ONLINE SUPPLEMENT

**Linear Mixed-Effects Distributed Processing Models for Reading Fixations**

Accompanying

**Toward a Perceptual-Span Theory of Distributed Processing in Reading:**

**A Reply to Rayner, Pollatsek, Drieghe, Slattery, & Reichle (2007)**

Reinhold Kliegl

University of Potsdam, Potsdam, Germany

Running Head: Lme model supplement

Address for correspondence:

Reinhold Kliegl (email: kliegl@rz.uni-potsdam.de)

Department of Psychology, University of Potsdam

Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

phone: +49 331 977 2868; fax: +49 331 977 2793

**Linear Mixed-Effects Distributed Processing Models for Reading Fixations**

The Supplement describes (1) data selection and terminology for the analyses of single-fixation and gaze duration patterns as well as relevant descriptive statistics for these two measures; (2) a description of the steps towards building the *lme* model for single-fixation durations; (3) an annotated output of the *lmer* program (Bates & Sarkar, 2006) for the fit of the final model; (4) a description of the *lme* model for gaze durations, mainly highlighting the differences to the single-fixation duration fits and including a description of the two three-factor interactions involving lexical status and skipping status of words n, n-1, and n+1 for gaze durations (i.e., analogous to Figure 2 for single-fixation durations in the Reply); (5) the *lmer* output for the final model for gaze durations.

*Data selection and terminology*

Fixations, actually saccades delineating fixations, were detected in both eyes. 71,097 fixations fixations assigned to the same word in both eyes were included in the single-fixation analysis; durations were taken from right-eye measurements. These are the data reported in the Reply. For the gaze duration analysis, I included all right-eye fixation durations, yielding 121,094 gaze durations (after removing gaze durations longer than 1 s). They comprise 103,031 single-fixation cases and 16111, 1648, 261, 39, and 4 two-fixation to six-fixation cases, respectively. For multiple-fixation cases, incoming saccade amplitude is taken from the first fixation and outgoing saccade amplitude from the last fixation; relative gaze position in word is computed as the mean of the fixation positions. Finally, gaze duration is the sum of fixation durations.

The general framework, illustrated in Figure 1 of the Reply, yields 2^5 different fixation patterns. They can be precisely specified with the following terminology: CW refers to content words and FW to function words n-1, n, and n+1; (fx) to fixated and (sk) to skipped words n-1 or words n+1; word n, the carrier of the dependent variable, is always fixated. For example, the baseline pattern of three content words fixated in sequence is coded: CW(fx)—CW—CW(fx). If word n is a function rather than a content word, the code is: CW(fx)—FW—CW(fx). If, in addition, word n+1 is a skipped function word rather than a fixated content word, the code is: CW(fx)—FW—FW(sk). Finally, variables that are ignored in contrasts that code specific interactions are marked with "*". For example: *(*)—CW—CW(sk) selects fixations on content words, irrespective of the lexical status and skipping status of word n-1, but requires that a content word n+1 is skipped during the next saccade; *(*)—*—*(*) represents the complete data base of first-pass single fixations.

Comparisons are restricted to interactions of terms relating to neighboring words, that is (1) combinations of word n-1 and word n and (2) combinations of word n+1 and word n; the former focuses lag effects and the latter successor effects. Table 4 (continuing the count from the Reply) summarizes corresponding means, standard deviations, absolute and relative number of patterns for single-fixation durations; Table 5 contains the analogous information for gaze durations. The number of fixations entering the contrasts with lexical status and skipping status of words n, n-1, and n+1 are summarized for single-fixation durations in the left part of Table 6 and for gaze durations in the right part. There is no major discrepancy between the relative frequencies of single-fixation durations and gaze durations.

We also need a terminology when referring to frequency, predictability, and length effects in the regression model. To this end, I prefix the effect with n, n-1, or n+1. For example, th n-predictability effect refers to the regression coefficient for the predictability of the fixated word, similarly n-1-predictability and n+1-predictability effects refer to the corresponding effects of word n-1 and word n+1 on the fixation duration on word n. Frequency effects are a bit more complicated because they are centered and nested within levels of lexical status to circumvent the collinearity with lexical status. Therefore, there are actually six frequency effects (see Figure 1). For example, the "n-1-CW-frequency effect", refers to the effect due to the frequency of word n-1 if word n-1 is a content word; analogously, the "n+1-FW-frequency" refers to the effect of frequency of word n+1 if word n+1 is a function word. Starting with the KNE model (see also Table 1 in Kliegl, Nuthmann, & Engbert, 2006), the following steps lead to the linear mixed-effect model referenced for statistical tests in the Reply.

*Specification of linear mixed-effects model (lme) for single-fixation duration (binocularly constrained)*

*Modification of baseline model (model 0).* Two consistently non-significant predictors are removed [i.e., length(n+1) and frequency(n) x frequency(n+1), see Table 1]. An important change is the inclusion of skipping status of word n-1 and word n+1 and of lexical status of word n-1, word n, and word n+1, as well as the specification of the frequency of the three words as centered and nested within levels of lexical status (see Figure 1). The specification of log word frequencies as nested with levels of lexical status (CW-frequency, FW-frequency) eliminates the collinearity between frequency and lexical status; interactions involving frequency are specified both for CW-frequency and FW-frequency. Also quadratic and cubic trends for frequencies of word n and predictability of word n-1, as well as a dummy code for six-letter and seven-letter words are added to the model to take care of obvious nonlinearities with these predictors (see Figure 3 in Kliegl et al., 2006). Continuous predictors are centered within readers to allow for future extension of the model with cross-level interaction terms (e.g., individual differences in skipping rate). Finally, based on residual analyses, fixation durations are log-transformed for these analyses. In line with general practice, main effects or simple interactions that are part of a higher-order interaction are kept in the model, even if they are not significant. After elimination of four unreliable frequency-related interaction terms, the model contains 32 predictors (i.e., fixed effects, including the intercept) plus two estimates for the variance between readers and for the variance of residuals (i.e., random effects). The REML deviance for this model is 8428 (AIC: 8494; BIC: 8797; logLik: -4214; ML deviance: 8108). All effects reported in the original KNE model are still significant.

*Inclusion of moderator variables (model 1).* In model 0, lexical status of word n (i.e., referred to as variable x in the model output, see below), word n-1 (variable x1), and word n+1 (variable x2) as well as skipping status of word n-1 (variable s1) and word n+1 (variable s2) were added as five dichotomous variables (see Figure 1). Content words (CW) were coded as zero and function words (FW) as one; fixated words were coded as zero and skipped words as one. This dummy coding implies that the reference model of moderators is the fixation pattern involving three successively fixated content words (i.e., called triplet-constrained patterns in KNE); the intercept codes the mean fixation duration for these fixations; i.e. CW(fx)—CW—CW(fx). In this extension, eight interactions of moderators of neighboring words (x1*s1, x1*x, s1*x, x1*x*s1; x*s2, x*x2, s2*x2, x*x2*s2) are included. The REML deviance for model 1 is 7592 (AIC: 7674; BIC: 8050; logLik: -3796; ML deviance: 7204); the improvement over model 1 is significant with $\chi(8) = 904$, $p<2.2e-16$. The two three-factor interactions are significant (x*x1*s1, b=-0.027, SE=0.010, t=-2.6; x*x2*s2, b=-0.077, SE=0.011, t=-6.8), qualifying the other six two-factor interactions. The two interactions are displayed in Figure 2a and 2b, respectively, and interpreted in the Reply. In additional analyses, the remaining 18 interaction terms among the five binary variables were checked; their inclusion failed to improve the model fit according to the Bayes-Schwartz Information Criterion (i.e., BIC increased rather than decreased); also these terms did not change the pattern of significance for the other effects.

*Inclusion of interactions of moderators and predictabilities, CW-frequencies, and FW-frequencies.* The inclusion of predictability and nested frequency effects occurs in four steps. In the first step, interactions with predictability are included and non-significant terms removed (model 2). This model represents a significant improvement in goodness of fit over model 1 [$\chi(17) = 347$, $p<2.2e-16$]. In the second step, interactions with CW-frequencies are added to model 2 and non-significant terms are removed, leading to model 2a. In model 2a, the n+2-predictability effect is no longer significant (b=0.002, SE=.001, t=1.4). Note that CW-frequency variables cannot interact with the lexical status of the word under which they are nested, but they can interact with lexical status of the other two words. Likewise, interactions with FW-frequencies are added to model 2 and, again, non-significant interactions are removed, leading to model 2b. Finally, models 2a and 2b are merged into model 3. Model 3 represents an improvement over model 2a [$\chi(12) = 167$, $p<2.2e-16$] and model 2b [$\chi(9) = 409$, $p<2.2e-16$]. Its REML deviance is 7011 (AIC: 7893; BIC: 7893; logLik: -3505; ML deviance: 6274). This model contains 78 fixed effects (incl. the intercept) and 2 random effects (i.e., variance of mean fixation durations across readers and variance of residual error). In model 3, the n+1-predictability baseline effect (b=0.002, SE=.001, t=1.7) and the 3-factor interaction x1*s1*x (b=-0.021, SE=0.012, t=-1.7) dropped below the 5%-level of significance. The intra-class correlation was .22.

*Inclusion of varying intercepts for sentences (model 4).* Both participants and sentences are random variables. In psycholinguistic research, separate analyses are often reported for subjects and items, a practice that has come under criticism because *lme* offers within a single analysis a straightforward alternative with many advantages, especially with respect to statistical power in unbalanced designs and violations of compound symmetry (Baayen, 2004, in press; Quéne & van den Bergh, 2004). In the *lmer* program, subjects and sentences can be specified as crossed random effects. Mean fixation durations of subjects and mean fixation durations of sentences are assumed uncorrelated and normally distributed.

Allowing for between-sentence variance significantly improves the goodness of fit of the model [$\chi(1) =723$, $p<2.2e-16$]. Its REML deviance is 6280 (AIC: 6440; BIC: 7174; logLik: -3140; ML deviance: 5552). Theoretically interesting fixed effects do not change.

*Inclusion of interindividual differences in selected effects (model 5).* In a final step, the three effects: relative fixation position in word, n-predictability, and n-content-word-frequency are estimated as varying across readers, that is, we test whether there are reliable interindividual differences in these effects. In *lme* one does not estimate these effects directly for each individual but one estimates their variances and covariances across individuals, assuming a normal distribution and a mean of zero (around the estimated fixed effect; see Pinheiro & Bates, 2000, for technical details). The inclusion of these three random effects improved the goodness of fit; $\chi2(3) = 583$, $p < 2.2e-16$. There was no further improvement when covariances between these effects were allowed to vary freely; $\chi2(3) = 3.4$, $p=.33$. The REML deviance for the final model is 5695 (AIC: 5861; BIC: 6622; logLik: -2847; ML deviance: 4968).

*Annotated output of lmer program for fit of single-fixation durations (final model 5).*

This section describes mainly Table 7 with the annotated *lmer* output of random-effect and fixed-effect estimates from model 5. The table lists estimates of random effects in the first section and estimates of fixed effects in the second section. Fixed effects are sorted into different fixation patterns derivable from Figure 1. The absolute and relative number of fixations contributing to each pattern is listed next to the block label. There is a short annotation for each significant term in the model; non-significant terms are usually in the model because they are part of higher-order interactions. These terms are marked as "+".

The first block of fixed effects, CW(fx)—CW—CW(fx), contains parameter estimates for the baseline model of three content words and no skipping. Therefore, only CW-frequency effects (variable g) are listed. The coefficients in this block serve as reference for corresponding coefficients of different fixation patterns, that is they are needed for the interpretation of moderated interaction terms, as illustrated next.

The second block, *(*)—FW—*(*), contains patterns with fixations on function words. The predictors listed in this block, must contain the variable "x" (i.e., word n == FW) or the variable "h" (i.e., frequency nested within function word n). In this analysis, “x” is not significant, meaning that the difference between fixation durations in the baseline pattern (3 content words, no skipping) do not differ from the fixation durations on function words. “h” is significantly negative, meaning that the frequency effect is obtained for fixated function words. The other effects in this block are interactions with x. These coefficients represent changes in slope relative to the baseline pattern. For example, the significant positive coefficient for g1:x (i.e., n-1-CW-frequency effect with fixations on a FW) implies that the lag-frequency effect of the baseline pattern (g1 = -0.0321) is significantly reduced if word n is a function rather than a content word (i.e., g1 + g1:x = -0.0321 + 0.0187 = - 0.0143).

The next six blocks inform about fixation patterns relating to word n-1 or interactions of word n-1 and word n. Specifically the third block specifies patterns with a function word in position n-1; `[FW(*)—*—*(*)]`. Then come fixation patterns (4) with skipped words n-1 `[*(sk)—*—*(*)]`, (5) with functions words both in position n-1 and position n `[FW(*)—FW—*(*)]`, (6) with word n-1 skipped and a function word n `[*(sk)—FW—*(*)]`, (7) with a skipped function word n-1 `[FW(sk)—*—*(*)]`, and (8) with a skipped function word n-1 and a function word n `[FW(sk)—FW—*(*)]`. The final six blocks inform about analogous fixation patterns relating to word n+1 or interactions of word n+1 and word n.

*Specifying the lmer model for gaze durations*

A report about distributed processing effects on gaze durations is included because many psycholinguists use this measure. Also, on the one hand, single-fixation durations are rare on long and low-frequency words; on the other hand, gaze durations are rare on short and high-frequency words (Kliegl et al., 2006; Rayner et al., 2007). In perspective, single-fixation and gaze durations must be integrated into a single, coherent, non-redundant analysis framework, capturing the rich dynamics of reading. For example, we already documented for two-fixation cases that foveal, lag, and parafoveal effects differ for intra-word forward and intra-word regressive patterns (Kliegl et al., 2006). The goal must also be to reduce the redundancy in reports of eye-movement studies. For example, in the traditional definition, first-fixation durations contain single-fixation durations and both are contained in gaze durations, which in turn are highly correlated with the number of fixations. Separate analyses for each of these measures ignore these built-in dependencies and, therefore, can be viewed as problematic to data analysts. Finally, this is only a first report. There are qualitative differences (i.e., changes in significance of predictors and in signs of regression coefficients) between the analyses of single-fixation and gaze durations. I use the various single-fixation duration models as reference for comparison but I do not "optimize" the specification for gaze-duration models. I will comment on differences with some bearing on the topics of the Reply. In general, however, I will not attempt to resolve these inconsistencies here, because, as indicated above, we are working on a new analysis framework that allows pinpointing the sources of differences. Model building followed the same sequence as described for single-fixation durations.

*GD-Model 0.* The baseline gaze-duration (gd) model 0 contains the same predictors as the single-fixation (sf) model 0. The main difference to the sf-model 0 is that neither the n+1-CW-frequency, nor the n+1-FW-frequency effects are significant for gazes. In the final model, after including contrast of lexical status and skipping status as well as interactions with predictability and CW-frequency and FW-frequency, there are fixation patterns with significant n+1-CW-frequency and n+1-FW-frequency effects (see Table 8). The n+1-predictability effect is significantly positive, as in the sf-model 0.

*GD-Model 1.* The extension with moderator variables yields two significant higher-order interactions. Figure 3a depicts the 3-factor interactions between lexical status of word n-1, lexical status of word n, and skipping status of word n-1 (b=-0.0263, SE=0.0086, t=-3.1). As for single-fixation durations, skipping "benefit" is observed only for gazes on a function word following a skipped content word n-1 and there is no skipping cost for gazes on content words after a skipped content word (Figure 3a, left panel). Gazes after skipped function words exhibit similar skipping cost for fixations on content and function words n (Figure 3a, right panel). In contrast, for singe-fixation durations, skipping cost was larger for fixations on content than on function words n (Figure 2a, right panel). Figure 3b depicts the 3-factor interaction between lexical status of word n+1, lexical status of word n, and skipping status of word n+1 (b=-0.055, SE=0.009, t=-5.5). The main difference to single-fixation durations is the absence of skipping cost for gazes on function words prior to skipped content words (see Figure 3b and 2b, left panel). This is the prime evidence for lexical preprocessing of a content word from a preceding function word, in agreement with Radach's (1996) word-group hypothesis. Apparently, evidence for this hypothesis is “lost” in gaze durations. In summary, gaze durations show the expected skipping costs for fixations of skipped words only for skipped function words n-1. There is no skipping cost, only skipping “benefit”, for gaze durations before skipped words.

*GD-Model 2.* Inclusion of interactions with predictability reverses the sign of the main effect of n+1-predictability (b=-0.0038, SE=0.001, t=-3.3). The n+1-predictability effects are highly positive, however, if word n is a function word or if word n+1 is a function word. The effect is negative if both word n and word n+1 are function words. This is exactly the pattern described for single-fixation durations in the Reply. Thus, selection effects associated with single fixations are not responsible for the n+1-predictability effects.

*GD-Model 3, GD-Model 4, and GD-Model 5.* Inclusion of interactions with CW-frequency and FW-frequency (gd-model 3), allowing varying intercepts for sentences (model 4) and varying slopes for relative fixation position, n-predictability, and n-CW-frequency leads to the final model 5. The estimates of this model are listed in Table 8. Comparison with corresponding values in Table 7 reveals similarities and differences in the overall pattern. These differences need to be sorted out. Clearly, however, gaze durations are also susceptible to influences from properties of neighboring words.

Supplement References

Baayen, H. (2004). Statistics in psycholinguistics: A critique of some current gold standards. *Mental Lexicon Working Papers.*

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

Bates, D., & Sakar, D. (2006). *lme4: Linear mixed-effect models using S4 classes.* R package version 0.995-2.

Blouin, D.C., & Riopelle, A.J. (2005). On confidence intervals for within-subject designs. *Psychological Methods, 10,* 397-412.

Harrell Jr, F.E. (2006). Hmisc: Harrell miscellaneous. R package version 3.0-12.

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**, 12-35.

Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated-measures designs: a tutorial. *Speech Communication 43,* 103-121.

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

Radach, R. (1996). *Blickbewegungen beim Lesen [Eye movements during reading].* Muenster: Waxmann.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T.J., & Reichle, E.D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General, 134,* xxx-xxx.

Sarkar, D. (2006). lattice: Lattice graphics. R package version 0.13-10.

Wickham, H. (2005). *reshape*: Flexibly reshape data. R package version 0.7.

Wickham, H. (2006). *Reshaping data with reshape package.* Manuscript.

## Acknowledgement

I thank Petra Grüttner and Nicole Stietzel for research assistance. The research was supported by grants KL 955/3 and KL 955/6 from Deutsche Forschungsgemeinschaft. The analyses reported in this Supplement were carried out in the *R* system for statistical computing (version 2.3.1; R Development Core Team, 2006) under the GNU General Public License (Version 2, June 1991). I used the *lmer* program of the *lme4* package (version 0.995-2, Bates & Sakar, 2006) for estimating fixed and random coefficients of *lme* models. Other packages I used are *reshape* (Wickham, 2005, 2006), *lattice* (Sarkar, 2006), and *Hmisc* (Harrell, 2006). Correspondence: Reinhold Kliegl, Department of Psychology, University of Potsdam, Karl-Liebknecht-Str. 24-25, Potsdam-Golm, Germany, email: kliegl@uni-potsdam.de.

Table 4. **Single-fixation duration** (mean, standard deviation, absolute and relative frequency) broken down by lexical status of word n and either lexical status and skipping status of words n-1 (top) or lexical status and skipping status of words n+1 (bottom)

| LS(n-1) | LS(n) | Sk(n-1) | M | SD | N | % |
|---|---|---|---|---|---|---|
| CW | CW | fx | 220 | 67 | 24401 | 34 |
| | | sk | 224 | 65 | 4187 | 6 |
| | FW | fx | 216 | 71 | 9970 | 14 |
| | | sk | 208 | 64 | 3454 | 5 |
| FW | CW | fx | 188 | 61 | 16476 | 23 |
| | | sk | 214 | 60 | 7019 | 10 |
| | FW | fx | 200 | 68 | 3183 | 4 |
| | | sk | 213 | 63 | 2407 | 3 |
| **LS(n+1)** | **LS(n)** | **Sk(n+1)** | **M** | **SD** | **N** | **%** |
| CW | CW | fx | 211 | 67 | 25309 | 36 |
| | | sk | 198 | 62 | 3046 | 4 |
| | FW | fx | 210 | 68 | 12263 | 17 |
| | | sk | 222 | 75 | 2234 | 3 |
| FW | CW | fx | 214 | 66 | 13501 | 19 |
| | | sk | 203 | 62 | 10227 | 14 |
| | FW | fx | 212 | 67 | 2585 | 4 |
| | | sk | 205 | 61 | 1932 | 3 |

Note. LS = lexical status, Sk = skipping status, CW = content word, FW = function word, fx =fixated, sk = skipped. N of single-fixations is 71,097; only fixations assigned to same word in both eyes; durations taken from right eye.

Table 5. **Gaze duration** (mean, standard deviation, absolute and relative frequency) broken down by lexical status of word n and either lexical status and skipping status of words n-1 (top) or lexical status and skipping status of words n+1 (bottom)

| LS(n-1) | LS(n) | Sk(n-1) | M | SD | N | % |
|---|---|---|---|---|---|---|
| CW | CW | fx | 240 | 102 | 37363 | 31 |
| | | sk | 240 | 97 | 7558 | 6 |
| | FW | fx | 225 | 86 | 17488 | 14 |
| | | sk | 218 | 82 | 6063 | 5 |
| FW | CW | fx | 227 | 119 | 28551 | 24 |
| | | sk | 243 | 104 | 13954 | 12 |
| | FW | fx | 210 | 87 | 5640 | 5 |
| | | sk | 227 | 91 | 4477 | 4 |
| **LS(n+1)** | **LS(n)** | **Sk(n+1)** | **M** | **SD** | **N** | **%** |
| CW | CW | fx | 247 | 118 | 44946 | 37 |
| | | sk | 197 | 68 | 4431 | 4 |
| | FW | fx | 224 | 90 | 21888 | 18 |
| | | sk | 216 | 76 | 4128 | 3 |
| FW | CW | fx | 244 | 109 | 24434 | 20 |
| | | sk | 200 | 64 | 13615 | 11 |
| | FW | fx | 227 | 88 | 4454 | 4 |
| | | sk | 203 | 62 | 3198 | 3 |

Note. LS = lexical status, Sk = skipping status, CW = content word, FW = function word, fx =fixated, sk = skipped. N of gazes = 121,094; all first-pass right-eye fixations.

Table 6. Frequencies (percentages) of fixation patterns for single-fixation and gaze durations

| Pattern | | single-fixation duration | | gaze duration | |
|---|---|---|---|---|---|
| | | N | % | N | % |
| 1 | CW(fx)—CW—CW(fx) | 9630 | 13 | 15171 | 13 |
| 2 | *(*)—FW—*(*) | 19014 | 27 | 33668 | 28 |
| Word n/word n-1 patterns | | | | | |
| 3 | FW(*)—*—*(*) | 29085 | 41 | 52622 | 43 |
| 4 | *(sk)—*—*(*) | 17067 | 24 | 32052 | 26 |
| 5 | FW(*)—FW—*(*) | 5590 | 8 | 10117 | 8 |
| 6 | *(sk)—FW—*(*) | 5861 | 8 | 10540 | 9 |
| 7 | FW(sk)—*—*(*) | 9426 | 13 | 18431 | 15 |
| 8 | FW(sk)—FW—*(*) | 2407 | 3 | 4477 | 4 |
| Word n/word n+1 patterns | | | | | |
| 9 | *(*)—*—FW(*) | 28245 | 40 | 45701 | 38 |
| 10 | *(*)—*—*(sk) | 17439 | 25 | 25372 | 21 |
| 11 | *(*)—FW—FW(*) | 4517 | 6 | 7652 | 6 |
| 12 | *(*)—FW—*(sk) | 4166 | 6 | 7326 | 6 |
| 13 | *(*)—*—FW(sk) | 12159 | 17 | 16813 | 14 |
| 14 | *(*)—FW—FW(sk) | 1932 | 3 | 5123 | 4 |

Table 7. Annotated output of *lmer* program for model 5, fitting **single-fixation durations**

```
Random effects:
 Groups   Name        Variance    SD
 id       g             0.0003 0.0182 # variance of n-freq effect
 id       p             0.0004 0.0188 # variance of n-pred effect
 id       o             0.0135 0.1161 # variance of rel. pos. effect
 id       (Intercept) 0.0173 0.1315 # variance of fix-dur (subjects)
 sn       (Intercept) 0.0011 0.0332 # variance of fix-dur (sentences)
 Residual             0.0611 0.2472
number of obs: 71097, groups: id, 222; id, 222; sn, 144
```

```
Fixed effects:
             Estim.     SE     t  Label
(1) CW(fx)—CW—CW(fx); N=9630 (13%)
(Intercept)  5.3515  0.0103   521 # mean log fixation duration
a1           0.0198  0.0006    35 # incoming saccade amplitude
o           -0.1004  0.0108  -9.3 # rel. fixation position
o-sq        -0.3827  0.0185   -21 # IOVP curvature
a            0.0056  0.0006   9.2 # outgoing saccade amplitude

p           -0.0260  0.0028  -9.3 # n-predictability effect
l           -0.0504  0.0193  -2.6 # n-word-length effect
l67         -0.0300  0.0027   -11 # dummy for 6/7-letter words
g           -0.0033  0.0040  -0.8 # linear n-CW-freq effect
g-sq         0.0174  0.0023   7.6 # quadr. n-CW-freq effect
g-cb        -0.0114  0.0017  -6.6 # cubic n-CW-freq effect
l:g          0.1964  0.0211   9.3 # n-freq/length interaction

g1          -0.0321  0.0022   -14 # n-1-freq effect
p1          -0.0124  0.0031  -4.0 # linear n-1-pred effect
p1-sq       -0.0088  0.0026  -3.4 # quadr. n-1-pred effect
l1           0.1185  0.0171   6.9 # n-1-length effect

g2          -0.0035  0.0016  -2.2 # n+1-CW-freq effect
p2           0.0036  0.0016   2.3 # n+1-pred effect—NOT sign!

l:p2        -0.0564  0.0124  -4.5 # n+1-prd/n-lngth interaction

(2) *(*)—FW—*(*); N=19014 (27%)
x            0.0103  0.0068   1.5 # +
h           -0.0921  0.0068   -13 # linear n-FW-freq effect
h-sq         0.0023  0.0052   0.4 # quadr. n-FW-freq effect
h-cb         0.0149  0.0040   3.7 # cubic n-FW-freq effect
p:x          0.0071  0.0035   2.0 # n-pred effect

x:g1         0.0187  0.0033   5.6 # n-1-CW-freq effect
x:p1        -0.0041  0.0046  -0.9 # +
x:p1-sq      0.0083  0.0029   2.9 # +

x:p2         0.0090  0.0030   3.0 # n+1-pred effect
h:g1         0.0105  0.0035   3.0 # n-1-CW-freq x n-FW-freq
```

Table 7 (continued)

| **Word n/word n-1 patterns** | Estim. | SE | t | Label |
|---|---|---|---|---|
| **(3) FW(*)—*—*(*); N=29085 (41%)** | | | | |
| x1 | -0.1421 | 0.0055 | -26 | # n-1-FW effect! |
| h1 | -0.0290 | 0.0037 | -7.8 | # n-1-FW-freq effect |
| x1:p1 | 0.0059 | 0.0047 | 1.3 | # linear n-1-pred effect |
| x1:p1-sq | 0.0095 | 0.0030 | 3.2 | # quadr. n-1-pred effect |
| g:x1 | -0.0343 | 0.0057 | -6.0 | # linear n-CW-freq effect |
| g-sq:x1 | 0.0253 | 0.0035 | 7.3 | # quadr. n-CW-freq effect |
| g-cb:x1 | 0.0152 | 0.0030 | 5.1 | # cubic n-CW-freq effect |
| p:x1 | -0.0052 | 0.0033 | -1.6 | # n-pred effect |
| g:h1 | 0.0156 | 0.0032 | 4.9 | # n-CW-frq x n-1-FW-frq |
| **(4) *(sk)—*—*(*); N=17067 (24%)** | | | | |
| s1 | -0.0074 | 0.0080 | -0.9 | # + |
| g1:s1 | 0.0395 | 0.0037 | 11 | # n-1-CW-freq effect(inv.) |
| p1:s1 | -0.0198 | 0.0059 | -3.4 | # linear n-1-pred effect |
| p1-sq:s1 | 0.0142 | 0.0042 | 3.4 | # quadr. n-1-pred effect |
| g:s1 | -0.0073 | 0.0061 | -1.2 | # linear n-CW-freq effect |
| g-sq:s1 | -0.0203 | 0.0040 | -5.1 | # quadr. n-CW-freq effect |
| g-cb:s1 | 0.0110 | 0.0029 | 3.9 | # cubic n-CW-freq effect |
| p:s1 | 0.0108 | 0.0026 | 4.1 | # n-pred effect (inv.) |
| **(5) FW(*)—FW—*(*); N=5590 (8%)** | | | | |
| x:x1 | 0.0725 | 0.0086 | 8.4 | # + |
| x:h1 | 0.0070 | 0.0082 | 0.9 | # + |
| h:h1 | 0.0574 | 0.0095 | 6.0 | # n-FW-freq x n-1-FW-freq |
| h:x1 | 0.0605 | 0.0082 | 7.4 | # n-FW-freq effect |
| p:x:x1 | 0.0179 | 0.0057 | 3.2 | # n-pred effect |
| **(6) *(sk)—FW—*(*); N=5861 (8%)** | | | | |
| x:s1 | -0.0425 | 0.0093 | -4.6 | # + |
| h:s1 | 0.0159 | 0.0077 | 2.1 | # + |
| x:p1:s1 | 0.0302 | 0.0077 | 4.0 | # linear n-1-pred effect |
| x:p1-sq:s1 | -0.0172 | 0.0041 | -4.2 | # quadr. n-1-pred effect |
| **(7) FW(sk)—*—*(*); N=9426 (13%)** | | | | |
| x1:s1 | 0.1200 | 0.0079 | 15 | # + |
| h1:s1 | 0.0273 | 0.0070 | 3.9 | # + |
| x1:p1:s1 | 0.0333 | 0.0080 | 4.1 | # linear n-1-pred effect |
| x1:p1-sq:s1 | -0.0183 | 0.0046 | -4.0 | # quadr. n-1-pred effect |
| **(8) FW(sk)—FW—*(*); N=2407 (3%)** | | | | |
| x:x1:s1 | -0.0128 | 0.0124 | -1.0 | # see Figure 2a |
| h:x1:s1 | -0.0832 | 0.0130 | -6.4 | # n-FW-freq effect |
| x:h1:s1 | -0.0651 | 0.0138 | -4.7 | # n-1-FW-freq effect |

Table 7 (continued)

```
Word n/word n+1 patterns
                Estim.     SE        t   Label
(9)  *(*)—*—FW(*); N=28245 (40%)
x2           -0.0305  0.0033    -9.3 # +
h2           -0.0132  0.0032    -4.1 # n+1-FW-freq effect
x2:p2         0.0081  0.0023     3.6 # n+1-pred effect

(10) *(*)—*—*(sk); N=17439 (25%)
s2           -0.0080  0.0060    -1.3 # +
g2:s2        -0.0178  0.0039    -4.5 # n+1-CW-freq effect

(11) *(*)—FW—FW(*); N=4517 (6%)
x:x2          0.0152  0.0070     2.2 # +
h:x2          0.0496  0.0096     5.1 # n-FW-freq effect
x:x2:p2      -0.0311  0.0051    -6.2 # n+1-pred effect

(12) *(*)—FW—*(sk); N=4166 (6%)
x:s2          0.0695  0.0096     7.2 # +
h:s2          0.0602  0.0118     5.1 # linear n-FW-freq effect
h-sq:s2      -0.0285  0.0117    -2.4 # quadr. n-FW-freq effect
h-cb:s2      -0.0253  0.0082    -3.1 # cubic n-FW-freq effect

(13) *(*)—*—FW(sk); N=12159 (17%)
x2:s2        -0.0109  0.0066    -1.7 # +
h2:s2         0.0152  0.0052     2.9 # n+1-FW-freq effect

(14) *(*)—FW—FW(sk); N=1932 (3%)
x:x2:s2      -0.0640  0.0120    -5.4 # see Figure 2b
h:x2:s2      -0.0500  0.0156    -3.2 # n-FW-freq effect
```

*Index for variables*. id: subject id, sn: sentence id,

x: lexical status of word n, x1: lexical status of word n-1, x2: lexical status of word n+1,

g: frequency of CW n, g1: frequency of CW n-1, g2: frequency of CW n+1,

h: frequency of FW n, h1: frequency of FW n-1, h2: frequency of FW n+1,

p: predictability of word n, p1: predictability of word n-1, p2: predictability of word n+1,

l:1/length of word n, l1: 1/length of word n-1, a1: incoming sacc. amplitude, a: outgoing sacc. amplitude

o: relative fixation position, s1: skipping status of word n-1, s2: skipping status of word n+1,

-sq: quadratic term of variable, -cb: cubic term of variable, ":" denotes interaction

Table 8. Annotated output of *lmer* program for model 5, fitting **gaze durations**

```
Random effects:
 Groups   Name        Variance     SD
 id       g            0.00034 0.01855
 id       p            0.00041 0.02013
 id       o            0.00897 0.09472
 id       (Intercept)  0.01433 0.11973
 sn       (Intercept)  0.00136 0.03682
 Residual              0.07676 0.27706
number of obs: 121094, groups: id, 222; sn, 144
```

```
Fixed effects:
              Estim.    SE       t
(1) CW(fx)—CW—CW(fx); N=15171 (13%)
(Intercept)  5.3466  0.0094    569
a1           0.0244  0.0003     75
o           -0.0946  0.0084    -11
o-sq        -0.4663  0.0151    -31
a            0.0160  0.0004     41

p           -0.0458  0.0027    -17
l            0.2896  0.0131     22
l67         -0.0229  0.0023   -9.9
g           -0.0077  0.0047   -1.6
gsq          0.0090  0.0025    3.5
gcb         -0.0074  0.0021   -3.6

g1          -0.0264  0.0020    -13
p1          -0.0269  0.0020    -14
p1sq         0.0080  0.0010    8.1
l1           0.4615  0.0110     42

g2           0.0065  0.0015    4.5
p2           0.0035  0.0013    2.8

g:g1         0.0073  0.0016    4.6
l:p2        -0.0746  0.0071    -11
(2) *(*)—FW—*(*); N=33668 (28%)
x            0.0170  0.0057    3.0
h           -0.0953  0.0065    -14
hsq         -0.0167  0.0046   -3.6
hcb          0.0155  0.0040    3.9
p:x          0.0293  0.0031    9.4

x:g1         0.0103  0.0029    3.5
x:p1         0.0112  0.0023    4.9
x:p1sq      -0.0045  0.0011   -4.0

x:p2         0.0038  0.0023    1.7
h:g1         0.0130  0.0031    4.2
```

Table 8 (continued)

| **Word n/word n-1 patterns** | | | |
|---|---|---|---|
| | Estim. | SE | t |
| **(3) FW(*)—*—*(*); N=52622 (43%)** | | | |
| x1 | -0.1371 | 0.0048 | -29 |
| h1 | -0.0169 | 0.0030 | -5.6 |
| x1:p1 | -0.0172 | 0.0025 | -6.7 |
| x1:p1sq | 0.0075 | 0.0013 | 6.0 |
| g:x1 | -0.0403 | 0.0057 | -7.1 |
| gsq:x1 | 0.0145 | 0.0035 | 4.2 |
| gcb:x1 | 0.0288 | 0.0029 | 9.9 |
| p:x1 | -0.0366 | 0.0025 | -15 |
| g:h1 | 0.0090 | 0.0030 | 3.0 |
| **(4) *(sk)—*—*(*); N=32052 (26%)** | | | |
| s1 | -0.0585 | 0.0058 | -10 |
| g1:s1 | 0.0211 | 0.0032 | 6.7 |
| p1:s1 | 0.0080 | 0.0028 | 2.9 |
| g:s1 | 0.0033 | 0.0042 | 0.8 |
| gsq:s1 | 0.0013 | 0.0034 | 0.4 |
| p:s1 | 0.0228 | 0.0023 | 10 |
| **(5) FW(*)—FW—*(*); N=10117 (8%)** | | | |
| x:x1 | 0.0514 | 0.0068 | 7.6 |
| x:h1 | -0.0080 | 0.0069 | -1.2 |
| h:x1 | 0.0668 | 0.0069 | 9.6 |
| h:h1 | 0.0518 | 0.0083 | 6.2 |
| p:x:x1 | 0.0426 | 0.0041 | 10 |
| **(6) *(sk)—FW—*(*); N=10540 (9%)** | | | |
| x:s1 | -0.0261 | 0.0070 | -3.7 |
| h:s1 | 0.0156 | 0.0069 | 2.3 |
| p:x:s1 | -0.0288 | 0.0037 | -7.8 |
| **(7) FW(sk)—*—*(*); N=18431 (15%)** | | | |
| x1:s1 | 0.1222 | 0.0069 | 18 |
| h1:s1 | 0.0196 | 0.0059 | 3.3 |
| x1:p1:s1 | 0.0226 | 0.0039 | 5.8 |
| x1:p1sq:s1 | -0.0106 | 0.0015 | -7.3 |
| g:x1:s1 | 0.0081 | 0.0061 | 1.3 |
| gsq:x1:s1 | -0.0225 | 0.0056 | -4.1 |
| **(8) FW(sk)—FW—*(*); N=4477 (4%)** | | | |
| x:x1:s1 | -0.0301 | 0.0100 | -3.0 |
| h:x1:s1 | -0.0649 | 0.0111 | -5.8 |
| x:h1:s1 | -0.0556 | 0.0115 | -4.8 |

Table 8 (continued)

| **Word n/word n+1 patterns** | Estim. | SE | t |
|---|---|---|---|
| **(9) *(*)—*—FW(*); N=45701 (38%)** | | | |
| x2 | -0.0123 | 0.0035 | -3.5 |
| h2 | -0.0045 | 0.0029 | -1.6 |
| x2:p2 | 0.0016 | 0.0019 | 0.8 |
| g:x2 | 0.0050 | 0.0051 | 1.0 |
| gsq:x2 | 0.0037 | 0.0029 | 1.3 |
| gcb:x2 | 0.0018 | 0.0024 | 0.8 |
| p:x2 | 0.0140 | 0.0020 | 7.2 |
| **(10) *(*)—*—*(sk); N=25372 (21%)** | | | |
| s2 | -0.0494 | 0.0054 | -9.1 |
| g2:s2 | -0.0133 | 0.0035 | -3.8 |
| p:s2 | 0.0203 | 0.0027 | 7.5 |
| **(11) *(*)—FW—FW(*); N=7652 (6%)** | | | |
| x:x2 | 0.0059 | 0.0071 | 0.8 |
| h:x2 | 0.0330 | 0.0122 | 2.7 |
| hsq:x2 | -0.0484 | 0.0171 | -2.8 |
| hcb:x2 | -0.0312 | 0.0095 | -3.3 |
| x:x2:p2 | -0.0105 | 0.0042 | -2.5 |
| **(12) *(*)—FW—*(sk); N=7326 (6%)** | | | |
| x:s2 | 0.0317 | 0.0073 | 4.3 |
| h:s2 | 0.0511 | 0.0077 | 6.7 |
| p:x:s2 | -0.0312 | 0.0046 | -6.7 |
| **(13) *(*)—*—FW(sk); N=16813 (14%)** | | | |
| x2:s2 | 0.0035 | 0.0060 | 0.6 |
| h2:s2 | 0.0092 | 0.0048 | 1.9 |
| **(14) *(*)—FW—FW(sk); N=5123 (4%)** | | | |
| x:x2:s2 | -0.0266 | 0.0102 | -2.6 |
| h:x2:s2 | -0.0612 | 0.0126 | -4.8 |

*Index for variables*. id: subject id, sn: sentence id,

x: lexical status of word n, x1: lexical status of word n-1, x2: lexical status of word n+1,

g: frequency of CW n, g1: frequency of CW n-1, g2: frequency of CW n+1,

h: frequency of FW n, h1: frequency of FW n-1, h2: frequency of FW n+1,

p: predictability of word n, p1: predictability of word n-1, p2: predictability of word n+1,

l:1/length of word n, l1: 1/length of word n-1, a1: incoming sacc. amplitude, a: outgoing sacc. amplitude

o: relative fixation position, s1: skipping status of word n-1, s2: skipping status of word n+1,

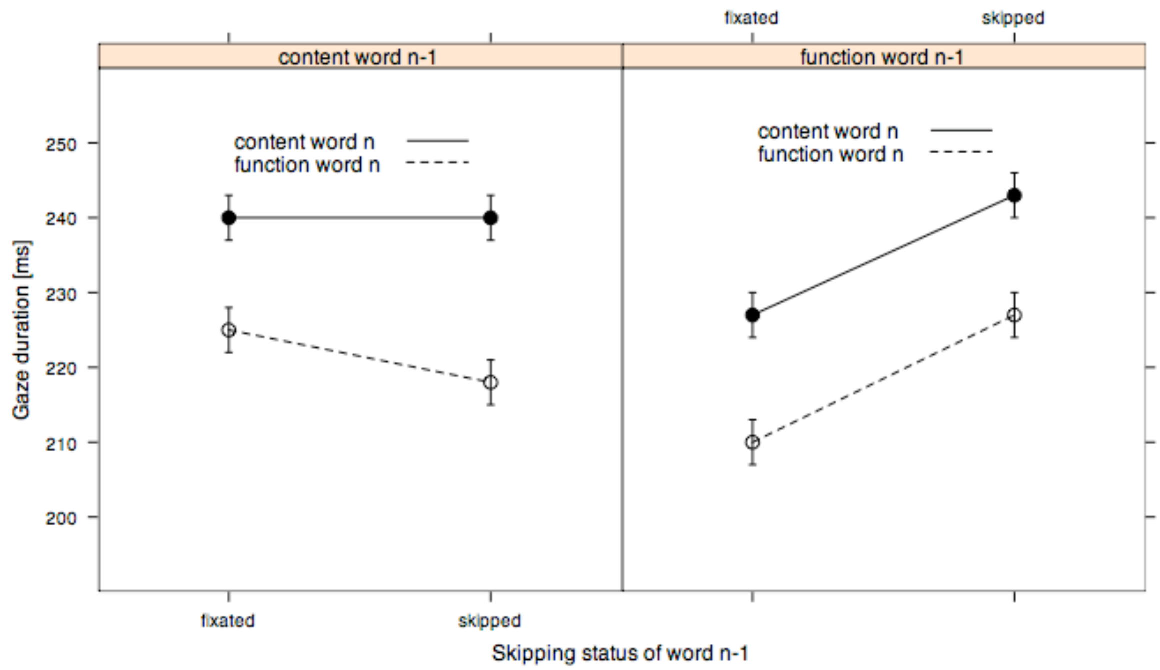
-sq: quadratic term of variable, -cb: cubic term of variable, ":" denotes interaction

Supplement Figure Caption

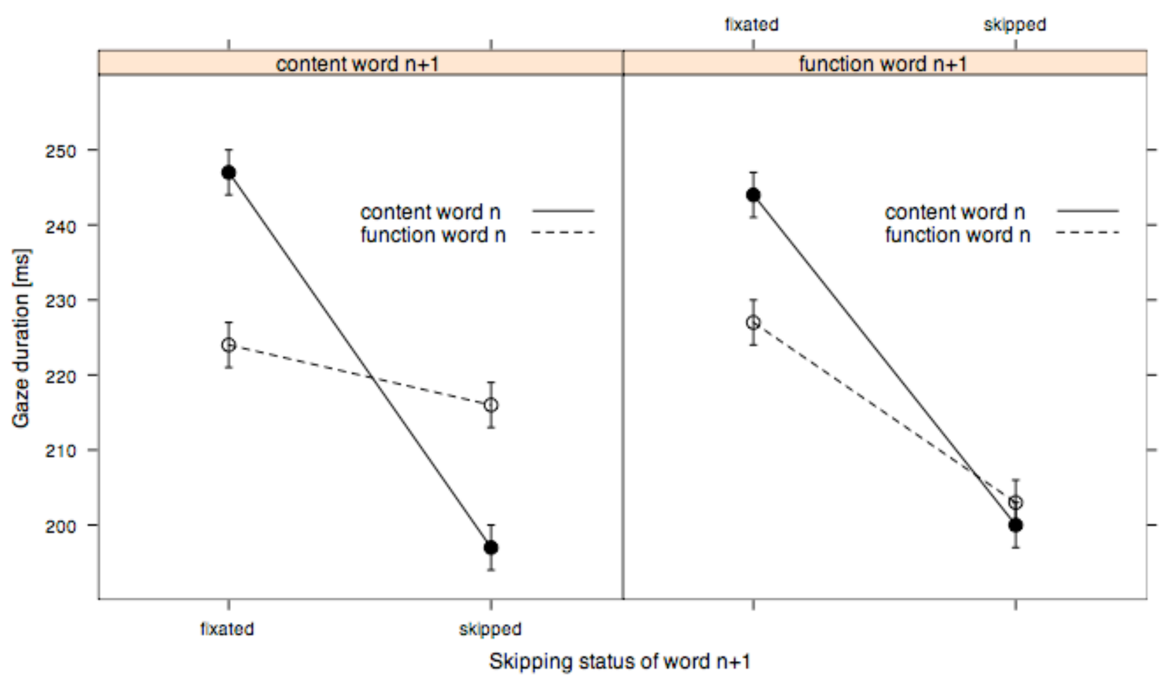
*Figure 3.* Two three-factor interactions for gaze durations. 99% confidence intervals are 3 ms, taken from *lme* model (Blouin & Riopelle, 2005). (a) Interaction of skipping status of word n-1, lexical status of word n-1, and lexical status of word n. Skipping "benefit" is observed only for gazes on a function word following a skipped content word n-1; no skipping costs for fixations on content words after a skipped content word. (b) Interaction of skipping status of word n+1, lexical status of word n+1, and lexical status of word n. There is no evidence for skipping "costs" for gazes prior to skipped words, rather there are very consistent skipping "benefits."

Figure 3

**(a)**


fixated
skipped
content word n-1
function word n-1
content word n
function word n
250
240
230
220
210
200
Gaze duration [ms]
fixated
skipped
Skipping status of word n-1

**(b)**


fixated
skipped
content word n+1
function word n+1
content word n
function word n
250
240
230
220
210
200
Gaze duration [ms]
fixated
skipped
Skipping status of word n+1