This is a preprint of an article whose final and definitive form was published in the Journal of Experimental Psychology: Learning, Memory, and Cognition.

DOI: 10.1037/xlm0000142

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

Matthew J. Abbott[1], Bernhard Angele[2], Y. Danbi Ahn[1], and Keith Rayner[1]

[1] Department of Psychology, University of California, San Diego

[2]Psychology Research Centre, Faculty of Science and Technology, Bournemouth University

Correspondence to:

Matthew J. Abbott

Department of Psychology, University of California, San Diego

9500 Gilman Dr., MC 0109

La Jolla, CA 92093-0109

Email: mabbott@ucsd.edu

Word count: 6,071 words in text and references

Running Head: Word predictability and skipping

Abstract

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.

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). While 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, syntactic context information, 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 night," the parafoveal preview information available to readers about the word *dim* could be (1) the word *dim* itself as a control condition, (2) a random letter preview (*fda*), or (3) 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 in the sentence context. On the other hand, if readers only rely on the parafoveal information to guide their skipping decision without taking the sentence syntactic context into account, they should be quite likely to skip the article *the* since 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 to 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 syntactic context has little effect on word skipping. However, this conclusion is in direct contradiction to the many experiments described above that 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 context largely did not constrain the target word, which might explain the absence of sentence context effects. The present study tested this hypothesis as we manipulated both the 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 UC 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 eyetracker. Sentences were displayed in 14pt 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 x 2 preview). Each sentence included a 3-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). Pre-target 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, twenty workers recruited through Amazon Mechanical Turk participated in a cloze norming task in exchange for payment. They were asked to report what they thought the next word in the sentence should be after reading the sentences up until to the target word (e.g., “Jane used the scissors to carefully ___”). 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 forty experimental sentences were embedded among sixty filler sentences. Each sentence was presented on the computer screen individually. Following a three-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 4ms on average (range 0-7ms) 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 pre-target word, the target word, and the post-target word. 4.3% of trials were eliminated due to blinks or track loss on the target word or on adjacent fixations, as were 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%). 3.7% of fixations shorter than 80 ms or longer than 1000 ms were eliminated. Comprehension accuracy was high (97% correct).

We report linear mixed models (LMMs) with subjects and items as crossed random effects (Baayen, Davidson, & Bates, 2008). These analyses were chosen instead of ANOVAs due to the uneven cell sizes that result from word skipping, where 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 where 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 α-level (*t*-values are interpreted in the same way). LMM results are presented in Tables 2, 4, and 6.

Pre-target word

Condition means for the pre-target word are presented in Table 1. Readers fixated on the pre-target 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 *z*s < .65). There was a small but reliable effect such that single fixation durations on the pre-target word (84% of first-pass fixations) prior to a skip were shorter than prior to making a fixation on the target (210ms vs. 217ms; $b = 0.02$, 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. Gaze duration showed a similar pattern with shorter fixations prior to skipping (234ms vs. 239ms), but this effect did not reach significance ($t = 0.05$).

Target word skipping

Condition means for the target word skipping rates are presented in Table 3 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' single fixation durations (41% of first-pass fixations) were shorter following correct target word previews than *the* previews (228ms vs. 248ms; b = -0.03, SE = 0.02, t = -2.06). The same pattern was present in gaze duration, with shorter fixations following valid previews (234ms vs. 257ms; 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 gaze durations (t = 0.07), and constraint did not interact with preview in this measure (t = -0.34). Go-past times were also shorter following correct vs. *the* previews (262ms vs. 288ms; b = -0.04, SE = 0.02, t = -2.53). In go-past time 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 ($z = -0.55$), and there were no significant interactions amongst any of the fixed factors (all *z*s < 1.41).

Post-target word

Condition means for the post-target word are presented in Table 5. It is clear that skipping over the target word (versus fixating it) disrupted processing on the post-target word, with longer fixations following a skip in single fixation duration (61% of first-pass fixations; 232ms vs. 207ms), gaze duration (274ms vs. 229ms), and go-past time (381ms vs. 258ms). 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 to valid target word previews. See Figure 3 for a depiction of the nature of this interaction in go-past time. In addition to affecting fixation times on the post-target word, skipping affected regression rates out of this region; readers made more regressions out of the post-target 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 compared to 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 to 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 non-words 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: (1) a word is "guessed" such that the duration of *L1* is set to 0ms with probability equal to its cloze predictability; (2) the 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: (1) *the*-skipping is driven primarily by its frequency of occurrence (i.e., there is no systematic effect of target word constraint on *the*-skipping); (2) *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 as compared to 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 to 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 to 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. While 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.

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 versus 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 this for this difference.

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 3-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 where the target word was highly constrained compared to 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 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: while readers spent longer fixating the post-target region when it was skipped compared to 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: (1) 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); (2) 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. While 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. Since 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 since 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 et al. (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.

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) is unique to very short (e.g., 3-letter) words or generalizes to longer words. It may be that lexical properties like word frequency play a larger role in skipping shorter (e.g., 3-letter) compared to longer (e.g., 5-6 letter) words. There is some direct support for this idea within the E-Z Reader model, since the rate of parafoveal processing is limited by the mean eccentricity of the parafoveal word from the point of fixation. Since 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., & Rayner, K. (2013). Processing "the" in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 649–662.

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*, in press.

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

Bates, D., Maechler, M., & Bolker, B. (2013). Lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-2. 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.

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.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20,* 641–655.

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

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.

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

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.

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review, 12,* 132–138.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*(17), 2201-2217.

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

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.

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.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

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

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*, *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., 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 older and younger readers. *Psychology and Aging*, *21*, 448-465.

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.

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., & 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., & Drieghe, D. (2013). Using E-Z Reader to examine word skipping during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1331-1320.

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., Pollatsek, A., & Rayner, K. (2012). Using EZ Reader to simulate eye movements in nonreading tasks: A unified framework for understanding the eye–mind link. *Psychological review*, *119*, 155-185.

Acknowledgments

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 manuscript. 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.

Footnotes

1. There was one exception where 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).

3. The model fit to skipping data on the pre-target 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 post-target word, and the resulting model including additive random effects slopes for preview and predictability by subjects and preview by items.

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 1. *Condition means for the pre-target word*.

		Skip	SFD	GD	SFD		GD	
Constraint	Preview	All cases			Target subsequently fixated	Target subsequently skipped	Target subsequently fixated	Target subsequentl
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.

Table 2. *Linear mixed effects model analyses on the pre-target 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. SFD = single fixation duration; GD = gaze duration; *SE* = standard error; *t* = test statistic (*b/SE*). Cells marked in bold represent $|t| \geq 1.96$.

Table 3. *Condition means for the target word.*

		Skip	SFD	GD	Go-Past	Regressions In	
Constraint	Preview	All cases				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. SFD = single fixation duration; GD = gaze duration; *SE* = standard error.

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. SFD = single fixation duration; GD = gaze duration; *SE* = standard error; *t* = test statistic (*b/SE*). Cells marked in bold represent $|t| \geq 1.96$.

Table 5. Condition means on the post-target 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)

Table 6. Linear mixed effects model analyses for the post-target word.

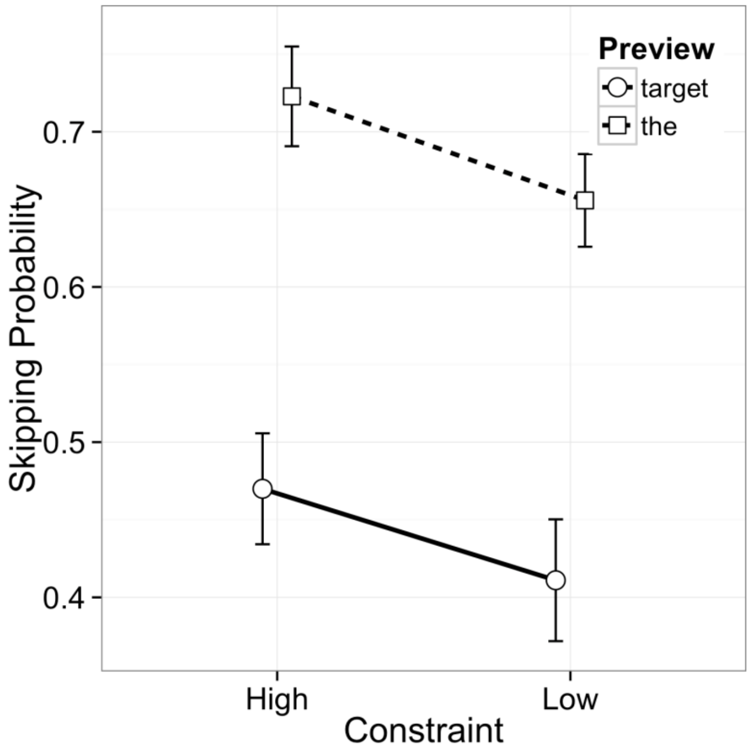
	Duration Measures									Binomial Measures					
	SFD			GD			Go-Past Time			p(Skip)			Regressions Out		
Effect	*b*	SE	*t*	*b*	SE	*t*	*b*	SE	*t*	*b*	SE	*z*	*b*	SE	*z*
(Intercept)	5.33	0.02	**349.31**	5.24	0.02	**277.58**	5.59	0.03	**205.49**	-1.13	0.27	**-4.17**	-2.37	0.31	**-7.77**
Fixed Effects															
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

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).

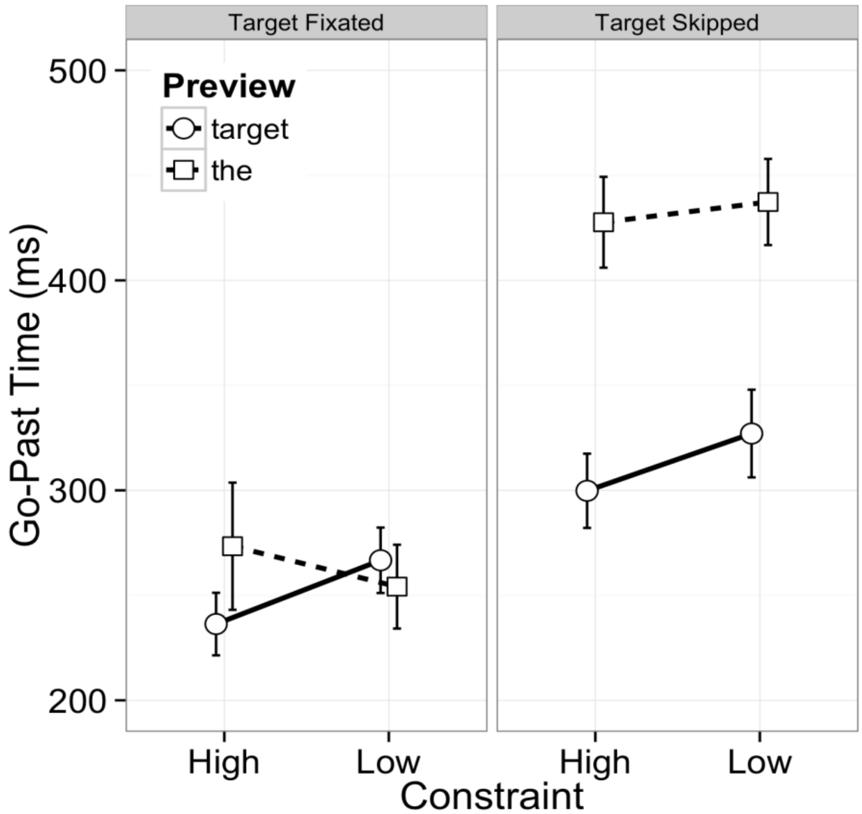
Constraint	**Preview**	**Sentence**
High	target	Jane used the scissors to carefully ¦ cut scraps of paper into neat shapes
High	*the*	Jane used the scissors to carefully ¦ the scraps of paper into neat shapes
Low	target	Jane used the machine to carefully ¦ cut scraps of paper into neat shapes.
Low	*the*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

Figure 2. *Target word skipping*.



Note: Error bars represent standard error of the mean.

Figure 3. *Go-past time on the post-target word*.



Note. Error bars represent standard error of the mean.

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

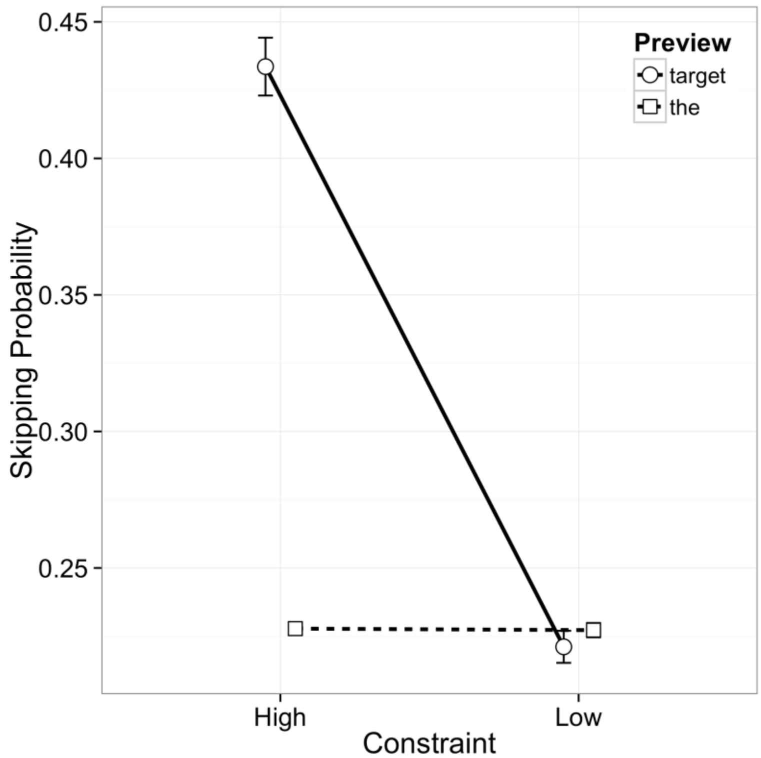


Figure 5. *Simulation 2: Target skipping probabilities by condition.*

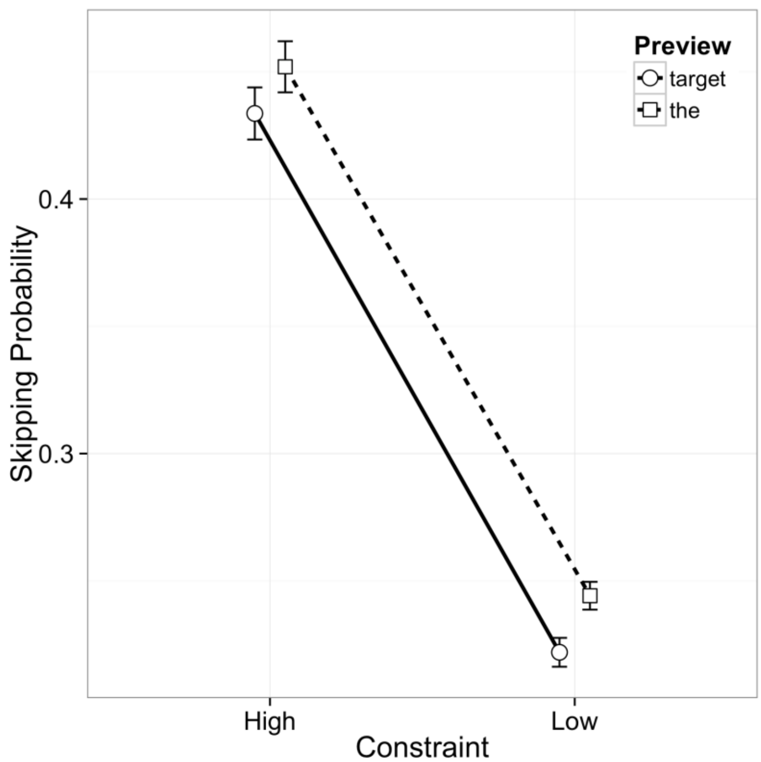
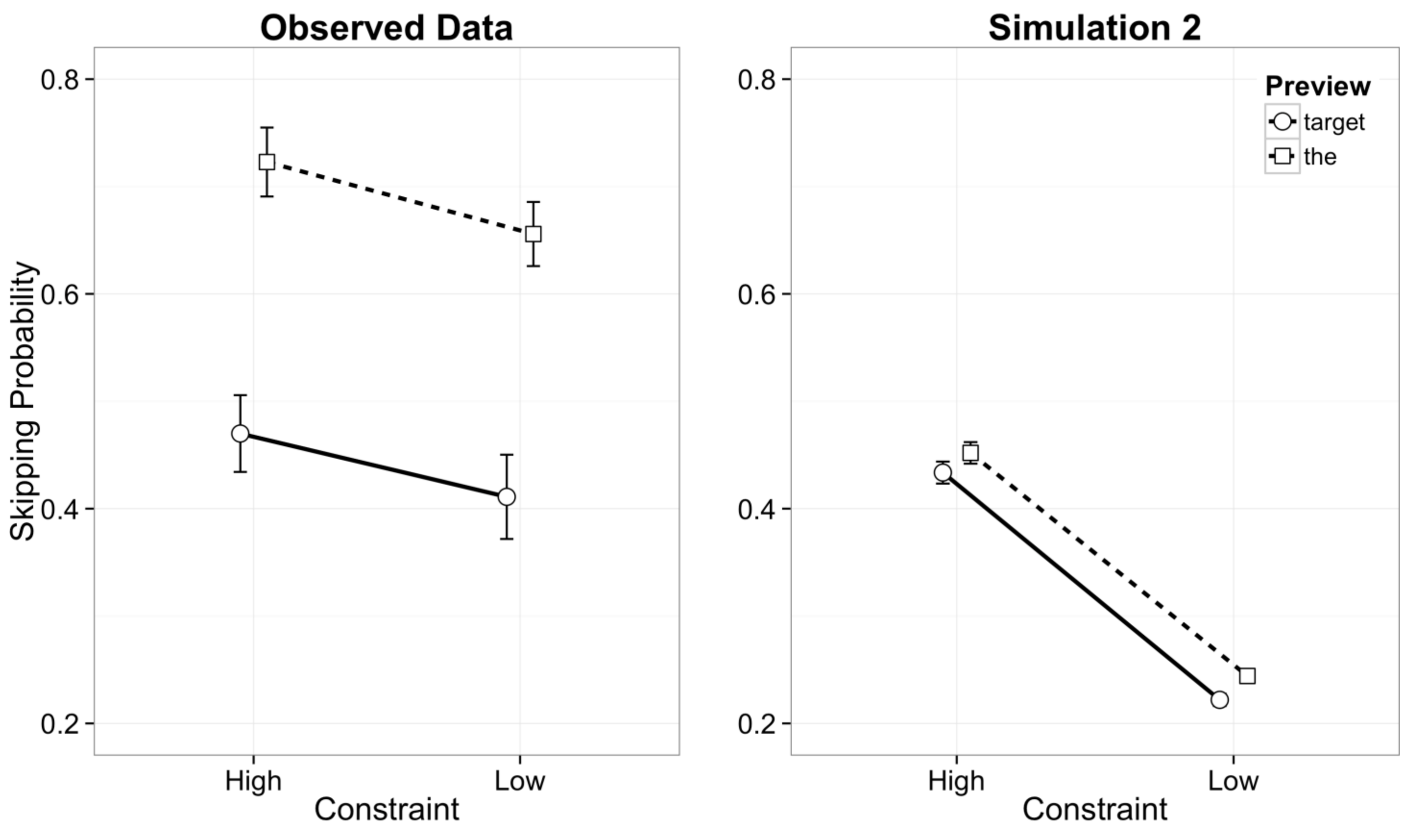


Figure 6. *Skipping probabilities for the target words from the observed data (left panel) and from Simulation 2 (right panel).*



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. 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.

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
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