Published in final edited form as:

J Exp Psychol Learn Mem Cogn. 2008 July ; 34(4): 1001–1010. doi:10.1037/0278-7393.34.4.1001.

Effects of Context on Eye Movements When Reading About Possible and Impossible Events

Tessa Warren,

Department of Psychology and Learning Research and Development Center, University of Pittsburgh;

Kerry McConnell, and

Department of Psychology and Learning Research and Development Center, University of Pittsburgh;

Keith Rayner

Department of Psychology, University of California, San Diego

Abstract

Plausibility violations resulting in impossible scenarios lead to earlier and longer lasting eye movement disruption than violations resulting in highly unlikely scenarios (K. Rayner, T. Warren, B. J. Juhasz, & S. P. Livsedge, 2004; T. Warren & K. McConnell, 2007). This could reflect either differences in the timing of availability of different kinds of information (e.g., selectional restrictions, world knowledge, and context) or differences in their relative power to guide semantic interpretation. The authors investigated eye movements to possible and impossible events in real-world and fantasy contexts to determine when contextual information influences detection of impossibility cued by a semantic mismatch between a verb and an argument. Gaze durations on a target word were longer to impossible events independent of context. However, a measure of the time elapsed from first fixating the target word to moving past it showed disruption only in the real-world context. These results suggest that contextual information did not eliminate initial disruption but moderated it quickly thereafter.

Keywords

language comprehension; eye movements; plausibility

Data about the time course of use of different kinds of information has been a primary source of evidence in debates about the architecture of the language comprehension system (Grodner, Gibson, & Watson, 2005; Rayner, Carlson & Frazier, 1983; Trueswell, Tanenhaus, & Garnsey, 1994). In the present study, we used eye movement data to investigate the relative time course of use of two kinds of information during reading: (a) context and (b) the match between semantic features on a verb and an argument. There has been considerable previous investigation of the time course over which contextual information can influence the selection of (a) a meaning for an ambiguous word (Duffy, Morris, & Rayner, 1988; Swinney, 1979;

Correspondence concerning this article should be addressed to Tessa Warren, Learning Research and Development Center, University of Pittsburgh, 607 LRDC, 3939 O'Hara Street, Pittsburgh, PA 15260. E-mail: tessa@pitt.edu.

⁶Two items had small differences in the posttarget region (e.g., a *he* vs. *she*). Analyses that could have been affected (total time, analyses from the posttarget region) were run without these two items and showed the same patterns as when they were included.

⁷One participant had no observations in the posttarget region for one of the conditions. This missing value was replaced with the grand mean for the region in that condition.

Tanenhaus, Leiman, & Seidenberg, 1979) and (b) a syntactic parse (Binder, Duffy, & Rayner, 2001; Rayner et al., 1983; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Trueswell et al., 1994). The question of whether and when contextual information can influence linguistic processing has been central to the debate about whether the language processing system is composed of separable modules that use only limited kinds of information (Fodor, 1983) or is an interactive system that can use any available information at any given time.

Multiple experiments have suggested that all meanings of ambiguous words are initially accessed regardless of context (Swinney, 1979; Tanenhaus et al., 1979; cf. Duffy et al., 1988). This has been taken as evidence that early word processing occurs in a module encapsulated from contextual or world knowledge. Recent studies have weighed in on this debate with evidence from an early stage of semantic processing: the integration of the semantic features of an argument and a predicate. Hagoort, Hald, Bastiaansen, and Petersson (2004) compared violations of argument/predicate semantic features (e.g., assigning an impossible attribute to an object, like “sour” to a train) with violations of world knowledge (e.g., a false statement). Event-related potentials (ERPs) to the two kinds of violations occurred along the same timeline, and functional magnetic resonance imaging (fMRI) data suggested they were processed in similar brain areas, inconsistent with the existence of a spatially distinct fast-acting linguistically specialized module.

Eye movement studies paint a slightly different picture. In an experiment similar to the present one, Rayner, Warren, Juhasz, and Liversedge (2004) examined eye movements to plausibility violations. They found that severe anomalies, most of which were caused by argument/predicate semantic feature mismatches, were associated with disruption as soon as the target word was encountered. Less severe violations of world knowledge (in this case, unlikely events like using an axe to chop carrots) produced disruption only in middle to late measures on the target word.¹ Warren and McConnell (2007) followed up on Rayner et al. but controlled the severity of world knowledge violations, while varying the presence of an argument/predicate semantic feature mismatch. Their stimuli (Sentences 1–3) described plausible events, extremely implausible but possible events, and extremely implausible impossible events. Impossible events were cued by a semantic violation between a verb and its object, which served as the target word.

1. The man used a strainer to drain the thin spaghetti yesterday evening. (Possible-plausible.)
2. The man used a blow-dryer to dry the thin spaghetti yesterday evening. (Possible-implausible.)
3. The man used a photo to blackmail the thin spaghetti yesterday evening. (Impossible-implausible.)

Disruption was evident in the impossible-implausible condition compared with the possible-implausible condition as early as the first fixation on the target word (*spaghetti* above) and continued in the posttarget region. In contrast, the two possible conditions only began to differ later, in the go-past (or regression path duration) measure on the target word, an index of how long it took for the eyes to move past this word. Disruption in this comparison had dissipated by the posttarget region. The finding of earlier and longer lasting disruption to impossible events than highly implausible events is consistent with the possibility that semantic feature matching between verbs and arguments occurs before world knowledge about event plausibility is available. Alternatively, a delay in the effects of world knowledge as compared with those of semantic feature information could reflect the slower availability of world

¹Joseph et al. (2008) replicated Rayner et al. (2004) and found that 10- to 14-year-old children showed similar patterns to adults (with some delayed effects).

knowledge or a tendency of semantic interpretation to proceed from coarse-grained to fine-grained (Sanford & Garrod, 2005) and, therefore, be consistent with either architecture.

Evidence based on contextual manipulations is particularly strong in this debate, as there is no concern that the influence of contextual information is delayed relative to the access of semantic feature information on a word. Context's ability to influence eye movements in advance of upcoming linguistic input has been demonstrated in both visual world (Kamide, Altmann, & Haywood, 2003; Sedivy, Tanenhaus, Chambers, & Carlson, 1999) and reading (Ehrlich & Rayner, 1981; Rayner & Well, 1996) experiments. Because context can affect the processing of an upcoming word (via predictability), it must be available before the semantic features of that word are accessed.

Using a contextual manipulation, Nieuwland and van Berkum (2006) found evidence suggesting that early semantic feature processing is not performed within an encapsulated module. In their study, participants listened to six-sentence texts that had either an inanimate or animate protagonist associated with predicates requiring an animate argument (e.g., "... consoled the yacht/sailor"). ERPs were measured on the protagonist noun phrase in the first, third, and fifth sentences. An N400 was evident for inanimate as compared with animate protagonists in the first sentence; however, this N400 effect was attenuated in the third sentence and absent in the fifth sentence. A second experiment verified that the N400's elimination was not due to lexical repetition and also showed that an N400 could be induced for an inanimate argument/predicate combination if the inanimate argument had previously been associated with multiple animate predicates (e.g., an N400 on "the peanut was salted" at the end of a love story starring peanuts).

The fact that context eliminated N400s for a semantic feature mismatch by the fifth encounter with such a violation suggests that context was able to change the semantic features that were accessed for the argument from default features on the first encounter to contextually appropriate features on later encounters. Note that this impact of context on lexical access or early lexical integration goes beyond simply modulating the activation of different sets of semantic features, as may occur with ambiguous words (Duffy et al., 1988), and instead demonstrates the overriding of default semantic features associated with an unambiguous noun.

Nieuwland and van Berkum's (2006) evidence strongly militates against a modular account of lexical access. However, by the time the N400 was eliminated in their study, the target noun had previously appeared four times with predicates that required it to be interpreted as animate. This suggests that it took considerable repetition for context to override the noun's semantic features. This leaves an open question of whether such overriding can happen on a reader's first encounter with a word. Also, it is important to investigate this question using other methodologies. Although ERPs provide a continuous measure, the relationship between processing disruption and the latency of the corresponding ERP component may not be as direct as that between processing disruption and the timing of eye movement disruption (Bornkessel & Schleewsky, 2006). Eye movement studies of plausibility violations have demonstrated disruption as early as 200–300 ms after the eyes fixate a semantically incompatible word (Rayner et al., 2004; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007; Warren & McConnell, 2007), and with eye tracking, it is possible to detect both the initiation and resolution of disruption even if they occur on the same word. If the eye movement record provides a more direct measure of early processing disruption than do ERPs (Sereno & Rayner, 2003), it will be important to corroborate the ERP results with eye tracking.

An experiment by Filik (2008) addressed both of these questions, but it did not resolve them. Filik tested sentences like "the mouse picked up the dynamite" (and a plausible control sentence) in either a cartoon context ("Tom and Jerry" in this example) or a real-world context.

There were no reliable fixation time effects on the initial encounter with the target word, but in the posttarget region, the implausible condition showed disruption only in the real-world context. This finding is consistent with a strong effect of context, but the timing of the effect is late compared with other eye movement studies of plausibility violations (Rayner et al., 2004; Staub et al., 2007; Warren & McConnell, 2007). This long detection latency may be related to the fact that most of the plausibility violations in Filik's items were not caused by mismatches between the semantic features on an argument and a predicate. Instead, many of the violations were due to an improbable agent carrying out an action, like chickens destroying a factory, a cat opening a can, or a small bird grabbing a stick. This fact limits the strength of conclusions that can be drawn about modularity on the basis of Filik's data because the violations in her items may have been signaled by world knowledge about the likely participants in events rather than by semantic feature mismatches of arguments to their predicates. This possibility is supported by the relatively late detection of violations, as Warren and McConnell (2007) found that strong anomalies cued by world knowledge are detected later than ones cued by argument/predicate feature mismatches.

To further address the questions left open by Nieuwland and van Berkum (2006) and by Filik (2008), we monitored eye movements, which provide detailed information about the full time course of language interpretation during normal reading (Rayner, 1998), across short texts in four conditions: (a) a real-world context with a possible event, (b) a real-world context with an impossible event, (c) a fantasy or specialized context with an event that is possible in the real world, and (d) a fantasy or specialized context with an event that is impossible in the real world (but possible in this context). In these texts, the target word that caused the event to be possible or impossible in the real world did not appear previously in the text. If semantic feature matching between a verb and its arguments is encapsulated from contextual information, then there should be an early main effect of impossibility on disruption at the target word such that impossible conditions show more disruption than possible conditions. If, however, contextual information can modify the default semantic features initially associated with a word (e.g., from – animate to + animate), then there should be an interaction such that there is early disruption in the real-world context/impossible condition but none in the other conditions.

Method

Participants

Forty-four undergraduates (native English speakers with normal or corrected-to-normal vision) at the University of Pittsburgh participated for course credit.²

Apparatus

An Eyelink 1000 tracker (SR Research Ltd., Toronto, Ontario, Canada), with spatial resolution of less than a 30-min arc and a sampling rate of 1 ms, was used to monitor participants' right eyes during reading. Participants viewed the stimuli binocularly on a monitor 63 cm from their eyes; approximately three characters equaled 1° of visual angle.

Materials

The materials consisted of 24 passages in a 2 × 2 design, crossing the factors of context (real world vs. fantasy³) and event type (possible vs. impossible in the real world). In the passages (see Appendix A for examples), the clause that described the target event had the verb *used*,

²One was not included in analyses because they only had one observation for one of the experimental conditions.

³Note that the labels are generalizations to simplify exposition; a contextual manipulation that was not fantasy was used for some items. In these cases, there was a highly specialized or unusual real-world context, for example, someone preparing a Bugs Bunny balloon for the Thanksgiving Day parade.

followed by an instrument (*microwave* in the possible event), an infinitival verb (*to heat* in the possible event), an adjectival noun phrase (*the tough bread*), and then continued with an additional clause or modifier (*and made a tasty-looking tuna melt*). The adjectival noun phrase and following two words were the same across conditions. The target word, the locus of the impossibility in the impossible event, was the noun of the adjectival noun phrase (*bread*) and a minimum of five characters long. Items were designed not to have strong semantic relationships between the critical words.⁴ The full set of materials is in Appendix B.

The 24 experimental passages were combined with 66 filler passages. Experimental passages were displayed such that analyzed regions appeared on the same line, as close to the center of the screen as possible. Three filler passages contained anomalies. Conditions were counterbalanced across four presentation lists using a Latin square design.

Procedure

The experiment lasted 40–55 min. A chinrest and forehead rest minimized head movements. Participants were asked to read normally, for comprehension, and told that after some passages, they would need to answer a yes–no comprehension question; these questions occurred after 80% of the passages. After the participant received instructions on the format of the experiment, the tracker was calibrated (and later recalibrated as necessary). Participants controlled the onset and offset of the passage displays by pressing a key.

Results

Comprehension rates were high (85%). Track losses, blinks, and incomplete trials necessitated that approximately 22% of trials be excluded from analysis. Fixations shorter than 80 ms that fell within 1.5 characters of another fixation were combined into the longer fixation. Additionally, data points more than 2.5 standard deviations from the mean within a condition were eliminated, excluding less than 2% of the remaining data. Data loss affected all conditions similarly. Approximately 79% of all analyzed regions were fixated during the first pass across all conditions.

Three analysis regions were defined: the *pretarget region* (the determiner and adjective preceding the target noun); the *target region*; and the *posttarget region* (the word after the target noun, or if that word contained fewer than five characters, the two words following the target noun). Five measures were computed over each region: *first fixation duration* (the duration of the first fixation on a region during first pass reading), *gaze duration* (the sum of all fixations from first entering a region during first pass reading until leaving it), *go-past time* (also called *regression path duration*, the sum of all fixations, including regressive fixations, from first entering a region during first pass reading until leaving it to the right), *regressions out* (the percentage of times a regression was launched from a region during first pass reading), and *total time* (the sum of all fixations on a region). The first two measures index early processing, whereas go-past is a mixture of early and slightly later processing. We subjected data to analyses of variance (ANOVAs), using participants (*F1*) and items (*F2*) as random factors.

⁴To verify this, we examined the South Florida word association norms (Nelson, McEvoy, & Schreiber, 1998) and the Edinburgh Association Test (Kiss, Armstrong, Milroy, & Piper, 1973). In the first set of norms, four target words (*water*, *pancake*, *event*, and *program*) had entries for the preceding noun in the sentence with an average forward association value of .15 and a backward association value of .06. To verify that these items were not responsible for the data pattern, we reanalyzed the data without them, and the data pattern was unchanged. The second set of norms included more of the target words. However, according to these norms, the probability of choosing the target word as an associate when given the first noun in the sentence (*microwave* and *book* in Appendix A) was less than 2%. Likewise, when given the verb (*heat* and *teach* in Appendix A), the probability of choosing the target word as an associate was less than 1% for both the possible-event sentences and the impossible-event sentences. Thus, the association between the two nouns and between the noun and the verb in the sentences was very low.

Pretarget Region

There were no significant differences between the conditions on either of the early measures (see Table 1). There was an interaction that was reliable by participants and marginal by items in go-past region, $F(1, 42) = 6.0, p < .05$; $F(1, 23) = 4.01, p = .06$. Plausibility norms (see below) indicated that this mirrors an effect in offline plausibility ratings at this point in the sentence. The total time measure showed a reliable interaction as well as main effects, context: $F(1, 42) = 25.1, p < .001$; $F(1, 23) = 11.8, p < .005$; plausibility: $F(1, 42) = 21.3, p < .001$; $F(1, 23) = 15.2, p < .001$; Context $\times$ Plausibility: $F(1, 42) = 31.4, p < .001$; $F(1, 23) = 25.8, p < .001$.

Target Region

First fixation durations (see Table 2) showed a marginal effect of possibility in the items analysis, $F(1, 42) = 1.4$; $F(1, 23) = 3.9, p = .06$, which became fully reliable in gaze duration, $F(1, 42) = 4.7, p < .05$; $F(1, 23) = 12.6, p < .01$. Specifically, first fixations were 12 ms longer in the impossible than in the possible conditions, and gaze durations were 24 ms longer. However, the go-past measure yielded a significant interaction, $F(1, 42) = 15.9, p < .001$; $F(1, 23) = 8.2, p < .01$. Pairwise comparisons confirmed that the real-world impossible condition showed more disruption than each of the other three conditions (all $ps < .001$). This interaction drove reliable main effects of context, $F(1, 42) = 12.8, p = .001$; $F(1, 23) = 9.3, p < .01$, and possibility, $F(1, 42) = 14.3, p < .001$; $F(1, 23) = 12.9, p < .005$. More important, there was no reliable difference between the possible and impossible fantasy conditions ($ts < 1$), indicating that the fantasy context overrode the initial difficulty associated with interpreting an event that was impossible in the real world. Regressions out showed only a reliable main effect of context; readers made more regressions in the real-world context, $F(1, 42) = 15.0, p < .001$; $F(1, 23) = 11.6, p < .005$. This regression effect was not predicted and is difficult to interpret. The same interaction and main effects that were apparent in go-past time also appeared in total reading times, Context $\times$ Possibility: $F(1, 42) > 50, p < .001$; $F(1, 23) = 17.7, p < .001$; context: $F(1, 42) = 50, p < .001$; $F(1, 23) = 28.7, p < .005$; possibility: $F(1, 42) > 50, p < .001$; $F(1, 23) = 20.0, p < .001$. As in go-past, the two fantasy conditions did not differ reliably ($ts < 1$).

Before discussing the posttarget region, it is important to address the possible contributions that overall passage plausibility and the predictability of the target word may have played in these results. To evaluate the overall plausibility of the items, participants (who did not participate in the eye-tracking experiment) rated the naturalness of passages on a scale ranging from 1 (*very natural*) to 5 (*very unnatural*). Of the participants, 20 rated passages truncated before the target word, and 21 rated the same passages truncated after the target word (see Table 3 for means). For the passages ending at the pretarget word, there was an interaction between context and possibility, $F(1, 19) = 21.5, p < .001$; $F(1, 23) = 4.9, p < .05$. Pairwise comparisons indicated this was due to the real-world impossible condition being rated less natural than the other conditions (all $ps < .05$, except the real-world impossible and fantasy possible by-items comparison; $p = .06$). Critically, there were no reliable differences between the fantasy impossible condition and either of the possible conditions ($ps > .26$). This suggests that the interaction in go-past in the pretarget region was likely spurred by naturalness differences at that point in the paragraph but that there was nothing in this region to presage the early disruption that appeared in the fantasy impossible condition in the next region. In the passages including the target word, ANOVAs revealed a reliable interaction, $F(1, 20) = 116.7, p < .001$; $F(1, 23) = 57.8, p < .001$, that drove reliable main effects, context: $F(1, 20) = 13.3, p < .005$; $F(1, 23) = 11.7, p < .005$; possibility: $F(1, 20) = 43.2, p < .001$; $F(1, 23) = 110.7, p < .001$. Again, pairwise comparisons confirmed that the real-world impossible condition was rated reliably less natural than all other conditions ($ps < .001$). The real-world possible condition was rated reliably more natural than the other three conditions ($ps < .005$), and the two fantasy

conditions did not differ ($ps > .3$). These results suggest that, as designed, the real-world impossible condition became extremely implausible (4.3 on a scale of 1–5) at the target word. More important, also as designed, the fantasy impossible condition did not become implausible at the target word, and in fact was numerically rated more natural than the fantasy possible condition.

To determine the predictability of the target words, 20 participants completed a cloze task in which they were given full experimental items, up to but not including the target word, and asked to generate a candidate next word that would fit the paragraph. Participants completed two versions of each item, one from each context and possibility status, as there was little overlap between items from different contexts and possibility statuses before the precritical region (see Table 3 for means). Target words were not very predictable overall, but they were more predictable in the possible than in the impossible conditions. In order to evaluate how the factors of predictability, context, and possibility jointly contributed to the gaze duration effect, we used a linear mixed-effects model (Baayen, 2008; Baayen, Davidson, & Bates, in press). We carried out these analyses using the lme4 package for linear mixed effects models (Bates, 2005), which is part of the R statistical computing package (R Development Core Team, 2007). Participants and items were included as crossed random factors. The fixed factors included in the model were context, possibility, the interaction between context and possibility, and cloze predictability. Both possibility ($t = 2.298, p < .05$) and cloze predictability ($t = 2.118, p < .05$) had reliable effects⁵ on gaze duration, unlike context or the interaction between context and possibility ($ps > .6$). These analyses indicated that the model that optimally accounted for the gaze duration data included parameters for both possibility and predictability, meaning that the main effect of possibility cannot be reduced to a predictability effect.

Posttarget Region

Table 46,7 shows the data. For first fixation, $FI(1, 42) = 7.3, p = .01$; $F2(1, 23) = 5.1, p < .05$, and gaze duration, $FI(1, 42) = 5.0, p < .05$; $F2(1, 23) = 2.2, p = .15$, there were main effects of possibility (only reliable by participants in gaze duration). There were reliable interactions in go-past, percentage of regressions out, and total time such that the real-world impossible condition showed more disruption than did every other condition, go-past: $FI(1, 42) = 12.9, p = .001$; $F2(1, 23) = 10.6, p < .005$; regressions out: $FI(1, 42) = 4.6, p < .05$; $F2(1, 23) = 3.6, p = .07$; total: $FI(1, 42) = 31.3, p < .001$; $F2(1, 23) = 14.2, p < .001$. These interactions were accompanied by main effects of possibility, such that impossible conditions were longer, go-past: $FI(1, 42) = 19.8, p < .001$; $F2(1, 23) = 21.5, p < .001$; regressions out: $FI(1, 42) = 7.8, p < .01$; $F2(1, 23) = 15.3, p < .001$; total: $FI(1, 42) = 34.9, p < .001$; $F2(1, 23) = 20.6, p < .001$. Total time also showed an effect of context, $FI(1, 42) = 6.5, p = .015$; $F2(1, 23) = 4.7, p = .04$. The two fantasy conditions did not differ reliably in the go-past or later measures ($ts < 1$), suggesting that the fantasy context overrode the initial difficulty associated with interpreting an event that would have been impossible in the real world.

Discussion

The present study indicated that the earliest disruption to impossible events was unaffected by context. Hints of disruption on the target word in first fixation for the impossible as compared with the possible conditions in both contexts became reliable disruption in gaze duration. However, a fantasy context quickly moderated disruption to impossible events. In go-past and subsequent measures on the target word, disruption was only evident in the real-world impossible condition. This pattern of results suggests that context can be strong enough to

⁵The p values were estimated by Monte Carlo simulations as described by Baayen (2008).

eliminate the unnaturalness of an impossible event in offline and relatively late online measures, yet not be able to eliminate initial online difficulty.

As noted earlier, a study by Filik (2008) yielded effects similar to those we obtained in the go-past measure, but in her study, the effects emerged in the posttarget region rather than at the word causing the plausibility violation. Although our study and Filik's study both suggest that a fantasy/cartoon context can quickly modulate disruption related to anomalous events, our study showed early disruption to anomalous events regardless of context, whereas Filik's did not. Two comments seem relevant. First, Filik did not report go-past time, so direct comparisons are difficult. Second, as noted earlier, most plausibility violations in Filik's study were signaled by world knowledge about the likely participants in events rather than by semantic feature mismatches, which may have contributed to the relatively late detection of the violations. Consistent with the present results, in a series of studies similar to the present study and to Filik (2008), Ferguson and Sanford (2008) found that a counterfactual context was unable to eliminate early disruption to an implausible event but did eliminate it in later measures.

Another point that we would stress is that we suspect that the results of our study regarding the possible conditions are not due to semantic priming. That is, it could be argued that *microwave* serves as a better prime for *bread* than *book* does. Or, that the noun-verb combination of *microwave* and *heat* primes *bread* more than does the combination of *book* and *teach*. Indeed, the three content words in the possible conditions have more semantic overlap than do the words in the impossible conditions according to the BEAGLE framework (Jones & Mewhort, 2007). However, word association norms (see Footnote 4) provided little evidence consistent with a priming account. Furthermore, Carroll and Slowiaczek (1986) found that although words that were strong associates (like *king-queen*) yielded priming effects reflected in gaze durations during reading, semantically related but nonassociated words (like *ambassador-queen*) did not. As we noted earlier, the critical words in our sentences (*microwave-bread*) were not semantic associates and were much more like the nonassociative words in Carroll and Slowiaczek's study. Finally, it would be difficult to explain the interaction found in the go-past measure using a simple priming account.

Returning to the main focus of the present study, although the pattern of results was consistent with a modular account of language processing, it can also be explained by a nonmodular account. Nieuwland and van Berkum's (2006) findings suggested that context is powerful enough to change the default semantic features accessed with a lexical item. However, our findings demonstrate a potential limitation of this power, namely, it may require previous exposure to a specific kind of semantic mismatch involving that lexical item. An account consistent with this is the *lexical reinterpretation model* discussed by Hess, Foss, and Carroll (1995) (see also the *graded saliency hypothesis*; Giora, 2002). According to this model, readers access the typical or context-independent semantics for a word upon their first encounter with it in a given context. After this initial retrieval, contextual information can modify active semantic features to be consistent with a particular referential instantiation of the word. Once this referential relation is established, subsequent encounters with the word activate the features associated with the specific referent rather than a context-independent semantics. For example, the default meaning accessed upon an initial encounter with the word *lobster* might be a sea crustacean, but context might quickly modify the meaning to be a seafood casserole. In some cases, it is possible that there may be processing costs associated with such meaning modification (Frisson & Frazier, 2005), whereas in other cases, these alternate meanings may already be lexicalized and require no additional processing (Frisson & Pickering, 1999).

Acknowledgements

This research was supported by National Institutes of Health (NIH) Grant HD048990 awarded to Tessa Warren and NIH Grant HD17246 awarded to Keith Rayner. We thank Charles Clifton and Adrian Staub for assistance with data analysis, Erik Reiche for many helpful discussions, the Pitt Reading and Language lab, Daniel Sterba and Amanda Virbitsky for help developing stimuli, and Sarah Heider and Alison Trude for assistance in data collection.

References

- Baayen, RH. Analyzing linguistic data: A practical introduction to statistics. Cambridge, England: Cambridge University Press; 2008.
- Baayen RH, Davidson DH, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*. in press
- Bates DM. Fitting linear mixed models in R: Using the lme4 package. *R News: The Newsletter of the R Project* 2005;5:27–30.
- Binder KS, Duffy SA, Rayner K. The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language* 2001;44:297–324.
- Bornkessel I, Schlesewsky M. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 2006;113:787–821. [PubMed: 17014303]
- Carroll P, Slowiaczek ML. Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition* 1986;14:509–522.
- Duffy SA, Morris RK, Rayner K. Lexical ambiguity and fixation times in reading. *Journal of Memory and Language* 1988;27:429–446.
- Ehrlich SF, Rayner K. Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior* 1981;20:641–655.
- Ferguson H, Sanford A. Anomalies in real and counterfactual worlds: An eye movement investigation. *Journal of Memory and Language* 2008;58:609–626.
- Filik R. Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition* 2008;106:1038–1046. [PubMed: 17524387]
- Fodor, JA. The modularity of mind. Cambridge, MA: MIT Press; 1983.
- Frisson S, Frazier L. Carving up word meaning: Portioning and grinding. *Journal of Memory and Language* 2005;53:277–291.
- Frisson S, Pickering MJ. The processing of metonymy: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition* 1999;25:1366–1383.
- Giora R. Literal vs. figurative language: Different or equal? *Journal of Pragmatics* 2002;34:487–506.
- Grodner D, Gibson E, Watson D. The influence of contextual contrast on syntactic processing: Evidence for strong interaction in sentence comprehension. *Cognition* 2005;95:275–296. [PubMed: 15788160]
- Hagoort P, Hald L, Bastiaansen M, Petersson KM. Integration of word meaning and world knowledge in language comprehension. *Science* 2004 April 16;304:438–441. [PubMed: 15031438]
- Hess DJ, Foss DJ, Carroll P. Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General* 1995;124:62–82.
- Jones MN, Mewhort DJK. Representing word meaning and order information in a composite holographic lexicon. *Psychological Review* 2007;114:1–37. [PubMed: 17227180]
- Joseph HSSL, Liversedge SP, Blythe HI, White SJ, Gathercole SE, Rayner K. Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology* 2008;61:708–723. [PubMed: 17853235]
- Kamide Y, Altmann GTM, Haywood SL. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 2003;49:133–156.
- Kiss, GR.; Armstrong, C.; Milroy, R.; Piper, J. An associative thesaurus of English and its computer analysis. In: Aitken, AJ.; Bailey, RW.; Hamilton-Smith, N., editors. *The computer and literary studies*. Edinburgh, Scotland: Edinburgh University Press; 1973. p. 153-165.

- Nelson, DL.; McEvoy, CL.; Schreiber, TA. The University of South Florida word association, rhyme, and word fragment norms. 1998. Retrieved June 10, 2007, from <http://www.usf.edu/FreeAssociation/>
- Nieuwland MS, van Berkum JJA. When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience* 2006;18:1098–1111. [PubMed: 16839284]
- Rayner K. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 1998;124:372–422. [PubMed: 9849112]
- Rayner K, Carlson MA, Frazier L. The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior* 1983;22:358–374.
- Rayner K, Warren T, Juhasz BJ, Liversedge SP. The effect of plausibility on eye-movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 2004;30:1290–1301.
- Rayner K, Well AD. Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review* 1996;3:505–509.
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2007. Retrieved October 3, 2007, from <http://www.R-project.org>
- Sanford A, Garrod S. Memory-based approaches and beyond. *Discourse Processes* 2005;39:205–224.
- Sedivy JC, Tanenhaus MK, Chambers CG, Carlson GN. Achieving incremental semantic interpretation through contextual representation. *Cognition* 1999;71:109–147. [PubMed: 10444906]
- Sereno SC, Rayner K. Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences* 2003;7:489–493. [PubMed: 14585445]
- Spivey MJ, Tanenhaus MK, Eberhard KM, Sedivy JC. Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology* 2002;45:447–481. [PubMed: 12480476]
- Staub A, Rayner R, Pollatsek A, Hyönä J, Majewski H. The time course of plausibility effects on eye movements in reading: Evidence from noun–noun compounds. *Journal of Experimental Psychology* 2007;33:1162–1169. [PubMed: 17983320]
- Swinney DA. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 1979;18:645–660.
- Tanenhaus MK, Leiman JM, Seidenberg MS. Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior* 1979;18:427–440.
- Trueswell JC, Tanenhaus MK, Garnsey SM. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language* 1994;33:285–318.
- Warren T, McConnell K. Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review* 2007;14:770–775. [PubMed: 17972747]

Appendix A Example Materials

Real-world context:

Patricia was teaching a kindergarten class. One day she caught a bully stealing another child's lunch. It turned out the bully had forgotten his lunch at home, so she brought him to the school kitchen, but all the food was stale.

Possible event:

Patricia used a microwave to heat the tough bread and made a tasty-looking tuna melt for him.

Impossible event:

Patricia used a book to teach the tough bread and made a tasty-looking tuna melt for him.

Fantasy context:

Harry Potter was up late reading his magic books. He decided to try out a new spell on some stale toast and got it to stand up and start reciting multiplication tables.

Possible event:

Harry used a microwave to heat the tough bread and made a tasty-looking tuna melt.

Impossible event:

Harry used a book to teach the tough bread and made a tasty-looking tuna melt.

Pretarget region: the tough

Target region: bread

Posttarget region: and made

Appendix B Experimental Items

1. Sarah was headed to her best friend's wedding. She left her house early to make sure she got there in time but was slowed down by an accident. On the way, she noticed the wrapping paper on her present was torn. Sarah used {a shortcut to avoid/an adhesive to glue} the annoying traffic and just made it to the wedding.

Cat Woman was headed to Spiderman's wedding. She left early in her catmobile to make sure she got there in time but was slowed down by an accident. She considered using the oil slick function on her car but decided that she should stick the cars in their place and drive around them all. She used {a shortcut to avoid/an adhesive to glue} the annoying traffic and just made it to the wedding.

2. Greg was working as a waiter at a fancy restaurant. Today, a foreign couple celebrating their first anniversary came in for dinner. He took their appetizer order. Greg used {a plate to serve/some music to entertain} the French cheese and placed it on the table with some red wine.

Peter Pan had invited Tinkerbell over for dinner and was preparing in the kitchen. Despite a healthy portion of Pixie dust, the salad yawned, and the other food was not in the mood for guests. He thought maybe he should do something to wake the food up. Peter Pan used {a plate to serve/some music to entertain} the French cheese and placed it on the table with some red wine.

3. Kate wanted to give her boyfriend a nice surprise and invited him to come over for dinner. She wanted to cheer him up because he had been really stressed out. After he came in and sat down, Kate used {a fork to serve/her hands to button-up} the thin spaghetti and hoped that his day would get better.

Witch Hazel was headed home from work one day when she came across a poor little pasta boy who was being picked on by the meatballs. She told the meatballs to go away and saw that the pasta-boy was lying in the snow with his coat open. The witch went over and used {a fork to serve/her hands to button-up} the thin spaghetti and hoped that his day would get better.

4. Carla invited her friends to a dinner party on Saturday night. She was fixing Mexican food, and it took a lot longer than she expected. Her friends became very hungry and

a little drunk and started arguing. Carla used {a dish to serve/some music to calm} the steaming beans and rushed the enchiladas to the table.

Samantha the witch invited her friends to a dinner party on Saturday night. She was fixing Mexican food, but things were not going well. The peppers and rice were picking fights with the other ingredients. Samantha used {a dish to serve/some music to calm} the steaming beans and rushed the enchiladas to the table.

5. Kent was working the night shift at the ER. He was called to attend to a surgery and grabbed a new white coat out of his locker on his way to the operating room. That morning, Kent had used {an iron to flatten/a tranquilizer to relax} the stiff collar so it was more comfortable.

Severus Snape, Professor of Potions at Hogwarts, was preparing for his class in the morning. He went to put on his cape and noticed that it was upright in the back of his closet, crying because the cleaners had left it stained and used too much starch. Snape used {an iron to flatten/a tranquilizer to relax} the stiff collar so it was more comfortable.

6. Trudy decided to stop hiking and set up camp for the night. She started to make dinner and set up her bedding. Trudy used {a knife to chop/a pump to inflate} the small carrots that she was carrying. She boiled them in water with some beans and made a satisfying soup.

Trudy worked for the Macy's Thanksgiving Day Parade. She had to make sure the balloons didn't have any leaks. She would cut out any portion of the balloon with leaks and replace it. After checking the Bugs Bunny balloon's arms, Trudy used {a knife to chop/a pump to inflate} the small carrots that he was carrying. {Ø/They were leaky and needed to be patched.}

7. Patricia was teaching a kindergarten class. One day she caught a bully stealing another child's lunch. It turned out the bully had forgotten his lunch at home, so she brought him to the school kitchen, but all the food was stale. Patricia used {a microwave to heat/a book to teach} the tough bread and made a tasty-looking tuna melt {Ø/for him}.

Harry Potter was up late reading his magic books. He decided to try out a new spell on some stale toast, and got it to stand up and start reciting multiplication tables. Harry used {a microwave to heat/a book to teach} the tough bread and made a tasty-looking tuna melt.

8. Jordan was hosting a family reunion at his cottage on the ocean. He was planning a really special lunch on Friday because everyone was going to leave on Saturday. On Thursday night, Jordan used {a pot to boil/a straw to drink} a large lobster as he made a batch of fresh rolls.

A gourmet chef got into a car accident a week ago and broke his jaw in ten places. He couldn't eat solid food, and had to liquefy his food in a blender. On Sunday, he wanted to taste a new shipment of seafood, so the chef used {a pot to boil/a straw to drink} a large lobster as he made a batch of fresh rolls.

9. Jenny wanted to be a naturalist when she grew up. She loved nature and spent a lot of time out in the field behind her house, collecting samples of plants and bugs. One day Jenny used {a net to catch/a riddle to trick} an unlucky butterfly that she kept in a jar for a few days. When she found out it was an endangered species, she released it.

During her travels in Wonderland, Alice was captured by an evil fairy. The fairy lived in a spectacular palace and was attended by swarms of insects. Large beetles guarded

Alice's cell, and dragonflies brought her food through a window. One day, Alice used {a net to catch/a riddle to trick} an unlucky butterfly that she kept in a jar for a few days. When it promised to help her escape, she let it go.

10. Meg's great aunt Harriet was finally coming home from a long stay at the hospital. Meg went over to Aunt Harriet's house that morning and saw that all of the little ceramic miniatures on the shelves were dirty, and there was a huge spider web in the corner. Meg used {a duster to clean/a spray to repel} the ugly figurines and saved the day.

Buzz Lightyear and the other toys were voting on the next mayor of the toy chest. All of a sudden, the lid of the toy chest opened and nasty-looking GI Joe figurines started flooding into the chest. The toys panicked and grabbed everything they could to defend themselves. Buzz used {a duster to clean/spray to repel} the ugly figurines and saved the day.

11. The cathedral in an old French town was undergoing renovations. Pigeons had been living in the rafters, and some of the stained glass was broken. However, on Easter Sunday, the town decided to celebrate Mass in the cathedral instead of in the chapel they had been using temporarily. A priest used {a bucket to carry/a rag to polish} the holy water carefully so that it would stay pure on the walk over to the cathedral.

At the coronation ceremony for the new pope, no expense was spared. The Vatican was decorated with beautiful flowers, and all of the walls were newly painted. As part of the ceremony, the pope blessed an ice sculpture in the cathedral. A priest used {a bucket to carry/a rag to polish} the holy water carefully so that it would stay pure.

12. John and his band were sitting in their studio, looking over the contract that a record company had just offered them. They thought this contract wasn't as good as their last one, and they were trying to figure out what to do. John used {a pick to strum/a computer to fax} an old guitar before he realized everyone was looking at him.

John and his band were sitting in their studio, looking through materials for album covers. They were supposed to send the record executives the album cover they wanted when they had all decided. John was impressed with some vintage photos of rock-and-roll instruments. He used {a pick to strum/a computer to fax} an old guitar before he realized that everyone was looking at him.

13. Frank and his kids stopped at a store to pick up food for a camping trip. After checking out, Frank tried to carry the heavy bags, but two ripped and food scattered everywhere. He asked his children to help him with the bags, but they were playing tag in the parking lot and refused. Frank used {a cart to transport/some chocolate to lure} the bothersome groceries across the parking lot.

Harry Potter went to the supermarket to get some food. He got some of his favorite soups and ravioli, but it turns out that they were naughtier than usual. They ran away in the store, and once he got them under control, the noodles picked fights with the sauce, and all was chaos. In the end, he used {a cart to transport/some chocolate to lure} the bothersome groceries across the parking lot.

14. Hilary was spending the summer in Paris, trying to perfect her French. After touring a cathedral, she noticed a man fishing in the pond under the hanging branches of a big tree. Hilary went over to the pond and used {a pencil to sketch/a joke to amuse} the old willow while her friends walked by.

Harry Potter and his friends needed to go to the enchanted forest, but big guardian trees waved their hanging branches threateningly at anyone who tried to leave the school grounds. But Harry knew that the tree was vain and would focus its attention

on anyone who made it feel special. Harry used {a pencil to sketch/a joke to amuse} the old willow while his friends walked by.

15. Kenny was preparing for his wife's birthday party. He decorated their house with flowers and put on a nice suit. After everything was almost ready, he checked on the present he had bought for her and hidden in the pantry. Kenny used {a cloth to polish/a pot to brew} a large diamond before she came home.

Kenny worked as an apprentice to a mean old witchdoctor. The witchdoctor's wife was always wearing all kinds of jewelry that her husband would conjure out of plain materials. The witchdoctor had instructed Kenny, and asked him to make a present for his wife's anniversary. Kenny used {a cloth to polish/a pot to brew} a large diamond before she came home.

16. Lindsey was late, as usual, to her job with the gardener. After she got there, she noticed that in her rush she had forgotten to put on her makeup. She kneeled down on the ground and used {a brush to apply/some compost to fertilize} the dark mascara in the corner of the garden.

Glinda the witch was working in her organic makeup garden. She pruned the glitter bush and watered the red lipstick, which was coming in quite full this year. Then she planted some new seeds and used {a brush to apply/some compost to fertilize} the dark mascara in the corner of the garden.

17. Dan worked in a lab studied birds. He had to stay late one night because some eggs were about to hatch. Around midnight, he got really hungry and decided to reheat some leftovers from his breakfast on the lab hot plate. Dan checked the bird eggs and used {a spatula to flip/an incubator to hatch} a small pancake before sitting down to eat some sausage.

A mad scientist was working late in his laboratory. His current project was to make food that would reproduce itself, so his monster army wouldn't go hungry when he sent it to conquer the world. He had already succeeded in making loaves of bread that grew baby loaves. The mad scientist used {a spatula to flip/an incubator to hatch} a small pancake before sitting down to eat some sausage.

18. Todd took his business partner hunting in the mountains. He owed his partner a lot of money and felt really guilty about it. After a long day of tracking deer, they still hadn't caught anything and decided to rest by a pond with lots of ducks. Todd used {a trap to catch/a check to reimburse} a large goose that emerged from the reeds.

Daffy Duck had borrowed a lot of money from the waterfowl brotherhood society. Now he was scared because Donald had told him the society was run by the mafia. One night Daffy heard something rustling outside his house and panicked. Daffy used {a trap to catch/a check to reimburse} a large goose that emerged threateningly from the reeds.

19. Ted and his friends were playing some ball one day on a cold and wet Saturday afternoon. Ted wanted to be the quarterback, and he was the best, so everyone let him. But because today was so rainy and cold, he was having a hard time not fumbling. He tried to run it, and he used {his hands to grasp/his lungs to inhale} the damp football, hoping that he wouldn't fumble it this time.

Superman, Batman, and the rest of the Superheroes' Pigskin Club got together to play a little ball. Superman, despite his strength, had a hard time not fumbling when it was raining. But this week, he had practiced breathing in super-hard, sucking the ball into his chest where he couldn't drop it. It was snapped, thrown to Superman, and he used

{his hands to grasp/his lungs to inhale} the damp football, hoping that he wouldn't fumble it this time.

20. Dave was making dinner for himself one night and started to reminisce about his high school days. His greatest moment was when he was crowned prom king. It was nice to think about that moment when he found out he won, but he could not remember who the queen was. Dave finished chopping an onion, and then he used {a photograph to refresh/a pan to fry} the old memory that was so sweet.

Vlad the sorcerer was excited to feast on his favorite food, which was the mind stuff of humans. He loved to eat broiled dreams with a subtle cream sauce, and the sugar-covered hopes that his mother used to bake for him were sinfully delectable. He used a spell to extract the main ingredient from an unsuspecting villager, and then the sorcerer used {a photograph to refresh/a pan to fry} the old memory that was so sweet.

21. Gerald was on his way to see Ozzy Osbourne play at the Mellon Arena. They were saying that Ozzy was supposed to be doing some nasty and distasteful things tonight onstage. Along his way, he got hungry and stopped and got some ice cream. He got to the arena and used {his ticket to attend/tongue to lick} the off-color event, knowing that he would regret it later.

Carl decided to go strange new art exhibit in which people were not only to look at the art, but taste it as well. Carl had a germ phobia, so he knew it would be hard to taste art with other peoples' spit on it. He peeked into the exhibit and saw a photo taken of Ozzy Osbourne playing a venue in Arizona. Carl used {his ticket to enter/tongue to lick} the off-color event, knowing that he would regret it later.

22. George and his friends were sitting around the campfire talking and watching the stars. It was getting chilly, but they planned to stay up and watch the night sky. Around midnight, they unpacked their astronomy equipment. George used {a telescope to watch/a sponge to absorb} a bright comet in the night sky.

Zeus had let himself go lately, and decided that it was time to do some house cleaning. Since his house was the universe, this was a tough task. First he painted some of the supernovas, and swept out some of the old worn-out stars. Then he went to work on a particularly troublesome spot, and he used {a telescope to watch/a sponge to absorb} a bright comet in the night sky.

23. Joan was at the carnival with her grandson Tommy. Tommy was anxious and wanted to do and get into everything. There was a clown selling toys and cheap carnival items, and Tommy saw what he wanted. He begged and begged his grandmother, and she used {some money to buy/some candy to spoil} the little balloon before he started to cry.

Bonzo the clown was acting out a routine with some blow-up toys that he made himself. Before anyone knew it, Bonzo had made an entire blow-up family. Bonzo made a grandmother who, he said, was always too nice to her grandchildren. Bonzo acted out a story, in which she used {some money to buy/some candy to spoil} the little balloon before he started to cry.

24. Jeremy worked as a software engineer for Apple. He had been working hard on one piece of software that he could not quite finish the way he wanted. He was very upset and wasn't talking to anyone. One night when he was with his girlfriend, he realized what he needed to do. Jeremy used {a computer to finish/his arms to hug} the new program and felt very happy.

While Jeremy, a software engineer, was sleeping one night, he dreamt that all of the software he had made in the past had a physical form. Some of them were giving him

high fives, while others just stuck to handshakes. Something he recently wrote stopped in front of him. Jeremy used {his computer to finish/his arms to hug} the new program and felt very happy.

Table 1
Average Fixation Times on the Pretarget Region in Milliseconds

Region	First fix		Gaze		Go-past		% Reg out		Total	
	M	SD	M	SD	M	SD			M	SD
Real-world possible	202	33	258	58	371	152	23		387	134
Real-world impossible	217	37	255	56	476	340	24		762	470
Fantasy possible	214	34	259	65	412	156	26		427	180
Fantasy impossible	209	31	266	82	398	141	24		418	201

Note. reg out = regressions out.

Table 2
Average Fixation Times on the Target Region in Milliseconds

Region	First fix		Gaze		Go-past		% Reg out		Total	
	M	SD	M	SD	M	SD			M	SD
Real-world possible	198	34	213	40	304	81	24		230	70
Real-world impossible	211	28	246	47	462	210	31		504	221
Fantasy possible	202	40	216	44	284	105	13		238	90
Fantasy impossible	214	30	240	47	300	85	16		260	84

Note. reg out = regressions out.

Table 3
Norming Results: Average Naturalness Ratings on a 1–5 Scale Prior to and Including the Target Word and Cloze Proportions at the Target Word

Region	Naturalness prior to target word		Naturalness including target word		Cloze proportions	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Real-world possible	2.61	0.65	2.10	0.59	0.25	0.30
Real-world impossible	3.36	0.66	4.29	0.57	0.004	0.02
Fantasy possible	2.87	0.65	2.87	0.58	0.32	0.32
Fantasy impossible	2.65	0.66	2.69	0.69	0.07	0.11

Table 4
Average Fixation Times on the Posttarget Region in Milliseconds

Region	First fix		Gaze		Go-past		% Reg out		Total	
	M	SD	M	SD	M	SD			M	SD
Real-world possible	213	31	278	48	329	67	12		342	95
Real-world impossible	229	37	283	53	629	253	30		543	176
Fantasy possible	208	33	265	60	386	213	14		385	135
Fantasy impossible	219	32	285	68	426	196	19		386	140

Note. reg out = regressions out.