

## **UC Merced**

# **Proceedings of the Annual Meeting of the Cognitive Science Society**

### **Title**

Rational eye movements in reading combining uncertainty about previous words with contextual probability

### **Permalink**

<https://escholarship.org/uc/item/1mm294p6>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 32(32)

### **ISSN**

1069-7977

### **Authors**

Bicknell, Klinton

Levy, Roger

### **Publication Date**

2010

Peer reviewed

# Rational eye movements in reading combining uncertainty about previous words with contextual probability

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

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

## Abstract

While there exist a range of sophisticated models of eye movements in reading, it remains an open question to what extent human eye movement behavior during reading is adaptive given the demands of the task. In this paper, we help to answer this question by presenting a model of reading that corrects two problems with a rational model of the task, Mr. Chips (Legge, Klitz, & Tjan, 1997). We show that the resulting model is closer to human performance across two measures, supporting the idea that many components of eye movement behavior in reading can be well understood as a rational response to the demands of the task.

**Keywords:** computational modeling; rational analysis; eye movements; reading

## Introduction

Choosing when and where to move one's eyes during reading is one of the most complicated skilled tasks humans perform. While there are a number of computational models achieving good numerical fits on eye movement data from reading (e.g., Reichle, Pollatsek, & Rayner, 2006; Engbert, Nuthmann, Richter, & Kliegl, 2005), it is still unclear to what extent the complex behaviors observed are rational responses to the demands of the problem itself and to what extent they arise from the idiosyncrasies and restrictions of human cognition. Legge, Klitz, and Tjan (1997) started to answer this question with Mr. Chips, a model which predicts eye movements that approximate an optimal solution to one formalization of the task of reading. Legge et al. pointed out that their model's behavior exhibits a number of patterns also found in human reading, providing evidence for understanding those behaviors as rational responses to the task. Despite its success, however, the Mr. Chips model oversimplifies two important aspects of the problem of reading, and also has empirical problems accounting for human reading behavior in two domains. In this paper, we propose a model extending Mr. Chips that removes these two oversimplifications to make the model's task more similar to that faced by humans. We show that the resulting model also remedies the two empirical deficiencies in Mr. Chips, further supporting the notion that many aspects of human reading behavior can be explained as rational responses to the demands of reading.

The essentials of the problem of making eye movements in reading are determining how long to leave the eyes in a given spot and – when a reader decides to move them – where to go. These decisions are made sequentially to produce the alternating sequence of fixations (relatively stable periods) and saccades (movements) that characterizes the eye

movement record. The past 30 years have seen a proliferation of experimental studies investigating this topic, which have answered a number of low-level questions such as the nature of the perceptual span and constraints on saccade latency as well as questions concerning the relationship between eye movements and higher-level cognitive processes such as the effect of word frequency and predictability (see Rayner, 1998 for an overview). Sophisticated computational models have been developed based on these findings, the most well-known of which are E-Z Reader (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle et al., 2006) and SWIFT (Engbert et al., 2005). Both E-Z Reader and SWIFT assume that lexical processing (or word recognition) is the primary driver for eye-movements in reading, and both have enjoyed considerable success, in large part because they achieve very good fits to eye movement data from reading in a number of contexts, using a relatively small number of parameters. Despite their empirical strength, they fail to illuminate the reason why human reading behavior looks the way it does in one crucial respect – the extent to which it resembles a rational response to the problem posed by reading.

One leading approach for answering such questions is that of rational analysis (Anderson, 1990), a paradigm in which one formalizes the goals and cognitive and physical constraints relevant to a problem and develops a model of optimal behavior under those condition. To the extent that the behavior of the model is similar to that of humans, this provides a new way of understanding the reason why human behavior looks the way it does – it is the best way to solve the problem. The relationship between rational models and models such as E-Z Reader and SWIFT is well understood in terms of Marr's (1982) levels of analysis. Marr distinguishes three levels of mutually-constraining analyses that can be performed on cognitive processes: the *computational* level, which specifies the nature of the computation being performed, the information relevant to solving it, and the way to combine that information to solve it; the *algorithmic* level, which specifies the representation for the input and output and the algorithm by which the agent goes about solving it; and the *implementational* level, which specifies how the representations and algorithm are realized neurally. In these terms, rational models generally provide answers at the computational level of analysis. Models such as E-Z Reader and SWIFT help us to understand the algorithmic level, but cannot answer questions about the extent to which human reading is rational.

Legge et al. (1997) presented a computational level analy-

sis of reading, formalizing the central task – as in E-Z Reader and SWIFT – as one of serial word identification. They presented the Mr. Chips model, which approximates optimal behavior under their formalization, and shows a number of similarities with human reading behavior. Here, we point out two problems with their model of reading. First, their model takes the task to be to identify a string of independent words rather than a coherent sequence, i.e., their model does not make use of linguistic context, which experimental work suggests that humans use (McDonald & Shillcock, 2003). Second, it assumes that the task of the reader is to identify each word with complete certainty, yet recent evidence suggests that readers maintain uncertainty as to the identities of previous words (Levy, Bicknell, Slattery, & Rayner, 2009). In addition to these problems in their model’s design, the model also makes incorrect predictions for two relatively basic measures of eye movements in reading: saccade sizes and word skipping rates. The model we present fixes these two design problems by including linguistic context and using a flexible word identification criterion, and results in improved performance in accounting for human saccade sizes and word skipping rates.

The plan of the remainder of the paper is as follows. First, we describe the details of the Mr. Chips model, along with its empirical successes and failures. Next, we describe our extension of the Mr. Chips model, and finally present two experiments showing that fixing each of the two design problems results in performance more like humans.

### Mr. Chips

The task of reading in the Mr. Chips model (Legge et al., 1997) is one of planning saccades for serial word identification. That is, the model works by gathering visual input from the current fixation location and using that visual input to plan a saccade. That saccade is then executed (with some motor error), visual input is obtained from the new location, and the cycle repeats. When one word is identified with 100% confidence, identification of the next word begins. Thus, the only decision the model makes is where to move the eyes next. There are just three sources of information relevant to making that decision. Visual input and knowledge of the language are combined to identify words, and knowledge of the motor error in the system assists in the planning problem. Since it forms the basis for our model, we describe the Mr. Chips model here in detail, discussing in turn each of the sources of information and then the algorithm by which the model combines them to choose saccades. To match the description of our model later, we use a notation a bit different than Legge et al. to describe Mr. Chips.

#### Information sources

**Visual input** The visual input in Mr. Chips consists of the veridical identities of the nine characters centered on the fixated character (representing the visual fovea), as well as partial information about the four characters on either side of this range (representing the visual periphery). This partial information is simply word boundary information, indicating

whether each character is part of a word or not (e.g., a space). The number of characters in each of these ranges was chosen to be representative of the perceptual span for readers of English, known to be around 17–19 characters (Rayner, 1998).

**Language knowledge** The model’s knowledge of language consists of simply word frequency information, i.e., a unigram model. Note that this means the model cannot make use of the linguistic context to aid in word identification.

**Motor error** The final component of the model’s knowledge of the task is that of motor error, the distribution of a saccade’s landing position given the intended target position the model chooses. In Mr. Chips, the  $i$ th landing position  $\ell_i$  is normally distributed around the  $i$ th intended target position  $t_i$  with a standard deviation of 30% of the intended distance<sup>1</sup>

$$\ell_i \sim \mathcal{N}(t_i, (0.3 \cdot |t_i - \ell_{i-1}|)^2). \quad (1)$$

### Model

We now give the algorithm that the Mr. Chips model uses to select the intended target for the next saccade. First, note that given the visual input obtained by the model from the first to the  $i$ th fixation  $\mathcal{I}_1^i$  and the word frequency information, the model can calculate the posterior probability of any possible identity of a word  $w$  that is consistent with the visual input by normalizing its probability from the language model by the total probability of all visually consistent identities,

$$p(w|\mathcal{I}_1^i) = \frac{\chi(\mathcal{I}_1^i, w)p(w)}{\sum_{w'} \chi(\mathcal{I}_1^i, w')p(w')} \quad (2)$$

where  $\chi(\mathcal{I}, w)$  is an indicator function with a value of 1 if  $w$  is consistent with the visual input  $\mathcal{I}$  and 0 otherwise, and  $p(w)$  is the probability of  $w$  under the language model.

To identify a given word, the model selects the saccade target  $\hat{t}_i$  that, on average, will minimize the entropy in this distribution, i.e., that is expected to give the most information about the word’s identity

$$\hat{t}_i = \operatorname{argmin}_{t_i} E [H(w|\mathcal{I}_1^i)|t_i, \mathcal{I}_1^{i-1}] \quad (3)$$

$$= \operatorname{argmin}_{t_i} \sum_{\mathcal{I}_i} H(w|\mathcal{I}_1^i) p(\mathcal{I}_i|t_i, \mathcal{I}_1^{i-1}). \quad (4)$$

That is, the minimum can be found by calculating the conditional entropy produced by each possible new input sequence and weighting those entropies by the probability of getting that input sequence given a choice of target location. In information theory (Cover & Thomas, 2006), conditional entropy is standardly defined as

$$H(w|\mathcal{I}_1^i) = - \sum_w p(w|\mathcal{I}_1^i) \log p(w|\mathcal{I}_1^i). \quad (5)$$

<sup>1</sup>In the terminology of the literature, this model has only ‘random’ motor error (variance) and not ‘systematic’ motor error (bias), under the assumption that an optimal model would just compensate for any systematic problems with its motor control system.

The second term in the formula for  $\hat{t}_i$ , the probability of a particular visual input given a target location and previous input, is given by marginalizing over possible landing positions

$$p(\mathcal{I}_i|t_i, \mathcal{I}_1^{i-1}) = \sum_{\ell_i} p(\ell_i|t_i) p(\mathcal{I}_i|\ell_i, \mathcal{I}_1^{i-1}) \quad (6)$$

and then possible words

$$p(\mathcal{I}_i|\ell_i, \mathcal{I}_1^{i-1}) = \sum_w p(\mathcal{I}_i|\ell_i, w) p(w|\mathcal{I}_1^{i-1}). \quad (7)$$

Putting these together, we have that  $\hat{t}_i$  is selected as

$$\operatorname{argmin}_{t_i} \sum_{\mathcal{I}_i} H(w|\mathcal{I}_i^i) \sum_{\ell_i} p(\ell_i|t_i) \sum_w p(\mathcal{I}_i|\ell_i, w) p(w|\mathcal{I}_1^{i-1}). \quad (8)$$

That is, we can calculate the expected conditional entropy for each possible value of  $t_i$  by summing over all possible inputs, whose probabilities are given by summing over all possible identities of the word and landing positions. To see that this sum ranges over a finite number of values, note first that there are only a finite number of possible word identities  $w$  to sum over. Given the possible word identities, there are only a finite number of landing positions  $\ell_i$  for which the visual information could possibly help in identifying the word – any landing positions outside this range will not produce any reduction in entropy. Since there is a single visual input  $\mathcal{I}_i$  for each combination of landing position and word identity, this summation is over a finite range. To ensure finiteness of the search to find the value of  $t_i$  that produces the minimum entropy, Mr. Chips only searches those within the range of the  $\ell_i$  that could give some information about the current word. In case of ties, the model selects the furthest position to the right.

### Comparing Mr. Chips to humans

Legge, Hooven, Klitz, Mansfield, and Tjan (2002) present a number of ways in which the behavior of the Mr. Chips model is similar to human reading behavior. The model produces behavior that replicates a number of human findings in word skipping rates, initial fixation locations on words, and refixation rates. The result for word skipping rates – where word skipping for the model is defined as never having any of the word’s characters as the centrally fixated character – is that longer words are skipped less often, and the slope of the relationship between word length and skipping rate has a very similar slope for the model as for humans. For initial word fixation locations, or landing positions, the model replicates the human behavior of most commonly landing at or just to the left of the word’s center, and also the fact that the landing position shifts toward the left as the launch site of the saccade shifts further to the left. For refixations, the model mimics human behavior in showing the proportion of refixations to increase with word length, and in addition, within a given word length class, the model refixates low frequency words a higher proportion of the time than high frequency ones. Finally, as a function of landing position, refixations are the least likely for the model, as for humans, when the initial landing position is near the center of the word.

As noted above, however, it is also the case that the model exhibits some behavior very different from that of human readers. For example, the model’s average saccade length is just 6.3 characters, noticeably lower than that for humans, who are around 8 (Rayner, 1998). Second, although, as mentioned, the slope of the relationship between word skipping rates and word length has a similar slope for the model as for humans, the model skips far fewer words than humans do.<sup>2</sup> In short, judging by these two measures, a rational model that is using all the information available and expensively calculating the best saccades to reduce entropy in word identification appears to be reading slower than humans do.

In rational analysis, the fact that an ‘optimal’ model is performing worse than humans (here in terms of speed) suggests two likely problems: (a) the model is not making use of all the information that humans use or (b) the model’s computational goal is not the same as the one that humans are solving. As suggested above, we argue that in this case both reasons are partially to blame. Since it has only word frequency information as its model of language, the Mr. Chips model cannot make use of linguistic context to aid in word identification, while there is evidence that humans make heavy use of it. The model also assumes that the goal is to identify each word with 100% confidence, but experiments suggest that humans do not. In the next section, we modify the Mr. Chips model to include some information about linguistic context and a flexible identification confidence criterion.

### Extending Mr. Chips

The model described here generalizes the Mr. Chips model in three ways. First, it can use an arbitrary language model as its source of language knowledge, and thus make use of information about the linguistic context in word identification, solving the first problem with Mr. Chips we pointed out above. Second, it can move on to the next word after it achieves a flexible level of certainty about the current word’s identity, solving the second problem. Finally, our model also allows for the standard deviation of the motor error to be an arbitrary linear function of the intended distance, allowing us to incorporate a more realistic motor error function. We describe the model in the same format as we described the Mr. Chips model, first describing its sources of information, and then its algorithm for selecting saccade targets.

#### Information sources

**Visual input** The visual input component is unchanged from the original Mr. Chips model.

**Language knowledge** The model’s knowledge of language is represented by an arbitrary language model that can generate string prefix probabilities, e.g., an  $n$ -gram model or a

<sup>2</sup>The graph given in Legge et al. (2002) appears to show remarkably similar word skipping rates between the model and humans, but that graph is from the sole simulation in the paper for which Legge et al. assumed no motor error. When motor error is included, the skipping rates are significantly lower for the model than for humans, as shown in Figure 1.

probabilistic context-free grammar (*PCFG*). Such models can capture the between-word dependencies needed for the model to make use of linguistic context in word identification.

**Motor error** In our model as in Mr. Chips, the  $i$ th landing position is normally distributed around the  $i$ th target location, except that the standard deviation is an arbitrary linear function of the intended distance

$$\ell_i \sim \mathcal{N}(t_i, (\beta_0 + \beta_1 |t_i - \ell_{i-1}|)^2) \quad (9)$$

allowing for the use of a more realistic motor error function. Experiments in this paper use the one from SWIFT (Engbert et al., 2005;  $\beta_0 = 0.87$ ,  $\beta_1 = 0.084$ ).

### Algorithm

As in the original Mr. Chips model, at any given point in time, the model is working to identify one word. However, this revised model considers the goal of identifying this word achieved when the marginal probability of some identity for the word given the visual input exceeds a predefined threshold probability  $\alpha$ . This flexibility requires the algorithm to be substantially modified to allow for uncertainty about previous word identities and the use of linguistic context. We denote the sequence of words as  $W$ , where the first word is  $W_1$ .

Because every word in Mr. Chips was identified with complete certainty, the model always knew precisely at which position the next word to be identified began, and its goal was always to identify this next word. Now that the model has uncertainty about the identities of previous words, however, the goal must be changed. In the revised model, the reader is always focused on some character position  $n$ , and its goal is to identify whether some word  $W_{(n)}$  begins at that position, and if so, which one, with confidence exceeding  $\alpha$ . Once the model has achieved this goal, it then chooses a new character position  $n$  via a procedure whose description we leave for later. To be explicit about this goal, we slightly update our original equation for choosing  $\hat{t}_i$ , swapping  $w$  out for  $W_{(n)}$ ,

$$\hat{t}_i = \operatorname{argmin}_{t_i} \sum_{\mathcal{I}_i} H(W_{(n)} | \mathcal{I}_i^i) p(\mathcal{I}_i | t_i, \mathcal{I}_1^{i-1}) \quad (10)$$

where the conditional entropy is calculated assuming that some word does in fact begin at position  $n$ . The fact that our language model can now make use of linguistic context means that the equation for finding the probability of the current word given some visual input (Equation 2) must also be changed to marginalize over identities of the preceding words

$$p(W_{(n)} | \mathcal{I}_1^i) = \sum_{W_1^{(n-1)}} p(W_{(n)} | \mathcal{I}_1^i, W_1^{(n-1)}) p(W_1^{(n-1)} | \mathcal{I}_1^i). \quad (11)$$

These probabilities of strings consistent with the visual input are again given probabilities according to their probability in the language model normalized by the probability of all other consistent strings (cf. Equation 2)

$$p(W | \mathcal{I}_1^i) = \frac{\chi(\mathcal{I}_1^i, W) p(W)}{\sum_W \chi(\mathcal{I}_1^i, W) p(W)}. \quad (12)$$

The second term in Equation 10 is expanded as in Mr. Chips by marginalizing over possible landing positions

$$p(\mathcal{I}_i | t_i, \mathcal{I}_1^{i-1}) = \sum_{\ell_i} p(\ell_i | t_i) p(\mathcal{I}_i | \ell_i, \mathcal{I}_1^{i-1}), \quad (13)$$

but now to incorporate information about the linguistic context, we must next marginalize over possible full sentence strings instead of possible words

$$p(\mathcal{I}_i | \ell_i, \mathcal{I}_1^{i-1}) = \sum_W p(\mathcal{I}_i | \ell_i, W) p(W | \mathcal{I}_1^{i-1}). \quad (14)$$

If we make the simplifying assumption that the model does not consider possible future input about words that are after  $W_{(n)}$ , this sum can again be finitely computed for a given  $t_i$  by a relatively straightforward dynamic programming scheme. The range of possible values of  $t_i$  to search through also grows relative to Mr. Chips, because the model must consider not only any position that can give visual input about  $W_{(n)}$  itself, but also positions that can give information about any position of uncertainty, since that may indirectly help to identify  $W_{(n)}$  through linguistic context. In the case where the language model is an  $n$ -gram model, it can be shown that the minimum value of  $t_i$  that can contribute toward helping to identify  $W_{(n)}$  only extends back to the first uncertain character after the most recent string of  $n - 1$  words for which the model has no residual uncertainty. Having established the method of selecting a saccade to identify  $W_{(n)}$ , we next give a description of the full algorithm of the model, including how to select  $n$ .

The model always begins reading by focusing on identifying  $W_{(0)}$ . Once the probability of some identity for  $W_{(0)}$  is greater than  $\alpha$ , all the possible identities of  $W_{(0)}$  that have not been ruled out by visual input are combined into a set of possible ‘prefixes’. Each of these prefixes has a conditional probability given the visual input, and each one predicts that the next word in the sentence begins at a particular position (i.e., two characters past the end of that string). Thus, the set of prefixes specify a probability distribution over the possible positions at which the next word begins. The model simply selects the most likely such position as the next character position  $n$  to focus on identifying  $W_{(n)}$ .

Now in the general case, the system has a set of prefixes together with their conditional probabilities given the visual input, and a position  $n$ , which it is trying to identify the word beginning at. It plans and executes saccades according to the formula for  $\hat{t}_i$ , and after getting each new piece of visual information, the model rules out not only possible candidates for the current word, but also possible prefix strings, and renormalizes both distributions. The model’s attempt to identify  $W_{(n)}$  can now end in one of two ways: (a) the model’s confidence in some identity of  $W_{(n)}$  exceeds the confidence threshold  $\alpha$  or (b) the model eliminates all possible candidates for  $W_{(n)}$  and thus knows that no word begins at that position. In the former case, the model creates all possible concatenations of prefixes ending 2 characters prior to  $W_{(n)}$  (i.e., prefixes whose next word begins at  $n$ ) with all the possible identities of

$W_{(n)}$ , and adds these new strings to the set of prefixes. Then, in both cases, it removes those original prefixes whose next word begins at  $n$  from the set. Note that this update of the list of prefixes leaves unaffected prefixes that are incompatible with a word beginning at position  $n$ , but still compatible with visual input. Finally, since the set of prefixes again gives a distribution over the starting position of the next word, the model selects the most likely new  $n$  and the cycle continues.

### Experiment 1

With our new model in place, we can now describe the two experiments we performed to test our hypotheses about the reasons for the Mr. Chips model’s performance being below that of humans in terms of average saccade length and word skipping rates. In Experiment 1, we test the hypothesis that one of the reasons that its performance was below humans is due to its assumption that the goal of the reader is to identify each word with 100% confidence. Specifically, we compare the performance of our model using a 100% criterion vs. a 90% criterion. The former is equivalent to Mr. Chips except for the more realistic motor error function, so for ease of exposition, we will refer to it simply as Mr. Chips.

#### Methods

**Confidence criterion** We set  $\alpha = 1.0$  to replicate Mr. Chips, and  $\alpha = 0.9$  for the model using a slightly lower confidence criterion to trigger moving on to the next word.

**Language model** Both models used a unigram language model, smoothed with Kneser-Ney under default parameters (Chen & Goodman, 1998; equivalent to add- $\delta$  smoothing for a unigram model). As in Legge et al. (2002), the models were trained on a 280,000 word corpus of *Grimms’ Fairy Tales*, containing 7503 unique words. This corpus was normalized by Legge et al. to convert all letters to lowercase, remove all punctuation other than apostrophes, convert all numbers to their alphabetic equivalents, and remove all gibberish words.

**Text** Following Legge et al. (2002), we tested the models by simulating the reading of 40,000 words of text generated from the language model, ensuring that the reading models had a normative probability model for the text they were reading.

#### Results

The results for mean saccade size for both models are given in the top two rows of Table 1. As shown in the table, using a criterion of 90% increases the average size of saccades, bringing it a bit closer to the human estimate of about 8 characters. The results for word skipping rates for the two models are plotted as the lower two lines in Figure 1. The results show a modest increase in word skipping rates across almost all word lengths for the new model with a lower criterion, bringing it closer to human performance.

#### Discussion

Although the gain is modest, the results of Experiment 1 show that changing the goal of the model to one more sim-

Table 1: Mean saccade size (and std. error) for each model

Model	Mean saccade size
Mr. Chips	6.7 (.012)
Without context, 90% criterion	7.1 (.013)
With context, 90% criterion	7.5 (.014)
Humans	$\approx 8$ (Rayner, 1998)

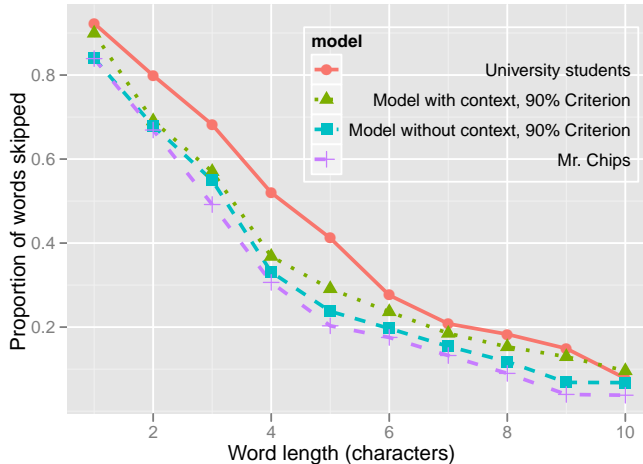


Figure 1: Proportion of words skipped by word length for each model. In all cases, the standard error of the mean for the Normal approximation to the Binomial distribution is smaller than the symbols. The human data is from Rayner and McConkie (1976) and has no standard errors.

ilar to that of human readers, i.e., relaxing the assumption that words need to be identified with 100% certainty, alters the performance of the model across two measures to look more like human performance. Such a result adds some support to the idea that the relevant human behavior is well understood as a rational response to the demands of the task. It is also worth pointing out that the resulting model maintains and uses uncertainty about previous input, something for which most models of sentence processing do not allow.

### Experiment 2

In Experiment 2, we test the effect of allowing the model to use the linguistic context as another source of information for word identification. Specifically, we compare the previous two models to one that includes a 90% confidence criterion as well as a simple bigram model of linguistic context.

#### Methods

The methods are the same as Experiment 1, except that the new model uses a bigram language model, again smoothed with Kneser-Ney under default parameters.

#### Results

The results for average saccade length for the new model is given in the third row of Table 1. As shown in the table, giving the model information about linguistic structure increases the

average size of saccades a bit more, bringing it still closer to the human estimate of 8 characters.

The results for word skipping rates for the new model is plotted as the second line in Figure 1. This new model gives an even larger increase in word skipping rates across all word lengths, on top of the increase seen previously, bringing it more in line with human results. Skipping rates are 30% closer to humans than the previous 90% criterion model on average, and for some word lengths are up to 75% closer.

## Discussion

The results of this experiment show a case in which making more of the information that is available to a human reader also available to a rational model causes its behavior to more closely approximate human performance. Together with the previous result, this supports the notion that some aspects of reading are well understood as a rational response to the structure of the problem.

## General Discussion

In this paper, we presented a new rational model of reading based on Mr. Chips, but which fixes two problems with it – it uses information about linguistic context in word identification and a flexible identification criterion. Experiment 1 showed that the model’s performance looks more like humans’ when the model’s goal is shifted to one more like that of humans, 90% confidence in each word. Experiment 2 showed the model’s behavior looks even more like humans’ when the model can use information that is used by humans: linguistic context. Taken together, these results suggest that many facets of human reading behavior can be well explained as resulting from a rational solution to the problem of reading. This model adds to the growing literature on rational process models, exploring the extent to which human performance can be viewed as rational agents across a wide variety of complex behaviors, such as multiple object tracking (Vul, Frank, Alvarez, & Tenenbaum, 2009) and online change detection (Brown & Steyvers, 2009).

It is the case, however, that many aspects of human reading behavior cannot in principle be explained by a model such as those described in this paper. This is because much of what we know about human reading behavior is about fixation durations, and these models have no notion of duration. They cannot have a notion of duration because visual input is veridical categorical information about a range of characters, so that there is no reason to stay at a given location for more than one timestep. Reichle and Laurent (2006) overcome this problem by making the simplifying assumption that required processing times on words are a function only of their length.

We believe, however, that the way forward for rational models of reading is to incorporate a model of noisy visual input, so that the model can make decisions about fixation durations to get more or less visual information. In other work (Bicknell & Levy, 2010), we explore the use of such models to answer questions that are impossible to ask with non-rational models of reading such as when and why should

between-word regressions be made, and how should reading behavior change as accuracy is valued more or less relative to speed.

## Acknowledgments

We are very grateful to Gordon Legge for sharing the corpus used in the original Mr. Chips experiments. This work also benefited greatly from useful discussion with Jeff Elman, Tom Griffiths, Andy Kehler, Keith Rayner, and Angela Yu, and feedback from the audience at the 84th Annual Meeting of the Linguistic Society of America. The research was partially supported by NIH Training Grant T32-DC000041 from the Center for Research in Language at UC San Diego to K.B., by a research grant from the UC San Diego Academic Senate to R.L., and by NSF grant 0953870 to R.L.

## References

- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- 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*. Uppsala, Sweden: Association for Computational Linguistics.
- Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, 58, 49–67.
- 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.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777–813.
- Legge, G. E., Hooen, T. A., Klitz, T. S., Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Research*, 42, 2219–2234.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: an Ideal-Observer model of reading. *Psychological Review*, 104, 524–553.
- 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*, 106, 21086–21090.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.
- McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14, 648–652.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K., & McConkie, G. W. (1976). What guides a reader’s eye movements? *Vision Research*, 16, 829–837.
- Reichle, E. D., & Laurent, P. A. (2006). Using reinforcement learning to understand the emergence of “intelligent” eye-movement behavior during reading. *Psychological Review*, 113, 390–408.
- 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., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, 7, 4–22.
- Vul, E., Frank, M., Alvarez, G., & Tenenbaum, J. (2009). Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22* (pp. 1955–1963).