How Preview Space/Time Translates into Preview Cost/Benefit for Fixation Durations during Reading

Reinhold Kliegl[a], Sven Hohenstein[a], Ming Yan[a], & Scott A. McDonald[b]

University of Potsdam, Germany[a] & University of Strathclyde, Glasgow, Scotland[b]

Running head: Preview space/time effects on reading fixations

Corresponding author:

Reinhold Kliegl

Department of Psychology, University of Potsdam

Address: Universität Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany

Email: kliegl@uni-potsdam.de

Tel.: +49 331 997 2868

ABSTRACT

Eye-movement control during reading depends on foveal and parafoveal information. If the parafoveal preview of the next word is suppressed, reading is less efficient. A linear mixed model (LMM) re-analysis of McDonald (2006) confirmed his observation that preview benefit may be limited to parafoveal words that have been selected as the saccade target. Going beyond the original analyses, in the same LMM, we examined how the preview effect (i.e., the difference in single-fixation duration [SFD] between random-letter and identical preview) depends on the gaze duration on the pretarget word and on the amplitude of the saccade moving the eye onto the target word. The two key results were: (i) The shorter the saccade amplitude (i.e., the larger preview space), the shorter a subsequent SFD with an identical preview; this association was not observed with a random-letter preview. (ii) However, the longer the gaze duration on the pretarget word, the longer the subsequent SFD on the target, with the difference between random-letter string and identical previews increasing with preview time. A third pattern—increasing cost of a random-letter string in the parafovea associated with shorter saccade amplitudes—was observed for target gaze durations. Thus, LMMs revealed that preview effects, which are typically summarized under "preview benefit", are a complex mixture of preview cost and preview benefit and vary with preview space and preview time. The consequence for reading is that parafoveal preview may not only facilitate, but also interfere with lexical access.

How Preview Space/Time Translate into Preview Cost/Benefit for Fixation Durations during Reading

Fixation durations in reading are sensitive to a remarkably broad spectrum of processes relating to perceptual and oculomotor constraints as well as to attention, word recognition, and language comprehension. Many of these processes occur in parallel, but presumably some of them depend on the completion of others, introducing serial-order constraints. The prototype example for fundamental processes running in parallel during reading is the undisputed evidence that we process a word before we fixate it (e.g., Kennedy, 2000a, 2000b; Rayner, 2009). Much of the evidence for, and presumably also against, parallel processing has been obtained with the boundary paradigm in which a target word becomes available contingent on the direction of gaze (McConkie & Rayner, 1975; Rayner, 1975). In the incorrect preview condition of this paradigm, initially the position of a target word is occupied by a random-letter string; the target is displayed only when the eyes cross an invisible boundary before the target. Fixation durations on the target are up to 30 ms longer in this condition compared to normal reading (i.e., the identical preview condition). This saving of fixation time during normal reading is called preview benefit.

How to distinguish preview benefit and preview cost?

Of course, referring to the savings in subsequent fixation time as “preview benefit” presupposes that fixations following a random-letter preview serve as an adequate zero baseline. If the random-letter string incurred any processing cost (e.g., some kind of disturbance of normal processing due to the unfamiliarity of the preview), observed preview benefit as usually measured might actually result from preview cost due to invalid preview information or a combination of genuine preview benefit and such preview cost. In other words, there is an open question as to whether or not a random-letter preview serves as an adequate baseline for the computation of preview benefit.

One approach to shed light on this issue is to examine how preview benefit changes with *preview space* and *preview time*. By *preview space*, we refer to the incoming saccade length prior to the fixation on target word. If the launch site is close to the boundary there is much preview space, because more of the preview word falls into the perceptual span in comparison with a far launch site. By *preview time*, we refer to the elapsed time that a parafoveal word is visible before its post-boundary version is fixated.

Figure 1 displays three scenarios: (a) constant preview benefit (or preview cost) across preview space/time (along with a main effect of preview space/time), (b) increasing preview benefit because target fixation durations after a correct preview decrease as preview space/time increases, whereas fixations durations are constant after an incorrect preview, and (c) sum of preview benefit and preview cost because fixation durations decrease with increasing preview space/time after a correct preview, but also increase after an incorrect preview, as a consequence of greater interference. Statistically, (a) depicts a main effect of preview condition (i.e., fixation durations are shorter after identical compared with incorrect previews) and a main effect of preview space/time (i.e., fixation durations are longer for more remote launch sites; e.g., Heller & Müller, 1983) or after short pretarget fixation durations; (b) and (c) depict two possible types of interaction between preview condition and the covariate preview space/time.

------- Insert Figure 1 about here -------

These three scenarios are not typically distinguished in current research on eye movements during reading; any difference between incorrect and correct preview is referred to as "preview benefit". Discussions of benefits and costs and the difficulties of distinguishing between them have been published in the context of priming research (e.g., Jonides and Mack, 1984). We propose that coefficients estimated in linear mixed models (LMMs) provide inferential statistics that distinguish between these scenarios (among others), affording a distinction between preview benefit, preview cost, and effects that might best be described as resulting from a mixture of both.

Interference effects such as those listed in Figure 1c are especially important for computational models of eye-movement control during reading, none of which has implemented interference from parafoveal previews. Therefore, we view our analyses as setting the stage for the next generation of models.

Effects of preview space

There is already evidence that preview benefit is larger for close launch sites from the original publication (McDonald, 2006). Preview benefit as measured by gaze durations (GDs; the sum of fixation durations on a word during first-pass reading) was highly significant for launch sites of four characters or less, but, despite increasing constraints on visual acuity, it was still significant for saccades launched 9 or 10 letters before the target word. The pattern resembled the one in Figure 1b (i.e., GD decreases for correct preview space, but is stable for incorrect preview space).

Effects of preview time

Preview time may have effects similar to preview space. Schroyens, Vitu, Brysbaert, and d'Ydewalle (1999) examined the dependence of identical preview benefit on preview time. They presented a sequence of three words with an invisible boundary between the first and second word of the triad, manipulating preview of word N+1 during pre-boundary fixations on word N, and reported larger preview benefits on target word N+1 with increasing pretarget duration. The generalizability of these findings was questioned on account of the special kind of reading task used. White, Rayner, and Liversedge (2005) partitioned their data from a boundary paradigm study according to the median for participants and conditions, and did not obtain any significant modulation effect of preview time on preview benefit.

The first evidence for a change of preview benefit across preview time, operationalized as single-fixation duration (SFD; cases in which a word is inspected with exactly one fixation) on the pretarget word was reported by Yan, Risse, Zhou, and Kliegl (2012) in a Chinese reading study. For first-fixation duration (FFD; the duration of the first fixation on a word, irrespective of whether it was the only fixation or the first of several fixations) on the target word, preview benefit was stable across preview time (as illustrated in Figure 1a). For GDs, the difference between preview conditions increased significantly with preview time, but neither the increase for unrelated previews nor the decrease for related previews was significant (see Figure 1c). Interestingly, there was a significant increase in the refixation rate for unrelated previews with increasing preview time, indicating a preview cost for this alternative measure of GD. Finally, an interaction of preview time and a contrast between unrelated and semantically-related previews was also reported for reading Chinese (Yan, Zhou, Shu, & Kliegl, 2011); in this study, target FFDs were constant for unrelated previews but decreased for semantically-related previews. Thus, this study represents the scenario shown in Figure 1b, that is, increasing semantic preview benefit with preview time.

Effects of pretarget and target word-frequency

Preview benefit/cost may not only depend on preview space or time, but also on properties of the pretarget or target word (e.g., frequency). There is evidence compatible with a dynamical modulation of the perceptual span in response to foveal load. In the case of a high-frequency pretarget word (implying a low foveal load), the perceptual span may be wider, giving rise to a larger preview benefit (or cost) compared with a low-frequency pretarget word (e.g., Henderson & Ferreira, 1990; Kliegl, Nuthmann, & Engbert, 2006). This concept has recently been implemented in a computational model of eye-movement control (Engbert & Kliegl, 2011). In agreement with this proposal, Yan, Kliegl, Shu, Pan, and Zhou (2010; see also Yan, Risse et al., 2012) provided evidence that preview benefit can be observed even on the second word after the boundary during Chinese reading, if the first word after the boundary is of high frequency.

Methodological issues

Statistical tests of the moderating influence of preview space and time as well as of pretarget and target word frequency on preview benefit/cost require an appropriate data-analytic framework. We propose that linear mixed models (LMMs) are well suited for this objective (see Yan, Risse et al., 2012, for examples). Since Yan, Risse et al. (2012) used Chinese sentences, they were unable to provide a test of the preview benefit based on the classic contrast of random-letter vs. identical previews that is used in many studies with alphabetic scripts. Thus, whether or not the preview time modulation effect demonstrated by Yan, Risse et al. (2012) is also observed during reading of alphabetic scripts needs to be established. McDonald (2006) employed the classic contrast, but used repeated-measures multiple regression analysis (rmMRA) rather than LMM for statistical inference. Moreover, due to restrictions associated with rmMRA, he specified separate models for different hypotheses rather than testing them simultaneously in a single model.

A re-analysis of McDonald's (2006) preview benefits, obtained with the classic contrast of random-letter string vs. identical preview, is well suited for an LMM re-analysis because the experiment also involved a theoretically challenging non-standard boundary contrast. As shown in Figure 2, in addition to the condition with the boundary placed after the pre-target word, McDonald included a condition with the boundary placed in the middle of the pretarget word. All two-fixation cases with one fixation before and one fixation after the boundary on the pretarget word had been crossed were analyzed. The question was whether a preview benefit is observed if the post-boundary fixation is located on the second half of the pretarget word rather than on the target word.

------- Insert Figure 2 about here -------

McDonald (2006) found no reliable evidence of a preview effect for this condition. He also tested two interactions of preview time and preview condition, using the first fixation duration (i.e., the one before the boundary change) and the second fixation duration (i.e., the one after the boundary change) on the pretarget word as covariates in subsidiary analyses of the mid-word boundary condition. Neither of the two interactions was significant. The null effect associated with this manipulation has been interpreted as evidence against the proposal that preview benefit may be obtained from word N+2 (Rayner, Juhasz, & Brown, 2007). LMMs retain statistical power better than rmMRAs in unbalanced designs—an unavoidable characteristic of eye-movement studies. Will the numeric trends in favor of a preview benefit in this experiment be returned as significant in an LMM?

METHOD

This is a re-analysis of an already published experiment (McDonald, 2006). We describe the main features of the experiment, judged to be relevant for an appreciation of the present article; for further technical details we refer to the original article.

Subjects

Sixteen young adults (12 female; median age 20 years, range: 19–37; native English speakers; normal or corrected-to-normal vision) participated in the experiment.

Apparatus

Sentences were displayed on a single line at mid-screen height on a 22-inch Iiyama VisionMaster Pro 514 monitor. Sentences were presented in black, boldface, 15-point Courier New font on a white background; each character was 12 pixels wide at a screen resolution of 1024 x 768 pixels. The refresh rate of the monitor was 120 Hz. The display change triggered by the eye crossing the invisible boundary was accomplished in 8.5 ms on average.

Eye movements were recorded with an SR EyeLink II video-based head-mounted eye tracker using monocular viewing of the right eye, sampled at 500 Hz. Accuracy of gaze position was checked before every trial, and either drift correction or recalibration performed if necessary. The viewing distance was 75 cm, with one letter subtending 0.36 degrees of visual angle. A chin rest minimized head motion and enforced a constant viewing distance.

Procedure

Each subject read 160 sentences incorporating 7-letter target nouns following a relatively neutral lead in, rendering the target non-predictable from the context. Target frequency (according to the 100-million word British National Corpus counts, Burnage & Dunlop, 1992) ranged from 0.3 per million to 5.7 per million (mean = 3.1, SD = 1.2). The pretarget word was a 9- or 10-letter adjective (mean frequency = 1.2, range: .7–1.8); this was an optimal configuration for eliciting two fixations on the pretarget according to analyses of eye-movement corpora.

Figure 2 displays the sequence of events. After the eye was detected on a fixation marker at the left side of the screen, the marker was distinguished and the sentence was displayed. As long as the eye was to the left of the invisible boundary, the target word was replaced by itself in the correct preview condition and a random, lower-case letter string in the invalid preview condition. After, the eye had crossed the boundary, the preview was replaced by the actual target word.

Each subject read 40 of the 160 target words in each of four (2 preview x 2 boundary type) conditions. Assignment of targets to the four design cells was counterbalanced across subjects. Each subject read sentences in a different random order. They answered simple yes/no questions after 40 of the sentences.

Data analysis

Average comprehension accuracy was 82%. After removing cases with very long fixation durations (cut-off 800 msec: one case rejected) or blinks on the target noun, 98% of the data remained. For inclusion in the analyses of mid-word boundary condition the pretarget adjective had to be fixated exactly twice, with the first fixations located to the left and the second fixation to the right of the boundary. Only 25% of the trials in this condition met this criterion (N in mid-word condition = 310; N in post-word condition = 1032).

Inferential statistics are based on a linear mixed model (LMM) specifying subjects and items as crossed random factors, using the *lmer* program of the *lme4* package (Bates & Maechler, 2010) in the R environment for statistical computing and graphics (R Development Core Team, 2010). Effects larger than twice their standard errors are interpreted as significant at the 5% level (i.e., given the number of subjects and especially the large number of observations for each subject, the *t*-statistic [i.e., M/SE] effectively corresponds to the *z*-statistic). Analyses of residuals and inspection of duration distributions strongly suggested that log-transformation was required to meet LMM assumptions. Therefore, we used log-transformed durations for LMMs.

For assessment of relative differences in goodness of fit, the *lmer* program provides the Akaike Information Criterion (AIC; values decrease with increasing goodness of fit), the Bayesian Information Criterion (BIC; values decrease with increasing goodness of fit), the log likelihood (logLik; values increase with goodness of fit), and, in the case of model comparisons, the $\chi 2$-distributed likelihood ratio and its associated *p*-value. The AIC (= -2 logLik + 2 n_{param}) and BIC (= -2 logLik + n_{param}log N_{obs}) values correct the log-likelihood statistic for the number of estimated parameters and, in the case of BIC, also for the number of observations, that is, we use them as a guide to prevent overfitting during the process of model selection. Comparisons between models with different fixed effects are based on maximum likelihood (ML) statistics; comparisons between models with different random effects are based on restricted maximum likelihood statistics (REML); comparisons involve only strictly nested models (Pinheiro & Bates, 2000). Estimates of final models are always based on REML statistics. For graphics we used the *ggplot2* package (Wickham, 2009).

RESULTS

McDonald's (2006) primary analyses comprised four rmMRAs, using GD and

FFD as dependent measures in both the mid-word and post-word boundary conditions. Besides the preview factor, the log frequency of the target word, quadratic landing position, and incoming saccade amplitude were included as covariates (centered around their mean) for the rmMRA within subjects. The main results indicated significant preview effects for GDs and FFDs, but only for the post-boundary condition. The analyses did not test for any interactions between the experimental design factors or between these factors and the covariates.

In the re-analysis we specified two LMMs, one with log(GD) and one with log(FFD) as the dependent variable, and included both preview and boundary conditions and their interaction as design factors for a total of 1342 measures on the target word. Most importantly, we also included the interaction of these design factors with each of the three covariates (log frequency of target word, quadratic landing position, and incoming saccade amplitude). Finally, subjects (N=16) and items (N=160) were included as crossed random factors, yielding estimates of variance components for mean GD across subjects and across items.

Modeling gaze durations

In the full model with log(GD) as the dependent variable, none of the interactions involving frequency and landing position were significant. Dropping the associated six interaction terms did not lead to a significant drop in goodness of fit [logLikelihood $\Delta\chi 2$ (6 df) = 6.4, $p = 0.38$]; also both AIC and BIC were smaller for the simple model (AIC: 555 vs 550; BIC: 654 vs 617). Table 1 shows fixed-effect estimates, standard errors and *t*-values for this reduced model; estimates of the square root of the three variance components (standard deviations) and goodness-of-fit statistics are also provided.

------- Insert Table 1 about here -------

The critical result was a significant three-factor interaction between preview (PRV), boundary condition (BND), and preview space, operationalized as incoming saccade amplitude (ISA; *t*=-2.34). The interaction is displayed in Figure 3 (i.e., partial effects based on the conditional modes for the contributing factors after removal of effects due to differences between subjects and differences between items as well as effects due to frequency and landing position).[1]

------- Insert Figure 3 about here -------

There was a large difference between the new mid-word (left panel) and the classic post-word (right panel) boundary conditions. The results agree with McDonald's (2006) conclusion that there is no evidence for a preview benefit when a refixation on the second half of the pretarget word preceded the saccade to the target word, and the display change occurs before the refixation. For the classical post-word boundary condition, however, the GD difference between incorrect and correct preview increased with preview space. Thus, the preview benefit reported in McDonald's primary analyses (2006, Table 3) is actually a combination of preview

---

[1] We also specified a model with varying subject-related effects for preview and boundary effects and parameters for the correlation between these variance components. This model converged but some of the correlations were very close to 1. Thus, we suspect that the sample size is not large enough to support such a complex model. The critical three-factor interaction shown in Figure 3 was still significant in this model.

benefit due to decreasing GD for correct and increasing GD for incorrect previews. There was some visual indication of this interaction in a supplementary analysis (see McDonald, 2006, Figure 2, left panel) in which post-word boundary and single-fixation cases from the mid-word boundary conditions were pooled.

The inference that the preview effect is due to a sum of preview benefit and cost was supported by a post-hoc re-parameterization of the fixed-effects part of the model (i.e., specifying a model with the same complexity and overall goodness of fit, but with the effect of preview specified as nested within levels of the boundary condition). For the classic post-word boundary condition, the effect of preview and the interaction of preview conditions and preview space (i.e., the divergence of slopes) were highly significant (preview: $t = 7.0$; preview x amplitude: $t = -3.2$). Additional tests indicated that as a function of decreasing incoming saccade amplitude, there was a reliable decreasing trend in gaze duration (t = 2.2) as well as in refixation rate (z = 1.9, $p = .060$) for correct preview (i.e., preview benefit); there was also a reliable increasing refixation rate (z = 2.8, $p = .006$) for incorrect preview (i.e., preview cost) although the effect in gaze duration only approached significance in a numerical trend ($t = 1.7$). Neither of these effects was significant for the mid-word boundary condition (preview: $t = 1.3$; preview x amplitude: $t = 0.9$).

Modeling first-fixation durations

The LMM for log(FFD) was developed analogously to the one for log(GD). In this analysis, none of the three-way interactions between covariates and the experimental-design factors preview and none of the two-way interactions involving landing position and frequency were significant. Dropping the associated seven interaction terms did not significantly reduce goodness of fit [log-likelihood $\Delta\chi 2$ (7 df) = 7.3, *p* = 0.40]; also both AIC and BIC were smaller for the simple model (AIC: 506 vs 499; BIC: 605 vs 561).

In Figure 4, we again display the interaction between incoming saccade amplitude and preview condition, conditional on the two boundary conditions for FFDs. Due to the comparatively small number of fixations in the mid-word boundary condition, the profiles in the two panels of the figure were not sufficiently different to reject the null hypothesis for the three-factor interaction, but three two-factor interactions were significant (preview x amplitude: *t* = 2.51; boundary x amplitude: *t* = 2.17; preview x boundary: *t* = 2.40; see right part of Table 1).

------- Insert Figure 4 about here -------

Again, with post-hoc re-parameterization of the fixed-effects part of the LMM (specifying preview as nested within levels of the boundary condition), interpretation of the pattern of results is straightforward. For the classic post-word boundary condition, the main effect of preview and the divergence of the preview conditions with increasing preview space were significant (preview: *t* = 5.1; preview x amplitude: *t* = -2.8). Note that, in contrast to GDs, FFDs for an incorrect preview in the post-word boundary condition did not increase with decreasing saccade amplitude (see Figure 4). The GD result is due to a higher refixation rate in the case of an incorrect preview. Neither of these effects was significant in the mid-word boundary condition (preview: *t* = 0.4; preview x amplitude: *t* = -0.2). [The correct-preview slope was significant (*t* = 2.9) in the post-word condition panel, but the incorrect-preview slope was not significant (*t* = 0.7); in the mid-word condition, only the amplitude slope was significant (*t* = 3.1).]

Effects of preview time

Preview effects may not only depend on preview space but also on preview time. So far we reported analyses with preview space as a covariate for dependent variables GD and FFD to stay as close as possible to McDonald's (2006) original publication. In our experience, however, if there is a sufficiently large number of them, SFDs yield the most consistent picture of effects across studies, presumably because they are less "contaminated" with corrections due to saccadic inaccuracies than multiple fixation cases. Therefore, for the remainder of this article we switch to SFD on the target as dependent variable. In the Appendix we provide results of the final model, including both preview space and preview time as covariates, for each of the three dependent variables SFD, GD, and FFD with either GD or FFD as covariate for preview time. SFD cannot be used as covariate because there was a restriction of at least two fixations on the pretarget word for the midword boundary condition. In the following, we report results for both a baseline (including only preview space as a covariate) and an extended LMM (including both preview space and preview time as covariates). The selection of trials with single fixations on target words and fixation time on the pretarget word reduced the number of observations from 1342 to 1131 (N mid-word = 278; N post-word = 853).

Baseline LMM for SFDs. The baseline LMM comprises the same terms as the

LMM reported for GDs and FFDs above. In Table 2 (left part) we report the coefficient estimates with preview specified as nested within the boundary condition. Again, as shown in the top part of Figure 5, there was an overall significant increase of fixation time with longer incoming saccades (t = -4.8), and a significant increase in the preview effect with shorter saccade amplitudes for the post-word condition (interaction: t = -3.9), but not for the mid-word condition (interaction: t = 0.3). The non-significant slope associated with SFDs was similar to the one reported for FFDs (see Figure 4).

SFDs yielded the same results as GDs with respect to the significant frequency effect (t = -2.6), but they were like FFDs as far as the negative quadratic effect of landing position is concerned (t = -4.5). Thus, both SFDs and FFDs yielded the expected inverted optimal viewing-position effect (long fixation durations when fixating the word’s center but short fixation durations when fixating its beginning or end).

------- Insert Table 2 and Figure 5 about here -------

Extended LMM for SFDs. The critical question in this analysis is whether adding a preview-time covariate provides further evidence for a modulation of the preview effect. To this end, we added centered log gaze duration on the pretarget word and its three interactions with preview and boundary conditions to the LMM. The goodness of fit improved significantly [log-likelihood $\Delta\chi 2$ (4 df) = 29.7, $p < .001$]; in addition, both AIC and BIC were smaller for the extended model (AIC: 181 vs. 159; BIC: 246 vs. 245). The right part of Table 2 displays the estimates of the re-parameterized

LMM coefficients. With one inconsequential exception (i.e., the saccade x boundary interaction), the pattern of common effects is very similar for baseline and extended LMM. The description of results provided above holds. The addition of pretarget GD as a covariate showed an overall dependence of target SFD on pretarget GD (t = 3.2). This positive relationship between pretarget gaze and target SFD is further evidence for a lag-1 autocorrelation of fixation durations during reading (McDonald, 2005). Most importantly, the analysis also yielded an interaction between pretarget GD and preview condition in the post-word boundary condition (t = 2.1), but not in the mid-word boundary condition (t = -0.8). The interaction is displayed in the bottom part of Figure 5. Focusing on the classical post-word boundary condition, SFDs on the target word are generally longer for long preview times. Most importantly, the preview effect increases with pretarget gaze duration, and this increase is driven more strongly by increasingly longer SFDs after random-letter than after identical previews (the associated slope is significantly positive for SFDs, t = 5.92, FFDs, t = 5.42, and GDs, t = 4.49). This pattern suggests that long pretarget GDs lead to preview cost rather than preview benefit in the classical post-word boundary paradigm.

DISCUSSION

An LMM re-analysis of McDonald's (2006) experiment with a classical post-word and an innovative mid-word boundary manipulation yields agreement with the central conclusion drawn in the original publication. The re-analysis, however, also demonstrates the potential associated with LMMs to unveil a much richer dynamic of parafoveal processing and associated preview effects than has been assumed and discussed in past research. This was achieved mainly by taking into account the effects of covariates of preview space and preview time that are necessary components of natural reading, which are largely beyond experimental control. Before we discuss implications of the results for the role of preview space/time during reading we address a few methodological concerns relating to the use of LMMs.

LMMs, rmMRAs, and ANOVAs

Advantages of LMM over traditional F1/F2-ANOVAs or rmMRAs are the following. They provide better statistical power in the case of a highly unbalanced design, typical of eye-movement research (e.g., Baayen, 2008; Quené & van den Bergh, 2004, 2008). They allow simultaneous estimation of crossed between-subject and between-item variance components (e.g., Baayen, Davidson, & Bates, 2008; Kliegl, Masson, & Richter, 2010), replacing separate F1- and F2-ANOVAs for subjects and items with a single analysis. They seamlessly handle mixtures of factors with discrete levels and continuous covariates. Moreover, as one is not forced to carry out analyses separately for each subject (as in rmMRA), comprehensive models containing a comparatively large number of covariates can be specified in a LMM (e.g., Kliegl, 2007). Finally, LMMs simultaneously generate estimates for experimental effects and—given sufficiently large samples—the correlations between these experimental effects (Kliegl, Wei, Dambacher, Yan, & Zhou, 2011). For example, subject-related variance/covariance parameters are informative with regard to reliable differences between subjects with respect to preview effects and their correlation with subjects' mean SFD; item-related variance parameters inform about the variance across items with respect to their sensitivity to preview effects.

These advantages of LMMs over ANOVAs justifiably increase the chances of detecting significant effects and, perhaps unsurprisingly, there is concern, often by "defenders" of null results, but also by journal editors, that LMMs are actually anti-conservative or too powerful. In our opinion, the agreement in statistical inference between ANOVAs and LMMs is much more impressive than their divergence for cognitive or psycholinguistic experiments with reaction times as dependent variables (e.g., Kliegl, 2007; Kliegl et al., 2010; Kliegl et al., 2011), but we note that there are some differences as far as boundary experiments with small effect sizes are concerned (e.g., Kliegl, Risse, & Laubrock, 2007).

Additional complications have recently been noted as far as the inclusion of random effects is concerned. In our experience, such discrepancies are almost always a consequence of fitting models that are too complex for the data at hand. For example, with the size of the current sample (N=1131), we cannot really expect to obtain reliable estimates of variances associated with experimental effects or correlation parameters of the LMM. Therefore, we stayed with a model including only variance components for subject and item means, but data and programs are available for tests of different and more complex models than we considered. Clearly, we are only at the beginning of employing this statistical procedure to the analyses of experiments.

Of course, we did not foresee in detail the results of the present analyses. This study was partly of an exploratory nature, especially with respect to the complex three-factor interactions relating to preview interference and benefits. Nevertheless, the results suggest dynamics of word recognition and oculomotor control compatible with *a priori* held beliefs. The results are not conclusive but, in our opinion, they will generate useful hypotheses for future experiments.

No significant evidence for preview benefit in the mid-word boundary condition

The re-analysis replicated the classical preview effect and yielded no evidence for a preview effect under conditions in which the target word was not the saccade target. These results are in agreement with the original publication (McDonald, 2006). Despite greater statistical power and the inclusion of several theoretically motivated covariates, the re-analyses with LMMs yield no significant evidence against the proposition that the upcoming word has to be selected as the target for the next saccade to engage in processing this word.[2]

McDonald (2006) took this result as evidence for the proposition that parafoveal processing leading to a preview benefit requires that the parafoveal word is the saccade target. Therefore, data from the mid-word condition were included in the analyses conditional on one fixation before and one fixation after the boundary on the pretarget word. The assumption is that a refixation on the pretarget word (i.e., immediately after the mid-word boundary) constitutes evidence that the pretarget word, not the target word was the saccade goal. McDonald (2006) also showed that preview benefit was obtained when saccades went from the first part of the pretarget

---

[2] We note that obviously switching from rmMRA to LMM is not *always* a guarantee of significant effects.

directly to the target word. Thus, the null effect was not due to the distance between the pretarget and target fixation locations.

Given the large number of letters (i.e., 4 or 5) after the midword boundary and the requirement that a fixation occurs in this region, McDonald's (2006) study has also been discussed as the first instance of an analogy to an N+2 boundary study in which preview of the word after the next is denied until the eye crosses the second word before the target word (Rayner et al., 2007). As with McDonald (2006), most studies have failed to demonstrate reliable evidence supporting N+2 preview benefits (Angele et al., 2008; Kliegl et al., 2007; Rayner et al., 2007), but there are exceptions for short German words in position N+1 (Risse & Kliegl, 2011) and single-character Chinese words in position N+1 (Yan et al., 2010).

There are three aspects of the mid-word boundary paradigm that may render it premature to accept the null hypothesis. First, due to the definitional constraint of observing two fixations in a specific order, the number of trials in the mid-word boundary condition was 25% lower than in the post-word boundary condition. Hence, the non-significant results in analyses with the former may be partly due to reduced statistical power.

Second, in the mid-word boundary paradigm, the critical display change is triggered by an intra-word forward saccade, whereas an N+2 boundary display change is triggered by an inter-word saccade from word N to word N+1 or word N+2. Refixations may result as a consequence of normal reading (Engbert, Nuthmann, Richter, & Kliegl, 2005), with special provisions if the first saccade missed the intended word due to oculomotor error (Engbert, et al., 2005; Nuthmann, Engbert, & Kliegl, 2005; Reichle, Warren, & McConnell, 2009). They have also been conceptualized as events preplanned before the primary saccade as a constant movement relative to the length of the fixed word and regardless of the initial fixation landing position (Vergilino-Perez & Beauvillain, 2004). With the exception of the normal-reading case, these accounts allow that an intra-word refixation saccade is somehow different from an inter-word primary saccade. Therefore, the null effect in the mid-word boundary condition may not be directly taken as evidence against preprocessing of word N+2.

Third, if there is validity to notions of a dynamical modulation of the perceptual span (Henderson & Ferreira, 1998; Kliegl et al., 2006), McDonald's (2006) mid-word boundary realizes a condition that is typical of a narrow focus of attention on the pretarget (i.e., low-frequency long adjective read in two fixations). Nevertheless, McDonald's (2006) proposal to link preview benefit and selection of the saccade target remains viable and deserves further experimental attention.

Preview effects depend on preview space

The new insight gained by the LMM re-analysis is that preview effects depend on the distance between pretarget- and target-word fixation locations in the classical preview manipulation contrasting an identical with a random-letter preview. Although, as reviewed in the introduction, such effects have been reported before, these earlier results were obtained in post-hoc analyses within the respective papers or in experiments with non-classical boundary conditions (e.g., Slattery et al., in press).

Specifying launch site as covariate in the LMM in the present study, we demonstrated a continuous preview-space effect on identical preview benefit in the context of a single analysis for the post-word boundary condition (see Figure 4, right panel, for FFD; Figure 5, top right, for SFD).

We interpreted the increasing gap between incorrect and correct previews for FFD and SFD as a function of preview space as *preview benefit* because there was not much change for the incorrect preview for this covariate. Of course, this is not the only perspective for these data. One could also argue that the mid-word boundary functions provide the appropriate baseline for the computation of cost or benefit, because there is no significant evidence for an effect of preview type in this condition (i.e., the lefts panel in Figures 3 and 4 and the top-left panel in Figure 5). In each of these panels, both functions are very similar to those estimated for correct preview in the post-word boundary condition. Thus, in a way, if we accept this null result of effect of preview type, we are forced to use the correct-preview condition as baseline for cost/benefit inferences, because we observe the decrease of GD, FFD, and SFD with preview space even in the absence of valid parafoveal information. Contrary to FFD and SFD, we observe an increase of GD for incorrect previews in the post-word boundary condition with decreasing saccade amplitude (see Figure 4). This increase was due to an increase in refixation rate. Such an increase with decreasing preview space may reflect an interference with processing of the target word due to earlier or ongoing processing of the random letter string.

The results are in line with Yan, Risse et al. (2012). The random-letter previews may not serve as an ideal baseline from which to measure preview benefit, because in part it may be due to processing cost induced by processing of random letters. We need to keep in mind that the term "preview benefit" is really a combination of benefit and cost (see Jonides & Mack, 1984, for a discussion of the intricacies associated with picking appropriate baselines in the context of priming experiments; they apply equally to the distinction between preview cost and benefit). Relative to this baseline, are we not forced to interpret the increasing gap between the two preview conditions with increasing preview space as *preview cost* due to interference from the parafoveal random-letter string rather than as *preview benefit* due to the correct preview of the later target word?

The influence of preview space is not only relevant with respect to benefit from a parafoveal preview for subsequent fixations on a target word, but also for parafoveal-on-foveal (POF) effects, that is, properties of the preview word influencing the fixation on the preboundary word. The functional role of preview space for POF effects was also featured in a controversy between Drieghe, Rayner, and Pollatsek (2008) and Kennedy (2008). The question is whether close launch site effects are nothing but evidence of mislocated fixations (i.e., fixations that were intended for the target word but due to undershoot landed on the pre-boundary word). The argument is that in the presence of such oculomotor error, attention may still be located at the saccade goal. Indeed, most lexical POF effects are restricted to fixation cases in which the eyes were close to the target word, which are likely to be mislocated fixations due to saccadic undershoot with the intended landing position on the target word (see

Rayner, 2009 for a review).

Due to a larger preview space associated with fixations close to the target word, more letters of the preview fall into the perceptual span. This enables more efficient parafoveal processing. This hypothesis is further supported by two experiments. Results from both Inhoff, Radach, Starr, and Greenberg (2000) and Yan, Sommer, and Guo (2011) suggest that fixations close to word borders may not be the main contributor to the POF effect and, in agreement with Kennedy (2008), not all POF effects can be reduced to artifacts of mislocated fixations. On the basis of our present re-analysis, we propose that the LMM framework employing continuous covariates may offer a promising new way of investigating the vexed relationship between parafoveal-on-foveal effects, preview benefit, and preview cost.

In reading experiments employing gaze-contingent display changes, participants may differ with respect to their awareness of these changes. This potentially influences fixations following the crossing of the boundary. When subjects are asked to detect a parafoveal display change in addition to reading a sentence for meaning, accuracy is much higher when the eyes land closer to the to-be-changed word and higher accuracy is associated with longer fixation times on the post-boundary word (Slattery, Angele, & Rayner, in press). There were also independent modulations of this pattern by the type of preview (nonword, word, case change). Of course, it is not clear how many of these effects are due to the dual-task character of the signal-detection task employed in these experiments, but, in general, the results are in agreement with our study in that trials with different pretarget fixation landing positions may lead to different preview benefits and the traditional way of averaging all trials may prevent us from achieving a complete understanding of the nature of parafoveal information processing. Clearly, there is need for more research on the functional role of display-change awareness in boundary experiments.

Preview effects depend on preview time

In the present study, we found that the difference between SFDs following a random-letter string and SFDs following an identical preview increased with preview time in the post-word boundary condition. Importantly, SFDs increased under both preview conditions, but more strongly in the case of random-letter previews. Therefore, this increase in “preview benefit” is actually due to an increase in “preview cost” associated with parafoveal nonwords. Figure 1 sketches various expectations about how preview time and the type of preview may interact to generate preview benefit. The scenario in the lower right panel of Figure 5 is not compatible with any of these scenarios. Does this dependency between parafoveal preview time and the size of the preview effect on a target word agree with previous research?

Inhoff, Eiter, and Radach (2005) as well as Hohenstein et al. (2010) reported evidence for an influence of the temporal availability of parafoveal previews on successive fixations on the target word. In recent studies employing the boundary paradigm (Yan, Risse et al., 2012; Yan, Zhou et al., 2011), pretarget fixation duration was included as a covariate, and analyses revealed interactions between the time readers fixated the pretarget word—conceptually identical with preview time—and the preview type. Yan, Risse, et al. (2012) used the term *preview cost* to describe the fact that GDs on target words increased numerically in an unrelated preview condition as preview duration increased; the increase in preview cost was significant for refixation rate. Both an increase in refixation rate and an increase in SFD with preview time likely reflect interference with the processing of the target word due to earlier or ongoing processing of the unrelated or nonword preview.

As McDonald (2005) pointed out in the context of a corpus analysis, if several words are processed in parallel, cumulative preview benefit could be expected to increase with the time a target word resides in the perceptual span. This argument, however, can only be applied to benefit from an identical preview condition because it assumes that integration across saccades is restricted to correct parafoveal information. In the boundary paradigm using non-identical previews, various kinds of “incorrect” parafoveal preview overlap with the subsequently displayed foveal target. It is plausible that during pre-boundary fixations we also accumulate diverging evidence from non-identical previews, which may interfere with later target word identification.

There is already evidence for even more complex scenarios. A crossover from preview benefit to preview cost has been demonstrated using non-identical related previews. In a re-analysis of Yan et al.’s (2009) Chinese reading data, Yan, Risse et al. (2012) found facilitation due to semantic preview of the target only with preview fixations shorter than 217 ms; semantically related previews acted like unrelated previews with long preview fixations. The authors propose that the accumulation of information specific to the meaning of the semantically related preview word interferes with lexical access of the target word. Thus, with a short preview time a semantically related preview word is initially as beneficial for later processing as an identical preview, but with long preview time it causes as much interference as an unrelated preview word. In summary, such results along with those from the present study, provide clear evidence that the classically observed "preview benefit" might actually consist of a complex mixture of preview benefit and preview cost.

Implications for computational models of eye-movement control during reading

Currently available computational models, such as the sequential-attention shift E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle et al., 2009) or variants of it (Engbert & Kliegl, 2001) and models built on the assumption of processing gradients, such as the SWIFT model (Engbert et al., 2005) and Glenmore (Reilly & Radach, 2006), are not ready to reproduce results at the level of detail reported here. As far as we know, previous applications of computational models to the classic boundary paradigm (e.g., Reichle, et al., 1998; Kliegl & Engbert, 2003) implicitly assumed genuine preview benefit (i.e., panel b in Figure 1). The present results, along with other research discussed above, reveal a differentiated picture of preview effects comprising preview benefit, preview cost, or a mixture of both as a function of preview space and preview time. If computational models are to meet the challenge to reproduce key results relating to the boundary paradigm, they will have to deal not only with facilitation, but also with interference of lexical access triggered by parafoveal information.

REFERENCES

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.

Baayen, R. H. (2008) *Practical data analysis for the language sciences with R*. Cambridge, MA: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390-412.

Bates, D. M., Maechler, M., & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes (R package version 0.999375-42) [Computer software]. Available at http://cran.r-project.org/web/packages/lme4

Bates, D. M., & Maechler, M. (2010). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-32. [Computer software]

Box, G.E.P. (1979). Robustness in the strategy of scientific model building. In R. L. Launer & G. N. Wilkinson (eds.), *Robustness in statistics.* New York: Academic Press

Burnage, G., & Dunlop, D. (1992). Encoding the British National Corpus. In J.M. Aarts, P. de Haan, & N. Oostdijk (Eds.), *English language corpora: Design, analysis, exploitation* (pp. 79–95). Amsterdam, the Netherlands: Rodopi.

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology, 61,* 1239-1249.

Engbert, R., & Kliegl, R. (2001). Mathematical models of eye movements in reading: a possible role for autonomous saccades. *Biological Cybernetics*, *85*, 77-87.

Engbert, R., & Kliegl, R. (2011). Parallel graded attention models of reading. In: S.P. Liversedge, I.D. Gilchrist, & S. Everling (Eds.), *Oxford Handbook of Eye Movements* (pp. 787-800). Oxford, UK: Oxford University Press.

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

Heller, D., & Müller, H. J. (1983). On the relationship between saccade size and fixation duration in reading. In R. Groner, C. Menz, D. F. Fisher, & R. A. Monty (Eds.), *Eye Movements and Psychological Processes*, Hillsdale, N.J.: Erlbaum (pp. 287-302).

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.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36,* 1150-1170.

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 979–995.

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In: A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), Reading as a perceptual process (pp. 221–246). Oxford: Elsevier.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin, 96*, 29-44.

Kennedy, A. (2008). Parafoveal-on-foveal effects are not an artifact of mislocated fixations. *Journal of Eye Movement Research,* 2(1):2, 1-10.

Kennedy, A. (2000a). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429-456.

Kennedy, A. (2000b). Attention allocation in reading: sequential or parallel. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (eds.), *Reading as a perceptual process* (pp. 193-220). Oxford: Elsevier.

Kliegl, R. (2007). Towards a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, & Reichle (2007). *Journal of Experimental Psychology: General, 138,* 530-537.

Kliegl, R., & Engbert, R. (2003). How tight is the link between lexical processing and saccade programs? *Behavioral and Brain Sciences, 26,* 491-492.

Kliegl, R., & Engbert, R. (in press). Evaluation a computational model of eye-movement control during reading. In J. H. Wolf (ed.), *Models, simulations, and the reduction of complexity*. Berlin-New York: Verlag der Akademie. DeGruyter.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18,* 655-681.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135,*13-35.

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n+2. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1250-1255. doi:10.1037/0096-1523.33.5.1250

Kliegl, R., Wei, P., Dambacher, M., Yan, M., & Zhou, X. (2011). Experimental effects and individual differences in Linear Mixed Models: Estimating the relationship between spatial, object, and attraction effects in visual attention. *Frontiers in Quantitative Psychology and Measurement.* doi: 10.3389/fpsyg.2010.00238

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578-586.

McDonald, S. A. (2005). Parafoveal preview benefit in reading is not cumulative across multiple saccades. *Vision Research, 45*, 1829-1834.

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Research, 46,* 4416-4424.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*, 2201-2217.

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59,* 413-425.

Quené, H., & Van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communications, 43,* 103-21.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-Plus.* New York: Springer.

R Development Core Team (2010). *R: A language and environment for statistical computing*. Wien: R Foundation for Statistical Computing. [Computer software]

Rayner, K. (1975). The perceptual span and peripheral cues during reading. *Cognitive Psychology, 7*, 65-81.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word n+2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception & Performance, 33,* 230-245. doi:10.1037/0096-1523.33.1.230

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125-157.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher-level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, 16, 1–20.

Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, *7*, 34–55.

Risse, S., & Kliegl, R. (2011). Investigating age differences in the perceptual span with the N+2-boundary paradigm. *Psychology and Aging, 26,* 451-460.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology*, *52A*, 1021-1046.

Slattery, T.J., Angele, B., & Rayner, K., (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance*, *37,* 1924-1938.

Vergilino-Perez, D., & Beauvillain, C. (2004). The ability of the saccadic system to change motor plans in scanning letter strings. *Psychonomic Bulletin and Review, 11*, 332-337.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, *12*, 466-471.

Wickham, H. (2009). *ggplot2. Elegant graphics for data analysis.* Dordrecht: Springer.

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word n+1 modulates preprocessing of word n+2. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1669-1676. doi: 10.1037/a0019329

Yan, M., Richter, E., Shu, H., & Kliegl, R. (2009). Chinese readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review, 16,* 561-566.

Yan, M., Risse, S., Zhou, X. & Kliegl, R. (2012). Preview fixation duration modulates identical and semantic preview benefit in Chinese reading. *Reading and Writing: An Interdisciplinary Journal*.

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sub-lexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

AUTHOR NOTE

Data and scripts for the analyses can be retrieved at the Potsdam Mind Research Repository (http://read.psych.uni-potsdam.de/pmr2/). The research was supported by Deutsche Forschungsgemeinschaft grants KL 955/15 and KL 955/18. Address for correspondence: Reinhold Kliegl, Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24/25, 14476 Potsdam, Germany, email: kliegl@uni-potsdam.de.

Table 1. LMM statistics for the log(GD) and log(FFD) measures

Variable	log(Gaze Duration)			log(First Fixation Duration)		
	Estimate	*SE*	*t-value*	*Estimate*	*SE*	*t-value*
Mean GD	5.560	0.030	189.75	5.578	0.024	233.13
LP	0.532	0.174	3.06	-1.232	0.172	-7.17
FRQ	-0.019	0.008	-2.38	-0.012	0.007	-1.73
ISA	0.014	0.005	2.70	0.028	0.005	5.63
PRV	-0.043	0.010	-4.29	-0.023	0.009	-2.43
BND	-0.042	0.010	-4.16	-0.018	0.010	-1.76
ISA:PRV	0.004	0.005	0.80	0.010	0.004	2.51
ISA:BND	0.013	0.005	2.57	0.010	0.005	2.17
PRV:BND	0.021	0.010	2.06	0.023	0.010	2.40
ISA:PRV:BND	-0.011	0.005	-2.34	-	-	-
Variance components		*SD*			*SD*	
Subjects		0.106			0.083	
Words		0.070			0.039	
Residual		0.284			0.284	
Goodness of fit						
Log Likelihood		-299			-270	
REML deviance		597			541	

Note: LP: landing position squared; FRQ: centered log frequency of target; ISA: centered incoming saccade amplitude; PRV: preview condition (incorrect/correct); BND: boundary condition (post-word/mid-word); N of observations: 1342, N of subjects: 16; N of words: 160; ":" indicates interaction between factors or covariate.

Table 2. LMM statistics for log(SFD) in the baseline and extended models

	Baseline model			Extended model		
	Estimate	*SE*	*t-value*	*Estimate*	*SE*	*t-value*
Mean GD	5.590	0.026	213.18	5.581	0.024	227.17
LP	-0.894	0.201	-4.46	-0.859	0.199	-4.32
FRQ	-0.019	0.007	-2.62	-0.017	0.007	-2.44
ISA	0.023	0.005	4.75	0.024	0.005	4.93
BND	0.054	0.019	2.81	0.082	0.025	3.31
ISA:BND	-0.025	0.009	-2.65	-0.017	0.010	-1.82
PRV | MID	0.004	0.017	0.23	0.014	0.023	0.62
PRV | PST	0.060	0.009	6.71	0.060	0.009	6.71
ISA:PRV |MID	0.003	0.008	0.33	0.004	0.008	0.48
ISA:PRV | PST	-0.019	0.005	-3.90	-0.017	0.005	-3.50
GD	-	-	-	0.095	0.030	3.20
GD:BND	-	-	-	0.023	0.058	0.40
GD:PRV | MID	-	-	-	-0.046	0.054	-0.84
GD:PRV | PST	-	-	-	0.042	0.020	2.06
Variance components		*SD*			*SD*	
Subjects		0.094			0.081	
Words		0.045			0.044	
Residual		0.252			0.250	
Goodness of fit						
Log Likelihood		-112			-107	
REML deviance		224			214	

Note: LP: landing position squared; FRQ: centered log frequency of target; ISA: centered incoming saccade amplitude; PRV: preview condition (incorrect/correct); BND: boundary condition (PST: post-word/MID: mid-word); GD: centered log(gaze duration) on pretarget; N of observations: 1131; N of subjects: 16; N of words: 160; ":" indicates interaction between factors or covariate; "|" indicates "given", e.g. GD:PRV | PST: gaze x preview interaction given a post-word boundary condition.

Figure Captions

Figure 1. Three scenarios for a modulation of preview effects by preview space/time. (1) Preview effect (benefit or cost) is invariant across preview space/time. (2) Preview benefit increases with preview space/time. (3) Preview effect is a sum of preview benefit and preview cost.

Figure 2. An example contingent-change display sequence, for the mid-word and post-word boundary conditions. The vertical bar indicates the position of the display change-triggering boundary. The first line of each condition shows the display while the eye is to the left of the boundary; the second line is displayed immediately once the eye crosses to the right of the boundary. In both examples, two fixations (indicated by ‘x’) are made on the pretarget word followed by one fixation on the target. (Source: McDonald, 2006, *Vision Research,* Elsevier).

Figure 3. Gaze duration as a function of incoming saccade amplitude by correct/incorrect preview for the mid-word (left panel) and post-word (right panel) boundary conditions after removing the effects of target-word frequency, the quadratic effect of landing position, and between-subject and between-item variance. Error bands show 95% confidence intervals.

Figure 4. First-fixation duration as a function of incoming saccade amplitude by correct/incorrect preview for the mid-word (left panel) and post-word (right panel) boundary conditions after controlling for target-word frequency, the quadratic effect of landing position, and removing between-subject and between-item variance. Note the 3-factor interaction is not significant. Error bands show 95% confidence intervals.

Figure 5. Two 3-factor interactions in the extended model of single-fixation durations (SFD; see Table 2): *Top*: SFD as a function of incoming saccade amplitude by correct/incorrect preview for the mid-word (left panel) and post-word (right panel) boundary conditions after removing the effects of target-word frequency, the quadratic effect of landing position, log(gaze duration) on pretarget word, and between-subject and between-item variance. *Bottom*: As above, but with log(gaze duration) on the pretarget word as a covariate, and removing effect of incoming saccade amplitude. Ranges of pretarget GD differ between mid-word and post-word condition because mid-word pretarget GD comprised two fixations by definition. Error bands show 95% confidence intervals.

Figure 1

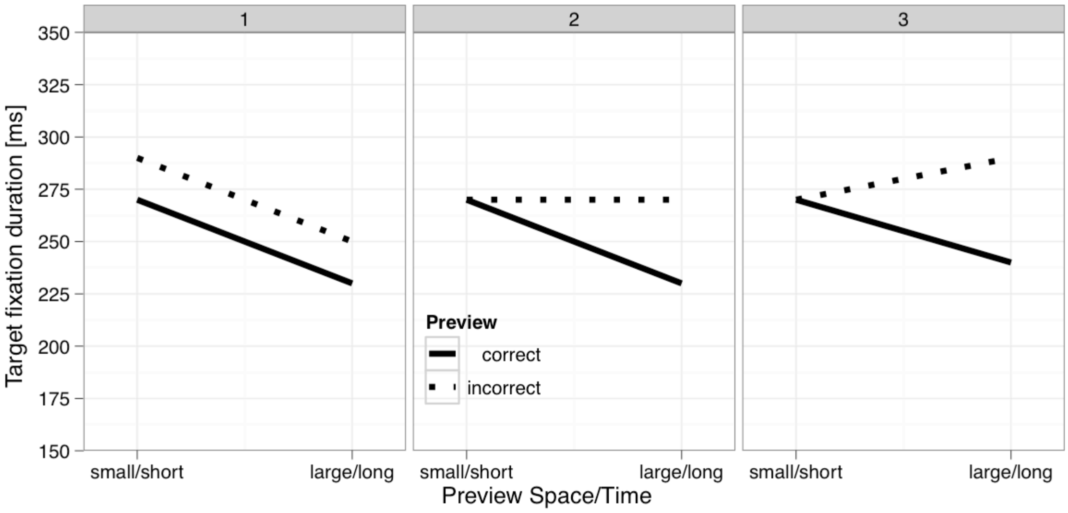
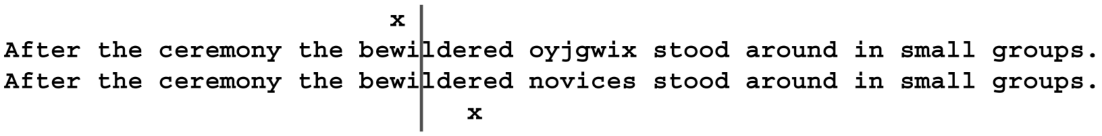


Figure 2

Mid-word boundary, incorrect preview condition:



Post-word boundary, incorrect preview condition:

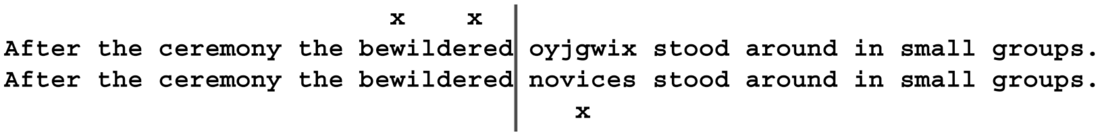


Figure 3

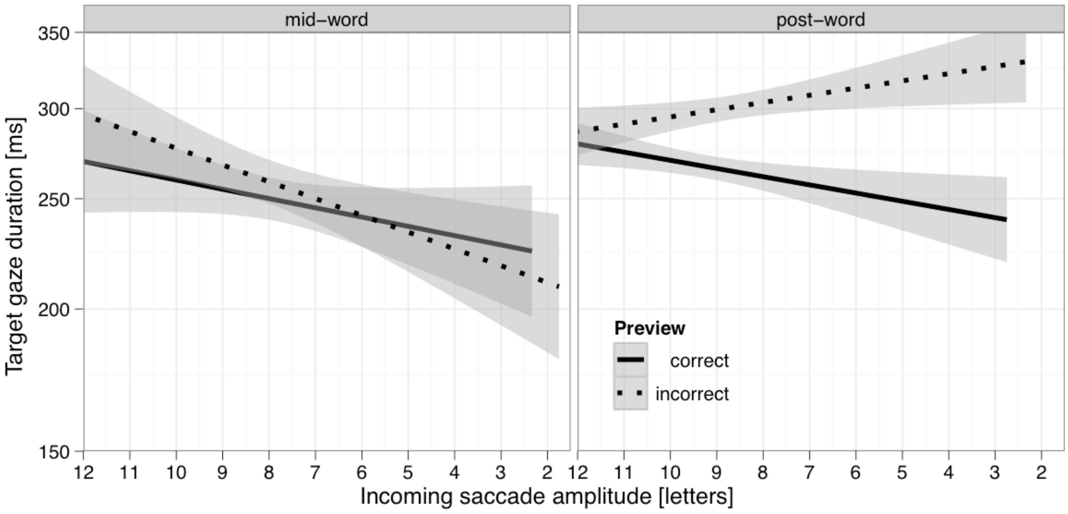


Figure 4

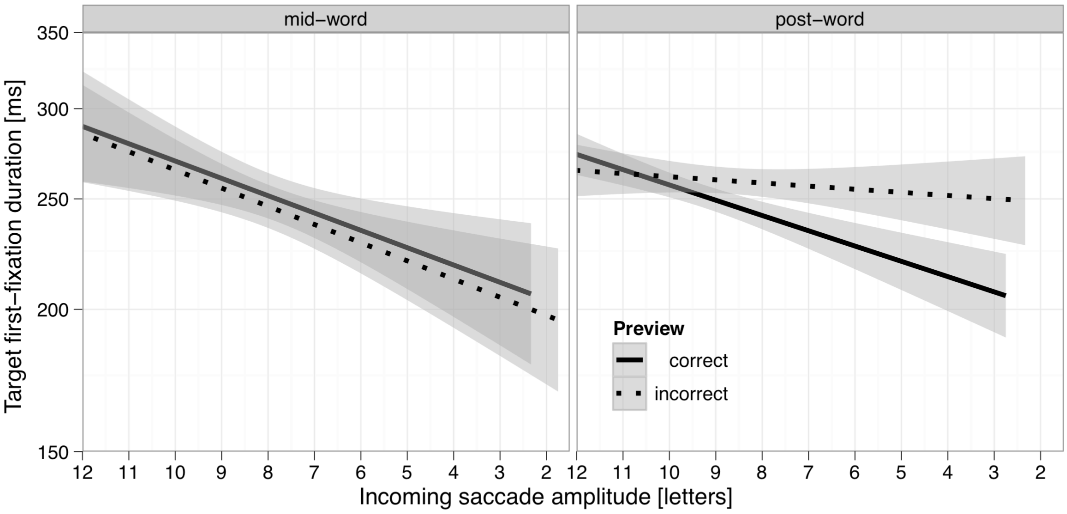
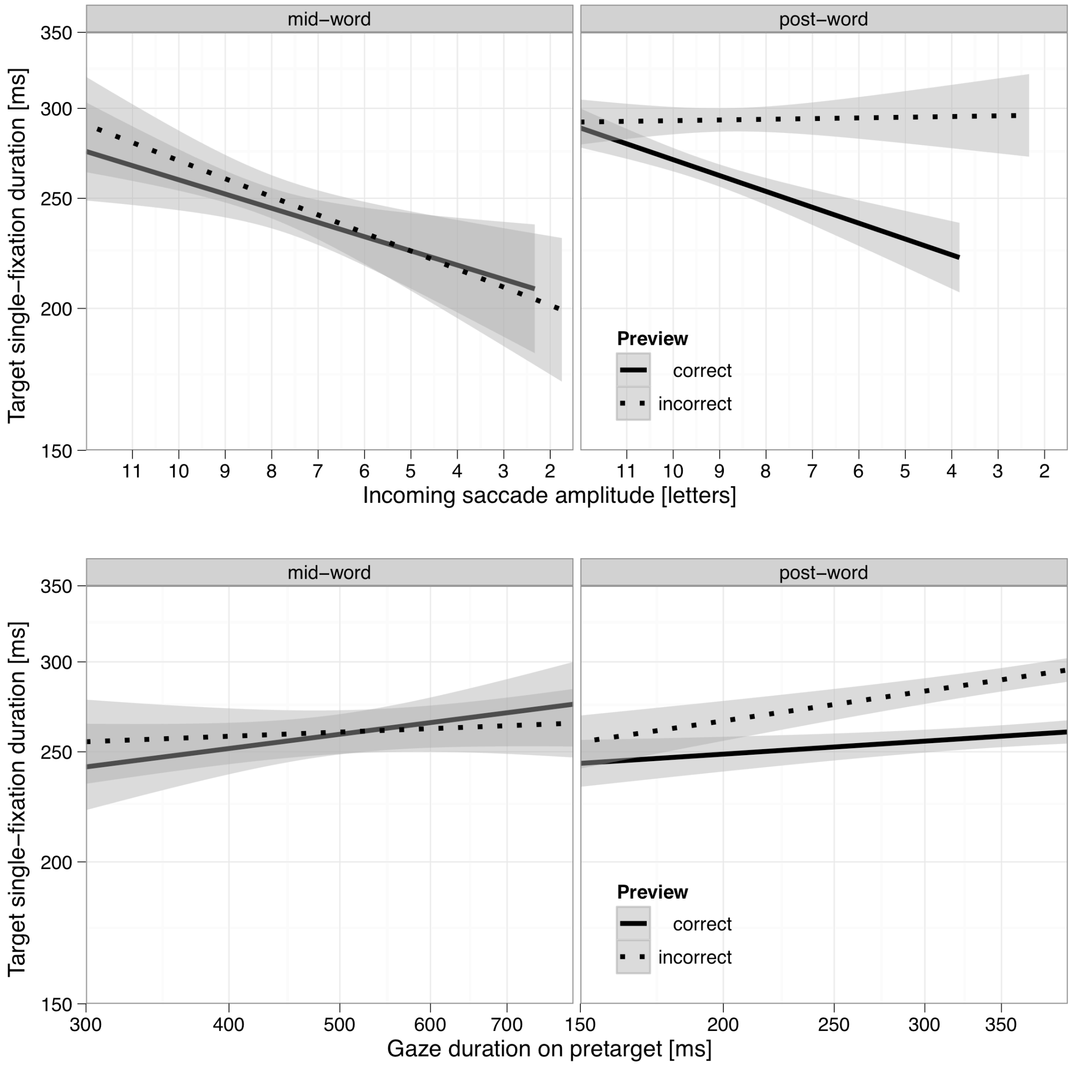


Figure 5


mid–word
post–word
Target single–fixation duration [ms]
350
300
250
200
150
11 10 9 8 7 6 5 4 3 2
Incoming saccade amplitude [letters]
Preview
correct
incorrect
mid–word
post–word
Target single–fixation duration [ms]
350
300
250
200
150
300 400 500 600 700
150 200 250 300 350
Gaze duration on pretarget [ms]
Preview
correct
incorrect

APPENDIX: Alternative Combinations of Covariates and Dependent Variables

Table A1. LMM statistics using pre-target GDs as a covariate in extended model for SFD, FFD, and GD as dependent variables

	Variable								
	target SFDs (N.obs = 1131)			target FFDs (N.obs = 1321)			target GDs (N.obs = 1321)		
	Estimate	*SE*	*t-value*	*Estimate*	*SE*	*t-value*	*Estimate*	*SE*	*t-value*
Mean GD	5.581	0.025	227.17	5.563	0.024	235.47	5.588	0.029	191.88
LP	-0.859	0.199	-4.32	-1.139	0.172	-6.62	0.602	0.177	3.41
FRQ	-0.017	0.007	-2.44	-0.013	0.007	-1.83	-0.019	0.008	-2.36
ISA	0.024	0.005	4.93	0.029	0.005	5.66	0.014	0.005	2.66
BND	0.082	0.025	3.31	0.065	0.027	2.45	0.106	0.027	3.86
ISA:BND	-0.017	0.010	-1.82	-0.010	0.010	-1.06	-0.019	0.010	-1.88
PRV | MID	0.014	0.023	0.62	0.017	0.025	0.68	0.032	0.025	1.25
PRV | PST	0.060	0.009	6.7	0.044	0.009	4.83	0.064	0.009	6.92
ISA:PRV |MID	0.004	0.008	0.48	0.000	0.009	0.05	0.009	0.009	1.04
ISA:PRV | PST	-0.017	0.005	-3.50	-0.011	0.005	-2.33	-0.014	0.005	-2.77
GD	0.095	0.030	3.20	0.096	0.031	3.06	0.064	0.032	1.96
GD:BND	0.023	0.058	0.4	0.030	0.061	0.49	-0.001	0.063	-0.02
GD:PRV | MID	-0.046	0.054	-0.84	-0.053	0.057	-0.93	-0.049	0.058	-0.83
GD:PRV | PST	0.042	0.020	2.06	0.039	0.021	1.87	0.042	0.021	1.97
Variance components		*SD*			*SD*			*SD*	
Subjects		0.081			0.074			0.098	
Words		0.044			0.032			0.066	
Residual		0.250			0.280			0.284	
Goodness of fit									
Log Likelihood		-107			-255			-295	
REML deviance		214			510			591	

Note: LP: square of landing position; FRQ: centered log frequency of target; ISA: centered incoming saccade amplitude; PRV: preview condition (incorrect/correct); BND: boundary condition (PST: post-word/MID: mid-word); GD: centered log(gaze duration) on pretarget; N of subjects: 16; N of words: 160; ":" indicates interaction between factors or covariate; “|” indicates “given”, e.g. GD:PRV | PST: gaze x preview interaction given a post-word boundary condition. Covariates were centered on the observations entering the respective analyses.

Table A2. LMM statistics using pre-target FFDs as a covariate in extended model for SFD, FFD, and GD as dependent variables

	Variable								
	target SFDs (N.obs = 1131)			target FFDs (N.obs = 1321)			target GDs (N.obs = 1321)		
	Estimate	*SE*	*t-value*	*Estimate*	*SE*	*t-value*	*Estimate*	*SE*	*t-value*
Mean GD	5.590	0.025	224.30	5.574	0.023	243.67	5.598	0.029	195.42
LP	-0.902	0.200	-4.51	-1.230	0.172	-7.16	0.544	0.176	3.09
FRQ	-0.019	0.007	-2.69	-0.014	0.007	-1.99	-0.019	0.008	-2.41
ISA	0.024	0.005	4.91	0.029	0.005	5.78	0.014	0.005	2.68
BND	0.052	0.019	2.70	0.034	0.020	1.71	0.084	0.020	4.1
ISA:BND	-0.028	0.009	-2.98	-0.022	0.010	-2.26	-0.026	0.010	-2.55
PRV | MID	0.003	0.017	0.20	0.003	0.018	0.2	0.018	0.018	1
PRV | PST	0.058	0.009	6.54	0.043	0.009	4.73	0.061	0.009	6.72
ISA:PRV |MID	0.005	0.008	0.59	0.002	0.009	0.2	0.008	0.009	0.93
ISA:PRV | PST	-0.019	0.005	-4.00	-0.013	0.005	-2.73	-0.015	0.005	-3.2
FFD	0.084	0.031	2.70	0.089	0.032	2.75	0.053	0.033	1.58
FFD:BND	-0.017	0.061	-0.28	-0.044	0.064	-0.69	0.033	0.065	0.5
FFD:PRV | MID	0.062	0.054	1.15	0.096	0.057	1.69	0.017	0.058	0.29
FFD:PRV | PST	0.056	0.029	1.96	0.055	0.029	1.91	0.030	0.029	1.01
Variance components		*SD*			*SD*			*SD*	
Subjects		0.088			0.078			0.102	
Words		0.042			0.033			0.068	
Residual		0.252			0.282			0.284	
Goodness of fit									
Log Likelihood		-114			-261			-298	
REML deviance		227			522			595	

Note: LP: square of landing position; FRQ: centered log frequency of target; ISA: centered incoming saccade amplitude; PRV: preview condition (incorrect/correct); BND: boundary condition (PST: post-word/MID: mid-word); FFD: centered log(first-fixation duration) on pretarget; N of subjects: 16; N of words: 160; ":" indicates interaction between factors or covariate; “|” indicates “given”, e.g. FFD:PRV | PST: first fixation duration x preview interaction given a post-word boundary condition. Covariates were centered on the observations entering the respective analyses.