0.2 300 0.4 350 0.6 400

0.2 0.4 0.6 0.8

cond
- None
- Right

Landing position

2 4 6

2 4 6

cond
- None
- Right

Landing position

Figure 2: Effect of shift on gaze duration and refixation rates by actual (post-shift) landing site for Experiment 1. Error bars show the standard error of the mean, computed after aggregation by subjects. Note that effects mediated by launch site have not been parceled out of this figure.

The effects of shift on word processing controlling for actual landing site. The second tests whether this effect is completely mediated by launch site (as predicted by the oculomotor account) by including launch site as an unordered, categorical control predictor. To assess significance for the linear gaze duration models, we report the $t$ statistic. For datasets of this size, this statistic will be approximately normally distributed (Baayen, Davidson, & Bates, 2008), meaning that $|t| > 1.96$ indicates a significant effect ($p < .05$) and $1.64 < |t| < 1.96$ indicates a marginal one (.10 $p < .05$). For logistic refixation models, we report the $z$ statistic, which has the same interpretation, and also give effect sizes in logits, which is the difference in the log-odds of making a refixation between conditions (Agresti, 2002).

Results

The effects of shift are plotted in Figure 2, aligned by actual, post-shift landing position. There is an effect of shift on gaze duration: gaze durations are estimated to be significantly faster ($-19$ ms, $t = -3.1$) when the eyes would originally have landed further into the word (i.e., after a rightward shift). Refixations are estimated to be 0.2 logits less likely after a rightward shift, but this is not significant ($z = -1.1$). In analyses including launch site as a control predictor, the effect of shift on gaze durations was reduced to an insignificant 10 ms ($t = -1.4$), and the effect on refixations remained similar ($-0.2$ logits, $z = -1.1$).

Discussion

The results of Experiment 1 confirmed that original landing position does provide some information about upcoming word processing, as cases in which the eyes would have landed further forward in the word result in 19 ms shorter gaze durations. While this result is predicted by the cognitive account, the oculomotor account can also predict it, but only to the extent that it is completely mediated by launch site. In analyses controlling for launch site, the results from this experiment were unclear, however, and there was only an insignificant trend for gaze durations to be 10 ms shorter when the eyes would have landed further forward in the word. Thus, the results of this experiment are consistent with both models.

Experiment 2

One limitation of the design of Experiment 1 is that it is possible that the effects we saw on gaze duration and refixation rate were merely low-level responses to shifting the sentence rather than true effects of prior processing of the target word. To allay this concern, in Experiment 2, we tested both right and left shifts of the sentence. While a simple, low-level response to the detection of a shift may be expected to affect eye movement measures similarly for leftward and rightward shifts, the cognitive account of saccade targeting makes opposite predictions for these two conditions. By the same logic as described for Experiment 1, this account predicts that gaze duration and refixation rates should be reduced in the rightward shift condition relative to the no shift condition when aligning on actual landing site. This is because the saccades in the rightward shift condition were directed further into the word, which on this account is caused by readers having performed more parafoveal processing. Analogously, this account predicts that these measures should be increased in the leftward shift condition relative to the no shift condition when aligning on actual landing site, since the leftward shift saccades were directed further back in the word than those in the no shift condition. The oculomotor account once again makes the same predictions as the cognitive account, but again requires that these effects be solely mediated by launch site. Experiment 2 thus allows us to test two predictions. First, if the simple, low-level shift effect is correct, we should find similar patterns of data for leftward and rightward shifts. Second, if we instead find opposite patterns of data for leftward and rightward shifts (as outlined above), examining whether these effects are solely driven by launch site will allow us to distinguish between the oculomotor and cognitive accounts of saccade targeting.

Methods

Experiment 2 was identical to Experiment 1 with two exceptions. First, while 40 subjects were again included in our analysis, 7 were excluded (for reasons explained above). Second, while sentences remained static, once again, in half of trials,
they shifted 3 characters to the right in one quarter of trials and shifted 3 characters to the left in the remaining quarter. An example is given in Figure 1.

**Analysis** Analysis was similar to that in Experiment 1 except that we now separately analyze the effect of left and right shift, comparing each to the no shift condition. Because we again exclude cases in which the eyes would have skipped the target or were “thrown off” the target by the shift, this means that the data from the right shift condition are at actual landing sites 1–4 and the data from the left shift condition are at actual landing sites 4–7. We thus analyzed data from only these landing site ranges in each analysis.

**Results**

The effects of shift are plotted in Figure 3, aligned by actual, post-shift landing position. For right shifts, the effect on gaze durations and refixation rates is again estimated to be in the predicted direction (−7 ms, −0.3 logits), but neither effect is significant (t = −0.7; z = −1.5). For left shifts, the effects are in the opposite direction, and are larger (20 ms, 0.6 logits) and significant (t = 3.7; z = 6.1).

As before, we also performed analyses in which launch site is a control predictor, to determine whether these effects are exclusively mediated by launch site. For right shifts, this analysis revealed insignificant effects on gaze duration (3 ms, t = 0.3) and refixation rates (−0.3 logits, z = −1.4). The effect of left shifts controlling for launch site was estimated to be slightly smaller than when not controlling for launch site (18 ms, 0.5 logits), but still robust (t = 2.9; z = 4.5).

To gain more power to assess the possible effects of rightward shifts, we performed a further, post-hoc analysis on the pooled data from for landing positions 1–4 from Experiments 1 and 2. This analysis revealed a significant effect of rightward shifts for gaze duration (−16 ms, t = 3.2) and a marginal trend for refixation rate (−0.2 logits, z = −1.8).

An analysis controlling for effects of launch site revealed an insignificant 6 ms trend on gaze durations (t = −1.0) and a marginal effect on refixation rates (−0.2 logits, z = −1.8).

**Discussion**

This experiment revealed, first, that leftward and rightward shifts produced opposite patterns of results, contrary to the predictions of the simple, low-level shift detection account: gaze duration and refixation rate were lower and higher in the rightward and leftward shift conditions respectively as compared with the static control condition, although this was only significant for the leftward shift condition. Because both the cognitive and the oculomotor accounts predicted this pattern of data, we also analyzed the data when controlling for launch site, a factor that should, according to the oculomotor view, entirely account for these results. These analyses revealed results more consistent with the cognitive account than the oculomotor account. In the leftward shift condition, gaze duration and refixation rate were significantly elevated even when controlling for launch site. For rightward shifts, pooling data across the two experiments also provided suggestive evidence in favor of the cognitive account of saccade targeting, suggesting that the effect was not entirely driven by launch site.

**Conclusion**

In summary, we described two alternative accounts of how readers decide where, precisely, to aim their eyes when planning a saccade to an upcoming word. According to the cognitive account, readers send their eyes further into a word after having parafoveally processed it more. According to the oculomotor account, readers always target the center of a word, but are subject to systematic error, which is a function of launch site. We presented evidence in favor of the cognitive account from two sentence-shift experiments. As predicted by the cognitive account, the word processing measures of gaze duration and refixation rate suggested that readers perform less subsequent processing of a word when they would have landed further into it, and more subsequent processing of a word when they would have landed further back, controlling for actual landing site. This was a significant effect for rightward shifts in Experiment 1 and for leftward shifts in Experiment 2. Crucially, we found evidence that this effect was not fully mediated by launch site, as required by the oculomotor account. When controlling for launch site, in Experiment 2, the effect of leftward shifts was fully reliable, and when
pooling data across Experiments 1 and 2, the effect of rightward shifts was marginal for refixation rate. This evidence thus suggests that readers decide where to target their eyes within a word based on how much processing of the word they have accomplished, and not just based on the current position of their eyes. Such an account requires that the details of saccade targeting are sensitive to ongoing cognitive, linguistic processing.

There is, however, one way in which the oculomotor account may still be able to accommodate these findings. As was pointed out above (see Expt. 1, Analysis), some trials included in our analysis may represent unintentional fixations on the target word: failed attempts (a) to refixate the pretarget word and (b) to skip the target word. If these trials represent a substantial portion of our data, the oculomotor account could also predict our findings, since (a) failed refixations would tend to land at the beginning of the target word and represent cases in which the target word was not yet the focus of processing and (b) failed skips would tend to land at the end of the target word and represent cases in which the target word was already processed. Further analyses will be required to determine whether the likely rates of such possibilities would be sufficient to render this account of our data plausible.

The data are certainly consistent, however, with the view that character-level saccade targeting is under cognitive control. Specifically, these results are predicted by an account in which readers send their eyes further into a word when they have obtained more parafoveal preview of it. Since this effect is not mediated by launch site, this means that where a reader’s eyes land in a word provides information about how much they processed the word on that particular trial, which is not only a function of the location of their eyes on the previous fixation. If this account is correct, it would support the notion that fine-grained eye movements decisions in reading are tightly linked to the details of ongoing linguistic processing, suggesting that readers do not merely rely on heuristic strategies to guide their eyes. More generally, our results support a view in which humans optimize the fine details of their behavior to maximize their efficiency in linguistic tasks such as reading (Bicknell & Levy, 2010; Lewis, Shvartsman, & Singh, in press) and in cognition more broadly.

Acknowledgments

This research was supported by NIH grant T32-DC000041 from the Center for Research in Language at UC San Diego to K.B., NSF grant 0953870 to R.L., and NIH grant R01-HD065829 to R.L. and K.R. We thank Araceli Cervantes, Georgina Chen, Tiffany Lai, Hannah Johansen, and Vinny Uy for assistance with data collection and stimuli development.

References