Published in final edited form as:
Q J Exp Psychol (Hove). 2010 August ; 63(8): 1516–1532. doi:10.1080/17470210903380822.

Eye movements when reading implausible sentences: Investigating potential structural influences on semantic integration

Nikole D. Patson and **Tessa Warren**
University of Pittsburgh

Abstract

The disruption that occurs in response to reading about implausible events in unambiguous sentences can be informative about the time course of semantic interpretation (e.g. Hagoort, Hald, Bastiaansen, & Petersson, 2004; Nieuwland & Van Berkum, 2006; Warren & McConnell, 2007). Two eye-tracking studies used implausible sentences to investigate whether local factors like the structural relationships and the distance between words cueing a plausibility violation influence how quickly those words are integrated into a global semantic interpretation. Experiment 1 suggested that eye-movement disruption was unaffected by the number of words intervening between the words cueing the implausibility. Experiment 2 demonstrated that eye-movement disruption to implausibility occurred along the same time course regardless of whether the words cueing the implausibility were in a theta-assigning relation or not. These results suggest that these local structural factors do not influence how quickly new words are integrated into a semantic representation, but rather the global event representation determines the time course over which implausibility is detected.

The disruption that occurs in response to reading about implausible events in unambiguous sentences can be informative about the time course of semantic interpretation (e.g. Cook & Myers, 2004; Filik, 2008; Garrod & Terras, 2000; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Murray, 2006; VanBerkum, Hagoort, & Brown, 1999; Warren & McConnell, 2007; see Kuperberg, 2007 for a review of related event-related potential (ERP) work). The current experiments used this strategy to investigate whether the structural and semantic relationships between words in a sentence influence the speed with which they are integrated into a semantic interpretation (cf. Forster & Ryder, 1971; Forster & Olbrei, 1974). Specifically, Experiment 1 investigated whether the number of words intervening between the words cueing implausibility influenced the difficulty of semantic integration, given that linear distance has been shown to make syntactic integration more difficult (e.g. Gibson, 1998). Experiment 2 tested whether readers' eye movements showed earlier disruption to an implausible event if the implausibility was cued by a conflict between words that shared a closer structural relation (e.g., a theta-assigning relation of the kind that holds between verbs and their arguments) compared to a conflict between words that did not participate in such a structural relation.

This investigation is important because it both addresses unresolved questions about what have been termed local versus global influences during semantic interpretation (e.g. Morris, 1994; Hess, Foss & Carroll, 1995; Cook & Myers, 2004) and influences our interpretation of previous experiments. One such experiment, Rayner, Warren, Juhasz, and Liversedge (2004), provides background for this study. They used items like (1–3) and found that the

Correspondence to: Tessa Warren, 607 LRDC, 3939 O'Hara St. Pittsburgh, PA 15260, tessa@pitt.edu Phone: (412) 624-7460.

severity of a plausibility violation is reflected in the latency and magnitude of the eye-movement disruption it incurs.

1. John used a knife to chop the large carrots for dinner.
2. John used an axe to chop the large carrots for dinner.
3. John used a pump to inflate the large carrots for dinner.

Disruption was earlier and stronger on the target word *carrots* when it was severely implausible (as in 3) as opposed to moderately implausible (as in 2), relative to a baseline (1). Warren and McConnell (2007) found the same pattern across similar sentences when (2) and (3) were similarly implausible, but (3) contained a selectional restriction violation and (2) did not. However, in both studies the violation in (3) occurred in a mismatch between a verb and its argument (*inflate* and *carrots* above) rather than between two arguments of a verb (*axe* and *carrots* above) as in (2). Additionally, more words intervened between the mismatching verb and argument than between the mismatching two arguments. If: (1) semantic interpretation is faster within theta-assigning relations, (2) semantic interpretation within a proposition occurs prior to interpretation across propositions, and/or (3) the distance between the words generating an implausibility affects the latency of disruption, these factors could explain Rayner et al. (2004) and Warren and McConnell (2007)'s effects.

These factors, namely the presence of a theta-assigning relation and the linear distance between words, represent potential structurally-relevant local mechanisms influencing semantic interpretation. Previous investigations into the influence of global versus local factors on semantic interpretation have generally focused on local mechanisms hypothesized to operate via lexical representations, such as semantic association (e.g. Morris, 1994; Hess et al., 1995). However, there are multiple reasons to expect that local structural mechanisms might also have effects on semantic interpretation.

First, if language comprehension occurs more quickly within theta-assigning relations than outside of them, words within a theta-assigning relation might be integrated into a semantic interpretation more quickly than words outside of one. Supporting this possibility, multiple experiments have demonstrated that argument phrases, which are assigned thematic roles, are processed more quickly than adjunct phrases, which are not (Clifton, Speer & Abney, 1991; Speer & Clifton, 1998; Shutze & Gibson, 1999; Kennison, 1999). This argument advantage holds for implausible arguments and adjuncts (Boland, 2005; Speer & Clifton, 1998), implicating some mechanism other than semantic predictability, perhaps faster syntactic processing. However, semantic predictability could also play a role. Ferretti, McRae, and Hatherell (2001) found that the lexical representations of verbs include world knowledge about their likely arguments. If semantic predictions for arguments are automatically activated early during verb processing, then mismatches between these predictions and the semantic features of the argument that actually appears might be detected earlier than in cases where the mismatches arose outside of theta-assigning relations and no predictions were made.

Second, if the difficulty of retrieving words during the process of sentence comprehension is related to how recently they have been encountered (e.g. Gibson, 1998), then it might take longer to form an interpretation requiring the joint activation of two words that are further apart in a sentence. Although there is evidence that increasing the number of sentences between an antecedent and an inconsistent anaphor lowers the probability that readers detect the inconsistency (Rayner, Chace, Slattery, & Ashby, 2006), the question of whether distance affects the detection of intra-sentential inconsistencies has not been investigated systematically. But there is considerable evidence that the difficulty of building syntactic dependencies between words increases as the linear distance between them increases (e.g.

Altmann, Garnham, & Henstra, 1998; Gibson, 1998, 2000; Gibson & Warren, 2004; Grodner & Gibson, 2005; Hawkins, 1994). Because syntax and semantics are tightly coupled, some of this evidence could suggest that distance affects the difficulty of building basic semantic or referential relationships. In fact, Warren and Gibson (2002, 2005) have proposed that the most appropriate metric for calculating syntactic distance may be referentially-based. If implausibility is less likely to be detected or is detected later when the words that cue it are farther apart, it might suggest that the memory demands of integrating the meanings of non-adjacent words into an event representation are parallel to those of building syntactic or thematic dependencies.

There is also evidence consistent with both thematic role and distance mechanisms, namely the finding that readers often fail to notice global syntactic violations in sentences in which the local propositions are coherent (Tabor, Galantucci, & Richardson, 2004). Because propositions generally follow argument structure (see Perfetti & Britt, 1995 for discussion), elements within theta-assigning relations will usually appear within the same proposition but arguments of different verbs generally will not. Additionally, words within the same proposition will usually be closer together. Thus the fact that readers sometimes only require local propositional coherence when they comprehend a sentence could indicate an important role for either or both of the proposed local mechanisms under discussion.

Although we have just reviewed considerable evidence suggesting there is reason to expect local structurally-related factors to influence semantic processing and the detection of implausibility, it is also possible that global factors might outweigh these local factors. There have been multiple demonstrations that global context can drive early semantic interpretation even when pitted against local context. For example, Cook and Myers (2004) found that initial eye-movement disruption associated with reading about a locally implausible event (e.g. a busboy taking an order in a restaurant) can be eliminated with the introduction of a prior global justification (e.g. the restaurant was short-staffed), although disruption was apparent in later measures. Filik (2008) found that eye-movement disruption to implausible sentences can be completely eliminated in cartoon contexts, like when the sentence *The mouse picked up the dynamite* is presented in a Tom and Jerry context (but cf. Warren, McConnell, & Rayner, 2008). The local factors manipulated in these experiments (i.e. script or world knowledge) are different from the more structural ones investigated in the current experiments, but in these cases the global context seems to drive at least the early stages of semantic interpretation.

Additionally, there is reason to think that semantic processing may not be subject to the same mechanisms that affect syntactic processing. Whereas syntactic representations are assumed to be hierarchical, semantic representations are not. This could have implications for the process of re-accessing earlier-processed material and might mean that the distance between words may not affect semantic processing in the same way it does syntax. Finally, although features of earlier-processed words can affect both the syntactic and semantic interpretations of upcoming words (e.g. Hare, Elman, Tabaczynski, & McRae, 2009; Patson & Ferreira, 2009), these effects are almost ubiquitous in semantics, where most words can have a spectrum of shades of lexical meaning. This is important because if an initial word influences the interpretation of subsequent words, then a conflict that seems to arise between two non-adjacent words might actually arise more locally. For example, in *John used the axe to chop the large carrots*, if *axe* influences the sense of *chop*, then the conflict may not arise between *axe* and *carrots*, but instead between a chopping-with-an-axe event and *carrots*.

The two studies in this paper test whether local structural factors, namely the presence of a theta-assigning relation and differences in the number of intervening words between words

cueing an implausible event, affect semantic processing, or whether semantic processing is entirely determined by global factors like the degree of implausibility of the entire event. Experiment 1 focused on potential effects of the number of words intervening between the words cueing the implausibility, whereas Experiment 2 investigated the impact of the presence or absence of theta-assigning relations between them.

Experiment 1

Experiment 1 was designed to investigate whether readers are slower to form semantic interpretations that require the joint activation of two words that are further apart in a sentence. Work in text processing has suggested that increasing the distance across which an element must be retrieved generally increases the difficulty of retrieval (e.g. O'Brien, Raney, Albrecht, & Rayner, 1997; Rayner et al., 2006). However, the distance manipulations in these experiments have generally involved the insertion of multiple sentences rather than a few extra words. Dickey and Thompson (2004) investigated whether intra-sentential distance affected anomaly judgments, although their distance manipulation also varied the necessity of computing a filler-gap dependency. Relevant to the current study, their unimpaired college aged participants were marginally slower to detect violations between words that were farther apart. As discussed previously, Rayner et al. (2004)'s finding that readers showed less and later eye-movement disruption to less-severe implausibility violations arising from words that were further apart is also consistent with the hypothesis that intra-sentential distance might affect semantic integration. However, Stewart, Pickering, and Sturt (2004) found a similar pattern in sentences in which distance didn't vary. This conflicting pattern of results leaves open the possibility that greater distance may delay the detection of implausibility and may slow semantic integration in general.

Experiment 1 had a 2×2 design crossing the distance between the words cueing a plausibility violation with the presence/absence of a plausibility violation. If distance affects the time course of semantic interpretation and the latency of violation detection, there should be earlier and/or more eye-movement disruption in the local implausible condition than the non-local implausible condition. The plausible sentences will provide baselines to control for any differences between the local and non-local conditions unrelated to semantic integration. Therefore if there is an interaction between distance and plausibility, with more and/or earlier disruption in the local implausible condition than the non-local implausible condition and less/little difference between the plausible conditions, it would suggest that distance affects the latency of violation detection and the time course of semantic interpretation. If distance does not affect these processes, then there should be a main effect of plausibility, but no effect of distance.

Participants

Forty eight University of Pittsburgh undergraduates participated for course credit. All were native English speakers with normal or corrected-to-normal vision.

Apparatus

An Eyelink 1000 eye-tracker monitored the gaze location of participants' right eyes during reading. The eye-tracker has a spatial resolution better than 30-min of arc and samples gaze location every millisecond. Participants viewed stimuli binocularly on a monitor 63 cm from their eyes; approximately 3 characters equaled 1 degree of visual angle.

Materials

The materials consisted of 32 items with a 2 × 2 design crossing plausibility (implausible vs. plausible) and distance (local vs. non-local). In the implausible conditions, the implausibility arose between the instrument and the patient. In the local conditions, the instrument and the patient were separated by two words. In contrast, in the non-local condition, the instrument and the patient were separated by four words.

(4) After illustrating the research results in a poster, David asked for help. (Local Plaus)

(5) After the research results were illustrated in a poster, David asked for help. (NL Plaus)

(6) After illustrating the research results in a mosaic, David asked for help. (Local Implaus)

(7) After the research results were illustrated in a mosaic, David asked for help. (NL Implaus)

Every item began with a subordinator (*After*). In the local conditions, this was followed by a verb and its patient (*illustrating the research results*), a phrase ending in the target word (*mosaic* or *poster* above), and then a final main clause (*David asked for help*). All words in the main clause were the same across conditions. In the non-local conditions, the subordinator was followed by the patient from the local condition (*the research results*), a passive verb phrase (*were illustrated*), a phrase ending with the target word, and then a final main clause. The target word was generally an instrument (*mosaic* or *poster* above), and in the implausible conditions it was implausible given the patient in the sentence, e.g. many things can be naturally illustrated in a mosaic, but not research results. To increase its likelihood of being fixated, the target word was a minimum of 5 characters long. Target nouns were identical in length (M= 6.84 characters) and matched in frequency (Kuchera & Francis, 1967; implausible M = 5.59; plausible M = 5.75; $t(31) = 0.80$, *n.s.*1). To verify that the items implemented the plausibility manipulation correctly, a rating study was run with ten participants who did not participate in the eye-tracking experiment. Participants rated the naturalness of the event in the initial clause of the sentence (e.g. illustrating research results in a poster or illustrating research results in a mosaic for the example above) on a scale of 1–5, with 1 being very natural and 5 very unnatural. Items were counterbalanced across two presentation lists, so each participant saw one condition of each item. As designed, the plausible conditions were rated as being reliably more natural (M = 1.8) than the implausible conditions (M = 3.88; $t_1(9) = 11.7$, $p < .01$; $t_2(31) = 13.3$, $p < .01$). Individual items' ratings appear in the appendix.

The 32 experimental items were combined with 90 filler items. Conditions were counterbalanced across four presentation lists using a Latin square design. After 45% of the sentences, participants answered a yes/no comprehension question. Half of these required a "yes" response.

Procedure

The experiment lasted approximately 35 – 45 minutes. A chinrest and forehead rest minimized head movements. Participants were asked to read normally, for comprehension, and were told that after some passages they would need to answer a yes/no comprehension question. After the participant was seated at the eye tracker and received experimental instructions, the tracker was aligned and calibrated. It was recalibrated as necessary.

[1]All reported t-tests were two-tailed.

Results

Three analysis regions were defined. The pre-target region was the word preceding the target noun unless it was less than five characters long, in which case the pre-target region was the two words preceding the target noun. The post-target region was the word following the target noun unless it was less than five characters long, in which case the post-target region was the two following words.

Comprehension rates were high (Mean=87.2%, SD=8.9%). Approximately 11% of trials were excluded from analysis due to track losses, blinks, and incomplete trials. Fixations shorter than 80ms, which were not within 1.5 characters of a previous or subsequent fixation, were eliminated. Approximately 93% of all analyzed regions were fixated during the first pass across all conditions.

Five eye-movement measures were computed (Rayner, 1998). *First fixation duration* is the duration of the first fixation on a region during first pass reading. *Gaze duration* is 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*) is the sum of all fixations from first entering a region during first pass reading until leaving it to the right, including regressive fixations. *Regressions out* is the percentage of times a regression was launched from a region during first pass reading. These measures are generally interpreted as reflecting first-pass reading time; although go-past time includes regressions to earlier regions of a sentence; it is a measure of how long it took to move forward in the text past a given region. Finally, *total time* is the sum of all fixations on a region (and combines first and second pass reading time). In all experiments, data were analyzed with repeated measures ANOVAs using participants (*F1*) and items (*F2*) as random factors.

Pre-target Region

There were no effects on gaze duration, go-past time, or regressions out (See Table 1; *p*s>.05). There was a main effect of plausibility on first fixation durations such that they were longer in the implausible than plausible conditions ($F1(1, 47)=4.74$, $p<.05$, $F2(1, 31)=9.14$, $p<.01$), however this effect is likely due to the fact that in some of the items, the pretarget region differed slightly between the plausible and implausible conditions (e.g., *some* was changed to *an*). In analyses excluding those eight items, this effect disappeared, *F*s < 1. Total reading time, which includes re-reading generated by difficulty later in the sentence, also showed a main effect of plausibility with more time spent on the implausible conditions than the plausible conditions ($F1(1, 47)= 18.27$, $p<.01$, $F2(1,31)= 8.49$, $p<.01$).

Target Region

The experimental manipulations did not affect first fixation duration, gaze duration, or first pass regressions out (See Table 2; *p*s>.05). Go-past showed a main effect of plausibility which was fully reliable by participants but marginal by items ($F1(1,47)=9.18$, $p<.01$, $F2(1,31)=3.30$, $p=.079$). Total times showed a main effect of plausibility ($F1(1,47)=16.33$, $p<.01$, $F2(1,31)=9.60$, $p<.01$), with longer times for the implausible conditions as compared to the plausible conditions. Critically, there were no interactions in any measure: first fixation ($F1(1, 47)=.01$, $p>.1$, $F2(1, 31)=.46$, $p>.1$); gaze duration ($F1(1, 47)=.04$, $p>.1$, $F2(1, 31)=.01$, $p>.1$); go-past ($F1(1, 47)=1.88$, $p>.1$, $F2(1, 31)=.76$, $p>.1$); regressions out ($F1(1, 47)=.09$, $p>.1$, $F2(1, 31)=.03$, $p>.1$); total time ($F1(1, 47)=.18$, $p>.1$, $F2(1, 31)=.01$, $p>.1$).

Post-target Region

First fixation, gaze duration, and total time were unaffected by the experimental manipulation (see Table 3; *p*s>.05). There were main effects of both manipulations in go-past times, but these were reliable only in analyses by participants. Go-past was longer in implausible than plausible conditions (*F1*(1,47)=20.99, *p*<.01, *F2*(1,31)=2.70, *p*>.1), and longer in local than non-local conditions (*F1*(1,47)=4.08, *p*<.05, *F2*(1,31)=0.74, *p*>.1). There was a main effect of plausibility in first pass regressions out such that more regressions were made out of the implausible conditions than the plausible conditions (*F1*(1,47)=14.99, *p*<.01, *F2*(1,31)=10.02, *p*<.01). Critically, there was no interaction between plausibility and locality in any measure: first fixation (*F1*(1, 47)=.53, *p*>.1, *F2*(1, 31)=.64, *p*>.1); gaze duration (*F1*(1, 47)=1.09, *p*>.1, *F2*(1, 31)=1.01, *p*>.1); go-past (*F1*(1, 47)=2.69, *p*>.1, *F2*(1, 31)=2.44, *p*>.1); regressions out (*F1*(1, 47)=2.31, *p*>.1, *F2*(1, 31)=2.56, *p*>.1); total time (*F1*(1, 47)=2.31, *p*>.1, *F2*(1, 31)=2.11, *p*>.1).

Target and Post-target combined region

In order to potentially clarify some of the effects that were not fully reliable in either the target or post-target regions, supplemental analyses were undertaken on a combined region made up of both the target and post-target regions. Across this longer region, first fixation and gaze duration remained unaffected by plausibility (*p*s>.05). However, the plausibility effect was fully reliable in go-past (*F1*(1,47)=30.06, *p*<.01, *F2*(1,31)=6.90, *p*<.05), first pass regressions out, (*F1*(1,47)=7.37, *p*<.01, *F2*(1,31)=6.06, *p*<.05), and total times (*F1*(1,47)=19.15, *p*<.01, *F2*(1,31)=8.67, *p*<.05), with longer times for and more regressions from the implausible conditions. There were no reliable interactions in any measure, all *Fs* < 1, *ps* > .1.

Summary of Experiment 1

The findings from Experiment 1 can be summarized as follows: disruption to implausible conditions began to be evident in the go-past measure on the target word and the post-target region, and led to more re-reading of the pre-target and target regions, resulting in longer total reading times in those regions. The time course of this disruption and the pattern of measures in which it appeared were very similar to the pattern in Rayner et al. (2004)'s moderately implausible condition. There was little indication that readers were better or faster at detecting violations in the local condition than in the non-local condition; evidence for this would have taken the form of an interaction between locality and plausibility. Although the numerical pattern of go-past times in the post-target region was in this direction, only main effects were statistically reliable. The results of Experiment 1 suggest that a moderate increase in the number of words intervening between two words cueing a plausibility violation has little or no effect on the speed or success of semantic interpretation.

Experiment 2

Experiment 2 used eye tracking to test whether the presence of a theta-assigning relation between two elements in a sentence causes a semantic mismatch between them to be detected more quickly. The degree of mismatch (implausibility) was kept moderate, so that effects of the presence or absence of a theta-relation could be evaluated independently from the presence of a selectional restriction violation (cf. Warren & McConnell, 2007). Items had the following form:

(8) Bryan used a bottle to feed the hungry infant yesterday morning. (Control)

(9) Bryan used a bottle to fight off the hungry infant yesterday morning. (Theta relation)

(10) Bryan used a trough to feed the hungry infant yesterday morning. (Non-theta relation)

(8) is a plausible baseline condition. In (9), the target word (*infant*) is a relatively implausible argument of the verb (*fight off*), and they appear within the same local proposition. In (10), the target word (*infant*) is a relatively plausible argument of the verb (*feed*), but is relatively implausible with respect to the instrument for this feeding event (*trough*). In this condition, *trough* and *infant* are also unlikely to be represented within the same local proposition (see Perfetti & Britt, 1995 for discussion). Note that the number of words intervening between the words cueing the implausibility also varies between conditions. In the theta-relation condition, two words intervene between *fight off* and *infant*, but in the non-theta-relation condition four words intervene between *trough* and *infant*. However, given the results of Experiment 1, we expect no effect of distance on the current results.

If the presence of a theta-assigning relation between the words cueing implausibility speeds the detection of that implausibility, or if readers build propositions locally before incorporating them into a larger semantic representation, then readers should show earlier eye-movement disruption in (9) than in (10), relative to the baseline in (8). If the presence of a theta-assigning relation has no effect on semantic processing, eye movement disruption should be similar in (9) and (10), but greater than (8).

Method

Participants

Thirty new participants from the same population as Experiment 1 participated in Experiment 2.

Apparatus

The same apparatus was used as in Experiment 1.

Materials

The materials were 27 items with three conditions: a baseline condition, an implausible condition where the violation occurred within a theta relation and an implausible condition where the violation did not occur within a theta relation. An example (8–10) appears above. In the theta-relation condition, the implausibility arose between the verb and its patient; in the non-theta-relation condition, the implausibility arose between the instrument and the patient. Every item began with a proper noun followed by the verb *used*, an instrument (*a bottle* in the control condition above), an infinitival verb (*to feed*), an adjectival noun phrase (*the hungry infant*), and a final adjunct phrase (*yesterday morning*). All words following the infinitival verb were the same for each item across conditions. The theta-relation condition differed from the control condition only in the verb, while the non-theta-relation condition differed only in the instrument. The target word was the noun of the adjectival noun phrase (*infant* above), and the locus of the violation in the theta relation and non-theta-relation conditions. To increase the likelihood that it would be fixated, it was at least 5 characters long.

The 27 experimental items were combined with 41 filler items and 40 items from an unrelated experiment involving relative clauses. Five plausible filler items had the main verb *used*, but continued with a syntactic structure different from the one in the experimental items. Conditions were counterbalanced across three presentation lists using a Latin square design. After 45% of the sentences, participants answered a yes/no comprehension question. Half required a “yes” response.

Norming Studies—Because it was critical for the logic of the experiment that the verb and target noun combination be implausible in the theta-relation condition, a norming study was run to verify this. Twenty participants who did not participate in the eye-tracking experiment rated the naturalness of the event described by the verb-object combination in either the theta-relation condition or the other two conditions (e.g., *fighting off a hungry infant* or *feeding a hungry infant*) on a scale of 1 (very natural)– 7 (very unnatural). Conditions were counterbalanced across two presentation lists. ANOVAs indicated that the verb-object combination was rated reliably less natural in the theta-relation condition (M = 5.33) than the control/non-theta-relation condition (M = 2.28) ($F1(1,19) = 319.01$, $p<.01$, $F2(1,26) = 100.28$, $p<.01$), demonstrating that there was a local implausibility between the verb and object in the theta-relation condition as compared to the other conditions.

Additionally, the two implausible conditions were designed to be equally implausible. To verify this, 27 University of Pittsburgh students who did not participate in the main experiment rated the naturalness of the events in the items truncated after the target word on a scale of 1 (very natural)– 5 (very unnatural). Conditions were counterbalanced across three presentation lists. ANOVAs indicated a reliable main effect of condition, ($F1(2,54) = 153.16$, $p<.01$, $F2(2,52) = 234.95$, $p<.01$). Pair-wise comparisons indicated that the implausible conditions (theta relation M= 3.63; non-theta relation M = 3.77) were rated as being similarly unnatural ($t_1(26) = .54$, n.s., $t_2(26) = .86$, n.s.), and both less natural than the control (M= 1.37), (theta relation vs. control: $t_1(26) = 16.14$, $p<.01$, $t_2(26) = 15.03$, $p<.01$; non-theta relation vs. control: $t_1(26) = 18.52$, $p<.01$, $t_2(26) = 15.32$, $p<.01$).

Procedure

The procedure was the same as in Experiment 1.

Results

Three analysis regions were defined. The pre-target region was the determiner and adjective preceding the target noun. The post-target region was the following word unless it was less than five characters long, in which case the post-target region was the two following words.

Comprehension rates were high (Mean = 94%, SD = 3.2%). Approximately 7% of trials were excluded from analysis due to track losses, blinks, and incomplete trials. Fixations shorter than 80ms which were not within 1.5 characters of a previous or subsequent fixation were eliminated. Approximately 93% of all analyzed regions were fixated during the first pass across all conditions. The same eye movement measures were computed and analyzed as in Experiment 1.

Pre-target Region

The experimental manipulations did not affect first fixation duration, gaze duration, go-past time, or regressions out (See Table 4; ps>.05). However, total reading time, which includes re-reading generated by difficulty later in the sentence, was affected ($F1(2,58) = 14.05$, $p<.01$, $F2(2,52) = 9.14$, $p<.01$). Both implausible conditions had longer total times than the control (theta relation vs. control: $t_1(29) = 5.52$, $p<.01$, $t_2(26) = 4.09$, $p<.01$; non-theta relation vs. control: $t_1(29) = 2.04$, $p<.05$, $t_2(26) = 2.40$, $p<.05$); additionally, the theta-relation condition had longer total times than the non-theta-relation condition ($t_1(29) = 2.99$, $p<.01$, $t_2(26) = 2.18$, $p<.05$).

Target Region

First fixations were numerically, but not reliably (ps>0.05), longer in the implausible conditions than in the control condition (see Table 5). This effect became reliable in gaze

duration ($F1(2,58) = 3.32$, $p<.05$, $F2(2,52) = 6.27$, $p<.01$). Pair-wise comparisons indicated that gaze duration was longer in each implausible condition than in the control (theta relation vs. control: $t_1(29) = 2.33$, $p<.05$, $t_2(26) = 3.01$, $p<.01$; non-theta relation vs. control: $t_1(29) = 2.01$, $p<.05$, $t_2(26) = 3.01$, $p<.01$), but the two implausible conditions did not differ ($p>.05$). Go-past showed the same pattern ($F1(2,58) = 3.41$, $p = .08$, $F2(2,52) = 6.23$, $p<.05$), with the implausible conditions not differing ($p>.05$) but each being longer than the control condition (theta relation vs. control: $t_1(29) = 2.45$, $p<.05$, $t_2(26) = 3.78$, $p<.01$; non-theta relation vs. control: $t_1(29) = 2.19$, $p<.05$, $t_2(26) = 3.20$, $p<.01$). First pass regressions out were unaffected by the experimental manipulation (all ps>.05). Similar to earlier measures, total time was affected ($F1(2,58) = 13.85$, $p<.01$, $F2(2,52) = 18.64$, $p<.01$), with longer times for the implausible conditions as compared to the control (theta relation vs. control: $t_1(29) = 5.36$, $p<.01$, $t_2(26) = 5.07$, $p<.01$; non-theta relation vs. control: $t_1(29) = 2.61$, $p<.05$, $t_2(26) = 3.77$, $p<.01$). Additionally, the theta relation and non-theta-relation conditions diverged in total time, with theta relation longer ($t_1(29) = 2.61$, $p<.05$, $t_2(26) = 3.26$, $p<.01$).

Post-target Region

First fixation and gaze duration were unaffected by the experimental manipulation (see Table 6; ps>.05). However go-past was affected ($F1(2,58) = 5.34$, $p<.01$, $F2(2,52) = 5.01$, $p<.05$). Go-past was longer in the theta-relation condition than the other two conditions (theta relation vs. control: $t_1(29) = 2.62$, $p<.05$, $t_2(26) = 2.23$, $p<.05$; theta relation vs. non-theta relation: $t_1(29) = 2.71$, p<.05, $t_2(26) = 2.72$, $p<.05$), which did not differ ($ps>.05$). The same pattern held for first pass regressions out ($F1(2,58) = 9.60$, $p<.01$, $F2(2,52) = 8.70$, $p<.01$), with the most regressions from the theta-relation condition (theta relation vs. control: $t_1(29) = 4.56$, $p<.01$, $t_2(26) = 3.45$, $p<.01$; theta relation vs. non-theta relation: $t_1(29) = 2.63$, $p<.05$, $t_2(26) = 3.18$, $p<.01$), but no difference between the non-theta relation and control conditions ($ps>.05$). Total time showed the same pattern ($F1(2,58) = 4.67$, $p<.05$, $F2(2,52) = 3.35$, $p<.05$), with longest times in the theta-relation condition (theta relation vs. control: $t_1(29) = 2.66$, $p > .05$, $t_2(26) = 2.30$, $p<.05$; theta relation vs. non-theta relation: $t_1(29) = 2.42$, $p<.05$, $t_2(26) = 2.19$, $p>.05$), and no difference between the non-theta relation and control conditions ($ps>.05$).

Summary of Experiment 2

Experiment 2 indicated that the latency of implausibility detection is relatively insensitive to whether the words cueing the implausibility share a theta-assigning relation or not, appear within the same proposition or not, and are separated by two or four words. In this experiment both implausible conditions first diverged from the control condition in gaze duration on the target word. This pattern continued for go-past on the target, but in subsequent measures including total time on the target and go-past and regressions out on the post-target region, the two implausible conditions began to diverge, with the theta-relation condition showing more disruption.

General Discussion

The two experiments reported in this paper converge in showing that local structural factors like the number of words intervening between two words cueing implausibility, the presence of a theta relation between two words cueing implausibility, and the co-presence of those words within a local proposition, do not affect the initial eye-movement disruption associated with reading a moderately implausible sentence. Disruption associated with plausibility violations was slightly weaker and was apparent in fewer measures in Experiment 1 than Experiment 2. Given that the difference in off-line naturalness ratings between conditions was similar in both experiments, this difference in the degree to which

the eye movement record was affected may simply reflect the high levels of variability that seem to characterize eye movements to higher-level language phenomena across experiments (Clifton, Staub, & Rayner, 2007; Murray, 2000).

Experiment 1's finding that small differences in the number of words intervening between the words cueing a semantic implausibility did not affect the magnitude of initial disruption contrasts with evidence from syntactic integration (e.g. Grodner & Gibson, 2005). Grodner and Gibson argue that syntactic distance effects are related to decay or interference resulting from the approximately serial nature of linguistic input. The current results indicate that this kind of seriality is not enough to drive distance effects; instead syntactic distance effects likely reflect characteristics of the storage and manipulation of hierarchical representations.

The findings of Experiment 2 suggest that a potentially local operation like theta-assignment is not done in isolation from the global context. Although the verb-object propositions in the theta-relation condition (e.g., *fight off the hungry infant*) were less natural than in the non-theta relation/control conditions (e.g., *feed the hungry infant*), both implausible conditions showed the same initial time course of eye-movement disruption and did not diverge until relatively late measures. This suggests that context, in this case the instrument referred to before the verb, may bias the verb's sense and influence its semantic fit with an upcoming object. This finding is related to recent work suggesting that the agent of a verb can affect its interpretation (e.g. Hare et al., 2009; Patson & Ferreira, 2009), because it demonstrates that instruments can also affect interpretation. The lack of word- or proposition-based distance effects or theta-assigning effects in Experiments 1 and 2 suggests that none of these factors drove Rayner et al. (2004) or Warren and McConnell (2007)'s effects.

One possible explanation for the sometimes greater and longer-lasting disruption in the theta-relation condition in Experiment 2 could come from the local propositional coherence account discussed in the introduction. Considerable work suggests that when the parser cannot form a globally coherent interpretation it will settle for a set of locally coherent structures (Tabor et al., 2004; Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira & Patson, 2007). Although the current data suggest that readers do not initially build local propositions and only subsequently integrate them, local propositional coherence may be important to later processes attempting to make sense of an implausible discourse model. It may be more difficult to evoke a model for an implausible proposition, e.g. imagining a situation wherein one would fight off an infant, than to evoke one for an implausibility that comes about through the conjunction of two plausible propositions (e.g., imagining a situation wherein an unusual instrument would be used in an otherwise natural event).

The results reported in this paper are consistent with a growing body of work suggesting that the early stages of semantic interpretation are not solely determined by local relations and representations, but are influenced by world and contextual knowledge (e.g. Cook & Myers, 2004; Filik, 2008; Garrod & Terras, 2000; Hagoort et al, 2004; van Berkum et al., 1999). In the experiments in this paper, early semantic processing was unaffected by variations in the word- and proposition-based distance between two words as well as the presence or absence of a theta-assigning relation between them. Instead, early eye-movement disruption seemed to reflect the critical event's overall naturalness, suggesting that the entire event representation is affecting early processing. This indicates that the structurally-relevant local factors investigated here are less important than the global context in early semantic interpretation. It is important to note, however, that there is evidence that other structurally-relevant local factors like the presence of strong semantic mismatches (i.e., selectional restriction violations) may initially override the global context during semantic interpretation (Warren, et al., 2008). This suggests that both local and global factors

influence early semantic interpretation, perhaps with the global context dominant in the absence of a severe local violation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH grants HD048990 and HD053639. We thank Kerry McConnell, Amanda Virbitsky, Erik Reichle, the Pitt Reading and Language lab, and the audience of Psychonomics 2007; as well as Rhonda McClain and Alison Trude for assistance in data collection.

References

Altmann G, Garnham A, Henstra J. Late closure in context. Journal of Memory and Language. 1998; 38:459–484.

Boland JE. Visual arguments. Cognition. 2005; 95:237–274. [PubMed: 15788159]

Christianson K, Hollingworth A, Halliwell JF, Ferreira F. Thematic roles assigned along the garden-path linger. Cognitive Psychology. 2001; 42:368–407. [PubMed: 11368528]

Clifton C Jr, Speer S, Abney S. Parsing arguments: Phrase structure and argument structure as determinants of initial parsing decisions. Journal of Memory and Language. 1991; 30:251–271.

Clifton, C., Jr; Staub, A.; Rayner, K. Eye movements in reading words and sentences. In: van Gompel, RPG.; Fisher, MH.; Murray, WS.; Hill, RL., editors. Eye movements: a window on mind and brain. Amsterdam: Elsevier, North-Holland; 2007. p. 341-371.

Cook A, Myers JL. Processing discourse roles in scripted narratives: The influence of context and world knowledge. Journal of Memory and Language. 2004; 50:268–288.

Dickey MW, Thompson CK. The resolution and recovery of filler-gap dependencies in aphasia: Evidence from on-line anomaly detection. Brain and Language. 2004; 88(1):108–127. [PubMed: 14698736]

Ferreira F, Patson ND. The 'Good Enough' approach to language comprehension. Language and Linguistics Compass. 2007; 1:71–83.

Ferretti TR, McRae K, Hatherell A. Integrating verbs, situation schemas, and thematic role concepts. Journal of Memory and Language. 2001; 44:516–547.

Filik R. Contextual override of pragmatic anomalies: Evidence from eye movements. Cognition. 2008; 106:1038–46. [PubMed: 17524387]

Forster KI, Olbrei I. Semantic heuristics and semantic analysis. Cognition. 1974; 2:319–347.

Forster KI, Ryder L. Perceiving the structure and meaning of sentences. Journal of Verbal Learning and Verbal Behavior. 1971; 10:285–296.

Garrod S, Terras M. The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. Journal of Memory and Language. 2000; 42:526–544.

Gibson E. Linguistic complexity: Locality of syntactic dependencies. Cognition. 1998; 68:1–76. [PubMed: 9775516]

Gibson, E. The dependency locality theory: A distance-based theory of linguistic complexity. In: Miyashita, Y.; Marantz, A.; O'Neil, W., editors. Image, language, brain. Cambridge, MA: MIT Press; 2000. p. 95-126.

Gibson E, Warren T. Reading time evidence for intermediate linguistic structure in long- distance dependencies. Syntax. 2004; 7:55–78.

Grodner D, Gibson E. Consequences of the serial nature of linguistic input for sentential complexity. Cognitive Science. 2005; 29:261–290. [PubMed: 21702774]

Hagoort P, Hald L, Bastiaansen MCM, Petersson KM. Integration of word meaning and world knowledge in language comprehension. Science. 2004; 304:438–440. [PubMed: 15031438]

Hare M, Elman JL, Tabaczynski T, McRae K. The wind chilled the spectators, but the wine just chilled: Sense, structure, and sentence comprehension. Cognitive Science. 2009; 33(4):610–628. [PubMed: 19750146]

Hawkins, JA. A Performance Theory of Order and Constituency. New York, New York: Cambridge University Press; 1994.

Hess DJ, Foss DJ, Carroll P. Effects of local and global context on lexical processing during language comprehension. Journal of Experimental Psychology: General. 1995; 124:62–82.

Kennison SM. Processing agentive “by”-phrases in complex event and nonevent nominals. Linguistic Inquiry. 1999; 30:502–508.

Kucera; Francis, WN. Computational Analysis of Present-Day American English. Providence: Brown University Press; 1967.

Kuperberg GR. Neural mechanisms of language comprehension: challenges to syntax. Brain Research. 2007; 1146:23–49. [PubMed: 17400197]

Morris RK. Lexical and message-level sentence context effects on fixations times in reading. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1994; 20(1):92–103.

Murray, WS. Sentence processing: Issues and measures. In: Kennedy, A.; Radach, R.; Heller, D.; Pynte, J., editors. Reading as a Perceptual Process. Oxford: Elsevier; 2000. p. 649-664.

Murray WS. The nature and time course of pragmatic plausibility effects. Journal of Psycholinguistic Research. 2006; 35:79–99. [PubMed: 16502145]

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]

O’Brien E, Raney GE, Albrecht JE, Rayner K. Processes involved in the resolution of explicit anaphors. Discourse Processes. 1997; 23:1–24.

Patson ND, Ferreira F. Conceptual plural information is used to guide early parsing decisions: Evidence from garden-path sentences with reciprocal verbs. Journal of Memory and Language. 2009; 60(4):464–486. [PubMed: 20161247]

Perfetti, CA.; Britt, MA. Where do propositions come from?. In: Weaver, CA.; Mannes, S.; Fletcher, CR., editors. Discourse Comprehension: Essays in Honor of Walter Kintsch. Hillsdale, NJ: Lawrence Erlbaum Assoc; 1995.

Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin and Review. 1998; 124:372–422.

Rayner K, Chace KH, Slattery TJ, Ashby J. Eye movements as reflections of comprehension processes in reading. Scientific Studies of Reading. 2006; 10(3):241–255.

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.

Shutze CT, Gibson E. Argumenthood and English prepositional phrase attachment. Journal of Memory and Language. 1999; 40:409–431.

Speer SR, Clifton C. Plausibility and argument structure in sentence comprehension. Memory & Cognition. 1998; 26:965–978.

Stewart AJ, Pickering MJ, Sturt P. Using eye movements during reading as an implicit measure of the acceptability of brand extensions. Applied Cognitive Psychology. 2004; 18:697–709.

Tabor W, Galantucci B, Richardson D. Effects of merely local syntactic coherence on sentence processing. Journal of Memory and Language. 2004; 50:355–370.

van Berkum JJA, Hagoort P, Brown CM. Semantic integration in sentences and discourse: Evidence from the N400. Journal of Cognitive Neuroscience. 2000; 11:657–671. [PubMed: 10601747]

Warren T, Gibson E. The influence of referential processing on sentence complexity. Cognition. 2002; 85:79–112. [PubMed: 12086714]

Warren T, Gibson E. Effects of NP-type on reading English clefts. Language and Cognitive Processes. 2005; 20:751–767.

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]

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

Table 1

Average fixation times (standard deviations) on the pre-target region in Experiment 1 in milliseconds.

	First Fix	Gaze	Go-Past	% Reg out	Total Time
Local Plaus	212 (60)	275 (82)	326 (118)	.11 (.19)	414 (190)
NL Plaus	225 (48)	283 (84)	340 (119)	.11 (.15)	404 (140)
Local Implaus	229 (46)	285 (72)	323 (95)	.10 (.13)	506 (168)
NL Implaus	232 (57)	285 (89)	318 (112)	.09 (.15)	488 (256)

Q J Exp Psychol (Hove). Author manuscript; available in PMC 2011 August 1.

Table 2

Average fixation times (standard deviations) on the target region in Experiment 1 in milliseconds.

	First Fix	Gaze	Go-Past	% Reg out	Total Time
Local Plaus	242 (50)	304 (81)	408 (144)	.22 (.17)	412 (136)
NL Plaus	239 (45)	301 (80)	396 (120)	.26 (.22)	401 (112)
Local Implaus	241 (45)	290 (62)	427 (196)	.26 (.25)	482 (154)
NL Implaus	240 (45)	291 (62)	481 (229)	.32 (.23)	487 (179)

Table 3

Average fixation times (standard deviations) on the post-target region in Experiment 1 in milliseconds.

	First Fix	Gaze	Go-Past	% Reg out	Total Time
Local Plaus	232 (44)	277 (64)	325 (137)	.07 (.10)	396 (130)
NL Plaus	228 (48)	278 (95)	315 (117)	.07 (.11)	421 (154)
Local Implaus	235 (57)	301 (99)	430 (179)	.18 (.22)	435 (143)
NL Implaus	224 (42)	279 (70)	356 (157)	.12 (.18)	433 (150)

Table 4

Average fixation times (standard deviations) on the pre-target region in Experiment 2 in milliseconds.

	First Fix	Gaze	Go-Past	% Reg out	Total Time
Control	222 (47)	363 (86)	401 (111)	5 (10)	443 (153)
Theta relation	220 (38)	354 (82)	407 (115)	8 (9)	582 (175)
Non-theta relation	227 (45)	359 (83)	413 (106)	8 (9)	493 (162)

Table 5

Average fixation times (standard deviations) on the target region in Experiment 2 in milliseconds.

	First Fix	Gaze	Go-Past	% Reg out	Total Time
Control	219 (36)	266 (63)	314 (99)	12 (14)	336 (128)
Theta relation	233 (35)	294 (55)	359 (98)	14 (11)	485 (140)
Non-theta relation	232 (39)	285 (65)	362 (129)	16 (12)	407 (172)

Table 6

Average fixation times (standard deviations) on the post-target region in Experiment 2 in milliseconds.

	First Fix	Gaze	Go-Past	% Reg out	Total Time
Control	236 (57)	410 (138)	689 (324)	29 (22)	486 (163)
Theta relation	239 (52)	381 (110)	912 (349)	48 (22)	562 (229)
Non-theta relation	252 (61)	386 (101)	686 (340)	34 (26)	592 (189)