Biol. Cybern. 85, 77–87 (2001)

Biological
Cybernetics

Mathematical models of eye movements in reading: a possible role for autonomous saccades

Ralf Engbert[1,2], **Reinhold Kliegl**[1]

[1] Institut für Psychologie, Universität Potsdam, PF 601553, 14415 Potsdam, Germany
[2] Zentrum für Dynamik komplexer Systeme, Universität Potsdam, PF 601553, 14415 Potsdam, Germany

Received: 21 March 2000 / Accepted in revised form: 10 January 2001

Abstract. An efficient method for the exact numerical simulation of semi-Markov processes is used to study minimal models of the control of eye movements in reading. When we read a text, typical sequences of fixations form a rather complicated trajectory – almost like a random walk. Mathematical models of eye movement control can account for this behavior using stochastic transition rules between few discrete internal states, which represent combinations of certain stages of lexical access and saccade programs. We show that experimentally observed fixation durations can be explained by residence-time-dependent transition probabilities. Stochastic processes with this property are known as semi-Markov processes. For our numerical simulations we use the minimal process method (Gillespie algorithm), which is an exact and efficient simulation algorithm for this class of stochastic processes. Within this mathematical framework, we study different forms of coupling between eye movements and shifts of covert attention in reading. Our model lends support to the existence of autonomous saccades, i.e., the hypothesis that initiations of saccades are not completely determined by lexical access processes.

1 Introduction

Investigating eye movements in reading may be looked upon as a case study for the more general problem of scanning of visual scenes with higher structural complexity. Typical eye movements in reading form a rather complicated trajectory – almost like a random walk (van Kampen 1981; Gardiner 1990; Gillespie 1992). For theoretical approaches to the control of eye movements, this random walk can be approximated by a series of fixations. As an illustrative comparison (Reichle et al. 1998), we may think of reading as a "slide show" where the "slides" (words) are fixated for about 200 to 250 ms, separated by executions of saccades (20 to 40 ms).

The experimentally observed random walk over words is the consequence of saccades from word$_n$ to word$_k$. For $k = n$, word$_n$ is *refixated*, which is typically observed for low-frequency words. If $k > n + 1$, word$_{n+1}$ is *skipped*. Word skipping is very likely for highly frequent words like articles. Considering only forward saccades and refixations, $k \geq n$, is motivated by the hypothesis that this form of eye movements may represent a "default" model of eye movement control. Using this conjecture, a successful class of models, the so-called E-Z Reader 1 to 5, has been proposed recently (Reichle et al. 1998). This model has also been extended to include initial fixation locations and refixations (Reichle et al. 1999). In mathematical models of this class, a number of internal states is used to represent combinations of different stages of lexical access and saccade programs. A stochastic sequence of internal states is related to a certain sentence. The sequence is generated by stochastic transition rules to account for the observed statistical properties of eye movements.

In this paper, we propose a "minimal" model of eye movement control in reading to explore modifications with respect to different assumptions of the E-Z Reader modeling framework (Reichle et al. 1998). Since all our alternative assumptions relate to a fundamental level of model design, the analysis is restricted to the same level of approximations as E-Z Reader 1. As a consequence, experimental data used here are *gaze durations*, defined as the sum of first fixation duration and all potential refixations, and probabilities for word skipping. Possible extensions of our model are proposed in Sect. 5.

The model proposed here differs in three major aspects from the E-Z Reader model. First, a fundamental assumption in mathematical models of eye movement control is related to the coupling of two subsystems: eye movements and shifts of visual attention. In the E-Z Reader models, lexical processes induce shifts of covert attention. The end of a familiarity check, i.e., a

Correspondence to: R. Engbert
(e-mail: engbert@rz.uni-potsdam.de)

preprocessing stage of lexical access, initiates a saccadic motor program, while a shift of (covert) attention to the next word is induced by full lexical access. In our model, we assume that both the saccadic motor program and the covert shift of attention are initiated at the end of lexical access of the foveal word, i.e., a *common* mechanism, which synchronizes both subsystems.

Second, in the E-Z Reader models, saccade initiation is completely governed by lexical processes. There is no saccade without successful preprocessing. In our model, we investigate the consequences of relaxing this hypothesis of a strong coupling between lexical access and eye movements. In particular, we demonstrate that autonomous saccades, i.e., saccade programs, which are initiated without a lexical trigger signal, will lead to rich and psychologically plausible dynamic behavior in a minimal model.

Finally, the third difference to the E-Z Reader framework is methodologically motivated. We use a generalized approach for stochastic transition rules, which is based on the concept of transition probability rates (van Kampen 1981; Gardiner 1990). Using this framework, we replace assumptions on the distribution of residence times – like gamma distributions (Reichle et al. 1998) – by an equation for transition probability rates on the level of internal states of the model. In particular, we show that the experimentally observed statistical properties of eye movements in reading can be explained by semi-Markov processes (Gillespie 1978) with residence-time-dependent transition probability rates.

We start with an introduction to semi-Markov processes and its exact numerical simulation in Sect. 2. In Sect. 3 we focus on a simple model for word skipping, which is then extended to modeling distributions of fixation duration (Sect. 4). In Sect. 5 we will address some other differences between our model and the E-Z Reader model.

2 Semi-Markov processes

In the class of models of eye movement control discussed here, a finite number of internal states $S_1, S_2, S_3, \ldots$ is used to describe different stages of processing of words and eye movements. Since transitions between adjoining states are defined by stochastic transition rules, a "random walk" over the internal states is performed when we use such a model to process a sentence. Different runs of the model yield different realizations of both the internal random walk and the observed random walk over words (i.e., the series of fixations).

2.1 Residence-time-dependent transition probability rate

An important concept for stochastic models of random processes is the transition probability rate. To start with a general framework for stochastic transitions, we use the following transition rule (Gillespie 1978): if the system is in the state S_m at time t, having arrived there at time $t-\tau (\tau \geq 0)$, the probability that it will step to some other state S_n in the next infinitesimal time interval $(t, t+\mathrm{d}t)$ is

$$W_{nm}(\tau)\mathrm{d}t \quad . \tag{1}$$

The most common assumption is that the transition probability rate does not depend on residence time τ, i.e., $W_{nm}(\tau) = W_{nm} = \text{const.}$ In this case the random walk is a Markov process (van Kampen 1981). The more general case (1), where the transition probability rate depends on residence time, is referred to as a semi-Markov process. Some techniques for the analysis of Markov processes can be exploited for the study of semi-Markov processes. As an example, Gillespie (1977) derived a generalized master equation for these processes. For numerical simulations of semi-Markov processes, an exact and efficient algorithm has been developed (Gillespie 1978). In the following we discuss why we need to implement residence-time-dependent transition probabilities in our model of eye movements in reading.

2.2 Pausing-time distribution

A fundamental concept for Monte Carlo simulation techniques is the transition probability density function $P(\tau, n|m, t)$. The probability that the system steps *next* to state S_n in the time interval $(t+\tau, t+\tau+\mathrm{d}\tau)$, if it arrived in the state S_m at time t, is given by $P(\tau, n|m, t)\mathrm{d}\tau$. The transition probability density function can be written as a product of two other functions,

$$P(\tau, n|m, t) = \pi(n, m)\psi(\tau|m) \quad , \tag{2}$$

where $\pi(n, m)$ is the stepping probability from state S_m to state S_n, and $\psi(\tau|m)$ is the probability density function for the pausing time τ in the state S_m.

The stepping probability can be calculated from the relative transition probabilities at time τ,

$$\pi(n, m) = \frac{W_{nm}(\tau)}{W_m(\tau)} \quad . \tag{3}$$

The probability density function for the pausing time is obtained from the transition probability rate (1),

$$\psi(\tau|m) = W_m(\tau) \exp\left\{ -\int_0^\tau W_m(\tau')\mathrm{d}\tau' \right\} \quad , \tag{4}$$

where we used the definition $W_m(\tau) = \sum_n W_{nm}(\tau)$ of the transition probability for a transition from S_m to any other state S_n (for details see Gillespie 1978).

In the case of a Markov process, where $W_m(\tau)$ is a constant, the pausing time is exponentially distributed: $\psi(\tau|m) = w_m \mathrm{e}^{-w_m \tau}$. As an important property of the exponential distribution, the maximum of ψ is at $\tau = 0$. The residence times of different internal states S_j (with sub-processes $S_k \mapsto S_l$) sum to fixation durations. Therefore, corresponding distributions of fixation durations are qualitatively of the exponential type. A typical experimentally observed distribution of fixation durations

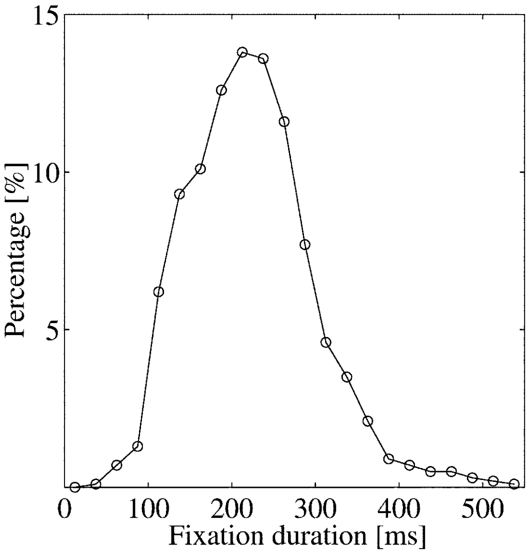

Fig. 1. A typical frequency distribution of fixation durations in reading (after Rayner 1998)

shows, however, a sharp maximum of relative frequency (probability) at 200 ms (Fig. 1). Therefore, the experimentally observed distribution of fixation durations cannot be explained by a Markov model in a straightforward way.

To avoid these restrictions of the Markov case, we use the more general approach for the transition probability rate (1). Compared to a constant W_m, we now discuss the next more complicated case: a transition probability which increases linearly with τ. As an additional parameter, we introduce a refractory period τ_0 with vanishing transition probability rate. These two assumptions are described by

$$W_m(\tau) = \begin{cases} 0 & \text{if } \tau < \tau_0 \\ w_m(\tau - \tau_0) & \text{if } \tau \geq \tau_0 \end{cases} \quad . \tag{5}$$

Following the general relation (4) between transition probability rate and pausing time distribution, (5) leads to the following probability density function for the pausing time,

$$\psi(\tau|m) = \begin{cases} 0 & \text{if } \tau < \tau_0 \\ w_m(\tau - \tau_0)\exp\left(-\frac{w_m}{2}(\tau - \tau_0)^2\right) & \text{if } \tau \geq \tau_0 \end{cases} \quad , \tag{6}$$

which is a single-humped distribution and qualitatively in agreement with the experimentally observed percentage of fixation durations (Fig. 1). In the following, the refractory time τ_0 is written as a proportion of mean pausing time μ_τ

$$\tau_0 = \phi\mu_\tau \quad \text{with} \quad 0 \leq \phi < 1 \quad . \tag{7}$$

Some remarks on this assumption are necessary: Our approach (7) permits independent variation of the variance and mean value of the residence time. Modifications to this relation could be necessary to capture additional statistical measures. Alternatively, a specification of additional states of the model could lead to comparable results. In the case of a fixed value of ϕ, (7) is in agreement with the assumptions of Reichle et al. (1998) in the E-Z Reader models (see below).

Given a mean value μ_τ of the pausing time, we have to compute a corresponding value for the total transition probability rate W_m for state S_m to all adjoining states S_n. Using the probability distribution function for the pausing time (6), we calculate its mean value,

$$\mu_\tau = \tau_0 + \left(\frac{\pi}{2w_m}\right)^{1/2} \quad , \tag{8}$$

from which parameter w_m can be read directly,

$$w_m = \frac{\pi/2}{(\mu_\tau - \tau_0)^2} = \frac{\pi/2}{\mu_\tau^2(1-\phi)^2} \quad . \tag{9}$$

We now investigate how variation of ϕ influences the relation between the mean value μ_τ and the standard deviation σ_τ of the pausing time τ. Calculation of the second moment of the distribution (6) and substracting the square of the mean gives

$$\sigma_\tau^2 = E(\tau^2) - \mu_\tau^2 = \frac{4-\pi}{2w_m} \quad . \tag{10}$$

The variance does not depend on ϕ (or τ_0), since the refractory time simply shifts the distribution of τ. The ratio of standard deviation (from Eq. 10) to mean (Eq. 8) is independent of the transition probability parameter w_m, i.e.,

$$\frac{\sigma_\tau}{\mu_\tau} = \left(\frac{4-\pi}{\pi}\right)^{1/2} \cdot (1-\phi) \approx 0.52 \cdot (1-\phi) \quad . \tag{11}$$

Since in the E-Z Reader models σ_τ/μ_τ is fixed at one-third (Reichle et al. 1998), the corresponding value of ϕ is 0.36. Given a certain mean value for the pausing time in state S_m, and a relation between standard deviation and mean value, we can calculate the model parameters w_m and ϕ analytically using (9) and (11).

As stated before, the type of distributions (6) for the sub-processes determines the distribution of fixation durations. It is shown in Sect. 4 that (5) leads to realistic predictions on the distributions of fixation durations. In Sect. 3 we review the basic mechanisms for word skipping.

3 Modeling word skipping

The simplest model of how lexical access drives eye movements is a strictly serial one. When a currently fixated word is lexically processed, a saccade program to the next word is initiated. Its termination signals the execution of the saccade. This assumption leads to strong constraints on the available time for lexical access. The reason for a rather limited time for lexical access is that in such a model the lexical access time and the time required for programming a saccade to the next word sum to the fixation duration of a word. Therefore, the lexical access time can be calculated from the difference of fixation duration and saccade program time. From the viewpoint of adaptivity, parallel processing of lexical access and saccade program to the next word would be much more efficient. We will show in Sect. 5 that – given the experimentally observed fixation durations – the time available for lexical access increases considerably compared to the serial model. The assumption of parallel processes has the additional consequence that it predicts skipping of highly frequent words (Morrison 1984; Rayner and Pollatsek 1989).

3.1 A two-state model

Information processing is most efficient in the central 2° of the visual field, the *foveal* region. Acuity decreases in the *parafoveal* region, which extends out to 5°, and is even poorer in the peripheral region beyond the parafovea (Rayner 1998). Nevertheless, it will turn out to be a considerable advantage if parafoveal information is used during reading. For processing the word to the right of a currently fixated word, attention has to shift to the parafovea. In this case, the programming of saccade and lexical access are active simultaneously. Therefore, such models show parallel processing (Morrison 1984). Regardless of the number of internal states, the basic mechanisms that are performed by these models are *shifts of attention* and *eye movements*.

A minimal model for parallel processing of lexical access and saccade programming consists of two internal states (Fig. 2). In state 1, word$_n$ is fixated while lexical processing l_n is active. It is well known that lexical access time depends on word frequency,

$$l_n = l_b - l_m \log(F_n) \ , \tag{12}$$

where F_n is the frequency of word$_n$, and l_b and l_m are constant parameters. When lexical access is finished, the system switches to state 2. This transition implies a shift of attention to word$_{n+1}$.

The saccade program to the next word$_{n+1}$ starts in state 2. The saccade program is denoted by s_n; n is replaced by $n+1$ (update) when the transition is performed. Simultaneously, lexical access l_n starts, which is

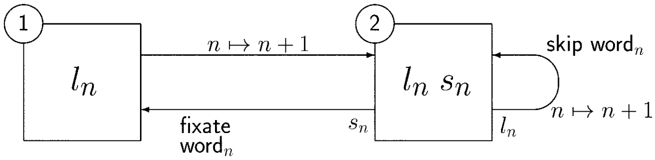

Fig. 2. The two-state model for word skipping. In state 1 (*left*) lexical access of word$_n$ is active while it is fixated. When lexical access is complete, the system steps to state 2 (*right*), i.e., a shift of attention occurs. The variable n is updated ($n \mapsto n+1$). In state 2, lexical access of word$_n$ starts by use of parafoveal information. Simultaneously, a saccade program to the same word is initiated. Depending on the frequency F_n of word$_n$, it can be skipped (l_n terminates first) or fixated (s_n terminates first)

possible due to the use of parafoveal information (by the shift of attention). Since the two sub-processes l_n and s_n of state 2 are in competition with each other, two transitions are possible. The saccade program is assumed to be independent of lexical properties of the text, with a mean value of the programming time of about 150 ms. If the saccade program is faster than lexical access, which is most likely for low-frequency words, the system switches to state 1. This transition signals the execution of a saccade (typical duration 15 to 40 ms). As a result, the fixation period of word$_n$ begins. In state 1, the lexical access of word$_n$ will be completed.

Alternatively, if lexical access is faster than the saccade program, a transition from state 2 to state 2 is performed. In this case, the saccade program to word$_n$ is canceled: there is no reason to fixate word$_n$, since it is already lexically processed. Instead a saccade program s_{n+1} and lexical access l_{n+1} of the next word are initiated (n is updated during the transition). As a consequence, word$_n$ is skipped. According to the model, this event will occur with higher probability for high-frequency words.

3.2 Simulations of the two-state model

For numerical simulations of the two-state model and its modifications, we use a corpus of sentences as previously discussed in Reichle et al. (1998). In an eye-tracking experiment participants read 48 sentences, each consisting of 8 to 14 words (Schilling et al. 1998). For model evaluation we compared numerical simulations of mean fixation duration and probability for word skipping with corresponding values observed in the experimental study.

Word frequency is the lexical parameter used for model simulations, as in (12). All words of the corpus were divided into five different frequency classes. Mean gaze duration and mean probability for skipping for these five classes are the experimental basis used here. Results from the model simulation are given in Fig. 3. Details of the simulation algorithm are discussed in Appendix A. For the execution of saccades we used gamma-distributed random numbers. The mean value was fixed at 25 ms and the standard deviation was fixed at one-third of the mean, as suggested by Reichle et al. (1998). The parameter estimation method is described in Appendix B, where mean fixation durations and skipping probabilities are used for the optimization procedure (Appendix B.1). Best-fit values for model parameters are $l_b = 271$, $l_m = 12.0$, $s_n = 170$, $\phi_l = 0.70$, and $\phi_s = 0.47$. The obtained minimal value for the deviation measure, (B3), is $\Delta_2 = 0.174$. Summarizing we can say that fixation durations as well as skipping probabilities are in good agreement with experimental data.

The most important restrictions for modeling different distributions of fixation durations arise from assumptions on the form of the transition probability rate (1). There are two important motivations for the linear τ-dependence chosen in our model. First, it is the next more complicated case compared to the Markov

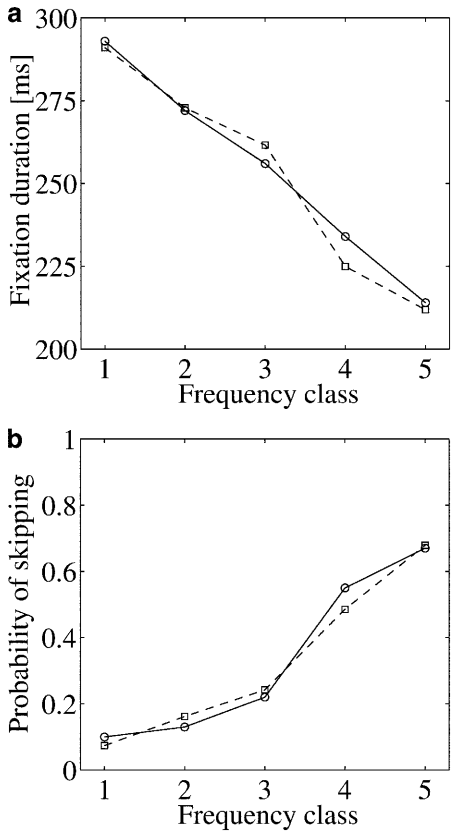

Fig. 3a,b. Fixation durations and probability of word skipping for model simulation (*dashed line*) and experimental data (*solid line*, from Reichle et al. 1998): **a** fixation durations show a linear dependence on the logarithm of word frequency – frequency classes represent effectively a logarithmic scale; **b** probability for word skipping increases with word frequency. In both panels, model simulations are in good agreement with experimental data

assumption ($W_m(\tau)$ = const.). Second, the resulting distribution (6) is qualitatively in agreement with typical experimentally observed distributions (Fig. 1); for (6), the probability decreases faster for increasing τ than in the case of gamma distributions.

While mean fixation durations and probabilites of word skipping can be explained with the two-state model, it fails to predict the correct distributions of fixation durations (see Fig. 6). In Sect. 4 we show how this shortcoming can be solved by increasing model complexity, i.e., by adding a third state to the model.

4 Modeling fixation durations

4.1 A three-state model

The next step in making the basic two-state model (Fig. 2) more flexible for reproducing distributions of fixation durations is to distinguish lexical access processes in states 1 and 2. Since acuity is decreased in the parafoveal region, we assume that lexical processing in state 2 is in a preliminary stage with different parameters: after the shift of attention, parafoveal information is required in state 2, while foveal information can be used in state 1. If lexical preprocessing in state 2, $l_n^\star = l_b^\star - l_m^\star \log(F_n)$, terminates, a transition to an additional state 3 occurs (Fig. 4). In this state, the saccade program to word$_n$ is canceled, i.e., word$_n$ is skipped, and a new saccade program to word$_{n+1}$ is initiated; at the same time, lexical access of word$_n$ is completed. We assume that lexical access time is independent of the state where it started. Therefore, lexical access time in states 1 and 3 is reduced by the amount of time lexical preprocessing (in state 2) is performed. As a consequence, it is most likely that lexical access l_n in state 3 is faster than s_{n+1}, since l_n is reduced by $l_n^\star$ (on the basis of mean values). Therefore, we do not add an additional transition for the case in which s_{n+1} terminates first.

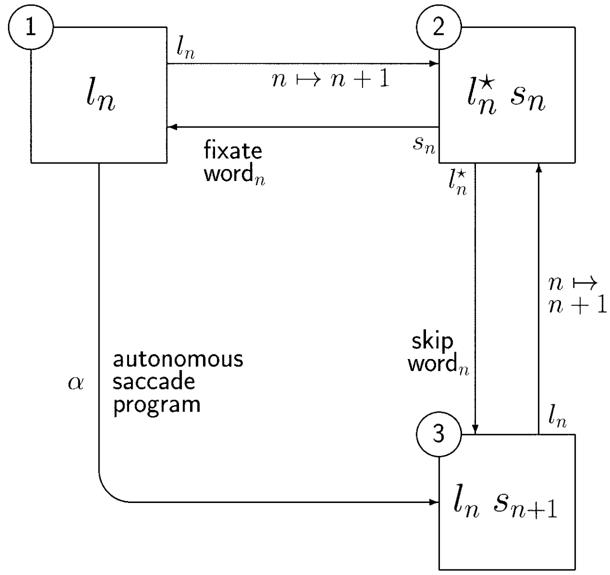

Fig. 4. The three-state model for the control of eye movements in reading. As a generalization of the two-state model (Fig. 2), lexical access parameters are different in states 1 and 2. To make the resulting model consistent with respect to lexical access parameters after a word has been skipped, an additional state 3 has to be assumed. The transition from state 1 to state 3 plays an important role for the coupling between lexical processes and saccade programs (see text)

4.2 Autonomous saccade programs

As an additional property of the three-state model, we consider a possible transition from state 1 to state 3: the "autonomous" initiation of a saccade program to word$_{n+1}$. The hypothesis behind the introduction of this transition is that the visual control system has some autonomy in programming a saccade. This assumption has important consequences for the coupling between shifts of attention and eye movement control. In extant theoretical models, eye movements are completely controlled by lexical processes, i.e., the initiation of a saccade program is determined by the familiarity check (Reichle et al. 1998). We now address the important question of whether we can relax the assumption of this strong form of coupling of the two sub-processes in order to explain the experimental data.

Table 1. Results of parameter estimation for two-state (*2*) and three-state (*3A*/*3B*) models. The star (*) indicates that model parameters were estimated from mean fixation durations and skipping probabilities only (Appendix B.1)

Model	Fitness	l_b	l_m	$l_b^\star$	$l_m^\star$	ϕ_l	s_n	ϕ_s	α	ϕ_a
2*	0.174	271	12.0	–	–	0.70	170	0.47	–	–
2	0.491	287	14.0	–	–	0.47	162	0.44	–	–
3A	0.348	255	9.5	188	9.9	0.47	113	0.69	–	–
3B	0.312	280	14.7	187	12.6	0.28	102	0.79	233	0.47

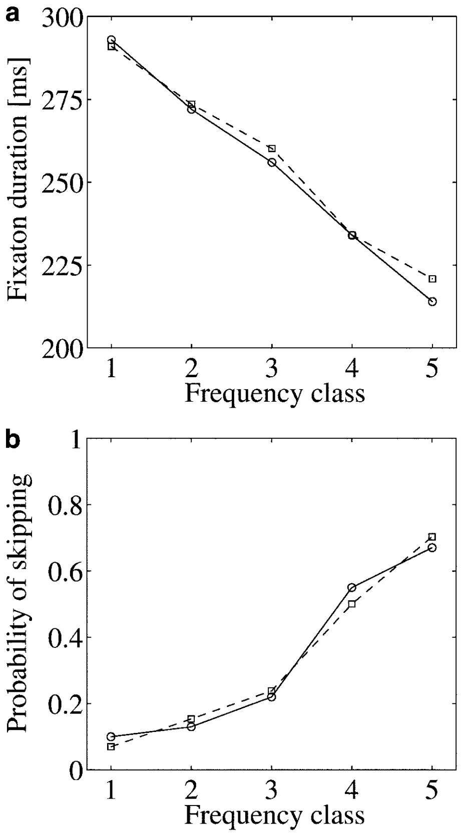

Fig. 5a,b. Fixation durations **a** and probability of word skipping **b** for numerical simulations (*dashed line*) of the three-state model (version B) and experimental data (*solid line*, from Reichle et al. 1998). Results are comparable to those of the two-state model (see Fig. 3)

This investigation is carried out by a comparison of the performance of different versions of the three-state model. In model 3A, we exclude transitions from state 1 to state 3, which is equivalent to $a \to \infty$. In model 3B, we assume that the mean residence time to start a saccade program s_{n+1} in state 1 is finite and constant, $a = \text{const.} > 0$. In this case, the transition from state 1 to state 3 is comparable to other transitions in the model. Like the mean time s_n to program a saccade to word$_n$, the parameter a is independent of word frequency F_n.

We now discuss the results obtained from model simulations. Using the same corpus of sentences as for the two-state model, the performance of the models is compared with respect to mean fixation durations and distributions as well as to the probability of skipping, for five different frequency classes (Eq. B5). The results are summarized in Table 1.

A first glance at Table 1 shows that the results for both versions (A and B) of the three-state model are comparable. This robustness of estimates for the parameters is a hint for the structural stability of the models: small modification do not lead to qualitative changes. This property emphasizes the psychological plausibility of the models. For a more detailed comparison between the two- and three-state models, the results for model 3B are presented in Fig. 5 (mean fixation duration and probability for word skipping) and 6 (distribution of fixation durations for 5 different classes of word frequency). Generally, model simulations are in good agreement with experimental data.

All models discussed here represent the same class of models, but with differing model complexity, e.g., number of internal states and parameters. Generally, model performance, i.e., fitness, increases with model complexity. This a non-trivial result, because the coupling between lexical access and saccade programs decreases from model 3A to model 3B. While in model 3A only lexical access can trigger the initiation of a saccade program, in model 3B the saccade programs can start spontaneously. Additionally, we investigated a saccadic sub-system with a periodic forcing, i.e., a periodic variation of the probability for the initiation of saccade programs, which turned out not to destabilize model performance. This numerical control study provides further evidence on the robustness of our results.

A further remark concerns the two-state model: while the two-state model can explain the pattern of mean fixation duration and probability for word skipping, it fails to predict the correct distribution of fixation durations. The corresponding best fitness value obtained in our simulations was $\Delta_3 = 0.491$ (see Eq. B5), which is a clear indication that the two-state model is too limited to explain the distribution of fixation durations.

The most important consequence of the introduction of an autonomous saccade program is related to predictions about preview benefit. This is a qualitatively new property of model 3B that cannot be achieved by model 3A. We discuss these results in Sect. 4.3.

4.3 Analysis of preview benefit

It is a well-known experimental observation that lexical processing time of a word is shorter when there was a preview of the word in the parafovea (Rayner 1998). Following Reichle et al. (1998), a key problem in minimal models of eye movement control in reading is

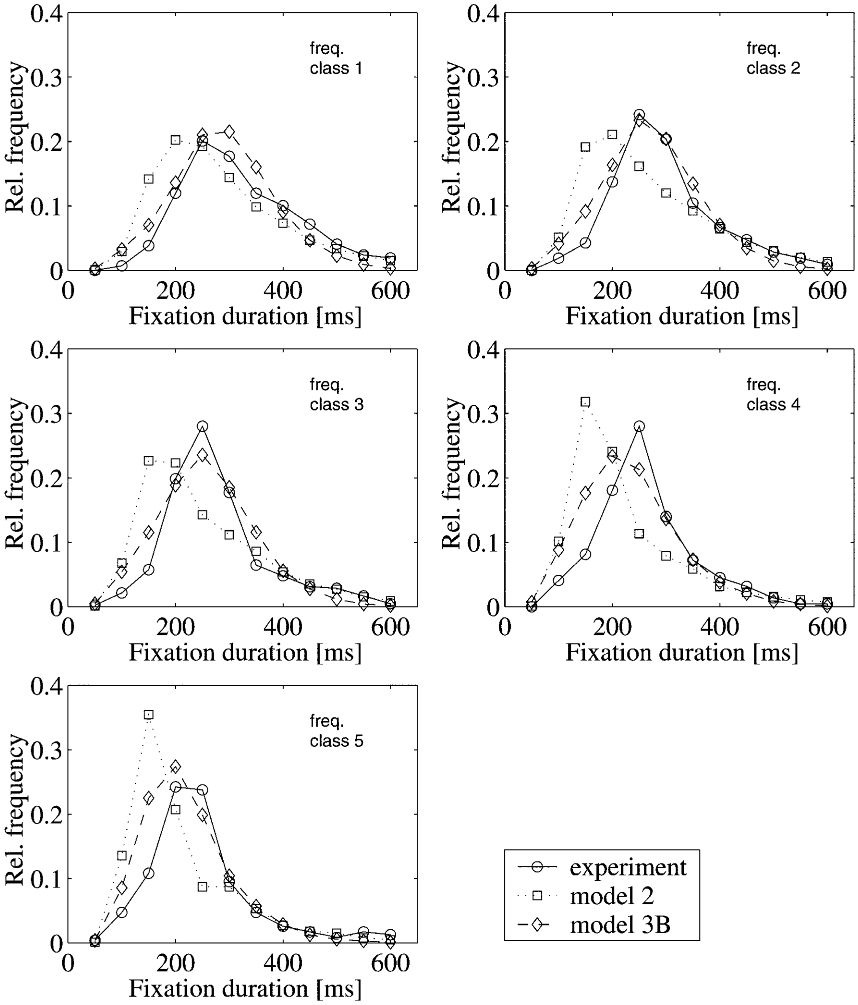

Fig. 6. Distribution of fixation durations for five different word frequency classes (experiment, *solid lines*; model 2, *dotted*; model 3B, *dashed*). Despite its performance with respect to mean fixation durations (Fig. 3a), the two-state model fails to reproduce the distributions of fixation durations. Simulations of the three-state model are in good agreement with experimental data (after Reichle et al. 1998)

to explain that preview benefit is modulated by foveal difficulty. Therefore, an important success of the E-Z Reader models is to provide a mechanism for this effect by introducing two stages of lexical processing: a familiarity check and lexical completion.

In contrast to Reichle et al. (1998), our three-state model 3B shows that preview benefit depends on its inherent dynamic behavior with a single lexical processing stage (Fig. 4). If lexical access of word$_n$ starts in state 1, then there are two possible transitions (to state 2 and state 3, respectively), whose stepping probabilities are modulated by word frequency. In the case of a foveal high-frequency word, a transition to state 2 will most likely occur. Preview benefit is, therefore – to a first approximation – given by the time required for programming the saccade to word$_{n+1}$. Alternatively, a low-frequency word$_n$ in the fovea induces a high probability for a transition to state 3 via an autonomously triggered saccade. If this happens, a saccade program to word$_{n+1}$ is initiated before lexical access of word$_n$ is completed. The consequence for the amount of preview benefit is that when the system finally steps to state 2 (there is no alternative), preview benefit is reduced by the amount of time that the system has spent in state 3. Thus, our model provides an explanation of how preview benefit can be modulated by foveal processing difficulty, in terms of its dynamics.

On the assumption that preview benefit is minimal if the previous word is skipped, we confirmed these considerations about the underlying qualitative dynamics with numerical simulations. Thus, the difference between fixation durations on words following a skipped word and those following a fixated word should indicate a preview benefit increasing with word frequency. The difference between the solid lines in Fig. 7 reveals that this was indeed the case for model 3B. The amount of preview benefit obtained is in the same order of magnitude as in the study by Reichle et al. (1998). The upper limit for preview benefit in our model is the mean time required to program a saccade, s_n, since preview benefit is given by the residence time in state 2 (Fig. 2). The dependence of preview benefit on word frequency was not observed for model 3A (i.e., the difference between the dashed lines was roughly constant). The primary source of the increase of preview benefit with word frequency in model 3B is due to the introduction of autonomous saccades in this model leading to an increase of the fixation duration for low-frequency words if the previous word was fixated (two bottom lines in Fig. 7). Since this effect is caused by transitions from state 1 to

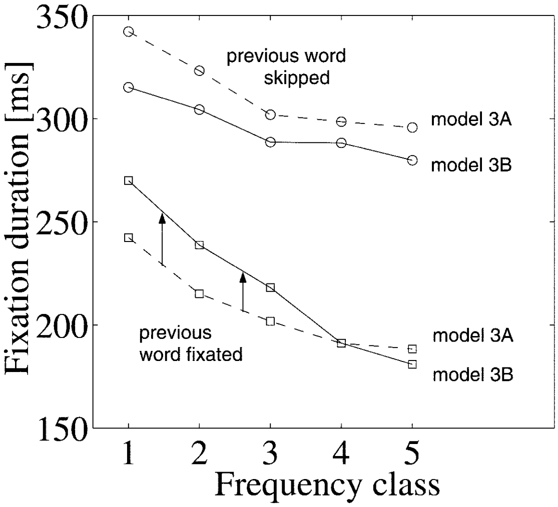

Fig. 7. Preview benefit in models 3A (*dashed lines*) and 3B (*solid lines*), averaged over 500 realizations. The *bottom two lines* represent fixation durations as a function of word frequency class for all cases, in which the previous word was fixated. Corresponding mean fixation durations are shorter than in those cases where the previous word was skipped (*top two lines*), i.e., there was minimal preview. It is an important consequence of the introduction of autonomous saccades that preview benefit is modulated by word frequency (see text)

state 3 (Fig. 4), the mean residence time in state 2 is reduced. As a consequence, this reduction of the preview benefit is highest for foveal low-frequency words. As expected, there is no or very little frequency-dependent modulation of fixation durations due to autonomous saccades if the previous word is skipped (top two lines in Fig. 7).

Finally we note that there are phenomena not covered by the E-Z Reader model as well as our present model, such as the influence of the difficulty of the last word fixated on fixation duration on the word currently fixated (Kennedy 2000), or extraction of information to the left of the fixated word (Binder et al. 1999; see also Kennedy et al. 2000 for a recent review). A discussion of these issues is beyond the scope of the current study; our aim here is to demonstrate that complex dynamic behavior emerging from simple mathematical models can explain experimental findings in a psychologically plausible way.

5 Discussion

Theoretical models of eye movement control successfully account for several statistical aspect of eye movement pattern during reading (Reichle et al. 1998). A key model assumption relates to the coupling of eye-movement preparation and shift of visual attention. In the E-Z Reader framework (Reichle et al. 1998), detailed mechanisms of lexical processing are proposed (Liversedge and Findlay 2000). This class of models is based on the assumption that saccade programming is strictly governed by lexical processes. Following Deubel et al. (2000), however, "the alternative notion that low-level oculomotor processes might be playing the dominant role in eye movement control during reading remains a serious possibility". Based on the analysis of initial landing positions, Reilly and O'Regan (1998) compared different word-targeting strategies as an example of this alternative approach to eye movement control.

In this study we restricted our analysis to minimal models investigating a new form of coupling between eye movements and visual attention, where saccade programs can be initiated both by lexical processes and autonomously. Using a three-state model, we showed that relaxing the assumption that lexical processing completely governs the initiation of saccade programs leads to an increase in model performance. Therefore, it seems promising to study further variants or extensions of our three-state model with the aim of analyzing the relation between shifts of attention and eye movements. Our preliminary results, however, suggest that models without a strong coupling between lexical processes and saccade programs may still be a viable alternative to the successful E-Z Reader models (Reichle et al. 1998).

As a consequence of the introduction of an autonomous saccade program, even the three-state model shows complex dynamic behavior, which leads to psychologically plausible results on preview benefit. An important property of this model is that it provides an explanation of how preview benefit is modulated by foveal processing difficulty in terms of its dynamics.

An interesting parameter in models of eye movement control in reading is lexical processing time. Fixation duration as a function of word frequency follows approximately the relation $t_n = 303 - 8.26\log(F_n)$ (Fig. 3a, solid line). In a strictly serial model, explained at the beginning of Sect. 3, lexical and saccade processing time simply sum to t_n. Assuming a mean saccade program duration of $s = 100.0$ ms yields a lexical processing time $l_n^s = t_n - s = 203 - 8.26\log(F_n)$. Based on the same experimental data, all models discussed here provide a significant increase in (available) processing time. This increase is due to parallel processing of saccades and lexical access (Morrison 1984). Furthermore, our model suggests that gaze duration is a key measure of lexical processing time, since the slope parameters (−9.5 to −14.7) estimated in the model are in good agreement with the empirical slope for the regression of gaze duration on word frequency (−8.26).

Using the theoretical framework of semi-Markov processes, we investigated the role of stochastic transition rules for models of eye movement control. In the E-Z Reader framework (Reichle et al. 1998), pausing-time distributions are assumed to be gamma-distributed with a fixed relation between standard deviation and mean value. As an alternative, we analyzed residence-time-dependent transition probability rates. Besides the fact that transition probability rate is a more fundamental concept for stochastic processes (Gillespie 1978), an additional advantage of this approach is that it can be implemented by an exact algorithm for numerical simulations.

The implementation of stochastic models may result in considerable deviations from the exact results (Feistel 1977). To avoid these problems for the numerical simulations of our mathematical models, we used an exact algorithm (Gillespie 1978) which is a generalization of the minimal process method (Gillespie 1976). This method was proposed originally for numerical simulations of chemical reactions, but has been applied to a broad class of systems, e.g., from molecular biology (Elowitz and Leibler 2000), physiology (Fricke and Schnakenberg 1991), or population dynamics (Engbert and Drepper 1994). Therefore, we believe that our approach may be applicable to a variety of problems in the field of eye movement control.

As a further remark, the framework for stochastic simulation introduced here can also be used for numerical simulations of a broader class of models (Reichle et al. 1998). We restricted our analysis to minimal models in order to investigate the coupling of eye movements and shifts of visual attention in detail. For this reason, the three-state model is on the same level of abstraction as E-Z Reader 1. Possible extensions of this three-state model to account for refixations could be derived in close analogy to the development of E-Z Reader 3 to 5 by Reichle et al. (1998). In particular, a saccadic refixation program could be introduced as an additional subprocess in state 1 (Fig. 4). We note, however, that a more psychologically plausible explanation for the occurrence of refixations should provide a common mechanism for saccade, refixations, and regressions.

As a final remark, our results suggest that assumptions about stochastic properties of mathematical models may strongly influence the performance of the models as well as their complexity (e.g., the number of internal states necessary). A comparison between E-Z Reader 1 (eight internal states) with our three-state model suggests that – if these assumptions are too restrictive – model complexity (e.g., number of internal states) could be overestimated.

Acknowledgement. We thank André Longtin, University of Ottawa, for valuable discussions and Eric D. Reichle, Carnegie Mellon University, for providing us with the corpus of 48 sentences used in our simulations. R.K. thanks Jan Droesler, University of Regensburg, for the hospitality during a stay at his unit. This work was supported by Deutsche Forschungsgemeinschaft (DFG grants KL 955/3-1 and KL 955/3-2).

Appendix A: Numerical simulation – Gillespie algorithm

Efficient numerical simulation techniques for Markov processes (Gillespie 1976; Feistel 1977) have been extended to stochastic simulation of processes with residence-time-dependent transition probability. Details of the derivation of the algorithm can be found in Gillespie (1978).

The pausing-time distribution can be calculated from the transition probability rate (4). For numerical simulations, we have to create pseudo-random numbers according to this distribution. The general rule (Gillespie 1978) for the transformation of computer-generated random numbers r to τ is given by

$$\int_0^{\tau} W_m(\tau')\mathrm{d}\tau' = \log\left(\frac{1}{r}\right) \ , \qquad \text{(A1)}$$

where r is a random number from the unit-interval uniform distribution (Press et al. 1988). In our case (5), this relation is invertible, so that $\tau(r)$ can be calculated analytically. Realizations of the pausing time τ are obtained from

$$\tau(r) = \tau_0 + \sqrt{\frac{2}{w_m}\log\left(\frac{1}{r}\right)} \ , \qquad \text{(A2)}$$

where τ_0 is the refractory time. Details of the simulation algorithm are discussed in Gillespie (1978).

As a typical case for the models studied in this paper, there are transitions between states with two competing processes. As an example, consider state 2 in the two-state model (Fig. 2): lexical access l_n of word$_n$ and the saccade program s_n to the same word are active at the same time. If the saccade program terminates before lexical access is complete, then the transition from state 2 to state 1 is performed and fixation of word$_n$ starts. When the system arrives in state 1, lexical access of word$_n$ has already been active for a time $\tilde{\tau}$, which is, in this case, the pausing time in the previous state 2. Therefore, the transition probability in state 1 is $W(\tau + \tilde{\tau})$, where $\tilde{\tau}$ is the time period for which lexical access has already been active. We now extend (A2) to the case of two competing sub-processes with one and two adjoining states.

A.1. Transitions with one adjoining state

Let us define the difference $\tau_1 = \tau_0 - \tilde{\tau}$ of refractory time τ_0 (5) and $\tilde{\tau}$ of the process. The pausing time at state S_1 can be transformed from unit-interval uniform random numbers as follows,

$$\tau_1 \geq 0: \quad \tau = \tau_1 + \sqrt{\frac{2}{w_1}\log\left(\frac{1}{r}\right)} \ , \qquad \text{(A3)}$$

$$\tau_1 < 0: \quad \tau = \tau_1 + \sqrt{\tau_1^2 + \frac{2}{w_1}\log\left(\frac{1}{r}\right)} \ . \qquad \text{(A4)}$$

A.2. Transitions with two adjoining states

In the case of two possible transitions, like state 2 of the two-state model (Fig. 2), the stochastic simulation of the pausing time is more complicated. We consider two competing sub-processes with $\tau_k = \tau_0^{(k)} - \tilde{\tau}^{(k)} (k = 1, 2)$, where $\tau_0^{(k)}$ are refractory times and $\tilde{\tau}^{(k)}$ are the time periods for which the processes have already been active. For simplicity, let us assume $\tau_1 \leq \tau_2$. The pausing time of such states can be computed as follows:

Case 1. $\tau_1 \le 0,\ \tau_2 \le 0$

$$\tau = \frac{w_1\tau_1 + w_2\tau_2}{w_1 + w_2} + \sqrt{\left(\frac{w_1\tau_1 + w_2\tau_2}{w_1 + w_2}\right)^2 + \frac{2}{w_1}\log\left(\frac{1}{r}\right)} \tag{A5}$$

Case 2. $\tau_1 \le 0, \quad \tau_2 > 0$

(a) $\tau \le \tau_2 \quad \tau = \tau_1 + \sqrt{\tau_1^2 + \frac{2}{w_1}\log\left(\frac{1}{r}\right)}$

(b) $\tau > \tau_2$

$$\tau = \frac{w_1\tau_1 + w_2\tau_2}{w_1 + w_2} + \sqrt{\left(\frac{w_1\tau_1 + w_2\tau_2}{w_1 + w_2}\right)^2 - \frac{w_2\tau_2^2}{w_1 + w_2} + \frac{2}{w_1}\log\left(\frac{1}{r}\right)} \tag{A6}$$

Case 3. $\tau_1, \tau_2 > 0,\ \tau_1 < \tau_2$

(a) $\tau \le \tau_2 \quad \tau = \tau_1 + \sqrt{\frac{2}{w_1}\log\left(\frac{1}{r}\right)}$

(b) $\tau > \tau_2$

$$\tau = \tau_1 + \frac{w_2\delta\tau}{w_1 + w_2} + \sqrt{\left(\frac{w_2\delta\tau}{w_1 + w_2}\right)^2 - \frac{w_2(\delta\tau)^2}{w_1 + w_2} + \frac{2}{w_1 + w_2}\log\left(\frac{1}{r}\right)} \ , \tag{A7}$$

where $\delta\tau = \tau_2 - \tau_1$.

Appendix B: Parameter estimation – genetic algorithm approach

For the estimation of model parameters we use a genetic algorithm (GA) approach (Holland 1992; Mitchell 1996). For each sentence, 500 stochastic realizations of the model were run with a new set of pseudo-random numbers. For the genetic algorithm we used a population of 100 combinations of parameter values which were iterated over 1000 generations. Several runs of the GA were used to test the reliability of the estimates for the model parameters. The parameter values given in Table 1 represent mean values over nine runs of the GA. A separate simulation was performed to produce the data shown in Fig. 3, 5, and 6. A single run for the parameter estimation took approximately 48 hours CPU time on a SUN Ultra 10 computer. The fitness function used for the GA optimization method is discussed below.

B.1 Mean fixation duration and skipping probability

The 536 words of the corpus (48 sentences) are divided into five different frequency classes, as suggested in Reichle et al. (1998). The frequencies in the different classes are 0–10 (class 1), 11–100 (class 2), 101–1000 (class 3), 1001–10000 (class 4), and 10001–∞ (class 5).

With the two-state model we aim to explain the experimental results of mean fixation duration and probability of word skipping. The mean fixation duration for words of class k obtained from model simulations is denoted by $T(k)$; $\sigma_T(k)$ is the corresponding standard deviation. This is compared with the experimentally observed value $T^\circ(k)$. The deviation of simulated mean fixation durations from observed mean fixation durations is defined as

$$\Delta_T = \sum_{k=1}^{5}\left(\frac{T^\circ(k) - T(k)}{\sigma_T(k)}\right)^2 \ , \tag{B1}$$

i.e., the sum of squared differences over five different frequency classes.

As the second measure of model performance we use the probability of word skipping. $p(k)$ is the probability of word skipping in class k; the experimental value is denoted by $p^\circ(k)$. Analogously to (B1), the deviation between simulated and observed probabilities of word skipping can be defined as

$$\Delta_p = \sum_{k=1}^{5}\left(\frac{p^\circ(k) - p(k)}{\sigma_p(k)}\right)^2 \ , \tag{B2}$$

where $\sigma_p(k) \equiv \sqrt{p(k)(1 - p(k))}$.

Combining these two terms gives a possible fitness function that is used for the genetic algorithm parameter estimation method,

$$\Delta_2 = (\Delta_T + \Delta_p)^{1/2} \ . \tag{B3}$$

B.2 Distribution of fixation durations

For the three-state model, we include a measure for the deviation of simulated data from the observed distribution of fixation durations,

$$\Delta_D = \sum_{k=1}^{12}(h^\circ(k) - h(k))^2 \ , \tag{B4}$$

where $h(k)$, $(h^o(k))$ are the relative frequencies of simulated (observed) fixation durations in a distribution over 12 bins (from Reichle et al. 1998). Therefore, the fitness function is modified to

$$\Delta_3 = (\Delta_T + \Delta_p + \Delta_D)^{1/2} \ . \tag{B5}$$

References

Binder KS, Pollatsek A, Rayner K (1999) Extraction of information to the left of the fixated word in reading. J Exp Psychol Hum Percept Perform 25: 1162–1172

Deubel H, O'Regan JK, Radach R (2000) Attention, information processing, and eye movement control. In: Kennedy A, Radach R, Heller D, Pynte J (eds) Reading as a perceptual process. Elsevier, Oxford

Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335–338

Engbert R, Drepper FR (1994) Chance and chaos in population biology – models of recurrent epidemics and food chain dynamics. Chaos Solitons Fractals 4: 1147–1169

Feistel R (1977) Betrachtung der Realisierung stochastischer Prozesse aus automatentheoretischer Sicht. Wiss. Z. WPU Rostock 26: 663–670

Fricke T, Schnakenberg J (1991) Monte-Carlo simulation of an inhomogeneous reaction-diffusion system in the biophysics of receptor cells. Z Phys B 83: 277–297

Gardiner CW (1990) Handbook of stochastic methods. Springer, Berlin, Heidelberg, New York

Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22: 403–434

Gillespie DT (1977) Master equations for random walks with arbitrary pausing time distributions. Phys Lett A 64: 22–24

Gillespie DT (1978) Monte Carlo simulation of random walks with residence time dependent transition probability rates. J Comput Phys 28: 395–407

Gillespie DT (1992) Markov processes. Academic, San Diego

Holland JH (1992) Adaptation in natural and artificial systems. MIT Press, Cambridge, Mass

van Kampen NG (1981) Stochastic processes in physics and chemistry. North-Holland, Amsterdam

Kennedy A (2000) Parafoveal processing in word recognition. Q J Exp Psychol A 53: 429–455

Kennedy A, Radach R, Heller D, Pynte J (eds) (2000) Reading as a perceptual process. Elsevier, Oxford

Liversedge SP, Findlay JM (2000) Saccadic eye movements and cognition. Trends Cogn Sci 4: 6–14

Mitchell M (1996) An introduction to genetic algorithms. MIT Press, Cambridge, Mass

Morrison RE (1984) Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades. J Exp Psychol Hum Percept Perform 10: 667–682

Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1998) Numerical Recipes in C. Cambridge University Press, Cambridge

Rayner K (1998) Eye movements in reading and information processing: 20 years of research. Psychol Bull 124: 372–422

Rayner K, Pollatsek A (1989) The psychology of reading. Prentice-Hall, Englewood-Cliffs, N.J

Reichle ED, Pollatsek A, Fisher DL, Rayner K (1998) Toward a model of eye movement control in reading. Psychol Rev 105: 125–157

Reichle ED, Rayner K, Pollatsek A (1999) Eye movement control in reading: accounting for initial fixation locations and refixations within the E-Z Reader model. Vision Res 39: 4403–4411

Reilly RG, O'Regan JK (1998) Eye movement control during reading: a simulation of some word-targeting strategies. Vision Res 38: 303–317

Schilling HEH, Rayner K, Chumbley JI (1998) Comparing naming, lexical decision, and eye fixation times: word frequency effects and individual differences. Mem Cogn 26: 1270–1281