

# Why long words take longer to read: the role of uncertainty about word length

**Klinton Bicknell**

Department of Psychology  
University of California, San Diego  
9500 Gilman Drive #109  
La Jolla, CA 92093-0109  
kbicknell@ucsd.edu

**Roger Levy**

Department of Linguistics  
University of California, San Diego  
9500 Gilman Drive #108  
La Jolla, CA 92093-0108  
rlevy@ucsd.edu

## Abstract

Some of the most robust effects of linguistic variables on eye movements in reading are those of word length. Their leading explanation states that they are caused by visual acuity limitations on word recognition. However, Bicknell (2011) presented data showing that a model of eye movement control in reading that includes visual acuity limitations and models the process of word identification from visual input (Bicknell & Levy, 2010) does not produce humanlike word length effects, providing evidence against the visual acuity account. Here, we argue that uncertainty about word length in early word identification can drive word length effects. We present an extension of Bicknell and Levy's model that incorporates word length uncertainty, and show that it produces more humanlike word length effects.

## 1 Introduction

Controlling the eyes while reading is a complex task, and doing so efficiently requires rapid decisions about when and where to move the eyes 3–4 times per second. Research in psycholinguistics has demonstrated that these decisions are sensitive to a range of linguistic properties of the text being read, suggesting that the eye movement record may be viewed as a detailed trace of the timecourse of incremental comprehension. A number of cognitive models of eye movement control in reading have been proposed, the most well-known of which are E-Z Reader (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003)

and SWIFT (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005). These models capture a large range of the known properties of eye movements in reading, including effects of the best-documented linguistic variables on eye movements: the frequency, predictability, and length of words.

Both models assume that word frequency, predictability, and length affect eye movements in reading by affecting word recognition, yet neither one models the process of identifying words from visual information. Rather, each of these models directly specifies the effects of these variables on exogenous word processing functions, and the eye movements the models produce are sensitive to these functions' output. Thus, this approach cannot answer the question of *why* these linguistic variables have the effects they do on eye movement behavior. Recently, Bicknell and Levy (2010) presented a model of eye movement control in reading that directly models the process of identifying the text from visual input, and makes eye movements to maximize the efficiency of the identification process. Bicknell and Levy (2012) demonstrated that this rational model produces effects of word frequency and predictability that qualitatively match those of humans: words that are less frequent and less predictable receive more and longer fixations. Because this model makes eye movements to maximize the efficiency of the identification process, this result gives an answer for the reason why these variables should have the effects that they do on eye movement behavior: a model that works to efficiently identify the text makes more and longer fixations on

words of lower frequency and predictability because it needs more visual information to identify them.

Bicknell (2011) showed, however, that the effects of word length produced by the rational model look quite different from those of human readers. Because Bicknell and Levy's (2010) model implements the main proposal for why word length effects should arise, i.e., visual acuity limitations, the fact that the model does not reproduce humanlike word length effects suggests that our understanding of the causes of word length effects may be incomplete.

In this paper, we argue that this result arose because of a simplifying assumption made in the rational model, namely, the assumption that the reader has veridical knowledge about the number of characters in a word being identified. We present an extension of Bicknell and Levy's (2010) model which does not make this simplifying assumption, and show in two sets of simulations that effects of word length produced by the extended model look more like those of humans. We argue from these results that uncertainty about word length is a necessary component of a full understanding of word length effects in reading.

## 2 Reasons for word length effects

The empirical effects of word length displayed by human readers are simple to describe: longer words receive more and longer fixations. The major reason proposed in the literature on eye movements in reading for this effect is that when fixating longer words, the average visual acuity of all the letters in the word will be lower than for shorter words, and this poorer average acuity is taken to lead to longer and more fixations. This intuition is built into the exogenous word processing functions in E-Z Reader and SWIFT. Specifically, in both models, the word processing rate slows as the average distance to the fovea of all letters in the word increases, and this specification of the effect of length on word processing rates is enough to produce reasonable effects of word length on eye movements: both models make more and longer fixations on longer words – similar to the pattern of humans – across a range of measures (Pollatsek, Reichle, & Rayner, 2006; Engbert et al., 2005) including the duration of the first fixation on a word (first fixation duration), the duration

of all fixations on a word prior to leaving the word (gaze duration), the rate at which a word is not fixated prior to a fixation on a word beyond it (skip rate), and the rate with which a word is fixated more than once prior to a word beyond it (refixation rate).

There are, however, reasons to believe that this account may be incomplete. First, while it is the case that the average visual acuity of all letters in a fixated word must be lower for longer words, this is just because there are additional letters in the longer word. While these additional letters pull down the average visual acuity of letters within the word, each additional letter should still provide additional visual information about the word's identity, an argument suggesting that longer words might require less – not more – time to be identified. In fact, in SWIFT, the exogenous word processing rate function slows as both the average and the sum of the visual acuities of the letters within the word decrease, but E-Z Reader does not implement this idea in any way. Additionally, a factor absent from both E-Z Reader and SWIFT, is that the visual neighborhoods of longer words (at least in English) appear to be sparser, when considering the number of words formed by a single letter substitution (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004), or the average orthographic Levenshtein distance of the most similar 20 words (Yarkoni, Balota, & Yap, 2008). Because reading words with more visual neighbors is generally slower (Pollatsek, Perea, & Binder, 1999), this argument gives another reason to expect longer words to require less – not more – time to be read.

So while E-Z Reader and SWIFT produce reasonable effects of word length on eye movement measures (in which longer words receive more and longer fixations) by assuming a particular effect of visual acuity, it is less clear whether a visual acuity account can yield reasonable word length effects in a model that also includes the two opposing effects mentioned above. Determining how these different factors should interact to produce word length effects requires a model of eye movements in reading that models the process of word identification from disambiguating visual input (Bicknell & Levy, *in press*). The model presented by Bicknell and Levy (2010) fits this description, and includes visual acuity limitations (in fact, identical to the visual acuity function in SWIFT). As already mentioned, how-

ever, Bicknell (2011) showed that the model did not yield a humanlike length effect. Instead, while longer words were skipped less often and refixated more (as for humans), fixation durations generally fell with word length – the opposite of the pattern shown by humans. This result suggests that visual acuity limitations alone cannot explain the positive effect of word length on fixation durations in the presence of an opposing force such as the fact that longer words have smaller visual neighborhoods.

We hypothesize that the reason for this pattern of results relates to a simplifying assumption made by Bicknell and Levy’s model. Specifically, while visual input in the model yields noisy information about the identities of letters, it gives veridical information about the number of letters in each word, for reasons of computational convenience. There are theoretical and empirical reasons to believe that this simplifying assumption is incorrect, that early in the word identification process human readers do have substantial uncertainty about the number of letters in a word, and further, that this may be especially so for long words. For example, results with masked priming have shown that recognition of a target word is facilitated by a prime that is a proper subset of the target’s letters (e.g., *blcn*–*balcon*; Peressotti & Grainger, 1999; Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006), providing evidence that words of different length have substantial similarity in early processing. For these reasons, some recent models of isolated word recognition (Gomez, Ratcliff, & Perea, 2008; Norris, Kinoshita, & van Casteren, 2010) have suggested that readers have some uncertainty about the number of letters in a word early in processing.

If readers have uncertainty about the length of words, we may expect that the amount of uncertainty would grow proportionally to length, as uncertainty is proportional to set size in other tasks of number estimation (Dehaene, 1997). This would agree with the intuition that an 8-character word should be more easily confused with a 9-character word than a 3-character word with a 4-character word. Including uncertainty about word length that is larger for longer words would have the effect of increasing the number of visual neighbors for longer words more than for shorter words, providing another reason (in addition to visual acuity limitations) that

longer words may require more and longer fixations.

In the remainder of this paper, we describe an extension of Bicknell and Levy’s (2010) model in which visual input provides stochastic – rather than veridical – information about the length of words, yielding uncertainty about word length, and in which the amount of uncertainty grows with length. We then present two sets of simulations with this extended model demonstrating that it produces more humanlike effects of word length, suggesting that uncertainty about word length may be an important component of a full understanding of the effects of word length in reading.

### 3 A rational model of reading

In this section, we describe our extension of Bicknell and Levy’s (2010) rational model of eye movement control in reading. Except for the visual input system, and a small change to the behavior policy to allow for uncertainty about word length, the model is identical to that described by Bicknell and Levy. The reader is referred to that paper for further computational details beyond what is described here.

In this model, the goal of reading is taken to be efficient text identification. While it is clear that this is not all that readers do – inferring the underlying structural relationships among words in a sentence and discourse relationships between sentences that determine text meaning is a fundamental part of most reading – all reader goals necessarily involve identification of at least part of the text, so text identification is taken to be a reasonable first approximation. There are two sources of information relevant to this goal: visual input and language knowledge, which the model combines via Bayesian inference. Specifically, it begins with a prior distribution over possible identities of the text given by its language model, and combines this with noisy visual input about the text at the eyes’ position, giving the likelihood term, to form a posterior distribution over the identity of the text taking into account both the language model and the visual input obtained thus far. On the basis of the posterior distribution, the model decides whether or not to move its eyes (and if so where to move them to) and the cycle repeats.

### 3.1 Formal problem of reading: Actions

The model assumes that on each of a series of discrete timesteps, the model obtains visual input around the current location of the eyes, and then chooses between three actions: (a) continuing to fixate the currently fixated position, (b) initiating a saccade to a new position, or (c) stopping reading. If the model chooses option (a), time simply advances, and if it chooses option (c), then reading immediately ends. If a saccade is initiated (b), there is a lag of two timesteps, representing the time required to plan and execute a saccade, during which the model again obtains visual input around the current position, and then the eyes move toward the intended target. Because of motor error, the actual landing position of the eyes is normally distributed around the intended target with the standard deviation in characters given by a linear function of the intended distance  $d$  ( $.87 + .084d$ ; Engbert et al., 2005).<sup>1</sup>

### 3.2 Language knowledge

Following Bicknell and Levy (2010), we use very simple probabilistic models of language knowledge: word  $n$ -gram models (Jurafsky & Martin, 2009), which encode the probability of each word conditional on the  $n - 1$  previous words.

### 3.3 Formal model of visual input

Visual input in the model consists of noisy information about the positions and identities of the characters in the text. Crucially, in this extended version of the model, this includes noisy information about the length of words. We begin with a visual acuity function taken from Engbert et al. (2005). This function decreases exponentially with retinal eccentricity  $\varepsilon$ , and decreases asymmetrically, falling off more slowly to the right than the left.<sup>2</sup> The model obtains visual input from the 19 character positions with the highest acuity  $\varepsilon \in [-7, 12]$ , which we refer to as the perceptual span. In order to provide the model with information about the current fixation position within the text, the model also obtains veridical in-

formation about the number of word boundaries to the left of the perceptual span.

Visual information from the perceptual span consists of stochastic information about the number of characters in the region and their identities. We make the simplifying assumption that the only characters are letters and spaces. Formally, visual input on a given timestep is represented as a string of symbols, each element of which has two features. One feature denotes whether the symbol represents a space ( $[+SPACE]$ ) or a letter ( $[-SPACE]$ ), an important distinction because spaces indicate word boundaries. Symbols that are  $[+SPACE]$  veridically indicate the occurrence of a space, while  $[-SPACE]$  symbols provide noisy information about the letter’s identity. The other feature attached to each symbol specifies whether the character in the text that the symbol was emitted from was being fixated ( $[+FIX]$ ) or not ( $[-FIX]$ ). The centrally fixated character has special status so that the model can recover the eyes’ position within the visual span.

This visual input string is generated by a process of moving a marker from the beginning to the end of the perceptual span, generally inserting a symbol into the visual input string for each character it moves across (EMISSION). To provide only noisy information about word length, however, this process is not always one of EMISSION, but sometimes it inserts a symbol into the visual input string that does not correspond to a character in the text (INSERTION), and at other times it fails to insert a symbol for a character in the text (SKIPPING). Specifically, at each step of the process, a decision is first made about INSERTION, which occurs with probability  $\delta$ . If INSERTION occurs, then a  $[-SPACE]$  identity for the character is chosen according to a uniform distribution, and then noisy visual information about that character is generated in the same way as for EMISSION (described below). If a character is not inserted, and the marker has already moved past the last character in the perceptual span, the process terminates. Otherwise, a decision is made about whether to emit a symbol into the visual input string from the character at the marker’s current position (EMISSION) or whether to skip outputting a symbol for that character (SKIPPING). In either case, the marker is advanced to the next character position. If the character at the marker’s cur-

<sup>1</sup>In the terminology of the literature, the model has only random motor error (variance), not systematic error (bias). Following Engbert and Krügel (2010), systematic error may arise from Bayesian estimation of the best saccade distance.

<sup>2</sup>While we refer to this function as visual acuity, it is clear from its asymmetric nature that it has an attentional component.

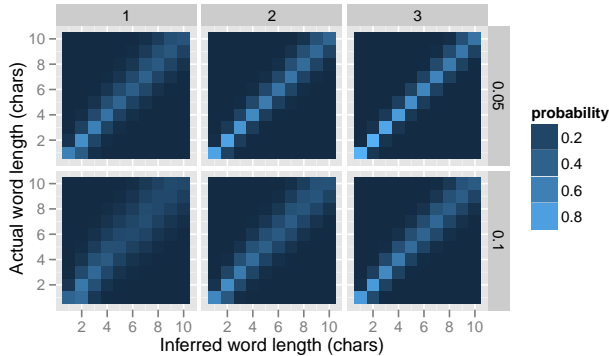


Figure 1: The expectation for the posterior distribution over the length of a word for actual lengths 1–10 after the model has received 1, 2, or 3 timesteps of visual input about the word, for two levels of length uncertainty:  $\delta \in \{.05, .1\}$ . These calculations use as a prior distribution the empirical distribution of word length in the BNC and assume no information about letter identity.

rent position is [+SPACE] or [+FIX], then EMISSION is always chosen, but if it is any other character, then SKIPPING occurs with probability  $\delta$ .

A [-SPACE] symbol (produced through EMISSION or INSERTION) contains noisy information about the identity of the letter that generated it, obtained via sampling. Specifically, we represent each letter as a 26-dimensional vector, where a single element is 1 and the others are zeros. Given this representation, a [-SPACE] symbol contains a sample from a 26-dimensional Gaussian with a mean equal to the letter’s true identity and a diagonal covariance matrix  $\Sigma(\epsilon) = \lambda(\epsilon)^{-1}I$ , where  $\lambda(\epsilon)$  is the visual acuity at eccentricity  $\epsilon$ . We scale the overall processing rate by multiplying each rate by  $\Lambda$ , set to 8 for the simulations reported here.

Allowing for INSERTION and SKIPPING means that visual input yields noisy information about the length of words, and this noise is such that uncertainty is higher for longer words. Figure 1 gives a visualization of this uncertainty. It shows the expectation for the posterior distribution over the length of a word for a range of actual word lengths, after the model has received 1, 2, or 3 timesteps of visual input about the word, at two levels of uncertainty. This figure demonstrates two things: first, that there is substantial uncertainty about word length even after three timesteps of visual input, and second, that this uncertainty is larger for longer words.

- (a)  $m = [.6, .7, \mathbf{.6}, .4, .3, .6]$ : Keep fixating (3)
- (b)  $m = [.6, .4, \mathbf{.9}, .4, .3, .6]$ : Move back (to 2)
- (c)  $m = [.6, .7, \mathbf{.9}, .4, .3, .6]$ : Move forward (to 6)
- (d)  $m = [.6, .7, \mathbf{.9}, .8, .7, .7]$ : Stop reading

Figure 2: Values of  $m$  for a 6 character text under which a model fixating position 3 would take each of its four actions, if  $\alpha = .7$  and  $\beta = .5$ .

### 3.4 Inference about text identity

The model’s initial beliefs about the identity of the text are given by the probability of each possible identity under the language model. On each timestep, the model obtains a visual input string as described above and calculates the likelihood of generating that string from each possible identity of the text. The model then updates its beliefs about the text via standard Bayesian inference: multiplying the probability of each text identity under its prior beliefs by the likelihood of generating the visual input string from that text identity and normalizing. We compactly represent all of these distributions using weighted finite-state transducers (Mohri, 1997) using the OpenFST library (Allauzen, Riley, Schalkwyk, Skut, & Mohri, 2007), and implement belief update with transducer composition and weight pushing.

### 3.5 Behavior policy

The model uses a simple policy with two parameters,  $\alpha$  and  $\beta$ , to decide between actions based on the marginal probability  $m$  of the most likely character  $c$  in each position  $j$ ,

$$m(j) = \max_c p(w_j = c)$$

where  $w_j$  indicates the character in the  $j$ th position. A high value of  $m$  indicates relative confidence about the character’s identity, and a low value relative uncertainty. Because our extension has uncertainty about the absolute position of its eyes within the text, each position  $j$  is now defined relative to the centrally fixated character.

Figure 2 illustrates how the model decides among four possible actions. If the value of  $m(j)$  for the current position of the eyes is less than the parameter  $\alpha$ , the model continues fixating the current position (2a). Otherwise, if the value of  $m(j)$  is less than the

parameter  $\beta$  for some leftward position, the model initiates a saccade to the closest such position (2b). If no such positions exist to the left, the model initiates a saccade to  $n$  characters past the closest position to the right for which  $m(j) < \alpha$  (2c).<sup>3</sup> Finally, if no such positions exist, the model stops reading (2d). Intuitively, then, the model reads by making a rightward sweep to bring its confidence in each character up to  $\alpha$ , but pauses to move left to reread any character whose confidence falls below  $\beta$ .

## 4 Simulation 1: full model

We now assess the effects of word length produced by the extended version of the model. Following Bicknell (2011), we use the model to simulate reading of a modified version of the Schilling, Rayner, and Chumbley (1998) corpus of typical sentences used in reading experiments. We compare three levels of length uncertainty:  $\delta \in \{0, .05, .1\}$ . The first of these ( $\delta = 0$ ) corresponds to Bicknell and Levy’s (2010) model, which has no uncertainty about word length. We predict that increasing the amount of length uncertainty will make effects of word length more like those of humans, and we compare the model’s length effects to those of human readers of the Schilling corpus.

### 4.1 Methods

#### 4.1.1 Model parameters and language model

Following Bicknell (2011), the model’s language knowledge was an unsmoothed bigram model using a vocabulary set consisting of the 500 most frequent words in the British National Corpus (BNC) as well as all the words in the test corpus. Every bigram in the BNC was counted for which both words were in vocabulary, and – due to the intense computation required for exact inference – this set was trimmed by removing rare bigrams that occur less than 200 times (except for bigrams that occur in the test corpus), resulting in a set of about 19,000 bigrams, from which the bigram model was constructed.

#### 4.1.2 Optimization of policy parameters

We set the parameters of the behavior policy  $(\alpha, \beta)$  to values that maximize reading efficiency.

<sup>3</sup>The role of  $n$  is to ensure that the model does not center its visual field on the first uncertain character. For the present simulations, we did not optimize this parameter, but fixed  $n = 3$ .

We define reading efficiency  $E$  to be an interpolation of speed and accuracy,  $E = (1 - \gamma)L - \gamma T$ , where  $L$  is the log probability of the true identity of the text under the model’s beliefs at the end of reading,  $T$  is the number of timesteps before the model stopped reading, and  $\gamma$  gives the relative value of speed. For the present simulations, we use  $\gamma = .1$ , which produces reasonably accurate reading. To find optimal values of the policy parameters  $\alpha$  and  $\beta$  for each model, we use the PEGASUS method (Ng & Jordan, 2000) to transform this stochastic optimization problem into a deterministic one amenable to standard optimization algorithms, and then use coordinate ascent.

#### 4.1.3 Test corpus

We test the model on a corpus of 33 sentences from the Schilling corpus slightly modified by Bicknell and Levy (2010) so that every bigram occurred in the BNC, ensuring that the results do not depend on smoothing.

#### 4.1.4 Analysis

With each model, we performed 50 stochastic simulations of the reading of the corpus. For each run, we calculated the four standard eye movement measures mentioned above for each word in the corpus: first fixation duration, gaze duration, skipping probability, and refixation probability. We then averaged each of these four measures across runs for each word token in the corpus, yielding a single mean value for each measure for each word.

Comparing the fixation duration measures to humans required converting the model’s timesteps into milliseconds. We performed this scaling by multiplying the duration of each fixation by a conversion factor set to be equal to the mean human gaze duration divided by the mean model gaze duration for words with frequencies higher than 1 in 100, meaning that the model predictions exactly match the human mean for gaze durations on these words.

## 4.2 Results

Figure 3 presents the results for all four measures of interest. Looking first at the model with no uncertainty, we see that the results replicate those of Bicknell (2011): while there is a monotonic effect of word length on skip rates and refixation rates in the same direction as humans, longer words receive

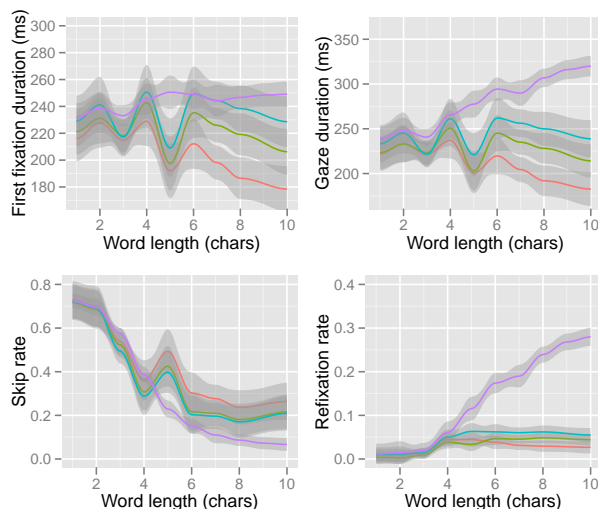


Figure 3: Effects of word length in three version of the full model with  $\delta = 0$  (red),  $\delta = 0.05$  (green), and  $\delta = 0.1$  (blue) on first fixation durations, gaze durations, skip rates, and refixation rates compared with the empirical human data for this corpus (purple). Estimates obtained via loess smoothing and plotted with standard errors.

shorter fixations in the model, opposite to the pattern found in human data. As predicted, adding length uncertainty begins to reverse this effect: as uncertainty is increased, the effect of word length on fixation durations becomes less negative.

However, while these results look more like those of humans, there are still substantial differences. For one, even for the model with the most uncertainty, the effect of word length – while not negative – is also not really positive. Second, the effect appears rather non-monotonic. We hypothesize that these two problems are related to the aggressive trimming we performed of the model’s language model. By removing low frequency words and bigrams, we artificially trimmed especially the visual neighborhoods of long words, since frequency and length are negatively correlated. This could have led to another inverse word length effect, which even adding more length uncertainty was unable to fully overcome. In effect, extending the visual neighborhoods of long words (by adding length uncertainty) may not have much effect if we have removed all the words that would be in those extended neighborhoods. In addition, the aggressive trimming could have been responsible for the non-monotonicities apparent in the model’s predictions. We performed another set of

simulations using a language model with substantially less trimming to test these hypotheses.

## 5 Simulation 2: model without context

In this simulation, we used a unigram language model instead of the bigram language model used in Simulation 1. Since this model cannot make use of linguistic context, it will not show as robust effects of linguistic variables such as word predictability (Bicknell & Levy, 2012), but since here our focus is on effects of word length, this limitation is unlikely to concern us. Crucially, because of the model’s simpler structure, it allows for the use of a substantially larger vocabulary than the bigram model used in Simulation 1. In addition, using this model avoids the problems mentioned above associated with trimming bigrams. We predicted that this language model would allow us to obtain effects of word length on fixation durations that were actually positive (rather than merely non-negative), and that there would be fewer non-monotonicities in the function.

### 5.1 Methods

Except the following, the methods were identical to those of Simulation 1. We replaced the bigram language model with a unigram language model. Training was performed in the same manner, except that instead of including only the most common 500 words in the BNC, we included all words that occur at least 200 times (corresponding to a frequency of 2 per million; about 19,000 words). Because of the greater computational complexity for the two models with non-zero  $\delta$ , we performed only 20 simulations of the reading of the corpus instead of 50.

### 5.2 Results

Figure 4 presents the results for all four measures of interest. Looking at the model with no uncertainty, we see already that the predictions are a substantially better fit to human data than was the full model. The skipping and refixation rates look substantially more like the human curves. And while the word length effect on first fixation duration is still negative, it is already non-negative for gaze duration. This supports our hypotheses that aggressive trimming were partly responsible for the full model’s negative word length effect.



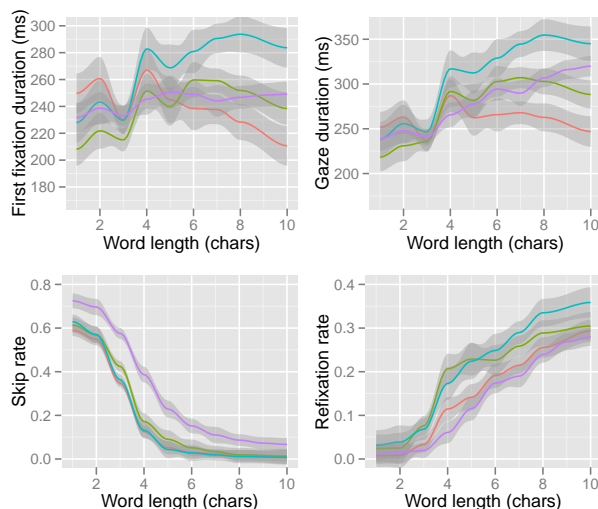


Figure 4: Effects of word length in three version of the model without context (unigram model) with  $\delta = 0$  (red),  $\delta = 0.05$  (green), and  $\delta = 0.1$  (blue) on first fixation durations, gaze durations, skip rates, and refixation rates compared with the empirical human data for this corpus (purple). Estimates obtained via loess smoothing and plotted with standard errors.

Moving on to the models with uncertainty, we see that predictions are still in good agreement with humans for skip rates and refixation rates. More interestingly, we see that adding length uncertainty makes both durations measures relatively positive functions of word length. While the overall size of the effect is incorrect for first fixation durations, we see striking similarities between the models predictions and human data on both duration measures. For first fixations, the human pattern is that durations go up from word lengths 1 to 2, down from 2 to 3 (presumably because of ‘the’), and then up to 5, after which the function is relatively flat. That pattern also holds for both models with uncertainty. For gaze duration, both models more or less reproduce the human pattern of a steadily-increasing function throughout the range, and again match the human function in dipping for word length 3. For gaze durations, even the overall size of the effect produced by the model is similar to that of humans. These results confirm our original hypothesis that adding length uncertainty would lead to more humanlike word length effects. In addition, comparing the results of Simulation 2 with Simulation 1 reveals the importance to this account of words having realis-

tic visual neighborhoods. When the visual neighborhoods of (especially longer) words were trimmed to be artificially sparse, adding length uncertainty did not allow the model to recover the human pattern.

## 6 Conclusion

In this paper, we argued that the success of major models of eye movements in reading to reproduce the (positive) human effect of word length via acuity limitations may be a result of not including opposing factors such as the negative correlation between visual neighborhood size and word length. We described the failure of the rational model presented in Bicknell and Levy (2010) to obtain humanlike effects of word length, despite including all of these factors, suggesting that our understanding of word length effects in reading is incomplete. We proposed a new reason for word length effects – uncertainty about word length that is larger for longer words – and noted that this reason was not implemented in Bicknell and Levy’s model because of a simplifying assumption. We presented an extension of the model relaxing this assumption, in which readers obtain noisy information about word length, and showed through two sets of simulations that the new model produces effects of word length that look more like those of human readers. Interestingly, while adding length uncertainty made both models more humanlike, it was only in Simulation 2 – in which words had more realistic visual neighborhoods – that all measures of the effect of word length on eye movements showed the human pattern, underscoring the importance of the structure of the language for this account of word length effects.

We take these results as evidence that word length effects cannot be completely explained through limitations on visual acuity. Rather, they suggest that a full understanding of the reasons underlying word length effects on eye movements in reading should include a notion of uncertainty about the number of letters in a word, which grows with word length.

## Acknowledgments

This research was supported by NIH grant T32-DC000041 from the Center for Research in Language at UC San Diego to K. B. and by NSF grant 0953870 and NIH grant R01-HD065829 to R. L.



## References

- Allauzen, C., Riley, M., Schalkwyk, J., Skut, W., & Mohri, M. (2007). OpenFst: A general and efficient weighted finite-state transducer library. In *Proceedings of the Ninth International Conference on Implementation and Application of Automata, (CIAA 2007)* (Vol. 4783, p. 11-23). Springer.
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*, 283–316.
- Bicknell, K. (2011). *Eye movements in reading as rational behavior*. Unpublished doctoral dissertation, University of California, San Diego.
- 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.
- Bicknell, K., & Levy, R. (2012). Word predictability and frequency effects in a rational model of reading. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Bicknell, K., & Levy, R. (in press). The utility of modelling word identification from visual input within models of eye movements in reading. *Visual Cognition*.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- Engbert, R., & Krügel, A. (2010). Readers use Bayesian estimation for eye movement control. *Psychological Science*, *21*, 366–371.
- 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.
- Gomez, P., Ratcliff, R., & Perea, M. (2008). The Overlap model: A model of letter position coding. *Psychological Review*, *115*, 577–601.
- Grainger, J., Granier, J.-P., Farioli, F., Van Assche, E., & van Heuven, W. J. B. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 865–884.
- Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Mohri, M. (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, *23*, 269–311.
- Ng, A. Y., & Jordan, M. (2000). PEGASUS: A policy search method for large MDPs and POMDPs. In *Uncertainty in Artificial Intelligence, Proceedings of the Sixteenth Conference* (pp. 406–415).
- Norris, D., Kinoshita, S., & van Casteren, M. (2010). A stimulus sampling theory of letter identity and order. *Journal of Memory and Language*, *62*, 254–271.
- Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics*, *61*, 691–706.
- Pollatsek, A., Perea, M., & Binder, K. S. (1999). The effects of “neighborhood size” in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1142–1158.
- 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.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models.

- Behavioral and Brain Sciences*, 26, 445–526.
- Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26, 1270–1281.
- Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15, 971–979.