

Why readers regress to previous words: A statistical analysis

Klinton Bicknell & Roger Levy
{kbicknell, rlevy}@ucsd.edu

UC San Diego, Department of Linguistics
9500 Gilman Drive #108, La Jolla, CA 92093-0108 USA

Abstract

While the major models of eye movement control in reading propose very different mechanisms for the generation of saccades to previous words, there has been relatively little empirical data to distinguish these hypotheses. Here we provide a systematic statistical analysis of the factors that elicit these saccades in a corpus of eye movements. We show that the results are contrary to the predictions of a number of accounts, and provide new evidence to discriminate among the models.

Keywords: eye movements; reading; regressive saccades; logistic regression

Introduction

The control of the eyes during reading is one of the most complex everyday tasks humans face, as efficient performance requires the rapid integration of complex information from visual, motor, and linguistic sources. In recent decades, our knowledge of the determinants of how readers select saccade targets and decide how long to fixate particular locations has grown immensely (for reviews see Rayner, 1998, 2009). The focus of this work has been on understanding the most common ways in which the eyes move through the text: progressive saccades to a word further in the text and refixations of the current word. As such, our understanding of instances in which the eyes move back to a previous word (between-word regressions) is still one of the most poorly developed facets of theories of eye movement control in reading. While not the most common type of saccade, between-word regressions are still a regular property of the eye movement record, consistently occurring on between 1 in 10 and 1 in 20 saccades (and as high as 30% of all saccades for some readers, Radach & McConkie, 1998). Thus, it is a striking state of affairs that the major models of eye movement control in reading (e.g., Pollatsek, Reichle, & Rayner, 2006; Engbert, Nuthmann, Richter, & Kliegl, 2005) propose very different reasons for making between-word regressions, each of which is intimately tied to the model's view of the nature of eye movements in reading. One of the best ways to tease apart the various models of eye movement control, then, is to gain a better understanding of between-word regressions. To date, however, there has been relatively little empirical data analysis to distinguish these various theories of why a reader would move their eyes to a previous word. Thus, the present work seeks to remedy this situation by providing a systematic analysis of the factors that elicit between-word regressions, and in so doing to provide a new source of evidence distinguishing among models of eye movement control in reading.

While there are many proposals as to why a reader would regress to a previous word, it is a common belief among researchers who have diverse opinions on the primary cause

of between-word regressions (e.g., Reichle, Warren, & McConnell, 2009; Engbert et al., 2005) that some proportion of them are made in response to overshooting a target word due to motor error, and making a between-word regression to return to the originally targeted word. Supporting this view is evidence that between-word regressions increase following word skipping (Vitu & McConkie, 2000). Thus, to better distinguish between competing theories of between-word regressions, we focus our discussion here only on cases in which the target of the between-word regression was not skipped (about half of all between-word regressions in the dataset we examine below). The question for the theories then becomes one of why a reader would regress to a previously fixated word.

Theories of between-word regressions

We consider here five classes of theories of the causes of between-word regressions to unskipped words. For each theory, we highlight the predictions it makes for which variables will influence the rate of between-word regressions. For concreteness, we discuss these predictions for the case of a regression from the n th word in a given sentence (word n) to the previous word $n - 1$. We can group the factors to which regressions are predicted to be sensitive into three categories: (a) properties of word n (i.e., its length, frequency, or predictability), (b) those same properties of word $n - 1$, and (c) motor properties (i.e., length of the previous saccade or the position of the eyes on word n or previously on word $n - 1$).

Corrective One possibility for most theories is that some between-word regressions (even when word $n - 1$ was not skipped) could still be corrective. This could happen if the saccade that landed on word n was intended to be a refixation of word $n - 1$. This is quite plausible since the most common pattern of refixations is that first a word's beginning is fixated and then its end (Rayner, Sereno, & Raney, 1996).

This account predicts that properties of word $n - 1$ should be relevant, since there should be more regressions when refixations are more likely (lower frequency or predictability, longer length of word $n - 1$). In addition, motor properties should be relevant: regressions should be more likely when the eyes land closer to the beginning of word $n - 1$ (since refixations are more likely) and closer to the beginning of word n (since that is where failed refixations would land). Properties of word n are not predicted to matter, except insofar as they correlate with properties of the preceding input.

Oculomotor strategy Another possibility is that regressive saccades could be initiated as part of an oculomotor strategy (O'Regan & Lévy-Schoen, 1987; Yang & McConkie, 2001).

On this account, readers have learned that it is generally beneficial to launch a regression in response to a particular configuration of visuomotor variables. For example, a strategy could be to regress after a particularly long saccade, or after skipping a word. While the precise details of these strategies have not been well worked out for cases in which a word was not skipped, the crucial prediction is that regressions produced by an oculomotor strategy cannot be influenced by linguistic properties like frequency and predictability.¹

Incomplete lexical processing A range of models argue that regressions can sometimes be produced by incomplete lexical processing (in reading, generally taken to be synonymous with word recognition). There are two possible accounts of how this could happen. In a serial word processing model, in which readers attend a single word at a time, it could occur when a reader accidentally moves their eyes away from a word too early, and then moves them back to continue processing it more efficiently (Vitu, McConkie, & Zola, 1998).² However, it is in attention gradient models, in which readers attend multiple words simultaneously, that this account has been better developed (Engbert, Longtin, & Kliegl, 2002; Engbert et al., 2005; Reilly & Radach, 2006). There, if the processing of the previous word was too short (relative to the word's length, frequency, and predictability), then its activation can become higher than that of the current word or future words, which can trigger a between-word regression.

Both of these accounts predict that factors that increase the difficulty of word $n - 1$ (longer length, lower frequency and predictability) should increase the number of regressions (for a given fixation duration on word $n - 1$). In addition, the attention gradient models predict that regressions should be more likely when word n is easier (shorter length, higher frequency and predictability), since its activation level will thus be less of a competitor. The serial model predicts that linguistic properties of word n will not have an effect.

Integration failure It has been well documented that strong garden path sentences (i.e., sentences with temporarily ambiguous words or syntactic structures, which are initially strongly biased towards the incorrect interpretation) often elicit between-word regressions at the disambiguating region (Frazier & Rayner, 1982). A common explanation for this finding is that integration of the disambiguating word into prior context fails, and readers must regress to previous words for reprocessing. Although most evidence for regressions in this situation comes from experimental manipulations with strong, artificial garden path sentences, it may be that weaker garden paths (which are not consciously perceptible) never-

theless sometimes cause integration failure and elicit regressions through this mechanism. It should be noted, however, that a large number of such garden paths would be required to produce between-word regressions on 5-10% of saccades.

Since the difficult disambiguation region in a garden path is by definition unpredictable, this account predicts more regressions when word n is less predictable (the opposite of the prediction made by the attention gradient incomplete lexical access account). In addition, it has often been found that garden path regressions occur on following words, thus this account also predicts more regressions on word n when word $n - 1$ is less predictable. It would not obviously predict the length of either word or the eyes' landing position on either word to be relevant.

Confidence falling The final theory of between-word regressions we discuss here was suggested recently by Bicknell and Levy (2010). In this account, readers maintain uncertainty about the identities of previous words and update that uncertainty as input from new words further downstream becomes available (Levy, 2008), a proposition that has some recent empirical support (Levy, Bicknell, Slattery, & Rayner, 2009). The model proposes that when a new word fits relatively poorly with what the reader believed the prior context to be, and relatively better with an alternative visually similar possibility, the reader's confidence in the identity of the prior context will be reduced. In this situation, it becomes useful to get more visual information about the prior context, and thus make a between-word regression.

The predictions of this account combine predictions of the incomplete lexical processing and integration failure accounts. Confidence is more likely to fall about words whose confidence was lower to begin with, which predicts that factors that slow processing of word $n - 1$ (longer length, lower predictability) will increase regressions, as in the incomplete lexical processing account. Like the integration failure account, however, this account predicts that an unpredictable word n (since it fits poorly with the prior context) will be more likely to cause confidence to fall, and the word's length should be irrelevant. The prediction for predictability is actually more subtle, however, since not every word that is unpredictable given a particular context will be more predictable given some other context. We will return to this point later.

Previous empirical evidence

There is relatively little empirical evidence regarding which factors between-word regressions are sensitive to in the case that the regression target word was not skipped. Most of the existing work (which has not generally controlled for skipping) has looked for effects of linguistic properties of the targets of regressions. For example, readers were found to regress more to words of low predictability with no significant additional effect of frequency (Rayner, Ashby, Pollatsek, & Reichle, 2004; Kliegl, Grabner, Rolfs, & Engbert, 2004). However, these studies failed to control for skipping, and thus the results could be confounded, if, for example, regressions

¹Of course, in models such as Yang and McConkie (2001), higher level language processing can sometimes intervene, so that not all regressions in this model would be produced by an oculomotor strategy. For these cases however, the regression must be produced by one of the other accounts described.

²Note however that in the most well developed serial model of eye movements in reading, *E-Z Reader* (Reichle et al., 2009), this could not happen because a saccade is only initiated to leave a word after all visual processing is completed.

are triggered more towards words that were unintentionally skipped. The one study that has investigated the effects of properties of word $n - 1$ on between-word regressions specifically in the case in which word $n - 1$ wasn't skipped found that there were more between-word regressions when word $n - 1$ was lower frequency and longer (Vitu & McConkie, 2000). However, Vitu and McConkie did not have predictability in their model, and thus could not distinguish between effects of predictability and frequency. More crucially, however, their analysis could not determine whether there were independent effects of frequency and length (since longer words are also generally less frequent). Thus, while there is good evidence that some properties of the regression target are implicated for the case in which the regression target was not skipped, it is not clear whether it is the length, frequency, or predictability of the word that attracts regressions.

Evidence about whether properties of word n make regressions more likely is still more scarce. The one reported result is that Kliegl et al. (2004) found that words that were of lower frequency and lower predictability were more likely to have a regressive saccade begin on them. However, to the best of our understanding Kliegl et al. did not control for whether a word was fixated at all, and thus these results are confounded with word skipping, since a word that was skipped could by definition not have a regression begin on it.

The evidence to date for effects of motor properties on between-word regressions also comes from Vitu and McConkie (2000). They found that, contrary to the overall tendency for regressions to be more likely following longer saccades (Buswell, 1920; Vitu et al., 1998), in the case that word $n - 1$ was not skipped, regression rates decrease with following longer saccades. In addition, Vitu and McConkie reported a non-significant trend for regressions to decrease as word $n - 1$ was fixated further from its beginning.

Thus, while Vitu and McConkie's (2000) results demonstrate that cases in which word $n - 1$ was not skipped pattern very differently from cases in which it was, it is presently far from clear what the determinants of between-word regressions are in this condition. There is evidence that properties of word $n - 1$ are relevant, but it is unclear which ones, and we know almost nothing about whether properties of word n are relevant. As noted above, many of the theories of regressions make predictions for factors such as the position of the eyes in word $n - 1$ or word n , yet a reliable effect of such variables has not been found. As a result, all five classes of theories mentioned above are still quite tenable explanations for making a regression to an unskipped word. Finally, as pointed out by Vitu and McConkie, it is important to realize that many of these factors are highly correlated with one another (for example, the landing position within a word and its length and frequency), and thus strong evidence that a variable is relevant for regressions can only be made using a model controlling for effects of correlated factors. Thus, the goal of the analysis reported in this paper was to simultaneously and systematically test for effects of a range of variables – including

properties of words n and $n - 1$ as well as motor variables – on the rate of between-word regressions in cases in which the regression target had not been skipped.

Method

Corpus and dataset

In this analysis we modeled the rate of between-word regressions in a large corpus of eye movements in reading, the Dundee corpus (Kennedy & Pynte, 2005). This corpus is comprised of the eye movement record of 10 individuals each reading 50,000 words of British newspaper editorials. In order to have a more controlled dataset, we focus only on predicting regressions from word n to the previous word $n - 1$ (which account for about 70% of between-word regressions in our dataset), so that no other words intervene, and (for reasons mentioned above) only in the case that word $n - 1$ was not initially skipped. Specifically, we predict whether each saccade that originated on some word n was a regression or not, in the case that (a) the previous saccade originated on word $n - 1$, (b) neither words n nor $n - 1$ were previously fixated, (c) no word beyond n was previously fixated (i.e., first pass reading), (d) neither the previous fixation (on word $n - 1$) nor the next fixation following the saccade in question were the first or last on a line nor detected as a blink, and (e) the saccade was not a regression going back further than word $n - 1$. Each saccade meeting these criteria was thus categorized as a regression if it went to word $n - 1$, or a non-regression if it was a re-fixation of word n or a progressive between-word saccade. Finally, we excluded cases in which words n or $n - 1$ were not in the British National Corpus (see below), had punctuation (including all non-alphabetic characters), or were the first or last words in a line, as well as any case in which the fixation on word n or $n - 1$ had been on the space preceding the word. This resulted in a dataset of 33569 saccades, of which 1362 or 4% were regressions.³

Analysis

We fit a generalized linear mixed-effects regression with a logit link function (Pinheiro & Bates, 2000; Agresti, 2002) to the data using the `lme4` package (Bates & Maechler, 2010). The fixed effects in the model included the factors discussed above: properties of words n and $n - 1$ (their log-transformed frequency, predictability, and length) as well as motor properties (the log-transformed length of the previous saccade and the landing positions on both words).⁴ In addition, the model included fixed effects for the length of

³This is lower than the overall rate of 5-10% mentioned previously because of the exclusion of between-word regressions going back further than word $n - 1$.

⁴The length of word $n - 1$, the landing positions on both words, and the saccade length form a linearly dependent set such that the fourth is completely determined given the values of the other three, and thus including these four variables in a single model directly would be impossible. In our case, this is not a problem because three of the four are log-transformed, removing the linear dependence. The fact that multi-collinearity still exists between them, however, means that the estimates of their effects may be conservative.

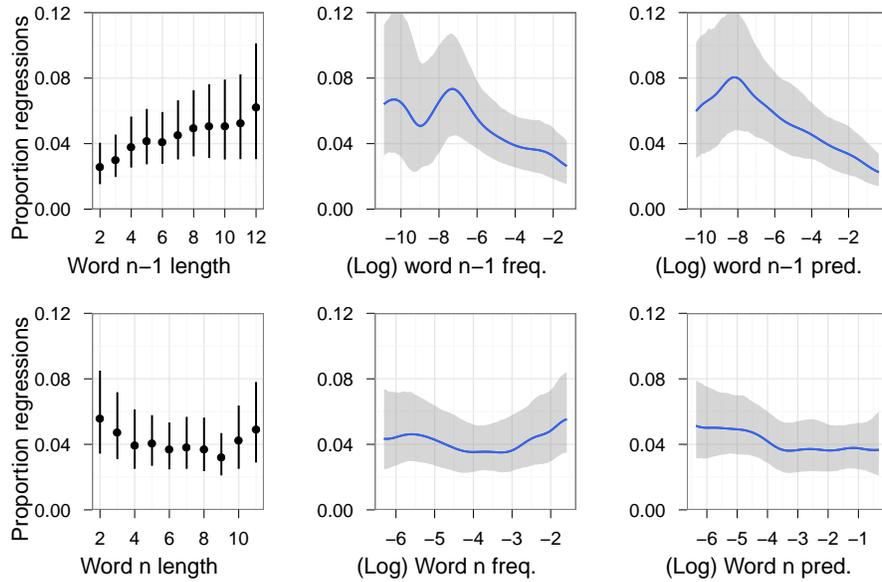


Figure 1: Marginal effects of length, frequency, and predictability of words n and $n - 1$ on proportion of regressions to word $n - 1$, shown for the middle 95% of the range of each variable. Proportion of regressions was estimated using Gaussian kernel regression with standard deviation equal to 1/15th of this range. The 95% confidence intervals are hierarchically bootstrapped from 1000 dataset replicates (Efron & Tibshirani, 1993).

the fixations on both words and random intercepts, but not random slopes, for each participant. (Models including random slopes for the nine predictors of interest failed to converge.) Frequency and predictability were estimated by unigram and trigram language models trained on the British National Corpus, smoothed with modified Kneser-Ney smoothing (Chen & Goodman, 1998). Because coefficient estimates in models without random slopes for participants can be anti-conservative for datasets in which there is real between-participant variability in effect sizes, we performed statistical tests by bootstrapping instead of using the standard normal-theory statistics (Efron & Tibshirani, 1993).⁵ Specifically, we obtained p -values and 95% confidence intervals for each coefficient from 2500 replicates of hierarchical bootstrapping, clustered by participant (Davison & Hinkley, 1997).⁶

Results

The marginal effects of the properties of words n and $n - 1$ are plotted in Figure 1. The results of the regression reveal significant effects of all three properties of word $n - 1$: regressions were more frequent when it was longer, more frequent, and less predictable ($ps < .0008$). Regressions also increased

when word n was less predictable ($p < .0008$), but were not sensitive to its length or frequency ($ps > .3$). Finally, regressions were less likely as the length of the previous saccade increased ($p < .02$) and as the landing position on word $n - 1$ was further from the beginning ($p < .01$). The relative contributions of each of these factors to the likelihood of a regression in the full model is visualized in Figure 2.

Additional analysis

Prior to discussing the implications of these results for the theories under discussion, we note that having frequency and predictability effects in opposite directions for word $n - 1$ (i.e., increased regressions for less predictable but more frequent words) is not an obvious prediction of any theory. The graphs in Figure 1, however, show that the marginal effects of frequency and predictability for both words are in the same direction (increased regressions for less predictable and less frequent words) when not both in the model together. This pattern appears in the marginal effects because of the high correlations between frequency and predictability ($r = .89$ and $r = .73$ for words $n - 1$ and n). Despite this, the regression model reports highly significant effects of the predictability and frequency of word $n - 1$ in opposite directions.

This pattern of opposite effects of frequency and predictability may be well understood in terms of the confidence falling account. As noted above, this account predicts that not every unpredictable word should cause more regressions, but only unpredictable words which are more predictable for alternate possible contexts, since they cause confidence about the true context to fall. Thus, one rough measure of

⁵Bootstrapping in this case also avoids the potential problems that normal-theory statistics are not completely valid when using the Laplace approximation to the model likelihood surface and that the Wald test becomes conservative when the data are very near to 0 or 1, as is the case for our dataset.

⁶We denote by $p < .0008$ cases in which the estimate of a coefficient in all bootstrap replicates is on the same side of zero, since if a single replicate had been on the other side of zero, the probability would be one half of 1/2500, or .0008.

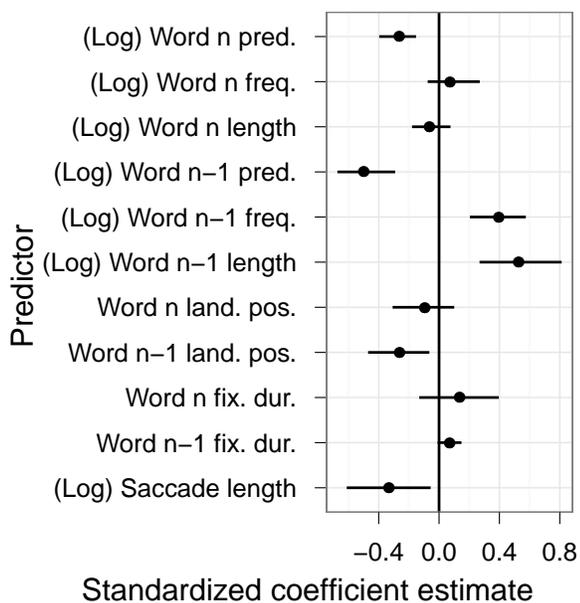


Figure 2: Estimates and 95% confidence intervals of the predictor coefficients, standardized to be on the same scale to visualize the relative contributions of each factor to the full model. (Standardization was performed by multiplying the actual coefficient by the standard deviation of the predictor.)

the amount by which a word causes a reader’s confidence to change about the preceding context (denoted Δ_c) is the change in log probability of the true context after seeing the word, relative to the context’s overall likelihood:

$$\Delta_c = \log p(c|w) - \log p(c)$$

Thus, we analyzed a new model in which we replaced predictability for both words with this measure of the change in confidence about the prior context given each word. It turns out, in fact, that this measure is equivalent to the difference between the word’s log-transformed frequency and predictability

$$\begin{aligned} \Delta_c &= \log p(c|w) - \log p(c) = \log \frac{p(w|c)p(c)}{p(w)} - \log p(c) \\ &= \log p(w|c) - \log p(w) \end{aligned}$$

and because of this relationship, the new model is simply a reparameterization of the former model (i.e., the fit and predictions will be identical), in which the coefficient for Δ_c is identical to the previous coefficient for predictability and the new coefficient for the additional effect of frequency is the sum of the previous coefficients for predictability and frequency, as can be seen in the following equation (where f and p denote frequency and predictability):

$$\beta_1 p + \beta_2 f = \beta_1 (p - f) + (\beta_1 + \beta_2) f$$

(All other coefficient values will remain the same.)

In this new model, the effect of Δ_c is such that there are significantly more regressions the further confidence falls on either word ($ps < .0008$).⁷ There are no additional effects of the frequency of either word ($ps > .1$), providing support to the notion that confidence falling may be a useful way to understand the opposite effects of predictability and frequency.

Discussion

We discuss the implications of this pattern of results for each of the five classes of theories separately.

Corrective It seems unlikely that corrective saccades are driving any effects for this dataset because of the directionality of the effect of landing position on word $n - 1$. As mentioned above, landing positions closer to the beginning of word $n - 1$ should result in more attempted refixations of the word’s end, and thus more unintentional fixations on the early part of word n . Furthermore, of course, the corrective account could not predict the effect of the predictability of word n .

Oculomotor strategy It is similarly unclear how an oculomotor strategy could account for these findings, since the linguistic properties of both words have strong effects. Also, we note that the particular strategy of making regressions after especially long saccades directly conflicts with our data.

Incomplete lexical processing The incomplete lexical processing account correctly predicts the effects in our data of predictability and length of word $n - 1$. As noted above, however, these accounts predict either that regressions should not be sensitive to properties of word n (serial models) or that they should increase when word n is easier (gradient models). Thus, the finding that regressions increase when word n is less predictable poses a problem for these accounts.

Integration failure The integration failure account is supported by the fact that more regressions are made as either word becomes less predictable. The fact that the length of word $n - 1$ also had a significant effect is problematic. One possible explanation for this is that some portion of the effect we obtained may be an artifact of the way we constructed our dataset. Specifically, by considering only between-word regressions made to word $n - 1$, we may have excluded more regressions which targeted word $n - 1$ but which landed on a word prior to $n - 1$ in the case that word $n - 1$ was shorter.⁸

Confidence falling The confidence falling account correctly predicts the effects of the predictability of words n and $n - 1$ as well as the effect of the length of word $n - 1$. As revealed by the additional analysis including Δ_c , the pattern of

⁷This is of course necessarily the case, since Δ_c has the same coefficient as predictability had previously.

⁸To investigate this possibility, we performed a similar analysis predicting all intra-line between-word regressions, and not only those to word $n - 1$. The results showed the effect of the length of word $n - 1$ to be marginal ($p = .07$), supporting the notion that our censored dataset may be responsible for a large part of the effect, but still hinting that the relationship may exist apart from censoring.

predictability and frequency effects that was found can be interpreted as some evidence that the amount by which a word makes confidence fall is a key determinant in whether a reader will make a regressive saccade.

Conclusion

We distinguished five classes of models of regressions to previously fixated words, and tested these accounts by performing a systematic statistical analysis of such regressions in a large eye movement corpus. The results of our analysis provide some of the clearest evidence to date about the variables contributing to between-word regressions. The analysis reveals strong effects of linguistic properties, and thus are hard to accommodate in purely corrective or oculomotor strategy accounts. In addition, the fact that there are more regressions to the previous word when the current word is less predictable is counter to the predictions of incomplete lexical processing models. Both integration failure and confidence falling accounts are consistent with the present data, but the facts that (1) the length of the previous word appears to matter and that (2) the opposing effects of frequency and predictability can be understood as falling confidence, suggest that the confidence falling account may find more support in the present data.

More generally, our results demonstrate that obtaining more detailed knowledge of the factors contributing to between-word regressions can distinguish between models of eye movements in reading which otherwise make very similar predictions for progressive saccades. Specifically, SWIFT (Engbert et al., 2005) makes regressions via the incomplete lexical processing account, which appears to make the wrong predictions for our results, while *E-Z Reader 10* (Reichle et al., 2009) makes use of integration failures, which is consistent with our data. Finally, we note that the empirical success of the confidence falling account, which follows from a very different class of reading model than the others considered, suggests that gaining a better understanding of regressions may have important consequences for our understanding of eye movement control in reading in general.

Acknowledgments

This work benefited from useful discussion with Keith Rayner. The research was supported by NSF grant 0953870 and NIH grant R01-HD065829 to R.L.

References

Agresti, A. (2002). *Categorical data analysis* (Second ed.). New York, NY: John Wiley & Sons.

Bates, D. M., & Maechler, M. (2010). *lme4: Linear mixed-effects models using Eigen and Eigen*. R package version 0.999375-37.

Bicknell, K., & Levy, R. (2010). A rational model of eye movement control in reading. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics ACL* (pp. 1168–1178). Uppsala, Sweden: Association for Computational Linguistics.

Buswell, G. T. (1920). *An experimental study of the eye-voice span in reading*. Supplementary Educational Monographs, No. 17. Chicago: University of Chicago.

Chen, S. F., & Goodman, J. (1998). *An empirical study of smoothing techniques for language modeling* (Tech. Rep. No. TR-10-98). Cambridge, MA: Computer Science Group, Harvard University.

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. New York: Cambridge University Press.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, 42, 621–636.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777–813.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178–210.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45, 153–168.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, 16, 262–284.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing* (pp. 234–243). Honolulu, Hawaii: Association for Computational Linguistics.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 21086–21090. (Correction in: *Proceedings of the National Academy of Sciences of the United States of America*, 107, 5260)

O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 363–383). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed effects models in S and S-Plus*. New York: Springer Verlag.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, 52, 1–56.

Radach, R., & McConkie, G. W. (1998). Determinants of fixation positions in words during reading. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 77–100). Amsterdam: Elsevier.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.

Rayner, K. (2009). The 35th Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457–1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 720–732.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: a comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1188–1200.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, 16, 1–21.

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, 7, 34–55.

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 301–326). Amsterdam: Elsevier.

Vitu, F., McConkie, G. W., & Zola, D. (1998). About regressive saccades in reading and their relation to word identification. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 101–124). Amsterdam: Elsevier.

Yang, S.-N., & McConkie, G. W. (2001). Eye movements during reading: a theory of saccade initiation times. *Vision Research*, 41, 3567–3585.