The adaptive nature of eye-movements in linguistic tasks: How payoff and architecture shape speed-accuracy tradeoffs

Richard L. Lewis and Michael Shvartsman
Department of Psychology
University of Michigan

Satinder Singh
Computer Science and Engineering
University of Michigan

We explore the idea that eye-movement strategies in reading are precisely adapted to the joint constraints of task structure, task payoff, and processing architecture. We present a model of saccadic control that separates a parametric *control policy space* from a parametric *machine architecture*, the latter based on a small set of assumptions derived from research on eye movements in reading (Engbert, Nuthmann, Richter, & Kliegl, 2005; Reichle, Warren, & McConnell, 2009). The eye-control model is embedded in a decision architecture (a machine and policy space) that is capable of performing a simple linguistic task integrating information across saccades. Model predictions are derived by jointly optimizing the control of eye-movements and task decisions under payoffs that quantitatively express different desired speed-accuracy tradeoffs. The model yields distinct eye-movement predictions for the same task under different payoffs, including single-fixation durations, frequency effects, accuracy effects, and list position effects, and their modulation by task payoff. The predictions are compared to—and found to accord with—eye-movement data obtained from human participants performing the same task under the same payoffs, but are found not to accord as well when the assumptions concerning payoff optimization and processing architecture are varied. These results extend work on rational analysis of oculomotor control and adaptation of reading strategy (Bicknell & Levy, 2010b; Norris, 2009; Wotschack, 2009; McConkie, Rayner, & Wilson, 1973) by providing evidence for adaptation at low levels of saccadic control that is shaped by quantitatively varying task demands and the dynamics of processing architecture.

Introduction

We present a set of novel eye-tracking experiments and computational models that explore the idea that eye-movement strategies in reading are adaptations to the joint constraints of processing architecture, task structure, and task payoff. To our knowledge, these experiments provide the first evidence for the modulation of low-level saccadic control in a se-

quential reading task by the experimental manipulation of an external payoff—a payoff that quantitatively specifies a speed-accuracy tradeoff. The computational models provide the first systematic analysis of how observed eye-movements in reading might be determined by such task payoffs interacting with the constraints of the processing architecture, including oculomotor dynamics and representation noise. The key components of this analysis are a model with multiple control (strategy or policy) parameters that performs the complete task, and an exploration of how changes in both external payoff and processing architecture lead to changes in achieved payoff and optimal strategies, and thus changes in the predicted behaviors.

Task effects in psycholinguistics

Task goals and context have long been known to have major effects on human performance in psycholinguistic experiments (for an early analysis see

the seminal chapter by Forster, 1979). For example, in the area of single-word lexical processing, there are robust differences in how frequency and other important effects are manifest in *naming* vs. *lexical decision* tasks (e.g., Grainger, 1990). Task context in the form of experimental list composition and goal manipulation via instructional emphases have significant effects, and have received detailed theoretical treatments (Wagenmakers, Ratcliff, Gomez, & McKoon, 2008).

There is also a small but growing line of empirical work demonstrating task effects on eye-movements in reading. For example, McConkie and colleagues (McConkie et al., 1973) have shown that participants tend to read longer when anticipating more difficult questions (for example, questions of a factual nature), as well as when they were financially incentivized to answer the questions correctly. More recently, Rayner and Raney (Rayner & Raney, 1996) have shown that the lexical frequency effect is eliminated when subjects read words in search of a target word rather than reading for comprehension. Finally, Wotschack (2009) found that increasing the frequency and difficulty of comprehension questions, as well as instructing the participants to proofread, led to slower reading speeds. But while this prior work manipulates task type and difficulty, there has not been a manipulation of quantitative speed-accuracy tradeoffs of the kind we pursue here.

In addition, recent work on visual attention in both linguistic and nonlinguistic contexts indicates that attention strategies are strongly shaped by prevailing task goals (e.g., Rothkopf, Ballard, Hayhoe, & Regan, 2007; Ballard & Hayhoe, 2009; see Salverda, Brown, & Tanenhaus, 2011 for a recent review). In general, effects of strategic adaptation penetrate all levels of human performance (Newell, 1973), from the most elementary perceptual decisions (Tanner & Swets, 1954) to more complex multi-tasking scenarios (Meyer & Kieras, 1997; Howes, Lewis, & Vera, 2009).

One of our guiding hypotheses is therefore that eye-movement strategies in reading are shaped by task goals. Just as there are no fixed visual search strategies, neither are there fixed cognitive or eye-movement control strategies in reading. But we also assume that there are relatively fixed aspects of the cognitive and oculomotor architecture (Reichle, Rayner, & Pollatsek, 2003; Engbert et al., 2005)

that define the space of possible processing strategies and give shape to the payoff surfaces that map strategies to expected payoff.

The theoretical challenge: From task and payoff through architecture to behavior

We face a challenge in bridging the gap between high level task goals and payoff, and the lowest levels of moment-to-moment behavioral control that make contact with eye-movement measures. Meeting this challenge demands a theoretical approach that provides an analytic means to investigate the effects of both task goals and architectural constraints on behavior. The broad foundations of the necessary approach were provided by early signal detection theory (SDT) (Tanner & Swets, 1954): a formal model that specifies parameters of strategic adaptation (in SDT, perceptual decision thresholds); a specification of the fixed processing constraints on performance (in SDT, the noise parameter, estimated by d'); and quantitative feedback on task performance that is used in both human experiments and in derivations of optimal adaptation by the model (in SDT, the payoff matrix for hits, misses, correct rejections, and false alarms). Ideal observer models built on SDT (Green & Swets, 1966; Geisler, 1989), and extensions to the dynamics of optimal decisions based on Bayesian sampling (Wald & Wolfowitz, 1948; Stone, 1960; Edwards, 1965; Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006) significantly extend this formulation in various ways, but retain its basic form.

The approach: Bounded optimal control

The most general form of the approach specifies machines and parametric policies capable of performing sequential decision making tasks, and a means to derive policies that optimize some measure of obtained payoff/reward—the *reinforcement learning* formulation (Sutton & Barto, 1998; Kaelbling, Littman, & Moore, 1996; Singh, Jaakkola, Littman, & Szepesvari, 2000). Our version of the approach as it is applied to both the human and computational experiments reported here is shown in schematic form in Figure 1 and includes the following components: (1) A linguistic task environment, the *List Lexical Decision Task* (LLDT), which requires determining whether a horizontal array of six-letter strings are all words or not (we

describe the LLDT in more detail below). (2) A machine (agent architecture) that can control both perception (via oculomotor decisions that determine saccade timing) and task-environment actions (via trial-level decisions that lead to simulated button presses indicating the response). (3) Machine constraints that embody assumptions about processing architecture (oculomotor dynamics and representation noise). (4) A set of distinct quantitative payoff functions that provide feedback on task performance and impose different speed-accuracy trade-offs (three payoffs we label *Accuracy*-, *Balanced*- and *Speed*-emphasis). (5) Eye-tracking experiments using these different payoff functions in which human participants are given cash bonuses based on their performance. (6) Computational experiments in which distinct optimal control policies are derived for the constrained machine under the different payoff functions, and under versions of the machine that vary architectural components of theoretical interest.

This approach meets the bridging challenge because it provides a way for specific task goals (expressed as payoff functions) to interact with machine constraints (through optimization) to yield detailed behavior. It also has the methodological virtue of reducing theoretical degrees of freedom in explaining behavior (Howes et al., 2009) because strategic parameters are optimized for task payoff, not fit to data. We refer to the approach as "bounded optimal" control to emphasize the role that processing architecture plays in defining the optimization problem (Russell, Subramanian, & Parr, 1993), but there is no special sense of "optimal" intended; it is simply an application of optimal sequential control.

This work is in the growing tradition of rational analysis (Anderson, 1990) approaches to language processing (Hale, 2011) and eye-movement control (Legge, Klitz, & Tjan, 1997; Bicknell & Levy, 2010a, 2012), and draws substantially on Bayesian sequential sampling and diffusion models of lexical processing (Norris, 2006, 2009; Wagenmakers et al., 2008) and mathematical models of eye-movement control in reading (Engbert et al., 2005; Reichle et al., 2003). What distinguishes our present work is the analytic and empirical focus (and associated novel results) on understanding how task-specific payoff and processing architecture jointly shape eye-movement behavior. We now introduce our task paradigm in more detail which
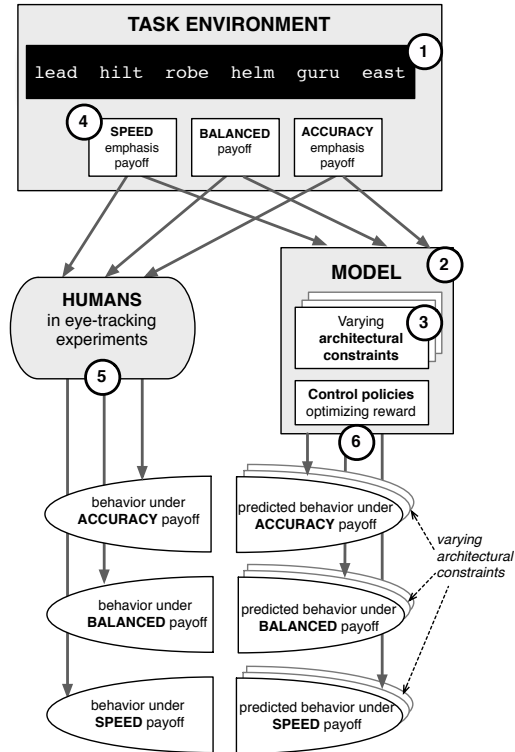

Figure 1. Overview of the approach. See the text for a detailed description keyed to the numbered elements.

will allow us to ground the main theoretical assumptions and model description.

The List Lexical Decision Task

The *List Lexical Decision Task* (LLDT) is a simple extension of a paradigm first introduced by Meyer and Schvaneveldt (1971). On each trial of the LLDT, participants are presented with a list of alphabetic character strings, and must make a single decision as to whether the list contains only words. The top of Figure 1 shows a typical trial. In the human and modeling experiments reported below, there are six strings in a horizontal array; each string is four letters long. There is at most one nonword per list and no words are repeated in the same list.

The LLDT is a simple task but has several desirable features for our purposes: (a) it is amenable to quantitative payoff manipulations and trial-by-trial

	Accuracy	Balanced	Speed
Incorrect penalty	-150	-50	-25
Speed bonus (per second under 5s)	8	6.7	5.7

Table 1

Quantitative payoffs given to both model and human participants. These payoff points translated into cash bonuses for the human participants.

feedback that differentially rewards speed and accuracy (discussed next); (b) it requires the control of serial visual attention; (c) it involves both perceptual control decisions and a separate trial-level decision, and thus poses a joint optimization problem over both sets of decisions; (d) it requires the application of (minimal) linguistic knowledge that can be approximated via corpus frequencies; (e) it is expected to lead to left-to-right reading and thus yield an eye-tracking record comparable to natural reading. Because all the words are four letters and selected independently, the LLDT is also expected to yield a high proportion of *single fixations* and clean estimates of frequency effects not confounded by length and predictability.

Three distinct payoffs

We evaluated both model and human participants according to three different payoff functions (specified in Table 1. The payoffs were designed to impose different speed-accuracy tradeoffs for a given level of success, and were all defined in terms of a bonus for speed and penalty for incorrect responses. The bonus was continuous at the millisecond level, starting at zero points for responses longer than 5s and rising by a different number of points per second for each payoff.

An optimal control model

Main theoretical assumptions

We can now state our three main theoretical assumptions:

1. Saccadic control is a "rise-to-threshold" system (Brodersen et al., 2008) conditioned on *task-specific decision variables* that reflect the accumulation and integration of noisy evidence over time. We model the evidence accumulation as Bayesian sequential sampling, and in our simple two-alternative task this is equivalent to a Sequential Probability Ratio Test (Wald & Wolfowitz, 1948).

2. The saccade thresholds are set to *maximize task-specific payoff*, but this is one part of a *joint optimization problem* that includes all other policy parameters that determine behavior in the task. In our model of the LLDT, this consists of a separate decision variable and threshold that determines the task-level response to the entire trial. These two thresholds together determine how long the model fixates on individual strings, how many strings it reads, and when and how it responds.

3. The shape of the payoff surface (and thus its maxima) over the multi-dimensional policy space is *determined jointly by the payoff function and properties of the perceptual and oculomotor system*, including saccade programming duration, eye-brain-lag, saccade execution duration, manual motor programming duration, and representational noise.

Overview

We provide a brief overview of a typical trial before focusing in on specific detail of each aspect of the model specification. See Figure 2 for a schematic diagram of the full model, and Figure 3 for simulated traces from two sample trials. On a given trial, the first fixation starts on the leftmost string. During each fixation, noisy information about the fixated string is acquired at every timestep, with some delay (the eye-brain-lag, VanRullen & Thorpe, 2001). This noisy information is used for updating the model's beliefs about the status of the current string as well as the trial as a whole. This means that information about the current word may affect the beliefs about other words, as a consequence of the constraint that each list has at most one nonword. (We later explore the implications of this in the model predictions and find support for these predictions in the human performance). The model receives no parafoveal input— i.e., it receives information from one word at a time. This is a reasonable approximation given the wide spacing of the strings in the human experiment (approximately 3.4 degrees of visual angle), and we found no empirical evidence for preview effects.

The sampling continues until either the string-level or the trial-level belief reaches some threshold, at which point either a saccade is initiated (if

Parameter	Mean	Std Deviation	Source
Eye-Brain Lag	50ms	15ms	(VanRullen & Thorpe, 2001)
Saccade programming time	125ms	37.5ms	E-Z Reader (Reichle et al., 2009)
Saccade execution time	40ms	12ms	Estimated from participants
Motor preparation and execution time	100ms	30ms	EPIC (Meyer & Kieras, 1997)
Trial onset detection and refixation	150ms	45ms	Prior estimate of short fixation and saccade
Sample duration	10ms	0	Nontheoretical discretization parameter
Gaussian sample noise	0	1.2	Standard deviation fit as described in text

Table 2

Model parameters; all are fixed in advance except sample noise, which is fit as described in the text. Means were taken from sources noted, with standard deviation fixed at 0.3 × mean.

	Search range	Final value	Source
Saccade Threshold	0.80, 0.85, 0.86, 0.87, 0.88, 0.89, 0.90, 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.995, 0.999, 0.9999, 0.99999	0.99 (Acc), 0.97 (Bal), 0.92 (Speed)	Maximizing payoff given task and architecture
Decision Threshold	0.80, 0.85, 0.90, 0.92, 0.94, 0.95, 0.96, 0.97, 0.98, 0.99, 0.999, 0.9999, 0.99999	0.999 (Acc), 0.999 (Bal), 0.99 (Speed)	Maximizing payoff given task and architecture

Table 3

Derived policy parameters. The saccade and decision thresholds were derived by evaluating the expected payoff achieved by the model with each combination of indicated policy values, and then choosing the pair of thresholds that maximize payoff; see Figure 4 for the derivation of the optimal policy.

the string-level threshold is reached), or a manual response is initiated (if the trial-level threshold is hit). We will refer to these thresholds as the *saccade threshold* and *decision threshold*. Information acquisition continues while the saccade or manual response is being programmed and until the saccade begins execution (with some visual persistence offset). Once saccade programming and execution is complete, the model fixates on the following string (if there are strings remaining), or initiates a response otherwise. Once motor programming and execution is complete the trial is over.

Oculomotor architecture and noise

The model's sequential perceptual inference mechanism is embedded in a simple oculomotor control machine, drawing upon current mathematical models of oculomotor control in reading. The delays noted above (eye-brain-lag, saccade programming

and execution times, and motor time) are drawn from gamma distributions, chosen for convenience because they are constrained to be positive and have been previously used in eye movement models (Reichle et al., 2009). For ease of interpretation, we will report the means and standard deviations of these parameters. They were converted to gamma distribution parameters by setting the shape parameter $k = \frac{\mu^2}{\sigma^2}$ and the scale parameter $\theta = \frac{\sigma^2}{\mu}$, where $\sigma = 0.3 \times \mu$. Means were taken from the existing literature (see Table 2), and mean saccade duration was estimated directly from our human participants.

Bayesian evidence integration and assumptions about prior beliefs (lexicon)

We assume that there is some noise in the perceptual information acquisition process and in the process of matching visual input to the lexicon. To overcome this noise, our model (and many ideal
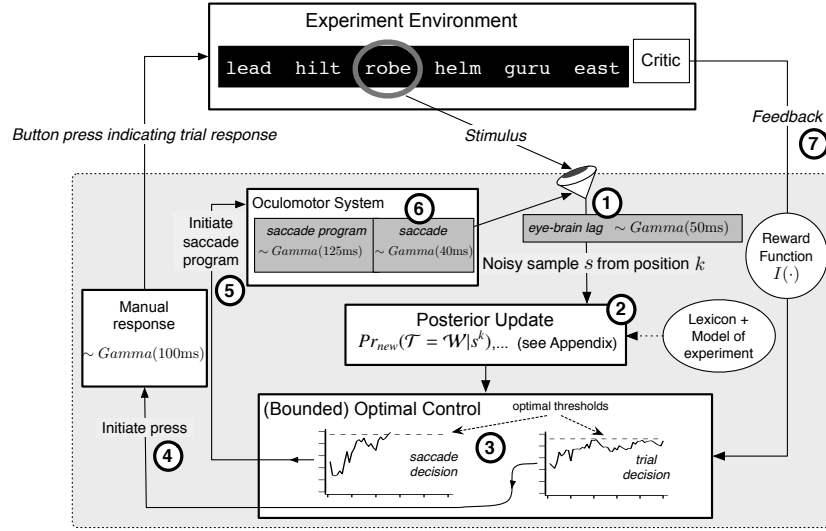
Figure 2.   Schematic diagram of the full model. Stimulus arrives from environment with some delay (eye-brain lag) (1), at which point a posterior update occurs (2), taking into account the new sample information plus prior beliefs (initialized from knowledge of the experiment and lexicon). Control conditioned on string-level and trial-level beliefs (3) may initiate manual (4) or oculomotor action (5) (button press or saccade). In the case of the latter, the eye moves to the next string with some delay (6) (saccade programming and execution). In the case of the former, response is made with some delay (motor programming and execution) and the model receives payoff feedback (7).
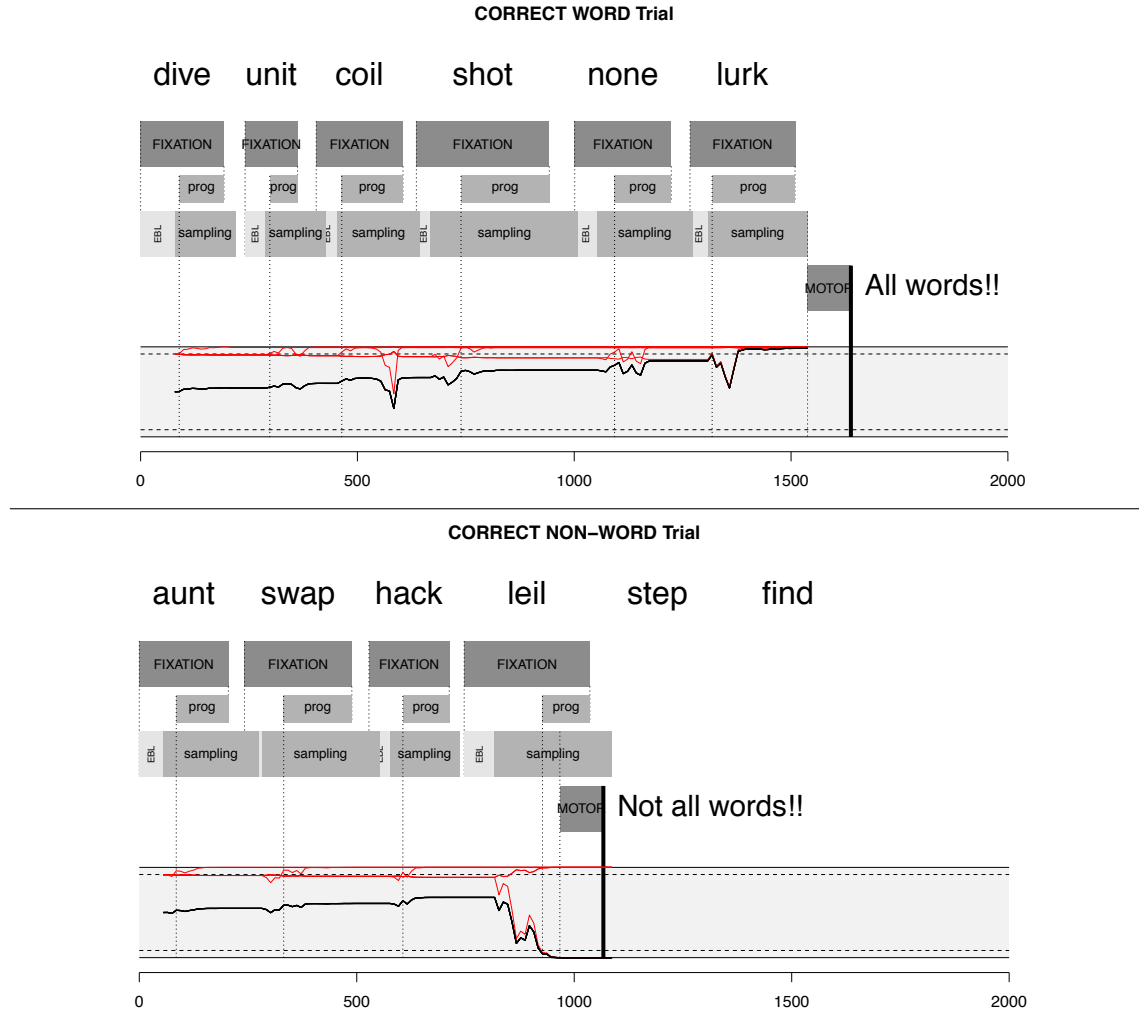
observer models) iteratively uses Bayes' update in combination with some prior belief to determine the probability distribution over the currently fixated string and the remaining strings in the trial.

The model maintains belief probabilities over the following items: (a) the probability distribution over all possible strings in the currently fixated position, (b) the probability of a nonword in each position, and (c) the probability that the current trial is a word trial (note that this is the complement of the sum over (b)). The prior over (a) is derived from corpus frequencies from the Brown Corpus (Kucera & Francis, 1967), the prior over (b) is the probability of a nonword trial divided by the number of positions, and the prior over (c) is the probability of a word trial. The simplifying assumption here is that the participants know all of the words in the experiment and thus would categorize all of the word strings correctly given sufficient time (we discuss below how this assumption may have contributed to a discrepancy between the accuracy achieved by the model and the humans).

Following Norris (2006), the model represents the string stimuli with a simple indicator vector cod-

ing. The true identity of each letter is represented as a vector of length 26 with a 1 in the position corresponding to this letter, and zeros elsewhere. Each string is represented as a 4×26 matrix with a row for each of the four letters in the string. This coding does not represent a deep theoretical commitment but is a convenient way to place strings in a representational space with plausible similarity relations. Samples are generated by adding mean-zero Gaussian noise to this representation; we discuss below how this noise parameter is set.

At each time step, upon receipt of a sample the model computes a multi-step Bayes update: first, it updates its belief of the probability distribution over strings in the current position; then, it uses this information to update its belief over the other positions and the trial. In doing so, it takes into account the fact that position-level nonword probabilities are not conditionally independent given the one-nonword restriction. However, in order to greatly decrease the computational cost of each update, the model allows for non-zero belief probabilities over lists with repeated word-strings. The larger lexicons that we can explore as a result are substantial

CORRECT WORD Trial

dive   unit   coil   shot   none   lurk

CORRECT NON-WORD Trial

aunt   swap   hack   leil   step   find

Figure 3.   Simulated model traces for a correct word trial and a correct nonword trial. At the top are the words in each trial. The filled rectangles show the timing and duration of fixation durations, saccade programming (prog), eye-brain-lag (EBL), sampling, and motor response preparation and execution. At the bottom is the random walk of the belief probabilities, with the bottom representing 0 and the top 1. The black line is the trial-level belief (and so starts at 0.5), and the red lines are the string-level beliefs (and so start at 0.82, the prior probability that a given string is a word). The solid horizontal black lines are the decision thresholds (the top one for a word-trial and the bottom one for a nonword-trial), the dashed horizontal lines are saccade thresholds.

enough that the probability of repeated words is extremely low and so we think that this simplification in the model is justified. The full mathematical detail of the update is included in the Appendix.

The model was tested with 50 different word and nonword lexicons of approximately 500 strings

each. The word lexicons always included the experimental words and an additional set of words drawn uniformly randomly from the set of 1,500 English four letter words represented in Kucera and Francis (1967); the nonword lexicons always included the experimental nonwords and an additional set of

nonwords in which letter bigrams were attested in the English word list. The model's performance is always evaluated on the words and nonwords from the human experiments, but for the model these strings are not distinguished in any way from the rest of the model's lexicon. Aggregating results across different model lexicons ensures that the results are not driven by a particular lexicon choice (though our experience with the modeling indicates the results are robust against this choice). Nevertheless, granting the model veridical knowledge of both words and nonwords in this way is a gross simplification of human subject knowledge; we consider some potential consequences of this below.

Trial decision and saccade thresholds

What remains is the process of conditioning control on the belief state detailed above. The control problem is essentially two nested optimal stopping (Wald & Wolfowitz, 1948; Stone, 1960; Edwards, 1961) problems: when to stop sampling from each string and move on to the next word, and when and how to respond and thus end the trial. The model uses two thresholds to implement these control decisions. The saccadic control is determined by a threshold defined over the evolving probability of a nonword (or word) in the current position, and the manual button (trial-level response) control is determined by a threshold over the evolving probability that the trial is a word or nonword trial. The model thus yields biased random walks (e.g., Ratcliff, 1978), in which the belief probabilities make a noisy rise or fall towards one of two symmetrical thresholds (Figure 3); crossing a threshold then triggers the corresponding action. The two thresholds (over saccade and manual decisions) embody the fundamental speed-accuracy tradeoff in the model. As the thresholds are set higher, the probability of making an error falls but the time to decision increases; as the thresholds fall, the opposite happens.

This dual-threshold policy space is a subset of the full space of possible policies, which consists of all actions available to the model (saccade, wait, respond yes, respond no) conditioned on all possible belief states. It is possible that better policies lie outside of the space we explored. As such, the policy space simplification may be taken as a theoretical commitment to a kind of computational bound on control. Although this may be a plausible assumption, future work must provide support for it

by comparing it explicitly to alternative models. For present purposes, the 2-dimensional space has the virtues of simplicity and computational tractability.

Understanding how the model makes predictions

Recall the fundamental theoretical challenge we identified in the Introduction: find a way to link high level task goals and payoff, through processing architecture, to the lowest levels of moment-to-moment behavioral control. We now describe how we can use the model to accomplish this, by making predictions on the List Lexical Decision Task under the three different payoff schemes. The methodology is to explore differences in the payoff surfaces, especially focusing on optimal and near-optimal policies, and their implications for behavior. More specifically, in what follows we (1) examine the relationship between policy and expected payoff; (2) examine the relationship between policy and behavior; and (3) examine the relationship between behavior and payoff. It is important to understand that these are not three separate computational steps in the modeling process but rather different ways of viewing the model's implications.

Fitting the noise parameter. The Gaussian noise added to the sample vectors is the one parameter that is not fixed in advance. It functions as a kind of scaling parameter in that increasing noise requires increasing the number of samples to obtain a given level of accuracy. We fit this parameter to the human data by computing optimal policies across a range of noise values and choosing the value that minimizes a simple error measure: root mean squared deviation from mean single fixation duration (SFDs, fixation durations on strings fixated only once) for the three payoff conditions. In this sense, the fixation durations act as our training set, and the model's remaining measures are the test set. Other choices, for example fitting SFD to only one of the conditions, makes little difference. Figure 9 (red curve) shows the resulting model error (computed here as root mean squared error) across a range of noise levels in the neighborhood of the minimum. Figure 9 also shows model errors for three other architectural variants discussed below. For the architecture described above and illustrated in Figures 2 and 3, the noise value providing the

best fit is 1.2 and this value is held constant across all predictions for all three payoffs.

The relationship between policy and payoff

Figure 4 provides views of the payoff surface in the two-dimensional policy space. In the first three panels, payoff is plotted against saccade threshold and each separate line corresponds to a separate decision threshold. In the fourth panel, payoff is plotted against decision threshold and each line corresponds to a separate saccade threshold—thus these are different views of the same 2-D payoff surface. Recall that a policy is simply a pair of threshold values (*saccade-threshold, decision-threshold*). The circled point at the top of each payoff plot represents the policy that yields the maximum expected payoff under this formulation of the policy space, and its value is given as the pair of numbers to the left of the point. The colored points represent policy points that are within 0.2 payoff units of the optimal. Values of each point are computed from means of 300K Monte Carlo trials; see the Appendix for details.

Consider the Accuracy payoff graph at the left. This graph indicates that there is a flat region of the payoff surface when saccade thresholds are below about 0.85; this corresponds to thresholds where the saccade program is initiated almost immediately upon fixation. There is a steep rise in payoff as saccade thresholds increase—because more samples are obtained and accuracies are increasing significantly—up to a maximum point near a threshold of 0.99, followed by a steep decline as the additional gain from increased accuracy diminishes and the time cost begins to dominate. This relationship holds for most of the good performing decision thresholds. The relationship between decision threshold and payoff has a similar but simpler profile over the range we explored: a steady increase in payoff as the decision threshold increases, followed by a steep drop as the time cost begins to dominate.

The Balanced and Speed payoffs have a similar profile as the Accuracy payoff—for the saccade thresholds, a flat region, a rise, and a sharp drop. But the qualitative shape differs considerably in the region of the maximum; payoff surface is considerably flatter for the Balance payoff, and very flat for Speed. Thus there is considerably more spread in the range of thresholds that perform within some close threshold of the optimum. The visualization of the payoff space suggests that we were more successful in separating the Accuracy condition from the other two than we were in separating Balanced and Speed from each other.

The graphs in Figure 4 provide a visualization of the fundamental basis of the model's link between task payoff and predicted behavior: they depict the nature of the adaptive space that the human participants must navigate according to the model. The overall differences in payoff levels are not important; what is important are differences in the optimal and near-optimal policies, and in the shape of the payoff surface. The next step is to relate these same policy points to predicted behavior directly.

The relationship between behavioral measures and payoff

Figure 5 plots SFD and overall trial response time against saccade threshold (we focus on these two measures for illustration; many other measures are obtained as shown in Figure 6, 7 and 8). The colored points correspond to the best policies identified in Figure 4. The relationships are clear: increasing saccade threshold increases single fixation duration, though it has almost no effect until thresholds are above 0.85, and past 0.99 the fixation times increase dramatically. The decision threshold lines are indistinguishable because task decision thresholds have no direct effect on single fixations. The middle graph shows that trial response times also increase as both saccade and decision thresholds increase. Finally, the right graph shows that increasing saccade threshold increases overall accuracy.

Finally, the third panel in Figure 5 visualizes the payoff surface directly in terms of one of the behavioral measures (SFD); here each line corresponds to a different payoff. In this space it is easy to see the speed-accuracy tradeoff in the upside-down U-shaped curves. On the left side of the curve, spending more time increases payoff because accuracy is increasing; on the right side, spending more time decreases payoff because the cost of the increased time outweighs the diminishing accuracy gains.

Model predictions and human experiment results

In what follows we describe the human eye-tracking experiments and walk through both the model predictions and human results at the trial-level and
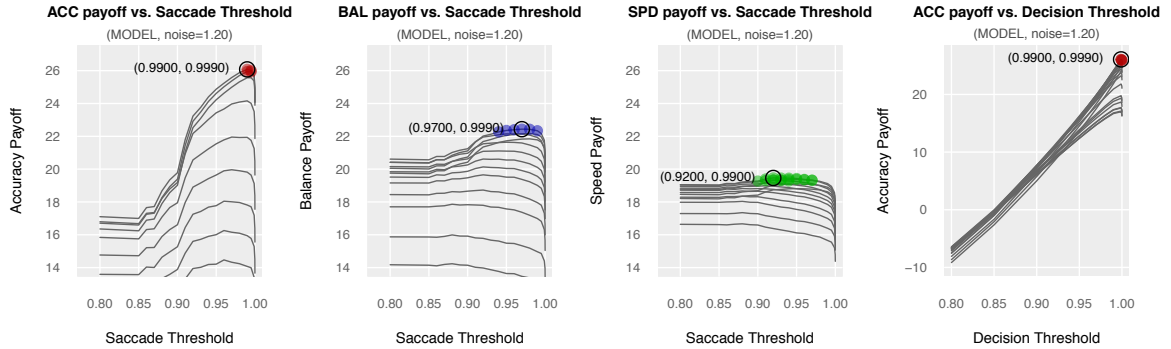
Figure 4. Expected payoffs generated by the model over the 2-D policy space defined by decision and saccade thresholds. The first three panels relate expected Accuracy, Balanced, and Speed payoff to saccade thresholds; each separate line corresponds to a separate decision threshold. The fourth (rightmost) panels relates the expected Accuracy payoff to decision thresholds; each separate line corresponds to a separate saccade threshold. (The Balanced and Speed payoffs, not shown, have rising but shallower slopes). The circled points represent the optimal policies, whose value is indicated at the left of the point. The colored points (*red* for Accuracy, *blue* for Balanced, and *green* for Speed) represent policies that are within 0.2 expected payoff units of the optimal point; thus the spread of these points in the Balanced and Speed payoffs reflects the flatter surfaces for those payoffs. Expected payoff values were computed over 300K simulated trials.
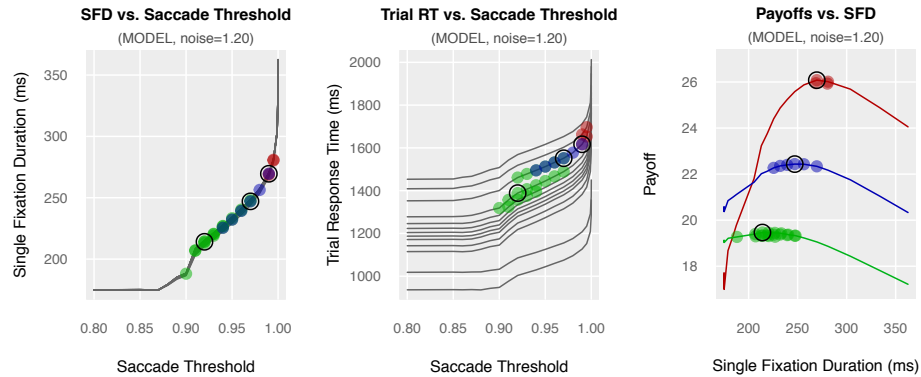
Figure 5. Examples of how thresholds determine behavior, and in turn how behavior and payoff are related. The first two panels show single fixation duration (SFD) and trial RT as a function of saccade threshold; the third panel shows expected payoff as a function of SFD. The circled points represent the optimal polices and the colored points (*red* for Accuracy, *blue* for Balanced, and *green* for Speed) represent policies within 0.2 payoff units of the optimal. See Figure 4 for the identification of these points.

string (word) level. Throughout it is important to understand the constrained nature of the model predictions. The threshold settings determine how the model behaves—the responses it makes, when it makes them, and how long it spends fixating on each string—and thus determine predictions for word and nonword trials, correct and incorrect trials, word and nonword fixation times, position ef-fects, accuracies, and so on. We report below multiple measures for each of the three conditions; with the single noise parameter fit to minimize error on single fixations durations as described above and held constant across the three payoff conditions.

Eyetracking experiment methods

Participants   Sixty-one members of the University of Michigan community participated in the experiments. Data from thirteen were unusable due to calibration problems, failure to complete the experiment, or equipment malfunctions, leaving a total of 48. Participants were given a baseline of \$10 for participation, plus a bonus of \$1 for each 1000 points they earned in the task.

Stimuli   Participants responded to 200 trials of the LLDT divided into 10 blocks, preceded by a 10-item practice block. There were two types of trials in each block: half of trials contained all-words lists, and the other half contained 5 words and one non-word in a randomly drawn position. Words were all 4 characters long and drawn from a 234-word subset of the Brown Corpus (Kucera & Francis, 1967), containing 117 high-frequency words (mean frequency count 239.2, SD 186.0) and 117 low frequency words (mean frequency count 5.6, SD 12.8). Nonwords were also all 4 characters long, and were drawn from a list of 53 nonwords pronounceable according to English phonotactics. While this means that participants saw each string more than once, the number of times a string was seen had no effect on fixation durations (effect of -0.64 ms per time seen, p=0.61).

Procedure   Each participant was assigned to one of the three payoff conditions used to make predictions from the model. They were not told the name of their payoff, only a quantitative description of the requirements (e.g. "You will receive a point for each 125 milliseconds by which your response is faster than 5000 ms (5 s). You will lose 150 points if your response is incorrect. You will get a \$1 bonus for each 1000 points.").

Items were presented on a CRT monitor in a 20pt Courier font, separated by 8 characters of whitespace. This resulted in each word covering 0.7 inches or 1.6 degrees of visual angle, and whitespace covering 1.48 inches or 3.4 degrees of visual angle at a distance of 25 inches from the screen. Each trial started with a fixation dot at the location of the first string. The entire six-string list would appear once subjects fixated, and the trial ended after subjects responded using a Cedrus response box. Eye movements were measured using an SR-Research Eyelink II head-mounted eye-tracker operating at 500Hz. Single-point drift correction was performed before every trial.

Statistical methods are given in the Appendix.

Trial level effects

Figure 6 shows the results for key trial-level measures: response times (on correct and incorrect trials), and percentage correct. The top row is the set of human results, and the bottom row is the set of model results from a set of policies at or near optimal (within 0.2 payoff units) plotted. Table shows condition means for human participants and model.

The key empirical result here is that the human participants show marginally decreased accuracies and response times across the Accuracy-Balanced-Speed payoff conditions: the accuracy condition results in slower response times (Table 6, contrast set (a)) and higher percentage correct (Table 6, contrast set (b)). The model predicts this trend because the optimal thresholds for Accuracy are higher (Table 3) than Speed, leading to slower but more accurate responses. There is a significant discrepancy in predicted accuracies that we address below. The model also correctly predicts that correct word trials will show slower responses than incorrect trials, with the converse holding in nonword trials (a reliable cross-over interaction, Table 6, contrast set (c)). This result is a consequence of the fact that 'all words' responses (correct or incorrect) tend to come after reading the full list, whereas 'not all words' responses tend to come after only reading a subset of words. Although this is not surprising behavior for the humans, the model need not have behaved this way: there are suboptimal strategies in the space we explored that set the decision threshold low enough that an 'all words' response are made before all strings are read, and ones that set it high enough that 'not all words' responses are not made until after the sixth string.

String level effects: Payoff, frequency, string type, and list position

Figure 7 shows the results of key string-level measures: single fixation duration across fixation types and by frequency class. We report single fixation durations both for brevity and because this is the measure that our model, currently lacking regressions, is able to quantitatively predict. 71% of the

Condition	Response Time (ms)		% Correct		Payoff	
	Human	Model	Human	Model	Human	Model
Accuracy	1667	1644	92	98	12.91	23.44
Balanced	1548	1546	87	97	16.08	21.52
Speed	1494	1455	88	95	16.65	18.89

Table 4
Trial-level measures; payoff reported is mean payoff per trial. See Table 6 for significance tests.
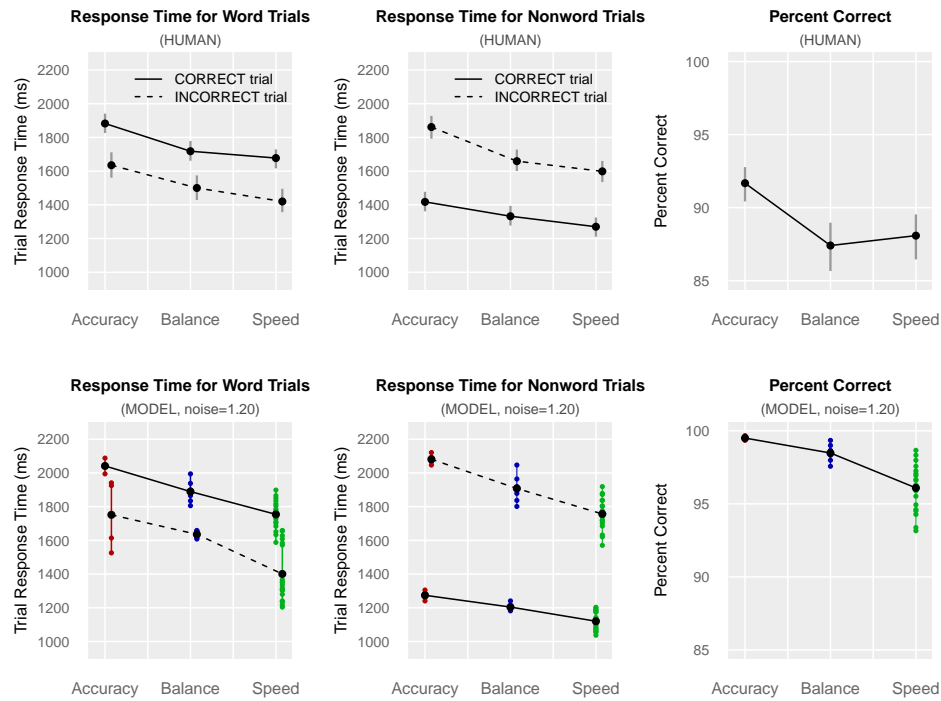

Figure 6. Empirical measures at the level of the trial for the full set of human participants and the computational model. The colored points represent predictions corresponding to the best-performing policies identified in Figure 4, the lines connect the means of this set. The error bars on the human data correspond to one standard error estimated from posterior densities of the mixed effects models.

strings were fixated only once, so other measures (e.g. first fixation and total fixation times) show the same patterns.

The key result here is that the human data is consonant with the model prediction of slower fixation durations in the Accuracy as compared to the Speed condition (Table 6, contrast set (d)). The model provides a straightforward explanation of this effect: when a payoff provides pressure to respond more

correctly, a higher saccade threshold will increase the amount of information acquired, increasing the likelihood of a correct response. Note that while the differences in trial response times is consistent with a difference in fixation durations in the model, this need not have been the case in the human data: the humans could have adapted via other strategies such as keeping their hands closer to the keys or minimizing mind wandering. Rather than (or in addition to)
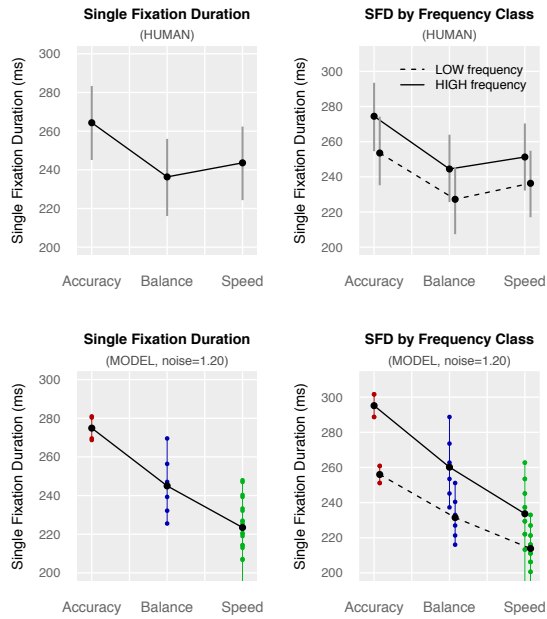
Figure 7. Single fixation durations and frequency effects for human participants and model. The colored points represent predictions corresponding to the best-performing policies identified in Figure 4, the lines connect the means of this set. We show the frequency effect here as means of low and high frequency bins (median-split) but all statistical models used continuous predictors. The error bars on the human data correspond to one standard error estimated from posterior densities of the mixed effects models.

these other adaptations, humans show evidence of adapting their moment-to-moment saccade timing to the differential speed-accuracy pressures.

The model also correctly predicts an effect of log frequency on fixation durations: fixations on highly frequent words are faster (Table 6, contrast set (d)). Norris (2006) has shown that this is a consequence of an otherwise unconstrained ideal observer model making decisions on single words, but it was not a necessary outcome with the additional oculomotor constraints we introduced. The effect can disappear with sufficiently low saccade thresholds, even though the model can still perform the task. Indeed, in the Speed condition, many good-performing thresholds nearly make the frequency effect disappear. The reason is that thresholds can

be set that are near the prior belief in expectation, so a saccade program is initiated immediately, before samples are taken. There is therefore little or no opportunity for frequency to affect fixation duration. But samples continue to arrive during the saccade programming time so the task can be performed. The model also predicts a larger frequency effect in the Accuracy condition as compared to the Speed condition, for the same reason: thresholds are set lower in the Speed condition and fewer samples are obtained during the pre-saccade-programming stage of sampling that affects fixation durations and thus frequency effects. The human data are consistent with this effect, but the interaction is not reliable.

The model makes other interesting and more subtle predictions attested in the human data (Figure 8, left two columns). First, nonwords are read more slowly than words. In the model, this is a consequence of the fact that the prior probability that any given string is a word is much higher than a nonword. It therefore takes more evidence (more sampling time) to reach the nonword threshold. Furthermore, the word-nonword difference is predicted to be larger in the Accuracy condition than Balance and Speed, an effect that appears in the human data as a reliable interaction between condition and string type. This reflects a nonlinear effect of distance-to-threshold on the number of samples required in expectation.

In addition, the effect of trial accuracy is different for words and nonwords (another interaction): fixation durations on words are about the same in correct and incorrect trials, but the model predicts that nonwords are read more *quickly* in incorrect trials, and the human data shows this pattern (Table 6, set (g)). In the model this arises because the word threshold is closer than the nonword threshold, and so incorrect random walks to the word threshold happen more quickly than correct random walks to the nonword threshold.

Finally, the model predicts that strings in later positions are read somewhat more slowly than strings in early positions for all three payoff conditions; this effect, though tiny, is also reliable in the human data (Table 6, contrast set (e)). This is another consequence of the list-level Bayesian update. The reason is a somewhat counterintuitive property of the probabilistic structure of the task: as evidence is accumulated identifying strings as words in the list, the probability of an all-words list increases—
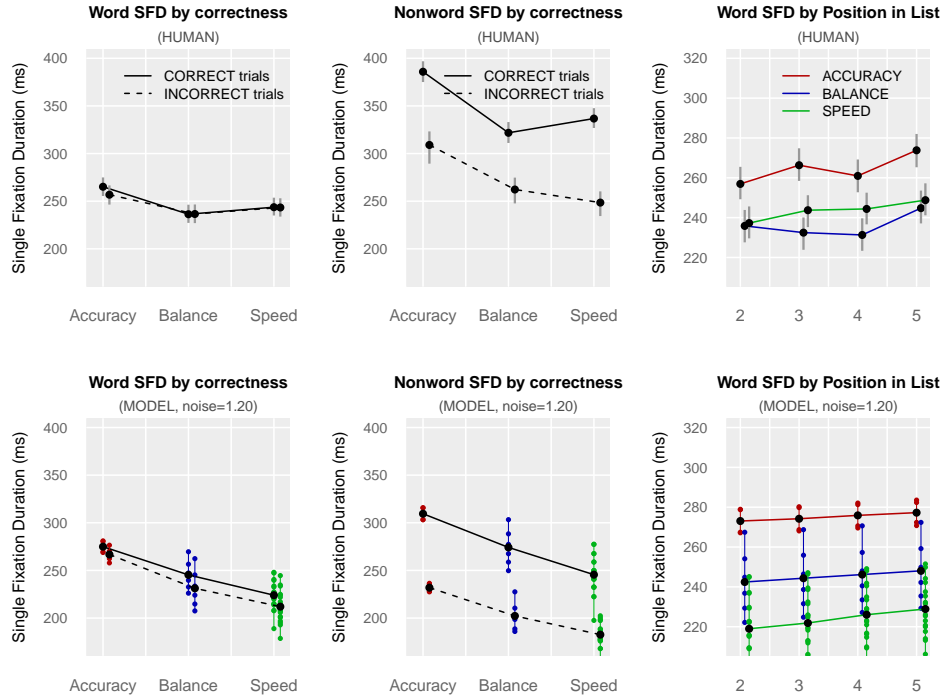
Figure 8. Single fixation durations for word and nonword, by correctness (left two columns), and single fixation durations for words by position in the list (rightmost column), for the full set of human participants and the computational model. The colored points represent predictions corresponding to the best-performing policies identified in Figure 4, the lines connect the means of this set. The error bars on the human data correspond to one standard error estimated from posterior densities of the mixed effects models.

but the probability that any one of the individual strings remaining is a nonword *increases* slightly. Thus, the prior belief that each remaining string is a word is slightly lower, and in expectation additional samples are required to hit the word threshold.

The accuracy discrepancy

The key discrepancy between model predictions and human performance is that the model achieves higher accuracy than the human participants (and consequently achieves a higher total payoff, especially in the accuracy condition). It is important to note that this discrepancy cannot be straightforwardly addressed by simply increasing the noise parameter—because the model would adjust its thresholds to maintain higher accuracy levels (and in doing so increase fixation durations and response

		Single Fixation Duration	
Condition	String Type	Human	Model
Accuracy	Nonword	377	331
	Word	264	275
Balanced	Nonword	308	294
	Word	236	242
Speed	Nonword	318	257
	Word	244	215

Table 5
Single Fixation Durations for word and nonword strings. See Table 6 for significance tests.

times). Thus there is a genuine discrepancy that cannot be explained by the present model.

Contrast set	Effect	Estimate	p
(a)	Condition on RT	-180.36	0.09
(b)	Condition on % Correct (logit)	-0.40	0.08
(c)	Trial Type (word vs nonword) on RT	**-355**	**<0.001**
	Correctness (correct vs incorrect) on RT	**-91**	**<0.001**
	Correctness x trialtype interaction on RT	**-596**	**<0.001**
(d)	Condition on SFD	**-21**	**0.04**
	Frequency on SFD	**-4.45**	**<0.001**
	Frequency x Condition interaction on SFD	1.64	0.29
(e)	Position on SFD	**3.26**	**0.007**
	Position x Condition interaction on SFD	-0.65	0.44
(f)	String (word vs nonword) type on SFD	**97**	**<0.001**
	String type x Condition interaction on SFD	**-61**	**0.05**
(g)	String (word vs nonword) x Correctness (correct vs incorrect) on SFD	**71.2**	**<0.001**

Table 6
Coefficient estimates and p-values calculated using a likelihood ratio test between two linear models identical except for the presence of the tested predictor. Lines separate different linear models. Condition was coded as a set of orthogonal contrasts; reported here is the Accuracy vs. Speed contrast.

We suggest two possible reasons for this discrepancy, both of which may be explored with additional modeling. Human participants are given a baseline of $10 at the end of their participation in addition to their performance-based cash bonus. If participants are maximizing reward *rate*, this $10 may be more valuable in 30 minutes than the $10 plus bonus is in 40 or 45 minutes. Similar additional speed pressures may result from temporal discounting.

Second, the model has veridical knowledge of which strings are classified as words and which are nonwords. But some of the very low-frequency words on the experiment list (e.g. *bard* and *nigh*) may simply not be in the participant's lexicons, and no amount of additional sampling will overcome such errors. These therefore represent a class of incorrect responses which are not possible for the model to make. Future work can address this by explicitly testing the full list of strings without time pressure post-experiment, and exclude trials with words not known by the participant.

How does architecture shape adaptation?

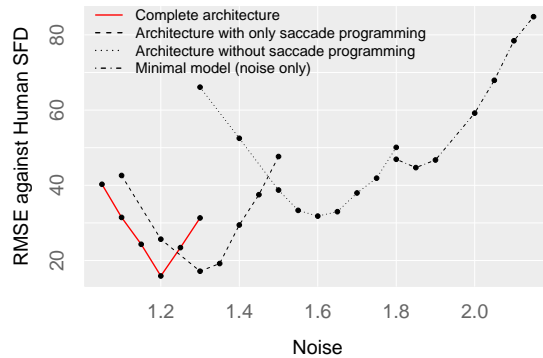We have shown that the present model, by optimizing policies for the different payoffs, and thereby

Figure 9. Root mean squared error of model predictions (against mean single fixation duration for the three payoff conditions) for four architectural variants. For each architecture, new optimal control policies are derived. In red is the complete architecture explored above and includes saccade programming, eye-brain-lag, and saccade execution. The minimal model dispenses with these oculomotor constraints. The other two models explore the effect of including excluding the saccade programming.
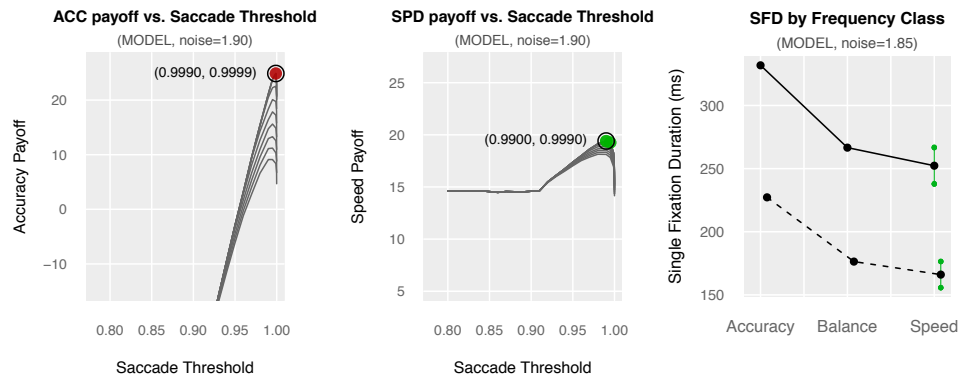
Figure 10. The first two panels on show the payoff surface and optimal policies (for Accuracy and Speed) for the *minimal model* that results from eliminating the oculomotor constraints; compare to Figure 4. The rightmost panel shows one example of the resulting behavioral predictions. The predicted magnitude of both payoff effects and frequency effects increase, yielding poorer fit to human data. (See Figure 9 for source of the noise parameter.)

its behavior, predicts a detailed empirical signature of adaptation across the payoffs that is largely attested in the human subjects, at the level of single fixation duration. We now make an initial attempt to address the following questions: Does the fixed architecture—the dynamics of the oculomotor system—shape adaptive behavior, and is there supporting evidence for this in the human data?

We begin the exploration by introducing a variant of the model that contains no ocumolotor dynamics; we label this the *minimal model* because it dispenses with eye-brain lag, saccade programming time, and saccade execution times. With this minimal model we find new best-fitting noise parameters and new optimal policies (Figure 9); the exploration is thus not simply a "lesioning" of the standard model.

Space precludes a full analysis but we focus here on some key outcomes of eliminating the architectural constraints. Figure 10 reveals a striking difference in the nature of the payoff surface for the minimal model: the peaks are sharper and the optimal thresholds are higher. There are qualitative changes to the behavioral predictions as well, including a larger payoff effect and larger frequency effect (right panel).

Our hypothesis was that the poorer fit is due primarily to the elimination of saccade programming time: as saccade programming time increases, there is greater pressure to reduce the saccade threshold, leading to shorter sampling durations during the initial stage in which the samples can affect fixation

durations. To test this hypothesis we explored two additional architectural variants: one which maintained only the saccade programming time (thus eliminating eye-brain lag and saccade execution), and one which kept eye-brain lag and saccade execution and eliminated saccade programming. Our hypothesis was partially supported: the saccade-programming-only architecture accounted for the human data nearly as well as the full architecture, with the exception that the log frequency effects were still significantly larger in the model than in the human data. The fit to human data provided by the architecture with the eye-brain-lag and saccade execution was between the extremes (Figure 9).

Comparison of LLDT to sentence reading

The LLDT was designed as a simple testbed for the empirical and theoretical exploration of optimal control models that will be extended to reading comprehension. But what is the relation between LLDT and reading sentences? We had two principal design goals to establish some empirical and theoretical similarities: (i) empirically, the task was designed to yield a sequential left-to-right fixation scanning pattern as in reading for comprehension, but with a high proportion of single-fixations; and (ii) theoretically, the task was designed to require the integration of information obtained from multiple lexical items across saccades.

The first goal is empirical, and given the eye-tracking record on the three payoff variants, the LLDT meets this goal: the scanning pattern was left-to-right; 71.2% of strings in our analysis were fixated once; and there were regressions originating in 7.5% of strings in the analysis, consistent with recent reading experiments[1].

The second goal is primarily a theoretical one, but with empirical implications. The goal is satisfied because LLDT does require integrating information obtained from multiple strings—for example, an alternative sequential lexical decision task, where a yes/no response must be made on each word, does not satisfy this requirement. Although the information integration required is minimal, the sequential nature of the left-to-right evidence accumulation and the nature of the probabilistic constraints in the task did give rise to contextual constraints that differ across positions of the list. This led to the prediction of a small increase in fixation durations across the list, attested in the human data.

We believe the LLDT is a useful first step in developing quantitative models of task effects in reading because it permits us to explore novel issues concerning optimal adaptation to the joint constraints of task payoff and architecture, but in a sequential eye-movement setting where the task and prior linguistic knowledge sources are clearly defined. And because the task structure and prior knowledge sources have a principled and integral role in the model and its predictions, there are paths to pursuing the integration of high-level linguistic processing and eye-movement control that are not available to approaches not grounded in a rational control analysis. One clear possibility is to adopt generative language models from computational linguistics to take the place of the generative model of the simple list structure used here. Indeed Bicknell & Levy, 2012, 2010a are pursuing such a path, and such models might be combined with the nested eye and task-control architecture explored here to produce optimal control models of reading that are tuned to specific tasks.

Summary of major results and broader implications

Let us take stock by revisiting the major theoretical challenge and key hypotheses, and summarizing the major results. The intent was to explore, computationally and empirically, the idea that eye-movement control in linguistic tasks is shaped jointly by specific task context, payoff, and fixed processing architecture—an idea motivated in part by a considerable body of work in psycholinguistics and visual attention (Forster, 1979; Wagenmakers et al., 2008; Rayner & Raney, 1996; Wotschack, 2009; McConkie et al., 1973; Reichle, Reineberg, & Schooler, 2010; Rothkopf et al., 2007; Ballard & Hayhoe, 2009; Salverda et al., 2011). The challenge was providing a way to bridge the gap between high level task goals/payoff, and the lowest levels of moment-to-moment behavioral control that make contact with eye-movement measures. We believe the approach pursued here meets the bridging challenge because it provides a way to explore how specific task goals (expressed as payoff functions) interact with machine constraints (through optimization) to yield detailed behavior. This approach then motivates the associated empirical methodology of running the same task under distinct quantitative payoffs (Tanner & Swets, 1954).

We introduced a simple linguistic task, the List Lexical Decision Task, that requires the application of lexical knowledge and the integration of information obtained from a sequence of left-to-right eye-movements. We presented an optimal control model of this task that embodies three key theoretical hypotheses: (1) saccadic control is conditioned on task-specific decision variables that reflect the accumulation and integration of noisy evidence over time; (2) these saccade thresholds are set to maximize task-specific payoff, but this optimization is one part of a joint optimization problem that includes all other policy parameters that determine behavior in the task; (3) the shape of the payoff surface (and thus its maxima) is determined jointly by the payoff function and properties of the perceptual and oculomotor system. The model permitted the exploration of these hypotheses computationally by optimizing control policies for three distinct payoffs imposing different speed-accuracy tradeoffs. Once a single noise parameter is fit, the model executing these optimal policies makes decisions about when to move the eyes, and how and when to respond

---

[1] E.g. Engbert et al., 2005 shows regression out probability between 6% and 1%; Reichle et al., 2009 shows regression out probability between about 8 and 11%, and Levy, Bicknell, Slattery, & Rayner, 2009 shows regression out probability between about 10 and 20%.

to the overall trial, thus generating dozens of behavioral measures, including fixation durations, for each of the distinct payoff conditions.

We ran eye-tracking experiments on the LLDT in a between-subject payoff manipulation. The detailed empirical predictions of the model were largely supported in the human results, including the key result that single fixation durations were modulated by payoff in ways predicted by the optimal control model. This provides some support for the first two theoretical hypotheses. We explored the third theoretical hypothesis—that the eye-movement behaviors are adapted not only to the task payoff but are shaped by the specific architectural constraints of the oculomotor system—by varying the architectural assumptions in the computational model and re-deriving optimal policies for these varied architectures. The optimal policies do clearly differ in the model when these constraints are changed. The modified architectures without saccade programming time provide poorer fits to the human data: in particular, they overestimate the size of the payoff effects and frequency effects.

The data from these experiments cannot provide unambiguous support for the effects of architecture on optimal policy and thus behavior because there is no experimental manipulation of the human architecture. But it is clear that, given the assumptions of this model, the human data are better accounted for by assuming that the control policies have indeed adapted to the saccade programming delay.

There are many interesting directions that may now be pursued. One especially promising and novel direction is building models that explain individual differences in reading strategies as bounded-optimal adaptations to individually-varying architectural constraints. This approach was demonstrated by Howes et al. (2009) in their individual differences models of PRP (Psychological Refractory Period) tasks, where differences in dual-task costs were explained as signatures of near-optimal adaptations to low-level processing characteristics (stage durations and motor noise), not as differences in an underlying dual-tasking capacity. It is also possible to explore effects of differences in internally modulated speed-accuracy tradeoffs—essentially differences in *intrinsic reward*. For example, the modulation of frequency effect by payoff suggests a way to understand the increase in lexical frequency effect on reading times in older adults

(Laubrock, Kliegl, & Engbert, 2006). Our results show a similar increase in younger adults under a payoff which emphasizes accuracy; it is known that older adults' tend to emphasize accuracy more so than young adults (Rabbitt, 1979; Smith & Brewer, 1995; Starns & Ratcliff, 2010).

Other useful directions include scaling the model to more complex sentence-level tasks as indicated above, as well as incorporating richer sets of cognitive constraints (such as short-term memory (e.g. Lewis, Vasishth, & Van Dyke, 2006; Lewis & Vasishth, 2005)). These will of course introduce formidable computational and empirical challenges. But given the ubiquity of adaptive effects at all levels of human performance, we see this combination of optimal (rational) control analysis and mechanistic architecture modeling as a necessary part of understanding language and cognition.

References

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum.

Ballard, D. H., & Hayhoe, M. M. (2009). Modelling the role of task in the control of gaze. *Visual Cognition*, *17*(6-7), 1185-1204.

Bicknell, K., & Levy, R. (2010a). Rational eye movements in reading combining uncertainty about previous words with contextual probability. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*.

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

Bicknell, K., & Levy, R. (2012). The utility of modeling word identification from visual input within models of eye movements in reading. *Visual Cognition*, 1–18.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006, October). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700-765.

Brodersen, K. H., Penny, W. D., Harrison, L. M., Daunizeau, J., Ruff, C. C., Duzel, E., et al. (2008). Integrated bayesian models of learning and decision making for saccadic eye movements. *Neural Networks*, *21*(9), 1247–1260.

Edwards, W. (1961, January). Behavioral decision theory. *Annual Review of Psychology*, *12*, 473–98.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction-times,

and human information-processing. *Journal of Mathematical Psychology*, *2*(2), 312-329.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005, October). Swift: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777-813.

Forster, K. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to merrill garrett.* Hillsdale, NJ: Lawrence Erlbaum.

Geisler, W. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, *96*(2), 267.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, *29*(2), 228–244.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Hale, J. T. (2011). What a rational parser would do. *Cognitive Science*, *35*, 399–443.

Howes, A., Lewis, R. L., & Vera, A. H. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, *116*(4), 717–751.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, *4*, 237-285.

Kucera, H., & Francis, S. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Laubrock, J., Kliegl, R., & Engbert, R. (2006, January). SWIFT explorations of age differences in eye movements during reading. *Neuroscience and biobehavioral reviews*, *30*(6), 872–84.

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr Chips: An ideal-observer model of reading [Proceedings Paper]. *Psychological Review*, *104*(3), 524-553.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009, December). 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*(50), 21086-21090.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*, 375–419.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*, 44–54.

McConkie, G. W., Rayner, K., & Wilson, S. J. (1973). Experimental manipulation of reading strategies. *Journal of Educational Psychology*, *65*, 1.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, *104*, 3–65.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*, 22–34.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In Chase & W. G. (Eds.), *Visual information processing.* New York: Academic Press.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal bayesian decision process. *Psychological Review*, *113*(2), 327-357.

Norris, D. (2009). Putting it all together: A unified account of word recognition and reaction-time distributions. *Psychological Review*, *116*(1), 207-219.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed Effects Models in S and S-Plus.* Springer.

R Development Core Team. (2011). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from `http://www.R-project.org/` (ISBN 3-900051-07-0)

Rabbitt, P. (1979, May). How old and young subjects monitor and control responses for accuracy and speed. *British Journal of Psychology*, *70*(2), 305–311.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, *3*, 245–248.

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*(4), 445-+.

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010, Sep). Eye movements during mindless reading. *Psychological Science*, *21*, 1300–10.

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

Rothkopf, C. A., Ballard, D. H., Hayhoe, M. M., & Regan, O. (2007). Task and context determine where you look. *Journal of Vision*, *7*, 1–20.

Russell, S. J., Subramanian, D., & Parr, R. (1993). Provably bounded optimal agents. In *Ijcai'93: Proceedings of the 13th international joint conference on artifical intelligence* (pp. 338–344). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Salverda, A. P., Brown, M., & Tanenhaus, M. K. (2011). A goal-based perspective on eye movements in visual-

world studies. *Acta Psychologica*, *137*(2), 172–180.

Singh, S., Jaakkola, T., Littman, M. L., & Szepesvari, C. (2000, Mar). Convergence results for single-step on-policy reinforcement learning algorithms. *Machine Learning*, *38*(3), 287-308.

Smith, G. A., & Brewer, N. (1995). Slowness and age: Speed-accuracy mechanisms. *Psychology and Aging*, *10*(2), 238–247.

Starns, J. J., & Ratcliff, R. (2010, June). The effects of aging on the speed-accuracy compromise: Boundary optimality in the diffusion model. *Psychology and aging*, *25*(2), 377–90.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*(3), 251-260.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: The MIT press.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*(6), 401.

VanRullen, R., & Thorpe, S. J. (2001, May). The Time Course of Visual Processing: From Early Perception to Decision-Making. *Journal of Cognitive Neuroscience*, *13*(4), 454–461.

Wagenmakers, E., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, *58*(1), 140–159.

Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, *19*(3), 326–339.

Wotschack, C. (2009). *Eye movements in reading strategies: How reading strategies modulate effects of distributed processing and oculomotor control*. Univ.-Verl.

Appendix

Monte carlo simulation details.   For the modeling results reported below, we computed expected values for the measures (payoffs, reaction times, single fixation duration, etc.) by taking means of these values over 300,000 simulated model trials, where the words and nonwords in the trials were drawn from the experimental list used for human participants using the distribution described above. As in the human experiments the probability of a word trial was 0.50 (and thus a nonword trial 0.50). Across the 300,000 trials there were 50 different word and nonword lexicons of approximately 500 strings each as described in the text.

We do not report statistical tests on the empirical measures that the model produces: at 300k simulated trials for each noise and policy setting, the confidence intervals around the reported means are negligible.

Statistical Methods.   Data analysis on the human data was carried out using mixed effects regression (Pinheiro & Bates, 2000) using the `lme4` package for the R environment for statistical computing (R Development Core Team, 2011). For inference, models with maximal random-effects structures were fit: in trial-level analyses of condition this included by-participant and by-trial random intercepts, and by-trial random slopes (by-subject random slopes are not necessary because ours was a between-subjects design). In string-level analysis of condition this additionally included random slopes and intercepts of word and list position. In string-level analysis of frequency this included random slopes and intercepts of position but only random intercept of word (since frequency is a between-word factor). In string-level analysis of position this included a random slope and intercept of word. Linear models were fit to all timing measures, and a logit model was fit to accuracy. The first and last strings in each trial were excluded from analysis, as were strings appearing after a nonword.

Response times and single fixation durations (SFDs) that were farther than 3 standard deviations from the mean of those respective measures were removed. Some additional response times had to be removed due to a bug in response collection code.

Hypothesis tests were conducted using a single pair of normalized orthogonal contrasts. The first contrast, and the one of theoretical interest, is the contrast between the accuracy and speed payoffs (i.e. coding accuracy and speed as ±0.5 and balanced as 0). The second contrast is included for orthogonality but is not theoretically informative, so we do not report it here. This contrast design allowed us to retain the balanced condition for purposes of improving our error estimates, increasing overall power. We treated log frequency and position as numeric linear predictions for the purposes of those hypothesis tests. The effect of each contrast is reported as the p value of a likelihood ratio test (using the Chi squared distribution) comparing two models identical except for the presence of the contrast set of theoretical interest.

The belief update.   Here we detail the Bayesian belief update. Some definitions: A trial consists of a presentation of a list of l strings of h letters $S^1 S^2 \ldots S^{l-1} S^l$. Let the set of all word strings be denoted $\{W_i\}_{i=1}^n$ and the set of nonword strings be denoted $\{N_i\}_{i=1}^m$. Let $\mathcal{T}$ denote the multinomial random variable for trial type; it can take on mutually exclusive and exhaustive values $\mathcal{W}$ for word-trial, and $\mathcal{N}^k$ for a nonword trial with the nonword being at position $1 \leq k \leq l$. Let $Pr(\mathcal{T} = \mathcal{W})$ denote the probability of a word trial and $Pr(\mathcal{T} = \mathcal{N}^k)$ denote the probability of a non-word trial with nonword in position k. We will assume that for all k, i, and $Pr(S^k = W_i | \mathcal{T} = \mathcal{W}) = Pr(S^k = W_i | \mathcal{T} = \mathcal{N}^{j \neq k})$ and we will denote these equal

quantities by $Pr(S^k = W_i | \mathcal{T} \neq \mathcal{N}^k)$ for probability of a word string W_i given that this is a trial that cannot have a nonword at position k, and finally $Pr(S^k = N_i | \mathcal{T} = \mathcal{N}^k)$ for the probability of nonword N_i at position k given that it is a trial with a nonword at position k.

For each string W (similarly for N), let μ_{ij} be the $h \times 26$ matrix of 1's and 0's representing the indicator coding for the string, where each row i corresponds to a position in the string and each column corresponds to a letter of the alphabet. $\mu_{ij} = 1$ if the ith position in W has the jth letter, otherwise $\mu_{ij} = 0$. A sample s from word W (or nonword N) is a $h \times 26$ real-valued matrix, where each element s_{ij} is sampled independently from a normal distribution with mean 1 or 0 as given by the indicator coding and standard deviation σ, the sample noise parameter:

$$s_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma^2).$$

The likelihood of a sample s given some word string W (or nonwordstring N) is:

$$Pr(s|W) = \prod_{i,j} f(s_{ij}; \mu_{ij}, \sigma^2)$$

where $f(x; \mu, \sigma^2)$ is the probability density function of the Gaussian distribution with mean μ and standard deviation σ.

Let a sample from location k be denoted s^k. On receiving sample s^k, the belief update proceeds as follows. First, we compute some intermediate variables:

$$Pr_{old}(s^k | \mathcal{T} \neq \mathcal{N}^k)$$
$$= \sum_{j=1}^{n} Pr(s^k | S^k = W_j, \mathcal{T} \neq \mathcal{N}^k) Pr_{old}(S^k = W_j | \mathcal{T} \neq \mathcal{N}^k)$$
$$= \sum_{j=1}^{n} Pr(s^k | S^k = W_j) Pr_{old}(S^k = W_j | \mathcal{T} \neq \mathcal{N}^k)$$
$$Pr_{old}(s^k | \mathcal{T} = \mathcal{N}^k)$$
$$= \sum_{j=1}^{m} Pr(s^k | S^k = N_j, \mathcal{T} = \mathcal{N}^k) Pr_{old}(S^k = N_j | \mathcal{T} = \mathcal{N}^k)$$
$$= \sum_{j=1}^{m} Pr(s^k | S^k = N_j) Pr_{old}(S^k = N_j | \mathcal{T} = \mathcal{N}^k),$$
$$Pr_{old}(s^k)$$
$$= Pr_{old}(s^k | \mathcal{T} \neq \mathcal{N}^k) Pr_{old}(\mathcal{T} \neq \mathcal{N}^k) +$$
$$+ Pr_{old}(s^k | \mathcal{T} = \mathcal{N}^k) Pr_{old}(\mathcal{T} = \mathcal{N}^k)$$

where $Pr_{old}(\mathcal{T} \neq \mathcal{N}^k) = Pr_{old}(\mathcal{T} = \mathcal{W}) + \sum_{j \neq k} Pr_{old}(\mathcal{T} = \mathcal{N}^j) = 1.0 - Pr_{old}(\mathcal{T} = \mathcal{N}^k)$.

Next, we update the string-level beliefs:

$$Pr_{new}(S^k = W_i | \mathcal{T} \neq \mathcal{N}^k, s^k)$$
$$= \frac{Pr(s^k | S^k = W_i, \mathcal{T} \neq \mathcal{N}^k) Pr_{old}(S^k = W_i | \mathcal{T} \neq \mathcal{N}^k)}{Pr_{old}(s^k | \mathcal{T} \neq \mathcal{N}^k)}$$
$$= \frac{Pr(s^k | S^k = W_i) Pr_{old}(S^k = W_i | \mathcal{T} \neq \mathcal{N}^k)}{Pr_{old}(s^k | \mathcal{T} \neq \mathcal{N}^k)}$$
$$Pr_{new}(S^k = N_i | \mathcal{T} = \mathcal{N}^k, s^k)$$
$$= \frac{Pr(s^k | S^k = N_i, \mathcal{T} = \mathcal{N}^k) Pr_{old}(S^k = N_i | \mathcal{T} = \mathcal{N}^k)}{Pr_{old}(s^k | \mathcal{T} = \mathcal{N}^k)}$$
$$= \frac{Pr(s^k | S^k = N_i) Pr_{old}(S^k = N_i | \mathcal{T} = \mathcal{N}^k)}{Pr_{old}(s^k | \mathcal{T} = \mathcal{N}^k)}$$

Finally, we update the trial level beliefs:

$$Pr_{new}(\mathcal{T} = \mathcal{W} | s^k)$$
$$= \frac{Pr_{old}(s^k | \mathcal{T} = \mathcal{W}) Pr_{old}(\mathcal{T} = \mathcal{W})}{Pr_{old}(s^k)} =$$
$$= \frac{Pr_{old}(s^k | \mathcal{T} \neq \mathcal{N}^k) Pr_{old}(\mathcal{T} = \mathcal{W})}{Pr_{old}(s^k)}$$
$$Pr_{new}(\mathcal{T} = \mathcal{N}^{j \neq k} | s^k)$$
$$= \frac{Pr_{old}(s^k | \mathcal{T} = \mathcal{N}^j) Pr_{old}(\mathcal{T} = \mathcal{N}^j)}{Pr_{old}(s^k)}$$
$$= \frac{Pr_{old}(s^k | \mathcal{T} \neq \mathcal{N}^k) Pr_{old}(\mathcal{T} = \mathcal{N}^j)}{Pr_{old}(s^k)},$$
$$Pr_{new}(\mathcal{T} = \mathcal{N}^k | s^k) =$$
$$= \frac{Pr_{old}(s^k | \mathcal{T} = \mathcal{N}^k) Pr_{old}(\mathcal{T} = \mathcal{N}^k)}{Pr_{old}(s^k)}$$

In order to make decisions, in addition to the probability that the trial is a word trial or not that we compute above, we also need the probability that the string at position k is a word or nonword, i.e., $Pr(S^k \in \cup_{i=1}^{n} \mathcal{W}_i)$ and $Pr(S^k \in \cup_{i=1}^{m} \mathcal{N}_i)$:

$$Pr(S^k \in \cup_{i=1}^{n} \mathcal{W}_i) = \sum_{i=1}^{n} Pr_{new}(S^k = W_i | \mathcal{T} \neq \mathcal{N}^k) Pr_{new}(\mathcal{T} \neq \mathcal{N}^k)$$

$$Pr(S^k \in \cup_{i=1}^{m} \mathcal{N}_i) = 1.0 - Pr_{new}(S^k \in \cup_{i=1}^{n} \mathcal{W}_i)$$

The full process then iterates on the next sample, with each Pr_{new} becoming the next Pr_{old}.