Journal of Memory and Language 88 (2016) 133–143

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Corrigendum

Corrigendum to "Do successor effects in reading reflect lexical parafoveal processing? Evidence from corpus-based and experimental eye movement data" [J. Mem. Lang. 79–80 (2015) 76–96] ☆

Bernhard Angele [a,*], Elizabeth R. Schotter [b], Timothy J. Slattery [c], Tara L. Tenenbaum [b], Klinton Bicknell [d], Keith Rayner [b]

[a] Psychology Research Centre, Faculty of Science and Technology, Bournemouth University, United Kingdom
[b] Department of Psychology, University of California San Diego, United States
[c] Department of Psychology, University of South Alabama, United States
[d] Department of Linguistics, Northwestern University, United States

ARTICLE INFO

Article history:
Received 9 December 2015
Available online 9 February 2016

Keywords:
Reading
Parafoveal-on-foveal effects
Lexical processing
Corpus analysis

ABSTRACT

In the past, most research on eye movements during reading involved a limited number of subjects reading sentences with specific experimental manipulations on target words. Such experiments usually only analyzed eye-movements measures on and around the target word. Recently, some researchers have started collecting larger data sets involving large and diverse groups of subjects reading large numbers of sentences, enabling them to consider a larger number of influences and study larger and more representative subject groups. In such corpus studies, most of the words in a sentence are analyzed. The complexity of the design of corpus studies and the many potentially uncontrolled influences in such studies pose new issues concerning the analysis methods and interpretability of the data. In particular, several corpus studies of reading have found an effect of successor word (n + 1) frequency on current word (n) fixation times, while studies employing experimental manipulations tend not to. The general interpretation of corpus studies suggests that readers obtain parafoveal lexical information from the upcoming word before they have finished identifying the current word, while the experimental manipulations shed doubt on this claim. In the present study, we combined a corpus analysis approach with an experimental manipulation (i.e., a parafoveal modification of the moving mask technique, Rayner & Bertera, 1979), so that, either (a) word $n + 1$, (b) word $n + 2$, (c) both words, or (d) neither word was masked. We found that denying preview for either or both parafoveal words increased average fixation times. Furthermore, we found successor effects similar to those reported in the corpus studies. Importantly, these successor effects were found even when the parafoveal word was masked, suggesting that apparent successor frequency effects

DOI of original article: http://dx.doi.org/10.1016/j.jml.2014.11.003

☆ Authors' note: Sadly, our co-author Keith Rayner passed away in the time between the publication of the original journal article and the discovery of the issue that made this corrigendum necessary. He was therefore not involved in the correction process. This corrigendum was written by the remaining authors.

* Corresponding author at: Bournemouth University, Psychology Research Centre, Faculty of Science and Technology, Talbot Campus, Fern Barrow, Poole BH12 5BB, United Kingdom.
E-mail address: bangele@bournemouth.ac.uk (B. Angele).

http://dx.doi.org/10.1016/j.jml.2015.12.005

may be due to causes that are unrelated to lexical parafoveal preprocessing. We discuss the implications of this finding both for parallel and serial accounts of word identification and for the interpretability of large correlational studies of word identification in reading in general.

The authors regret that an error has appeared in the article entitled "Do successor effects in reading reflect lexical parafoveal processing? Evidence from corpus-based and experimental eye movement data". This error affected the first 72 subjects we tested and resulted in the accidental exclusion of two thirds of the fixations made by those subjects. The fixations made by the remaining 56 subjects were unaffected. We also discovered an unrelated issue affecting the calculation of gaze durations: a number of refixation cases were inappropriately excluded for all subjects. This problem affected 20.4% of all gaze durations in the original analysis and was corrected in our re-analysis.

As a consequence if these issues, the means and test statistics reported in the original article are incorrect as they apply only to a subset of the data. Accordingly, all Tables (Tables 1–9), and Figs. 2 and 3 need to be updated. Additionally, Appendix A, the table of random effects in Appendix C, and Figs. A1 and A2 in Appendix A and need to be updated. Finally, a number of changes in the text of the Results section need to be made.

Number of observations

As the number of observations changed, the last two sentences in the first paragraph of the Results section need to be changed as follows:

The original text read:

"Despite the various exclusion criteria, there were 60,640 data points available for analysis of SFD and 64,210 available for analysis of GD in the corpus analysis. Assuming that most of the difference between SFD data points and GD data points is due to refixations, this corresponds to a 5.5% refixation rate. In the analyses using the experimental approach, there were 3922 data points available for analysis of SFD on the pre-target word, 6137 datapoints for analysis of SFD on the target word, 4114 data points for the analysis of GD on the pre-target word, and 6630 data points for the analysis of GD on the target word".

This must be changed to read:

"Despite the various exclusion criteria, there were 85,792 data points available for analysis of SFD and

Table 1
Properties of the experimental stimuli.

Measure	Mean	Median	*SD*	Minimum	Maximum
A: Properties of the words used in the sentence stimuli, excluding the first word in a sentence, the two last words in a sentence, and target words					
Word length	5.43	6	2.52	1	10
Word frequency	9763	67.24	17,222	0.19	42,825
Predictability (conditional trigram probability)	0.03	0	0.05	0	0.21
B: Properties of the target words					
Word length	6.5	6	1.51	5	10
Word frequency (high frequency condition)	152.47	103.77	151.49	13.77	1034
Word frequency (low frequency condition)	2.86	1.78	3.59	0.07	21.65
Predictability (high frequency condition)	<0.01	0	0.03	0	0.43
Predictability (low frequency condition)	<0.01	0	<0.01	0	0.02

Variable	Mean	*SD*
C: Properties of the dependent and the continuous independent variables in the corpus analysis for SFDs		
SFD	221.55	70.39
$n-1$ Frequency[a]	−3.32	1.44
n Frequency[a]	−3.59	1.32
$n+1$ Frequency[a]	−3.12	1.39
$n-1$ Predictability[a]	−2.70	1.61
n Predictability[a]	−3.00	1.53
$n+1$ Predictability[a]	−2.59	1.56
$n-1$ Length[b]	0.29	0.18
n Length[b]	0.24	0.13
$n+1$ Length[b]	0.31	0.19
Incoming saccade length[c]	7.78	6.15
Outgoing saccade length[c]	4.22	20.36
Fixation position[d]	0.45	0.28

[a] As log10(probability of occurrence).
[b] In characters, inverse.
[c] In characters.
[d] In proportion of word length, .5 = word center.

Table 2
Means for the $n+1$/$n+2$ mask conditions across the entire corpus. Standard deviations are in parentheses.

$n+1$	$n+2$	SFD	GD
Unmasked	Unmasked	201 (58)	214 (76.5)
Unmasked	Masked	205 (57.1)	218 (74.9)
Masked	Unmasked	233 (74.8)	253 (93.9)
Masked	Masked	250 (78.2)	270 (94.9)

99,381 available for analysis of GD in the corpus analysis. Assuming that most of the difference between SFD data points and GD data points is due to refixations, this corresponds to a 13.7% refixation rate. In the analyses using the experimental approach, there were 3820 data points available for analysis of SFD on the pre-target word, 6179 datapoints for analysis of SFD on the target word, 4266 data points for the analysis of GD on the pre-target word, and 7702 data points for the analysis of GD on the target word".

Random effects structure

In our re-analysis, we removed the random slope for word *n* frequency (i.e. the frequency of the current word) over word. This was necessary as the models as originally specified did not converge after the addition of the inappropriately excluded data. As word *n* frequency does not vary within words (each word always has the same frequency), this random slope would not be expected to capture any systematic variance, and removing it does not make a difference to the model results. It is important to note that having this slope in the model in the original analysis did not cause any problems in terms of interpretation – it merely did not capture any meaningful variance and made convergence more difficult to achieve. Using this model specification, and with the additional data, we could now fit all the models with the same random effects structure (unlike in the original analysis). Consequently, Sentences 4 through 6 in the paragraph on random effects structure in the Results section need to change as follows:

The original text read:

"Following this recommendation, we included random slopes for subjects and words only for the following effects which we deemed critically important: (1) the effect of $n+1$ frequency; (2) the effect of $n+1$ predictability; (3) the interaction between $n+1$ frequency and $n+1$ word length (necessarily also including the main effect of $n+1$ word length); (4) the interaction between $n+1$ frequency and *n* frequency (necessarily also including the main effect of *n* frequency). As we still had convergence problems even with these simplified models, we removed the correlations between the random effects. Removing random correlations reduces model complexity (and potentially, power), but does not lead to the significance tests being anticonservative. Even with this simplification, we could not fit a model for SFD that included both random slopes for the

Table 3
LMM results for single fixation duration in the corpus analysis. Only fixed effects are shown. Significant effects are represented in boldface.

Predictor	Estimate std.	Error	t value
(Intercept)	**5.44910**	**0.01180**	**461.825**
Predictors relevant to parafoveal processing			
Preview (unmasked)	**−0.17056**	**0.00313**	**−54.564**
Frequency *n*	**−0.03045**	**0.00559**	**−5.448**
1/Length *n*	0.00017	0.00591	0.029
Predictability *n*	**−0.02267**	**0.00280**	**−8.099**
Frequency *n + 1*	**0.01349**	**0.00520**	**2.593**
1/Length $n+1$	0.00664	0.00472	1.406
Predictability *n + 1*	**−0.02528**	**0.00444**	**−5.691**
Frequency $n+1$/Length $n+1$	−0.00543	0.00463	−1.174
Frequency n ∗ Frequency $n+1$	0.00333	0.00336	0.992
Frequency $n+1$/Length n	−0.00228	0.00443	−0.515
Predictability $n+1$/Length n	−0.00272	0.00379	−0.717
Preview (unmasked) ∗ Frequency *n*	**0.01449**	**0.00271**	**5.340**
Preview (unmasked) ∗ Frequency $n+1$	0.00246	0.00430	0.572
Preview (unmasked)/Length $n+1$	0.00643	0.00333	1.929
Preview (unmasked) ∗ Predictability $n+1$	0.00557	0.00372	1.496
Preview (unmasked) ∗ Frequency $n+1$/Length n	0.00070	0.00239	0.293
Preview (unmasked) ∗ Frequency n ∗ Frequency $n+1$	−0.00076	0.00235	−0.324
Preview (unmasked) ∗ Frequency $n+1$/Length $n+1$	−0.00015	0.00344	−0.042
Other predictors			
Frequency *n* − *1*	**0.01826**	**0.00308**	**5.930**
1/Length *n* − *1*	**0.01383**	**0.00274**	**5.041**
Predictability *n* − *1*	**−0.02215**	**0.00248**	**−8.930**
Incoming saccade length	**0.01768**	**0.00093**	**18.996**
Outgoing saccade length	**−0.00629**	**0.00091**	**−6.920**
Fixation position (linear trend)	**0.01343**	**0.00089**	**15.043**
Fixation position (quadratic trend)	**−0.01559**	**0.00093**	**−16.808**
Frequency $n-1$/Length $n-1$	0.00314	0.00259	1.216
Preview (unmasked)/Length *n*	**0.01470**	**0.00366**	**4.014**
Frequency n/Length n	−0.00120	0.00495	−0.242
Preview (unmasked) ∗ Frequency n/Length n	−0.00212	0.00269	−0.789

Table 4
LMM results for gaze duration in the corpus analysis. Only fixed effects are shown. Significant effects are represented in boldface.

Predictor	Estimate std.	Error	t value
(Intercept)	**5.49797**	**0.01252**	**439.233**
Predictors relevant to parafoveal processing			
Preview (unmasked)	**−0.18750**	**0.00330**	**−56.900**
Frequency *n*	**−0.03031**	**0.00589**	**−5.144**
1/Length *n*	**−0.04913**	**0.00599**	**−8.201**
Predictability *n*	**−0.02458**	**0.00298**	**−8.241**
Frequency $n+1$	0.00991	0.00530	1.872
1/Length $n+1$	0.00751	0.00484	1.551
Predictability *n* + 1	**−0.02404**	**0.00464**	**−5.179**
Frequency $n+1$/Length $n+1$	−0.00642	0.00463	−1.387
Frequency n * Frequency $n+1$	0.00512	0.00340	1.508
Frequency $n+1$/Length n	−0.00386	0.00442	−0.873
Predictability $n+1$/Length n	−0.00252	0.00399	−0.631
Preview (unmasked) * Frequency *n*	**0.01717**	**0.00287**	**5.975**
Preview (unmasked) * Frequency $n+1$	0.00154	0.00457	0.338
Preview (unmasked)/Length *n* + 1	**0.00826**	**0.00352**	**2.344**
Preview (unmasked) * Predictability $n+1$	0.00584	0.00396	1.476
Preview (unmasked) * Frequency $n+1$/Length n	0.00193	0.00257	0.753
Preview (unmasked) * Frequency n * Frequency $n+1$	−0.00129	0.00253	−0.509
Preview (unmasked) * Frequency $n+1$/Length $n+1$	−0.00346	0.00364	−0.950
Other predictors			
Frequency *n* − 1	**0.01609**	**0.00334**	**4.817**
1/Length *n* − 1	**0.01625**	**0.00298**	**5.446**
Predictability *n* − 1	**−0.02341**	**0.00270**	**−8.686**
Incoming saccade length	**0.01901**	**0.00099**	**19.134**
Outgoing saccade length	**−0.01029**	**0.00097**	**−10.634**
Fixation position (linear trend)	**0.00239**	**0.00096**	**2.484**
Fixation position (quadratic trend)	**0.01159**	**0.00098**	**11.880**
Frequency $n-1$/Length $n-1$	0.00024	0.00283	0.086
Preview (unmasked)/Length *n*	**0.01980**	**0.00394**	**5.023**
Frequency *n*/Length *n*	**0.02122**	**0.00482**	**4.406**
Preview (unmasked) * Frequency n/Length n	−0.00170	0.00277	−0.614

$n+1$ frequency by $n+1$ length and the $n+1$ frequency by n frequency interaction. In this case, we had to fit two separate models, one containing only $n+1$ frequency by $n+1$ length interaction and one containing only the $n+1$ frequency by n frequency interaction. When evaluating the effect of the $n+1$ frequency by n frequency interaction, we will report results from the latter model".

This must be changed to read:

"Following this recommendation, we included random slopes for subjects and words only for the following effects which we deemed critically important: (1) the effect of $n+1$ frequency; (2) the effect of $n+1$ predictability; (3) the interaction between $n+1$ frequency and $n+1$ word length (necessarily also including the main effect of $n+1$ word length); (4) for subjects only, the interaction between $n+1$ frequency and n frequency (necessarily also including the main effect of n frequency – as all instances of a word have the same n frequency, a random slope over words does not make sense in this case). As we still had convergence problems even with these simplified models, we removed the correlations between the random effects. Removing random correlations reduces model complexity (and potentially, power), but does not lead to the significance tests being anticonservative".

Effects of the mask manipulation

The paragraph describing the effects of the mask manipulation needs to change as follows:

The original text read:

"Just as in the analysis including only the mask manipulation, the full model showed a significant effect of the $n+1$ mask on SFD ($b = -0.1$, $SE = 0.004$, $t = -25.57$) and GD ($b = -0.1$, $SE = 0.0041$, $t = -24.92$), with fixation times being longer in the $n+1$ masked condition (mean SFD = 230 ms, mean GD = 235 ms) than in the $n+1$ unmasked condition (mean SFD = 208 ms, mean GD = 211 ms)".

Table 5
Mean SFD and GD on the pre-target word and the target word as a function of mask condition and target word frequency in the factorial analysis. Standard deviations are in parentheses.

Word n frequency	Word $n+1$ mask	Pretarget SFD	Pretarget GD	Target SFD	Target GD
High	Masked	226 (68.7)	240 (85.1)	256 (65.6)	276 (83.8)
High	Unmasked	194 (51.1)	203 (64.3)	205 (52.3)	218 (70.6)
Low	Masked	229 (72.2)	245 (86.1)	272 (73.9)	306 (101)
Low	Unmasked	191 (50.7)	202 (68.7)	223 (61.5)	247 (88.4)

Table 6
LMM results for single fixation duration on the pre-target word in the factorial analysis. Only fixed effects are shown. Significant effects are represented by boldface.

Predictor	Estimate std.	Error	t value
(Intercept)	**5.381**	**0.019**	**286.890**
***n + 1* Preview (masked vs. unmasked)**	**−0.163**	**0.012**	**−14.029**
Frequency $n + 1$ (low vs. high)	0.010	0.007	1.447
$n + 1$ Preview (masked vs. unmasked) * Frequency $n + 1$ (low vs. high)	−0.018	0.009	−1.918

Table 7
LMM results for gaze duration on the pre-target word in the factorial analysis. Only fixed effects are shown. Significant effects are represented by boldface.

Predictor	Estimate std.	Error	t value
(Intercept)	**5.410**	**0.019**	**287.500**
***n + 1* Preview (masked vs. unmasked)**	**−0.174**	**0.012**	**−14.498**
Frequency $n + 1$ (low vs. high)	0.013	0.008	1.669
$n + 1$ Preview (masked vs. unmasked) * Frequency $n + 1$ (low vs. high)	−0.022	0.011	−1.924

Table 8
LMM results for single fixation duration on the target word in the factorial analysis. Only fixed effects are shown. Significant effects are represented by boldface.

Predictor	Estimate std.	Error	t value
(Intercept)	**5.544**	**0.015**	**368.421**
***n* + 1 Preview (masked vs. unmasked)**	**−0.206**	**0.010**	**−20.386**
Frequency *n* (low vs. high)	**0.031**	**0.005**	**5.799**
***n + 1* Preview (masked vs. unmasked) * Frequency *n* (low vs. high)**	**0.014**	**0.007**	**2.058**

Table 9
LMM results for gaze duration on the target word in the factorial analysis. Only fixed effects are shown. Significant effects are represented by boldface.

Predictor	Estimate std.	Error	t value
(Intercept)	**5.617**	**0.017**	**328.545**
***n + 1* Preview (masked vs. unmasked)**	**−0.224**	**0.011**	**−20.018**
Frequency *n* (low vs. high)	**0.049**	**0.006**	**8.408**
$n + 1$ Preview (masked vs. unmasked) * Frequency n (low vs. high)	0.009	0.008	1.192

This must be changed to read:

"Just as in the analysis including only the mask manipulation, the full model showed a significant effect of the $n + 1$ mask on SFD ($b = -0.17$, $SE = 0.0031$, $t = -54.56$) and GD ($b = -0.19$, $SE = 0.0033$, $t = -56.90$), with fixation times being longer in the $n + 1$ masked condition (mean SFD = 241 ms, mean GD = 261 ms) than in the $n + 1$ unmasked condition (mean SFD = 203 ms, mean GD = 216 ms)".

Successor effects

The section describing the successor effects, which extends over several paragraphs, needs to change as follows:

The original text read:

"In total, we found three successor effects that could be considered PoF effects. First, we observed a significant effect of $n + 1$ predictability on SFD ($b = -0.019$, $SE = 0.0047$, $t = -4.02$) and GD ($b = -0.018$, $SE = 0.0047$, $t = -3.86$), with more predictable words $n + 1$ being associated with shorter fixation times on the currently fixated word. The interaction of $n + 1$ predictability and mask condition, however, did not reach significance (SFD: $b = 0.0036$, $SE = 0.0047$, $t = 0.76$; GD: $b = 0.0034$, $SE = 0.0048$, $t = 0.70$). Predictability involves more than just lexical information (syntactic and semantic properties of both the context and the word also influence the predictability of the word). Still, an $n + 1$ predictability effect like the one we found would commonly be interpreted as a high-level PoF effect. This $n + 1$ predictability effect showed a significant interaction with word n length in SFD ($b = -0.0095$, $SE = 0.0038$, $t = -2.53$), while the same interaction term was only marginally significant in GD ($b = -0.0073$, $SE = 0.0037$, $t = -1.96$). This interaction indicates that the $n + 1$ predictability effect was stronger when the currently fixated word n was long than when it was short. The direction of the $n + 1$ predictability effect (higher $n + 1$ predictability leads to shorter fixation times on n) was opposite to the direction of the effects observed by Kliegl et al. (2006), but in the same direction as the effect observed by Li et al. (2014). This may be due to the difference in language between these studies (German in Kliegl et al., Chinese in Li et al.), but it is also worth mentioning that our predictability measure (conditional trigram probability) is slightly different from the cloze predictability measures used by the other studies. As $n + 1$ conditional trigram probability was strongly correlated with $n + 1$ word frequency, it is possible that the collinearity made it difficult to distinguish predictability from frequency effects. Indeed, in a model without conditional trigram probability predictors, $n + 1$ frequency had a robust main effect in the $n + 1$ masked condition (SFD: $b = -0.011$, $SE = 0.0044$, $t = -2.40$; GD: $b = -0.011$, $SE = 0.0045$, $t = -2.56$).

When $n + 1$ predictability was included in the model, $n + 1$ frequency did not have an effect either when $n + 1$ was masked (SFD: $b = 0.0076$, $SE = 0.0056$, $t = 1.35$; GD: $b = 0.0061$, $SE = 0.0057$, $t = 1.08$) or when it was unmasked (SFD: $b = -0.0015$, $SE = 0.0056$, $t = -0.27$; GD: $b = -0.00088$, $SE = 0.0058$, $t = -0.15$). Also, there was a significant main effect of $n + 1$ length in the $n + 1$ masked condition (SFD: $b = 0.015$, $SE = 0.0047$, $t = 3.26$; GD: $b = 0.016$, $SE = 0.0048$, $t = 3.40$). This effect did not seem to be modulated by the $n + 1$ mask condition, as indicated by the non-significant interaction term between $n + 1$ length and

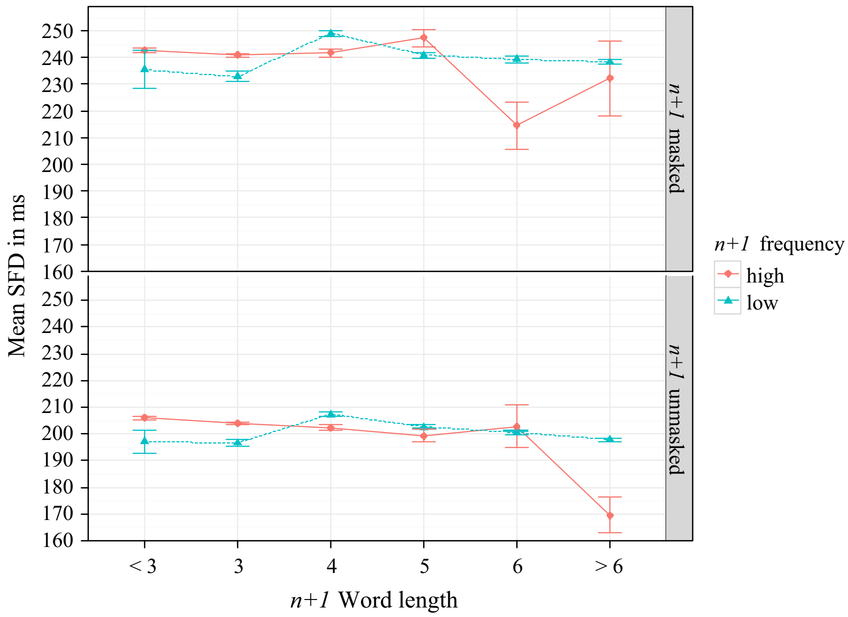

Fig. 2. $n+1$ Single fixation duration by $n+1$ word length, $n+1$ frequency, and $n+1$ mask condition (error bars show standard error over the entire data set).

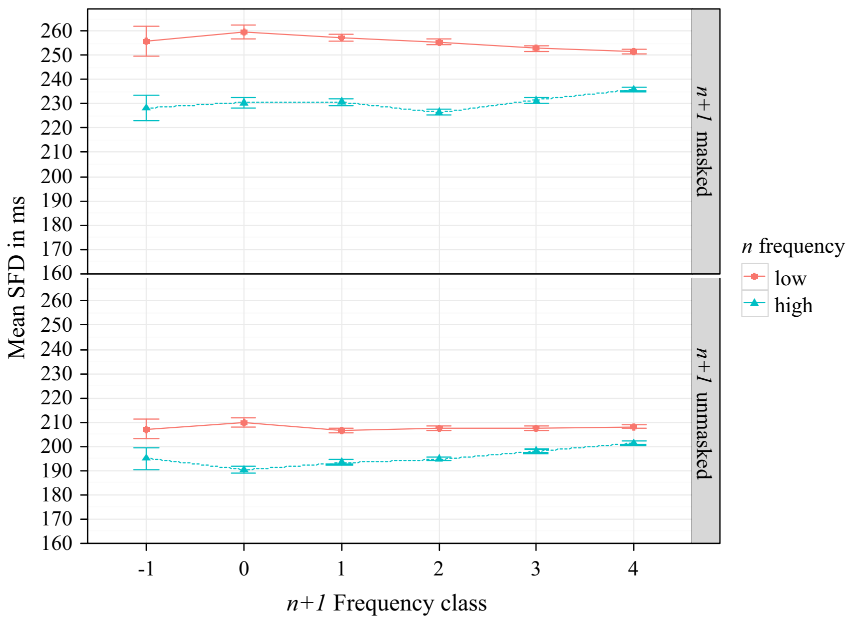

Fig. 3. Single fixation duration by n frequency, $n+1$ frequency, and $n+1$ mask condition (error bars show standard error over the entire data set). $N+1$ frequency class is determined by taking the log10 of $n+1$ frequency per million, rounding to the nearest integer.

preview (SFD: $b = 0.00059$, $SE = 0.0044$, $t = 0.14$; GD: $b = -0.0016$, $SE = 0.0045$, $t = -0.35$). This is not surprising as valid length information for words was available whether the words were masked or not. However, there was a significant interaction between $n+1$ frequency and $n+1$ length (SFD: $b = -0.017$, $SE = 0.0049$, $t = -3.42$; GD: $b = -0.017$, $SE = 0.0049$, $t = -3.41$), indicating that, in the $n+1$ masked condition, there was an $n+1$ successor frequency effect that differed for $n+1$ words of different lengths. Fig. 2 depicts this interaction on SFD (note that the GD pattern was very similar). Subjects appeared to spend more time fixating the current word when the upcoming word $n + 1$ was very short (1 or 2 letters) and of high frequency than when it was very short and of low frequency.

This effect may be caused by skipping. Kliegl and Engbert (2005) found that subjects tend to make shorter fixations before skipping short and high frequency words than before fixating them. Of course, there are not many low frequency words that are this short (there were actually only three words in the experimental sentences that fit this criterion, namely *my*, *us*, and *me*), which makes this finding quite specific. There was no frequency effect for 3-, 4-, and 5- letter $n + 1$ words. However, readers seemed to be sensitive to the frequency of $n + 1$ words with 6 or more letters, spending more time on the current word when $n + 1$ was long and of low frequency than when it was long and of high frequency. This finding replicates the effect found by Kliegl et al. (2006), although, in their study, the frequency effect was present independent of $n + 1$ word length, while we only observed the effect for long (6 letters or more) $n + 1$ words. This may be due to differences between English and German, which has more long words and fewer short words than English.

Additionally, $n + 1$ word frequency significantly interacted with the frequency of the current word n, such that the $n + 1$ frequency effect was stronger when word n was of lower frequency (SFD: $b = 0.0088$, $SE = 0.0036$, $t = 2.43$; GD: $b = 0.0086$, $SE = 0.0037$, $t = 2.31$). It is possible that whatever process drives the apparent $n + 1$ frequency effect only has a chance to influence fixation times on the current word if that word is relatively difficult to process and is therefore fixated longer.

Critically, there was no three-way interaction between the mask condition and the interaction between $n + 1$ length and $n + 1$ frequency (SFD: $b = 0.00031$, $SE = 0.0045$, $t = 0.07$; GD: $b = -5.2e{-}05$, $SE = 0.0047$, $t = -0.01$), suggesting that the $n + 1$ frequency by $n + 1$ length interaction was present both when $n + 1$ was masked and when $n + 1$ was unmasked. The same was true for the three-way interaction between the mask condition and the interaction between $n + 1$ frequency and n frequency (SFD: $b = -0.0019$, $SE = 0.0029$, $t = -0.64$; GD: $b = -0.0025$, $SE = 0.003$, $t = -0.83$), suggesting that the $n + 1$ frequency by n frequency interaction was also present both when $n + 1$ was masked and when $n + 1$ was unmasked. At first glance, the three effects we found would seem to replicate previous findings of successor frequency effects. However, these apparent successor effects were present both when $n + 1$ was unmasked and when $n + 1$ was masked (see Figs. 2 and 3). This is quite surprising as it suggests that the $n + 1$ frequency successor effect we observed on short words is caused by a process that is not dependent on lexical parafoveal processing".

This entire section must be changed to read:

"We found a number of successor effects that could be considered PoF effects. First, we observed a significant effect of $n + 1$ predictability on SFD ($b = -0.025$, $SE = 0.0044$, $t = -5.69$) and GD ($b = -0.024$, $SE = 0.0046$, $t = -5.18$), with more predictable words $n + 1$ being associated with shorter fixation times on the currently fixated word. The interaction of $n + 1$ predictability and mask condition, however, did not reach significance (SFD: $t = 1.50$; GD $t = 1.48$). Predictability involves more than just lexical information (syntactic and semantic properties of both the context and the word also influence the predictability of the word). Still, an $n + 1$ predictability effect like the one we found would commonly be interpreted as a high-level PoF effect. There was no significant interaction between this $n + 1$ predictability effect and word n length in SFD ($t = -0.72$) or GD ($t = -0.63$). The direction of the $n + 1$ predictability effect (higher $n + 1$ predictability leads to shorter fixation times on n) was opposite to the direction of the effects observed by Kliegl et al. (2006), but in the same direction as the effect observed by Li et al. (2014). This may be due to the difference in language between these studies (German in Kliegl et al., Chinese in Li et al.), but it is also worth mentioning that our predictability measure (conditional trigram probability) is slightly different from the cloze predictability measures used by the other studies.

In SFD, $n + 1$ frequency had a significant effect when $n + 1$ was masked ($b = 0.013$, $SE = 0.0052$, $t = 2.59$). Critically, there was no interaction between the mask condition and the $n + 1$ frequency effect (SFD: $t = 0.57$, suggesting that the $n + 1$ frequency effect was present both when $n + 1$ was masked and when $n + 1$ was unmasked. In GD, there was a marginal $n + 1$ frequency effect in the same direction ($t = 1.87$), again with no indication of an interaction between the mask condition and the $n + 1$ frequency effect ($t = 0.34$). The model for GD may lack power due to multicollinearity issues, as separate models fitted for the $n + 1$ masked condition and the $n + 1$ unmasked condition both show a significant effect of $n + 1$ frequency on GD ($n + 1$ masked: $b = 0.011$, $SE = 0.005$, $t = 2.20$; $n + 1$ unmasked: $b = 0.013$, $SE = 0.0056$, $t = 2.33$) even though the full model doesn't. In the full model, there also was a significant main effect of $n + 1$ length in the $n + 1$ masked condition (SFD: $t = 1.41$; GD: $t = 1.55$). This effect was modulated by the $n + 1$ mask condition (SFD (marginal): $t = 1.93$; GD: $b = 0.0083$, $SE = 0.0035$, $t = 2.34$). This is not surprising as valid length information for words was available whether the words were masked or not. There was no significant interaction between $n + 1$ frequency and $n + 1$ length (SFD: $t = -1.17$; GD: $t = -1.39$; see Fig. 2). There also was no significant interaction between $n + 1$ word frequency with the frequency of the current word n, although there was a trend in GD (SFD: $t = 0.99$; GD: $t = 1.51$, see Fig. 3).

At first glance, the effects we found would seem to replicate previous findings of successor frequency effects. However, these apparent successor effects were present both when $n + 1$ was unmasked and when $n + 1$ was masked (see Figs. 2 and 3). This is quite surprising as it suggests that the $n + 1$ frequency successor effect we observed on short words is caused by a process that is not dependent on lexical parafoveal processing".

Summary of the changes

The most important difference between the new analysis and our original analysis is that we now find a main of $n + 1$ frequency effect on SFD. On GD, this effect is marginal, but reaches significance in both preview conditions if we split the data by $n+1$ preview and analyze the unmasked and masked conditions separately. This suggests that, in the full analysis, we lose power due to multicollinearity between the $n+1$ frequency main effect and the $n+1$ frequency by preview interaction.

Overall, we find that the apparent lexical parafoveal-on-foveal effect is no longer modulated by word $n+1$ length and by word n frequency. Instead, it is present across all $n+1$ word lengths and word n frequencies. The interactions we found in the original analysis were likely the result of fixations on longer words being more likely and therefore also more likely to be inappropriately excluded, thereby changing the fixation record for long words more than the record for short words.

Effects on the pre-target word in the factorial analysis

The section about the effects on the pre-target word in the factorial analysis (with the exception of the last two sentences in that section) needs to change as follows:

The original text read:

"As observed in the corpus analysis, there was a significant effect of the $n+1$ mask on SFD ($b = -0.16$, $SE = 0.0082$, $t = -19.46$) and GD ($b = -0.16$, $SE = 0.0085$, $t = -19.31$), with SFD and GD on the pre-target word being shorter when $n+1$ was unmasked than when it was masked (SFD: $n+1$ masked: 227 ms, $n+1$ unmasked: 193 ms; GZD: $n+1$ masked: 229 ms, $n+1$ unmasked: 194 ms). Of course this effect can also be viewed as a standard preview benefit effect due to the fact that the pre-target words were almost always masked on an earlier fixation in the masked condition. That is, these pre-target words were at one point a word ($n+1$) for a fixation on the pre-pre-target word.

In the $n+1$ masked condition, there was no significant effect of the target word (i.e., $n+1$) frequency manipulation on either SFD ($b = -0.0076$, $SE = 0.0057$, $t = -1.34$) or GD ($b = -0.0096$, $SE = 0.0058$, $t = -1.65$) on the pre-target word. The lack of interactions between target frequency and the mask conditions (all $|t| < 1.5$) means that we did not find an $n+1$ frequency effect in the $n+1$ unmasked condition either".

This must be changed to read:

"As observed in the corpus analysis, there was a significant effect of the $n+1$ mask on SFD ($b = -0.16$, $SE = 0.012$, $t = -14.03$) and GD ($b = -0.17$, $SE = 0.012$, $t = -14.50$), with SFD and GD on the pre-target word being shorter when $n+1$ was unmasked than when it was masked (SFD: $n+1$ masked: 228 ms, $n+1$ unmasked: 193 ms; GZD: $n+1$ masked: 243 ms, $n+1$ unmasked: 202 ms). Of course this effect can also be viewed as a standard preview benefit effect due to the fact that the pre-target words were almost always masked on an earlier fixation in the masked condition. That is, these pre-target words were at one point a word ($n+1$) for a fixation on the pre-pre-target word.

In the $n+1$ masked condition, there was no significant effect of the target word (i.e. $n+1$) frequency manipulation on either SFD ($t = 1.45$) or GD ($t = 1.67$) on the pre-target word.

There was a marginally significant interaction between target frequency and the mask condition in SFD ($t = -1.92$) and GD ($t = -1.92$). However, the means show that neither mask condition shows a substantial $n+1$ frequency effect ($n+1$ masked: SFD, high frequency: 226, low frequency: 229; GD, high frequency: 240 ms, low frequency: 245 ms; $n+1$ unmasked: SFD, high frequency: 194 ms, low frequency: 191 ms; GD, high frequency: 203 ms, low frequency: 202 ms)".

Effects on the target word in the factorial analysis

The paragraph describing the effects of our manipulation on the target word in the factorial analysis needs to change as follows:

The original text read:

"We found the same effects of the $n+1$ mask (SFD: $b = -0.2$, $SE = 0.0063$, $t = -31.88$; GD: $b = -0.2$, $SE = 0.0066$, $t = -30.62$) as in the pre-target analysis (SFD: $n+1$ masked: 262 ms, $n+1$ unmasked: 214 ms; GZD: $n+1$ masked: 265 ms, $n+1$ unmasked: 194 ms). More importantly, there was a clear effect of target word frequency on SFD ($b = -0.032$, $SE = 0.0056$, $t = -5.82$) and GD ($b = -0.032$, $SE = 0.0058$, $t = -5.58$). Specifically, SFDs and GDs were shorter in the high frequency condition (SFD: 227 ms, GD: 230 ms) than in the low frequency condition (SFD: 246 ms, GD: 248 ms). The frequency effect was modulated by the $n+1$ mask condition, as the significant interaction shows (SFD: $b = -0.015$, $SE = 0.0063$, $t = -2.43$; GD, marginal: $b = -0.013$, $SE = 0.0066$, $t = -1.95$): when $n+1$ was masked, the frequency effect was slightly smaller (SFD, high frequency: 254, low frequency: 271; GD, high frequency: 256, low frequency: 274) than when $n+1$ was unmasked (SFD, high frequency: 205, low frequency: 225; GD, high frequency: 206, low frequency: 227)".

This must be changed to read:

"We found the same effects of the $n+1$ mask (SFD: $b = -0.21$, $SE = 0.01$, $t = -20.39$; GD: $b = -0.22$, $SE = 0.011$, $t = -20.02$) as in the pre-target analysis (SFD: $n+1$ masked: 264 ms, $n+1$ unmasked: 214 ms; GZD: $n+1$ masked: 291 ms, $n+1$ unmasked: 202 ms). More importantly, there was a clear effect of target word frequency on SFD ($b = 0.031$, $SE = 0.0053$, $t = 5.80$) and GD ($b = 0.049$, $SE = 0.0059$, $t = 8.41$). Specifically, SFDs and GDs were shorter in the high frequency condition (SFD: 229 ms, GD: 246 ms) than in the low frequency condition (SFD: 246 ms, GD: 276 ms). In SFD, the frequency effect was modulated by the $n+1$ mask condition, as the significant interaction shows (SFD: $b = 0.014$, $SE = 0.0066$, $t = 2.06$; not significant in GD: $t = 1.19$): when $n+1$ was masked, the frequency effect was slightly smaller (high frequency: 256, low frequency: 272 than when $n+1$ was unmasked (high frequency: 205 ms, low frequency: 223 ms)".

Summary of the results section

In the summary of the results section, the sentence describing the interaction of the $n+1$ frequency effect with word length needs to be deleted:

"Also, as described above, if there was an interaction between the $n + 1$ frequency effect and another word property (as our corpus analysis suggests) it is possible that the target words were selected in such a way that they happened to be associated with a very weak $n + 1$ frequency effect (e.g. by being medium-length words rather than covering the entire word length spectrum)".

Appendix A

The appendix needs to change as follows. The original text read:

"In this Appendix, all effects observed in the correlational approach LMM analyses that were not directly relevant to parafoveal preprocessing and therefore were not reported in the main text are reported.

Effects of fixation position. We observed very strong effects of fixation position on SFD and GD. The linear trend was positive and significant in both SFD ($b = 0.012$, $SE = 0.0011$, $t = 10.70$) and GD ($b = 0.013$, $SE = 0.0012$, $t = 11.23$). There also was a significant quadratic trend in SFD ($b = -0.019$, $SE = 0.0012$, $t = -16.25$) and GD ($b = -0.017$, $SE = 0.0012$, $t = -13.84$). The length of the incoming saccade showed a significant effect on both SFD ($b = 0.012$, $SE = 0.0012$, $t = 10.06$) and GD ($b = 0.011$, $SE = 0.0012$, $t = 9.43$), with longer fixation times when the incoming saccade was large, while there was no such effect for the length of the outgoing saccade (all $t < 1.96$).

Lag effects: effects of lexical properties of the word preceding the currently fixated word (word n − 1). We observed significant lag effects of frequency in both SFD ($b = 0.027$, $SE = 0.0036$, $t = 7.38$) and GD ($b = 0.026$, $SE = 0.0037$, $t = 6.86$), with words receiving longer fixations when they had been preceded by low-frequency words than when they had been preceded by high frequency words. This effect was modulated by the length of the preceding word, such that words preceded by short low-frequency words $n - 1$ were fixated longer than words preceded by long low-frequency words $n - 1$ (SFD: $b = 0.009$, $SE = 0.003$, $t = 3.00$; GD $b = 0.0066$, $SE = 0.0031$, $t = 2.14$. Finally, we observed a lag effect of predictability, with words showing lower SFD ($b = -0.028$, $SE = 0.003$, $t = -9.47$) and GD ($b = -0.028$, $SE = 0.0031$, $t = -9.01$) when the preceding word had been highly predictable from the context than when it had not been predictable.

Immediacy effects: effects of the properties of the currently fixated word (word n). There was a significant frequency effect on SFD ($b = -0.012$, $SE = 0.0055$, $t = -2.11$) but not GD ($b = -0.0048$, $SE = 0.0061$, $t = -0.78$), indicating that words received shorter fixations when they were of high frequency.

There also was a significant effect of predictability, with words showing shorter SFD ($b = -0.024$, $SE = 0.0032$, $t = -7.51$) and GD ($b = -0.026$, $SE = 0.0033$, $t = -7.95$) when a word was predictable from the sentence context than when it was not.

When $n + 1$ was masked, there was a significant main effect of word length on GD ($b = -0.012$, $SE = 0.0059$, $t = -2.05$), but not SFD ($b = -0.009$, $SE = 0.0059$, $t = -1.52$), and there was no significant interaction between word frequency and length (all $t < 1.96$). However, there was a significant interaction between mask condition and word length, indicating that word length did have an effect when $n + 1$ was unmasked (SFD: $b = 0.022$, $SE = 0.0047$, $t = 4.63$; GD: $b = 0.021$, $SE = 0.0048$, $t = 4.40$). There was a significant three-way interaction between mask condition, frequency, and word length (SFD: $b = -0.0073$, $SE = 0.0035$, $t = -2.09$; GD: $b = -0.0093$, $SE = 0.0036$, $t = -2.59$). This interaction seems to indicate that the reverse effect of word length in the unmasked condition was strongest when word *n* was of low frequency, while it was weaker for higher frequency words *n* and even reversed for extremely high frequency words *n*. It may seem a bit surprising that there was no strong main effect of word *n* length in the $n + 1$ masked condition in SFD, and that the effect of *n* length goes in the opposite direction (with shorter words receiving longer fixations) in the unmasked preview condition. However, it is important to keep in mind that the predictors in LMMs describe partial effects, that is, the effect of the independent variable in question when all other variables in the model are statistically controlled. In our raw data, longer words were associated with longer fixations, just as expected. Word *n* length was also highly correlated with word *n* frequency and predictability, though, with shorter words tending to be higher frequency and more predictable. With *n* frequency and predictability in the model (and therefore being statistically controlled), there either was no remaining partial effect of word length (in the masked condition) or that the remaining partial effect of word length was reversed (in the unmasked condition). This situation is known as a suppressor effect".

This must be changed to read:

"In this Appendix, we report all effects observed in the correlational approach LMM analyses that were not directly relevant to parafoveal preprocessing and therefore were not reported in the main text.

Effects of fixation position. We observed very strong effects of fixation position on SFD and GD. The linear trend was positive and significant in both SFD ($b = 0.013$, $SE = 0.00089$, $t = 15.04$) and GD ($b = 0.0024$, $SE = 0.00096$, $t = 2.48$). There also was a significant quadratic trend in SFD ($b = -0.016$, $SE = 0.00093$, $t = -16.81$) and GD ($b = 0.012$, $SE = 0.00098$, $t = 11.88$).

The length of the incoming saccade showed a significant effect on both SFD ($b = 0.018$, $SE = 0.00093$, $t = 19.00$) and GD ($b = 0.019$, $SE = 0.00099$, $t = 19.13$), with longer fixation times when the incoming saccade was large, and when the outgoing saccade was short (SFD: ($b = -0.0063$, $SE = 0.00091$, $t = -6.92$); GD: ($b = -0.01$, $SE = 0.00097$, $t = -10.63$)).

Lag effects: effects of lexical properties of the word preceding the currently fixated word (word $n - 1$). We observed significant lag effects of frequency in both SFD ($b = 0.018$, $SE = 0.0031$, $t = 5.93$) and GD ($b = 0.016$, $SE = 0.0033$, $t = 4.82$), with words receiving longer fixations when they had been preceded by low-frequency words than when they had been preceded by high frequency words. There was no interaction between $n - 1$ frequency and $n - 1$ length (SFD: $t = 1.22$; GD $t = 0.09$. Finally, we observed a lag effect of predictability, with words showing lower SFD ($b = -0.022$, $SE = 0.0025$, $t = -8.93$) and GD ($b = -0.023$, $SE = 0.0027$, $t = -8.69$) when the preceding word had been highly predictable from the context than when it had not been predictable.

Immediacy effects: effects of the properties of the currently fixated word (word n). There was a significant frequency effect on SFD ($b = -0.03$, $SE = 0.0056$, $t = -5.45$) and GD ($b = -0.03$, $SE = 0.0059$, $t = -5.14$), indicating that words received shorter fixations when they were of high frequency.

There also was a significant effect of predictability, with words showing shorter SFD ($b = -0.023$, $SE = 0.0028$, $t = -8.10$) and GD ($b = -0.025$, $SE = 0.003$, $t = -8.24$) when a word was predictable from the sentence context than when it was not.

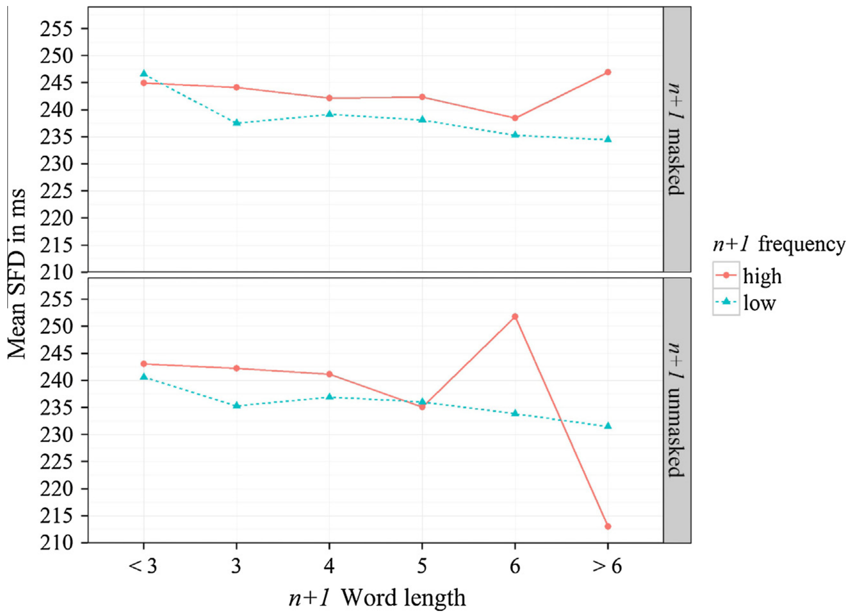

Fig. A1. Predicted partial effect of $n + 1$ word length, $n + 1$ frequency, and $n + 1$ mask condition on single fixation duration (from the fitted LMMs using the remef function by Hohenstein & Kliegl, 2014).

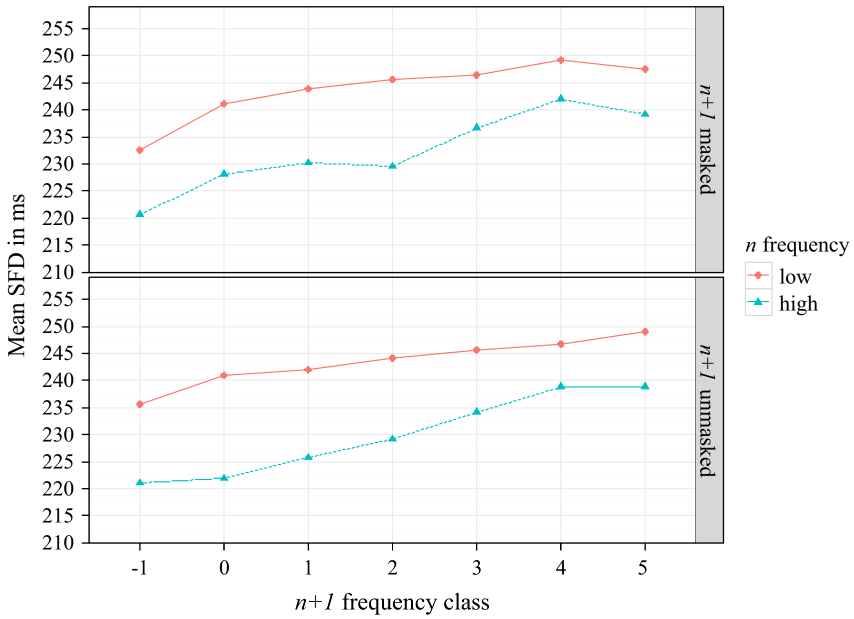

Fig. A2. Predicted partial effect of $n + 1$ frequency, n frequency, and $n + 1$ mask on single fixation duration (from the fitted LMMs using the remef function by Hohenstein & Kliegl, 2014). $n + 1$ frequency class is determined by taking the log 10 of $n + 1$ frequency per million, rounding to the nearest integer.

When $n + 1$ was masked, there was a significant main effect of word length on GD ($b = -0.049$, $SE = 0.006$, $t = -8.20$), but not SFD ($t = 0.03$), and there was no significant interaction between word frequency and length (all $t < 1.96$). However, there was a significant interaction between mask condition and word length, indicating that word length did have an effect when $n + 1$ was not masked (SFD: $b = 0.015$, $SE = 0.0037$, $t = 4.01$; GD: $b = 0.02$, $SE = 0.0039$, $t = 5.02$). The three-way interaction between mask condition, frequency, and word length was not significant (SFD: $t = -0.79$; GD: $t = -0.61$)". Figs. A1 and A2 need to be updated.

Appendix C

The table of random effects needs to be updated: Dependent variable: GD

No random correlations allowed.

Group	Random effect	*SD*
Word	(Intercept)	0.0616
	$n + 1$ length	0.0379
	$n + 1$ frequency	0.0302
	$n + 1$ predictability	0.0341
	$n + 1$ length by $n + 1$ *frequency*	0.018
Subject	(Intercept)	0.1302
	$n + 1$ length	0.0082
	$n + 1$ frequency	0.003
	n frequency	0.0067
	$n + 1$ predictability	0.0369
	$n + 1$ length by $n + 1$ *frequency*	0.0018
	n frequency by $n + 1$ *frequency*	0.0036
Residual		0.29121