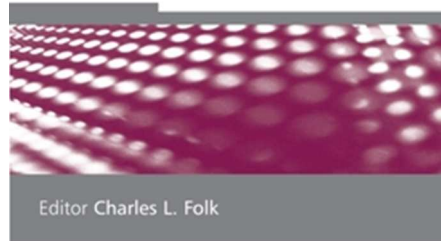


Visual Cognition



Psychology Press
Taylor & Francis Group

A theoretical analysis of the perceptual span based on SWIFT simulations of the n+2 boundary paradigm

Journal:	<i>Visual Cognition</i>
Manuscript ID:	VC-CONK 09.13.R2
Manuscript Type:	Full Article
Date Submitted by the Author:	30-Dec-2013
Complete List of Authors:	Risse, Sarah; University of Potsdam, Hohenstein, Sven; University of Potsdam, Kliegl, Reinhold; University of Potsdam, Engbert, Ralf; University of Potsdam,
Keywords:	eye movements, reading, computational modeling, perceptual span, preview

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12 **A theoretical analysis of the perceptual span based on SWIFT simulations of the n+2**
13 **boundary paradigm**
14

15
16 Sarah Risse, Sven Hohenstein, Reinhold Kliegl, & Ralf Engbert

17
18 *University of Potsdam, Germany*
19

20
21
22
23
24
25
26 Running head: SWIFT simulations on the perceptual span

27
28 Keywords: eye movements, reading, computational modeling, perceptual span, preview
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 Correspondence to:

45 Sarah Risse

46 Department of Psychology

47 Cognitive Psychology

48 University of Potsdam

49 Karl-Liebnecht-Strasse 24-25

50 D-14476 Potsdam, Germany

51 email: sarah.risse@uni-potsdam.de
52
53
54
55
56
57
58
59
60

Abstract

Eye-movement experiments suggest that the perceptual span during reading is larger than the fixated word, asymmetric around the fixation position, and shrinks in size contingent on the foveal processing load. We used the SWIFT model of eye-movement control during reading to test these hypotheses and their implications under the assumption of graded parallel processing of all words inside the perceptual span. Specifically, we simulated reading in the boundary paradigm and analyzed the effects of denying the model to have valid preview of a parafoveal word $n+2$ two words to the right of fixation. Optimizing the model parameters for the valid preview condition only, we obtained span parameters with remarkably realistic estimates conforming to the empirical findings on the size of the perceptual span. More importantly, the SWIFT model generated parafoveal processing up to word $n+2$ without fitting the model to such preview effects. Our results suggest that asymmetry and dynamic modulation are plausible properties of the perceptual span in a parallel word-processing model such as SWIFT. Moreover, they seem to guide the flexible distribution of processing resources during reading between foveal and parafoveal words.

Introduction

The perceptual span during reading is substantially larger than the word that is currently fixated. English readers, for example, obtain information from about 3-4 letters to the left of fixation and up to 14-15 letters to the right (McConkie & Rayner, 1975, 1976; Rayner, Well, & Pollatsek, 1980). The larger extent into the direction of reading suggests that readers rely substantially on information ahead of the eyes in order to efficiently coordinate their eye movements with the ongoing word-recognition processes. Moreover, the size of the perceptual span seems to adjust to the processing difficulty of the fixated word. With a difficult word in foveal vision the perceptual span becomes smaller allowing less preprocessing of the upcoming word in parafoveal vision (Henderson & Ferreira, 1990; Kennison & Clifton, 2005)—an effect recently implemented in a computational model (Schad & Engbert, 2012).

The boundary paradigm (Rayner, 1975; Rayner, McConkie, & Zola, 1980) is a powerful method to investigate the effect of parafoveal information on reading. While the readers' gaze is to the left of an invisible boundary located at the end of a given word n (word n is defined as the currently fixated word), the preview of the next word $n+1$ to the right of the boundary is masked (e.g., with random letters forming a nonword). Immediately after the eyes have crossed the boundary the preview is replaced by the target word. Reading times of the target word are reliably longer for these invalid preview cases than when the reader previewed the identical target word on fixations also before the boundary. Consequently, this difference has been termed *parafoveal preview benefit* (see Hyönä, 2011; Schotter, Angele, & Rayner, 2012, for reviews). This definition is consistent with the notion that the perceptual span also contains parafoveal words that are preprocessed before they are fixated.

However, many questions regarding the perceptual span remain unresolved. One such issue relates to how the words falling into the perceptual span are processed during each fixation, and reading models differ fundamentally in their assumptions on this matter. SWIFT as a fully implemented computational model of saccade generation during reading

1
2
3
4
5 postulates that attention is distributed across the full perceptual span, and all words inside
6 the span are processed in parallel (Engbert, Longtin, & Kliegl, 2002; Richter, Engbert, &
7 Kliegl, 2006; Schad & Engbert, 2012). While it is correct to say that SWIFT is a model of
8 parallel processing, it is important to note that processing is not equally distributed over a
9 large number of words. A more precise description of the SWIFT word processing
10 mechanism is that word processing rates are dependent on word position within the
11 perceptual span and on word length (Engbert, Nuthmann, Richter, & Kliegl, 2005), so that
12 word processing is often limited to one or two words at a time (see Appendix C; Engbert et
13 al., 2005). The assumption of graded parallel processing, however, is not undisputed as
14 there are other models such as E-Z Reader implementing attention allocation in the form of a
15 one-word spotlight that is sequentially shifted from word to word (Reichle, 2011; Reichle,
16 Pollatsek, Fisher, & Rayner, 1998; see also Schotter, Reichle, & Rayner, this volume).
17
18
19
20
21
22
23
24
25
26
27
28

29
30 Research on eye-movement control during reading demonstrated that computational
31 modeling is a useful scientific approach to directly test the plausibility of cognitive theories on
32 the perceptual span during reading by means of systematic comparisons between
33 experimental data and simulation studies. The full potential of theoretical models can only be
34 realized, however, if they are used to make predictions on data that were not considered for
35 model development and optimization or for parameter estimation. Following this approach,
36 we used the SWIFT model to simulate reading in the boundary paradigm for preview of word
37 $n+2$, that is the word beyond the word to the right of the invisible boundary (Kliegl, Risse, &
38 Laubrock, 2007; Risse & Kliegl, 2011), without fitting the model to the preview effects. We
39 will explain the modeling approach in more detail below. In contrast to word $n+1$, the next
40 word $n+2$ lies at the spatial limit of the perceptual span and parafoveal preprocessing should
41 be attenuated due to decreasing visual acuity at such parafoveal distances. It is important to
42 note, however, that models of saccade planning assume a perceptual component that
43 typically extends to words $n+1$ and $n+2$ (Engbert & Krügel, 2010; Krügel & Engbert, 2010;
44 Krügel, Vitu, & Engbert, 2012). The findings of word $n+2$ preview in the boundary paradigm
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 vary across experiments. While some studies did not find significant preview effects of word
5 $n+2$ (Rayner, Juhasz, & Brown, 2007; see also Angele, Slattery, Yang, Kliegl, & Rayner,
6 2008), other studies showed that, given a short two- or three-letter word $n+1$, some
7 information about the next word $n+2$ can be preprocessed already during fixations on the
8 preboundary word n if word $n+1$ is subsequently skipped (Angele & Rayner, 2011; Kliegl et
9 al., 2007; Risse & Kliegl, 2011, 2012; but see Radach, Inhoff, Glover, & Vorstius, 2013, for
10 effects on word $n+2$ even when word $n+1$ was fixated). In fact, $n+2$ preview effects occurred
11 mainly on the word that was fixated first after the boundary (Kliegl et al., 2007; Risse & Kliegl,
12 2011, 2012). Thus, a computational model should account for nonlocal patterns of preview
13 effects in the target region as they seem to reveal important spatiotemporal characteristics
14 about the foveal and parafoveal integration-processes across the perceptual span during
15 reading.
16
17
18
19
20
21
22
23
24
25
26
27

28 29 **The modeling approach**

30
31 Using a recently published version of the SWIFT model (Schad & Engbert, 2012; see also
32 Engbert et al., 2005) we tested whether a model based on distributed attention across the
33 perceptual span accounts for the specific pattern of $n+2$ preview effects observed in the
34 boundary paradigm. Following from the results reported above our main questions
35 concerned (1) whether the model generates an $n+2$ preview benefit on word $n+2$ after
36 skipping word $n+1$, (2) whether a preview benefit on word $n+2$ is absent when word $n+1$ was
37 previously fixated, and (3) whether the model shows an effect of $n+2$ preview on word $n+1$ if
38 it was fixated first. Moreover, it is often argued that parallel word processing naturally implies
39 crosstalk between processing foveal and parafoveal words simultaneously, resulting in
40 parafoveal-on-foveal effects of the not-yet-fixated word in fixation durations on the fixated
41 word (Drieghe, Rayner, & Pollatsek, 2008; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999;
42 see also Trukenbrod & Engbert, 2012, for a scanning task). However, whether the nonlinear
43 dynamics of the SWIFT model indeed generate such effects requires simulations if one
44 wants to go beyond speculation. Here, we checked whether simulations with SWIFT also
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 yield a parafoveal-on-foveal effect of previewing word $n+2$ on word n . Experimental evidence
5
6 for such a result is, however, mixed (i.e., parafoveal-on-foveal effects of word $n+2$ were
7
8 found only in Kliegl et al., 2007, and Risse & Kliegl, 2012, Experiment 1, but not Experiment
9
10 2).

11
12
13 A further goal was to use the SWIFT model and to test specific assumptions about
14
15 the perceptual span during reading. To this end, we first fit the model parameters to the
16
17 normal reading situation. In the boundary paradigm, normal reading occurs in the identical
18
19 preview condition because in this condition preview is provided during all fixations. The best-
20
21 fitting model parameters were then used to simulate reading in the invalid preview condition
22
23 in which the target word $n+2$ was not revealed until the eyes moved away from word n .
24
25 Critically we did not optimize original or even additional model parameters to capture the
26
27 preview effects in the boundary paradigm. Rather we only assumed that the model deals with
28
29 invalid cases, specifically the fact that there is a display change, in a principled way
30
31 (described below). Using the parameter values fitted to reading with valid preview and the
32
33 model's principled response to display changes, we explored how much of the preview
34
35 effects are already inherent in the present version of the SWIFT model. These simulations
36
37 were then compared with the experimental results, using a likelihood computation of the
38
39 experimental data given the SWIFT model. In addition, we checked whether the model
40
41 parameters stayed within a reasonable range given what we know about the perceptual span
42
43 from experimental research. We also used the estimates of the model parameters as a
44
45 source of information about how the model accounts for processing of parafoveal words. For
46
47 this purpose, we investigated the range of certain span parameters and assessed their
48
49 contribution to the processing of word $n+2$ within the model.
50
51

52 **The zoom lens model: SWIFT 3**

53
54
55 To account for the readers' when- and where-decisions during reading, computational
56
57 models of eye-movement control must provide an interface between cognitive and
58
59 oculomotor processes. With respect to cognition, the SWIFT model represents sentence
60

1
2
3
4
5 processing as a set of word-based lexical activations. This field of lexical activations (e.g.,
6 Erlhagen & Schöner, 2002) changes over time as a function of word processing. Word
7 recognition is implemented as a two-stage process: A word is processed when its lexical
8 activation increases from zero (i.e., no information about the word) to a maximum defining
9 the word's difficulty and decreases from this maximum back to zero (i.e., full identification of
10 the word). With respect to the oculomotor aspects of the model, SWIFT generates saccade
11 programs autonomously based on a random timer¹ (similar to Findlay & Walker, 1999; see
12 also Trukenbrod & Engbert, 2013). As illustrated in Figure 1, word processing and
13 oculomotor control are linked to each other via two routes: (1) The lexical activation of the
14 fixated word in foveal vision can inhibit the initiation of the next saccade program by inhibiting
15 the random timer. Thus, the activation of the fixated word influences the decision of when to
16 move the eyes. (2) The lexical activations of all words in the sentence determine the
17 probability with which each word is selected as the next saccade target. Thus, the lexical
18 activation of all words influences the decision of where to move the eyes. Highly activated
19 words are more likely to be fixated next whereas words with very low activation are likely to
20 be skipped and not fixated at all. Accordingly, the lexical-activation field in SWIFT enables
21 the full range of behaviors in the model. It allows forward saccades to the next word in
22 sequence and word skipping as well as refixations and regressions back to earlier words.
23 Simultaneously, it controls the processing-dependent prolongation of fixation durations
24 through inhibition of the random timer according to the foveal activation state.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

45
46 (Figure 1 about here)
47

48 How much information is processed during each fixation and how fast it is processed
49 is determined by the model's processing span. This can be regarded as the perceptual span
50 of the model and defines the region of letters that contribute to the increase or decrease of
51 lexical activation of the corresponding words. The processing span in the latest version of the
52 SWIFT model (SWIFT 3; Schad & Engbert, 2012) is implemented as an inverse parabolic
53 function that assigns a processing rate above zero to all letters falling below the curve. In
54
55
56
57
58
59
60

agreement with a gradual reduction of visual acuity beyond the fovea, the processing rate decreases with increasing eccentricity from the fixation position. Moreover, this model also implemented the idea of an attentional zoom lens (e.g., Eriksen & St. James, 1986; La Berge & Brown, 1989) such that the size of the processing span is modulated conditional on the foveal processing demand. As the lexical activation of the fixated word decreases, the right part of the span increases and more and more letters to the right of fixation fall inside the processing span. Such a dynamic span modulation conforms to findings of larger preview benefit in case of an easy compared to a difficult word before the boundary that have been attributed to immediate adjustments of the perceptual span with local processing load (Henderson & Ferreira, 1990; Kennison & Clifton, 2005). A dynamical adjustment of the perceptual span has also been used as an explanation of various parafoveal-on-foveal effects observed with multivariate analyses of reading using corpus analyses (Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006). The zoom lens model has been developed and tested in the context of differences in reading normal and randomly shuffled text. In particular, it can account for a rather unintuitive finding of a reversal of the word frequency effect in shuffled text reading (Schad, Nuthmann, & Engbert, 2010; Schad & Engbert, 2012).

The processing span in the present SWIFT variant

The processing span in SWIFT is modeled in two parts, a left-side extension δ_L and a right-side extension δ_R . In order to investigate the characteristics of SWIFT's processing span, we estimated three parameters that could be interpreted as three independent properties of the processing span. The right side of the span is dynamically modulated as a function of the lexical activation $a_k(t)$ of the foveal word k at time t , i.e.,

$$\delta_R = \delta_0 + \delta_1 \left(1 - \frac{a_k(t)}{A} \right), \quad (1)$$

where A denotes the maximum activation realized for the most difficult word in the corpus.

The parameter δ_0 denotes the constant portion of the processing span and defines the

smallest rightward extent at maximum foveal difficulty (i.e., focused span state). The parameter δ_1 indicates the strength with which the processing span dilates when the foveal processing load becomes smaller (i.e., defocused span state). Note that the dilation is set to zero as long as the current word's activation a_k has not reached the theoretical maximum of activation A . Thereafter, the dilation is proportionately increasing with decreasing activation a_k . The dilation is at its maximum when the current word's activation a_k decreased back to zero. The left side of the span is assumed to be constant and independent of any changes in lexical activations, i.e.,

$$\delta_L = \delta_0 \delta_2 \quad (2)$$

where the parameter δ_2 with $0 < \delta_2 < 1$ scales δ_0 to permit an asymmetry of the span even for maximum difficulty in foveal vision. This asymmetry parameter was absent in the SWIFT version developed by Schad and Engbert (2012), in which the asymmetry of the perceptual span was introduced by dynamic extension to the right only.

Given the span extent δ_L to the left and δ_R to the right, the processing rate λ for a letter of eccentricity ε relative to the fixation position at zero was determined by an inverse-parabolic relation according to the equation,

$$\lambda(\varepsilon) = \lambda_0 \begin{cases} 0 & : \varepsilon < -\delta_L \\ 1 - \varepsilon^2 / \delta_L^2 & : -\delta_L \leq \varepsilon < 0 \\ 1 - \varepsilon^2 / \delta_R^2 & : 0 \leq \varepsilon \leq \delta_R \\ 0 & : \delta_R < \varepsilon \end{cases} \quad (3)$$

Word-based processing rates at time t were then computed for each word n with length M_n from processing rates of each letter j following a nonlinear relation given by

$$\lambda_n(t) = (M_n)^{-\eta} \sum_{j=1}^{M_n} \lambda(\varepsilon_{nj}(t)). \quad (4)$$

Moreover, the processing span was normalized by $\lambda_0 = 3/(2(\delta_L + \delta_R))$, such that the total processing rate for all letters in the span was fixed at one. Note that this normalization has important implications. The broader the processing span, the smaller is the processing rate at the fixated position. This assumption relates to views of attention as a limited-capacity resource that can be allocated towards several processes simultaneously but at a cost of processing efficiency (Kahneman, 1973) and is also similar to the gradient shift hypothesis proposed by Inhoff, Eiter, and Radach (2005). Schad and Engbert (2012) showed with simulations that the reversed frequency effect observed in shuffled text reading (i.e., shorter rather than longer fixation durations on low-frequency compared to high-frequency words) was consistent with an increase in foveal processing efficiency when the processing span was small.

However, Schad and Engbert (2012) estimated the actual processing span within a larger perceptual region. They assumed that there exists a fixed-extent region of preprocessing (15 letters to the right of fixation as motivated by findings on the perceptual span size) in which words are preactivated slightly above zero. The model's processing span turned out to be substantially smaller than the total preprocessing region, rendering the processing span more comparable to the letter-identification span than the perceptual span (O'Regan, 1990; Underwood & McConkie, 1985). In order to test properties of the perceptual span instead, we implemented a model version in which words were preactivated not until they fell into the processing span. This forced the model to select the absolute size of its processing span such that it allowed an efficient preactivation of words to the right of fixation for becoming potential saccade targets and at the same time optimized the allocation of processing resources between foveal and parafoveal word processing. As a consequence, the best-fitting parameter values could be directly interpreted in analogy to properties of the perceptual span during reading. Estimating the modulation parameter δ_1 as zero would lead to a model with a constant span independent of any ongoing processing activities in foveal vision. Additionally estimating the asymmetry parameter δ_2 as one would suggest an

1
2
3
4 advantage for a symmetric rather than an asymmetric processing span with equal extent to
5 the left and to the right of fixation. Thus, the best-fitting parameter values speak to important
6 aspects of the processing span and inform us about what type of processing span is optimal
7 in a model with parallel-distributed attention such as SWIFT.
8
9

10 11 12 **Simulating the boundary paradigm**

13
14
15
16 Simulation of reading in the boundary paradigm requires assumptions about how the model
17 should behave when the boundary is crossed and the preview is replaced with the target
18 word. It should be noted that these assumptions imply no changes to the SWIFT model
19 architecture and are thus not counted as differences to SWIFT 3 as reported in Schad and
20 Engbert (2012). In order to simulate the occurrence of a display change in the invalid preview
21 condition, we made the following assumptions.
22
23
24
25
26
27

28
29 *Reset of lexical activation.* As parafoveal preview benefit is interpreted as resulting
30 from a head start of processing based on trans-saccadic integration of parafoveal information
31 (Inhoff, 1990; Inhoff & Tousman, 1990; see also Inhoff & Radach, this volume), it is
32 reasonable to assume that word recognition processes need to be restarted for the target
33 word if it differs from its parafoveal preview (i.e., in case of invalid preview). Particularly when
34 using random-letter nonwords for invalid previews as in the present study, there should not
35 be much useful information to be integrated across saccades. We therefore reset the lexical
36 activation of the target word $n+2$ to zero when the model first fixated a word to the right of the
37 boundary (i.e., to the right of word n ; for similar assumptions see Kliegl & Engbert, 2003;
38 Pollatsek, Reichle, & Rayner, 2006). The nonword preview of word $n+2$ in the invalid preview
39 condition was assigned to a processing maximum equal to the value A of the most difficult
40 word in all experimental sentences accounting for the fact that nonwords are infrequent and
41 therefore difficult to process.
42
43
44
45
46
47
48
49
50
51
52
53
54

55
56 *Saccade cancellation.* Moreover, the physical change of the display – although in
57 principle occurring during a saccadic movement – might affect the oculomotor system and
58
59
60

1
2
3
4 inhibit upcoming saccades (similar to Reingold & Stampe, 2004; Yang & McConkie, 2001).
5
6 Thus, we assumed that with a probability of 0.25 an ongoing saccade program was cancelled
7
8 when the model exceeded the boundary and when the saccade program was still in its early
9
10 labile stage (see also Nuthmann, Smith, Engbert, & Henderson, 2010). We will report the
11
12 simulation results (1) for a model with lexical reset only and (2) for a model with both lexical
13
14 reset and saccade cancellation. For simulations with different parameter values for the
15
16 activation reset and the saccade cancellation see Figure A2 and A3 in the Appendix.
17
18

19 **Summary of model adaptations**

20
21 In order to examine the perceptual span given parallel-distributed processing, we used
22
23 SWIFT 3 as a starting model. With respect to the architecture of the model, we mainly
24
25 changed two aspects compared to its predecessor. As the major goal was to estimate the full
26
27 size of the model's perceptual span, the preactivation of a constant 15-letter processing
28
29 region was omitted. We expected that δ_R would now cover the full extent of parafoveal
30
31 preactivation necessary for the SWIFT model to show adequate eye-movement behavior.
32
33 Moreover, to test the importance of a constant left-side asymmetry of the perceptual span for
34
35 a parallel-processing model like SWIFT an additional asymmetry parameter δ_2 was inserted.
36
37
38

39 (Table 1 about here)

40
41 Further differences to SWIFT 3 were resulting from fixing parameter values rather
42
43 than from changes in the model architecture. In order to reduce the total number of free
44
45 parameters, we fixed several model parameters at values reported for SWIFT 3 in Schad and
46
47 Engbert (2012) for normal text reading (see Table 1 for a complete overview). For example,
48
49 parameters determining saccade programming such as the mean duration for the labile and
50
51 nonlabile saccade-programming stages were taken from SWIFT 3 and thus did not differ
52
53 between model variants. However, we fixed two further parameters to different values
54
55 compared to SWIFT 3: (1) The global inhibition parameter (*ppf*) that slows down word
56
57 processing to the right of fixation proportional to the residual activation of words to the left of
58
59
60

1
2
3
4 fixation was set to zero and therefore rendered inactive. (2) The parameter for activation
5 transfer across saccades (τ) were set to one (i.e., no activation reset) such that the full
6 amount of parafoveally acquired information would be available in the next fixation. Both
7 changes were made to enable substantial parafoveal preprocessing in the model and to
8 allow processing to the right of fixation to be most effective on the model's dynamics.
9

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Other parameters were freely estimated for the present simulation but did not change
the architecture of the model. Specifically, we fit all parameters that were associated with the
model's processing span and those that were related to the lexical processing in the model.
The latter was optimized for the present stimulus material because it differed from that used
in Schad and Engbert (2012). Finally, the present model was getting along without
information about word predictability. Note that Schad and Engbert (2012) showed that
SWIFT 3 was able to perform adequately with the predictability parameter θ set to zero.
Moreover, predictability as typically measured in the cloze task refers to the upcoming word's
predictability (i.e., of word $n+1$), which seems to be of lesser importance when the interest is
in preprocessing of the preceding word $n+2$. How well the present SWIFT model accounts for
eye-movement statistics in reading (e.g., fixation durations and fixation probabilities for
different word lengths and word frequencies) is summarized in Figure A1 in the Appendix.

Results

Best-fitting model parameters

Parameters were estimated using a genetic algorithm procedure (see Mitchell, 1998). We ran
seven independent parameter estimations and computed the means and standard errors for
each of the nine free parameters. Table 1 summarizes the results for both free and fixed
parameter estimates and provides a comparison with the parameters of the zoom-lense
model (SWIFT 3) for normal text reading presented in Schad and Engbert (2012). As a first
result, the comparison of the best-fitting parameters showed that the constant span size was
estimated to be much larger than in the normal text reading condition investigated by Schad

1
2
3
4 and Engbert (2012). This is mainly due to the fact that SWIFT 3 used a preactivation zone of
5
6 15 letters to the right of fixation, while in the current version we neglected the preactivation
7
8 mechanism and limited processing to the span defined in Eqs. (1-3). As a result, the model
9
10 needed an increased size of the constant span parameter to generate sufficient parafoveal
11
12 information for the selection of saccade targets in the periphery. One implication of a larger
13
14 constant span is that more letters to the right of fixation are preprocessed on any fixation
15
16 irrespective of the difficulty of the ongoing foveal processing. At the same time, due to the
17
18 normalization of the inverse-parabolic processing span, foveal processing will always be
19
20 slower than in a model consisting of a smaller span and the fixated word will be assigned to a
21
22 smaller processing rate.
23

24
25 As a second result, the model showed on top of that some contribution of the dynamic
26
27 span modulation relative to foveal processing. The processing span could increase up to five
28
29 letters beyond the constant span size if the foveal processing demand dropped to zero
30
31 during a fixation. Finally, the asymmetry parameter was estimated to be substantially smaller
32
33 than one suggesting that even in its most focused state with minimum span size (i.e., in the
34
35 case of highest lexical activation of the fixated word) the optimal processing span was
36
37 asymmetric around the fixation position. Thus, even though the genetic algorithm was free to
38
39 select between different processing span types, the parameter estimates suggest that the
40
41 optimal processing span for the present SWIFT model consists of an asymmetric constant
42
43 part and an additional processing-related modulation of the right side. More importantly, the
44
45 estimates of the model's processing span were in a range that agrees well with what has
46
47 been reported in studies investigating the perceptual span experimentally. We will discuss
48
49 this in more detail below.
50

51 52 **Simulated $n+2$ preview effects in the boundary paradigm**

53
54 The goal of this study was to simulate reading in the boundary paradigm using the SWIFT
55
56 model and to compare the results with the experimentally observed $n+2$ preview effects.
57
58 Typically, experimental observations of $n+2$ preview effects are aggregated condition means
59
60

1
2
3
4 of fixation durations to approximate stable estimates of the fixation durations in the valid and
5 invalid preview condition. However, one experiment represents only a single realization from
6 the unknown distribution of true $n+2$ preview effects. With a computational model, we can
7 simulate many realizations and obtain a distribution of the means across different simulation
8 runs. In order to compare the distribution of simulated results to the experimental outcome,
9 we can then use likelihood computations to quantify how likely the experimental mean is
10 given the SWIFT model. Therefore, we ran 100 independent simulations with 68 virtual
11 subjects each reading 160 sentences, half of them with valid preview (i.e., the identical word)
12 and the other half with invalid preview of word $n+2$ (i.e., a random-letter nonword). The
13 simulated data was then compared to the empirical data of 30 young adults from the study of
14 Kliegl et al. (2007) and 38 young adults from the study of Risse and Kliegl (2011), all reading
15 the same sentence material in the same experimental setup of the $n+2$ boundary paradigm
16 comparing valid identical previews with invalid nonword previews.
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 (Figure 2 about here)
32
33

34 Figure 2 summarizes the present simulation results, and across rows shows the
35 results for word n before the boundary, word $n+1$ after the boundary, and the target word $n+2$
36 after word $n+1$ was previously fixated or skipped. The first two columns compare the
37 simulation results (distributions) and experimental means (vertical lines) on the level of
38 fixation durations (i.e., gaze durations) for the valid (green) and invalid (red) preview
39 condition. Across columns, we successively test the assumptions with respect to the display
40 change. The first column displays results from the model with lexical reset only (i.e., without
41 saccade cancellation), the second column shows corresponding results from the model
42 implementing both lexical reset and saccade cancellation after gaze-contingent display
43 change. The right column illustrates the mean differences between the valid and invalid
44 condition and therefore the size of the $n+2$ preview effects on each word in the target region.
45
46
47
48
49
50
51
52
53
54
55

56 The results clearly show that the SWIFT model generates preview effects of word $n+2$
57 in the boundary paradigm similar to what was observed in experiments. First, the model
58
59
60

1
2
3
4 produced a differential effect on the target word $n+2$ contingent on fixating or skipping word
5 $n+1$ (see bottom rows of panels of Figure 2): There was a substantial preview benefit on
6 word $n+2$ when word $n+1$ was skipped and a weaker effect when word $n+1$ was fixated. This
7 result is in good qualitative agreement with the experimental results, but, quantitatively, the
8 model clearly overestimated the magnitude of the preview benefit in both cases of skipping
9 and fixating word $n+1$ (i.e., the simulation distributions for valid and invalid preview are
10 further apart than the difference between the two corresponding vertical lines indicating the
11 experimental results).

20
21 Second, SWIFT also conformed well to the experimental results with respect to the
22 parafoveal-on-foveal effect of word $n+2$ (see top row of panels in Figure 2). There was
23 almost no difference in fixation durations between the valid and invalid preview condition on
24 word n before the boundary for both model variants with and without saccade cancellation.
25 Note that the experimental results exhibit a trend in the direction of such a parafoveal-on-
26 foveal effect which is small and not always significant (e.g., Risse & Kliegl, 2011).

33
34 Third, for fixations on word $n+1$ after the boundary, SWIFT produced a small
35 difference in fixation durations between the valid and invalid preview condition. This
36 difference was substantially smaller in the model variant without saccade cancellation than
37 the experimental effect. However, with the assumption that ongoing saccade programs are
38 cancelled with some probability after a display change, the distribution of fixation durations
39 for a nonword preview of word $n+2$ shifted significantly to the right and resulted in a
40 considerable increase in the $n+2$ preview effect on word $n+1$. Overall, both lexical reset and
41 saccade cancellation were thus required for a recovery of the qualitative profile of the
42 experimental effects.

52
53 This last result is an interesting finding because it suggests that nonlocal preview
54 effects after crossing the boundary can be viewed as effects that are driven significantly by
55 oculomotor inhibition due to the display change (O'Regan, 1990; Reingold & Stampe, 2004).
56 However, in a recent study we found fixation duration differences between two display-
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

change conditions when an easy preview changed to a difficult target word or vice versa (Risse & Kliegl, 2012). A difference between these two conditions suggests that such nonlocal $n+2$ preview effects (i.e., on word $n+1$ after the boundary) also involve higher-level cognitive processes and cannot be accounted for solely by low-level oculomotor inhibition. Moreover, saccade cancellation also increased fixation durations on word $n+2$ if word $n+1$ was skipped and thus contributed even more to the model's overestimation of preview benefit on word $n+2$.

In summary, fitting the model to data from normal reading only (i.e., with a valid preview of word $n+2$) and assuming a reset of lexical activation of word $n+2$ after its replacement (i.e., after an invalid preview of word $n+2$), the present version of the SWIFT model was able to account for effects observed from reading with invalid preview of word $n+2$ in the boundary paradigm. In some cases, the observed experimental means differed considerably from the overall simulated mean fixation durations for the two preview conditions (e.g., on the preboundary word n) and their likelihoods were very small given the SWIFT model. Yet the differences between the preview conditions were captured quite well by the model. Most critically, without directly fitting the preview effects, SWIFT simulations yielded larger preview benefit on word $n+2$ after skipping word $n+1$. With the additional assumption of saccade cancellation due to the display change, they also showed an effect on word $n+1$. However, the simulated $n+2$ preview benefit overestimated the experimental results by far. The latter type of misfit may well be due to somewhat ad-hoc assumptions about how the model was to respond to the display change specific for the boundary paradigm. We will return to this issue in the General Discussion.

Investigation of processing-span parameters

In this section, we report analyses on how the processing-span parameters account for parafoveal processing of word $n+2$ and what they contribute to the $n+2$ preview effects in the boundary paradigm. To this end, we focus on the preview benefit on word $n+2$ in the simpler model without saccade cancellation after display changes, taking also into account whether

1
2
3
4 word $n+1$ was fixated or skipped. We carried out a parameter-range investigation (similar to
5 a grid search) and computed the preview benefit for simulations with varying processing-
6 span parameters (i.e., constant part δ_0 , dynamical part δ_1 , and asymmetry δ_2 of the
7 perceptual span), keeping all other parameters constant.
8
9

10
11
12
13 (Figure 3 about here)
14
15

16 The most prominent result of these simulations was the difference in the evolution of
17 preview benefit on word $n+2$ after word $n+1$ was fixated or skipped. Increasing the span
18 parameters and thus the size of the processing span had a strong impact on the preview
19 benefit after word $n+1$ was skipped (see Figure 3, right column). However, when word $n+1$
20 was fixated the preview benefit on word $n+2$ was much weaker and started to increase with
21 much greater span sizes. Note that the absolute right part of the span consists of both the
22 constant span part δ_0 and the dynamic part δ_1 adding up to what is illustrated in the left
23 column of Figure 3. Interestingly, the single fixation durations (solid light-green lines) were
24 much less affected by the processing span size than the gaze durations (solid dark-green
25 lines) and showed a preview benefit maximally up to 15 ms at the largest span size.
26
27

28 The most interesting result, in our opinion, can be traced back to varying the
29 parameter of the constant part δ_0 of the processing span (see Figure 3, upper row).
30 Increasing δ_0 the parafoveal preview benefit first became increasingly large, reached
31 asymptote, and then slowly decreased again. That more parafoveal preview is obtained
32 when the span is larger seems trivial. But how can the preview benefit decrease again? The
33 illustration of the processing span (see Figure 3, left panel) suggests that with a large
34 processing span the SWIFT model assigns comparable processing rates to all words
35 residing in the span. The processing rates for each word are low and indicate overall slow
36 word-recognition at each location. In other words, lexical processing speed does not vary
37 systematically between words irrespective of whether they are in foveal or parafoveal vision.
38
39 When activation of word $n+2$ has been reset to zero after the display change and word $n+2$ is
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 then fixated, its activation will rise only slowly in foveal vision. In the case of identical preview
5 the activation of word $n+2$ will continue to rise after its fixation conserving the benefit from its
6 prior preprocessing. However, as foveal processing speed is slow, lexical activation might
7 not increase enough to exceed the processing maximum after which activation is decreasing
8 again. As a consequence, word $n+2$ may be activated higher after identical preview than
9 after a nonword preview, and this may result in a higher likelihood to inhibit the next saccade
10 program and prolong $n+2$ fixations in the case of valid preview rather than invalid preview.
11
12
13
14
15
16
17
18

19
20 In contrast, increasing the parameter value of the dynamic part δ_1 of the processing
21 span and thus the dynamic response of the processing span towards foveal processing
22 difficulty lead only to moderate changes in the preview benefit on word $n+2$ (see Figure 3,
23 middle row). Yet, although the parameter range tested here resulted in processing-span
24 sizes that were even larger than those covered in the analysis of the constant span part, the
25 according decrease in foveal processing efficiency did not affect the preview effect size as it
26 did affect it when δ_0 was varied. This suggests that the dynamic increase of the span leads
27 to a qualitatively different pattern of preview effects in the present model than a constant
28 span of the same size. The fact that the span dilates and broadens only after foveal
29 processing has reached the activation maximum and lexical activations are decreasing again
30 seems to prevent preview benefit from turning into preview cost.
31
32
33
34
35
36
37
38
39
40
41
42

43 Finally, changing the asymmetry parameter δ_2 had no strong effect on parafoveal
44 preview benefit when word $n+1$ was fixated (see Figure 3, lower row). The span's proportion
45 to the left side relative to the constant part of the right side did not contribute much to
46 parafoveal processing of word $n+2$. However, when word $n+1$ was skipped, increasing the
47 scaling parameter δ_2 from 0.1 to 1.0 and thus increasing the left side of the processing span
48 until it was equal to the right side in its most focused extent resulted in a constant decrease
49 of preview benefit on word $n+2$. As detailed above, a decrease in preview benefit is a result
50 of the normalization of the processing span leading to a trade-off between foveal and
51
52
53
54
55
56
57
58
59
60

1
2
3
4 parafoveal processing rates. Taken together, the present findings suggest that up to a certain
5 span size the SWIFT model can compensate the decrease in foveal processing rate by
6 increasing parafoveal processing rates. However, when the foveal processing rate falls
7 below a critical value it results in decreasing preview effects of parafoveal processing.
8
9
10
11

12 13 **General Discussion**

14
15 In summary, we showed that the present version of the SWIFT model generates reliable
16 preview benefit from word $n+2$ in the boundary paradigm without directly fitting the model to
17 such effects. To this end, we made the assumption that a display change causes a complete
18 reset of lexical activation. With the additional assumption that display changes cause
19 saccade cancellation with a 0.25 probability—a mechanism that has proven to be important
20 also in other models to explain results from gaze-contingent experiments (Nuthmann et al.,
21 2010)—SWIFT also produced a nonlocal preview effect of word $n+2$ in fixation durations on
22 word $n+1$. Therefore, in its present version, the SWIFT model could account for a highly
23 specific pattern of results in the $n+2$ boundary experiments, that is, a larger preview benefit
24 on the target word when the pretarget word $n+1$ was skipped rather than fixated and an
25 effect on word $n+1$ if this was fixated first.
26
27
28
29
30
31
32
33
34
35
36
37

38 **The role of the processing span parameters in the present SWIFT version**

39
40 Many aspects of the results have already been discussed above. In the following, we want to
41 focus only a few issues most relevant to the present simulation approach. To investigate the
42 role of the different processing-span parameters on parafoveal processing in the SWIFT
43 model we defined the parameters such that they tested three independent properties of the
44 processing span: a constant span part δ_0 , a dynamic adjustment into the direction of reading
45 δ_1 , and an additional leftward asymmetry δ_2 . In principle, the current implementation allowed
46 the possibility that the model required only a subset of these parameters to account for the
47 reading data. If δ_1 and δ_2 were estimated to be 0 and/or 1, respectively, this would indicate,
48 counter to our expectations, that there is no need to allow for a dynamic modulation of the
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4 perceptual span and/or that a symmetric processing span is sufficient when word-processing
5 is parallel in a reading model such as SWIFT. The results suggested, however, that all three
6 parameters contributed to an optimal model fit. A constant asymmetric processing span with
7 a broader extent to the right than to the left (i.e., the combination of δ_0 and δ_2) was not
8 sufficient for the SWIFT model to optimally account for the present eye-movement data
9 under conditions of identical preview. Word processing benefited from a dynamic modulation
10 of the processing span in response to foveal processing load, although this modulation might
11 seem rather small in the present simulations (about 5 letters at its maximum). However, it
12 was substantially larger than what was previously observed for normal text reading (Schad &
13 Engbert, 2012).
14
15
16
17
18
19
20
21
22
23
24

25
26 Otherwise, there was a high consistency with the SWIFT 3 model. This is particularly
27 striking as the parameters were estimated based on reading very different sentence material.
28 In contrast to the sentence corpus used in Schad and Engbert (2012), the experimental
29 sentences used in the present study each involved a three-word target region with a short
30 three-letter word $n+1$ before a medium-length target word $n+2$ ($M = 5$ letters). Thus, the
31 likelihood of encountering three-letter words in the experimental sentences was artificially
32 high due to requirements in the $n+2$ boundary paradigm. Finding strong agreement in
33 parameter estimates across reading material supports the validity and importance of the
34 dynamic adjustment of SWIFT's processing span in accounting for eye movements during
35 reading. In addition, the present simulation showed realistic values of parameter estimates of
36 the processing span. In fact, the estimated parameters resulted in a processing span with a
37 left extension of 3.9 letters and a maximum right extension of 19.5 letters (minimum 14.5
38 letters) and were thus very close to the experimental estimates of the perceptual span. Such
39 coherence between estimated model parameters and experimentally observed measures of
40 the perceptual span is an encouraging example of how eye-movement data may constrain
41 parameters in computational modeling to reasonable values and increase our understanding
42 of the complex theoretical concepts they are representing.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Overestimation of $n+2$ preview benefit in the present SWIFT simulation

The present simulations were based on the idea to not fit the model to the preview effects in the boundary experiment but to examine whether a model optimized for the normal reading condition would succeed in showing preview effects even from word $n+2$ at the spatial limits of the perceptual span. Qualitatively, the present SWIFT version well captured a complex pattern of local and nonlocal effects of previewing word $n+2$ in the boundary paradigm. Yet not surprisingly, there were also a few noteworthy discrepancies, particularly on the quantitative level between the experimental and simulated preview effects. The size of the preview effects of word $n+2$ were simulated reasonably well for fixations on word n and word $n+1$ (given saccade cancellation). However, simulated preview effects on the target word $n+2$ diverged seriously from the empirical observations. We propose two reasons.

This misfit may have been partially due to reading times on the preboundary word n that were overestimated by about 20 ms across all conditions. Overall however, the present simulations resulted in an adequate fit of fixation durations for different word lengths and word frequencies of the words used in the experimental sentences (see Figure A1, Appendix). Thus, it seems unlikely that the misfit of fixation durations on the preboundary word may have been a result of the absence of predictability information in the present version of the SWIFT model or of the fixing of parameters. We suspect that it reflects word-specific biases in the present sentence material such as parafoveal-on-foveal differences from an upcoming three-letter content or function word at position $n+1$. Such differences may currently not be adequately represented in the lexical activations of the SWIFT model and may originate from beyond linguistic word properties such as length and frequency.

In addition, the SWIFT model overestimated the fixation durations on the target word in case of invalid preview of word $n+2$ and skipping word $n+1$ by a similar magnitude of 20-30 ms compared to the experimentally observed viewing times. This overestimation in the model might be a consequence of how we simulated reading in the boundary paradigm with the SWIFT model. We implemented two assumptions: (1) We assumed that a display change

1
2
3
4 of a nonword preview to the target word would reset the lexical activation of word $n+2$ to
5 zero. This assumption implies that no information about the parafoveal preview could be
6 used on the next fixation after the boundary and word processing has to start completely
7 from zero. However, even nonword previews share the word-length information with the
8 target word. As the simulations in the Appendix (Figure A3, upper row) show, resetting the
9 activation of word $n+2$ to zero may have overestimated the disadvantage of having
10 preprocessed an invalid preview in parafoveal vision and thus inflated the nonword preview
11 cost. (2) We assumed that saccade programs were cancelled after a display change (i.e., in
12 the nonword preview condition) with a probability of 0.25. Notwithstanding the fact that
13 saccade cancellation led to a reasonable preview effect on word $n+1$ when it was fixated first
14 after the boundary, cancelling ongoing saccade programs also increased fixation durations
15 on word $n+2$ when it was fixated first (see simulations in Figure A3, lower row, Appendix). As
16 a consequence, this also added to the already overestimated fixation durations on word $n+2$
17 in case of a nonword preview.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 **Solutions for future simulations with the SWIFT model**

34
35
36 Both the lexical activation reset and the saccade cancellation with a probability of 0.25
37 contributed to the overestimation of preview cost on word $n+2$. Moreover, both parameters
38 were fixed and chosen for considerations of plausibility. However, the results in the Appendix
39 suggest that changing these values directly affects the size of the preview effects from word
40 $n+2$. One option for future simulations is, therefore, to estimate the optimal amount of rest
41 activation that should be preserved after the boundary as well as the optimal probability of
42 saccade programs that should be cancelled after a display change as two free parameters in
43 the model. However, such a direct fit of preview benefit or preview cost is beyond the
44 approach advocated here of exploring effects of parafoveal processing that are already
45 inherent in the SWIFT model architecture without fitting the model to them.
46
47
48
49
50
51
52
53
54
55

56
57 Alternatively, one could design experiments to inform about how much activation is
58 remaining after a display change and how many saccade cancellations seem reasonable. An
59
60

1
2
3
4 index for differences in activation reset may be approximated from studies comparing
5 preview benefit for different types of parafoveal previews (e.g., semantically anomalous,
6 unpredictable, or visually dissimilar previews as in Drieghe, Rayner, & Pollatsek, 2005). We
7 may also be able to obtain independent estimates of the probability of saccade cancellation
8 from saccadic inhibition experiments as reported in Reingold and Stampe (2004). Our
9 prediction is that simulations using such experimental parameter values will result in a
10 reduction of the preview benefit of word $n+2$ and thus eliminate the overestimation of fixation
11 durations after the display was changed.
12
13
14
15
16
17
18
19

20 21 **Summary**

22
23 The present simulations showed that parafoveal processing up to word $n+2$ is inherent to a
24 parallel model such as SWIFT without particularly fitting its parameters to the experimental
25 preview effects. Moreover, based on the normal reading condition, the SWIFT model favored
26 a processing span that was both asymmetric around the fixation position and dynamically
27 modulating its rightward extent with respect to the processing demand of the word in foveal
28 vision. The estimated size of SWIFT's processing span was consistent with experimentally
29 derived measures on the perceptual span during reading. Our simulation results confirm that
30 computational models provide an intriguing and compelling way to test different
31 conceptualizations of cognitive processes (i.e., related to different theoretical properties of
32 the perceptual span) that are experimentally difficult to obtain.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Angele, B., & Rayner, K. (2011). Parafoveal preprocessing of word n+2 during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1210-1220.
- Angele, B., Slattery, T., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Visual Cognition*, 16, 697-707.
- Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 954-969.
- Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *The Quarterly Journal of Experimental Psychology*, 61, 1239-1249. doi:10.1080/17470210701467953
- Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, 42, 621-636.
- Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777-813. doi:10.1037/0033-295X.112.4.777
- Engbert, R., & Krügel, A. (2010) Readers use Bayesian estimation for eye-movement control. *Psychological Science*, 21, 366-371.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention - a zoom lens model. *Perception and Psychophysics*, 40(4), 225-240.
- Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movements preparation. *Psychological Review*, 109, 545-572.
- 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.
- 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.

- 1
2
3
4 Hyönä, J. (2011). Foveal and parafoveal processing during reading. In: S.P. Liversedge, I.D.
5
6 Gilchrist, & S. Everling (eds.), *Eye Movements Handbook* (pp. 819-838). Oxford:
7
8 Oxford University Press.
9
- 10
11 Inhoff, A.W. (1990). Integrating information across eye fixations in reading. *Acta*
12
13 *Psychologica*, 73, 281–297. doi:10.1016/0001-6918(90)90027-D
14
- 15 Inhoff, A.W., Eiter, B.M., & Radach, R. (2005). The time course of linguistic information
16
17 extraction from consecutive words during eye fixations in reading. *Journal of*
18
19 *Experimental Psychology: Human Perception and Performance*, 31, 979–995.
20
- 21 Inhoff, A.W., & Radach, R. (2014). Parafoveal preview benefits during silent and oral reading:
22
23 Testing the parafoveal information extraction hypothesis. *Visual Cognition*, this volume.
24
- 25 Inhoff, A.W., & Tousman, S. (1990). Lexical integration across eye fixations in reading.
26
27 *Psychological Research*, 52, 330-337.
28
- 29 Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
30
- 31 Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision*
32
33 *Research*, 45, 153-168.
34
- 35 Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and
36
37 low working memory capacity readers: Implications for eye movement control. *Journal*
38
39 *of Experimental Psychology: Learning, Memory, and Cognition*, 21, 68–81.
40
- 41 Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In J. Hyönä, R. Radach, & H. Deubel
42
43 (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp.
44
45 103–117). Oxford, England: Elsevier.
46
- 47 Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the
48
49 influence of past, present, and future words on fixation durations. *Journal of*
50
51 *Experimental Psychology: General*, 135, 12-35.
52
- 53 Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects
54
55 from word n+2. *Journal of Experimental Psychology: Human Perception and*
56
57 *Performance*, 33, 1250–1255. doi: 10.1037/0096-1523.33.5.1250
58
- 59 Krügel, A., & Engbert, R. (2010) On the launch-site effect for skipped words during reading.
60
61 *Vision Research*, 50, 1532-1539.
- 62 Krügel, A., Vitu, F., & Engbert, R. (2012). A single space makes a large difference: Fixation
63
64 positions after skipping saccades. *Attention, Perception, & Psychophysics*, 74(8),
65
66 1556-1561.

- 1
2
3
4 LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification.
5 *Psychological Review*, 96(1), 101-124.
6
7
8 McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in
9 reading. *Perception & Psychophysics*, 17, 578–586.
10
11 McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin*
12 *of the Psychonomic Society*, 8, 365–368.
13
14 Mitchell, M. (1998). *An introduction to genetic algorithms*. MIT Press.
15
16 Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational
17 model of fixation durations in scene viewing. *Psychological Review*, 117, 382-405. doi:
18 10.1037/a0018924
19
20 O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and*
21 *their role in visual and cognitive processes* (pp. 395–453). Amsterdam: Elsevier.
22
23 Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring
24 the interface between cognition and eye-movement control. *Cognitive Psychology*, 52,
25 1–56. doi:10.1016/j.cogpsych.2005.06.001
26
27 Radach, R., Glover, L., Vorstius, C., & Inhoff, A. W. (2011). Contextual constraint and N+2
28 preview effects in reading. *Quarterly Journal of Experimental Psychology*, 66, 619-633.
29 doi: 10.1080/17470218.2012.761256
30
31 Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive*
32 *Psychology*, 7, 65–81. doi:10.1016/0010-0285(75)90005-5
33
34 Rayner, K., Juhasz, B.J., & Brown, S.J. (2007). Do readers obtain preview benefit from word
35 n+2? A test of serial attention shift versus distributed lexical processing models of eye
36 movement control in reading. *Journal of Experimental Psychology: Human Perception*
37 *and Performance*, 33, 230-245.
38
39 Rayner, K., McConkie, G.W., & Zola, D. (1980). Integrating information across eye
40 movements. *Cognitive Psychology*, 12, 206-226.
41
42 Rayner, K., Well, A.D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in
43 reading. *Perception & Psychophysics*, 27, 537-544.
44
45 Reichle, E.D. (2011). Serial-attention models of reading. In: S.P. Liversedge, I.D. Gilchrist, &
46 S. Everling (eds.), *Eye Movements Handbook* (pp. 767-786). Oxford: Oxford University
47 Press.
48
49 Reichle, E.D., Pollatsek, A., Fisher, D.L., & Rayner, K. (1998). Toward a model of eye
50 movement control in reading. *Psychological Review*, 105, 125-157.
51
52
53
54
55
56
57
58
59
60

- 1
2
3
4 Reingold, E. M., & Stampe, D. M. (2004). Saccadic inhibition in reading. *Journal of*
5 *Experimental Psychology: Human Perception and Performance*, 30, 194-211. doi:
6 10.1037/0096-1523.30.1.194
7
8
9 Richter, E., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems*
10 *Research*, 7, 23-33.
11
12 Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading.
13 *Psychology and Aging*, 26, 451-460. doi:10.1037/a0021616
14
15 Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word
16 n+2 in reading. *Journal of Experimental Psychology: Human Perception and*
17 *Performance*, 38, 1026-1042. doi: 10.1037/a0027735
18
19
20
21 Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus
22 normal text reading using the SWIFT model. *Visual Cognition*, 20(4-5), 391-421.
23 doi:10.1080/13506285.2012.670143
24
25
26 Schad, D., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly
27 shuffled texts. *Vision Research*, 50, 2600-2616.
28
29
30 Schotter, E.R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention,*
31 *Perception, & Psychophysics*, 74, 5-35.
32
33 Schotter, E.R., Reichle, E., & Rayner, K. (2014). Re-thinking parafoveal processing in
34 reading: Serial-attention models can explain semantic preview benefit and n+2
35 preview effects. *Visual Cognition*, this volume.
36
37
38
39
40 Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control
41 during reading: Foveal load and parafoveal processing. *The Quarterly Journal of*
42 *Experimental Psychology*, 52A, 1021-1046.
43
44
45
46 Trukenbrod, H.A., & Engbert, R. (2013). ICAT: A computational model for the adaptive
47 control of fixation durations. *Psychonomic Bulletin & Review* (in press).
48
49
50 Trukenbrod, H.A., & Engbert, R. (2012). Eye movements in a sequential scanning task:
51 Evidence for distributed processing. *Journal of Vision*, 12(1): 5, 1-12.
52
53
54 Underwood, N.R., & McConkie, G.W. (1985). Perceptual span for letter distinctions during
55 reading. *Reading Research Quarterly*, 20, 153-162.
56
57
58 Yang, S.-N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade
59 initiation time. *Vision Research*, 41, 3567-3585.
60

Acknowledgments

Parts of this research were presented at the 46th Annual Meeting of the Society for Mathematical Psychology, August 4-7, 2013, at the University of Potsdam, Germany, and at the 17th European Conference on Eye Movements, August 11-16, 2013, in Lund, Sweden. This work was funded by Deutsche Forschungsgemeinschaft via Research Group 868 "Computational Modeling of Behavioral, Cognitive, and Neural Dynamics" (Grants EN 471/7, and KL 944/14).

Footnotes

1. All dynamical variables such as the random timer and the evolution of lexical activations of each word in the sentence were modeled as independent discrete random walk processes (see Nuthmann, Smith, Engbert, & Henderson, 2010; Trukenbrod & Engbert, 2013). Parallel processing was approximated through randomly incrementing the random walks.

Appendix

Here we provide additional analyses and simulations with the SWIFT model that (1) illustrate the goodness-of-fit of the present model variant with respect to global and local reading statistics and (2) test the assumptions we made regarding the model's response to display changes.

SWIFT's performance: Fixation duration and fixation probability statistics

Figure A1 compares the experimental (exp: dotted lines) and simulated (sim: solid lines) fixation durations and fixation probabilities for different word-frequency classes (left panels) and different word lengths (right panels) across all words in sentences read in the condition with valid preview of word $n+2$. This contains the subset of data to which the model was fit. The results show that although the model in its present version managed without predictability information (i.e., with θ set to zero) and also differed in some other aspects from the predecessor model SWIFT 3 (e.g., the global inhibition (ppf) parameter was rendered inactive), the global reading statistics derived from the simulated eye movements were in overall good agreement with the experimental results.

- Figure A1 about here -

In particular, simulated fixation durations showed the expected decrease in fixation times when fixating words of higher frequency (i.e., words belonging to a higher frequency class) and increasing durations with increasing word length (i.e., words belonging to a higher length class). The relative decrease or increase was most pronounced for total reading times containing the durations of all fixations on a word including rereading, and this was so both for the simulated as well as the experimental data. However, while the SWIFT model captured the average durations of single fixations very well (i.e., cases in which a word was fixated only once during the first left-to-right reading of the sentence), it overestimated the duration of the first of multiple fixation cases substantially, specifically for words of the highest frequency class (10,000 or more occurrences per million) and of the two lowest

length classes (2-4 letter words). This corresponds nicely with the fact that each sentence contained a three-letter word $n+1$ of which half were very high-frequent function words that could even repeat across sentences (e.g., “and” or “for”). As a consequence, readers might have shown an even stronger processing facilitation of these words reflected in even shorter fixation durations than could be explained purely on the basis of the frequency and length effects covered by the SWIFT model.

As can be seen in the bottom row of Figure A1, the SWIFT model successfully reproduced highly specific word-length and word-frequency effects for different fixation probability measures. Although the probability of single fixations was generally underestimated by the model, it clearly showed the general trends in the data. Most noteworthy seems the inverted U-shaped relation between word length and single fixation probability showing the most single fixation cases for medium word lengths.

- Table A2 about here -

Table A2 shows the comparison of the local reading statistics for the three words n , $n+1$, and $n+2$ in the target region. Also here, the simulated means for fixation durations and fixation probabilities were highly consistent with the empirical findings. The largest differences may be seen on word $n+2$. Nevertheless, the SWIFT model seems to produce a similar ratio of skippings, single fixations, and refixation cases in the target region as obtained in the present $n+2$ -boundary experiments.

Display-change assumptions: Testing activation reset and saccade cancellation

Resetting the lexical activation of word $n+2$ to zero after the display change implies the assumption that word-recognition processes have to restart from scratch when the target word replaces the nonword preview of word $n+2$. However, it is likewise plausible that some low-level information can still be used after the display change (e.g., the word-length information) and some amount of lexical activation of word $n+2$ may be preserved. Moreover, studies varying the degree of information overlap between preview and target have shown

1
2
3
4 that different previews can lead to different amount of preview benefit (e.g., Drieghe, Rayner,
5 & Pollatsek, 2005). Figure A3 (upper row) illustrates how the lexical reset affects the $n+2$
6 preview effects in fixations after the boundary. As expected, the effect sizes increase the less
7 activation is maintained after the display change. However, this relation is much stronger in
8 gaze durations (darkgreen lines) and weaker in single fixation durations (lightgreen lines),
9 likely because higher activation of word $n+2$ impacts stronger on the saccade-target
10 selection (i.e., favoring refixations of word $n+2$) and modulations of the single fixation
11 duration due to foveal inhibition are more moderate.
12
13
14
15
16
17
18
19

20
21 - Figure A3 about here -
22
23

24 Increasing the number of saccade cancellations after a display change (while
25 saccade programs are still in their labile stage) shows a similar effect (see Figure A3, bottom
26 row). Preview effects of word $n+2$ linearly increase with increasing number of cancelled
27 saccades. In contrast to the activation reset, saccade cancellation influences gaze durations
28 and single fixation durations alike. Additional saccade reprogramming costs are independent
29 of the fixation type as saccades are cancelled with a certain unbiased probability determined
30 only by the occurrence of a display change. As a consequence, they only affect the first
31 fixation after crossing the boundary and do not influence the preview effects on word $n+2$
32 when word $n+1$ was previously fixated (see Figure A3, bottom row right panel).
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Summary of parameter values in the present SWIFT model.

	<i>Parameters</i>	<i>Symbol</i>	<i>Dynamical span modulation</i>			
			<i>M</i>	<i>SE</i>	<i>Range</i>	<i>Schad & Engbert (2012)</i>
Lexical parameters	Frequency, intercept	α	13.1	0.57	1-20	18.7
	Frequency, slope	β	0.07	0.02	0-1	0.71
	Global inhibition	ppf	0.0			0.069
Processing span	Span, constant	δ_0	14.5	0.31	1-40	1.64
	Span, dynamic - foveal	δ_1	5.04	0.45	0-40	2.15
	Span, asymmetry	δ_2	0.27	0.01	0.1-1	–
Visual parameters	Visual preprocessing	$pspan$	–			15.0
	Word length exponent	η	0.91	0.02	0-1	0.3
	Preprocessing factor	f	0.81			0.814
	Global decay	ω	0.01			0.01
	Transfer across saccades	τ	1.0			0.50
	Eye-mind lag	pcd	30.0			30.0
Saccade timing	Random timing (ms)	t_{sac}	236.5	1.03	150-300	221.2
	Inhibition factor	h	0.48	0.07	0-5	0.549
	Target selection weight	γ	1.0			1.0
Saccade programming	Labile stage (ms)	τ_{lab}	100.0			100.0
	Refixation factor	$refix$	0.7			0.7
	Mislocation factor	$misfac$	0.75			0.75
	Nonlabile stage (ms)	τ_{nlab}	50.0			50.0
	Latency modulation	κ_0	2.5			2.5
	Latency modulation	κ_1	0.3			0.3
	Saccade execution (ms)	τ_{ex}	30.0			30.0

Note: Means (*M*) of all model parameters. For the free parameters (bold), also standard errors (*SE*) and parameter boundaries (*Range*) are provided.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

Table A2. Experimental (exp) and simulated (sim) means for fixation duration and fixation probability measures on word n , $n+1$, and $n+2$ in the target region.

	Word n		Word $n+1$		Word $n+2$	
	exp	sim	exp	sim	exp	sim
1st fixation duration [ms]	187	198	186	216	196	206
2nd fixation duration [ms]	172	174	156	155	176	166
single fixation duration [ms]	207	206	202	204	210	200
total time [ms]	252	257	214	221	264	235
skipping probability [p]	.06	.17	.47	.45	.13	.27
single fixation probability [p]	.81	.60	.52	.50	.74	.59
refixation probability [p]	.13	.23	.01	.06	.13	.14
regression probability [p]	.06	.06	.09	.09	.11	.05

Figure Captions

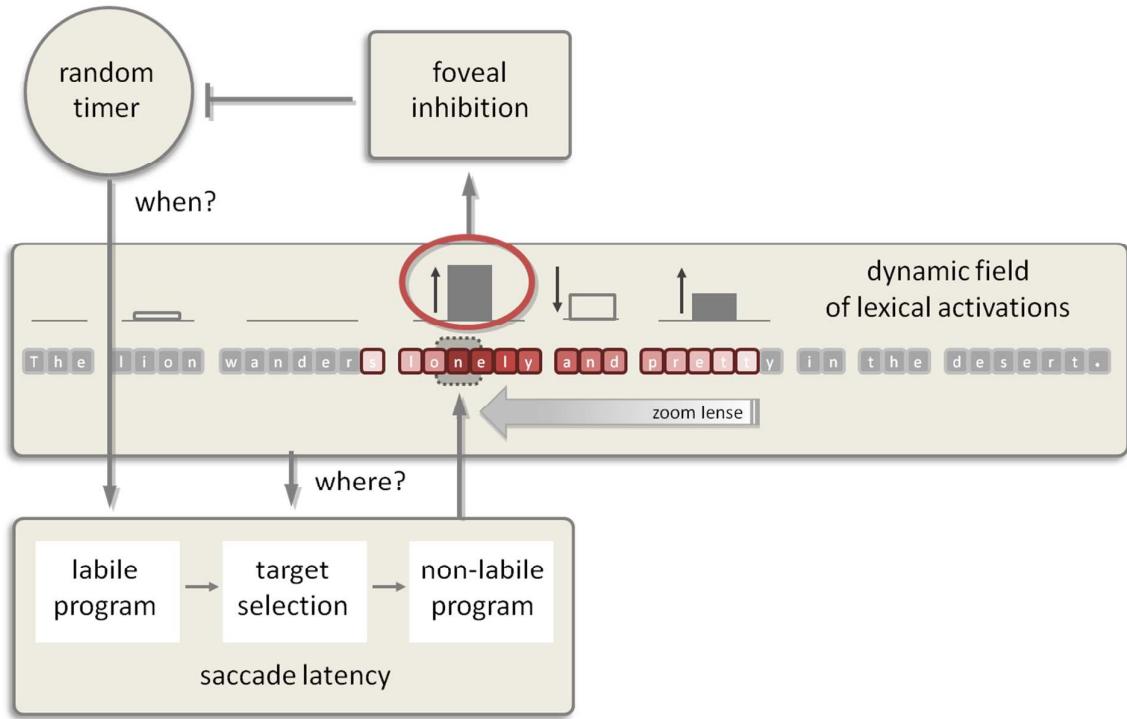
Figure 1: The SWIFT model. Illustration of components constituting the when- and where- pathways of saccadic control during reading. The model is illustrated at a stage at which lexical activation at the fixated word is increasing and the zoom-lense processing span is decreasing accordingly.

Figure 2: Summary of simulation results. Rows correspond to word n , $n+1$, and $n+2$. Note that for word $n+2$, we distinguish between cases when word $n+1$ was fixated or skipped. Illustrated is a comparison of simulations with and without saccade cancellation (due to display change) with results from two boundary experiments (Kliegl et al. 2007; Risse & Kliegl, 2011) testing conditions of valid (green) and invalid (red) preview. Vertical lines in left and middle column indicate the experimental means; probability densities are shown for the distribution of simulated fixation durations. The right column shows the likelihood for the experimental preview-benefit effect given the present simulation results. See text for more details.

Figure 3: The three panels in the left column show how varying parameter values for the constant part δ_0 (top row), the dynamic part δ_1 (bottom row), and the asymmetry of the perceptual span δ_2 (top row) affect the shape of the model's processing span at medium foveal activation. In each of these panels, the red curve indicates the processing span with the best-fitting parameter. The second and third columns illustrate how the preview benefit on word $n+2$ (difference between valid and invalid preview condition) with word $n+1$ being fixated (middle column) or skipped (right column) changes as a function of values for corresponding processing-span parameters. Horizontal dashed lines show the experimental preview benefit and solid lines the simulated preview benefit. Colors code two different fixation duration measures: single fixation durations (SFD; light green) and gaze durations (GD; dark green). Red vertical lines indicate the best-fitting parameter estimates.

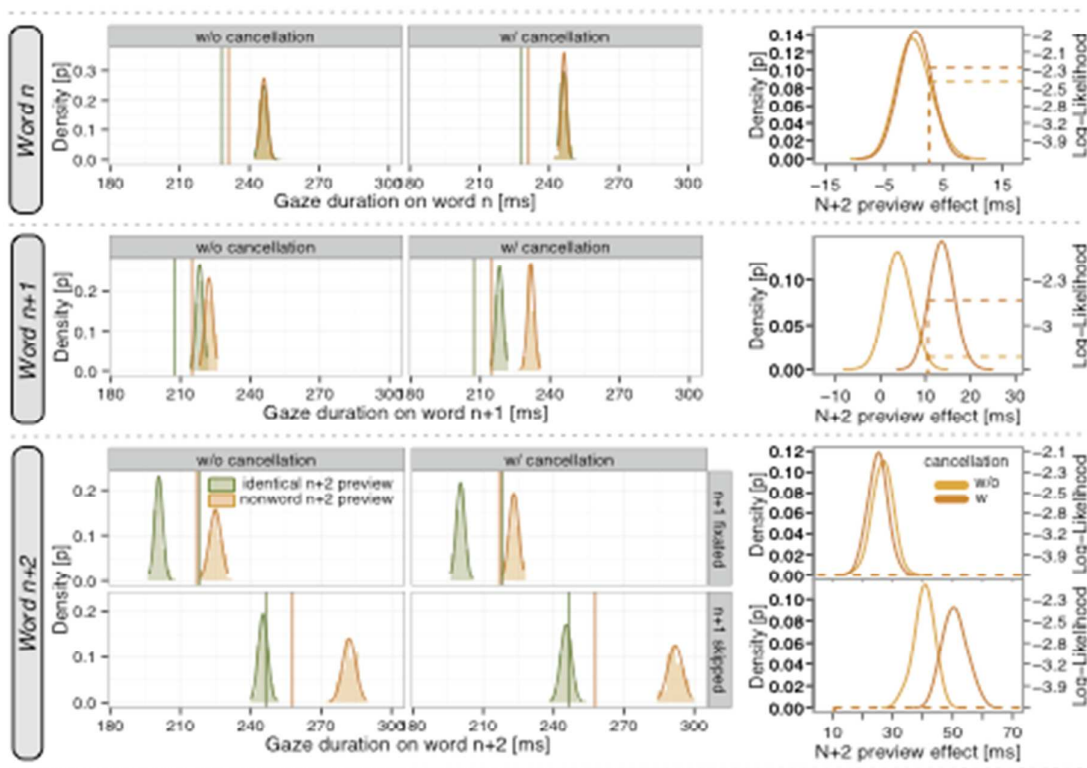
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1



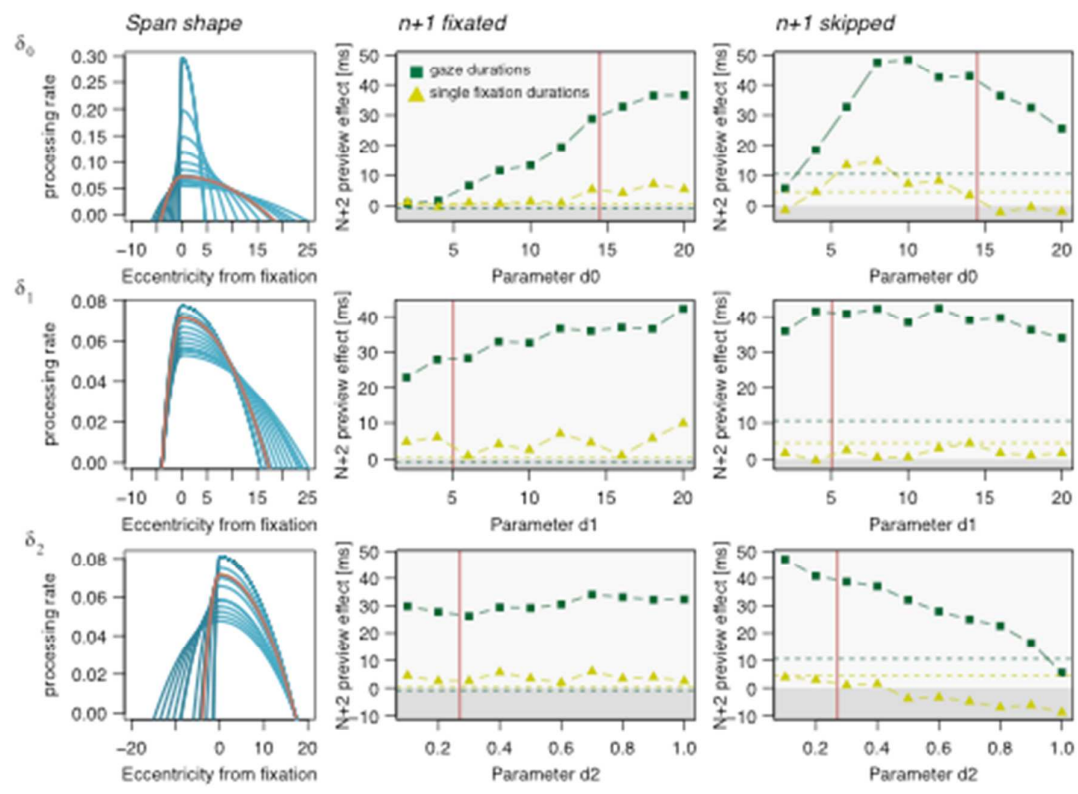
view Only

Figure 2



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 3



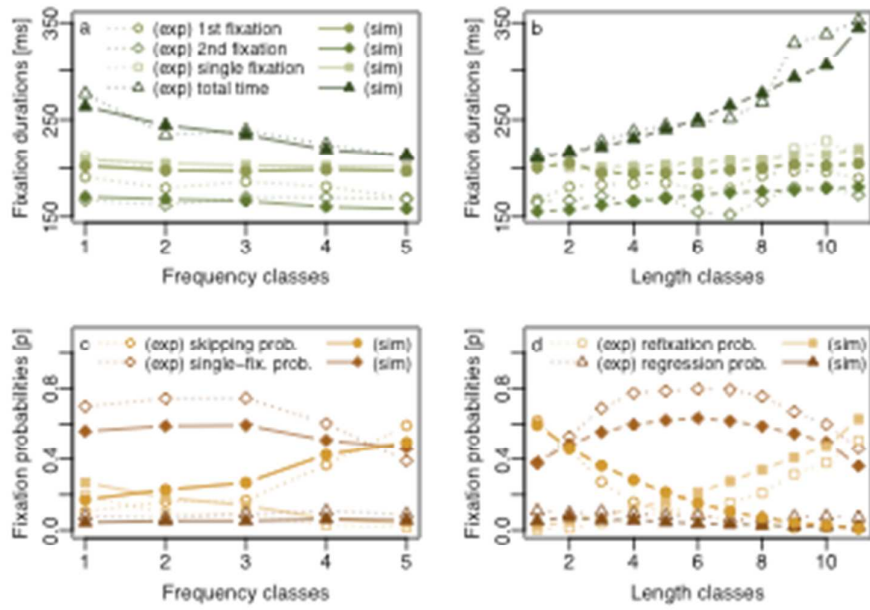
view Only

Figure Captions Appendix

Figure A1: Summary statistics of the SWIFT simulations of normal reading with identical preview in the $n+2$ boundary paradigm. The upper panels show the simulated (sim) and experimental (exp) data for various fixation duration measures as function of (a) word frequency class and (b) word length class. The lower panels show the comparison for various fixation probability measures for (c) word frequency class and (d) word length class.

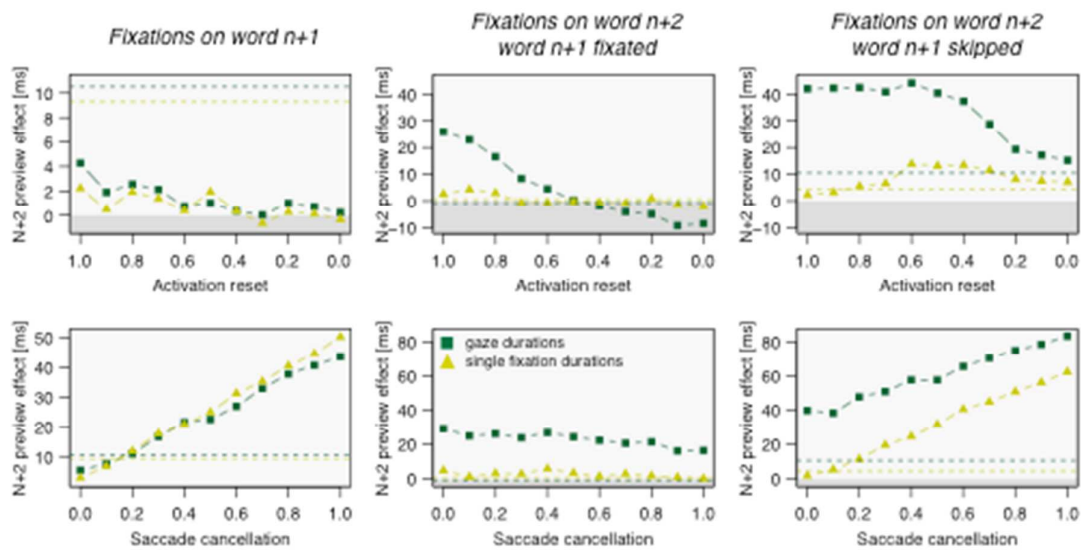
Figure A3: The influence of the display-change assumptions on the simulated preview effects of word $n+2$. The upper panels illustrate the variation on $n+2$ preview-effect size as a function of the activation reset after the display change. The lower panels show the change in preview effects according to increasing proportion of saccade cancellations in case of word replacements. Panels from left to right summarize the results for fixation durations on word $n+1$, on word $n+2$ after word $n+1$ was fixated, and after it was skipped. The dark-green solid lines show the simulated gaze durations, the light-green solid lines the simulated single-fixation durations, and dashed lines the respective experimental effect sizes.

Figure A1



Review Only

Figure A3



er Review Only