



# A dynamical model of saccade generation in reading based on spatially distributed lexical processing

Ralf Engbert<sup>a,\*</sup>, André Longtin<sup>b</sup>, Reinhold Kliegl<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Potsdam, P.O. Box 601553, 14415 Potsdam, Germany

<sup>b</sup> Department of Physics, University of Ottawa, P.O. Box 450, Ottawa Ont., Canada, K1N 6N5

Received 8 March 2001; received in revised form 22 October 2001

## Abstract

The understanding of the control of eye movements has greatly benefited from the analysis of mathematical models. Currently most comprehensive models include sequential shifts of visual attention. Here we propose an alternative model of eye movement control, which includes three new principles: spatially distributed lexical processing, a separation of saccade timing from saccade target selection, and autonomous (random) generation of saccades with foveal inhibition. These three features provide a common control mechanism for fixations, refixations, and regressions. Consequently, the model is called SWIFT (Saccade-generation with inhibition by foveal targets). Results from numerical simulations are in good agreement with effects of word frequency on single-fixation, first-fixation, and gaze durations as well as fixation and word skipping probabilities in first-pass analysis. The model inherently produces complex eye movement patterns including refixations and regressions due to its underlying dynamical principles. © 2002 Elsevier Science Ltd. All rights reserved.

**Keywords:** Reading; Eye movement; Saccade; Modeling

## 1. Introduction

Theoretical analyses of eye movement control in reading provide an important case study for the dynamical allocation of visual attention. During normal reading, lexical processing of words is the main task, which has to be performed under the restrictions arising from our eye movement system (Rayner, 1998). It is commonly agreed that the difficulty of lexical processing (e.g., related to word frequency, predictability of words by context) and constraints related to oculomotor control and attention allocation (e.g., saccade latencies, perceptual span) influence eye movements in reading. The time required to program a saccadic eye movement is of the same order of magnitude (150 ms) as lexical

processing time for a single word of average difficulty (150–300 ms). Given an average fixation duration of 220–280 ms in normal reading, it is clear that saccade programming and lexical processing must occur—at least partially—in parallel. There is considerable debate, however, about how these processes are coupled. In this article, a new model of eye movement control in reading, called SWIFT,<sup>1</sup> is proposed to investigate the dynamical interaction of lexical and oculomotor processes in detail. Before we present our modeling assumptions, we briefly review three theoretical approaches to the control of eye movements in reading under the labels “primary oculomotor control” (POC), “sequential attention shift” (SAS), and “guidance by attentional gradient” (GAG).<sup>2</sup> The SWIFT model we develop here has its primary affiliation with the last category but it draws heavily on notions developed in SAS models and embodies some (but not all) established notions of POC.

POC refers to the basic assumption that eye movements are driven by low-level factors. For example, Reilly

\* Corresponding author. Address: Center for Dynamics of Complex Systems, University of Potsdam, P.O. Box 601553, 14415 Potsdam, Germany.

E-mail addresses: engbert@rz.uni-potsdam.de (R. Engbert), alongtin@physics.uottawa.ca (A. Longtin), kliegl@rz.uni-potsdam.de (R. Kliegl).

URLs: <http://www.agnd.uni-potsdam.de/~ralf>, <http://www.science.uottawa.ca/phy/profs/Longtin>, <http://www.psych.uni-potsdam.de/people/kliegl/index-e.html>.

<sup>1</sup> Saccade-generation with inhibition by foveal targets.

<sup>2</sup> For a comprehensive overview of the state of this research we refer to the volume edited by Kennedy et al. (2000).

and O'Regan (1998) (see also O'Regan & Lévy-Schoen, 1987; O'Regan, 1990, 1992) assume that the eye is directed to the longest word in the area of about 20 characters to the right of fixation and that oculomotor errors (e.g., overshoot of close and undershoot of far targets) lead to characteristic within-word corrections required for word identification. McConkie, Kerr, and Dyre (1994) proposed a two-state transition model. In state 1, either fast saccades (i.e., saccades in which no information is processed from the point of fixation) occur with a very low probability that rises slowly according to a linear hazard function, or a transition occurs to state 2 with a more rapidly rising linear hazard function. In state 2 normal saccades are generated according to a random waiting time distribution. Psychologically the two states capture a distinction between primarily oculomotor (state 1) and primarily information uptake (state 2) processes. Such models are based on and provide good accounts of distributions of landing positions and the probability of word skipping as a function of the distance from which the saccade was launched (McConkie, Kerr, Reddix, & Zola, 1988) (see also Suppes, 1990, for a similar approach). Note that the significance of lexical processes is not denied, but they are assumed to play only a relatively minor modulating influence relative to effects of word length. Within this context we can assume that in normal reading many lexical-processing effects may eventually emerge as epiphenomena of a POC process (e.g., via the negative correlation between word length and word frequency). At present, however, there appears to be no computational variant of a POC model which explicitly models the "secondary" influence of lexical-access effects.

SAS models assume a close coupling of lexical processing, attentional shifts, and eye movements. These models derive from Morrison's (1984) proposal that a covert shift of attention to the next word occurs simultaneously with the initiation of a saccade program to the same word. Lexical processing of a (parafoveal) word starts at the same time the saccade program is initiated. If processing of the parafoveal word is completed before the saccade is executed (which is quite likely for easy or highly predictable words), the saccade is reprogrammed to the following word. This mechanism provides an elegant account of selective skipping of short high-frequency words. In a further development of this model, Reichle, Pollatsek, Fisher, and Rayner (1998) assumed that (1) lexical access occurs in two steps, namely a familiarity check and a subsequent process of lexical completion, and (2) there is a distinction between a labile and a subsequent nonlabile phase of saccade programming. Saccade programming and lexical access are synchronized at the completion of the familiarity check, where in parallel to lexical completion a new saccade is programmed to the next word. Saccades can be reprogrammed only during the labile phase. Covert shifts of

attention to the next word occur after lexical completion irrespective of the state of the saccade program. The partial independence of lexical and oculomotor programs in this fully computational model goes beyond the Morrison model and provides an account of mean fixation durations and word skipping probabilities in relation to word frequency. Moreover, it yields an elegant account of the negative effect of the difficulty of the foveal word on subsequent processing of the parafoveal word (i.e., preview benefit)—a result incompatible with the Morrison model. In a recent version, Reichle, Rayner, and Pollatsek (1999) incorporate oculomotor control parameters with which they account for distributions of within-word fixation positions as reported in the corpus used by Rayner (1998), thus extending the model to account for the database of POC models. Engbert and Kliegl (2001) reported a variant of such an SAS model in which the coupling of lexical and oculomotor program occurs at the end of lexical completion (i.e., when attention is moved covertly to the next word) avoiding the problem of programming a saccade to an unattended word which is difficult to reconcile with results from basic attention research (e.g., Kustov & Robinson, 1996; Deubel & Schneider, 1996; Kowler, Anderson, Doshier, & Blaser, 1995). These studies showed that accurate saccades require prior shifts of perceptual attention to a target. Our model also allows for autonomously triggered saccades as postulated by POC models. Preview benefit arises from the reduction of parafoveal processing due to these autonomous saccades.<sup>3</sup>

*Critique of SAS models:* The computational implementation of SAS models opens the possibility for falsifications which represents a major advantage over previous conceptualizations. Not the least due to the computational precision of such models, it has become apparent that there are a number of empirical and experimental observations which cannot be accommodated within this framework (see e.g., Deubel, O'Regan, & Radach, 2000; Kennedy, 2000b). Among these problems, there are three aspects which motivate our current computational study. First, current SAS models cannot account for the influence of the difficulty of the subsequent parafoveal word on the processing of the foveal word. Effects have been shown for informativeness of word beginnings (Inhoff, Starr, & Shindler, 2000; Kennedy, 2000a; Underwood, Binns, & Walker, 2000; Rayner, Fischer, & Pollatsek, 1998), word length and word frequency (Kennedy, 1998) and semantics (Inhoff, Radach, Starr, & Greenberg, 2000; Murray, 1998). Note, however, that at least with respect to word fre-

<sup>3</sup> In a methodological refinement the model also replaces assumptions on the distributions of residence times, like Gamma distributions (Reichle et al., 1998) with semi-Markov processes, a generalization of Markov processes with residence-time dependent transition probability rates (Gillespie, 1978).

quency the effects are controversial (Henderson & Ferreira, 1990; Rayner et al., 1998; Inhoff, Starr et al., 2000). Second, there is evidence that information can be picked up to the left of fixation (Binder, Pollatsek, & Rayner, 1999; Inhoff, Radach et al., 2000). The E-Z Reader model (Reichle et al., 1998) offers a solution to this via a refixation mechanism, but the behavioral consequences of this implementation (e.g., necessity of rereadings, longer fixations prior to regressions) have been criticized (Inhoff, Radach et al., 2000; Kennedy, 2000b). Third, current SAS models predict that fixations prior to a skipped word should be longer due to the time required for reprogramming the saccade, but Radach and Heller (2000) convincingly showed that this is simply not the case (see also Hogaboam, 1983; McConkie et al., 1994, Fig. 6).

In addition to these three empirical issues, current SAS models do not provide a common mechanism for saccades moving from word to word, skipping words in forward direction, regressing to previous words, and refixating the current word. For example, including a mechanism for refixations (requiring one additional rule) in the E-Z Reader framework increased the number of internal states of this model from 8–14. Such an increase in number of internal states, however, often limits our understanding of the resulting dynamics as even models with few state variables may show rich dynamical behavior. In this paper we present a model that generates all types of eye movements in reading as a consequence of three basic principles.

*GAG*: It appears that models assuming GAG might be compatible with empirical observations that are problematic for SAS models. Inhoff, Radach et al. (2000) proposed the following conceptualization which we summarize in five points: (1) All words within the range of effective vision are attended to and subjected to lexical analyses, not only one word at a time (Inhoff, Starr et al., 2000; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999). (2) Allocation of attention to different segments of a spatially contiguous array is determined from a gradient value. (3) Gradient values are a function of fixation location (i.e., fixated word gets a high value for visual resolution) and of success of linguistic analyses (i.e., higher values for words difficult to recognize). (4) Successful word recognition (or a subword unit or sequence of words) could lead to a gradient adjustment (i.e., a dynamic adjustment of center of attention). In particular, allocation of attention may be increased to nonrecognized neighbors and decreased to the identified unit. (5) Saccades are programmed towards the new center of attention. Such a model could account for effects of foveal difficulty on parafoveal processing and vice versa.

Clark (1999) published a connectionist model which had implemented some of the proposals by Inhoff, Radach et al. (2000). The model links shifts in spatial

attention (via a winner-take-all model integrating spatio-temporal characteristics of low level feature detectors) and the generation of saccadic eye movements. It accounts for saccadic latencies in various low-level oculomotor phenomena (i.e., the double-step paradigm and gap effects) and, of particular relevance for reading, for the modulation of saccadic accuracy as a function of target predictability and saccadic latency (Coëffé & O'Regan, 1987). Thus, this computational model relates primarily to POC models. Legge, Klitz, and Tjan (1997) also proposed a gradient-type model to determine the saccade that minimizes uncertainty about the current word in an ideal-observer model of reading. At each location the entropy contributed by each letter within the visual span is computed from three sources (visual, lexical and oculomotor information) and is used to determine the next eye movement. Interestingly, in this model regressions occur as an emergent phenomenon of the computational algorithm. The model is an ideal-observer model and, in this respect, its implications for human performance are to explore constraints in information uptake in an explicit manner. To our knowledge, so far no computational GAG model has been implemented fully quantitatively with an explicit focus on the coupling of lexical access and oculomotor control and a comparable scope of behavioral phenomena in analogy to the SAS models described above.

The SWIFT model which we propose here is basically a variant of GAG. This approach can also be motivated and constrained by recent neurophysiological evidence. Two observations are particularly pertinent for the present concerns. First, the motor map in the intermediate layers of the superior colliculus codes the spatial allocation of attention which serves as the target for saccade programs (Munoz & Wurtz, 1993a; Wurtz, 1996). Thus, there is some evidence for a coupling of attention shifts and the initiation of saccade programs to this area. Interestingly, Rizzolatti's (1983) premotor theory even suggests that covert attentional shifts and initiation of a saccade are identical processes. Second, in SAS models, saccade timing and saccade target selection are considered as inseparable processes. Recent neurophysiological findings suggest just the opposite and provide a distinction of a "when" from a "where" pathway in the preparation of eye movements (Carpenner, 2000; Wurtz, 1996).

For the development of our model, the underlying assumptions can be summarized as Principles I and II:

- *Principle I*: Lexical information processing is spatially distributed over an attentional window.
- *Principle II*: Saccade timing is separated from saccade target selection.

The formulation of a computational framework, which includes the possibility of parallel processing of

several words at a time, is a current challenge of eye movement research in reading (Starr & Rayner, 2001). We propose a fully quantitative model based on Principles I and II as a viable alternative to current SAS models to meet this challenge.

An additional assumption concerns the initiation of saccade programs. Here we assume a random timing mechanism, i.e. an underlying distribution of intersaccadic intervals. Since this property suggests that the time-course of saccades is independent of lexical processing, we call this random timing *autonomous* (Engbert & Kliegl, 2001). It will turn out, however, that a purely autonomous timing fails to predict that first-fixation durations depend on word frequency. This problem with random timing can be solved by introducing a foveal inhibition process which operates only in short episodes during fixation of difficult words. This assumption on saccade timing is a central property of SWIFT, which we formulate as

- *Principle III*: Saccade generation is an autonomous (random) process with inhibition by foveal targets.

SWIFT is motivated and mathematically formulated in the next three sections. We start with our assumptions on lexical processing. The second step is to explain our concept of an attentional window (Principle I). We then focus on the programming and execution of saccades (Principle II). Next, we explain inhibition of saccade programming by lexical processing and our assumption on execution of saccades (Principle III). We give an overview of the proposed model to sum up the working principles. Finally, we present results from model simulations and give an outlook on modeling regressions and more complex eye movement patterns.

## 2. Lexical processing: a two-level process

A key factor that drives eye movements during reading is lexical processing. Information located within 2° of visual angle (6–8 characters) around the fixation point is processed in foveal vision. Additionally, parafoveal preview (within 5° of visual angle) provides information on words to the right (and left) of the currently fixated word. To keep track of the time-evolution of the ongoing lexical processing, we associate a state variable to each word, which we will refer to as *lexical activity*. We denote the lexical activity of word<sub>*n*</sub> at time *t* by  $a_n(t)$ . The set of lexical activities of all words,  $\{a_n(t)\}$ , changes over time due to lexical processing. We will interpret the relative lexical activity (in relation to the activities of all words in a given sentence) as a measure of the probability that the word under consideration is selected as a saccade target. The time-evolution of the set of lexical activities,  $\{a_n(t)\}$ , leads to a

change of target selection probabilities which influence the eye's trajectory of fixations.

It is important to note the dynamical nature of the relation between the state of lexical processing of words and the movements of our eyes. As soon as we start to process a word, our knowledge about the text starts to increase. Therefore, the decision where and when to move the eyes changes over time and depends strongly on previous fixation locations. This historicity is inherent in the concept of lexical activities, which we propose in our model. Since historicity may be looked upon as a key property of dynamical systems, we emphasize dynamical aspects of eye movement control with our basic assumption on lexical activities.<sup>4</sup> As a result, the temporal evolution of lexical processing induces complex eye movement behavior including refixations and regressions, as long as a flexible target selection mechanism is specified.

We use a two-level process for lexical access. First, during a lexical preprocessing stage, lexical activity is increasing,  $da_n/dt > 0$ . As a consequence of this build-up of lexical activity, the probability to select the word as a saccade target is increasing. The end of the lexical preprocessing stage is reached at the maximum activity value,  $l_n$ , which represents the word's lexical difficulty and is related to word frequency  $f_n$ . As in several previous modeling approaches (Reichle et al., 1998; Engbert & Kliegl, 2001) we use the relation

$$l_n = \alpha - \beta \log(f_n). \quad (1)$$

When this maximum value of lexical activity is reached, lexical information processing continues in a second stage called *lexical completion*. During this lexical completion process, we assume a decreasing lexical activity, i.e.  $da_n/dt < 0$ , until lexical access is completed when  $a_n(t)$  reaches zero.

In SWIFT, a unimodal temporal evolution of lexical activities is used to capture the time-course of target selection probabilities. Before a word is selected as a saccade target, we assume that, by default, some preprocessing is performed in order to get some information on low-level properties of the word (e.g., word length). During this preprocessing, the increase of lexical activity results in an increased target selection probability. At a later stage, the probability to select a word as target for a saccade should decrease, as the lexical completion of the word approaches its end. This effect is reproduced by a unimodal time-evolution of lexical activities during lexical processing.

It will turn out later that the concept of lexical activities also provides a simple mechanism for complex eye movement patterns (refixations and regressions). If we assume that the fixation of the word has started and

<sup>4</sup> See Jackson (1991) for an introduction to dynamical systems.

lexical activity is still moderate after a certain time (as would occur, e.g., if the word is difficult to access), then a second selection of the word as a target is highly probable, which may cause a refixation in our model.

The lexical difficulty of a word also depends strongly on its context. A word may become predictable from previous words and this can facilitate lexical processing. We include a prediction process (Reichle et al., 1998), competitive to lexical access, by multiplying the lexical difficulty variables by the complement of the empirically observed probabilities  $p_n$  for predicting word<sub>*n*</sub> from the sequence of all previous words (1, 2, 3, ...,  $n - 1$ ) in a given sentence, i.e.

$$L_n = (1 - p_n)l_n. \quad (2)$$

The value  $L_n$  actually is a rough measure of the mean probability of selection as a target for a saccade.

### 3. Spatially distributed lexical processing

In accordance with decreasing visual acuity from fovea to parafovea, we propose that the lexical processing rate is highest in the center of the visual field (fovea) and decreases to the periphery. In most SAS models, two words are permitted to be processed from a certain fixation position  $k(t)$ , i.e. word<sub>*k*</sub> or word<sub>*k+1*</sub>,<sup>5</sup> the latter one due to a shift of attention. As discussed above, however, for these models there is no overlap in processing time, i.e. lexical processing is strictly serial. A key concept of our new model is spatially distributed processing (Principle I), motivated by experimental findings. First, Kennedy (1998, 2000a) was able to demonstrate that lexical processing is “distributed over a region larger than a single word.” Second, Binder et al. (1999) have shown that “readers often still attend to a word after it is skipped and that when readers fixated a word, they occasionally attend to the word after they have begun to fixate the next word.” These results suggest that the allocation of attention during reading is a more spatially distributed process than assumed in SAS models. Therefore, a model that includes spatially distributed processing seems psychologically plausible and may give important new insights on the dynamical interplay between allocation of visual attention and eye movement control.

Our basic assumption for the implementation of spatially distributed processing is that the lexical processing rate, denoted by  $\lambda > 0$ , is a function of the distance (eccentricity) from the current fixation position  $k(t)$  at time  $t$ . As a first-order approximation, we compute this eccentricity  $\epsilon$  in units of number of words, i.e. we assume a discrete space of word positions  $n$

( $n = 1, 2, 3, \dots, N_w$ , where  $N_w$  is the number of words in a given sentence). Since the fixation position  $k(t)$  changes as the eyes move, the eccentricity of word<sub>*n*</sub> is also a function of time,

$$\epsilon_n(t) = n - k(t). \quad (3)$$

The lexical processing rate is a function of the eccentricity,  $\lambda = \lambda(\epsilon)$  because of the structure of our visual field. The size of the visual span decreases from at least 10 letters in central vision to 1.7 letters at 15° eccentricity (Legge, Mansfield, & Chung, 2001), in good agreement with the corresponding reduction of reading speed (Chung, Mansfield, & Legge, 1998). To account for these psychophysical restrictions, we assume that the maximum of the lexical processing rate is reached at eccentricity  $\epsilon = 0$  (the fovea), i.e.  $\lambda(0) = \max\{\lambda\}$  and that lexical processing rate decreases with increasing eccentricity, i.e.  $\lambda(-1) < \lambda(0)$  and  $\lambda(0) > \lambda(1) > \lambda(2)$ . For simplicity, we fix  $\lambda(-1) = \lambda(1)$  in a first attempt to include spatial effects. Furthermore, we assume that the lexical processing rate vanishes if  $\epsilon < -1$  or  $\epsilon > 2$ , i.e. the attentional window has a length of four words, extending one word to the left and two words to the right. This asymmetry in the spatial extension of the attentional window is compatible with research on the perceptual span. These simplifying assumptions are necessary for a first study of parallel processing of several words in a computational model and may be looked upon as a first-order approximation. Obviously, in future research, modeling must account for spatial processing constraints at the level of both words and letters. For now, lexical processing of word<sub>*n*</sub> can be written as

$$\frac{da_n(t)}{dt} = \begin{cases} +f\lambda(\epsilon_n(t)) & \text{if } t < t_p \text{ (lexical preprocessing)} \\ -\lambda(\epsilon_n(t)) & \text{if } t \geq t_p \text{ (lexical completion)} \end{cases} \quad (4)$$

where  $f > 1$  is a lexical preprocessing factor and  $t_p$  is the time,<sup>6</sup> when the lexical activity first reaches the maximum value  $L_n$ . At the end of lexical completion, lexical activity reaches zero,  $a_n(t) = 0$ , and stays so throughout the simulation. We assume that lexical preprocessing is faster than lexical completion, i.e.  $f \geq 1$ , because it represents a preliminary stage of processing, including recognition of basic word properties like word length and initial trigram frequency.

We now illustrate the interplay between lexical processing (Eq. (4)) and word eccentricity (Eq. (3)) with two examples. Fig. 1(a) illustrates the time-course of lexical activity during lexical processing for three different eccentricities. Note that the eccentricity  $\epsilon$  influences the lexical processing rate  $\lambda$ , where a higher processing rate is reflected in a steeper slope in the lexical activity function  $a_n(t)$ . During model simulations, however, eye

<sup>5</sup> Or word<sub>*k+2*</sub>, if word<sub>*k+1*</sub> was skipped etc.

<sup>6</sup>  $t_p = \min\{t | a_n(t) = L_n\}$ .

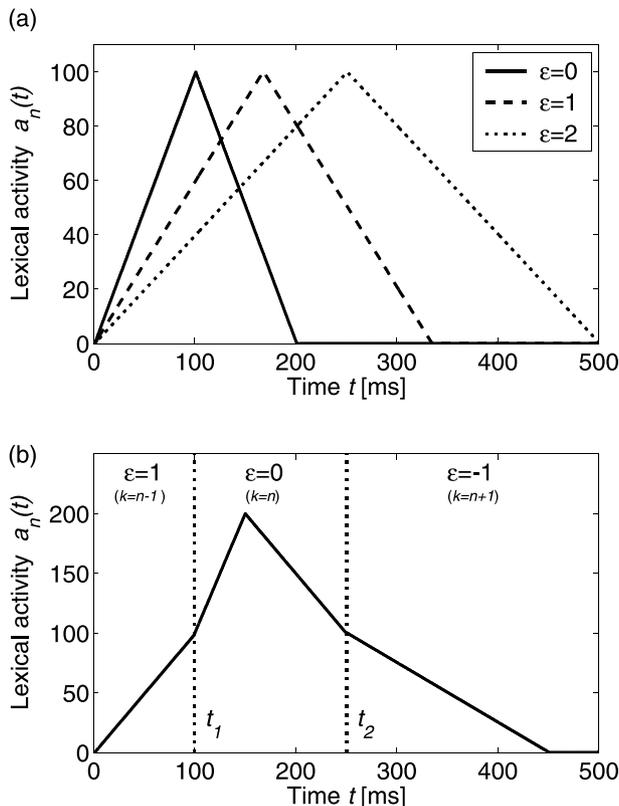


Fig. 1. Illustration of lexical processing in SWIFT. (a) Influence of eccentricity  $\epsilon$  on the time evolution of lexical activity (lexical processing rate:  $\lambda(0) = 1$ ,  $\lambda(1) = 0.6$ ,  $\lambda(2) = 0.4$ ; preprocessing factor:  $f = 1$ ; lexical difficulty:  $L_n = 100$ ). Note that the slopes of the lexical activity reflect the word's eccentricity. (b) Example of the time-course of lexical activity  $a_n(t)$  during eye movements ( $k(t) = n - 1$  for  $t < t_1$ ,  $k(t) = n$  for  $t_1 \leq t < t_2$ ,  $k(t) = n + 1$  for  $t \geq t_2$ ;  $\lambda(1) = 0.5$ ,  $\lambda(0) = 1$ ,  $\lambda(-1) = 0.5$ ;  $f = 2$ ;  $L_n = 200$ ). Eye fixation position  $k(t)$  and eccentricity  $\epsilon_n(t)$  of word<sub>*n*</sub> can be calculated from each other by Eq. (3).

fixation position  $k(t)$  changes over time. Since eye movements induce a change in the set of eccentricities  $\{\epsilon_n(t)\}$  (Eq. (3)), the lexical processing rate for a given word<sub>*n*</sub> typically changes strongly over time. This effect is illustrated in a simple example in Fig. 1(b). During lexical processing of word<sub>*n*</sub>, the eye moved from word<sub>*n-1*</sub> to word<sub>*n*</sub> at time  $t_1$  and from word<sub>*n*</sub> to word<sub>*n+1*</sub> at time  $t_2$ . The change in the eye fixation position coincides with different slopes in the time-evolution of the lexical activity  $a_n(t)$  of word<sub>*n*</sub>.

Our implementation of lexical processing is currently based on three assumptions. First, physical space is represented as a one-dimensional discrete space of word positions in the model. This rather coarse picture represents a first approximation to the inclusion of spatial effects into a theoretical model for eye movement control. In particular, this simplification is to be modified in an advanced version of the model to include an explanation of results on initial landing positions in word targeting (McConkie et al., 1988; Reilly & O'Regan, 1998; Reichle et al., 1999). A possible way to include the

effect of initial landing positions is to assume a set of *spatiotemporal functions*  $\{a_n = a_n(x, t)\}$  of lexical activity with a mean spatial length scale related to word length. An extension of our model in this way relates to recent theoretical concepts developed in the *dynamic neural field theory* (Thelen, Schönner, Scheier, & Smith, 2001) (see Discussion).

Second, the spatial distribution of lexical processing rates may be interpreted in terms of statistical allocation of attention or as parallel processing. In the statistical interpretation, the position of the eye  $k(t)$  determines the probability for the distribution of attention. In this regard, the allocation of attention could be included as an additional random process. A stochastic process for visual attention might have benefits on processing speed, since recent results on visual search suggest that an “anarchic” or random allocation of visual attention is faster than a volitional strategy (Wolfe, Alvarez, & Horowitz, 2000)—even though important features of reading are thought to be volitional at the highest level of processing. Alternatively, we may interpret spatially distributed lexical access as parallel processing, i.e. several words are lexically processed at a time. This assumption might have even stronger implications for related problems in reading. At the present level of development of the model, however, it seems neither necessary to choose nor possible to decide between these plausible interpretations.

Third, in a more comprehensive version of the model, the parameters of the attentional window could be allowed to depend on lexical activity, i.e.  $\lambda = \lambda(\epsilon, a_k(t))$ . This is a possible way to study a more dynamic interaction between lexical processing and eye fixation position. For example, when the reading material is difficult, the size of the perceptual span tends to be smaller than for a text of average difficulty (Henderson & Ferreira, 1990; Inhoff & Rayner, 1986; Balota, Pollatsek, & Rayner, 1985) and depends on the skill of the reader (Rayner, 1986). This effect could be accounted for by a sharper distribution of the attentional window parameters  $\lambda(\epsilon, a_k)$  for increasing foveal word difficulty  $a_k(t)$ . In the current study, however, we restrict our analysis to the basic architecture of the model.

As the eyes move over text, the lexical activities  $\{a_n(t)\}$  change over time as a function of the fixation position  $k(t)$ . An important feature of the model is that the lexical activity  $a_n(t)$  of word<sub>*n*</sub> determines its probability to be selected as a saccade target. Therefore, the model provides a mechanism to integrate processing of sensory information and movement preparation.

#### 4. Programming of saccades

The decision to program an eye movement is followed by the launch of a saccade after the saccade la-

tency period (roughly 150 ms). But since lexical processing, i.e. our knowledge about the words, is simultaneously changing the lexical activities  $\{a_n(t)\}$ , sometimes the movement of the eyes may be no longer desirable or, at least, may prefer a different target. As a consequence, the occurrence of saccade latencies has strong implications for the dynamics of eye movement control. The next two sections discuss how saccade programming is included in the model. First, we distinguish target selection and saccade timing and, second, we propose a lexical inhibition mechanism for saccade timing. Finally, we discuss how the execution of saccades is implemented.

#### 4.1. Where and when to move the eyes

Neurophysiological findings suggest a separation between target selection and saccade timing on a neural level (Carpenter, 2000; Findlay & Walker, 1999; Wurtz, 1996). We include this distinction of “where” from “when” in saccade programming as Principle II in our model, schematically illustrated in Fig. 2. First, a two-stage programming of saccades is motivated by results from double-step experiments (Becker & Jürgens, 1979). According to these findings, the target of a saccade can be modified even after the saccade program is initiated. We account for this effect by introducing a labile saccade program (like Reichle et al., 1998), during which

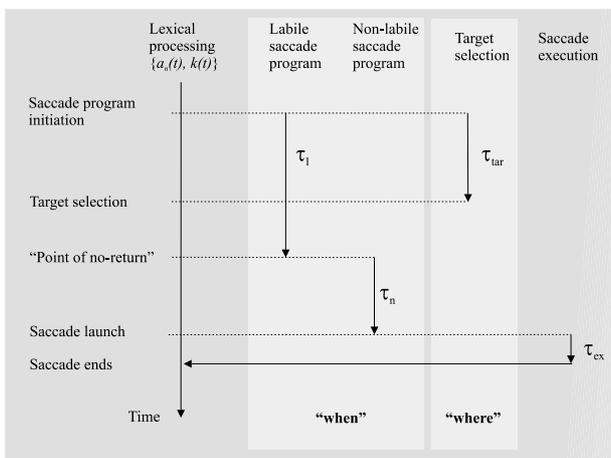


Fig. 2. Programming of saccades. A decision to initiate a saccade program at time  $t_1$  induces a labile saccade program. During this labile stage of saccade programming, a second initiation of a saccade program at time  $t_2 > t_1$  resets the time-course (cancellation). At the end of the labile stage (duration  $\tau_l$ ), the nonlabile stage is entered (duration  $\tau_n$ ). The selection of a saccade target is, in a different pathway, at a preliminary stage and may, therefore, be modified during the target selection program (duration  $\tau_{tar}$ ). When the nonlabile saccade program stage is entered, the saccade will be executed with the pre-selected target. At the end of the nonlabile stage, the saccade is launched. After saccade execution, the eyes have arrived at the new fixation position.

the saccade can be canceled. In contrast to E-Z Reader models we also allow target modification during this labile program.

Second, target selection is based on the set of lexical activities  $\{a_n(t)\}$ . The conditional probability  $\pi(n, t|k)$  for word<sub>n</sub> to be selected as a saccade target at time  $t$ , if the current fixation position is at word<sub>k</sub>, is given by its relative lexical activity, i.e.

$$\pi(n, t|k) = \begin{cases} \frac{a_n(t)}{\sum_{m=1}^{k+2} a_m(t)}, & \text{if } n \leq k + 2 \\ 0, & \text{if } n > k + 2 \end{cases} \quad (5)$$

if  $\sum_{m=1}^{k+2} a_m(t) > 0$ . This form of target selection is motivated by the idea that important target words are in an intermediate state of lexical processing with lexical activities  $a_n(t)$  close to their maximal values  $L_n$ . Whenever the denominator in Eq. (5) vanishes, we select as a saccade target the next closest word to the right of the current attentional window, which is not (or not completely) lexically processed (or terminate the simulation if all words are completely processed).<sup>7</sup>

Since the target selection process is not mandatorily coupled to saccade timing, the model provides a common mechanism for forward saccades, refixations and regressions. It is important to note that we do not assume that target selection is a process that occurs later in time than the initiation of the labile saccade program. Instead, we suggest that saccade targets can be modified during the target selection time interval  $\tau_{tar}$  (Fig. 2). For the implementation of target selection, we simply use the set of probabilities  $\{\pi\}$ , computed at the end of the target selection interval.

How do our assumptions relate to the E-Z Reader models (Reichle et al., 1998)? In E-Z Reader, the end of a lexical preprocessing stage, called familiarity check, initiates a saccade program to the next word. Hence, a saccade is always initiated with a fixed target. A cancellation of a saccade during the labile stage of the saccade program can occur; a change of the saccade target, however, necessarily influences the time course of the saccade. Here our assumption is more general, as a saccade can be canceled (“when” pathway) or its target can be modified (“where” pathway), and these two changes are independent of each other. Furthermore, target modification does not necessarily delay saccade onset (Becker & Jürgens, 1979).

<sup>7</sup> The case distinction restricts target selection to all words from the beginning of the sentence to the right border of the attentional window. Eq. (5) prevents oscillations between regressions and long forward saccades in sentences with high lexical difficulty. In an earlier version of our model, we used a simpler form of target selection,  $\pi(n, t) = a_n(t) / \sum_{m=1}^{N_w} a_m(t)$ , which leads to similar results for an analysis of first-pass eye movements.

#### 4.2. Inhibition of saccade programming

A straightforward assumption for the temporal control of saccades is that programming of a new saccade is triggered by the end of the current lexical access (Morrison's, 1984), motivated by the fact that text comprehension is the goal of normal reading. This principle of lexical or "high level" temporal control is also used in E-Z Reader models, in which, in contrast with the Morrison model, the initiation of a new saccade program is triggered by the end of a familiarity check, i.e. a preliminary stage of lexical access (Reichle et al., 1998). However, as shown in a recent theoretical study (Engbert & Kliegl, 2001), within the class of SAS models the strong assumption of lexical control can be relaxed by allowing for autonomous saccades, i.e. saccade programs that are not triggered by a lexical signal.

Here we assume an even stronger autonomy of the saccadic system, a "dumb" default strategy (Deubel et al., 2000) for saccade timing based on a preferred mean rate of saccades, with lexical processes intervening only if eye position and lexical processing state are threatening to desynchronize. In our model, we use a stochastic interval  $t_s$  (according to a predefined gamma distribution) between two subsequent decisions to program a new saccade. A new saccade program is started, if time  $t = t_s$ , where  $t$  denotes the time elapsed since the last start of a (labile) saccade program. Numerical simulations indicate that this type of random timing of saccades can already reproduce the basic dependence of gaze duration on word frequency (Fig. 6 and discussion in Section 6.3), since our target selection mechanism, Eq. (5), produces longer gaze durations for low-frequency words using several fixations (refixations). Random timing, however, must obviously fail to produce a frequency-dependent first fixation duration. Therefore, we assume that the stochastic interval  $t_s$  may be inhibited by (foveal) lexical activity  $a_k(t)$  (Principle III). This inhibition mechanism is implemented by an additive contribution of  $a_k(t)$  to the random interval  $t_s$ . In this case, a new labile saccade program starts, if

$$t = t_s + ha_k, \quad (6)$$

where  $h$  is a factor representing the strength of inhibition by foveal lexical activity  $a_k(t)$ . It is important to notice that  $a_k(t)$  rapidly approaches zero, when word <sub>$k$</sub>  is fixated. A variation of the inhibition factor  $h$  can be interpreted as changing the saccade timing mechanism between the two extreme cases of autonomous timing ( $h = 0$  and  $\text{var}(t_s) > 0$ ) and completely lexically controlled movements ( $h = \text{const} > 0$  and  $\text{var}(t_s) \rightarrow 0$ ). In Eq. (6), we have restricted the inhibition mechanism to influences from lexical processing of foveal words,  $a_k(t)$ . This is motivated by neurophysiological evidence that the time course of saccade generation arises from the

center of the visual field by a disengagement of the currently fixated site (Munoz & Wurtz, 1993a,b; Wurtz, 1996). Given the central role played by this inhibition process, our model is called SWIFT (Saccade-generation with inhibition by foveal targets). In a more general version of the model, inhibition could be extended to nonfoveal words (e.g., sum of lexical activity within the attentional window).

#### 4.3. Execution of saccades

Finally, we describe the execution of saccades in the model. As the eye is effectively blind during movements, acquisition of new information (i.e. lexical preprocessing) is stopped when saccadic eye movements start. In the case of reading experiments, Wolverton and Zola (1983) showed that eye-contingent display changes during saccades did not disrupt reading. We assume that lexical completion is unaffected during saccades, i.e. lexical completion continues while the eyes are performing saccadic movements.

### 5. Model overview

In this section, we summarize the model assumptions, before presenting results from model simulations in the next section. As already discussed in the introduction, the basic principles that motivated this study are spatially distributed lexical processing (Principle I), a separation between target selection and saccade timing (Principle II), and autonomous saccade generation with inhibition by foveal targets (Principle III). The global structure of SWIFT is summarized in Fig. 3 which

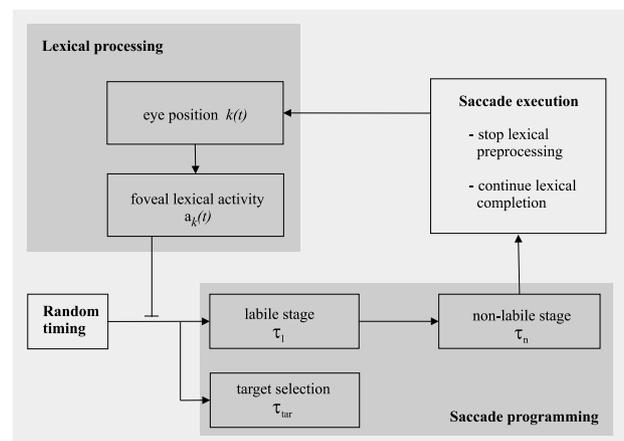


Fig. 3. Schematic diagram of SWIFT. The main subsystems are saccade programming and lexical processing. These two subsystems are coupled via a foveally-inhibited random timing system and a saccade execution system which moves the eyes during saccades.

Table 1  
Model parameters

	Parameter	Symbol	Value	Error <sup>a</sup>	Reference
Lexical parameters	Difficulty, intercept	$\alpha$	148.5	3.6	Eq. (1)
	Difficulty, slope	$\beta$	5.71	0.29	Eq. (1)
Processing rate <sup>b</sup>	Foveal	$\lambda(0)$	0.798	0.017	Eq. (4)
	Parafoveal	$\lambda(1, -1)$	0.077	0.017	Eq. (4)
	Parafoveal	$\lambda(2)$	0.048	0.009	Eq. (4)
	Preprocessing	$f$	62.5	5.8	Eq. (4)
Saccade parameters <sup>c</sup>	Random timing (ms)	$t_s$	187.1	2.6	Eq. (6)
	S.d. (relative)	$\rho$	0.239	0.021	
	Inhibition factor	$h$	50.3	14.1	Eq. (6)
	Labile stage (ms)	$\tau_l$	128.6	3.2	Fig. 2
	Nonlabile stage (ms)	$\tau_n$	41.6	4.7	Fig. 2
	Target selection (ms)	$\tau_{tar}$	112.1	7.2	Fig. 2

<sup>a</sup> Mean error obtained from five runs of the GA parameter estimation.

<sup>b</sup> The sum of the lexical processing parameters over the attentional window was normalized to one. Additionally, we assumed that  $\lambda(-1) \equiv \lambda(1)$  for simplicity. Therefore, the distribution of lexical processing rates is determined by two parameters,  $\lambda(0)$  and  $\lambda(1)$ , where  $\lambda(2) = 1 - \lambda(0) - 2\lambda(1)$ .

<sup>c</sup> Saccade execution  $t_{ex}$  was fixed at 25 ms with a standard deviation of 1/3 of the mean.

provides an overview of how the model acts functionally on the systems level. The main components are saccade programming, saccade execution, and lexical processing.

**Lexical processing:** While eye fixation position  $k(t)$  is at word <sub>$k$</sub>  at time  $t$ , lexical processing evolves according to Eq. (4). The parameter influencing the dynamical behavior is the lexical processing rate  $\lambda(\epsilon)$  as a function of eccentricity  $\epsilon$  (Eqs. (3) and (4)). Lexical activities  $a_n(t)$  are increasing during lexical preprocessing, until the maximum value  $L_n$  is reached. Then, the lexical completion stage starts and is completed when  $a_n(t) = 0$  again (Fig. 1).

**Saccade programming:** A saccadic eye movement program is initiated by an autonomous (random) timing system with an inhibition mechanism as a lexical control process (Eq. (6)). During the labile stage, an additional command to initiate a saccade program may cancel the current program and start a new program, i.e. the time of the labile program is set to zero. After a time interval  $\tau_{tar}$  starting from the initiation of the saccade program (timing scheme, Fig. 2), a saccade target is selected with a probability distribution derived from relative lexical activity  $\pi(n, t|k)$  (Eq. (5)) at time  $t$ . When the labile program terminates after time interval  $\tau_l$ , the nonlabile stage is entered. The saccade will be executed with the target selected after  $\tau_{tar}$ , when the nonlabile stage terminates (after duration  $\tau_n$ ).

**Saccade execution:** The end of the nonlabile stage of the saccade program determines the onset of the saccade, when lexical preprocessing is immediately paused, while ongoing lexical completion is not affected. After a random time interval  $\tau_{ex}$  eye movements stop (end of saccade), the eye fixation position  $k(t)$  is updated, and lexical preprocessing resumes. Lexical processing (Eq. (4)) then evolves under the influence of the new eye position.

## 6. Model simulations

In this section, we present results from numerical simulations with SWIFT. The central aim of our current study is to propose a viable and parsimonious alternative to SAS models of reading, which accounts for more recent physiological and psychological findings (see Introduction). Due to the underlying nonlinear interactions between lexical processing and eye movement control, a mathematical analysis of the model is beyond the scope of the current paper. Instead, we would like to demonstrate that the model reproduces the main statistical features of eye movements in reading. For all numerical investigations, a first-pass analysis of a reading study by Schilling, Rayner, and Chumbley (1998) is used as the empirical basis. Therefore, in principle a quantitative comparison of results presented here with those obtained from E-Z Reader simulations (Reichle et al., 1998) is possible.<sup>8</sup> Here we propose a computational alternative to SAS models, which quantitatively performs in a manner comparable to E-Z Reader, but which offers some qualitative advantages, which will be discussed later. The parameters (Table 1) of the SWIFT model used in the simulations are obtained from a genetic algorithm (GA) optimization technique (see Appendix A). For all words, estimated word frequencies and prediction probabilities were available. In our simulations, we use an integration time step of size  $\Delta t = 1$  ms for the continuous dynamics of lexical processing (Eq. (4)).

<sup>8</sup> Due to differences in parameter optimization procedures and recent modifications proposed for the underlying stochastic algorithms in SAS models (Engbert & Kliegl, 2001), a direct comparison of models as different as E-Z Reader and SWIFT requires a larger computational analysis which will be carried out in a future study.

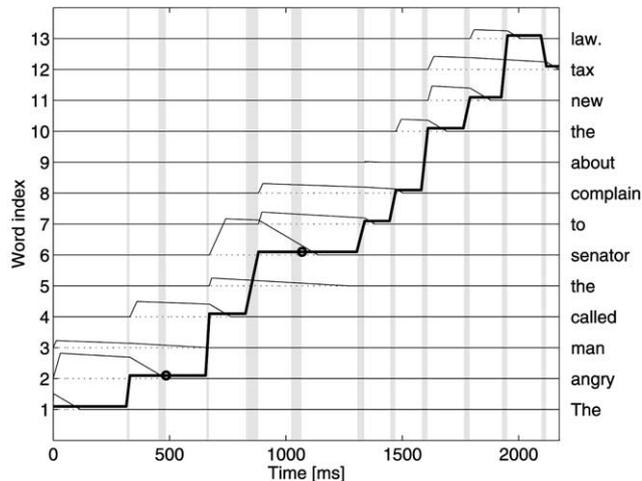


Fig. 4. A trajectory from a simulation of SWIFT. The set of lexical activities  $\{a_n(t)\}$  is plotted over time together with the eye movement position  $k(t)$  (bold line). The execution of saccades is indicated by the shaded regions. Words 3, 5, and 9 are skipped. The beginning of a refixation is indicated by the circle. A regression occurs to word<sub>12</sub>. Epochs with active inhibition (see Footnote 10) are marked by the bold lines on the time axis. The sentence and data on word frequencies and prediction probabilities are taken from the experimental study by Schilling et al. (1998).

### 6.1. Single trajectories

An example for a single run of the model is shown in Fig. 4. The evolution of the whole set of lexical activities  $\{a_n(t)\}$  is plotted over time. The eye's trajectory, i.e. fixation position  $k(t)$ , is given by the bold line. Executions of saccades are indicated by shaded regions. During a saccade, lexical preprocessing is paused and lexical completion is unaffected. We briefly describe some dynamical properties related to word skipping and refixations which are visible in the simulation example (Fig. 4).

**Word skipping:** The model generates two types of word skipping. The first type is illustrated on word<sub>3</sub>. The lexical activity of word<sub>3</sub> approaches zero (the word is completely processed in the parafovea) during fixation of word<sub>2</sub>. Therefore, the probability to select word<sub>3</sub> as a target for the next saccade tends to zero rapidly (Eq. (5)). A similar mechanism underlies skipping of word<sub>9</sub>, with the difference that this word is predictable from context, i.e. the maximum of lexical activity is close to zero (Eq. (2)). The underlying mechanism for this type of word skipping is comparable to the one in SAS models, except that there is no time cost for canceling a saccade (see Discussion). A second type of word skipping, however, occurs for word<sub>5</sub>. Here, the word is skipped due to the stochastic target selection process. Although the lexical activity of word<sub>5</sub> is nonzero, there is a higher probability to select word<sub>6</sub> as a target for the next saccade. A comparable skipping occurs for word<sub>12</sub>.

*Refixations* share the same underlying mechanism with forward saccades (and regressions) due to our target selection concept. When the eyes arrive at word<sub>2</sub> or word<sub>6</sub>, there is still lexical processing to do, since  $a_2, a_6 > 0$ . By chance, these words are selected as saccade targets, i.e. a refixation is prepared and (after saccade latency) executed.<sup>9</sup>

### 6.2. Statistical evaluation of model simulations

Stochasticity plays an important role in the control of eye movements. Random influences arise from several sub-processes: (i) saccade target selection, (ii) autonomous initiation of saccade programs, (iii) lexical processing time, (iv) saccade latencies, and (v) saccade execution. Following Reichle et al. (1998), the influence of noise can be included in lexical processing and saccade timing pattern during labile,  $\tau_l$ , and nonlabile,  $\tau_n$ , stages as well as during target selection,  $\tau_{tar}$  and lexical processing difficulty  $L_n$  (Eqs. (1) and (2)). The relation between mean values and standard deviations of the stochastic distributions are described by a single parameter  $\rho$ .

We present some results from a statistical analysis of simulations of the SWIFT model (Fig. 5). For the calculation of statistical averages, all 536 words of the corpus of 48 sentences (Schilling et al., 1998) were divided into five classes (see Appendix A). In the simulations, 1000 statistical realizations of eye movement trajectories (model runs) for the complete corpus of sentences (Schilling et al., 1998) were used for calculating statistical averages. Model parameters (Table 1) used in the numerical simulations were obtained by a GA optimization procedure (see Appendix A).

In Fig. 5(a), we compare mean first fixation durations, gaze duration and single fixation duration with the experimental results. The main effect of word frequency can be found in all three measures. Aside from lexical difficulty and predictability (Eqs. (1) and (2)), the dependence of fixation duration on word frequency is produced by the interplay of autonomous saccade timing with inhibition by foveal lexical difficulty (Eq. (6)). This mechanism will be studied in more detail in the next section.

The main patterns of fixation sequences and word skippings are displayed in Fig. 5(b). There are two main empirical regularities. First, probability for word skipping increases for increasing word frequency. In many cases, high-frequency words are lexically processed

<sup>9</sup> In the example shown here, the lexical completion process for word<sub>2</sub> has already terminated, when refixation of the word<sub>2</sub> starts. This is an interesting consequence of the delayed target selection process (with time delay  $\tau_{tar}$ , Fig. 2).

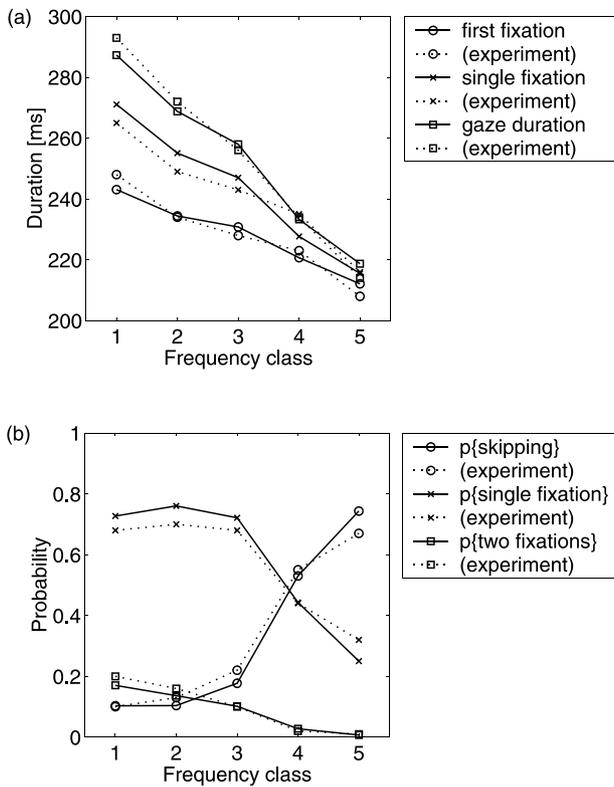


Fig. 5. Statistical evaluation of SWIFT performance. (a) First fixation duration, gaze duration, and single fixation duration as a function of word frequency class (averaged over 1000 statistical realizations from SWIFT simulations, i.e. 1000 simulations of the model over the same corpus of sentences but with different pseudo-random numbers). (b) Probabilities for word skipping, performing a single fixation, and making two fixations (computed from the same model runs as in (a)) as a function of word frequency class.

completely in the parafovea, i.e. these words are not fixated. Second, many difficult (low-frequency) words are processed with more than one fixation. As a consequence, the probability for making two fixations on a word increases with decreasing word frequency. In the model, this effect is a direct consequence of the assumption of autonomous saccade timing. During a fixation the initiation of a new saccade program is inhibited as long as foveal lexical activity  $a_k(t)$  is high (Eq. (6)). For difficult words, however, a new saccade program is initiated, although lexical completion of the currently fixated word is not finished. Due to the random target selection assumption (Eq. (5)), there is, in this case, a high probability to select a difficult word as a target for refixation.

Using numerical explorations of our model with different sets of parameters, we found that the patterns of fixations are robust with respect to parameter changes. In summary, measures for fixation durations as well as measures of the fixation patterns are in good agreement with experimental data.

### 6.3. The role of inhibition of saccade timing

At first sight, the assumption of autonomous saccade timing seems to be a rather strong constraint for explaining the complicated word frequency effects in eye movement patterns. Therefore, we analyzed the role of the inhibition mechanism for reproducing the main properties of the experimental data. An important experimental result was that first fixation duration depended on word frequency. This effect cannot be explained by autonomous saccade timing, which is demonstrated with numerical simulations of the model without inhibition ( $h = 0$ ) in Fig. 6(a), where we observe a shift of all variables to smaller values. The most important qualitative difference to Fig. 5(a), however, is that first fixation durations are no longer a function of word frequency. This dependence cannot be reproduced by random timing of saccades. Despite autonomous saccade timing, however, gaze duration, i.e. the sum of first fixation and refixations, is qualitatively in good agreement with the experimental data, since first fixation duration and refixations of random length sum up to the

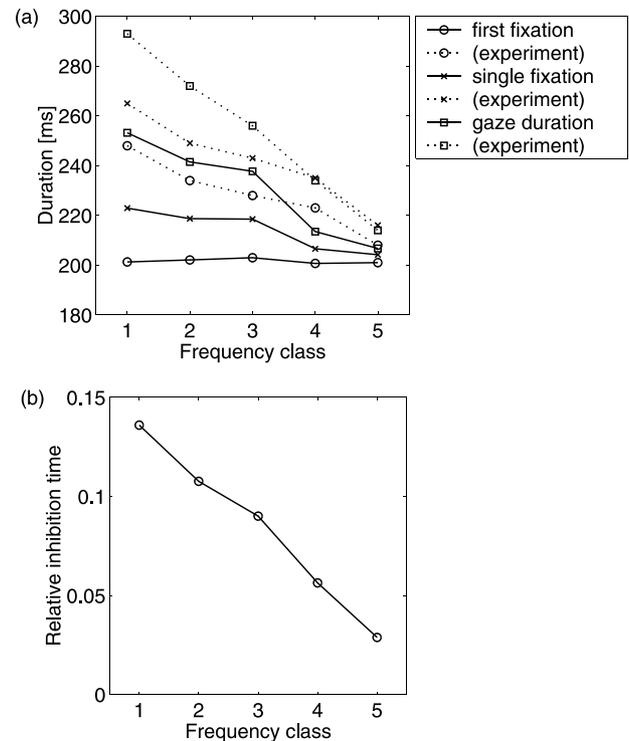


Fig. 6. Analysis of inhibition in SWIFT. (a) Model results as in Fig. 5(a), but without inhibition of saccade program initiation, i.e.  $h = 0$ . In this case, first fixation duration does not depend on word frequency. (b) Fraction of fixation time where inhibition was active (computed from the same realizations as in Fig. 5) as a function of word frequency class. The results show that, even for low-frequency words, inhibition represents only a slight modification of autonomous (random) saccade timing, since the proportion of time with inhibition is below 15% of fixation time.

required fixation time. More interestingly, single fixation duration also shows a strong dependence on word frequency. The explanation is that the corresponding curve in Fig. 6(a) represents data from all fixations for which random (autonomous) saccade timing fitted the required processing time by chance. In summary, the most important role of the inhibition mechanism is to explain the word-frequency dependence of first fixation durations in a framework of minimal eye movement control.

Concerning the debate on possible low-level control processes for eye movements during reading (Deubel et al., 2000) it is interesting to analyze how inhibition works in SWIFT. If inhibition would be active most of the time during fixation, the saccade timing mechanism could no longer be looked upon as an autonomous random process. Using numerical simulations (for model parameters see Table 1), we show that inhibition modifies the autonomous saccade timing only slightly (Fig. 6(b)). Even for low-frequency words, the proportion of time with active inhibition<sup>10</sup> is on average below 15% of the fixation duration. Therefore, despite the inhibition process, the model may still be called an autonomous and random saccade timing model. In summary, SWIFT is able to explain the experimentally observed fixation patterns with an autonomous saccade timing mechanism, inhibited by foveal word difficulty. These results show that a low-level control strategy is psychologically plausible if we permit high-level control processes to correct “dumb” default strategy during less than 15% of fixation time.

#### 6.4. Progress with regressions?

Resulting from the three basic Principles I–III, which we have used to derive our model, complex eye movement patterns are inherent to its dynamical behavior. During reading, about 10–15% of all fixations are regressions. Following Liversedge and Findlay (2000), one of the outstanding problems of eye movement research is what causes regressions. It is well known that text difficulty strongly influences the number of regressions (Rayner, 1998). Our theoretical study suggests that a portion of regressions might arise from incomplete lexical processing. An example for such a more complex eye movement trajectory is given in Fig. 7.<sup>11</sup> Two different types of regression can occur in our model. First, due to the target selection mechanism, any word within the attentional window is a potential saccade target, if its lexical activity  $a_n(t)$  is nonzero. The regression to word<sub>2</sub> is of this type, which we call “local” regression

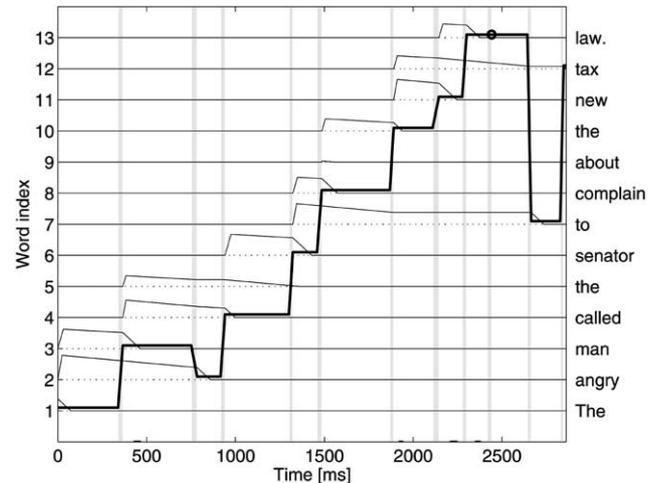


Fig. 7. Different types of regressions in SWIFT (see legend of Fig. 4 for an explanation of the plot). The first regression to word<sub>2</sub> is performed within the attentional window (a “local” regression); the second (“global”) regression to word<sub>7</sub> occurs in a later stage, when the eye position is already close to the end of the sentence. In SWIFT, both types of regressions are a consequence of incomplete lexical processing of word<sub>2</sub> and word<sub>7</sub>, respectively.

due to its limited length. In the E-Z Reader models, a local regression can occur as a special case of a re-fixation (Reichle et al., 1998). A second type, called “global” regression, can happen in our model at any time, if a word has not been processed completely (word<sub>6</sub> in the example) before it has exited the attentional window. Note that as soon as a word is out of the attentional window, the degree of incompleteness of lexical access is maintained. Furthermore, the relative contribution of its lexical activity, which is important for target selection (Eq. (5)) increases over time. It is an interesting prediction of our model that regressions may be due to incomplete lexical processing.

## 7. Discussion

We have presented a new model of eye movement control in reading on the assumption of GAG. The model, called SWIFT, operates with autonomous saccade timing, inhibited by lexical processing load in the fovea (Principle III). The main principles of the model are (Principle I) that lexical processing is spatially distributed over an attentional window and (Principle II) that saccade timing (“when”) is separated from saccade target selection (“where”). Spatially distributed lexical processing is suggested from recent experimental findings (e.g., Inhoff, Radach et al., 2000; Inhoff, Starr et al., 2000; Kennedy, 2000a,b). The separation of target selection from saccade timing is motivated by behavioral (Becker & Jürgens, 1979; Findlay & Walker, 1999) as well as neurophysiological findings (Carpenter, 2000, and refs. therein).

<sup>10</sup> Inhibition is active, if  $t_s < t < t_s + ha_k$  (see Eq. (6)).

<sup>11</sup> The differences between Figs. 4 and 7 arise from the stochasticity described earlier. The same values of the model parameters are used throughout this study.

As an inherent dynamical property, this model provides a common mechanism for forward saccades, refixations, and regressions. The model was developed as an alternative to SAS models, in particular the E-Z Reader model (Reichle et al., 1998; Reichle et al., 1999) and a later variant of this approach (Engbert & Kliegl, 2001). For ease of comparison we used the same corpus of data that was used in these earlier papers to validate the performance of the model. The model reproduces relations between word frequency class and means of durations of single fixations, first fixations and gaze as well as the probabilities of word-to-word, word skipping and refixation movements.

SWIFT goes beyond previous SAS models by providing a common mechanism for forward and backward movements and for processing words to the left of fixation. First, the occurrence of all types of eye movements, including refixations and regressions, is a natural consequence of Principles I and II. In the course of dynamics of processing words within the attentional window, any incompletely processed word is a potential target for the next saccade. Moreover, according to the model there is the possibility of a second type of regression to any incompletely processed word to the left of the fixation, including words that are outside the left border of the attentional window. The reason for this is that once the attentional window has moved beyond a certain word, there is no further lexical processing of this word. However, the word remains a potential target for a saccade with the current level of activation (lexical activity). Usually, this activation will be very low, and consequently, long regressions will be more likely towards the end of the sentence. Thus, we argue that, in addition to high-level control strategies (such as re-parsing which are not part of the current model), long regressions may also occur for purely lexical processing needs. This prediction should be experimentally testable in future work. We would like to mention that the model's behavior is in agreement with observations by Kennedy and Murray (1987) that there is a very precise coding of the spatial location of previous words.

Second, an empirical issue addressed with post-hoc analyses of the simulation data relates to the fixation durations prior to word skipping. The model does not show the large increase of fixation durations prior to word skipping reported for SAS models as a consequence of saccade canceling (Reichle et al., 1998: 100 ms in E-Z Reader 3, 173 ms in E-Z Reader 5; Engbert & Kliegl, 2001: 82 ms). Increased fixation duration before skipping is a qualitative property for SAS models, since a saccade target cannot be modified after initiation of a saccade program. The only way to change a target is to cancel the saccade and start a new saccade—with the cost of an increased latency. This issue is empirically controversial, but the size of the effect hints to a strong advantage for the SWIFT framework. Pollatsek, Ray-

ner, and Balota (1986) reported that fixations prior to word skipplings were 21 ms longer and Reichle et al. (1998) report an effect of 38 ms. Radach and Heller (2000) argued that there is no significant increase in fixation time but the small positive effects observed could be due to the selection of data without regressions. In our simulations based on the same material, we observed an increase of 17 ms for fixations prior to word skipplings with a maximum value of 25 ms for fixations on high frequency words prior to a word skipping in the SWIFT model.

The concept of lexical activities described here shows qualitative similarity to the *dynamic neural field theory* (Kopeck & Schöner, 1995; Schöner, Kopeck, & Erlhagen, 1997; Thelen et al., 2001). The basic underlying concept of this theory is the integration of sensory and memory information with movement preparation. In this respect, the inclusion of the evolution of lexical activities for target selection or, more generally, movement planning, may be looked upon as a simplified version of a dynamic neural field model. The theory of dynamic neural fields suggests that a considerable amount of the observed variability in motor behavior arises from the “coupling” between sensory information and movement preparation. The nature of this coupling might turn out to be less deterministic than previously suggested (Thelen et al., 2001). From this perspective, our model of eye movement control may be looked upon as a further example for such a complex interaction between sensory information and movement planning. As already discussed above, an inherent result from this interaction is the occurrence of long-range regressions.

Theoretical models may give new insights into the current debate on eye movement control strategies in reading (Starr & Rayner, 2001). While there is evidence for word frequency effects in statistical measures of eye movements (Rayner, 1998), the underlying control processes might be dominated by a “dumb” default strategy, that is, for example, random processes with a pre-defined distributions of inter-saccadic time intervals and saccade lengths (Reilly & O'Regan, 1998; McConkie et al., 1988). First, it is important to note that lexically-driven control processes, which are characteristic of SAS models, can be combined with random influences from a low-level default strategy. As an example, we have shown in a previous study that SAS models may be extended to include autonomous saccades which are initiated randomly without reference to the actual state of lexical processing (Engbert & Kliegl, 2001). From this result we conclude that the problem of drawing the distinction between the two basic control strategies, lexically-driven versus low-level default control, might be ill-posed.

Second, SWIFT provides an interesting new alternative to SAS models. Basically, the model combines

random saccade timing with a target selection process that is also random but whose probability distribution changes over time as a consequence of lexical processing.<sup>12</sup> We have introduced an inhibition process which corrects the pure random saccade timing to explain the observed word-frequency dependencies, in particular, the one for first fixation durations. Using numerical simulations we have also shown that the proportion of time during which the inhibition process works is below 15% of the fixation durations. In summary, our results suggest that a “dumb” default strategy with respect to saccade timing can indeed explain a large portion of statistical aspects of eye movements if lexically driven control processes are allowed to intervene from time to time. Nevertheless, our results also suggest that some strategy of highly efficient control (such as the foveal inhibition by lexical difficulty in the present model) needs to be incorporated in any model of eye-movement control during reading.

### Acknowledgements

We thank Jürgen Kurths for his contribution to establish our collaboration, Eric D. Reichle for providing us with the corpus of sentences for our simulations, Ralph Radach, Keith Rayner, Jochen Laubrock, Klaus Oberauer, and Ellen Grabner for valuable comments on the manuscript. This work was supported by Deutsche Forschungsgemeinschaft (DFG grants KL 955/3-1 and -2) and by NSERC (Canada). RE thanks the Department of Physics, University of Ottawa, for the hospitality during his stay, and the Centre de Recherches Mathématique, Université de Montréal, Canada, for additional financial support.

### Appendix A. Parameter estimation by genetic algorithm

For the estimation of model parameters we used a GA approach (Holland, 1992; Mitchell, 1996).<sup>13</sup> For each sentence, 1000 stochastic realizations of the model were run with a new set of pseudo-random numbers. For the GA we used a population of 50 combinations of parameter values which were iterated over approximately 600 generations. Several runs of the GA were used to test the reliability of the estimates for the model

parameters and to compute the mean errors of parameters (Table 1). A separate simulation was performed to produce the data shown in Fig. 5. A single run for the GA parameter estimation method took approximately 100 h of CPU time on a SUN Ultra 10 computer. The 536 words of the corpus (48 sentences) were divided into five different frequency classes, as suggested in Reichle et al. (1998). The frequencies per million in the different classes were 0–10 (class 1), 11–100 (class 2), 101–1000 (class 3), 1001–10,000 (class 4), 10,001+ (class 5). Cell counts in these classes ranged from 90 to 134.

The performance of the model is defined as a root mean square of the normalized errors of fixation durations and probabilities. Fixation durations obtained from model simulations are denoted by  $T_k^j$ , where the three different types of fixation durations (first fixation duration, single fixation duration and gaze duration) are labeled by the superscript  $j$  and the subscript  $k$  indicates the five frequency classes;  $\sigma(T_k^j)$  is the standard deviation of the distribution of fixation durations over all realizations. The experimentally observed value is given by  $\bar{T}_k^j$ . The deviation of simulated mean fixation durations from observed mean fixation durations can be written as

$$\Delta_T = \sqrt{\sum_{j=1}^3 \sum_{k=1}^5 \left( \frac{T_k^j - \bar{T}_k^j}{\sigma(T_k^j)} \right)^2}. \quad (\text{A.1})$$

The three different fixation probabilities, for skipping, single fixation, and two fixations, are denoted by  $p_k^j$  (with sub- and superscript as for the fixation durations). As estimates of standard deviations of probabilities, we use  $\sigma(p_k^j) \equiv \sqrt{p_k^j(1-p_k^j)}$ . The corresponding root mean square error is

$$\Delta_P = \sqrt{\sum_{j=1}^3 \sum_{k=1}^5 \left( \frac{p_k^j - \bar{p}_k^j}{\sigma(p_k^j)} \right)^2}. \quad (\text{A.2})$$

where  $\bar{p}_k^j$  are the experimentally observed probabilities. Adding the two error terms, Eqs. (A.1) and (A.2), gives the loss function that is minimized<sup>14</sup> by the GA method,

$$\Delta = \kappa \Delta_T + \Delta_P, \quad (\text{A.3})$$

where deviations in the mean fixation durations are weighted by a factor of  $\kappa = 50$ . The weighting turned out to be useful, since estimates of mean fixation durations appear to be more difficult to adjust to the empirical values than the pattern of fixation probabilities, which turn out to be a rather robust property of the SWIFT model.

<sup>12</sup> In preliminary simulations we established that alternative target-selection rules, e.g., selection of words with highest lexical activity or random selection among all lexically active words, i.e.  $a_m(t) > 0$ , can qualitatively reproduce the results in Figs. 6 and 7. More experimental work will be needed to determine which type of target selection is most appropriate. Model performance did depend critically on eccentricity-dependent lexical processing rates.

<sup>13</sup> This method was previously applied to obtain parameters of an SAS model (Engbert & Kliegl, 2001).

<sup>14</sup> The value  $F = 1/\Delta$  can be interpreted as a fitness function that is usually maximized in genetic algorithms (Mitchell, 1996).

## References

- Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17, 364–390.
- Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by mean of double step stimuli. *Vision Research*, 19, 967–983.
- Binder, K. S., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1162–1172.
- Carpenter, R. H. S. (2000). The neural control of looking. *Current Biology*, 10, R291–R293.
- Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research*, 38, 2949–2962.
- Clark, J. J. (1999). Spatial attention and latencies of saccadic eye movements. *Vision Research*, 39, 585–602.
- Coëffé, C., & O'Regan, J. K. (1987). Reducing the influence of non-target stimuli on saccade accuracy: Predictability and latency effects. *Vision Research*, 27, 227–240.
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, 36, 1827–1837.
- Deubel, H., O'Regan, J. K., & Radach, R. (2000). Attention, information processing, and eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.
- Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: A possible role for autonomous saccades. *Biological Cybernetics*, 85, 77–87.
- Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, 22, 661–721.
- Gillespie, D. T. (1978). Monte-Carlo simulation of random walks with residence time dependent transition probability rates. *Journal of Computational Physics*, 28, 395–407.
- Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 417–429.
- Hogaboam, T. W. (1983). Reading patterns in eye movements. In K. Rayner (Ed.), *Eye movements in reading*. New York: Academic Press.
- Holland, J. H. (1992). *Adaptation in natural and artificial systems*. Cambridge (Mass.): MIT Press.
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431–439.
- Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.
- Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics*, 62, 1474–1484.
- Jackson, E. A. (1991). *Perspectives of nonlinear dynamics*. Cambridge, UK: Cambridge University Press.
- Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade-off. In G. Underwood (Ed.), *Eye guidance in reading and scene perception*. Oxford: Elsevier.
- Kennedy, A. (2000a). Parafoveal processing in word recognition. *The Quarterly Journal of Experimental Psychology*, 53A, 429–455.
- Kennedy, A. (2000b). Attention allocation in reading: Sequential or parallel. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.
- Kennedy, A., & Murray, W. S. (1987). Spatial coding and reading: Some comments on Monk (1985). *Quarterly Journal of Experimental Psychology*, 39A, 649–718.
- Kennedy, A., Radach, R., Heller, D., & Pynte, J. (Eds.). (2000). *Reading as a perceptual process*. Oxford: Elsevier.
- Kopeck, K., & Schöner, G. (1995). Saccadic motor planning by integrating visual information and pre-information on neural dynamic fields. *Biological Cybernetics*, 73, 49–60.
- Kowler, E., Anderson, E., Doshier, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35, 1897–1916.
- Kustov, A. A., & Robinson, D. L. (1996). Shared neural control of attentional shifts and eye movements. *Nature*, 384, 74–77.
- Legge, G. E., Mansfield, J. S., & Chung, S. T. L. (2001). Psychophysics of reading XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, 41, 725–743.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review*, 104, 524–553.
- Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science*, 4, 6–14.
- McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research*, 28, 1107–1118.
- McConkie, G. W., Kerr, P. W., & Dyre, B. P. (1994). What are 'normal' eye movements during reading: Toward a mathematical description. In J. Ygge, & Lennerstrand (Eds.), *Eye movements in reading*. Oxford: Elsevier.
- Mitchell, M. (1996). *An introduction to genetic algorithms*. Cambridge (Mass.): MIT Press.
- Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 667–682.
- Munoz, D. P., & Wurtz, R. H. (1993a). Fixation cells in monkey superior colliculus: I. Characteristics of cell discharge. *Journal of Neurophysiology*, 70, 559–575.
- Munoz, D. P., & Wurtz, R. H. (1993b). Fixation cells in monkey superior colliculus: II. Reversible activation and deactivation. *Journal of Neurophysiology*, 70, 576–589.
- Murray, W. S. (1998). Parafoveal pragmatics. In G. Underwood (Ed.), *Eye guidance in reading and scene perception*. North-Holland: Amsterdam.
- O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Reviews of Oculomotor Research, vol. 4. Eye movements and their role in visual and cognitive processes*. Amsterdam: Elsevier.
- O'Regan, J. K. (1992). Optimal viewing positions in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading*. New York: Springer.
- O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance. XII. The psychology of reading*. England, Erlbaum: Hove.
- Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception and Psychophysics*, 40, 123–130.
- Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.

- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled reading. *Journal of Experimental Child Psychology*, 41, 211–236.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movements control. *Vision Research*, 38, 1129–1144.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125–157.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movements control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, 39, 4403–4411.
- Reilly, R. G., & O'Regan, J. K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research*, 38, 303–317.
- Rizzolatti, G. (1983). Mechanisms of selective attention in mammals. In J. P. Ewart, R. Capranica, & D. J. Ingle (Eds.). New York: Plenum.
- Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26, 1270–1281.
- Schöner, G., Kopeck, K., & Erlhagen, W. (1997). The dynamic neural field theory of motor programming: Arm and eye movements. In P. G. Morasso, & V. Sanguineti (Eds.), *Self-organization, computational maps, and motor control*. Amsterdam: Elsevier-North Holland.
- Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal Load and parafoveal processing. *Quarterly Journal of Experimental Psychology*, 52A, 1021–1046.
- Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences*, 5, 156–163.
- Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In E. Kowler (Ed.), *Reviews of oculomotor research, vol. 4. Eye movements and their role in visual and cognitive processes*. Amsterdam: Elsevier.
- Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24, 1–86.
- Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process*. Oxford: Elsevier.
- Wolverton, G. S., & Zola, D. (1983). The temporal characteristics of visual information extraction during reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes*. New York: Academic Press.
- Wurtz, R. H. (1996). Vision for the control of movements. *Investigative Ophthalmology & Visual Science*, 37, 2131–2145.
- Wolfe, J. M., Alvarez, G. A., & Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature*, 406, 691.