Title
The interaction of language processing and eye movement control during reading

Author
Abbott, Matthew James Hansen

Publication Date
2016

UNIVERSITY OF CALIFORNIA, SAN DIEGO

The interaction of language processing and eye movement control during reading

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Psychology

by

Matthew James Hansen Abbott

Committee in charge:

Professor Victor S. Ferreira, Chair
Professor David Barner
Professor Seana Coulson
Professor John M. Henderson
Professor Roger Levy
Professor John Wixted

2016

The dissertation of Matthew James Hansen Abbott is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2016

DEDICATION

To my advisor, Keith Rayner.

EPIGRAPH

Saccade is a fancy word for eye movement

—Keith Rayner

TABLE OF CONTENTS

Signature Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Figures vii

List of Tables viii

Acknowledgements ix

Vita xi

Abstract of the Dissertation xii

Chapter 1 Introduction 1
  1.1 Parafoveal processing and word skipping 2
  1.2 E-Z Reader 10 5
  1.3 Testing assumptions 12
  1.4 Summary 15

Chapter 2 Skipping syntactically illegal *the* previews: The role of predictability 20
  2.2 Method 22
  2.2 Results 23
  2.2 General Discussion 27

Chapter 3 Preferential skipping of function words during reading 34
  3.1 Introduction 34
  3.2 General Method 37
  3.3 Experiment 3.1 39
    3.3.1 Method 39
    3.3.2 Results and Discussion 40
  3.4 Experiment 3.2 45
    3.4.1 Method 45
    3.4.2 Results and Discussion 47
  3.5 General Discussion 51
  3.6 Appendix: Experimental stimuli from Experiment 3.1 56

Chapter 4 The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions 70
4.4 Introduction 71
4.4 Method 74
4.4 Results 76
4.4 Discussion 79

Chapter 5 General Discussion 84
5.1 Skipping as a hedged bet 84
5.2 Are short words treated differently? 86
5.3 Lexical versus postlexical processing 87
5.4 Implications for models of eye movement control 88
5.5 Conclusion 91

LIST OF FIGURES

Figure 2.1: Example stimuli. The display changed to the correct target word when readers moved their eyes to the right of the invisible boundary (dashed line). 23
Figure 2.2: Target word skipping. Error bars represent standard error of the mean. 27
Figure 2.3: Go-past time on the posttarget word. Error bars represent standard error of the mean. 27
Figure 2.4: Simulation 1: Target skipping probabilities by condition. 28
Figure 2.5: Simulation 2: Target skipping probabilities by condition. 28
Figure 2.6: Skipping probabilities for the target words from the observed data (left panel) and from Simulation 2 (right panel). 29

Figure 3.1: Target word skipping, Experiment 3.2. 48

Figure 4.1: Grand means and standard errors for binomial dependent measures on the target word. 76
Figure 4.2: Grand means and standard errors for skipping probability by word length and frequency (top panel) and plausibility (bottom panel). . . 77
Figure 4.3: Grand means and standard errors for fixation time measures on the target word. (For first fixation duration and gaze duration, error bars are obscured by the plotting symbols.) 78
Figure 4.4: Posterior distribution of the effects of frequency and plausibility and their interaction on first fixation duration estimated through MCMC sampling using BayesFactor. 80

LIST OF TABLES

Table 2.1: Condition means for the pretarget word. Note. Standard errors are in parentheses. SFD = single fixation duration; GD = gaze duration . . 24
Table 2.2: Linear mixed effects model analyses on the pretarget word. Note. Each column represents a model fit to one of the dependent variables. 24
Table 2.3: Condition means for the target word. Note. Standard error of the mean in parentheses. SFD = single fixation duration; GD = gaze . . 25
Table 2.4: Linear mixed effects model analyses for the target word. Note. Each column represents a model fit to one of the dependent variables. . . . 25
Table 2.5: Condition means on the posttarget word. Note. SFD = single fixation duration; GD = gaze duration. 26
Table 2.6: Linear mixed effects model analyses for the posttarget word. Note. Each column represents a model fit to one of the dependent variables. 27

Table 3.1: Mean natural log-transformed target word frequencies by condition and experiment as obtained from ELP. 38
Table 3.2: Target word subject means and standard errors by condition: Experiment 1. 42
Table 3.3: GLMM and LMM results: Target word, Experiment 1. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). 43
Table 3.4: Post-target word subject means and standard errors by condition: Experiment 1. 44
Table 3.5: GLMM and LMM results: Post-target word, Experiment 1. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). . . . 46
Table 3.6: Target word subject means and standard errors (in parentheses) by condition: Experiment 2. 47
Table 3.7: GLMM and LMM results: Target word, Experiment 2. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). 49
Table 3.8: Post-target word subject means and standard errors (in parentheses) by condition: Experiment 2. 50
Table 3.9: GLMM and LMM results: Post-target word, Experiment 2. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). All . 52

Table 4.1: Grand means for eye movement measures on the target word, by condition. Duration measures are in ms. Standard error of the mean . 76
Table 4.2: Analyses of variance with random error by subjects (F_1) and by items (F_2) for the dependent measures on the target word. 77
Table 4.3: Bayes factors for each of the nested models in the ANOVA design compared to the by-subjects or by-items random effects model. . . . 78
Table 4.4: Bayes factors for each of the nested models in the ANOVA design compared to the by-subject random effects only model for 79

Table 4.5: Mean and *SD* of posterior distribution for each parameter of interactive effects model of first fixation duration (by subjects) 69

ACKNOWLEDGEMENTS

First and foremost, this dissertation would not have been possible without the endless love and support of my wife Stèphanie, who sacrificed so much to help me through it. I feel it is very much her accomplishment as it is mine. My family (little Evie and Ruby included) and my parents (and in-laws) were amazingly supportive throughout, as were current and former members of the UCSD Rayner Lab (especially Patrick, Liz, Mallorie, Titus, Emily, Danbi, and Nathalie). I thank the members of my dissertation committee for being so supportive following the death of my advisor, and for putting in the effort to stick with the tight timeline I drew (especially Vic and Dave who coached me through it). I also thank my co-authors (especially Bernhard and Adrian, who often felt like surrogate advisors), Simon, the 199s who helped run my experiments, and everyone who ever listened to me talk about these data.

I am deeply grateful to my late advisor, Keith Rayner, who took a chance and invited me to be his student. I owe much of my development as a thinker to him. Although I am saddened that he did not see the works in this dissertation published, I think he would have said I did pretty good. I miss Keith's unique way of seeing the world, and his humor. While writing this dissertation, I often asked myself "What the crap would Keith say?" and tried to fill in accordingly.

Chapter 2, in full, is a reprint of the material as it appears in Skipping syntactically illegal *the* previews: The role of predictability (Abbott, Angele, Ahn, & Rayner, 2015). The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of material in preparation for publication: Preferential skipping of function words during reading (Abbott & Angele, in prep). The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material as it appears in The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions (Abbott & Staub, 2015). The dissertation author was the primary investigator and author of this paper.

VITA

2010	B.A. in Psychology *summa cum laude*, University of Massachusetts Amherst
2013	M.A. in Psychology, University of California, San Diego
2016	Ph. D. in Psychology, University of California, San Diego

PUBLICATIONS

2015 Abbott, M. J., & Staub, A. “The effect of plausibility on eye movements in reading: Testing E-Z Reader’s null predictions.” *Journal of Memory and Language*, 85, 76-87. doi:10.1016/j.jml.2015.07.002

2015 Abbott, M. J., Angele, B., Ahn, Y. D., & Rayner, K. “Skipping syntactically illegal *the* previews: The role of predictability” *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 41(6), 1703–1714. doi:10.1037/xlm0000142

2015 Rayner, K., Abbott, M. J., & Plummer, P. “Individual differences in perceptual processing and eye movements in reading. In *Handbook of Individual Differences in Reading: Text and Context.* New York, NY: Informa UK Limited.

2012 Staub, A., Abbott, M., & Bogartz, R. S. (2012). “Linguistically-guided anticipatory eye movements in scene viewing." *Visual Cognition*, 20, 922-946.

ABSTRACT OF THE DISSERTATION

The interaction of language processing and eye movement control during reading

by

Matthew James Hansen Abbott

Doctor of Philosophy in Psychology

University of California, San Diego, 2016

Professor Victor S. Ferreira, Chair

In this dissertation, I addressed three questions regarding the impact of language processing on eye movements in reading: (1) What information from the sentence context do readers recruit prior to skipping? (2) Are words like *the* unique with regard to skipping because they are so short? (3) Does postlexical integration strictly follow word identification? Throughout, I recruit the E-Z Reader 10 model (Reichle, Warren, & McConnell, 2009) for testable predictions. In Chapter 2, I demonstrate that the influence of the sentence context on skipping is limited to predictability, and not deeper semantic or syntactic processing. Readers skipped over invalid, contextually infelicitous parafoveal previews of *the* more often than valid previews of three-letter words, even when the target words were predictable. In Chapter 3, I show that this effect generalizes to longer function words like *that* and *there*, which highlights the role of parafoveal processing relative to oculomotor or contextual constraints (e.g., syntactic or semantic fit) in word skipping. Comparing Experiment 3.2 to Experiment 3.1, it is also evident that readers prefer to skip function words over content words, regardless of their length. The data presented in Chapters 2–3 pose a deeper question about the relationship between language processing and eye movement control that indicates a possible constraint on the architecture: Are word identification and postlexical integration isolated, separable processes? This issue is examined in Chapter 4 by jointly manipulating word frequency and plausibility. I demonstrate probabilistic evidence against an influence of plausibility on word skipping, and in favor of an additive relationship between frequency and plausibility (and against an interactive relationship), by computing Bayes factors. In summary, I paint a picture in which word skipping reflects a hedged bet that identification will ultimately be successful, and all postlexical processing (i.e., of syntax and semantics) follows lexical access.

Chapter 1

Introduction

Since the seminal work of Huey (1908), a great deal of attention has been paid to the study of eye movements in reading. Much of this work has been carried out by Rayner and colleagues (see Rayner, 1998, 2009 for reviews; see also Clifton et al., 2016 for an overview of Rayner's contributions to eye movement research) using advanced eyetracking technology. As Rayner found throughout his work, studying eye movements in reading carries multiple benefits: It provides a well-controlled environment in which to examine the constraints of the eye movement control system, and it provides a glimpse into the architecture of the language processing system. This dissertation supplies three main contributions to our theoretical understanding of eye movements in reading. The studies presented in Chapter 2 demonstrate that parafoveal processing prior to skipping over words is influenced by multiple sources of information but is limited in scope. The studies presented in Chapter 3 show that readers rapidly identify function words from the parafovea, and skip them preferentially to other words of the same length. The study presented in Chapter 4 builds on the preceding chapters by demonstrating strong evidence that the process of word identification strictly precedes integration into the broader sentence or discourse context. Taken together, these studies further characterize the extent of lexical processing when it takes place in parafoveal vision, and the manner in which postlexical integration follows lexical processing more broadly.

1.1 Parafoveal processing and word skipping

Much of this dissertation is concerned with a specific eye movement behavior during reading: Word skipping. It is well known that we do not fixate every word in a sentence, and that skilled readers of English skip approximately 25-30% of words during first-pass reading (Rayner, 1998, 2009). The words we skip tend to be short, highly frequent, or predictable from the preceding context. An important question is the extent to which words are processed parafoveally prior to being skipped (see Schotter, Angele, & Rayner, 2012, for a review of parafoveal processing in reading). Studying skipping can reveal, for example, the types of information available to readers about words before they are directly fixated (i.e., phonological information, syntactic or semantic information) and the extent to which that information can influence saccade decisions. Examining skipping is broadly important to validating theoretical frameworks like E-Z Reader (Reichle, Pollatsek, Fisher, & Rayner, 1998) which posit that word skipping is dependent on the completion of an early stage of word identification, which are in opposition to frameworks which assume that skipping decisions are affected more heavily by oculomotor constraints (McConkie, Kerr, Redix, & Zola, 1988).

It is clear that parafoveal information is used during reading to direct where the eyes move next. In English, spaces are used to determine word boundaries and to target saccades (Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998). Word length is also a strong predictor of fixation probability: Long words are more likely to be fixated than short words (Rayner & McConkie, 1976; Rayner, Sereno, & Raney, 1996; Rayner, Slattery, Drieghe, & Liversedge, 2011).

There is substantial evidence that lexical processing is involved in word skipping, as is directly assumed by computational models like E-Z Reader (Reichle et al., 1998; Reichle, Warren, & McConnell, 2009) and SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005). These models posit that factors like word length determine how quickly a word is identified. Much of the past work examining the influence of word identification on skipping used the gaze-contingent boundary paradigm (Rayner, 1975), in which the parafoveal preview of a target word is manipulated prior to direct fixation. In this paradigm, before readers' eyes cross an invisible boundary (typically in the character space preceding the target word) either a valid preview of the target word is displayed or an incorrect preview such as a different word, a non-word, or a mask (e.g., Xs); once the eyetracker detects a saccade crossing the boundary the preview is always replaced with the target word.

Several studies have found, for example, that readers are less likely to skip over parafoveal previews of non-words than words (e.g., Balota, Pollatsek, & Rayner, 1985; Drieghe, Rayner, & Pollatsek, 2005; Gordon, Plummer, & Choi, 2013). Skipping is also influenced by word frequency (Angele, Laishley, Rayner, & Liversedge, 2014; Angele & Rayner, 2013; Gollan et al., 2011). When matched on length, readers are more likely to skip high-frequency words than low-frequency words. Word frequency is a measure of how difficult a word is to identify (Scarborough, Cortese, & Scarborough, 1977) that also affects the duration of a readers' eye fixation on a word (Rayner, 1998, 2009), and so the presence of a frequency effect on skipping probability is evidence that lexical processing occurred during the prior fixation. Additionally, words that are more predictable from the prior context are skipped more often than less predictable words (Balota et al., 1985; Drieghe et al., 2005; Rayner et al., 2011).

Clearly, the effect of predictability indicates that readers' saccade decisions are also influenced by the sentence context to an extent. The most common measure used to assess the predictability of a target word is cloze probability (see Staub, Grant, Astheimer, & Cohen, 2015 for a review of the cloze task), an offline measure which is the proportion of participants who entered the correct target word when asked to provide a word given the preceding words in the sentence. Thus, we expect that cloze probability reflects the likelihood that a reader would guess the next word conditioned on having read and understood the prior words in the sentence. Readers skip over high cloze words more often than low-cloze words, but there appears to be no difference in skipping medium versus low-cloze words (Rayner & Well, 1996). The influence of the context on skipping seems to be limited to cases in which the target word is nearly obligatory.

Although the predictability effect suggests that the context can influence skipping decisions, it appears to be limited in scope. Angele and Rayner (2013) presented readers either correct parafoveal previews of three-letter words (e.g., *all*), or incorrect previews of the word *the* in syntactically infelicitous positions (e.g., She was sure she would *ace* all the tests; target word in italics). They found that readers skipped over *the* more often than the correct previews, even though *the* was inappropriate in context. Readers encountered disruption downstream, with longer fixation times and increased regressions made out of the post-target region. Angele et al. (2014) provided further evidence that this effect generalizes to other higher-frequency previews, and not just *the*. This indicates an interesting limitation of the influences of the sentence context on eye movement behavior. Are readers able to integrate information they receive from the parafovea with their current representation of the sentence context in order to make their decision to skip or fixate the next word?

Taken together, these results suggest that the types of integrative processes that are performed in order to “guess” a word and skip over it may be very different from those which inform a reader that a syntactically or semantically anomalous word has been encountered (which does not appear to affect skipping).

1.2 E-Z Reader 10

The E-Z Reader 10 model (Reichle et al., 2009) provides a relevant theoretical and computational framework in which to interpret the above findings, and to generate novel predictions. E-Z Reader (Reichle et al., 1998) assumes that words are processed serially (in contrast to models like SWIFT that assume parallel lexical processing; Engbert et al., 2005), and that attention can be decoupled from ongoing saccadic programming.

In the model, when the eyes land on word *n* the letters are processed in parallel across the visual field. Lexical processing then proceeds in two stages, L_1 and L_2. The completion of L_1 (the "familiarity check") indicates that word recognition is imminent and triggers the initiation of saccadic programming to the next word *n+1*, which also proceeds in two stages (a labile stage M_1 and a non-labile stage M_2). The completion of L_2 indicates lexical access, after which attention shifts covertly to word *n+1* and postlexical processing for word *n* proceeds. Lexical processing begins from the parafovea after the attention shift, and if the familiarity check (L_1) for the parafoveal word completes while the saccade program to word *n+1* is still labile, that program can be canceled and the word is skipped over.

As is apparent from the architecture, whether or not a word is skipped is critically dependent on the timing of L_1 for word *n+1* relative to ongoing saccadic programming. Equation 1.1 reveals E-Z Reader's assumption that a word's frequency of occurrence and predictability are solely responsible for the timing of early lexical processing. The upper branch of the equation states that a word may be "guessed" from the parafovea, with L_1 set to 0 ms, with probability (p) equal to that word's cloze probability (determined using an offline cloze completion task). The lower branch indicates that if a word is not guessed in this manner, L_1 is an additive function of frequency and predictability (where α_1 is overall lexical processing rate, and α_2 and α_3 are parameters that determine the relative influence of frequency and predictability on L_1).

$$L_1 = \begin{cases} 0 & \text{if } p \leq \text{predictability}_n \\ \alpha_1 - \alpha_2 \ln(\text{frequency}_n) - \alpha_3 \, \text{predictability}_n & \text{if } p > \text{predictability}_n \end{cases} \tag{1.1}$$

The impact of word length on lexical processing rate is accounted for by scaling the rate of L_1 by the position of the current fixation relative to the mean eccentricity of the letters of the word being processed (see Equation 1.2; ε is a free parameter that controls the extent to which limitations of acuity slow word identification). Longer words will have higher mean eccentricity on average, resulting in longer lexical processing rate, longer fixation times, and a reduced likelihood of being skipped.

$$t(L_1)' = t(L_1)\, \varepsilon^{\Sigma\, |\text{fixation - letter}|\, /\, N} \tag{1.2}$$

Postlexical integration in E-Z Reader 10

Critically, L_1 in E-Z Reader is intended to reflect readers' "hedged bet" that recognition is imminent. Taken literally, we are to assume that full recognition has not yet occurred (as this is accomplished with the completion of both L_1 and L_2). Processing that pertains to integrating a word's meaning into the broader syntactic or semantic context is delayed until lexical access has completed, within a postlexical integration stage (*I*), introduced in E-Z Reader 10 (Reichle et al., 2009).

E-Z Reader's integration stage (see Staub, 2011 for more detail and model simulations) is intended to capture all integrative process that occur after lexical access has completed. It is assumed that the difficulty of identification has no direct bearing on the difficulty of integration. Integration is largely intended to run in the background–it runs in tandem with the shift of attention to the next word and subsequent word identification processes, and generally has no detectable effect on eye movements. However, there are two ways in which I will have a direct effect on eye movements: (a) if I has not completed by the time the next word has been identified, attention shifts back to the source of difficulty (either resulting in longer fixation times on that word or regressions if the eyes had already moved on) or (b) integration fails outright with probability P_F, resulting in longer fixation times (if integration fails while the outgoing saccade program can still be cancelled) or regressions.

If we are to assume, then, that recognizing a syntactic anomaly would require accessing a word's meaning from memory, and failing to integrate it, then this architecture provides an explanation to the Angele and Rayner (2013) findings: *the* may be skipped so often because readers rarely have difficulty guessing its identity from the parafovea. Because integrative processing cannot begin until after lexical access has completed, and the signal to hold the eyes in place can only come after some additional delay, it seems unlikely that readers would have enough time to detect the syntactic anomaly and cancel the decision to skip before the eyes move on.

This critical dissociation is uniquely predicted by E-Z Reader 10. A question that remains unanswered is how integrative processing boosts activation for particular words, leading them to be, in the extreme, "guessed" and skipped over (or identified more rapidly), but does not slow early identification processes for unexpected words.

One possibility is that there is a threshold above which contextual constraint (in terms of cloze probability) will affect skipping. The findings reported by Rayner and Well (1996) appear to support this claim, as they found no difference in skipping medium versus low-cloze words (but did find a difference for high versus medium-cloze words). It may be that the effect of predictability on skipping is unique in that the context exerts an immediate effect and speeds the retrieval of a words meaning from memory (as is predicted by the upper branch of E-Z Reader's L_1 computation). It is not clear, though, whether predictable words are also integrated immediately (i.e., because their meaning is available right away), or if highly constrained words are subject to the same integrative processes as weakly constrained words. Indeed, it seems difficult to imagine a situation in which a word would be predictable but difficult to integrate (unless, for example, that word is highly constrained by the local context but is somehow infelicitous in the broader discourse).

The influence of word length

Of course, it is possible that very short words like *the* are skipped for reasons other than (or in addition to) word identification. Short saccades tend to be more error-prone than average length saccades (7-8 characters; Rayner, 1998. Indeed, Nuthmann, Engbert, and Kliegl (2005) demonstrated the skipping of very short words can be sometimes explained by mislocated fixations (i.e., a saccade overshooting a short word and landing on a subsequent longer word), due to saccadic range error (McConkie et al., 1988). Thus, words like *the* might be skipped more often than longer words for at least two reasons, because they are processed more quickly from the parafovea, and/or because of oculomotor constraints like saccadic range error.

By this logic, we might expect a factor like frequency which influences the rate of identification to interact with a factor like length, because the effect of word length is a mixture of cognitive and oculomotor constraints. More specifically, the effect that the rate of word identification has on skipping (i.e., by way of a word frequency manipulation) very short words may be obscured because short words are more likely to be skipped over accidentally since short saccades are hard to execute. For a word to be skipped over "accidentally" we would have to assume that the eyes intended to land on it, and so it had not been previously identified from the parafovea. If this is the case, then we would expect to see smaller effects of word frequency on the skipping of very short (e.g., three-letter) words than longer (e.g., four or five-letter) words (because something other than its frequency led it to be skipped). However, by that same token the effect of word frequency on skipping may be smaller for longer versus shorter words because processing rate is slowed for longer words that extend further out into the parafovea.

Independence of lexical and postlexical processing

The strict sequencing of lexical and postlexical processing assumed by E-Z Reader 10 leads to a third assumption: The effects that lexical and postlexical processing have should have independent effects across the eye movement record. The very earliest measure, skipping, should be affected by the difficulty of word identification, while it should not be affected by the difficulty of integration, and in duration measures their effects should be strictly additive; the difficulty of integration may affect the likelihood of making a regression, although the difficulty of identification should not. Using additive factors logic (Sternberg, 1969), we would assume that if word identification and postlexical integration processes overlapped, their effects would be interactive.

Reichle et al. (2009) do not commit to a specific linguistic model of postlexical processing, and assumes that difficulties with syntactic processing (e.g., garden-path sentences; Frazier & Rayner, 1982) or semantic processing (i.e., implausible or anomalous words; Rayner, Warren, Juhasz, & Liversedge, 2004) impact eye movements similarly. That is, a difficult to integrate word will either cause longer fixation times or regressions. Staub (2011) tested these assumptions with a series of E-Z Reader 10 simulations and experiments in which a word's frequency and syntactic attachment difficulty were manipulated. Consistent with E-Z Reader, when frequency and syntactic processing difficulty had effects in the same measure (as early as *first fixation duration*, the duration of a reader's first eye fixation on a word), their effects were never significantly interactive. Additionally, although attachment difficulty affected the rate with which readers made regressive eye movements (more regressions were made following temporarily ambiguous constructions), this measure was not significantly affected by word frequency.

1.3 Testing assumptions

Chapter 2

The patterns of data reported by Angele and colleagues appear to fit within the framework of E-Z Reader 10. If one assumes that word skipping is based on rapid parafoveal identification, then words should be skipped according to factors that affect processing rate like word frequency. How information from the sentence context does or does not affect skipping is critically important to an understanding of how very early word identification processes may differ from later integrative processing. We expect that predictability influences skipping because the integration of all words prior to a target word either allows readers to "guess" the target or process it more quickly (e.g., by preactivation of its representation in memory). Because the process of integration lags behind complete identification, we expect that readers will not be able to detect syntactic anomalies prior to making the decision to skip or fixate the target word. Thus, effects of the context on skipping should be limited to cases where the target is highly predictable.

The studies presented in Chapter 2 provide a strong test of these assumptions by manipulating parafoveal preview (correct three-letter word or *the*) and predictability (unpredictable or predictable target word) using example sentences like (1). Importantly, the false previews of *the* were never permitted in the sentence context.

1. (a) Jane used the scissors to carefully *cut*... (Predictable, valid preview)
   (b) Jane used the scissors to carefully *the*... (Predictable, *the* preview)
   (c) Jane used the machine to carefully *cut*... (Unpredictable, valid preview)

(d) Jane used the machine to carefully *the*... (Unpredictable, *the* preview)

If E-Z Reader is correct in that skipping entails rough parafoveal identification, then the target words should be skipped more often when they are predictable versus unpredictable in context. If the process of syntactic integration lags behind identification, then *the* should be skipped more often than any other three-letter word regardless of its syntactic fit because it is more frequent and easier to identify. If the predictability effect entails the comparison of expectations against parafoveal input (i.e., integration), then there should be an interaction such that skipping of *the* is reduced when the target word is predictable compared to when it is unpredictable.

Chapter 3

Building on the findings of Chapter 2, we examine parafoveal processing in more depth by considering the joint effects of word frequency and word length on skipping in Chapter 3. If skipping is based on a rough check of the identity of the parafoveal word, and is not dependent on any deeper integrative processing, then short and high frequency words should be skipped more often than long and low frequency words because both factors affect how rapidly a word is identified. The relationship between word length and frequency with regard to skipping is potentially complicated, though, by the fact that short words are more likely to be skipped over due to saccadic overshoot than are longer words. Understanding how cognitive processing versus oculomotor constraints influence eye movement behavior is important to validating theories like E-Z Reader that rely on word identification as the engine that moves the eyes through the text.

We examine this issue in Chapter 3 by examining word skipping with three, four, and five-letter words. If words like *the* are skipped so often because they are processed very rapidly from the parafovea, and the effect of length just slows that processing, we expect readers to skip words like *that* and *there* less often than *the*, but more often than lower-frequency words of the same length. If oculomotor constraints cause very short words to be skipped over more frequently than long words regardless of their frequency (or if the effect of length dramatically reduces the rate of parafoveal processing as length increases), we expected an interaction such that the frequency effect on skipping would differ for the different word lengths.

Chapter 4

Chapters 2 and 3 lay the foundation for a deeper theoretical question: Do lexical and postlexical processes occur in discrete stages? According to E-Z Reader 10 (Reichle et al., 2009) this is the case, and the model makes a number of specific predictions regarding how lexical and postlexical processes should unfold across time for any word *n*. Specifically, as in the preceding chapters, variables like frequency and predictability that affect the rate of lexical processing should affect the rate with which words are skipped, but variables that affect postlexical integration (i.e., syntactic attachment difficulty or plausibility) should not. Following Staub (2011), when factors that affect both frequency and integration are present in the same fixation time measure, their effects should be strictly additive. Finally, integration difficulty may affect the probability of making a regression away from the target word to previous material, but lexical processing will not. Chapter 4 builds on Staub (2011) by testing these hypotheses using factorial manipulations of frequency and plausibility as in (2), and by testing null main effects and interactions directly by computing Bayes factors (Edwards, Lindman, & Savage, 1963; Rouder, Morey, Speckman, & Province, 2012).

2. (a) The professor invited the *writer* to an important meeting. (HF, plausible)
    (b) The professor invited the *orator* to an important meeting. (LF, plausible)
    (c) The professor repaired the *writer* with a trusty old wrench. (HF, implausible)
    (d) The professor repaired the *orator* with a trusty old wrench. (LF, implausible)

1.4 Summary

In summary, the E-Z Reader model provides a testable framework for understanding the flow of information during reading and its resulting impact on eye movements. During reading, forward movements of the eyes are driven by word identification, and in some cases efficient processing of words in the parafovea can lead to word skipping. However, the early stage of identification accomplished before skipping is dissociated from the processing that takes place later on during postlexical integration. In the following sections, I detail three investigations of these claims. In the first investigation, I ask whether word skipping is driven by the ease of parafoveal identification, even when a different word is predicted by the context. In the second investigation, I ask whether the effect that frequency has on skipping remains constant across words of different length, or whether oculomotor constraints reduce the frequency effect for short words. Finally, the third study more deeply investigates whether the eye movement record supports an account under which lexical and postlexical processing occur in discrete stages.

References

Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1181-1203.

Angele, B., & Rayner, K. (2013). Processing the in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 649–662.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390.

Clifton, C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith rayner's 40 year legacy. *Journal of Memory and Language*, *86*, 1-19.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 954-969.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, *70*, 193-242.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). Swift: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777-813.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178–210.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186–209.

Gordon, P. C., Plummer, P., & Choi, W. (2013). See before you jump: Full recognition of parafoveal words precedes skips during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1032-1046.

Huey, E. (1908). *The psychology and pedagogy of reading: With a review of the history of reading and writing and of methods, texts, and hygiene in reading*. The Macmillan company.

McConkie, G. W., Kerr, P. W., Redix, M. D., & Zola, D. (1988). Eye movement control during reading: I. the location of initial eye fixations on words. *Vision Research*, *28*, 1107-1118.

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

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 817-833.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. doi: 10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K. (2009). The thirty-fifth sir frederick bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*, 1457–1506.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research*, *38*, 1129-1144.

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, *16*, 829–837.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 1188–1200.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 514-528.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning,* *Memory, and Cognition*, *30*, 1290-1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*, 504-509.

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*, 1-21.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default bayes factors for anova designs. *Journal of Mathematical Psychology*, *56*, 356-374.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 1–17.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, and Psychophysics*, *74*, 5-35.

Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General*, *140*, 407-433.

Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, *82*, 1-17.

Sternberg, S. (1969). The discovery of processing stages; extensions of donders' method. *Acta Psychologica*, *30*, 276-315.

Chapter 2

Skipping syntactically illegal *the* previews: The role of predictability

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/15/$12.00 http://dx.doi.org/10.1037/xlm0000142

Skipping Syntactically Illegal *the* Previews: The Role of Predictability

Matthew J. Abbott
University of California, San Diego

Bernhard Angele
Bournemouth University

Y. Danbi Ahn and Keith Rayner
University of California, San Diego

Readers tend to skip words, particularly when they are short, frequent, or predictable. Angele and Rayner (2013) recently reported that readers are often unable to detect syntactic anomalies in parafoveal vision. In the present study, we manipulated target word predictability to assess whether contextual constraint modulates *the*-skipping behavior. The results provide further evidence that readers frequently skip the article *the* when infelicitous in context. Readers skipped predictable words more often than unpredictable words, even when *the*, which was syntactically illegal and unpredictable from the prior context, was presented as a parafoveal preview. The results of the experiment were simulated using E-Z Reader 10 by assuming that cloze probability can be dissociated from parafoveal visual input. It appears that when a short word is predictable in context, a decision to skip it can be made even if the information available parafoveally conflicts both visually and syntactically with those predictions.

Keywords: word skipping, predictability, eye movement control, reading, parafoveal processing

One striking observation about skilled readers' eye movements is that they do not fixate every word in a sentence. Instead, quite a few words are skipped. This tends to occur mainly for short words, but longer words are skipped occasionally. In order to read efficiently, skilled readers must decide very quickly—within the first 150 ms of a fixation (Rayner, 1998, 2009)—whether to skip or to fixate the next word and then initiate the appropriate saccade program. During this time, readers have two sources of information to base their decision on: First, they can use parafoveal information that is available about the next word; second, they can use information about the sentence context they have previously read. However, it is not clear whether readers actually use both of these sources of information in making their skipping decisions.

It does seem clear that parafoveal information is important in deciding where to look next. That is, readers of English use parafoveal input in order to identify word boundaries and target words accurately (Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998). Although word length has a strong influence on fixation probability (Rayner & McConkie, 1976; Rayner, Sereno, & Raney, 1996), other properties of the parafoveal word also affect whether it is skipped or fixated: Articles and other closed-class words are skipped more often than open-class words of the same length such as three-letter verbs (Angele & Rayner, 2013; Drieghe, Pollatsek, Staub, & Rayner, 2008; Gautier, O'Regan, & Le Gargasson, 2000; O'Regan, 1979, 1980), and high-frequency words are skipped more often than low-frequency words (Gollan et al., 2011; Rayner et al., 1996). However, there is also much evidence that prior sentence context in the form of predictability of a target word influences skipping. If a word is highly predictable from the sentence context it is skipped more often than a less predictable word (Balota, Pollatsek, & Rayner, 1985; Drieghe, Rayner, & Pollatsek, 2005; Ehrlich & Rayner, 1981; Fitzsimmons & Drieghe, 2013; Rayner, Slattery, Drieghe, & Liversedge, 2011; Rayner & Well, 1996).

Another form of sentence context, namely, the syntactic context, has been examined in studies of word skipping. Specifically, Angele and Rayner (2013) tested whether parafoveal information or sentence syntactic context information had the stronger effect on word skipping by pitting the two influences against each other. They did this by using the gaze-contingent display change paradigm (Rayner, 1975) to manipulate the parafoveal information that readers received about a three-letter verb embedded in a sentence. For example, while fixating the word *always* in the sentence "They always dim the lights at

---

Matthew J. Abbott, Department of Psychology, University of California, San Diego; Bernhard Angele, Psychology Research Centre, Faculty of Science and Technology, Bournemouth University; Y. Danbi Ahn and Keith Rayner, Department of Psychology, University of California, San Diego.

This research was supported by Grant HD065829 from the National Institute of Child Health and Human Development, and the results were presented at the 27th Annual CUNY Conference on Human Sentence Processing in Columbus, Ohio, and the Psychonomic Society's 55th Annual Meeting in Long Beach, California. We thank Aleksandar Sheringov for assistance with data collection. We would also like to thank Erik D. Reichle and two anonymous reviewers for their comments on an earlier version of this article. In memoriam, we thank Keith Rayner for advising us in our graduate studies (and beyond) and for imparting his wisdom throughout this study and many others. We are very fortunate to have had the opportunity to work with such an astounding and influential scientist, and he is greatly missed.

Correspondence concerning this article should be addressed to Matthew J. Abbott, Department of Psychology, University of California, San Diego, 9500 Gilman Drive, MC 0109, La Jolla, CA 92093-0109. E-mail: mabbott@ucsd.edu

night," the parafoveal preview information available to readers about the word *dim* could be (a) the word *dim* itself as a control condition, (b) a random letter preview (*fda*), or (c) the article *the*; following "*They always*," the article *the* is syntactically illegal or infelicitous. Once the reader's saccade crossed an invisible boundary location, the preview changed to the target word (i.e., *dim*). If readers have access to syntactic context information while making their skipping decision, they should not attempt to skip a word that is clearly infelicitous given the previous words of the sentence. On the other hand, if readers only rely on the parafoveal information to guide their skipping decision, without taking the syntactic context into account, they should be quite likely to skip the article *the* because it is an extremely high-frequency function word. Angele and Rayner (2013) found that the latter was the case: Readers were virtually as likely to skip the infelicitous *the* previews as they were to skip *the* when it occurred in a felicitous position. Only after moving on to subsequent words did readers experience disruption when they had skipped the infelicitous *the* previews, evidenced by higher go-past times and regression rates compared with the control condition. Angele, Laishley, Rayner, and Liversedge (2014) showed that this effect was not limited to the article *the*, but also applied to short high-frequency open-class words. For example, readers were more likely to skip the parafoveal preview of a high-frequency word (*dog*) than the preview of a low-frequency word (*dim*), even when the sentence syntactic context only allowed the low-frequency word ("The increasingly *dim/dog* light made it hard to see").

Taken together, the studies by Angele and Rayner (2013) and Angele et al. (2014) indicate that sentence context has little effect on word skipping. However, this conclusion appears to contradict the many experiments that have found clear effects of word predictability on word skipping. One possible explanation for this apparent contradiction is that sentence context may only have an effect on fixation probability when it is highly constraining, that is, when the word to be skipped is highly predictable from the sentence context. In the studies by Angele and Rayner (2013) and Angele et al. (2014), the sentence largely did not constrain the target word, which might explain the absence of sentence context effects on skipping. The present study tested this hypothesis, as we manipulated both the parafoveal preview that readers received of a target word (identical vs. infelicitous *the*) and the degree of constraint of the sentence context it was embedded in. For example, the target word *cut* was either presented in a highly predictable context ("Jane used the scissors to carefully *cut* scraps of paper") or in an unconstrained context ("Jane used the machine to carefully *cut* scraps of paper").

If sentence context effects are limited to cases in which the target word is highly constrained, we expected to find an interaction between constraint and preview: Readers should be less likely to skip the infelicitous *the* previews in the high-constraint condition than in the low-constraint condition. If constraint has no effect on the influence of sentence context, on the other hand, skipping rates for the infelicitous *the* previews should not differ between the high- and the low-constraint conditions. The experiment is also simulated using E-Z Reader 10 (Reichle, Warren, & McConnell, 2009) to assess these different possibilities.

Method

Subjects

Forty-four University of California, San Diego students participated in the experiment for course credit. All were native English speakers with normal or corrected-to-normal vision, and were naïve to the purpose of the study.

Apparatus

Subjects' eye movements were recorded with a sampling rate of 1000 Hz using an SR Research EyeLink 1000 (SR Research, Toronto, Ontario, Canada) eyetracker. Sentences were displayed in 14-pt Courier New font on a Hewlett Packard p1230 CRT monitor with a refresh rate of 150 Hz. Viewing distance was approximately 60 cm and each character subtended about .3 degrees of visual angle. Only the right eye was recorded, although viewing was binocular.

Materials

Forty experimental sentence frames comprised four conditions (2 constraint × 2 preview). Each sentence included a three-character target word that was used as a verb[1] (e.g., "Jane used the scissors to carefully *cut* scraps of paper"; target word in italics; see the Appendix for the complete list of sentences). Pretarget words were chosen to be of sufficient length to avoid skipping (mean length 6.3 characters). We included two preview conditions: A correct preview of the target word (*cut*) and a *the*-preview containing an infelicitous preview of an article (see Figure 1). The *the*-preview always appeared in a position in the sentence in which it was syntactically inappropriate. On average, the frequency for the target words was 18 counts per million in the Francis and Kučera (1982) corpus.[2] For comparison, the frequency of *the* in this corpus is 69,971 counts per million.

We manipulated the constraint of the target word such that it was predictable or unpredictable depending on the context of the sentence while holding the position of the target word as constant as possible (see Figure 1; on average, the target word was word 9.4 in both the high- and low-constraint sentences). To determine the predictability of the target words, 20 workers recruited through Amazon Mechanical Turk participated in a cloze norming task in exchange for payment. They read the sentences up to the target word, with the target word and subsequent material deleted (e.g., "Jane used the scissors to carefully ___") and were asked to report what they thought the next word in the sentence should be. Subjects entered the target word 76.8% of the time in the high-constraint condition and 5.0% of the time in the low-constraint condition. They never entered *the* as their response.

Procedure

The 40 experimental sentences were embedded among 60 filler sentences. Each sentence was presented on the computer screen

---

[1] There was one exception in which the target word *run* was used as a noun instead of a verb (see the Appendix).

[2] Four of the target words do not have corresponding entries in the Francis and Kučera (1982) corpus.

Constraint	Preview	Sentence
High	target	`Jane used the scissors to carefully cut scraps of paper into neat shapes.`
High	*the*	`Jane used the scissors to carefully the scraps of paper into neat shapes.`
Low	target	`Jane used the machine to carefully cut scraps of paper into neat shapes.`
Low	*the*	`Jane used the machine to carefully the scraps of paper into neat shapes.`

Figure 1. Example stimuli. The display changed to the correct target word when readers moved their eyes to the right of the invisible boundary (dashed line).

individually. Following a 3-point calibration procedure, subjects were asked to read the sentences silently and to press a button when they finished reading. The gaze-contingent boundary paradigm (Rayner, 1975) was used to present either a correct preview of the target word (*cut*) or a preview of *the* that changed to the target word after readers' eyes crossed the invisible boundary (see Figure 1 for an example). The display change completed within 4 ms, on average (range 0 to 7 ms), once the tracker detected a saccade crossing the boundary. After 33 of the 100 sentences, subjects responded to a two-alternative forced-choice comprehension question.

Results

Skipping and regression probabilities, as well as fixation times, were analyzed for three regions: the pretarget word, the target word, and the posttarget word. Prior to analysis, we deleted 4.3% of trials because of blinks or track loss on the target word or on adjacent fixations, as well as 10.2% of trials with display changes that completed after fixation onset. Fixations shorter than 80 ms and within one character of an adjacent fixation were combined (0.3%). Finally, we deleted 3.7% of fixations shorter than 80 ms or longer than 1,000 ms in duration. Comprehension accuracy was high (97% correct).

We report linear mixed models (LMMs) with subjects and items as crossed random effects (Baayen, Davidson, & Bates, 2008). We chose these analyses in favor of ANOVAs because of the uneven cell sizes that result from word skipping, in which skipped words are treated as missing data in duration measures. LMMs were fit using the lmer function from the lme4 package (Version 1.1–7; Bates, Maechler, Bolker, & Walker, 2014) in the R Environment for Statistical Computing (Version 3.1.0; R Core Team, 2014). Generalized LMMs were fit (using a logit link) to two binomial dependent measures, skipping probability and probability of regressions out, and we report regression coefficients (b), standard errors, and z values for fixed effects and their interactions (t values are reported for duration measures). LMMs were fit to the following log-transformed fixation duration measures: single fixation duration (SFD; the mean duration of fixations on a word when that word received just one fixation in the first pass), gaze duration (GD; the sum of all fixations on a word before leaving it, including refixations), and go-past time (the sum of all fixations on a word before leaving it to the right, including regressions to previous words). Two fixed factors were included and contrasts were sum-coded (meaning that the intercept of the model is the grand mean of the dependent measure) for preview (*the* = 1, target = −1) and for constraint (high = 1, low = −1). An additional sum-coded factor indicating whether or not the target word was skipped (target fixated = 1, target skipped = −1) was also included in a number of the models to investigate how reading behavior is affected just prior to or following a skipping decision.

We determined the random effects structure of our models by starting with the maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013), but as the maximal models did not converge, we had to reduce the random effects structure. Reported models include random intercepts for subjects and items, random slopes for preview and predictability by subjects (as additive main effects with no interaction; an additional slope for target skipping was included as well for those models in which it was included as a fixed effect), as well as random slopes for preview and constraint by items (again with no interaction).[3] Because it is not clear how to compute the degrees of freedom for LMMs, we do not report p values, and use the two-tailed criterion $|z| \geq 1.96$ to correspond to a significance test at the 0.05 alpha level (t values are interpreted in the same way). Condition means for the dependent measures for the pretarget word are presented in Table 1, and the corresponding LMM

---

[3] The model fit to skipping data on the pretarget word that is reported estimates only random intercepts for subjects and items (no random slopes), because the model including random slopes indicated very high correlations between the random slopes and intercepts. The likelihood ratio test (LRT) indicated that removing the random slopes from the model did not hurt model performance ($p = 1.00$). The model fit to target word go-past time data failed to converge. The predictability term was dropped from the by-items random effects structure, the model was refit, and successfully converged. The same situation occurred when fitting the model to skipping data for the posttarget word, and the resulting model including additive random effects slopes for preview and predictability by subjects and preview by items.

Table 1
Condition Means for the Pretarget Word

		All cases			SFD		GD	
Constraint	Preview	Skip	SFD	GD	Target subsequently fixated	Target subsequently skipped	Target subsequently fixated	Target subsequently skipped
High	target	0.14 (0.03)	209 (5.3)	236 (6.7)	215 (6.0)	203 (7.0)	239 (8.6)	237 (11.2)
High	the	0.13 (0.02)	208 (5.5)	232 (7.4)	217 (11.2)	206 (5.9)	229 (12.3)	237 (9.3)
Low	target	0.13 (0.02)	211 (5.1)	232 (7.2)	212 (6.7)	209 (7.4)	230 (8.0)	244 (18.7)
Low	the	0.15 (0.02)	215 (5.1)	239 (8.1)	235 (14.2)	208 (5.9)	249 (14.5)	237 (8.9)

Note. Standard errors are in parentheses. SFD = single fixation duration; GD = gaze duration.

results are presented in Table 2, for the target word in Tables 3 and 4, and for the posttarget word in Tables 5 and 6.

Pretarget Word

Readers fixated on the pretarget word 86% of the time, and fixation probabilities were not affected by either the nature of the parafoveal preview or the predictability of the target word (all zs $<$ .65). There was a small but reliable effect such that SFDs on the pretarget word (84% of first-pass fixations) prior to a skip were shorter than prior to making a fixation on the target (210 ms vs. 217 ms; $b = 0.02$, $SE = 0.01$, $t = 2.16$), replicating a finding reported by Kliegl and Engbert (2005). This effect did not interact with either preview or constraint. GD showed a similar pattern with shorter fixations prior to skipping (234 ms vs. 239 ms), but this effect did not reach significance ($t = 0.05$).

Target Word Skipping

Target word skipping rates are presented in Table 5 and are presented graphically in Figure 2. Readers skipped over the target word on 56.9% of trials. This behavior was influenced both by the nature of the parafoveal preview as well as the constraint for the target word. Readers skipped previews of *the* more often than correct previews of the target word (0.69 vs. 0.44; $b = 0.62$, $SE = 0.07$, $z = 8.34$). Readers also skipped the word in the target position (either the correct target word or *the*) more often when the target word was highly constrained from the prior context than when it was not (0.60 vs. 0.53; $b = 0.16$, $SE = 0.06$, $z = 2.71$). These two factors did not significantly interact ($z = 0.50$). Interestingly, the lack of an interaction between constraint and preview suggests that constraint affected skipping equally for valid and invalid previews.

Target Word Fixations

Fixation time measures on the target word represent the remaining subset of the data when the target word was fixated. Here, readers' SFDs (41% of first-pass fixations) were shorter following correct target word previews than *the* previews (228 ms vs. 248 ms; $b = -0.03$, $SE = 0.02$, $t = -2.06$). The same pattern was present in GD, with shorter reading times following valid previews (234 ms vs. 257 ms; $b = -0.04$, $SE = 0.02$, $t = -2.31$). This reflects the benefit of having valid parafoveal information available prior to fixation (Rayner, 1975). There was no effect of constraint on GD ($t = 0.07$), and constraint did not interact with preview in this measure ($t = -0.34$). Go-past times were also shorter following correct versus *the* previews (262 ms vs. 288 ms; $b = -0.04$, $SE = 0.02$, $t = -2.53$). In this measure, there was no effect of constraint ($t = -0.20$) and no interaction between constraint and preview ($t = 0.55$).

Readers made considerably more regressions back to the target word when it was initially skipped than when it was fixated (0.85 vs. 0.07; $b = -2.28$; $SE = 0.20$; $z = -11.45$). The effect of parafoveal preview on regression rates was only marginal, with more regressions to the target following *the* previews than correct target previews (0.50 vs. 0.17; $b = 0.28$, $SE = 0.16$, $z = 1.69$). There was no effect of constraint on regression rates ($z = -0.55$),

Table 2
Linear Mixed Effects Model Analyses on the Pretarget Word

	p(Skip)			SFD			GD		
Effect	*b*	*SE*	*z*	*b*	*SE*	*t*	*b*	*SE*	*t*
(Intercept)	−2.39	0.24	**−9.94**	5.32	0.02	**265.24**	5.40	0.03	**207.93**
Fixed effects									
Skip Target	—	—	—	0.02	0.01	**2.16**	0.00	0.01	0.05
Preview	−0.02	0.08	−0.22	0.01	0.01	0.77	0.00	0.01	−0.03
Constraint	−0.01	0.08	−0.13	−0.02	0.01	–1.68	−0.01	0.01	−0.84
Skip Target × Preview	—	—	—	0.01	0.01	1.01	−0.01	0.01	−0.70
Skip Target × Constraint	—	—	—	0.00	0.01	0.05	−0.01	0.01	−0.68
Preview × Constraint	−0.05	0.08	−0.62	−0.01	0.01	−0.88	0.02	0.01	−1.76
Skip Target × Preview × Constraint	—	—	—	−0.01	0.01	–1.44	0.02	0.01	−1.62

Note. Each column represents a model fit to one of the dependent variables. Cells marked in bold represent $|t| \geq 1.96$. SFD = single fixation duration; GD = gaze duration; b = estimated effect size; SE = standard error; t = test statistic (b/SE).

Table 3
Condition Means for the Target Word

		All cases				Regressions in	
Constraint	Preview	Skip	SFD	GD	Go-past	Target initially skipped	Target initially fixated
High	target	0.41 (0.04)	221 (6.3)	226 (6.9)	261 (13.7)	0.74 (0.10)	0.06 (0.02)
High	the	0.72 (0.03)	249 (11.1)	256 (11.2)	289 (18.3)	0.83 (0.05)	0.10 (0.04)
Low	target	0.41 (0.04)	224 (6.8)	230 (7.4)	253 (10.6)	0.66 (0.09)	0.09 (0.02)
Low	the	0.66 (0.03)	242 (11.0)	254 (10.9)	293 (14.6)	0.88 (0.03)	0.11 (0.03)

Note. Standard error of the mean in parentheses. SFD = single fixation duration; GD = gaze duration.

and there were no significant interactions among any of the fixed factors (all zs < 1.41).

Posttarget Word

Condition means for the posttarget word are presented in Table 6. It is clear that skipping over the target word (vs. fixating it) disrupted processing on the posttarget word, with longer fixations following a skip in SFD (61% of first-pass fixations; 232 ms vs. 207 ms), GD (274 ms vs. 229 ms), and go-past time (381 ms vs. 258 ms). Critically, this effect interacted with parafoveal preview in all three measures, with a larger cost in terms of elongated fixation times after skipping *the* previews compared with valid target word previews. Figure 3 provides a depiction of the nature of this interaction in go-past time. In addition to affecting fixation times on the posttarget word, skipping affected regression rates out of this region; readers made more regressions out of the posttarget word when the target word had been skipped versus fixated (0.33 vs. 0.07; $b = -1.20$, $SE = 0.28$, $z = -4.27$). Regressions were also more frequent following *the* previews than target previews (0.36 vs. 0.11; $b = 0.79$, $SE = 0.21$, $z = 3.76$). Finally, readers regressed out of this region more often in sentences containing low-constraint compared with high-constraint target words (0.25 vs. 0.23; $b = -0.38$, $SE = 0.19$, $z = -2.00$).

E-Z Reader 10 Simulations

The primary result of our experiment is striking and unexpected: Readers skip illicit occurrences of the word *the* more often when a different word is highly constrained by the sentence compared with when one is not. Why should constraint affect the processing of an altogether different word? This is unexpected given the wealth of evidence that readers are less likely to skip low-constraint words in sentences designed to elicit a higher constraint alternative (e.g., Balota et al., 1985; Drieghe et al., 2005; Fitzsimmons & Drieghe, 2013; Rayner & Well, 1996), and are just about as likely to skip visually similar nonwords as other low-constraint or even semantically anomalous words (Drieghe et al., 2005). Clearly, words are processed to an advanced degree parafoveally prior to being skipped, perhaps even to the point of full recognition (Gordon, Plummer, & Choi, 2013). Furthermore, when a word is predictable in context, this pattern indicates a procedure wherein the system checks bottom-up parafoveal input against linguistic expectations before skipping. Given the extant data, our results demand explanation.

In order to uncover the mechanisms underlying our effects, we chose to simulate our experiment using the E-Z Reader model of eye movement control during reading (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle et al., 2009). In E-Z Reader, a word can be skipped if the fixated (foveal) word has been identified and the parafoveal word is nearly identified before the system is ready to move the eyes to the next word. This requires that the initial stage of lexical processing (L1, the "familiarity check") complete from the parafovea. The time required to complete L1 for a parafoveal word is influenced by the difficulty of that word in the following ways: (a) a word is "guessed" such that the duration of L1 is set to 0 ms with probability equal to its cloze predictability; and (b) the

Table 4
Linear Mixed Effects Model Analyses for the Target Word

	Duration measures									Binomial measures					
	SFD			GD			Go-past			p(Skip)			Regressions in		
Effect	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*	*b*	*SE*	*z*
(Intercept)	5.40	0.03	**184.04**	5.43	0.03	**192.84**	5.51	0.03	**175.48**	0.30	0.14	**2.15**	−0.53	0.15	**−3.45**
Fixed effects															
Skip target	—	—	—	—	—	—				—	—	—	−2.28	−0.20	**−11.45**
Preview	−0.03	0.02	**−2.06**	0.04	0.02	**2.31**	0.04	0.02	**2.53**	0.62	0.04	**8.34**	0.28	0.17	1.69
Constraint	0.01	0.04	0.56	0.00	0.02	0.07	−0.00	0.02	−0.20	0.16	0.06	**2.71**	−0.09	0.16	−0.55
Skip Target × Preview	—	—	—	—	—	—	—	—	—	—	—	—	−0.20	0.14	−1.368
Skip Target × Constraint	—	—	—	—	—	—	—	—	—	—	—	—	−0.21	0.15	−1.40
Preview × Constraint	−0.01	0.01	−0.67	0.00	0.01	0.34	−0.01	0.02	−0.55	0.03	0.06	0.50	−0.07	0.13	−0.58
Skip Target × Preview × Constraint	—	—	—	—	—	—	—	—	—	—	—	—	0.09	0.15	0.58

Note. Each column represents a model fit to one of the dependent variables. Cells marked in bold represent $|t| \geq 1.96$. SFD = single fixation duration; GD = gaze duration; *b* = estimated effect size; *SE* = standard error; *t* = test statistic (*b/SE*).

Table 5

Condition Means on the Posttarget Word

Constraint	Preview	Skip	SFD	GD	Go-past	Regressions out
				Target skipped		
High	target	0.16 (0.04)	225 (7.67)	253 (10.99)	300 (17.68)	0.12 (0.03)
High	the	0.22 (0.04)	243 (11.40)	290 (12.24)	428 (21.62)	0.44 (0.04)
Low	target	0.11 (0.03)	221 (9.55)	287 (19.15)	327 (20.88)	0.15 (0.04)
Low	the	0.18 (0.03)	238 (10.08)	285 (11.51)	437 (20.52)	0.44 (0.05)
				Target fixated		
High	target	0.57 (0.05)	205 (8.18)	217 (8.85)	236 (14.90)	0.01 (0.01)
High	the	0.44 (0.06)	212 (12.87)	230 (13.53)	273 (30.30)	0.07 (0.04)
Low	target	0.54 (0.05)	219 (9.28)	240 (9.80)	267 (15.58)	0.08 (0.04)
Low	the	0.51 (0.05)	198 (10.87)	214 (13.65)	254 (19.95)	0.12 (0.05)

Note. SFD = single fixation duration; GD = gaze duration.

duration of L1 is a function of word frequency, cloze probability, and mean eccentricity from the point of fixation. In terms of E-Z Reader, the effect of constraint on skipping any words in the observed data may be accounted for by the first mechanism, and skipping parafoveal words that look like *the* by the second (due to its very high frequency).

The question is how to incorporate cloze probability to model the joint effects of constraint and parafoveal preview on skipping. Our simulations were conducted to test two possibilities with regard to this issue: (a) *the*-skipping is driven primarily by its frequency of occurrence (i.e., there is no systematic effect of target word constraint on *the*-skipping); and (b) *the*-skipping is affected by the cloze probability of the target word (e.g., *cut*) constrained by the preceding sentence context. If E-Z Reader predicts that skipping of apparent *the*-previews is affected by the degree of constraint for the target word (e.g., *cut*), with increased *the*-skipping in sentences with high versus low constraint, it would suggest that the contribution of constraint to skipping is indeed independent from that of the ease of identifying the parafoveal word.

We conducted the following simulations using the E-Z Reader 10 model[4] (Reichle et al., 2009). Our approach was to use the model to estimate the rate at which readers would skip the preview word, given its length, frequency of occurrence, and cloze probability. Frequency values for all words used in the experiment were extracted from the Francis and Kučera (1982) corpus.[5] For convenience we assumed a cloze probability of 0 for all words other than those in the target position (see Staub, 2011). For each simulation, we used the default model parameters (as provided by Reichle, Pollatsek, & Rayner, 2012) and predicted data for 10,000 statistical subjects reading the sentences in all four conditions.

Simulation 1. In this simulation, we asked whether the very high frequency of the word *the* alone (69,971 counts per million compared with an average of 18 counts per million for the target words in our study) accounts for the rate at which it was skipped in our experiment. To do so, cloze probability for *the* in sentences with illicit previews was assumed to be 0, reflecting the fact that no subject entered *the* in the cloze norming study (for sentences with valid previews, the cloze probabilities equaled those collected in the norming study). Mean simulated skipping rates are presented in Figure 4. For the sentences with correct target word previews, skipping rates were greater for sentences with high- compared with low-constraint words (0.42 vs. 0.22). The model estimated that skipping *the*, however, was unaffected by constraint (0.23 for both high- and low-constraint conditions), and would actually be skipped less often than the correct target in constraining sentences (0.23 vs. 0.42).

Simulation 2. In the second simulation, we asked whether the frequency of the parafoveal word and the cloze probability for the expected word have additive effects on skipping. For both preview conditions, we assumed that the cloze probability for the preview word was that of the correct target word (e.g., for a target *die* with cloze 0.8, we assigned 0.8 cloze for *the* in the false preview condition). The assumption here is that reader expectations are entirely separate from the input received from the parafoveal word. The mean obtained skipping rates for the four different conditions are presented in Figure 5. The pattern here is quite clear: The model simulation replicated the additive effects on skipping in the observed data, with greater skipping for constrained words (0.44 for high- vs. 0.23 for low-constraint words) and slightly more skipping for *the* compared with valid target previews (0.35 vs. 0.33). Importantly, E-Z Reader does not predict an interaction between preview and constraint, similar to the observed data. The results from this simulation are plotted alongside the observed skipping data in Figure 6. Although the model appears to overestimate the role of constraint and underestimate the role of frequency in skipping these words, the basic pattern was successfully reproduced.

---

[4] E-Z Reader 10 is available as a Java program at http://www.erikdreichle.com/ezreader.html

[5] Four of the target words do not have corresponding entries in the Francis and Kučera (1982) corpus; these items were not included in the simulations. Frequencies were not available for 11 words (3%); these words were assigned the median frequency of words used in the experiment (937 counts per million). We note that this was done simply in order for the model to run simulations to completion. We did not replace missing frequency counts for the target words with the value that we used to replace other missing frequency counts because this value is much higher than the median frequency of the target words (18), and there is good reason to believe that the actual frequencies of these words would be quite low, given that they do not appear in the corpus.

Table 6

Linear Mixed Effects Model Analyses for the Posttarget Word

	Duration measures									Binomial measures					
	SFD			GD			Go-past time			p(Skip)			Regressions out		
Effect	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*	*b*	*SE*	*z*
(Intercept)	5.33	0.02	**349.31**	5.24	0.02	**277.58**	5.59	0.03	**205.49**	–1.13	0.27	**−4.17**	–2.37	0.31	**−7.77**
Fixed effects															
Skip target	−0.06	0.01	**−4.93**	−0.08	0.01	**−6.28**	−0.17	0.02	**−11.17**	1.16	0.09	**12.75**	–1.20	0.28	**−4.27**
Preview	0.01	0.01	0.65	0.02	0.01	1.13	0.09	0.02	**5.75**	−0.07	0.10	−0.71	0.79	0.21	**3.76**
Constraint	0.00	0.01	0.33	−0.01	0.01	−0.78	−0.03	0.02	−1.77	0.10	0.08	1.29	−0.38	0.19	**−2.00**
Skip target * preview	−0.03	0.01	**−2.66**	−0.03	0.01	**−2.36**	−0.07	0.02	**−4.78**	−0.17	0.09	−1.95	−0.06	0.20	−0.30
Skip target * constraint	−0.01	0.01	−0.78	−0.01	0.01	−0.41	−0.01	0.02	−0.96	−0.13	0.08	−1.61	−0.33	0.19	−1.75
Preview * constraint	0.00	0.01	0.31	0.02	0.01	1.39	0.01	0.02	0.73	−0.08	0.08	–1.04	0.16	0.18	0.91
Skip target * preview * constraint	0.00	0.01	0.39	0.00	0.01	0.02	0.00	0.02	0.15	−0.02	0.08	−0.20	0.09	0.17	0.53

Note. Each column represents a model fit to one of the dependent variables. Cells marked in bold represent | *t* | = 1.96. SFD = single fixation duration; GD = gaze duration; *b* = estimated effect size; *SE* = standard error; *t* or *z* = test statistic (*b*/SE).

These simulations using E-Z Reader 10 demonstrate that our pattern of skipping data can be explained if we assume that, prior to skipping *the*, readers took into account the predictability of the word that they expected to see in that position. Importantly, the model predicted this pattern without making any assumptions about the syntactic fit between the preview word and the preceding material. That E-Z Reader predicts this pattern makes sense when we consider that the model permits a word to be "guessed" based solely on its predictability. Note, though, that in highly constraining sentences, the model still predicts that *the* will be skipped slightly more often than the correct target words (0.45 vs. 0.43). The second route of the L1 computation, which assumes that when a word is not "guessed," its identification time is a function of its frequency and cloze probability, accounts for this difference.

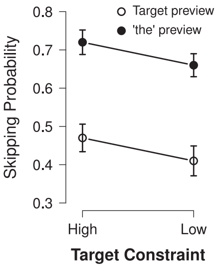

Figure 2. Target word skipping. Error bars represent standard error of the mean.

General Discussion

We manipulated target word constraint and parafoveal preview using the gaze-contingent boundary paradigm. The preview of the target word was either a correct preview of a three-character verb or an invalid preview of *the*, presented in sentences in which the correct target word was either predictable or unpredictable. Our goal was to determine whether contextual constraint modulates skipping infelicitous occurrences of *the*, following prior research that has shown that syntactic constraints do not modulate skipping decisions (Angele et al., 2014; Angele & Rayner, 2013). If readers primarily make use of the sentence context when it highly constrains the target word, we expected an interaction between constraint and preview, such that readers would be more likely to catch the violation and decide not to skip the false *the* previews in high- versus low-constraint sentences.

In our study, readers were more likely to skip parafoveal previews of *the* than valid previews of the target words (replicating Angele & Rayner, 2013). This is further evidence that readers are more likely to skip high-frequency words, regardless of whether the word fits in the context of the sentence (consistent with Angele et al., 2014). Readers also skipped the word in the target position (either the target word or *the*) more often in sentences in which the target word was highly constrained compared with when it was unconstrained. Thus, even when the sentence context sets a high expectation for a specific word, readers do not seem to integrate parafoveal content with higher level linguistic information, and, as a consequence, are still more likely to skip a parafoveal word that looks like *the* than one that looks like the expected word. This

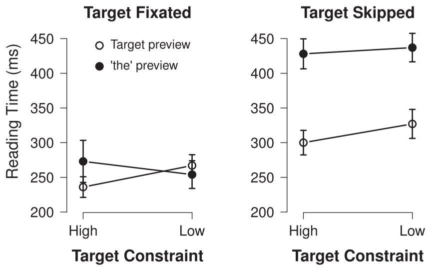

Figure 3. Go-past time on the posttarget word. Error bars represent standard error of the mean.

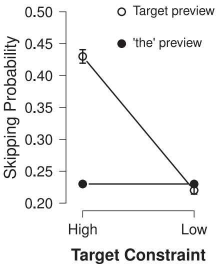

Figure 4. Simulation 1: Target skipping probabilities by condition.

suggests that the parafoveal preview is processed to an advanced degree (to the point that its frequency, or, at the very least, its orthographic familiarity, affects reading behavior), but not to the extent that its fit within the context is assessed. However, it is evident that the false *the* previews disrupted later processing, especially when they were skipped: Although readers spent longer fixating the posttarget region when it was skipped compared with when it was fixated, this difference was even greater when the preview word was *the*.

Given the novelty of the effect that our experimental manipulation had on skipping behavior, we sought to model this pattern of data using the E-Z Reader model. We conducted two simulations of our experiment using assumptions about the cloze probabilities for the word in the target position: (a) the cloze probability was set to 0 in the *the*-preview condition (corresponding to the actual cloze probability of *the* in that location; in the target preview condition, the cloze probability was the actual cloze probability of the target word); and (b) the cloze probability was always set to the cloze probability for the target word in the sentence, regardless of the preview condition (e.g., when *the* was the preview for *cut*, the cloze probability for *cut* was used). Simulation 2, but not Simulation 1, successfully replicated our pattern of skipping data. Thus, it appears that serial attention shift models of eye movement control are well equipped to handle our results if we assume that reader expectations are computed separately from the visual information they have in front of them. Indeed, this assumption is built in to the model as a component of the function that handles the initial stage of lexical processing. Although it is possible that parallel processing models such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) would also predict this pattern, quantitative simulations like those reported here would need to be conducted. Because fixation probabilities in SWIFT are modulated by the amount of parafoveal processing that a word has received (which is influenced both by a word's frequency and cloze probability), it seems possible that it could replicate this pattern as well.

The results of our study further contribute to issues regarding the interaction of parafoveal information and higher level cognitive processing. For example, the types of context that can influence skipping are limited to those that are highly predictable. This could reflect a limitation of the extent to which words are lexically processed prior to being fixated (i.e., prior to skipping in E-Z Reader, the reader makes a hedged bet that recognition is imminent, but the word is not yet fully lexically accessed), and that very early lexical processes are dissociated from later processes that involve syntactic integration (Staub, 2011). Predictability is unique, then, in that it appears to affect the same early word recognition processes as frequency unlike any other form of context. However, like other effects from the context on fixation durations (e.g., syntactic fit; Staub, 2011), the influences of predictability and frequency on skipping appear to be independent.

It is notable that regardless of the level of constraint for the target word, it was very often regressed back to it after having been skipped (regression rate 0.85 following a skip vs. 0.08 following fixation). This seems odd in light of our interpretation of the role of constraint: If constraint buys the reader confidence in the identity of an upcoming word, then they should be less likely to regress back to it when it is highly constrained (even if it is initially skipped). One possibility is that because the target words were typically the main verb of the sentence, adequate comprehension of the sentence necessitates fully processing it. Another possibility is that skipping short words is generally a "risky" (error-prone) behavior. Indeed, this suggestion has been made before (though not exclusively for short words). Rayner, Reichle, Stroud, Williams, and Pollatsek (2006) found that older readers skipped more words than younger readers, but as a consequence made more regressions. They modeled this pattern using E-Z Reader by modulating the parameter corresponding to the rate at which words are "guessed" from the parafovea to respond not just to cloze probability but also to frequency (such that high-frequency words can also be "guessed" in this manner). This could very well be exactly what happened in the present study: Readers skipped short words because they were often "guessed" based on their high frequency. However, if the result of this process is less time devoted to parafoveal processing, then it could be that these words are not in fact processed very deeply and require additional viewing time—hence, the increased regressions.

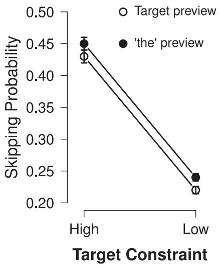

Figure 5. Simulation 2: Target skipping probabilities by condition.

It is important for future research to establish whether or not the effects reported here, and the general finding of greater skipping rates for higher frequency word previews (as in Angele et al., 2014; Angele & Rayner, 2013), are unique to very short (e.g., three-letter) words or generalizes to longer words. It may be that lexical properties like word frequency play a larger role in skip-

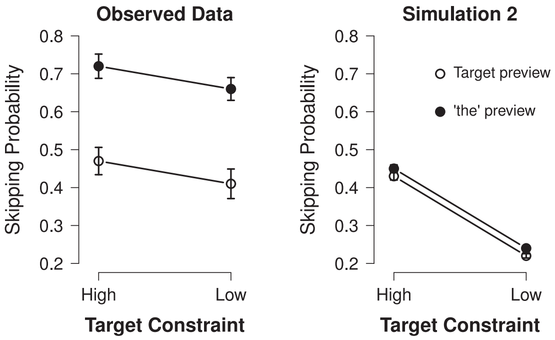

Figure 6. Skipping probabilities for the target words from the observed data (left panel) and from Simulation 2 (right panel).

ping shorter (e.g., three-letter) compared with longer (e.g., five- to six-letter) words. There is some direct support for this idea within the E-Z Reader model, because the rate of parafoveal processing is limited by the mean eccentricity of the parafoveal word from the point of fixation. Because longer words are, on average, further away from the point of fixation than shorter words, parafoveal processing will be slower. Though this accounts for why long words are unlikely to be skipped in general, it may also predict different effects of parafoveal preview frequency on skipping rates as length increases.

In summary, both contextual constraint and parafoveal information influence word-skipping behavior during reading independently. There seems to be no cross-talk between the two sources of information, even when there is a conflict between them. When the string in parafoveal vision is more easily identifiable than a highly predictable word of the same length that should appear in that position, readers do not appear to detect the mismatch between these sources of information. E-Z Reader predicts the same pattern if we assume that readers consider the cloze probability of the expected word in lieu of the word actually available in parafoveal vision. If they do detect the mismatch, it is likely at a stage of processing at which the decision to skip the upcoming word cannot be canceled. These results are important for our understanding of the extent to which higher level cognitive processing interacts with other aspects of lexical processing in parafoveal vision when making the decision to skip.

References

Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1181–1203. http://dx.doi.org/10.1037/a0036396

Angele, B., & Rayner, K. (2013). Processing the in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 649–662. http://dx.doi.org/10.1037/a0029294

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. http://dx.doi.org/10.1016/j.jml.2007.12.005

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390. http://dx.doi.org/10.1016/0010-0285(85)90013-1

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-7. Retrieved from http://CRAN.R-project.org/package=lme4

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1552–1560. http://dx.doi.org/10.1037/a0013017

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 954–969. http://dx.doi.org/10.1037/0096-1523.31.5.954

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning & Verbal Behavior, 20,* 641–655. http://dx.doi.org/10.1016/S0022-5371(81)90220-6

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777–813. http://dx.doi.org/10.1037/0033-295X.112.4.777

Fitzsimmons, G., & Drieghe, D. (2013). How fast can predictability influence word skipping during reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1054–1063. http://dx.doi.org/10.1037/a0030909

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghton Mifflin.

Gautier, V., O'Regan, J. K., & Le Gargasson, J. F. (2000). "The-skipping" revisited in French: Programming saccades to skip the article "les". *Vision Research, 40,* 2517–2531. http://dx.doi.org/10.1016/S0042-6989(00)00089-4

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General, 140,* 186–209. http://dx.doi.org/10.1037/a0022256

Gordon, P. C., Plummer, P., & Choi, W. (2013). See before you jump: Full recognition of parafoveal words precedes skips during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1032–1046. http://dx.doi.org/10.1037/a0030432

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review, 12,* 132–138. http://dx.doi.org/10.3758/BF03196358

O'Regan, J. K. (1979). Eye guidance in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics, 25,* 501–509. http://dx.doi.org/10.3758/BF03213829

O'Regan, J. K. (1980). The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception & Psychophysics, 28,* 112–117. http://dx.doi.org/10.3758/BF03204335

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 817–833. http://dx.doi.org/10.1037/0096-1523.8.6.817

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81. http://dx.doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422. http://dx.doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). The Thirty-Fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 62,* 1457–1506. http://dx.doi.org/10.1080/17470210902816461

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129–1144. http://dx.doi.org/10.1016/S0042-6989(97)00274-5

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research, 16,* 829–837. http://dx.doi.org/10.1016/0042-6989(76)90143-7

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21,* 448–465. http://dx.doi.org/10.1037/0882-7974.21.3.448

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200. http://dx.doi.org/10.1037/0096-1523.22.5.1188

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 514–528. http://dx.doi.org/10.1037/a0020990

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504–509. http://dx.doi.org/10.3758/BF03214555

R Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

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. http://dx.doi.org/10.1037/0033-295X.105.1.125

Reichle, E. D., Pollatsek, A., & Rayner, K. (2012). Using E-Z Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye-mind link. *Psychological Review, 119,* 155–185. http://dx.doi.org/10.1037/a0026473

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,* 1–21. http://dx.doi.org/10.3758/PBR.16.1.1

Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General, 140,* 407–433.

(*Appendix follows*)

Appendix

Sentences Used in the Experiment as Well as Corresponding Cloze Probabilities and Francis and Kučera (1982) Frequency Values (in Counts per Million) for the Target Words as Submitted to E-Z Reader 10

Sentence frame	Constraint	Sentence	Cloze	Frequency
1	High	If you are shot in the heart you will *die* very quickly unless you are treated.	0.8	73
1	Low	If you are shot in the toe you will *die* very quickly unless you are treated.	0.0	73
2	High	To win money at the horse races you must *bet* your money on the best horse.	0.8	20
2	Low	To win the most at the event you must *bet* your money very wisely.	0.0	20
3	High	If the price isn't listed on the menu you should *ask* your server when he comes.	1.0	128
3	Low	If the price isn't listed on the website you should *ask* your server when he comes.	0.0	128
4	High	I got a deep gash when a dog violently *bit* into my thigh.	0.8	101
4	Low	I got a deep gash when a man violently *bit* into my thigh.	0.0	101
5	High	There will be more old people as the population is set to rapidly *age* in coming years.	0.5	227
5	Low	There will be more poor people as the population is set to rapidly *age* in coming years.	0.0	227
6	High	To advance in the tournament you must *win* each game that you play.	0.8	55
6	Low	To advance in the system you must *win* each game that you play.	0.0	55
7	High	To compute the sum you must *add* all of the numbers together.	0.8	88
7	Low	To compute the result you must *add* all of the numbers together.	0.2	88
8	High	Jane's husband isn't truthful and will often *lie* about trivial things.	0.9	59
8	Low	Jane's husband isn't helpful and will often *lie* about trivial things.	0.1	59
9	High	John found his old suit but it no longer *fit* properly at all.	1.0	75
9	Low	John found his old watch but it no longer *fit* properly at all.	0.0	75
10	High	Jane used the scissors to carefully *cut* scraps of paper into neat shapes.	0.8	192
10	Low	Jane used the machine to carefully *cut* scraps of paper into neat shapes.	0.2	192
11	High	After a shower I grab a towel to quickly *dry* myself off before putting on clothes.	0.9	68
11	Low	After a jog I take a minute to quickly *dry* myself off before going back inside.	0.0	68
12	High	When your baby is hungry she will loudly *cry* until you come to feed her.	0.9	48
12	Low	When you cat is hungry she will loudly *cry* until you come to feed her.	0.0	48
13	High	When you are fasting you cannot *eat* for certain periods of time.	1.0	61
13	Low	When you are working you cannot *eat* for certain periods of time.	0.0	61
14	High	When he wanted to say yes, he would silently *nod* instead of saying "yes".	0.9	12
14	Low	When he wanted to say no, he would silently *nod* instead of shaking his head.	0.1	12
15	High	The author was not sure how his novel should *end* without upsetting his readers.	0.6	410
15	Low	The author was not sure how his day should *end* without his wife around.	0.2	410
16	High	When the lawn got too long my dad asked if I would *mow* it over the weekend.	0.8	*
16	Low	When it got too long my dad asked if I would *mow* our lawn over the weekend.	0.0	*
17	High	My pants had holes in them so I asked my mother to kindly *sew* them back together.	0.5	6
17	Low	My shoes had holes in them so I asked my mother to kindly *sew* them back together.	0.1	6
18	High	Using the telescope Jennifer *saw* galaxies and constellations in the sky.	0.5	352
18	Low	Using the device Jennifer *saw* galaxies and constellations in the sky.	0.0	352
19	High	They are going to an alpine resort where they will *ski* on some very steep slopes.	0.8	5
19	Low	They are going to a desert resort where they will *ski* some unique outdoor slopes.	0.0	5
20	High	I like to jog and will go for a short *run* after work every day.	0.6	212
20	Low	I like to swim and will go for a short *run* after work too.	0.1	212
21	High	Each month your landlord expects you to promptly *pay* your rent and bills.	1.0	172
21	Low	Each month your friend expects you to promptly *pay* your rent and bills.	0.4	172
22	High	To change your hair color you can have a stylist *dye* it any color you like.	0.6	*
22	Low	To change your hair style you can have a stylist *dye* it any color you like.	0.0	*
23	High	In a theater class you will learn to professionally *act* in a variety of situations.	0.7	283
23	Low	In the evening class you will learn to professionally *act* in a variety of situations.	0.0	283
24	High	As a kid I wore Velcro shoes until I could *tie* laces into a neat bow.	0.8	23
24	Low	As a kid I wore ugly shoes until I could *tie* laces into a neat bow.	0.0	23
25	High	If you are happy with your waiter you should *tip* twenty percent of the bill.	1.0	22
25	Low	If you are happy with your experience you should *tip* twenty percent of the bill.	0.0	22
26	High	My dog will lift his leg and quickly *pee* on every hydrant he sees.	0.6	3
26	Low	My cat will lift his tail and quickly *pee* on the floor when he is scared.	0.1	3
27	High	If you stick too much paper in the printer it will *jam* until it is cleared.	0.9	6
27	Low	If you put too much faith in the machine it will *jam* and break down.	0.0	6
28	High	With those strong oars, he could *row* his boat very far.	0.6	35

(*Appendix continues*)

Appendix (*continued*)

Sentence frame	Constraint	Sentence	Cloze	Frequency
28	Low	With those strong arms, he could *row* his boat very far.	0.0	35
29	High	A mirror next to a steamy shower will likely *fog* unless you run the fan.	1.0	25
29	Low	A door next to a running shower will likely *fog* unless you run the fan.	0.3	25
30	High	The girl saw the furry puppy and asked if she could *pet* her and give her a treat.	0.4	8
30	Low	The girl saw the furry spider and asked if she could *pet* her and give her a treat.	0.1	8
31	High	The impatient cows will loudly *moo* when they want to be milked.	0.9	*
31	Low	The impatient animals will loudly *moo* when they want to be milked.	0.0	*
32	High	I'm tired of renting, and would rather *own* a nice house or condo now.	0.7	772
32	Low	I'm tired of looking, and would rather *own* a nice house or condo now.	0.0	772
33	High	After the concert, the performers humbly *bow* to their audience.	0.7	15
33	Low	After the concert, the guests humbly *bow* to their Queen.	0.0	15
34	High	You should use the shovel to carefully *dig* your hole to plant the tree.	1.0	10
34	Low	You should use the tool to carefully *dig* your hole to plant the tree.	0.0	10
35	High	To clean the floor, fill a bucket and quickly *mop* until there are no spots.	0.6	3
35	Low	To clean the area, take the supplies and quickly *mop* until there are no spots.	0.0	3
36	High	If left out, apples and avocados will quickly *rot* and turn brown.	0.5	8
36	Low	If left out, certain food items will quickly *rot* and turn brown.	0.0	8
37	High	It is fun to take bubble wrap and joyously *pop* all of the bubbles on it.	0.9	8
37	Low	It is fun to take bubble gum and joyously *pop* all of the bubbles you make.	0.0	8
38	High	Instead of a kiss you could *hug* your children when they leave.	0.9	3
38	Low	Instead of a wave you could *hug* your children when they leave.	0.0	3
39	High	To make the doughnuts, heat some oil and quickly *fry* until they are golden-brown.	0.4	*
39	Low	To make the dish, heat your surface and quickly *fry* until the food is golden-brown.	0.1	*
40	High	An angry audience will often *boo* if the performer makes a racist comment.	0.8	1
40	Low	An engaged audience will often *boo* if the performer makes a racist comment.	0.0	1

Note. The target words are italicized. Items marked with an asterisk did not have frequency values in Francis and Kučera (1982) and were excluded from the E-Z Reader simulations.

Received March 4, 2015
Revision received April 16, 2015
Accepted April 16, 2015 ■

Chapter 2, in full, is a reprint of the material as it appears in Skipping syntactically illegal *the*-previews: The role of predictability. Abbott, Angele, Ahn, & Rayner (2015). The dissertation author was the primary investigator and author of this paper.

Chapter 3

Preferential skipping of function words during reading

3.1 Introduction

Readers do not fixate every word in a sentence, but occasionally skip over words, especially words that are short, highly frequent, or predictable from the sentence context (Rayner, 1998, 2009). An important question is what words are skipped over and what are the underlying mechanisms, as several theories of eye movements in reading directly assume that skipping reflects the identification of a word from the parafovea (as opposed to, e.g., an oculomotor tendency to move the eyes forward at a constant distance). Researchers have developed highly sophisticated computational models like E-Z Reader (Reichle, Pollatsek, Fisher, & Rayner, 1998) and SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) to explain these processes, but many questions remain unanswered. Establishing the relevant factors involved in skipping will inform these models by providing information as to the depth of processing that occurs prior to skipping. This article examines whether function words are skipped preferentially to words of other syntactic classes (Abbott, Angele, Ahn, & Rayner, 2015; Angele & Rayner, 2013; but see Angele, Laishley, Rayner, & Liversedge, 2014), or whether *the* is a special case. To anticipate the results, we find strong evidence that readers prefer to skip function words regardless of their length.

Several studies have found that readers skip over articles more often than short open-class (content) words (Abbott et al., 2015; Angele & Rayner, 2013; Drieghe, Pollatsek, Staub, & Rayner, 2008; Gautier, O'Regan, & Le Gargasson, 2000; O'Regan, 1979). O'Regan (1979) and Gautier et al. (2000) both found, for example, that readers skipped the French article *les* more often than other three-letter words. This may indicate that articles are skipped over because they are easy to process, but it may also indicate that function words are in some way processed differently.

Angele and Rayner (2013) used the boundary paradigm (Rayner, 1975) to present readers with sentences where, prior to their eyes' crossing an invisible boundary, they had a correct parafoveal preview of a three-letter word) or an incorrect preview of the article *the*. Importantly, *the* was anomalous in that position of the sentence, as the target words were always verbs (e.g., "She was sure she would *ace* all the tests"). Readers skipped *the*-previews at a much higher rate than valid previews. Disruption was observed downstream with longer *go-past time* (sum of all fixations on a word, including regressions to previous words, before moving to the right) and increased regressions following *the*-previews. These results appear to support the interpretation that *the* is skipped because it is so frequent, although it does not rule out other contributions.

Angele et al. (2014) reported another experiment using a similar design, using invalid parafoveal previews that were higher-frequency (HF) three-letter open-class (content) words instead of *the* (i.e., The increasingly *dog* light; where *dog* is the invalid preview for *dim*). Once more, the invalid previews were infelicitous in context. The findings were the same: HF previews were skipped more often than previews of lower-frequency words, regardless of sentential fit. However, the size of the skipping effect was smaller than the *the*-skipping effect. These results also support the view that word frequency is a critical factor to skipping. However, the smaller skipping effect for content words could be important, as it may reflect a more general tendency to skip over function words, perhaps because they are very easy to process, whether or not they fit in the context.

Here we address this issue in two experiments. We do so by taking into consideration word length, another critical factor to skipping (e.g., Rayner, Slattery, Drieghe, & Liversedge, 2011). Word length plays into skipping for a few different reasons. Short words may be skipped more often due to oculomotor constraints. For example, some short words are skipped because saccades overshoot their intended target (i.e., landing on a subsequent longer word; Nuthmann, Engbert, & Kliegl, 2005). Alternatively, short words may be skipped because they are processed more rapidly from the parafovea than are longer words due to the lower average eccentricity of each letter (as is assumed by models like E-Z Reader; Reichle et al., 1998) .

We tested the effect that word length has on skipping both content word (Experiment 1) and function word (Experiment 2) previews. As in Angele et al. (2014) we manipulated parafoveal preview for the target word so that readers had either valid preview of the target word or an incorrect preview of a HF infelicitous word (either a length-matched word of a different syntactic class, Experiment 1, or a function word, Experiment 2). This design thus allows us to test two questions: (1) Does word frequency affect skipping similarly for shorter versus longer words and; (2) Is function word skipping unique to *the* or does it generalize to, e.g., *that* and *there*?

If short words are skipped over more often than longer words as a result of parafoveal processing, then we would expect to observe effects of word length and word frequency, with longer words skipped less often than shorter words, and less frequent words skipped less often than more frequent words, because both factors the rate at which those words are processed from the parafovea. If oculomotor constraints like saccadic range error have a large effect on skipping in addition to the effect of parafoveal processing (i.e., some words are skipped over accidentally due to errors in execution), then we may observe an interaction between length and frequency. This may take on a couple of different forms. If very short (e.g., three-letter) words are accidentally skipped over more often than longer words, the effect that frequency has on skipping may be smaller for shorter versus longer words. If the effect of visual acuity more dramatically reduces the rate at which words are processed parafoveally, we may expect the opposite pattern, that frequency effects would be smaller for skipping longer versus shorter words.

If it is the case that readers do not use information from the syntactic context to guide skipping, we expect that the infelicitous previews will not reduce skipping, but will have downstream effects on fixation times and regression rates. If function words are skipped preferentially relative to other high-frequency words, then we expect much smaller skipping effects for content versus function word previews. If the context is not processed prior to skipping, we expect that contextual fit will not block the skipping of infelicitous function words.

3.2 General Method

Materials

180 sentences comprised six conditions (3 length X 2 parafoveal preview). 60 sentences containing three-letter target words and corresponding HF previews were adopted with permission from Angele et al. (2014). We selected four and five-letter target words from the HAL corpus (available online as a portion of the English Lexicon Project; Balota et al., 2007), taking care to match log-transformed word frequency

Table 3.1: Mean natural log-transformed target word frequencies by condition and experiment as obtained from ELP.

Preview	Three	Four	Five
Valid preview	8.50	8.72	8.62
HF content (Experiment 1)	12.22	10.37	10.34
Function word (Experiment 2)	16.96	15.48	14.09

across conditions. For the invalid preview condition, we selected four and five-letter HF nouns (the targets were always verbs). The illicit previews in the three-letter condition were higher in frequency than the other two conditions (see Table 3.1), but this is somewhat unavoidable given the ceiling on frequency for four and five-letter words. The only differences between Experiment 1 and Experiment 2 were the preview words; in Experiment 2 the preview words were always *the*, *that*, and *there*.

We wrote sentences containing the four and five-letter target words that were similar in style to Angele et al. (2014). The target word was in position 6.5 on average, and did not differ by condition. Pre-target words were of sufficient length to avoid skipping, and were on average 5.9 letters long (with no difference by condition). Sentences were written to be neutral prior to the target word (i.e., no direct semantic relationship with the preceding material), and to render the preview word syntactically infelicitous. An example sentence is: “We need to quickly *boil*/*beer* some water before adding in the pasta” (target word and infelicitous preview in italics; the full list of sentences is in the Appendix). Parafoveal preview was controlled using the boundary paradigm by programming an invisible boundary into the space preceding the target word. Once the eyetracker detected subjects’ eyes crossing the boundary, the preview was always replaced with the correct target word.

To assess target word constraint, we presented the sentences up to the target word, with the target and remaining material deleted, to ten workers recruited through Amazon Mechanical Turk. They were asked to provide the next word they thought best continued each fragment. The target words were weakly constrained (mean cloze probability = .06); mean cloze probability differed according to length, with higher cloze for three-letter targets versus four or five-letter targets (mean three-letter = .14, mean four-letter = .02, mean five-letter = .02).

Apparatus

An SR Research EyeLink 1000 eyetracker (SR Research, Toronto, Ontario, Canada) was used to sample subjects' eye position at 1000 Hz. Sentences were displayed in 14-pt Courier New font on a Hewlett Packard p1230 CRT monitor at 150 Hz. Viewing distance was approximately 60 cm, with approximately three characters per degree of visual angle. Viewing was binocular, but only the right eye was recorded.

3.3 Experiment 3.1

3.3.1 Method

Participants

Forty-eight University of California, San Diego, students participated for course credit. All participants were native English speakers with normal or corrected-to-normal vision (using soft contacts or glasses), and were naïve to the purpose of the study.

Procedure

Participants first completed a 3-point calibration procedure. Then participants read each of the 180 sentences (60 for each length target word; e.g., "The lamp cast an increasingly *dim* light"; target word in italics), and comprehension was assessed using yes/no questions following one third of the sentences. We used the gaze-contingent boundary paradigm (Rayner, 1975) to control the parafoveal preview of the target word, presenting either a valid preview (*dim*) or an incorrect preview of an infelicitous word (*dog*). Once the tracker detected a saccade crossing the invisible boundary, the display changed to the correct target word within 4 ms on average (range 0-7 ms).

3.3.2 Results and Discussion

Prior to analysis, we deleted 6.3% of trials with blinks or track loss on the target word or on adjacent fixations, and 12.6% of trials with display changes that completed after fixation onset on the target word. Fixations shorter than 80 ms and within one character of an adjacent fixation were combined, and remaining fixations shorter than 80 ms or longer than 1000 ms were deleted. Comprehension question accuracy was high (mean = 91%, SD = 3.5%). We excluded from our analysis trials where participants made a regressive eye movement away from the pre-target word during first-pass reading (7.4% of trials).

Our analysis focused on skipping probability for the target word in the different conditions. To determine effects of preview benefit and downstream effects of the preview condition, we report fixation time measures for the target word and for the post-target word (*gaze duration*, GZD, the sum of all fixations made on a word during first-pass reading, including refixations; go-past time, GPT), as well as regression probability (RO) from the target or post-target region.

We report generalized linear mixed-effects models (GLMMs) using a logit link for binomial measures (skipping and RO). For duration measures (GZD and GPT), we report linear mixed-effects models (LMMs). GLMMs and LMMs were fit using the *glmer* and *lmer* functions, respectively, from the *lme4* package (Version 1.1-8; Bates, Maechler, Bolker, & Walker, 2014) in R (Version 3.2.1; R Core Team, 2015). We included the preview effect as a categorical factor using sum-to-zero contrasts (Higher-

Frequency = 1, Target = –1), and length as a centered linear predictor (when we fit models including length as a three-level categorical factor the results were qualitatively the same). For post-target models we included an additional factor indicating whether the target word was skipped or fixated (Fixated = 1, Skipped = –1). Following Barr, Levy, Scheepers, and Tily (2013), we fit our models using the maximal random-effects structure with respect to the additive effects of length and preview. We included random subject intercepts and slopes for the additive effects of preview and length by subjects, and random item intercepts and slopes for the effects of preview. While this random-effects structure is not maximal with respect to the interaction between the two, the interaction is never close to significance even without the by-subjects random slope for that interaction. As it is difficult to determine the correct degrees of freedom for LMM models, we do not report p-values, and use $|z|$ or $|t| = 2.00$ as our .05-level cutoff for significance. Table 3.2 presents subject means for the different dependent measures and conditions for the target word, with corresponding model results in Table 3.3. Subject means for the post-target word are in Table 3.4, with corresponding model results in Table 3.5.

Target Word

Mean target word skipping rate was 31.6%. Word length significantly influenced skipping probability, with an increase in skipping from three to five-letter words (b = -.76, SE = .06, z = -12.56). The preview effect was not significant (z = -.49), and there was no significant interaction ($z < 1.45$), although three-letter HF infelicitous previews were skipped a bit more often than correct previews (M = .49 vs. M = .46).

When the target word was fixated, GZD was longer following infelicitous previews than correct previews (b = 13.09, SE = 1.79, t = 7.33). There was an effect of length in GZD wherein fixations were shorter for shorter versus longer words (b = -4.65,

Table 3.2: Target word subject means and standard errors by condition: Experiment 1.

Length	Preview	Skipping	GZD	GPT	RO
3	HF	0.49 (0.03)	267.79 (7.84)	306.61 (11.94)	0.11 (0.02)
3	Target	0.46 (0.03)	236.15 (6.51)	257.62 (8.04)	0.07 (0.01)
4	HF	0.32 (0.03)	260.67 (8.22)	290.31 (10.57)	0.08 (0.01)
4	Target	0.32 (0.03)	233.46 (5.88)	256.43 (7.39)	0.06 (0.01)
5	HF	0.18 (0.03)	257.02 (7.07)	292.52 (10.40)	0.12 (0.02)
5	Target	0.19 (0.03)	231.08 (6.23)	251.62 (9.82)	0.05 (0.01)

SE = 2.01, t = -2.31). GPT was longer following invalid versus valid previews (b = 20.51, SE = 2.95, t = 6.95), and was accompanied by more regressions from the target following invalid vs. valid previews (b = .42, SE = .11, z = 3.82).

Post-Target Word

The post-target word was fixated on 59.9% of trials; this happened more often when the target word was skipped (b = 2.14, SE = 0.08, z = 27.41), and when the target word was shorter versus longer (b = -0.35, SE = 0.14, z = -2.55). These two contrasts interacted; the post-target word was more likely to be skipped over when the target had been shorter versus longer (i.e., readers skipped two words more often when the second word was short; b = 0.30, SE = 0.07, z = 4.13). GZD was longer when the target was skipped versus fixated (b = -18.26, SE = 1.98, t = -9.23). GZD was longer following invalid vs. valid previews (b = 6.06, SE = 1.75, t = 3.47), and the preview and target skipping contrasts interacted; the preview effect on post-target GZD was larger when the target was skipped. The target skipping and target length contrasts interacted in GZD–post-target GZD was longer after skipping vs. fixating three and four-letter words, but this effect reversed following five-letter targets (b = 9.28, SE = 1.88, t = 4.93).

GPT followed a similar pattern as GZD. Reading times were longer after skipping the target word vs. fixating it (b = -48.95, SE = 3.80, t = -12.89), and were longer following invalid vs. valid previews (b = 17.54, SE = 3.69, t = 4.75). There was an

Table 3.3: GLMM and LMM results: Target word, Experiment 1. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). All $|t| or |z| \geq 2.00$ are marked in bold.

Effect	Skipping			GZD			GPT			RO		
	b	SE	*z*	*b*	SE	*t*	*b*	SE	*t*	*b*	SE	*z*
(Intercept)	-0.99	0.17	**-5.82**	248.50	5.99	**41.52**	278.11	8.11	**34.27**	-2.82	0.14	**-20.07**
Preview	-0.02	0.04	-0.49	13.09	1.79	**7.33**	20.51	2.95	**6.95**	0.42	0.11	**3.82**
Length	-0.76	0.06	**-12.56**	-4.65	2.01	**-2.31**	-6.34	3.39	-1.87	-0.09	0.09	-0.97
Preview X Length	-0.05	0.04	-1.44	-1.20	1.52	-0.79	-0.86	2.49	-0.35	0.07	0.08	0.87

Table 3.4: Post-target word subject means and standard errors by condition: Experiment 1.

Length	Preview	Skipping	GZD	GPT	RO
Target initially fixated					
3	HF	0.59 (0.03)	230.20 (12.26)	274.19 (20.69)	0.08 (0.02)
3	Target	0.64 (0.03)	234.70 (9.78)	269.97 (17.46)	0.08 (0.02)
4	HF	0.59 (0.03)	213.09 (6.46)	250.96 (11.79)	0.11 (0.03)
4	Target	0.57 (0.03)	217.86 (8.00)	240.38 (11.35)	0.07 (0.02)
5	HF	0.54 (0.03)	223.47 (8.09)	260.70 (10.92)	0.07 (0.02)
5	Target	0.59 (0.03)	222.09 (6.93)	246.82 (10.74)	0.05 (0.01)
Target initially skipped					
3	HF	0.12 (0.02)	276.63 (10.39)	409.37 (19.83)	0.35 (0.04)
3	Target	0.12 (0.02)	252.08 (10.17)	344.59 (17.85)	0.27 (0.03)
4	HF	0.04 (0.01)	266.48 (11.27)	358.22 (19.02)	0.32 (0.05)
4	Target	0.04 (0.01)	224.86 (8.59)	294.08 (19.57)	0.23 (0.04)
5	HF	0.11 (0.04)	219.66 (11.82)	338.29 (26.81)	0.34 (0.06)
5	Target	0.06 (0.02)	217.94 (12.00)	267.16 (19.37)	0.13 (0.04)

interaction between target skipping and preview in GPT, such that the preview effect was larger following skips versus fixations (b = -10.40, SE = 3.26, t = -3.19). Post-target GPT was longer following shorter versus longer targets (b = -12.70, SE = 5.59, t = -2.27). This contrast also interacted with target skipping– the effect of skipping shorter targets was larger than skipping longer targets (b = 15.69, SE = 3.47, t = 4.52). Readers regressed from the post-target word more often when the target was skipped versus fixated (b = -0.81, SE = 0.07, z = -11.64), and after invalid vs. valid previews (b = 0.36, SE = 0.08, z = 4.41).

In summary, we found strong effects of word length on skipping, and disruptive effects of the infelicitous preview manipulation on target and post-target word fixation times and regression rates. Interestingly, readers skipped over infelicitous previews at about the same rate as correct previews. That is, the invalid previews did not disrupt processing and affect skipping behavior. This is consistent with the hypothesis that readers do not use the context to guide skipping, and instead rely on a rough check of the word's identity. We test this more explicitly in Experiment 3.2 by using function words as infelicitous previews, which are very high in frequency and contain little semantic information.

3.4 Experiment 3.2

To further test our hypotheses, we conducted a second experiment using function word previews (*the*, *that*, and *there*). Here we expected larger effects of the preview manipulation, either due to function words' very high frequency, or because function words are skipped automatically. As in Experiment 3.1, if very short words like *the* are skipped more often than longer words due to saccadic range error, we expect an interaction such that the difference in skipping infelicitous previews versus valid previews will be smaller for short (three-letter) versus longer (five-letter) words. However, if skipping is driven primarily by parafoveal processing and not oculomotor constraints, we expect similar preview effects for shorter and longer words.

3.4.1 Method

Participants

Forty-eight University of California, San Diego, students who did not participate in Experiment 3.1 participated for course credit. All participants were native English speakers with normal or corrected-to-normal vision (using soft contacts or glasses), and were naïve to the purpose of the study.

Apparatus and Procedure

The apparatus and procedure were the same as in Experiment 3.1. The only difference was the introduction of *the*, *that*, and *there* as infelicitous previews.

Table 3.5: GLMM and LMM results: Post-target word, Experiment 1. *b*: Regression coefficient, SE: standard error, *t*: test statistic (b/SE). All $|t| or |z| \geq 2.00$ are marked in bold.

Effect	Skipping			GZD			GPT			RO		
	b	SE	*z*	*b*	SE	*t*	*b*	SE	*t*	*b*	SE	*z*
(Intercept)	-1.66	0.19	**-8.53**	227.78	5.71	**39.87**	284.10	11.19	**25.39**	-2.17	0.16	**-13.27**
Target Skip	2.14	0.08	**27.41**	-18.26	1.98	**-9.23**	-48.95	3.80	**-12.89**	-0.81	0.07	**-11.64**
Preview	0.04	0.07	0.53	6.06	1.75	**3.47**	17.54	3.69	**4.75**	0.36	0.08	**4.41**
Target Length	-0.35	0.14	**-2.55**	-5.12	3.32	**-1.54**	-12.70	5.59	**-2.27**	-0.17	0.08	**-2.07**
Target Skip X Preview	-0.13	0.07	-1.77	-5.25	1.67	**-3.14**	-10.40	3.26	**-3.19**	-0.07	0.06	-1.15
Target Skip X Target Length	0.30	0.07	**4.13**	9.28	1.88	**4.93**	15.69	3.47	**4.52**	0.06	0.06	0.99
Preview X Target Length	0.06	0.06	1.05	0.05	1.76	0.03	0.15	3.54	0.04	0.05	0.06	0.79
Target Skip X Preview X Target Length	-0.13	0.06	**-2.13**	3.16	1.70	1.86	1.66	3.29	0.51	-0.06	0.06	-1.04

Table 3.6: Target word subject means and standard errors (in parentheses) by condition: Experiment 2.

Length	Preview	Skipping	GZD	GPT	RO
3	HF	0.69 (0.03)	272.49 (10.47)	311.52 (15.51)	0.11 (0.03)
3	Target	0.49 (0.03)	242.18 (6.20)	274.97 (8.60)	0.10 (0.02)
4	HF	0.53 (0.04)	264.09 (8.61)	292.51 (10.69)	0.09 (0.02)
4	Target	0.36 (0.03)	226.75 (5.27)	253.88 (7.96)	0.09 (0.02)
5	HF	0.34 (0.04)	275.21 (8.42)	312.33 (12.47)	0.10 (0.01)
5	Target	0.21 (0.04)	229.31 (4.63)	261.02 (8.35)	0.08 (0.01)

3.4.2 Results and Discussion

Data processing followed the same procedure as Experiment 3.1. We deleted 6.4% of trials due to blinks or track loss on or around the target region, and an additional 12.9% of trials with improper display changes. We merged fixations that were 80 ms or shorter and within one character of an adjacent fixation, and deleted remaining fixations shorter than 80 ms or longer than 1000 ms. Comprehension was high (91.8%). As in Experiment 3.1, we excluded an additional 8.4% of trials where participants regressed from the pre-target word during first-pass reading. Table 3.6 presents subjects means for the different dependent measures and experimental conditions for the target word, with corresponding model results in Table 3.7. Subject means for the post-target word are in Table 3.8, with corresponding model results in Table 3.9.

Target Word

The target word skipping results are depicted graphically in Figure 3.1. The target word was skipped during first-pass reading on 41.5% of trials. There was a large effect of length such that shorter words were skipped more often than longer words (b = -0.76, SE = 0.06, z = -12.01). Readers skipped over infelicitous function word previews considerably more often than valid target word previews (b = 0.49, SE = 0.05, z = 9.38). There was no significant interaction between preview and length (z < 0.55).

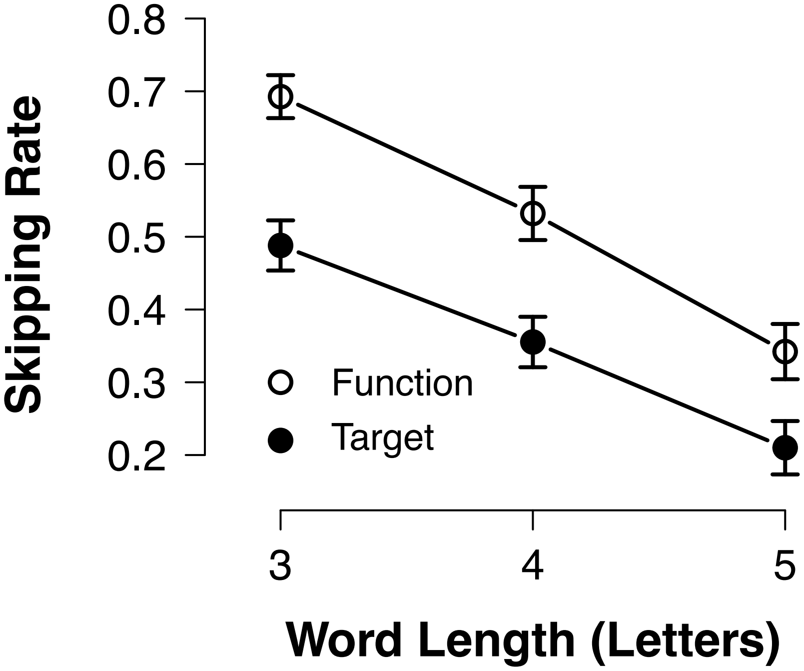

Figure 3.1: Target word skipping, Experiment 3.2.

When the target was fixated, GZD was longer following infelicitous function word previews than valid previews (*b* = 20.27, SE = 2.98, *t* = 6.80). There was no significant effect of length on target GZD, nor was there a significant interaction. GPT showed the same pattern, with longer reading times following infelicitous previews (*b* = 24.16, SE = 4.03, *t* = 6.01).

Post-Target Word

The post-target word was fixated on 64.1% of trials; it was fixated more often when the target word was skipped (*b* = 1.99, SE = 0.08, *z* = -17.35), and following invalid function word versus correct previews (*b* = 0.19, SE = 0.07, *z* = 2.78). The effect that skipping the target word had on fixating the post-target was larger following invalid versus valid target word previews (*b* = -0.34, SE = 0.07, *z* = -4.70). Fixating the post-target word was more common following shorter versus longer target words (*b* =

Table 3.7: GLMM and LMM results: Target word, Experiment 2. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). All $|t|or|z| \geq 2.00$ are marked in bold.

Effect	Skipping			GZD			GPT			RO		
	b	SE	*z*	*b*	SE	*t*	*b*	SE	*t*	*b*	SE	*z*
(Intercept)	-0.38	0.19	**-2.01**	253.06	5.69	**44.50**	286.22	7.85	**36.45**	-2.69	0.15	**-17.87**
Preview	0.49	0.05	**9.38**	20.27	2.98	**6.80**	24.16	4.02	**6.01**	-0.05	0.11	-0.43
Length	-0.76	0.06	**-12.01**	-2.71	2.38	-1.14	-2.59	4.03	-0.64	0.02	0.10	0.24
Preview X Length	-0.02	0.04	-0.54	2.71	1.89	1.43	1.45	3.28	0.44	0.06	0.08	0.74

Table 3.8: Post-target word subject means and standard errors (in parentheses) by condition: Experiment 2.

Length	Preview	Skipping	GZD	GPT	RO
Target initially fixated					
3	Function	0.58 (0.04)	230.01 (11.35)	265.50 (16.54)	0.09 (0.03)
3	Target	0.59 (0.04)	235.37 (9.63)	275.67 (13.47)	0.09 (0.02)
4	Function	0.50 (0.03)	219.32 (7.44)	248.80 (12.05)	0.07 (0.02)
4	Target	0.58 (0.04)	227.23 (8.63)	250.35 (10.94)	0.06 (0.02)
5	Function	0.51 (0.03)	233.67 (8.38)	264.21 (11.33)	0.10 (0.02)
5	Target	0.59 (0.03)	221.94 (6.71)	268.51 (18.74)	0.07 (0.02)
Target initially skipped					
3	Function	0.19 (0.03)	264.09 (10.39)	475.84 (16.23)	0.54 (0.04)
3	Target	0.12 (0.03)	265.38 (11.48)	365.84 (17.35)	0.25 (0.03)
4	Function	0.12 (0.03)	264.56 (9.61)	401.06 (15.43)	0.45 (0.04)
4	Target	0.07 (0.02)	241.48 (9.62)	295.99 (17.89)	0.16 (0.03)
5	Function	0.09 (0.03)	233.39 (7.10)	373.85 (18.62)	0.47 (0.05)
5	Target	0.12 (0.04)	216.88 (13.26)	314.96 (28.28)	0.27 (0.06)

-0.27, SE = 0.13, z = -2.14). The model revealed an interaction between target word skipping and target length, which indicates that the post-target word was most likely to be fixated when a four-letter target word had been skipped.

GZD was longer when the target word was skipped versus fixated (b = -17.35, SE = 2.27, t = -7.63). Target word length and target skipping interacted–the effect of skipping the target on GZD was restricted to the four-letter condition. GPT was longer after skipping versus fixating the target (b = -63.83, SE = 4.17, t = -15.32). GPT was longer following infelicitous function word previews compared to valid previews (b = 20.31, SE = 4.74, t = 4.29). Skipping the target modulated the preview effect, as the infelicitous function word preview effect was larger when the target was skipped versus fixated (b = -23.79, SE = 3.83, t = -6.21). GPT also longer following shorter versus longer targets (b = -17.92, SE = 6.46, t = -2.77). Target skipping also interacted with length, with longer GPT after skipping shorter versus longer targets (b = 11.86, SE = 3.90, t = 3.04). Regressions were more common after skipping versus fixating the target

(b = -1.85, SE = 0.15, z = -12.59), and following infelicitous function word previews versus valid previews (b = 0.45, SE = 0.09, z = 5.22). There was an interaction between target skipping and preview, with a considerably larger preview effect on regression rates when the target was skipped versus fixated (b = -0.30, SE = 0.07, z = -4.51).

3.5 General Discussion

We used the gaze-contingent boundary paradigm to manipulate parafoveal preview for three, four, and five-letter words. Readers were were given correct previews of three, four, or five-letter words, or previews of higher-frequency, length-matched contextually infelicitous words (content words of a different syntactic class, Experiment 3.1, or function words, Experiment 3.2). Our aims were to determine if word skipping is primarily driven by word identification processes (as is assumed by accounts like E-Z Reader; Reichle, Warren, & McConnell, 2009), with oculomotor constraints like saccadic range error (McConkie, Kerr, Redix, & Zola, 1988) playing a more minor role, and the extent to which readers process the syntactic context prior to skipping.

If oculomotor constraints result in much more frequent accidental skipping of short words, it was possible that the effects of preview frequency and length would interact such that the influence of frequency on skipping short words would be smaller than for longer words. On the other hand, if constraint is more heavily on acuity, which scales the rate of parafoveal processing, then it was possible that frequency would influence longer words less than shorter words. Following Angele and Rayner (2013), we expected that the syntactic context would not inform the decision to skip over a word.

In Experiment 3.1, skipping was driven primarily by word length, and there was no significant effect of the preview manipulation. Replicating the numeric pattern from Angele et al. (2014), HF infelicitous previews for three-letter words were skipped

Table 3.9: GLMM and LMM results: Post-target word, Experiment 2. b: Regression coefficient, SE: standard error, t: test statistic (b/SE). All $|t| or |z| \geq 2.00$ are marked in bold.

Effect	Skipping			GZD			GPT			RO		
	b	SE	*z*	*b*	SE	*t*	*b*	SE	*t*	*b*	SE	*z*
(Intercept)	-1.44	0.24	**-6.01**	233.66	6.14	**38.05**	307.19	10.63	**28.91**	-1.85	0.15	**-12.59**
Target Skip	1.99	0.08	**26.48**	-17.35	2.27	**-7.63**	-63.83	4.17	**-15.32**	-1.01	0.07	**-14.06**
Preview	0.19	0.07	**2.89**	2.78	2.43	1.14	20.31	4.74	**4.29**	0.45	0.09	**5.22**
Length	-0.27	0.13	**-2.14**	-6.18	3.52	-1.76	-17.92	6.46	**-2.77**	-0.14	0.09	-1.60
Target Skip X Preview	-0.34	0.07	**-4.70**	-2.10	2.10	-1.00	-23.79	3.83	**-6.21**	-0.30	0.07	**-4.51**
Target Skip X Target Length	0.28	0.07	**3.94**	4.81	2.16	**2.23**	11.86	3.90	**3.04**	-0.01	0.07	-0.09
Preview X Length	-0.06	0.05	-1.02	-1.24	2.05	-0.61	-7.22	4.18	-1.73	0.04	0.07	0.64
Target Skip X Preview X Target Length	0.00	0.05	0.07	-1.14	2.03	-0.56	3.14	3.74	0.84	0.09	0.06	1.42

more often than the valid previews, although this was not the case for four or five-letter previews. It could be that the difference in frequency between the preview and target words was not sufficiently large to promote more automatic skipping. Critically, this means that words were skipped at about the same rate regardless of their fit within the context. The preview manipulation had clear downstream effects: Reading times on the post-target region were longer following infelicitous previews, and this effect was magnified when readers skipped over the target word. Readers also regressed more often following invalid previews.

In Experiment 3.2 we tested whether function words are processed more automatically than content words, regardless of their length or syntactic fit, by using the function words *the*, *that*, and *there* for the invalid preview condition. Here we found large effects of preview on skipping at each word length, with no corresponding interactions. As in Angele and Rayner (2013) and Abbott et al. (2015), readers skipped *the* considerably more often than correct previews for contextually appropriate three-letter words, and did so as well for *that* and *there*. This result is important for at least two reasons: It highlights the importance of parafoveal processing compared to oculomotor constraints for word skipping and it contributes additional evidence that the sentence context seems to be irrelevant to the decision to skip over short and medium-length function words and content words.

Our results provide additional evidence that word skipping reflects a rough check of a word's identity, and not deeper semantic or syntactic processing. This is consistent with the architecture of E-Z Reader 10 (Reichle et al., 2009), which assumes that lexical processing and postlexical integration occur in serial processing stages. In E-Z Reader, words are skipped when an early stage of lexical processing completes from the parafovea before the eyes are ready to move on to the next word; this stage is influenced solely by word frequency and predictability. In the model, the integration stage (which is intended to capture all syntactic and semantic integration) begins after lexical processing has completed. Thus, by the model's design there is no clear mechanism for integration difficulty to disrupt word skipping, although it has relatively clear implications for downstream processing. Abbott and Staub (2015) demonstrated independent effects of word identification and integration difficulty across the eye movement record (including skipping) using a factorial manipulation of frequency and plausibility.

That the semantic or syntactic context does not appear to influence word skipping seems odd in light of the fact that predictability does affect skipping (e.g., Rayner et al., 2011). Abbott et al. (2015) demonstrated that the rate at which infelicitous *the* previews are skipped is affected by the predictability of the word at that position in the sentence. Readers skipped previews of *the* more often when a different word was predictable versus unpredictable, and this effect was additive with respect to the *the*-skipping effect. What, then, is the role of predictability? Perhaps it provides readers with a level of confidence that the upcoming word will be easy to identify, regardless of its actual identity. This would be broadly consistent with the idea of predictability as preactivation (Staub, Grant, Astheimer, & Cohen, 2015).

It remains an open question whether function words are skipped so often due to their very high frequency or due to some other mechanism. Function words may, for example, be more highly recognizable based on shape, allowing them to be more rapidly identified from the parafovea. Future research could manipulate aspects of parafoveal word shape to determine whether word shape (a rough visual cue) or letter identity (obtained via lexical processing) is more important to skipping.

In summary, readers prefer to skip over function words like *the*, *that*, and *there* over lower-frequency length-matched words regardless of their fit within the sentence context. Skipping these words is deleterious to processing (and yet readers cannot seem to help it) as readers fixate longer following these invalid previews and make more frequent regressions. A topic of future research is to examine how deeply infelicitous words are processed from the parafovea, and when the "error signal" is sent that an incorrect word has been skipped over. Further research should also aim to determine how function it is that function words are processed so automatically from the parafovea.

Chapter 3, in full, is a reprint of material in preparation for publication: Preferential skipping of function words during reading (Abbott & Angele, in prep). The dissertation author was the primary investigator and author of this paper.

3.6 Appendix: Experimental stimuli from Experiment 3.1

Target words are in *italics* and the invalid preview words are in parentheses. Items 1-60 have five-letter targets, items 61-120 have four-letter targets, and items 121-180 have three-letter targets (and were adopted from Angele et al., 2014, with permission from the authors). For Experiment 3.2, the invalid preview words were always either *the*, *that*, or *there* for three, four, and five-letter target words, respectively.

1. Some irritating drivers will *weave* (*shirt*) through traffic at erratic speeds.
2. Our daughter would often *drool* (*devil*) on her clothes and make a mess as a baby.
3. I like to slowly *munch* (*acorn*) some peanuts and other snacks while doing my homework.
4. You can use a special tool to remove particles that *cling* (*elite*) onto your jacket.
5. The little boy would *blush* (*creek*) every time a girl said hello to him.
6. It is possible that you might *drown* (*trend*) if you don't know how to swim properly.
7. The floors were so clean that we could *glide* (*bread*) across them when wearing socks.
8. Though it is somewhat taboo, some people will *spank* (*males*) their child when they do wrong.
9. The professor thinks his student should *dwell* (*woods*) on the issue before writing the conclusion.
10. Some animals will impatiently *stomp* (*blade*) around when they are hungry or bored.

11. The man thought that if he could silently *creep* (*pants*) by the guards he could take the jewels.
12. Every spring Jackie would *sniff* (*snail*) and sneeze when the pollen levels were high.
13. Though it is unlikely, the government could *seize* (*batch*) your assets in some situations.
14. Most people will *blink* (*pearl*) when something is coming toward their eyes.
15. As kids my brother and I would *shove* (*fruit*) each other around in jest all the time.
16. While you are driving you must *steer* (*teeth*) your car to avoid hitting trees and other vehicles.
17. You will most likely *bleed* (*cloth*) after receiving the vaccine in your left arm.
18. The kids will sometimes *sneak* (*chess*) around the house to avoid being seen by their parents.
19. It is very rude to blatantly *stare* (*depth*) at other people in public settings.
20. If you are not careful you could *choke* (*chest*) on your food and end up in the hospital.
21. The punk kids would rudely *slash* (*guest*) tires in the parking lot by the school.
22. The couple didn't know they could *grind* (*peace*) fresh herbs using their coffee grinder.
23. Before my daughter could *crawl* (*queen*) she could barely even roll over.
24. We hoped that the young actor would *shine* (*sheet*) in his first roll in a major film.

25. The dog was trained to quickly *fetch* (*trial*) tennis balls in exchange for tasty treats.
26. At parties we would often *crank* (*noise*) up the music until the walls were shaking.
27. We hoped the home team would *crush* (*money*) the opposition and win the game.
28. We hoped that the celebration would *cease* (*crime*) soon after they arrived.
29. My parents will always *bless* (*clock*) the meal before we are allowed to eat.
30. The doctor will usually *weigh* (*cards*) his patient at the beginning of a checkup.
31. Emily was concerned that she could *scare* (*cards*) her roommates with her sleep walking.
32. The new advertisement might *boost* (*faith*) tourism in the old and forgotten town.
33. I was taught not to loudly *swear* (*scene*) in public as it is likely to offend someone.
34. Though the infant could *shake* (*disks*) her rattle, she couldn't crawl yet.
35. We watched the kids lazily *float* (*miles*) around on inner tubes at the lake yesterday.
36. John couldn't quite *grasp* (*grasp*) the concepts in statistics and went to office hours for help.
37. Peter told his brother that he should *climb* (*birth*) to the top of the tree house to see the stars.
38. If you choose to openly *cheat* (*beach*) on your exam you may be expelled from school.
39. In certain places you must *yield* (*years*) and allow other drivers to go before you.

40. The boxer couldn't quite *dodge* (*facts*) out of the way in time and was punched in the face.

41. The men had to forcefully *knock* (*mouth*) down the door in order to get inside the house.

42. In the winter time we would *skate* (*style*) around on the icy pond in our back yard.

43. There's a risk that someone could *steal* (*mouse*) your belongings if your door is unlocked.

44. On weekends we will *chuck* (*blood*) tennis balls across the yard for our dog to chase.

45. Although Kate could *blame* (*brain*) her sister for the accident, she said it was her fault.

46. We hired someone who could *mount* (*suite*) our new television to the wall.

47. We learned how to correctly *solve* (*sales*) difficult math problems in our linear algebra class.

48. In the game everyone must *drink* (*child*) each time someone says a certain word.

49. I thought that I could *teach* (*death*) myself calculus but I couldn't find a good book.

50. Rob is skilled and will *catch* (*weeks*) every ball that is thrown to him.

51. If we are seen someone might *shoot* (*hours*) at us and start a small battle.

52. The couple wondered how they could *raise* (*month*) their children to be just like them.

53. We hoped the children would *sleep* (*smith*) through the night after our long day out.

54. The staff were trained to properly *serve* (*words*) their guests with a smile on their face.

55. Joe’s dad could probably *throw* (*house*) him five hundred pitches before he will hit the ball.

56. The man had to somehow *prove* (*years*) that he did not commit the horrible crime.

57. The toddler could not quite *reach* (*games*) the jar of cookies on top of the counter.

58. The couple decided they would *spend* (*world*) their savings on a down payment for a house.

59. The company wasn’t sure if they could *build* (*thing*) the house within their time constraints.

60. Carla was concerned that she would *break* (*money*) her leg when she went skiing for the first time.

61. Singing is not a skill that one would *deem* (*myth*) necessary for your survival.

62. John has a tendency to openly *brag* (*bull*) about his karaoke abilities.

63. I assume that you could *weld* (*fort*) the pieces together using a special tool.

64. When we dance we will *sway* (*nuts*) from side to side and look into each other’s eyes.

65. The soldiers decided to quickly *flee* (*odds*) from the approaching enemy.

66. Yesterday Claire decided to impulsively *pawn* (*lips*) off several old items.

67. Yesterday Liz decided she would *mash* (*folk*) sweet potatoes to make baby food.

68. The defenders had to somehow *fend* (*lane*) off the offensive to win the game.

69. In many cities weird people will *lurk* (*tube*) in back alleys and dimly lit areas.

70. On the weekends my friends often *coax* (*dice*) me to join them at bars and parties.

71. A good way to clean dishes is to thoroughly *soak* (*cell*) them in warm water and soap.

72. When my computer dies I will probably *moan* (*disc*) for a while before I buy a new one.

73. On Saturdays we like to leisurely *hike* (*fees*) through the desert and the mountains.

74. Your heart might *melt* (*math*) when you see the picture of our newborn baby.

75. Sometimes you must *peel* (*boat*) fruits before you can eat them.

76. Mike will sometimes *rant* (*rice*) about his day if things don’t go his way.

77. Some say that you should *chew* (*bird*) your food completely before you swallow.

78. Mary thought she would *rave* (*bass*) about last night’s show but she really hated it.

79. It isn’t a good idea to blindly *lend* (*lake*) your expensive cell phone to a stranger.

80. I was afraid that I would *spew* (*font*) the food out of my mouth if I laughed too hard.

81. After many washes the colors will often *fade* (*bits*) from your clothing.

82. Julie wasn’t sure how she would *cope* (*wars*) with the recent loss of her grandmother.

83. We need to quickly *boil* (*beer*) some water before adding in the pasta.

84. Other drivers like it when you promptly *warn* (*wood*) them of turns and lane changes.

85. The mother likes to gently *wipe* (*bell*) her baby's mouth after each feeding.

86. It will take some time to fully *heal* (*goal*) the wound after the tragic accident.

87. All night the infant would *toss* (*tree*) and turn until she learned to calm herself down.

88. If you don't *stir* (*self*) the ingredients together, the dish will not cook properly.

89. The little boy likes to slowly *lick* (*hill*) his ice cream to savor it as much as possible.

90. The trick is to somehow *lure* (*loss*) fish toward your hook if you want to catch them.

91. Before they can properly *swim* (*cars*) many children simply wade in shallow water.

92. It is recommended that you should *cite* (*soul*) each source when writing research papers.

93. If you are going to be late to work I would *urge* (*role*) you to contact your supervisor.

94. The girl would often *skip* (*tool*) through the park on sunny afternoons.

95. A good friend will *dare* (*hall*) you to do something you've never done before.

96. Many people will quietly *pray* (*town*) before they eat each meal.

97. At the gym many people *lift* (*wife*) weights and run on treadmills to get in shape.

98. When I get home I usually *dump* (*laws*) the contents of my pockets on the counter.
99. The baby will often *grab* (*ball*) any object within her field of vision.
100. If you very quickly *spin* (*feet*) around in circles you are likely to become dizzy.
101. Jack is so happy that he could *sing* (*hair*) praises from the top of a mountain.
102. Kate decided to carefully *pour* (*path*) the contents of the beaker down the drain.
103. It would be great if we could *hire* (*hour*) more assistants to work in our lab.
104. Some of the new soldiers *drag* (*king*) behind during drills until they learn to keep up.
105. Some lucky people will *earn* (*tech*) millions of dollars in their lifetime.
106. There's a good chance that you would *burn* (*door*) your hand if you touch a hot stove.
107. The young girl will often *hide* (*mode*) from her parents when she has done something wrong.
108. It is not a good idea to simply *quit* (*ways*) projects that you find very difficult to finish.
109. Kim will most often *hang* (*luck*) clothes immediately after taking them out of the dryer.
110. Some people think that they might *fail* (*food*) a final exam despite extensive studying.
111. To get in you must *push* (*guys*) the button that opens the heavy door.
112. John is paranoid and thinks that people *seek* (*lots*) him out to verbally attack him.

113. It is important that you fully *shut* (*road*) down the lab computers at the end of the day.

114. Mark was afraid he would *gain* (*ones*) lots of weight if he ate pizza every day.

115. John hoped that his son would *grow* (*data*) into a strong and intelligent young man.

116. First you must slowly *fill* (*none*) the bucket with water before you add the soap.

117. It is possible that you might *hurt* (*news*) yourself if you fall down the stairs.

118. Skilled combat medics rapidly *tend* (*kind*) wounded soldiers on the battlefield.

119. It is a tough economy and many people *lose* (*days*) their jobs due to budget cuts.

120. I am thinking that I might *join* (*life*) a few clubs on campus to meet new people.

121. There was a massive *ant* (*all*) infestation in the old house.

122. He found the shoes a tiny *bit* (*buy*) too small.

123. This weekend, I saw a great *ape* (*aim*) and her new-born baby at the zoo.

124. The children would always *beg* (*boy*) their mother for ice cream.

125. The karate teacher threw his students down on the large *mat* (*but*) multiple times.

126. Watching the flowers *bud* (*ago*) was magical.

127. In order to solve the math problem, one must *sum* (*she*) up all of the variables.

128. The famous singing *duo* (*dry*) performed for the entranced audience.

129. He would never *don* (*did*) the red hat for the formal ceremony.

130. The lamp cast an increasingly *dim* (*dog*) light over the large room as its fuel ran out.

131. She sharpened her pencil down to a tiny *nub* (*not*) in the middle of class.

132. I have a little kangaroo that will *hop* (*his*) on command.

133. To turn the light switch *off* (*far*) he had to jump very high.

134. The fine gentleman pulled out his impressively *fat* (*for*) stack of bills in order to pay for dinner.

135. They slept in the small *cot* (*yet*) because their beds were not ready.

136. The boy was very *shy* (*gas*) when he performed on stage.

137. I promised myself that I would go to the local *gym* (*get*) every day this week.

138. The burglar will *gag* (*god*) the frightened old woman.

139. The lake was so big we needed an extra *oar* (*had*) so we could get around faster.

140. My grandmother told me she would *hem* (*has*) that dress for me.

141. The novice player *hit* (*her*) the ball over the fence.

142. During our road trip a huge *bug* (*him*) splattered onto our window.

143. The mother *hen* (*hot*) watched over her chicks in the coop.

144. During the game of tag, the smart little girl *hid* (*how*) where no one could find her.

145. Because it was empty, she refilled the freezer's *ice* (*its*) cube tray with water.

146. Her friend motioned for her to wipe her lower *lip* (*lay*) because she had sauce on it.

147. She liked standing on the soft floor *pad* (*met*) in the bedroom.

148. My dog loves taking a long *nap* (*new*) when its hot outside.

149. The fisherman pulled in his heavy *net* (*nor*) and discovered that he had caught a swordfish.

150. The husband would *hog* (*now*) all of the blankets at night.

151. When he bought the chemistry *kit* (*got*) it didn't come with instructions.

152. They knew that much *fun* (*men*) was to be had at the carnival.

153. The couple *own* (*one*) a large house in the country.

154. A bright shimmering *orb* (*our*) appeared in the night sky.

155. She would not eat even a single *pea* (*pay*) off of her plate.

156. The nice man would kindly *nod* (*per*) every time someone entered the store.

157. The priest bought an expensive *pew* (*put*) for his church.

158. They heard the high-pitched *cry* (*set*) of the wolf in the night.

159. The vegetables began to immediately *rot* (*six*) after she had bought them.

160. The man would peek over his fence and surreptitiously *spy* (*war*) on his neighbors.

161. Because she is lactose intolerant, she drinks *soy* (*you*) milk.

162. I spilled juice on the precious *rug* (*win*) when she bumped my arm.

163. She had to forcefully *tug* (*two*) the shirt to get it out from under the dresser.

164. I wore my purple *wig* (*why*) the other day, and received a lot of compliments.

165. She had a giant *wad* (*who*) of gum in her mouth.

166. She hit her golf ball off of the small *tee* (*was*) and watched it soar into the distance.

167. She rode in her boyfriend's *van* (*via*) to the concert.

168. I like to relax in my large *tub* (*use*) at the end of a long day.

169. I saw the little foxes scurry into their *den* (*tea*) as I came walking by.

170. I left a large *tip* (*sat*) for the helpful waiter.

171. My grandmother has a small *urn* (*tax*) holding her beloved dog's remains.

172. The campers stayed away from the wolf *pup* (*are*) because they knew the mom was near.

173. They circled the remote *bay* (*act*) a few times before debarking.

174. The critics expressed *awe* (*aid*) at the orchestra's inspiring rendition of Beethoven's fifth.

175. The beginner took lessons from the golf *pro* (*sun*) so he could get better.

176. As the tiny, furry *cub* (*key*) rolled around on the ground, the mother bear watched him closely.

177. The popular girl at school started a novel *fad* (*try*) by wearing her socks inside-out.

178. The old man would talk to his small *cat* (*art*) as if it was human.

179. The captain told the sailor to stand at the very *aft* (*gun*) of the boat.

180. The woman was very *coy* (*sun*) with the younger gentleman.

References

Abbott, M. J., Angele, B., Ahn, Y. D., & Rayner, K. (2015). Skipping syntactically illegal the-previews: The role of predictability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 000–000.

Abbott, M. J., & Staub, A. (2015). The effect of plausibility on eye movements in reading: Testing e-z reader's null predictions. *Journal of Memory and Language*, *41*, 1703-1714.

Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1181-1203.

Angele, B., & Rayner, K. (2013). Processing the in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 649–662.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., ... Treiman, R. (2007). The english lexicon project. *Behavior Research Methods*, *39*, 445-459.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278.

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using eigen and s4. (r package ver. 1.1-7). *Vienna, Austria: R Foundation for Statistical Computing*.

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1552-1560.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). Swift: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777-813.

Gautier, V., O'Regan, J. K., & Le Gargasson, J. F. (2000). 'The-skipping'revisited in french: Programming saccades to skip the article 'les'. *Vision research*, *40*(18), 2517–2531.

McConkie, G. W., Kerr, P. W., Redix, M. D., & Zola, D. (1988). Eye movement control during reading: I. the location of initial eye fixations on words. *Vision Research*, *28*, 1107-1118.

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

O'Regan, J. K. (1979). Eye guidance in reading: Evidence for the linguistic control hypothesis. *Perception and Psychophysics*, *25*, 501–509.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. doi: 10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K. (2009). The thirty-fifth sir frederick bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *62*, 1457–1506.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 514-528.

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*, 1-21.

Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, *82*, 1-17.

Chapter 4

The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions

Journal of Memory and Language 85 (2015) 76–87

Contents lists available at ScienceDirect

Journal of Memory and Language

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

The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions

Matthew J. Abbott [a,*], Adrian Staub [b]

[a] Department of Psychology, University of California, San Diego, United States
[b] Department of Psychological and Brain Sciences, University of Massachusetts Amherst, United States

ARTICLE INFO

Article history:
Received 5 May 2015
revision received 23 July 2015

Keywords:
Eye movements
Bayesian inference
Lexical processing

ABSTRACT

The E-Z Reader 10 model of eye movements in reading (Reichle, Warren, & McConnell, 2009) posits that the process of word identification strictly precedes the process of integration of a word into its syntactic and semantic context. The present study reports a single large-scale ($N = 112$) eyetracking experiment in which the frequency and plausibility of a target word in each sentence were factorially manipulated. The results were consistent with E-Z Reader's central predictions: frequency but not plausibility influenced the probability that the word was skipped over by the eyes rather than directly fixated, and the two variables had additive, not interactive, effects on all reading time measures. Evidence in favor of null effects and null interactions was obtained by computing Bayes factors, using the default priors and sampling methods for ANOVA models implemented by Rouder, Morey, Speckman, and Province (2012). The results suggest that though a word's plausibility may have a measurable influence as early as the first fixation duration on the target word, in fact plausibility may be influencing only a post-lexical processing stage, rather than lexical identification itself.

Introduction

Comprehending a sentence involves identifying each word as it is encountered. In addition, each word must be integrated into a representation of the sentence's structure and meaning. A range of language processing models (e.g., Friederici, 2002; Lewis & Vasishth, 2005; Reichle, Warren, & McConnell, 2009) have proposed architectures in which word identification, on the one hand, and integration of a word into its syntactic and semantic context, on the other, occur in separate, serially ordered stages. This distinction between lexical and post-lexical processing is not universally endorsed, however. For example, it is well established that the amplitude of the N400 waveform, which occurs in event-related potential (ERP) experiments in response to each word in a sentence, is increased when a word is implausible in its context (e.g., *He spread the warm bread with socks*; Kutas & Hillyard, 1980), and this effect has been interpreted as resulting from contextual modulation of the difficulty of lexical access itself (e.g., Lau, Phillips, & Poeppel, 2008).

The present study is concerned with the question of whether the effects of a word's plausibility in context on eye movements during reading are consistent with a staged architecture that would attribute these effects to a post-lexical integration process. Several previous studies have examined the effects of implausibility on eye movements (e.g., Matsuki et al., 2011; Rayner, Warren, Juhasz, & Liversedge, 2004; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007; Warren & McConnell, 2007; Warren, McConnell, & Rayner, 2008). Rayner et al. (2004) presented

---

* Corresponding author at: Department of Psychology, University of California, San Diego, 9500 Gilman Dr., MC 0109, La Jolla, CA 92093-0109, United States.

E-mail address: mabbott@ucsd.edu (M.J. Abbott).

http://dx.doi.org/10.1016/j.jml.2015.07.002

readers with sentences such as (1) below, labeling the target word (in italics) as either plausible (a), implausible (b), or anomalous (c); it is a matter of debate whether the violations involved in (b) and (c) are different in kind or merely in degree (Matsuki et al., 2011).

1. (a) John used a knife to chop the large *carrots* for dinner.
   (b) John used an axe to chop the large *carrots* for dinner.
   (c) John used a pump to inflate the large *carrots* for dinner.

Rayner et al. observed increased gaze duration (the sum of the durations of all eye fixations on the word before leaving it for the first time) on the target word in the anomalous condition compared to the plausible condition. In comparing the implausible and plausible conditions, a significant difference first emerged in a measure that is usually regarded as reflecting later processing, go-past time (the sum of all fixation durations beginning with the first fixation on the word until the eyes go past the word to the right, including any regressive fixations).

Subsequent studies, however, have shown effects of plausibility as early as the first fixation duration on a target word (i.e., the duration of the eyes' very first fixation on the word; Matsuki et al., 2011; Staub et al., 2007), a measure that is typically affected by a word's frequency (e.g., Rayner & Duffy, 1986; Staub, White, Drieghe, Hollway, & Rayner, 2010) and its predictability (i.e., cloze probability; e.g., Rayner & Well, 1996; Staub, 2011b). Staub et al. (2007) manipulated whether the first noun of a noun–noun compound (see 2; target word in italics) was plausible (a) or implausible (b) as the head of a noun phrase at that point in the sentence. In fact, the critical word was the first constituent of a compound, and the sentences were globally plausible.

2. (a) The new principal visited the *cafeteria* manager at the end of the school day.
   (b) The new principal talked to the *cafeteria* manager at the end of the school day.

Staub et al. observed longer reading times on the target word when it was an implausible head noun in several measures (first fixation duration, gaze duration, and go-past time). These effects did not carry over onto the subsequent noun of the compound, though other plausibility studies in which the sentences are globally implausible have found continued downstream effects (e.g., Rayner et al., 2004; Warren & McConnell, 2007).

The previous literature has demonstrated, then, that effects of a word's plausibility may appear very rapidly indeed, in some cases on the earliest reading time measures. The question we address here is whether these very early effects of plausibility may still be regarded as emerging at a post-lexical stage of processing. We address this question by considering the specific predictions for post-lexical effects that emerge from the E-Z Reader 10 model of eye movements in reading (Reichle et al., 2009). E-Z Reader 10 is the first version of the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003; see Reichle, 2011, for a review of the development of E-Z Reader) to account for the influence of syntactic and semantic integration difficulty on eye movements. Though prior versions modeled the influence of a word's predictability (i.e., cloze probability), this factor was treated as influencing the same stages of lexical processing that are influenced by word frequency. E-Z Reader 10 proposes that both syntactic processing difficulty and the difficulty of semantic interpretation in context influence an explicitly post-lexical Integration (*I*) stage. The model proposes that effects in the eye movement record will appear if the process of integration lags sufficiently far behind lexical processing of the next word, or if this process fails outright.

Simulations performed by Staub (2011a) illustrated a number of specific empirical predictions that emerge from the E-Z Reader 10 architecture. First, in contrast to lexical factors such as word frequency, an effect of integration difficulty is not necessarily predicted to appear in reading times on the critical word; it may show up only downstream of that word. Second, when lexical processing difficulty and integration difficulty both influence reading time measures on the critical word such as first fixation duration and gaze duration, they are predicted to do so in a strictly additive manner. Third, the probability of a regressive eye movement from the target word may also be influenced by integration difficulty, but lexical processing difficulty should not influence this measure. In a series of experiments in which a word's frequency and its syntactic fit were factorially manipulated, Staub (2011a) largely confirmed these predictions.

The integration stage of E-Z Reader 10 is intended to capture both syntactic processing difficulties of the kind explored by Staub (2011a) and semantic interpretation difficulties of the kind encountered in the plausibility experiments described above. Thus, precisely the same predictions should hold with respect to plausibility: In experiments in which plausibility does influence the early reading time measures of first fixation duration and gaze duration, this effect should be strictly additive with the effect of word frequency; and plausibility may influence the probability of regressive eye movement, but frequency should not. It is important to note that the identification of both syntactic and semantic integration with a single post-lexical processing stage is intended as a simplification; Reichle et al. (2009) make clear that this is not meant to preclude a more finely articulated conception of how various aspects of syntactic, semantic, and pragmatic processing may proceed and may interact. E-Z Reader 10 is clearly committed, however, to the idea that these forms of processing occur only after the lexical processing stages that are influenced by, e.g., word frequency.

An additional prediction of the E-Z Reader 10 framework that was not discussed by Staub (2011a) relates to word skipping. It is well established that while a word's length exerts the largest influence on the probability that it is skipped rather than directly fixated by the eyes, both a word's frequency and its predictability also exert an influence on word skipping (e.g., Drieghe, Rayner, & Pollatsek, 2005; Rayner, Slattery, Drieghe, & Liversedge, 2011). E-Z Reader posits that word skipping takes place when an initial stage of lexical processing (the L1 stage) has been completed during parafoveal viewing (i.e., while the previous word is still being fixated), and while the planned saccade into the word can still be cancelled. Both frequency and predictability influence the duration of the L1 stage, and therefore the probability that L1 will complete in time for a word to be skipped. But because the integration stage of E-Z Reader 10 is fully post-lexical, occurring not only after L1 but also after a second stage of lexical processing, L2, difficulty at the integration stage will not directly influence the probability that a word is skipped.[1] Manipulations of syntactic attachment difficulty and plausibility should therefore not influence word skipping, even if they influence reading time measures as early as first fixation duration.

Many plausibility studies (Rayner et al., 2004; Staub et al., 2007; Warren & McConnell, 2007) have not reported skipping rates for the target words. An exception is Matsuki et al. (2011), who did report a significant effect of plausibility (or in their terminology, *typicality*) on skipping probability. However, the interpretation of this result is complicated by the fact that the typical and atypical targets in Matsuki et al. (2011) differed substantially in predictability, with mean cloze probability of 20.8% for typical targets and 1% for atypical targets. It is well established that predictability influences word skipping (e.g., Drieghe et al., 2005). Recent studies by Angele and colleagues (Abbott, Angele, Ahn, & Rayner, 2015; Angele, Laishley, Rayner, & Liversedge, 2014; Angele & Rayner, 2013) support the contention that skipping is determined by the ease of lexical identification of the upcoming word rather than by integration difficulty. These studies used the boundary paradigm (Rayner, 1975) to display a three-letter word in the location of the upcoming word that was different from the word that would ultimately appear, once the eyes fixated this location. Angele and Rayner (2013) and Angele et al. (2014) found that when this parafoveal preview word was high in frequency it tended to be skipped, even when it was syntactically or semantically illicit. This is as predicted by E-Z Reader 10; the skipping decision is made on the basis of lexical processing, before the fit of the word in the sentence can be evaluated.

Several of the predictions of E-Z Reader that we have just discussed are predictions of null effects or null interactions. In the present study, we manipulate the frequency and the plausibility of a target word. E-Z Reader predicts that (a) frequency but not plausibility will influence the probability that the word is skipped; (b) frequency will not influence the probability of a regressive saccade from the word, though plausibility may; and (c) if both variables influence reading time measures they will have additive effects, i.e., there will be a null interaction. As the null is a theoretically important hypothesis in this context (see, e.g., Gallistel, 2009; Rouder, Speckman, Sun, Morey, & Iverson, 2009, for arguments as to the theoretical importance of null effects), the traditional null hypothesis significance testing (NHST) framework is problematic. If the probability of an effect as large as one observed in an experiment is sufficiently small (i.e., $p < .05$) given the null hypothesis of no effect, we reject this null hypothesis. But when this threshold is not reached, we do not claim to have evidence in favor of the null. On this approach, the absence of an effect is not itself evidence of absence. A theory that predicts an absence of an effect, however elegant and well motivated by data and argument, is then "on the wrong side of the null," (Rouder, Morey, Speckman, & Province, 2012) in the sense that traditional statistical inference cannot support it (Gallistel, 2009; Raftery, 1995; Rouder et al., 2009).

An alternative to NHST that has recently received increased attention from experimental psychologists involves directly assessing the relative support for null and non-null statistical models by computing a Bayes factor (Edwards, Lindman, & Savage, 1963; Gallistel, 2009; Kass & Raftery, 1995; Nathoo & Masson, 2015; Rouder et al., 2009, 2012; Wagenmakers, 2007). A Bayes factor is the ratio of the data's marginal likelihood under two models. In the present context, these would be a statistical model in which the value of a specific parameter – either a main effect or an interaction – is posited to be zero, and a model on which this parameter is not restricted to a null value. If the two models are assumed to have equal prior probability (i.e., the two models are assumed, in advance of considering the data, to be equally likely), the Bayes factor is then the ratio of the posterior probabilities of the two models. Values near 1 indicate similar marginal likelihoods, and values much larger than 1, or fractional values much smaller than 1, provide evidence in favor of one of the models. It has been suggested that a Bayes factor greater than 3.2 represents "substantial" evidence for one model over another, greater than 10 represents "strong" evidence, and greater than 100 represents "decisive" evidence (Jeffreys, 1961; Kass & Raftery, 1995). However, it is arguably a central virtue of Bayes factors that investigators can bring their own evidentiary standards to bear once a numerical Bayes factor has been calculated, eliminating the need for a specific convention that supports a binary decision.

The Bayes factor relies on a computation of the likelihood over the entire parameter space of each model, with the likelihood weighted by the prior probability that the model assigns to different values of each parameter, i.e., the likelihood is integrated over the prior distribution of the parameters. A model is therefore penalized for assigning prior probability to parameter values under which the data are very unlikely. Critically, this penalizes a more

---

[1] In fact, there is a complex, indirect mechanism by which integration difficulty can reduce skipping. In cases in which word n would have been skipped based on early completion of the L1 stage, but word n is difficult to integrate, this difficulty will occur concurrently with the planning of the saccade into word $n + 1$. This can prevent the saccade into word $n + 1$, and on some proportion of these trials, word n is then fixated. However, simulations show that this mechanism comes into play rarely enough that it would have a measurable effect only when word n is very short, and therefore would be very frequently skipped. When word n is five or more letters, as in the present experiment, this mechanism results in a numerical reduction in skipping rate of only one or two percent even at extreme levels of integration difficulty.

complex model as a function of its vagueness compared to a null model, which is committed to the position that the value of the parameter in question is precisely zero. Thus, the Bayes factor can favor a null model, as it can indicate that the data are more likely under a null model than under a model that assigns substantial probability to non-zero values of a critical parameter or vector of parameters. This distinguishes the Bayes factor from the likelihood ratio test (LRT), which has entered wide use in the context of fitting mixed-effects models in psycholinguistics (e.g., Baayen, 2008). The LRT compares the likelihood of the data under a specific parameter setting for each model, i.e., that which results in the maximum likelihood of the data. The maximum likelihood under the more complex model will always be at least as high as under the less complex model, providing no direct way of supporting the less complex model.

Here we illustrate the use of Bayes factors in an ANOVA design, as recently implemented by Rouder et al. (2012).[2] We believe this is novel in the context of studies of eye movements in reading, and perhaps in the broader psycholinguistic context. The method is implemented in a convenient R package, *BayesFactor* (Morey, Rouder, & Jamil, 2015), which uses Monte Carlo sampling to integrate the likelihood function over a model's prior distribution over parameter values. A critical aspect of this approach involves the development of default priors for the parameters of ANOVA models, making use of Cauchy priors on effect sizes (Johnson, Kotz, & Balakrishnan, 1994) and a Jeffreys prior on the variance (Jeffreys, 1961). For ANOVA designs, one can compare each of several nested linear models (i.e., a model containing one fixed effect, two additive fixed effects, two additive fixed effects and their interaction, etc.) against a model that estimates only the random effects (i.e., subject or item) variance. Nested models can also be directly compared to one another by taking the ratio of their Bayes factors, permitting direct comparison of an additive versus interactive model, or for any pair of models that share a common denominator in their Bayes factor computations.[3]

Here we report a single, large-scale ($N = 112$) experiment in which subjects' eye movements were monitored as they read sentences like those in (3) and (4), which contained a target word that was manipulated with respect to its frequency and its plausibility given the preceding context. The target word was always the direct object noun in a simple one-clause sentence. In the implausible conditions, this noun violated the animacy requirements imposed by the preceding verb; the verb in these conditions either required an inanimate theme and the target noun was animate (as in 3), or the verb required an animate theme and the target was inanimate (as in 4). The material following the critical word was allowed to vary between plausible and implausible conditions, as all analyses focused on first pass reading of the critical word.

3. (a) The professor invited the *writer* to an important meeting. (HF, plausible)
   (b) The professor invited the *orator* to an important meeting. (LF, plausible)
   (c) The professor repaired the *writer* with a trusty old wrench. (HF, implausible)
   (d) The professor repaired the *orator* with a trusty old wrench. (LF, implausible)

4. (a) The man noticed the *journal* was missing from his desk. (HF, plausible)
   (b) The man noticed the *stapler* was missing from his desk. (LF, plausible)
   (c) The man angered the *journal* by placing it in the drawer. (HF, implausible)
   (d) The man angered the *stapler* by placing it in the drawer. (LF, implausible)

We used Bayes factors to assess the evidence for the critical null effects predicted by the staged architecture of E-Z Reader 10: a null effect of plausibility on skipping rate; a null interaction between frequency and plausibility in reading time measures; and a null effect of frequency on the probability of a regression from the critical word.

Method

Participants

A total of 112 students at the University of Massachusetts Amherst participated in the study in exchange for course credit. All were native speakers of English, had normal or corrected-to-normal vision, and were naïve to the purpose of the study. Data from 13 additional participants were discarded due to excessive track loss or poor calibration. The 112 subjects ran during three testing periods in different semesters, with 35 in the first group, 45 in the second, and 32 in the third. The only meaningful differences in the procedures for the three sets of subjects involved the filler materials with which the critical items were intermixed, as described below.

Materials

Sixty-two items like (3) and (4) above were initially created. An off-line norming study was then carried out. Between 17 and 19 naïve participants who did not participate in the eyetracking study rated the plausibility of each

---

[2] Relations between Bayes factors and model comparison measures such as the Akaike Information Criterion (AIC; Akaike, 1973) and the Bayesian Information Criterion (BIC; Schwarz, 1978) are discussed in Kass and Raftery (1995). Wagenmakers (2007) illustrates the calculation of the BIC from ANOVA output, and the calculation of a Bayes factor based on a difference in BIC. Nathoo and Masson (2015) extend this method to repeated measures designs, and also demonstrate how posterior distributions may be calculated based on the BIC; however, these authors recommend the Rouder et al. (2012) method for factorial designs.

[3] The BayesFactor package also enables Bayesian comparison of linear mixed effects models (e.g., Baayen, Davidson, & Bates, 2008) varying in both fixed and random effects structure, for continuous dependent measures; comparison of mixed-effects logistic regression models for binary dependent measures has not yet been implemented. We do not illustrate this analysis here, as the simple factorial design of the present experiment is well suited to a Bayesian extension of the traditional two-ANOVA strategy in psycholinguistics, separately considering subjects (F1) and items (F2) as random effects. See Singmann, Klauer, and Kellen (2014) for an instructive example of a Bayesian mixed-effect analysis in the domain of research on reasoning.

version of each item, with no participant rating more than one version of an item. Participants were provided with the sentence through the critical word, which was italicized, and were instructed to rate the fit of the italicized word in the sentence, with the following instructions: "If the word is a natural, sensible continuation of the sentence, rate it a 5. If the word makes no sense as a continuation of the sentence, rate it a 1." Based on the results of this norming session, 43 of the original 62 item sets were selected for inclusion as experimental items. (The items are available upon request.) Two inclusion criteria were imposed: the mean rating was required to be 3 or above for each of the plausible conditions, and below 3 for the implausible conditions; and within the item set, it was required that there was at least a full point difference in the mean rating of each of the two plausible conditions and each of the two implausible conditions. All 62 item sets were included in the eyetracking experiment, but the data from the 19 item sets that did not satisfy these criteria were never analyzed.

The HF and LF words were matched on length within each item, varying from 5 to 8 characters. Of the 43 critical items, 7 had a five-letter target, 13 a six-letter target, 17 a seven-letter target, and 6 an eight-letter target. The mean Subtlex (Brysbaert & New, 2009) frequency of the HF words in occurrences per million was 56.64, with mean log frequency of 3.17 ($sd = 1.06$; range from 1.82 to 6.46), while the mean for the LF words was 2.01, with mean log frequency of .55 ($sd = .76$; range from .02 to 3.74).[4] Within an item, the minimum frequency difference was 1.19 log units. The plausibility ratings for each of the four conditions were: HF, plausible ($M = 4.33$, $sd = .49$; range from 3.17 to 5.00); LF, plausible ($M = 3.94$, $sd = .48$; range from 3.06 to 4.94); HF, implausible ($M = 1.50$, $sd = .41$; range from 1.00 to 2.68); LF, implausible ($M = 1.52$, $sd = .31$; range from 1.06 to 2.24). Thus, while there was essentially no difference in the plausibility ratings for the HF and LF implausible conditions, there was a significant difference for the HF and LF plausible conditions ($t(42) = 4.32$, $p < .001$). This difference was expected, given that plausibility ratings for LF words are likely to be influenced by frequency itself, independent of plausibility. This interpretation of the relationship is reinforced by the finding that for the LF words, the plausibility ratings in the putatively plausible condition were positively correlated with log frequency, $r = .36$, $p = .019$. No such correlation was evident for the HF words in the plausible condition, $r = -.08$, $p = .60$, or for either the HF or LF words in the implausible condition (HF: $r = -.08$, $p = .60$; LF: $r = -.004$, $p = .98$).

The predictability of the target words in the 43 critical items was also assessed by means of a cloze task. Twenty Amazon Mechanical Turk workers, who were paid six USD, read the contexts preceding the target words, and were asked to enter the next word that came to mind as a continuation of the sentence. Each participant read one context from each item, providing 43 continuations overall, and each context was completed by 10 participants. Neither target word was ever entered in the implausible condition,[5] and the low frequency target was never entered in the plausible condition. The high frequency target was supplied eight times overall, for a mean cloze probability of 1.9%. However, five of these eight hits occurred in a single item ("The agent interrogated the *suspect*"), and for 40 of the 43 items the target was never provided.

The 62 original item sets were arranged into four lists so that each subject read fifteen or sixteen items in each of the four experimental conditions. Because of the exclusion of 19 of the original items from analysis, subjects read slightly unbalanced sets of the 43 critical items, with each subject reading at least 9 items in each condition. The 62 items were randomly intermixed with different unrelated filler items in each of the three experiment sessions. In the first session, there were 48 fillers, while there were 72 in the second and third sessions. These fillers were materials from other experiments; in the second and third sessions, but not the first, some of these filler sentences also involved implausibility (20 of the 72 filler items). In all three sessions, 42% of the filler sentences were followed by simple two-choice comprehension questions designed to ensure that participants were reading for comprehension. In the first session, all subjects achieved at least 71% correct on these questions. For the second and third sessions we do not consider performance on questions following relative clause fillers that were designed to be especially difficult; leaving these aside, all subjects achieved at least 80% correct.

Procedure

Subjects were tested individually. Eye movements were recorded using an EyeLink 1000 (SR Research, Toronto, ON, Canada) eyetracker, interfaced with a PC computer. The sampling rate was 1000 Hz. Subjects were seated 55 cm from a CRT monitor on which the sentences were displayed. At this distance the resolution of the eyetracker was substantially less than one character. Only the movement of the right eye was recorded.

All critical sentences were displayed on a single line, in 12-point Monaco font in session 1, and in 11-point Monaco font in sessions 2 and 3. Before the experiment began, each subject was instructed to read for comprehension, though it was noted that some sentences might be 'a little weird' and that subjects should try to understand these as well as possible. A calibration procedure was performed at the start of the experiment and as needed between trials. The subject triggered each sentence by fixating a box at the left edge of the monitor. The experiment lasted approximately 30 min. The experiment was implemented with the EyeTrack software, and initial stages of data analysis were carried out with EyeDoctor and EyeDry (http://blogs.umass.edu/eyelab/software/).

---

[4] Note that because Subtlex frequency is reported in occurrences per million words, values for LF words may be less than 1, which results in a negative log. We add 1 to the Subtlex frequency prior to taking the log, resulting in positive log values.

[5] Due to an error in the script, the data for one implausible context ("The teacher motivated the____") had to be discarded. We think it is safe to assume that the target words *cabinet* and *divider* would have 0 cloze probability here, as in the 42 other implausible contexts.

Track loss or other error resulted in deletion of 6.3% of trials, leaving 4512 trials for inclusion in the analysis. Individual eye fixations less than 80 ms in duration and within one character of a previous or subsequent fixation were incorporated into this neighboring fixation. Other fixations of less than 30 ms in duration were deleted. No other data trimming was carried out.

Results

Analyses focused on the critical noun in each sentence. We computed first fixation duration, gaze duration, and go-past time. We also determined for each trial whether the reader skipped over, rather than fixated, the word during first pass reading, and whether the reader made a regressive eye movement out of the word rather than exiting it to the right on first pass. We organize our analyses according to the time-course, beginning with skipping probability.

Grand means and standard errors for each measure are presented in Table 1; the data for the binomial dependent measures are graphically depicted in Fig. 1 (and in more detail for skipping in Fig. 2) and for temporal measures in Fig. 3. To test for the statistical reliability of the effects of frequency, plausibility, and their interaction, we performed two types of analyses: analyses of variance (ANOVAs) with participants (F_1) and items (F_2) as random effects (see Table 2); and computation of Bayes factors for each of the nested linear models in the ANOVA design compared to a random-effects only model (see Table 3). For reading time measures, ANOVAs were conducted over participant/item means, while for the binomial measures, ANOVAs were conducted over logit-transformed proportions (adding .5 to both numerator and denominator counts; e.g., Barr, 2008).

All analyses were conducted in the R Environment for Statistical Computing (Version 3.1.3; R Core Team, 2015), and Bayesian analyses were carried out using the BayesFactor package (Version 0.9.11-1; Morey et al., 2015). For each set of data submitted to an ANOVA, we fit an anovaBF object, using 100,000 Monte Carlo iterations. Our initial fits used the default scale value of ½ for the Cauchy priors on effect size; below we also report fits with alternate priors. The marginal likelihood of each nested model in the ANOVA design is compared to that of the random effects only model, producing the Bayes factors presented in Table 3.

Table 1
Grand means for eye movement measures on the target word, by condition. Duration measures are in ms. Standard error of the mean is in parentheses. *FFD* = first fixation duration, *GZD* = gaze duration.

	Skipping	FFD	GZD	Go-past	Regressions
Plausible					
HF	.088 (.008)	234 (2.4)	274 (3.8)	345 (6.8)	.172 (.012)
LF	.050 (.007)	250 (2.7)	334 (5.5)	420 (8.1)	.185 (.012)
Implausible					
HF	.080 (.008)	243 (2.5)	294 (4.1)	391 (8.1)	.202 (.013)
LF	.048 (.006)	257 (3.0)	340 (5.7)	461 (9.5)	.220 (.013)

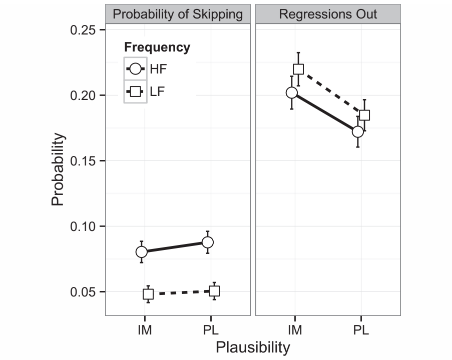

Fig. 1. Grand means and standard errors for binomial dependent measures on the target word.

Skipping probability

The target word was skipped only 6.7% of the time. As expected, the ANOVAs revealed a significant effect of frequency on the probability of skipping, which was 8.4% for HF and 4.9% for LF target words. Neither the effect of plausibility nor the interaction effect approached significance.

The Bayes factors computed for each nested model compared to the null "random-effects only" model favor a model including frequency but no other factor. The Bayes factor for this model compared to the null model is 175 by subjects and 1933 by items; the likelihood of the data under the by-subjects model including a frequency effect is 175 times that of the likelihood under the null model. On the other hand, the Bayes factor associated with the plausibility model is .11 by subjects and .02 by items. These values below 1 indicate that the likelihood is greater under the null model than under one that assumes only an effect of plausibility. The reciprocal of this Bayes factor gives the odds in favor of the null model, which are approximately 9:1 in the by-subjects analysis.

Since the nested models in the ANOVA design share a common denominator in their Bayes factor calculations (i.e., all are compared against the null model), we can also directly compare the various models; the Bayes factor for the comparison of any two of these models is simply the ratio of their Bayes factors in the comparison with the null model. For example, we can ask whether there is any evidence in favor of a model including additive frequency and plausibility effects compared to the frequency-only model. This is not the case: The Bayes factor for this comparison is .11 by subjects and .37 by items; by taking the reciprocal, we see that the frequency-only model of the by-subjects data is favored by 9:1 odds.

One possible concern is that because the rate of word skipping was fairly low overall, due to the fact that the target words were relatively long, a small effect of plausibility on skipping may be difficult to detect, as there were

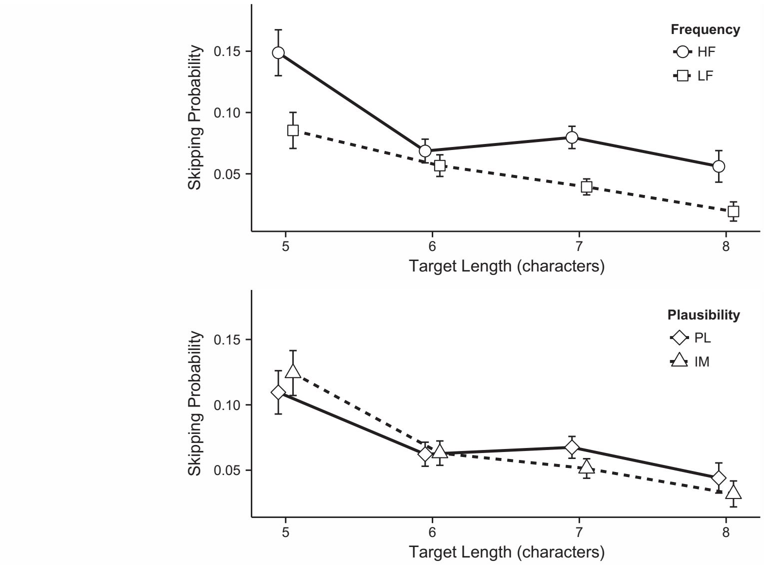

Fig. 2. Grand means and standard errors for skipping probability by word length and frequency (top panel) and plausibility (bottom panel).

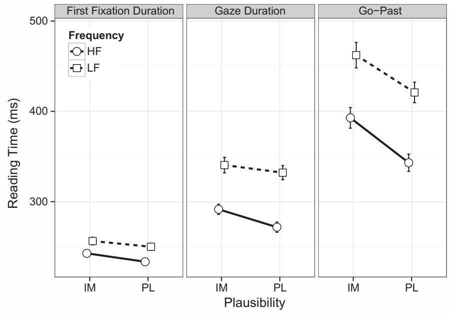

Fig. 3. Grand means and standard errors for fixation time measures on the target word. (For first fixation duration and gaze duration, error bars are obscured by the plotting symbols.)

Table 2
Analyses of variance with random error by subjects (F_1) and by items (F_2) for the dependent measures on the target word.

Variable	$F_1(1,111)$	p	$F_2(1,42)$	p
Skipping probability				
Frequency	16.00	<.001	17.80	<.001
Plausibility	.05	.82	1.75	.19
Freq × Plaus	.12	.73	.14	.71
First fixation duration				
Frequency	35.14	<.001	23.56	<.001
Plausibility	7.52	<.01	7.53	<.01
Freq × Plaus	.33	.57	.36	.55
Gaze duration				
Frequency	129.60	<.001	38.65	<.001
Plausibility	9.32	<.01	10.46	<.01
Freq × Plaus	1.57	.21	4.42	.04
Go-past time				
Frequency	65.99	<.001	28.82	<.001
Plausibility	27.59	<.001	15.64	<.001
Freq × Plaus	.29	.59	.15	.70
Regression probability				
Frequency	.81	.37	3.05	.09
Plausibility	4.46	.04	3.06	.09
Freq × Plaus	.06	.81	.03	.87

few skips even in the plausible conditions. Fig. 2 illustrates, however, that though the target words ranged only from five to eight letters, there was a clear effect of word length on skipping probability, and HF words were more likely to be skipped than were LF words at every level of word length. Thus, the experiment had the power to detect not only the effect of frequency on skipping, but also an effect of word length that obtains even within the restricted range of lengths that were used. On the other hand, while there is a trend toward more skipping of plausible than implausible 7- and 8-letter words, there is no such trend for 6-letter words, and for 5-letter words the trend actually goes in the opposite direction. In other words, the lack of any plausibility effect on skipping is in fact clearest for shorter words, which are skipped more often.

Table 3
Bayes factors for each of the nested models in the ANOVA design compared to the by-subjects or by-items random effects model. Values greater than 10,000 are expressed in scientific notation.

Model	Skipping probability	First fixation	Gaze duration	Go-past	Regression probability
Freq					
Subj	175.5	2.31×10^5	1.19×10^{23}	2.02×10^{12}	.16
Item	1933.5	2.41×10^4	4.07×10^{12}	3.02×10^6	.41
Plaus					
Subj	.11	4.62	2.38	6187.4	.69
Item	.02	4.63	.69	24.3	.01
Freq + Plaus					
Subj	18.98	1.54×10^6	1.00×10^{24}	1.26×10^{17}	.12
Item	714.9	1.99×10^5	6.20×10^{12}	2.87×10^8	.01
Freq * *Plaus*					
Subj	2.85	2.58×10^5	2.92×10^{23}	2.04×10^{16}	.02
Item	1.07	5.14×10^4	3.35×10^{12}	6.70×10^7	.02

First fixation duration

The ANOVAs revealed significant effects of both word frequency and plausibility on first fixation duration, by subjects and by items. The interaction between frequency and plausibility did not approach significance. As in the analysis of the skipping data, the Bayes factors for each of the nested models in the design compared to the null model reflect the patterns of significance in the ANOVAs. For example, the Bayes factor is 4.6 or greater for each of the individual fixed effects models compared to the null model. Given that the frequency effect is numerically larger than the plausibility effect, it is unsurprising that the Bayes factors for the frequency model are much larger than for the plausibility model. The additive and interactive effects models also have very high Bayes factors. However, our main question with regard to these temporal measures is whether the data are better accounted for by a model that includes a frequency × plausibility interaction term, or by an additive model. We directly compare the additive versus interactive models by taking the ratios of their Bayes factors. The Bayes factor for this comparison is 5.97 by subjects and 3.88 by items, favoring the additive model.

Gaze duration

The pattern of effects on gaze duration, as revealed by ANOVAs, largely mirrors the pattern for first fixation duration. There were significant effects of both frequency and plausibility by subjects and by items. The interaction did not approach significance by subjects, but was significant by items ($p = .04$). In the Bayesian by-subjects analysis, the additive model was the clearly favored model, by a Bayes factor of 8.34 over the frequency-only model, and by a Bayes factor of 3.44 over the interactive model. In the by-items analysis, the Bayes factor in favor of the additive model over the frequency-only model was 1.54, and the Bayes factor in favor of the additive model over the interactive model was 1.85. Thus, while both the by-subjects and by-items Bayesian analyses favor the additive model, this conclusion is clearly stronger in the case of the by-subjects analysis.

The by-items Bayesian analysis, which provided equivocal evidence for the additive model over the frequency-only model, on the one hand, and the interactive model, on the other, may be regarded as inconsistent with the patterns of significance revealed by ANOVA, which revealed both a significant plausibility effect and a significant interaction. This apparent discrepancy, however, simply demonstrates that a rejection of the null, based on $p < .05$, may occur even when the Bayes factor reveals only anecdotal evidence against the null, or even evidence in favor of the null (e.g., Rouder & Morey, 2011; Wetzels et al., 2011).

Go-past time

The results for go-past time were similar to first fixation duration: there were clear additive effects of both frequency and plausibility, with no significant interaction. The Bayesian analysis supports the additive relative to the interactive effects model, with Bayes factors by subjects and items of 6.20 and 4.28, respectively.

Regressions out

There was a significant effect of plausibility on regression probability by subjects ($p = .04$), which was marginal by items ($p = .09$), and a marginal effect of word frequency on regression probability by items ($p = .09$), though not by subjects ($p = .37$). The interaction between the two factors never approached significance. The Bayesian analysis clearly favored the null model over the model that included a frequency effect. The Bayes factors for the comparison of the null model and the plausibility model were close to 1, and therefore indecisive between the two models.

Alternate priors

A common concern regarding Bayesian analyses is their dependence on specific assumptions regarding prior distributions. The value of a Bayes factor will depend on precisely how the priors for non-null models are specified. As noted above, the BayesFactor package assumes a

Table 4
Bayes factors for each of the nested models in the ANOVA design compared to the by-subject random effects only model for three measures, using three Cauchy prior scale parameters, corresponding to prior distributions on the effect sizes that are "narrow" (.25), "medium" (.5), or "wide" (1.0). Note that models with the medium scale setting were re-fit for this demonstration, resulting in small differences in values from Table 3.

Model	Skipping probability			First fixation			Regression probability		
	Narrow	Medium	Wide	Narrow	Medium	Wide	Narrow	Medium	Wide
Freq	218.9	177.74	106.80	2.29×10^5	2.30×10^5	1.58×10^5	.30	.16	.08
Plaus	.20	.11	.05	6.90	4.71	2.59	1.18	.69	.37
Freq + Plaus	45.30	19.03	5.75	2.23×10^6	1.52×10^6	5.75×10^5	.35	.12	.03
Freq + Plaus + Freq * Plaus	12.83	2.87	0.45	6.58×10^5	2.54×10^5	4.97×10^4	.10	.02	.00

Cauchy prior on effect size. The Cauchy distribution is a *t* distribution centered on a value of 0, so relatively small effect sizes will have high probability density compared to large effect sizes. The scale of the Cauchy distribution can be adjusted to assess the impact of expectations for a relatively larger or relatively smaller effect size on the resulting Bayes factor, which we illustrate by manipulating the scale parameter by a factor of 2 in either direction (the default scale is ½). A scale of ¼ adjusts the spread of the Cauchy prior so that even more probability density is associated with very small effects (so, if an effect is small, the Bayes factor will more strongly favor the non-null model); a scale of 1 adjusts the spread so that more probability is associated with larger effects (so, if an effect is small, the null model will be more strongly preferred). Table 4 illustrates these patterns for the by-subjects analysis of skipping probability, first fixation duration, and regression probability. It is clear that the critical results from the present study do not depend on the spread of the Cauchy prior. The frequency-only model is always far-and-away the preferred model of the skipping data; the additive model is always clearly the preferred model of the first fixation data; and for the regressions data, the Bayes factor is never decisive between the null model and the plausibility-only model.

Posterior distributions

Finally, it is possible to estimate the posterior probability distribution for each of the parameters in a model by means of MCMC sampling using a Gibbs sampler (Geman & Geman, 1984), again within the BayesFactor package. This full posterior distribution may then be directly inspected, allowing inference regarding likely and unlikely values of each parameter.

We illustrate these posterior distributions for the parameters of the by-subjects interactive first fixation duration model, using the default Cauchy scale, based on Markov Chain Monte Carlo (MCMC) simulation with 100,000 iterations (see http://bayesfactorpcl.r-forge.r-project.org/#mixed for details of this routine in R). We do not discard burn-in samples. Table 5 provides the means and standard deviations for the posterior distributions for each parameter of the model, as well as the boundaries of the 95% highest density interval (HDI) for each parameter. The posterior distributions for the three critical effects are shown in Fig. 4. (For simplicity, we plot only the posterior for the effect of an LF word, as the posterior for an HF word is a mirror image on the opposite side of 0, and similarly for the plausibility and interaction effects.)

Table 5
Mean and *SD* of posterior distribution for each parameter of interactive effects model of first fixation duration (by subjects), based on MCMC sampling, and boundaries of 95% highest density interval.

Variable	Mean	*SD*	2.5% Quantile	97.5% Quantile
Mu	245.70	3.25	239.31	252.09
Freq-HF	−7.38	1.33	−9.98	−4.77
Freq-LF	7.38	1.33	4.77	9.98
Plaus-IM	3.80	1.32	1.23	6.40
Plaus-PL	−3.80	1.32	−6.40	−1.23
Freq:Plaus-HF & IM	.70	1.29	−1.82	3.25
Freq:Plaus-HF & PL	−.70	1.29	−3.25	1.82
Freq:Plaus-LF & IM	−.70	1.29	−3.25	1.82
Freq:Plaus-LF & PL	.70	1.29	−1.82	3.25

Discussion

The aim of the present study was to assess whether the effects of plausibility on eye movements in reading may be viewed as post-lexical effects, even though they appear in measures as early as first fixation duration on the target word. To address this question, we considered specific empirical predictions that emerge from the E-Z Reader 10 model (Reichle et al., 2009), which treats plausibility effects as arising post-lexically. E-Z Reader predicts that: (a) frequency but not plausibility should influence word skipping; (b) plausibility may influence the probability of a regression from a word, but frequency should not; (c) frequency and plausibility should always display additive effects on reading time measures (i.e., null interactions). Because conventional NHST cannot provide support for hypotheses of null effects or null interactions, we computed Bayes factors to determine the degree of evidence favoring null or alternative hypotheses.

The experiment confirmed that frequency, but not plausibility, influences the probability that a word will be skipped. There was a higher skipping rate for HF compared to LF words regardless of the word's plausibility. Bayes factors substantially favored a frequency-only model of skipping probability over a model that also included a plausibility effect. This result is consistent with the previous finding (Angele & Rayner, 2013; Angele et al., 2014) that readers tend to skip a short, high-frequency parafoveal

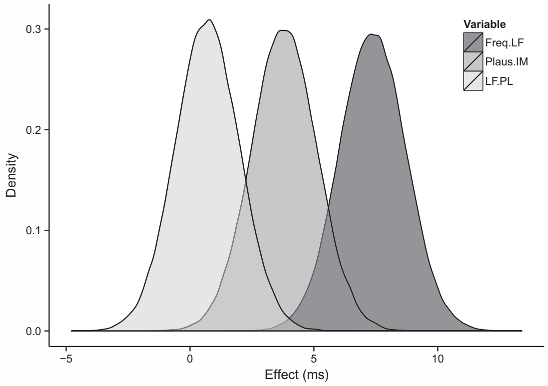

Fig. 4. Posterior distributions of the effects of frequency and plausibility and their interaction on first fixation duration estimated through MCMC sampling using BayesFactor.

word regardless of whether it is a grammatical or plausible sentence continuation.

Frequency and plausibility exerted additive effects on all three first-pass reading time measures: first fixation duration, gaze duration, and go-past time. With the exception of the by-items gaze duration analysis, an additive model was favored over an interactive model by a Bayes factor of greater than 3; in the by-items gaze duration analysis the additive model was still favored, but only weakly.

The use of Bayes factors allows for a clear advance over the inferential strategy in Staub (2011a). Staub conceded, for example, that the failure to obtain significant interactive effects on reading time measures between word frequency and syntactic attachment difficulty could not directly support the hypothesis of no interaction. By contrast, the Bayes factors in the present study provide positive evidence for a null effect of plausibility on skipping and null interactions between plausibility and frequency on reading time measures.

The presence of a clear plausibility effect on first fixation duration, together with evidence of an absence of an effect on skipping probability, confirms a critical dissociation predicted by E-Z Reader 10. In the E-Z Reader framework, word skipping occurs when an initial stage of lexical identification (L1) is completed prior to the word being directly fixated. The duration of the L1 stage is influenced by both frequency and predictability, but not, in the E-Z Reader framework, by plausibility. Even though the earliest reading time measures may be influenced by plausibility, the present study shows that the very earliest measure of all – the probability that the eyes skip over the word rather than directly fixating it – is not. Thus, this study provides evidence that word frequency influences eye movement control before plausibility does.

It remains to be seen how an eye movement model such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005), in which multiple words may be processed in parallel, would handle the dissociation between lexical and post-lexical processing that E-Z Reader appears to account for. At present, SWIFT makes no attempt to account for the influence of sentence-level processing on eye movements in reading (Schad & Engbert, 2012). SWIFT does provide an account of predictability effects, but like E-Z Reader, SWIFT treats predictability as influencing lexical processing itself. Effects of syntactic processing difficulty and semantic integration difficulty have not been discussed in the context of the SWIFT model.

Effects of both frequency and plausibility on the probability of a regression were small (i.e., a 1.5% difference between the means of the two frequency conditions and a 3.3% difference between the means of the plausibility conditions). While the Bayes factor analysis provided the expected evidence against a frequency effect, it provided no clear evidence either for or against a plausibility effect.

Inspection of previous plausibility studies shows that while there is a clear trend across studies for an effect of plausibility on regressions, the effects are rarely statistically significant. Staub et al. (2007) reported a non-significant trend, and Warren and McConnell (2007) reported a trend that was significant by subjects ($p = .05$), but not by items. Rayner et al. (2004) did not report regression rates. The simulations reported by Staub (2011a) showed that E-Z Reader 10 predicts increased regressions due to integration difficulty only with very specific values of the model's two integration parameters: the probability of outright integration failure must be relatively high, but time to complete the integration process must be short. Thus, the lack of a clear plausibility effect on the probability of a regression is not itself problematic for E-Z Reader. However, it is important to note that E-Z Reader predicts that integration difficulty will affect reading times on the target word under just the same set of parameter settings that result in an effect on the probability of a regression. Thus, the present data pattern, with clear plausibility effects on reading times but an unclear effect on the regression rate, is somewhat unexpected.

This apparent discrepancy raises an important point about the interpretation of Bayes factors. A Bayes factor is the ratio of the marginal likelihood of the data under two models. As noted in the Introduction, if the two models have prior odds of 1 (i.e., the two models are regarded as equally probable, in advance of the data), the posterior odds will be identical to the Bayes factor itself, as the posterior odds are simply the product of the Bayes factor and the prior odds. But if the two models are not regarded as equally likely in advance, the posterior odds are not identical to the Bayes factor. For example, Rouder and Morey (2011) determined that the Bayes factor in favor of one of the ESP hypotheses in Bem (2011) is about 40. However, given that the prior odds for this ESP hypothesis, compared to the null, should arguably be very small indeed, even a 40-fold increase in these odds may not be nearly enough for the posterior odds actually to favor ESP. In the present context, a Bayes factor near 1 for the test of the plausibility effect on regressions will results in similar posterior odds only if the prior odds are also near 1. If, on the other hand, a researcher concludes from the previous literature that there is likely to be a real effect of plausibility on the probability of regression – with odds of 3-to-1, say – the equivocal results from the present study will have the effect of simply carrying forward these prior odds, resulting in posterior odds of, still, 3-to-1.

Conclusion

The present study addressed the question of whether the effects of plausibility on eye movements in reading may be attributed to post-lexical processing, despite their early time course. The answer is yes: The detailed data pattern in an experiment manipulating both word frequency and plausibility was consistent with plausibility having its effect at a post-lexical integration stage, as proposed by E-Z Reader 10 (Reichle et al., 2009). First, while word frequency influenced the probability that a word was skipped, the plausibility of the word in context did not influence this probability. Second, the two variables had additive, not interactive, effects on reading time measures. These conclusions were based on computation of Bayes factors, using the default priors and sampling methods developed by Rouder et al. (2012). The present study strongly supports the recent trend in experimental psychology toward the use of Bayes factors as a means of providing direct support for null hypotheses, when they are of theoretical importance.

Acknowledgments

Matthew Abbott's work on this project was partially funded by a gift from Microsoft to Keith Rayner. Thanks to Brian Dillon and Titus von der Malsburg for helpful feedback, to Sarah White and two anonymous reviewers for helpful comments on an earlier version of this manuscript, and to Richard Morey and Jeff Rouder for technical help regarding the BayesFactor package.

References

Abbott, M. J., Angele, B., Ahn, Y. D., & Rayner, K. (2015). Skipping syntactically illegal "the"-previews: The role of predictability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. (Advance online publication). <http://dx.doi.org/10.1037/xlm0000142>.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *2nd international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Angele, B., Laishley, A., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews on and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1181–1203.

Angele, B., & Rayner, K. (2013). Processing "the" in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 649–662.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: 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. http://dx.doi.org/10.1016/j.jml.2007.12.005.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology, 100*, 407.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 954–969. http://dx.doi.org/10.1037/0096-1523.31.5.954.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review, 70*, 193–242.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6*, 78–84. http://dx.doi.org/10.1016/S1364-6613(00)01839-8.

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review, 116*, 439–453.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 721–741.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York: Oxford University Press.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). Continuous univariate distributions (Vol. 2, 2nd ed.). New York: Wiley.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*, 773–795.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203–205.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de) Constructing the N400. *Nature Reviews Neuroscience, 9*, 920–933.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science: A Multidisciplinary Journal, 29*, 375–419. http://dx.doi.org/10.1207/s15516709cog0000_25.

Matsuki, K., Chow, T., Hare, M., Elman, J. L., Scheepers, C., & McRae, K. (2011). Event-based plausibility immediately influences on-line language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 913–934.

Morey, R. D., Rouder, J. N., & Jamil, X. (2015). *BayesFactor: Computation of Bayes factors for common designs. Version 0.9.11*. <http://bayesfactorpcl.r-forge.r-project.org/>.

Nathoo, F. S., & Masson, M. E. (2015). Bayesian alternatives to null-hypothesis significance testing for repeated-measures designs. *Journal of Mathematical Psychology*. (Advance online publication). <http://dx.doi.org/10.1016/j.jmp.2015.03.003>.

R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–164.

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

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 514–528.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1290–1301.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*, 504–509.

Reichle, E. D. (2011). Serial attention models of reading. *Oxford Handbook on Eye Movements*, 767–786.

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., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–476. http://dx.doi.org/10.1017/S0140525X03000104.

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*, 1–21. http://dx.doi.org/10.3758/PBR.16.1.1.

Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review, 18*, 682–689.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356–374. http://dx.doi.org/10.1016/j.jmp.2012.08.001.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*, 391–421.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Singmann, H., Klauer, K. C., & Kellen, D. (2014). Intuitive logic revisited: New data and a Bayesian mixed model meta-analysis. *PLoS One, 9*, e94223.

Staub, A. (2011a). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General, 140*, 407–433.

Staub, A. (2011b). The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin & Review, 18*, 371–376.

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun–noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1162–1169.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1280–1293.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problem of p values. *Psychonomic Bulletin & Review, 14*, 779–804.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review, 14*, 770–775.

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1001–1010.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology an empirical comparison using 855 t tests. *Perspectives on Psychological Science, 6*, 291–298.

Chapter 4, in full, is a reprint of the material as it appears in The effect of plausibility on eye movements in reading: Testing E-Z Reader's null predictions. Abbott and Staub (2015). The dissertation author was the primary investigator and author of this paper.

Chapter 5

General Discussion

In Chapter 1, I introduced three testable assumptions regarding language processing during reading and its impact on eye movements. These predictions were largely derived from the E-Z Reader 10 model (Reichle, Warren, & McConnell, 2009). I tested the extent to which information from the sentence context versus information from the parafovea influences word skipping, the relative importance of parafoveal processing versus oculmotor constraints to this behavior, and the serial nature of word identification and postlexical integration. Now I will discuss what we can learn from the studies presented in Chapters 2–4, how the results support the assumptions we made in Chapter 1, and the implications for computational models of eye movement control.

5.1 Skipping as a hedged bet

The experiment presented in Chapter 2 used the boundary paradigm (Rayner, 1975) to test the extent to which word predictability and the ease of parafoveal identification influence word skipping relative to deeper syntactic or semantic processing. Specifically, I tested whether readers would skip over infelicitous instances of *the* even when the target word (a different three-letter word) is predictable in the context. Our results demonstrate quite clearly that although skipping is influenced both by predictability and ease of identification, the skipping decision is made without regard to its fit within the syntactic context (following Angele, Laishley, Rayner, & Liversedge, 2014; Angele & Rayner, 2013). We found that false previews of *the* were skipped more often than correct previews of the target words, and found a similar effect of predictability in both invalid *the* and valid preview conditions. Readers skipped over instances of *the* even when it ultimately hindered later processing.

These results suggest that the benefit of predictability may not necessarily come about from preprocessing the word from the parafovea (which is a typical interpretation of the predictability effect; Balota, Pollatsek, & Rayner, 1985; Schotter, Angele, & Rayner, 2012). This seems odd in light of evidence that readers are less likely to skip over previews of visually similar non-words than correct previews (e.g., *livor* as a preview for *liver*; Balota et al., 1985; Drieghe, Rayner, & Pollatsek, 2005). Balota et al. (1985) first presented evidence that predictability and visual similarity interacted, such that the effect of visual similarity was smaller when the target words were predictable versus unpredictable. This would have been consistent with the view that readers could "fill in" conflicting details about the parafoveal word (i.e., a misspelling) based on the context, but Drieghe et al. (2005) did not replicate this finding. They found instead finding that visually similar non-word previews of predictable words were skipped at about the same rate as other types of non-word previews. The authors suggested that differences in the quality of visual displays across the twenty year gap between studies could account for the discrepancy in results (i.e., poorer displays may have prompted larger effects of words versus non-words on eye movements).

Instead, contextual constraint may reflect preactivation of a word (or perhaps words) based on the preceding context (Staub, Grant, Astheimer, & Cohen, 2015). Interestingly, this sort of preactivation appears to give a boost to processing (as reflected by increased skipping) even when the parafoveal word is not the preactivated word. This may only occur, though, when the parafoveal word is easier to process than expected. This would suggest that when processing is easy overall, the effect that predictability has on skipping is a baseline adjustment. Certainly, future research should aim to determine whether this behavior is somewhat unique to the word *the*. Chapter 3 is a first step into this territory.

5.2 Are short words treated differently?

Experiments 3.1 and 3.2 tested whether readers maintain a general preference to skip over words based on their ease of identification, rather than their fit within the context, or whether the effects we found in Chapter 2 are unique to very short words like *the* due to oculomotor constraints like saccadic range error (McConkie, Kerr, Redix, & Zola, 1988). If oculomotor constraints cause very short words like *the* to be skipped over more frequently than longer words, we expected an interaction in skipping such that three-letter words would display a reduced frequency effect relative to longer (e.g., five-letter) words. We used the boundary paradigm to present readers with valid previews of three, four, or five-letter words, or invalid previews of easier to identify contextually infelicitous higher-frequency words (either content words of a different syntactic class in Experiment 3.1, or the function words *the*, *that*, or *there* in Experiment 3.2). We further demonstrated that readers do not engage in deep syntactic or semantic processing prior to skipping, as skipping was not significantly affected by the contextually infelicitous nature of the invalid preview word in either experiment.

In Experiment 3.1, readers skipped over the invalid previews at about the same rate as correct previews. In Experiment 3.2, readers skipped over previews of function words more often than lower-frequency words even when they were inappropriate in the sentence context. This was true not only for short words like *the* but also longer words like *there*. Critically, the size of the frequency effect on skipping was similar for three, four, and five-letter words (i.e., there was no significant interaction between frequency and length). This indicates the importance of word identification over oculomotor constraints with regard to word skipping. A topic of future research is to further examine the extent to which this behavior is unique to function words, and how soon the “error signal" is detected that an incorrect word has been skipped.

5.3 Lexical versus postlexical processing

That word skipping appears to be driven primarily by word identification, and is unaffected by the fit with the preceding context, suggests a constraint on the architecture of the language processing system: The partitioning of lexical and postlexical processing. Staub (2011) demonstrated that E-Z Reader 10 incorporates a staged architecture that separates these processes into serially sequential stages, which lends to a number of specific predictions for eye movements. The model predicts that lexical processing, but not integration, will affect skipping, that integration but not lexical processing may affect the probability of making a regressive eye movement, and that the effects due to the difficulty of word identification and integration may appear in the same eye fixation measure (i.e., gaze duration), but that when they do the effects will be additive (critically, not interactive). We tested this explicitly in Chapter 4 by manipulating word frequency and plausibility and assessing the evidence in favor of null main effects and interactions by computing Bayes factors (Edwards, Lindman, & Savage, 1963; Rouder, Morey, Speckman, & Province, 2012).

Our results confirmed the predictions made by E-Z Reader: Word skipping was influenced by frequency but not plausibility, and the effects of frequency and plausibility were additive across all temporal eye movement measures (regression probability was not convincingly influenced by either factor, broadly consistent with the plausibility literature). Critically, Bayes factors indicated probabilistic evidence supporting a null effect of plausibility on skipping, and null interactions between the two factors in eye fixation measures. These results are important because they indicate a constraint on eye movement control (that the eyes will begin to move on after a rough check of a word's identity and not much deeper processing) and language processing more generally (that higher-order processing does not modulate the speed of word identification).

The findings counter results reported in the event-related potential (ERP) literature demonstrating modulation of the N400 waveform according to plausibility (e.g., Kutas & Hillyard, 1980), which was interpreted as contextual modulation of lexical access (Lau, Phillips, & Poeppel, 2008). Since, as yet, no model of eye movement control also makes predictions about ERP waveforms, it is difficult to reconcile these results. Here the coregistration of eye movements and ERPs (see Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011) may be particularly beneficial. If eye movements displayed a consistently additive pattern while the N400 displayed a consistently interactive pattern in response to the same stimuli (i.e., factorial manipulations of frequency and plausibility as in Chapter 4), it may suggest that the two types of measures index subtly different processes. Understanding these types of relationships is broadly important for determining eye movements can tell us versus ERPs.

5.4 Implications for models of eye movement control

This dissertation is grounded in a serial attention shift view of eye movements in reading, more specifically the E-Z Reader model. Here I consider whether the results presented in Chapters 2–4 necessitate any revision to the assumptions raised in Chapter

1 that were derived from E-Z Reader, and whether these results would be equally accommodated by models like SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) that assume parallel word processing.

E-Z Reader

E-Z Reader 10 (Reichle et al., 2009) is the first eye movement model to incorporate postlexical processing and assumes that it strictly follows lexical access. The timing of lexical processing is influenced by word frequency, predictability, and length. Equation 1.1 reveals the model's assumption that the duration of the first stage of lexical processing is affected by these properties for a singular word. The results presented in Chapter 2 indicate that the benefit from predictability can be dissociated from that which comes from parafoveal preprocessing. This suggests that contextual constraint gives readers a boost in confidence that the upcoming word will be more easily identified. I suggest the model be modified so that the cloze probability of the most predictable word at position n influences early identification processes (and not necessarily the cloze probability of the parafoveal word n). The studies from Chapter 3 largely support the assumptions of E-Z Reader, that word identification, and not oculomotor constraints, drive word skipping. Indeed, it is assumed that the trigger that sends the signal to move the eyes forward is the completion of L1, the timing of which is influenced by the mean eccentricity of the letters of the attended word in addition to factors like word frequency. Finally, Chapter 4 largely confirmed E-Z Reader 10's assumptions, that word identification and postlexical integration are sequential processes. The model may overestimate the impact that integration difficulty has on regression probability, but our results overall are in concert with the model's predictions.

Comparisons with other models

It is also important to consider whether the results reported in this dissertation are compatible with other models of eye movement control in reading, namely parallel processing models like SWIFT (Engbert et al., 2005). Parafoveal processing is the main cause of word skipping in SWIFT (Schad & Engbert, 2012), and so this model could likely predict the results that were presented in Chapter 2 (greater skipping of *the* even when inappropriate in context). As SWIFT assumes that the rate of lexical processing is affected by visual acuity (similar to the influence of eccentricity on processing rate in E-Z Reader), SWIFT could also probably account for the preferential skipping of function words presented in Experiment 3.2, and perhaps for the additive pattern of word length and function word skipping.

However, it is unlikely that SWIFT could account for the fully array of results presented in this dissertation, particularly those that are assumed to have been due to postlexical processing. SWIFTs architecture only directly incorporates factors which affect the rate of lexical processing such as frequency, predictability, and visual acuity. The most recent version of the SWIFT model (Schad & Engbert, 2012) incorporates a zoom-lens architecture in which the size of the attention span expands or contracts along with the difficulty of foveal (i.e., lexical) processing. There is, however, a parameter that reduces processing rate by a constant factor “during postlexical processing" (p. 397; Schad & Engbert, 2012), although the authors do not make explicit what factors influence this parameter other than foveal processing difficulty. The authors report that this version of SWIFT can correctly predict certain differences between the reading of normal and shuffled text (namely “slowed deactivation of words in shuffled-SWIFT”, p. 415), although no comment is made with regard to factors like plausibility or syntactic fit.

The critical difference, then, between E-Z Reader and SWIFT with regard to any postlexical processing is that in E-Z Reader it is quite explicit that the factors which influence postlexical integration are not the same as those that influence lexical processing. If postlexical processing in SWIFT is essentially a function of the difficulty of lexical processing, then it seems unlikely that it could fully account for the results presented in this dissertation, especially those in Chapter 4 with regard to frequency and plausibility. Future research using the SWIFT model is necessary in order to determine whether a parallel processing architecture makes similar predictions to E-Z Reader with regard to postlexical processing and its relationship to word identification.

5.5 Conclusion

In this dissertation I reported a number of results concerning the impact of parafoveal processing on early word identification, the extent to which readers can recruit information from the context prior to word skipping, and the staged nature of word identification and postlexical processing. Overall, the results of these experiments indicate that parafoveal processing is limited to a hedged bet that recognition is imminent, and that any deeper processing pertaining to syntactic or semantic integration is delayed to a later processing stage. This is consistent with the idea that parafoveal processing involves a rough check of the identity of the upcoming word, and does not necessarily reflect full identification. Interestingly, what contributes to the speed of parafoveal processing may not depend entirely on the identity of the parafoveal word, but the relative confidence that the upcoming word will be easy to identify.

References

Angele, B., Laishley, A. E., Rayner, K., & Liversedge, S. P. (2014). The effect of high- and low-frequency previews and sentential fit on word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1181-1203.

Angele, B., & Rayner, K. (2013). Processing the in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 649–662.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*(3), 364–390.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and eeg in natural reading: Analyses and review. *Journal of Experimental Psychology: General*, *140*, 552-572.

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 954-969.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review*, *70*, 193-242.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). Swift: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777-813.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203-205.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de) constructing the n400. *Nature Reviews Neuroscience*, *9*, 920-933.

McConkie, G. W., Kerr, P. W., Redix, M. D., & Zola, D. (1988). Eye movement control during reading: I. the location of initial eye fixations on words. *Vision Research*, *28*, 1107-1118.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81. doi: 10.1016/0010-0285(75)90005-5

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*, 1-21.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default bayes factors for anova designs. *Journal of Mathematical Psychology*, *56*, 356-374.

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the swift model. *Visual Cognition*, *20*, 391-421.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, and Psychophysics*, *74*, 5-35.

Staub, A. (2011). Word recognition and syntactic attachment in reading: Evidence for a staged architecture. *Journal of Experimental Psychology: General*, *140*, 407-433.

Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, *82*, 1-17.