Why Do Readers Answer Questions Incorrectly After Reading Garden-path Sentences?

Zhiying Qian (zqian3@illinois.edu)
University of Illinois at Urbana-Champaign, 405 N. Mathews Ave, Urbana, IL 61801 USA

Susan Garnsey(sgarnsey@psych.illinois.edu)
University of Illinois at Urbana-Champaign, 405 N. Mathews Ave, Urbana, IL 61801 USA

Abstract

Readers misinterpret garden-path sentences such as *While the man hunted the deer that was brown and graceful ran into the woods* as meaning *The man hunted the deer that was brown and graceful and the deer ran into the woods*. The "Good-enough" processing account proposes that misinterpretation occurs when readers are satisfied with the interpretation derived from the first-pass parse, and thus do not bother to fully reanalyze the sentence (Ferreira et al., 2001; Christianson et al., 2001). Such an account predicts that there should be more evidence of reanalysis at the disambiguating verb (*ran*) on trials with correct responses to the question *Did the man hunt the deer*?, than on those with incorrect responses. The present study tested this prediction using separate self-paced reading and event-related brain potential (ERP) experiments. Results from Experiment 1 (self-paced reading) showed no difference in the reading time at the disambiguating verb between trials that were answered correctly and those that were answered incorrectly. Experiment 2 (ERP) corroborated this finding by showing no difference in the amplitude of the P600 component elicited by the disambiguating verb in trials with correct responses and those with incorrect responses. However, results from a norming experiment showed that plausibility information significantly predicted question accuracy in both experiments. Overall, these results suggest that responses to questions intended to probe whether garden-path sentences are fully reanalyzed do not always answer that question, but can instead be determined primarily by the plausibility of the events described in that question.

Keywords: lingering misinterpretation; reanalysis, good-enough processing; ERPs, P600

Introduction

It has been well established that when reading sentences like (1), readers slow down at the main clause verb *ran*, because they have initially interpreted the noun phrase *the deer that was brown and graceful* as the object of the subordinate clause verb *hunted*. At the main clause verb *ran*, the parser realizes that *ran* lacks a subject and triggers reanalysis. This is termed the garden-path effect. Successful reanalysis would lead to the noun phrase being deleted from the object role of the subordinate clause verb *hunted* and attached to the main clause verb *ran* as its subject.

(1) While the man hunted the deer that was brown and graceful ran into the woods.
(2) Did the man hunt the deer?
(3) Did the deer run into the woods?

Traditional accounts of sentence processing, regardless of whether they are serial or parallel, assume that readers ultimately reach the correct interpretation after reading garden-path sentences such as (1). However, several studies have shown that readers do not always arrive at the correct interpretation (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Christianson, Williams, Zacks, & Ferreira, 2006; Ferreira, Christianson, & Hollingworth, 2001). Rather, they tend to misinterpret the sentence as meaning that *The man hunted the deer that was brown and graceful and the deer ran into the woods*, as evidenced by a high rate of erroneous "yes" responses to (2), although the correct interpretation licensed by syntax should be *The man hunted something unspecified and the deer that was brown and graceful ran into the woods*. Such misinterpretation arises because the initial misinterpretation derived from first-pass analysis (*the man hunted the deer*) persists.

One criticism of this interpretation of incorrect responses to questions like (2) after sentences like (1) concerns the possibility that misinterpretations may not entirely result from garden-pathing. In Ferreira et al. (2001), readers answered "yes" erroneously to (2) 73% of the time after reading (1) and 52% of the time after reading the comma-disambiguated version of (1): *While the man hunted, the deer that was brown and graceful ran into the woods*. Although the difference in question accuracy between ambiguous and unambiguous versions of (1) could be ascribed to garden-pathing, the fact that the error rate was still quite high for unambiguous sentences suggests that at least a portion of the misinterpretations did not result from garden-pathing. Readers may have answered questions based in part on inferences they made after reading this type of sentence: when a deer runs into the woods while a man is hunting, it seems highly likely that what the man is hunting is the deer that is mentioned.

To reduce the likelihood that readers would answer questions based on such inferences, Christianson et al. (2001) used sentences and questions such as (4) and (5).

(4) While Anna dressed the baby who was cute and small spit up on the bed.
(5) Did Anna dress the baby?

Unlike in (1), the subordinate clause verb *dressed* in (4) is a Reflexive Absolute Verb (RAT), which takes its subject as its object if no object is explicitly mentioned. Therefore, whereas in (1), correct reanalysis would result in the object of the subordinate verb (*hunted*) being unspecified and lead to the interpretation that *the man hunted something that is unknown*, correct reanalysis of (4) would result in the interpretation that *Anna dressed herself*, rather than somebody unknown. Christianson et al. showed that erroneous "yes" responses to (5) after reading comma-disambiguated version of (4) was much lower than after reading the ambiguous version (ambiguous: 57%; unambiguous: 12%). Taken together, readers' incorrect answers to comprehension questions after reading sentences like (1) and (4) indicates that the initial misinterpretation derived from first-pass analysis persists to the end of the sentence.

Several types of accounts have been proposed to explain lingering initial misinterpretations. Perhaps the parser builds a shallow (Clahsen & Felser, 2006; Frisson, 2009) or underspecified (Ferreira, Bailey, & Ferraro, 2002; Sanford & Sturt, 2002; Swets, Desmet, Clifton, & Ferreira, 2008) linguistic representation. Other accounts posit that memory traces of the initial misparse (Kaschak & Glenberg, 2004) or shallow semantic processing (Barton & Sanford, 1993) are responsible.

The major processing account that explains lingering misinterpretation is the Good-Enough Account, which proposes that reanalysis is not an "All-or-Nothing" process (see Fodor & Inoue, 1998; Lau & Ferreira, 2005, for similar explanations). Two possibilities could lead to lingering misinterpretation. The first is that when initial interpretation is semantically sensible, readers do not bother to fully reanalyze the syntactic structure of the sentence even though later information turns out to be incompatible with the structure that has already been built. The fact that readers answer "yes" correctly most of the time to questions like (3) that probed comprehension of the matrix sentence after reading (1) indicated that reanalysis started and was carried out to such an extent that the ambiguous noun (*the deer*) was successfully attached to the matrix clause verb. However, reanalysis was not completed, so the noun phrase was not successfully detached from the object role of the subordinate clause (Christianson et al., 2001; Ferreira et al., 2001). The second possibility is that reanalysis of the syntactic structure is completed, but interpretations from both the first-pass analysis and reanalysis linger. Both possibilities could lead readers to incorrectly answer "yes, the man hunted the deer". The present study aims to test the first possibility, i.e., lingering misinterpretation results from incomplete reanalysis.

Such an account predicts more evidence of reanalysis in trials that are interpreted correctly than those that are interpreted incorrectly. We seek to test this prediction by comparing reading times and the amplitude of the P600 ERP component at the disambiguating verb (*ran*) in sentences like (1) between correctly and incorrectly answered trials. Since more reanalysis effort should result in more time spent on the disambiguating verb, slower reading times for trials with correct responses than for those with incorrect responses would support the idea that incomplete reanalysis leads to lingering misinterpretation. The P600 ERP component, which is a positive-going deflection occurring 600-900 ms after the onset of the word that triggers it, has been found to be associated with reanalysis (Osterhout, Holcomb, & Swinney, 1994), so the prediction here was that P600 amplitude in response to the disambiguating verb should be larger in trials with correct responses than in those with incorrect responses, indicating more reanalysis leading to more correct responses..

Previous studies also showed that readers misinterpreted the sentence more often when the ambiguous noun was modified by post-nominal modification (*the deer that was brown and graceful*) than by pre-nominal or no modification (*the brown and graceful deer/deer*) (Ferreira & Henderson, 1991, 1998). The present study used post-nominal modification, because such sentences promote garden-pathing, maximize ambiguity effects at the disambiguating verb, making it more likely that reading times and P600 amplitude will show differences between correctly-answered and incorrectly-answered trials.

Experiment 1

Participants

Thirty-two native speakers of English (12 males; mean age 18.5; range 18-21), who were undergraduate students at the University of Illinois at Urbana-Champaign, participated for course credit. All had normal or corrected-to-normal vision and gave written informed consent.

Materials and Design

Forty sets of sentences such as (1) (repeated here as 6) were taken from Christianson et al. (2001), with each set containing an ambiguous and a comma-disambiguated version. Ambiguous nouns were followed by a post-nominal relative clause containing two adjectives.

(6) While the man hunted(,) the deer that was brown and graceful ran into the woods.

Experimental sentences were distributed over two lists using a Latin Square design, so that each participant saw equal numbers of items in each condition and only one version of each item. Each sentence was followed by a comprehension question that probed the comprehension of the subordinate clause (e.g., *Did the man hunt the deer*?).

One hundred and sixteen distractor sentences were added to each list for a total of 156 trials/list. There were three types of distractors: (1) unambiguous sentences with subordinate-matrix clause order (e.g., *While Jennifer held the cigar that was aged and expensive she told bad jokes.*, 40 sentences); (2) unambiguous sentences with matrix-subordinate clause order (e.g., *The mother comforted the toddler who was chubby and scared while the clown handed him a balloon.,* 40 sentences); and (3) items from a separate experiment with ambiguous and unambiguous sentences distributed over two lists (e.g., *As Jane and Mary met(,) the men from Florida drove past them*, 12 items; *While Anna dressed the baby(,) who was cute and small spit up on the bed*, 24 items). All distractors were followed by a comprehension question that targeted various parts of the sentences (except that questions to items from the separate experiment targeted comprehension of the subordinate clause). Answers to true distractors were half "yes" half "no" across the experiment.

Procedure

Participants sat in a dimly lit and sound-attenuated booth in front of a 23-inch LCD monitor. To make presentation mode comparable with the ERP experiment (Experiment 2), sentences were presented one word at a time in white font on a black background in the center of the screen. Each trial began with a "Ready" sign that stayed on the screen for one second. Each time participants pressed a button on a Cedrus-830 response box, the next word appeared to replace the previous word. Following each sentence, a comprehension question was presented and participants pressed one of two buttons to indicate their answers. Feedback about question accuracy was not given. However, a "Too Slow" feedback message was presented if participants did not make a response within four seconds. The entire experiment took approximately forty minutes to complete.

Results

Comprehension Accuracy All participants were above 80% in comprehension accuracy to distractor items (range: 80%-97%, mean: 90%), suggesting that they were paying attention to the sentences.

Comprehension accuracy to experimental sentences was analyzed using logistic regression to see whether ambiguity predicted response accuracy. Results showed that readers made significantly more erroneous "yes" responses following ambiguous sentences compared to unambiguous sentences (β=1.0, z=5.0, p<.001), as shown in Table 1.

Table 1: Percentage of incorrect responses to comprehension questions in Experiment 1.

Condition	Error Rate
Ambiguous	67%
Unambiguous	51%

Reading Times Prior to data analysis, word-by-word reading times that were faster than 100 millisecond (ms) or slower than 2000 ms were excluded, leading to a loss of 0.5% of data. Reading times above or below 2.5 standard deviations from each participant's mean were replaced by the cut-off value for that participant (Sturt, Pickering, & Crocker, 1999), and word-by-word reading times for sentences that participants failed to make response to the comprehension question were excluded from further analysis. To remove individual differences in reading speed, results presented below were based on length-corrected residual reading times computed separately for each participant by entering their reading times for each word in all sentences (including distractors) into a regression equation that took reading time as the dependent variable and word length as the independent variable, and then subtracted the predicted reading times from the actual reading times (Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994). Residual reading times were then analyzed using mixed-effect models that included ambiguity and reading time as fixed effects, and participant and item as random effects. Random effects structure was further determined using a forward-selection approach. Random slopes for each participant and each item that significantly improved the model fit in the likelihood ratio test (p<.05) were included in the model.

Results at the disambiguating region, which contained the disambiguating verb and the word immediately following it (e.g., *ran into*) showed a significant main effect of ambiguity (β=31.1, t=4.95, p<.001), with the ambiguous condition being read slower than the unambiguous condition (ambiguous: mean 449 ms; mean unambiguous: 419 ms), as illustrated below in Figure 1.

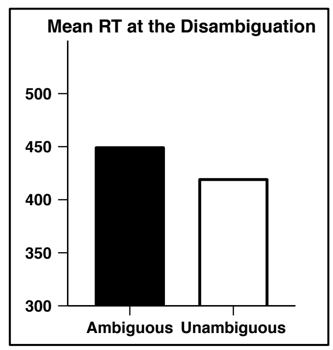

Figure 1: Reading time at the disambiguating region in Experiment 1, collapsing over question accuracy.

There was also a significant interaction between ambiguity and question accuracy (β=27.2, t=2.1, p<.05), which resulted because the disambiguating region was read more slowly in ambiguous sentences that were answered correctly than in those that were answered incorrectly (mean ambiguous correct: 455 ms; mean ambiguous incorrect: 445 ms), and this pattern was reversed in unambiguous sentences (mean unambiguous correct 409 ms; mean unambiguous incorrect 430 ms), as shown below in Figure 2. Pairwise comparisons showed that the differences in reading times between trials with correct answers and those with incorrect answers did not reach significance within either ambiguous or unambiguous conditions (ts<1, ps>.3). Analysis done for the region after the disambiguating verb showed similar results: there was no difference in the reading time between trials that were answered correctly and those that were answered incorrectly.

Overall, Experiment 1 did not show that slower reading time at the disambiguating region was associated with more correct responses in ambiguous sentences. Therefore, there was no evidence that more reanalysis effort was associated with more correct question responses, contrary to the hypothesis that lingering misinterpretation was resulted from incomplete reanalysis.

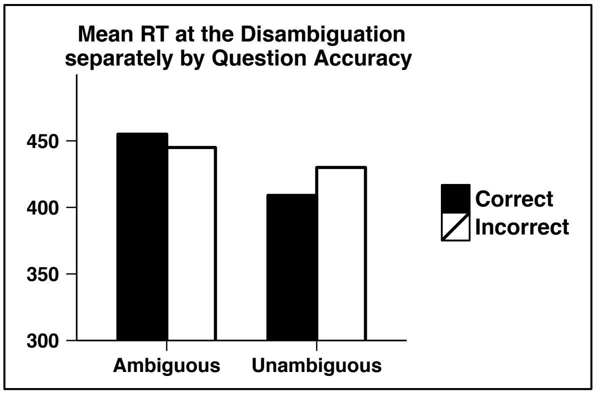

Figure 2: Reading time at the disambiguating region in Experiment 1, separately by response accuracy.

Experiment 2

Experiment 2 used event-related brain potentials (ERPs) to compare the amplitude of P600 evoked by the disambiguating verb in trials with correct responses and those with incorrect responses, since the P600 component has been found to be associated with revision processes in garden-path sentences (e.g., Osterhout et al. 1994), among other things. If trials with larger P600 amplitude at the disambiguating verb are associated with more correct responses, such a result will support that more reanalysis leads to less lingering misinterpretation.

Participants

Participants were fifty-five native speakers of English (26 males; mean age 19.2, range 18-22) who were undergraduate students at the University of Illinois at Urbana-Champaign. All were strongly right-handed as assessed by the Edinburgh inventory (Oldfield, 1971), had normal or corrected-to-normal vision and no neurological or psychiatric disorder. All gave written informed consent and received course credits or payment for taking part. Data from seven participants were excluded from analysis due to low comprehension accuracy to distractors.

Materials

Critical sentences in Experiment 2 were exactly the same as Experiment 1. There were 144 distractors, sixty-four of which were experimental sentences of an unrelated experiment that examined sentences with direct object/sentential complement ambiguity (e.g., *The bus driver warned the passengers would get too rowdy.*). As in Experiment 1, a comprehension question was asked after each sentence.

Procedure

Participants were seated comfortably in a dimly lit and sound-attenuating booth in front of a 23-inch LCD monitor. Each trial began with a fixation point, which stayed in the center of the screen for 500 milliseconds. Because eye movements cause artifacts that contaminate the EEG signal, sentences were presented word-by-word at the center of the screen in 26-point white Arial font on a black background, at a rate of 400 ms per word (300 ms text, 100 ms blank screen).

At the end of each sentence, a comprehension question was presented and participants responded by pressing one of two buttons on a Cedrus RB-830 response box. No feedback was given regarding question accuracy. A total of 184 trials were divided into four blocks, each beginning with four distractor items. Participants took a short break after each block and they were instructed to minimize blinking and body movement while the sentences were being presented. They were encouraged to blink between trials when necessary. A practice block with five sentences was given at the beginning. The recording session took about 40 minutes and the entire session lasted approximately 2 hours.

EEG Recording and Data Analysis

Continuous EEG was recorded from 27 Ag/AgCl sintered electrodes placed in an elastic cap (EasyCap, 10-10 system; Chatrian, 1985), referenced online to the left mastoid and re-referenced offline to the average of left and right mastoids: midline: Fz, Cz, Pz; lateral: AF3/4, F3/4, F7/8, FT7/8, FC3/4, C3/4, T3/4, CP3/4, T5/T6, P3/4, P5/6, PO7/8. Blinks and eye movements were detected with electrodes above and beneath the right eye and at the outer canthi of both eyes. EEG and EOG recordings were amplified by a Grass Model 12 amplifier and sampled at a frequency of 200 Hz. A 0.01-30 Hz analog bandpass filter was applied during online recording and a 0.1 Hz high-pass digital filter was applied offline. Impedances were maintained below 5kΩ.

Epochs were extracted from the continuous waveforms from 100 ms before the onset of the disambiguating verb through 1200 ms later. Epochs contaminated with artifacts were discarded, affecting 11% trials.

ERPs time-locked to the onset of the disambiguating verb relative to a 100 ms baseline immediately preceding it were obtained for each channel in each condition for each participant. The conventional time window of 600-900 ms after the onset of the disambiguating verb was chosen to capture effects on the P600 component. Window mean amplitudes were submitted to separate repeated-measures analyses of variance. One set of analyses included all lateral electrodes and another included just midline electrodes. The ANOVA including all lateral electrodes had four within-subject factors: two levels of ambiguity (Ambiguous, Unambiguous), two levels of question accuracy (Correct, Incorrect), three levels of electrode site anteriority (Frontal, Central, Posterior) and two levels of electrode site laterality (Left, Right). The ANOVA including just midline electrodes (Fz, Cz, Pz) consisted of the same three within-subject factors except that there was no laterality factor. The Greenhouse-Geisser correction was applied wherever necessary to correct for violations of sphericity (Greenhouse & Geisser, 1959). Corrected *p*-values and original degrees of freedom are reported. Grand average ERPs were digitally low-pass filtered at 10 Hz to smooth the waveforms for display, but analyses were performed before such filtering was applied.

Results

Comprehension accuracy Comprehension accuracy to distractor items was analyzed to evaluate whether or not participants were paying attention to the sentences. Similar to Experiment 1, average accuracy to distractors was 89%. Data from seven participants whose comprehension accuracy was below 80% were excluded from further analyses, leaving data from forty-eight participants.

Comprehension question accuracy for experimental sentences was analyzed using logistic regression. Results showed that, consistent with Experiment 1, readers made significantly more erroneous "yes" responses following ambiguous than unambiguous sentences (β=0.6, z=3.8, p<.01), as shown in Table 2 below.

Table 2: Percentage of incorrect responses to comprehension questions in Experiment 2.

Condition	Error Rate
Ambiguous	57%
Unambiguous	46%

ERP results Visual inspection revealed that responses to the disambiguating verb in the ambiguous condition were more positive than in the unambiguous condition at central-posterior sites, showing the usual scalp distribution of P600 effects, as shown in Figure 3.

However, there was no difference in the amplitude of the P600 between trials that were answered correctly and those that were answered incorrectly, as shown in Figure 4. These observations were confirmed by statistical analyses. ANOVA over all lateral electrodes showed a significant interaction between ambiguity and anteriority, F(2,94)=4.7, p<.05, which resulted because the ambiguity effect was significant at posterior electrodes, marginally significant at central electrodes, but not significant at frontal electrodes (Frontal: F(1,47)=.1, p=.8; Central: F(1,47)=3.5, p=.07; Posterior: F(1,47)=7.4, p<.01). There was no main effect or any interaction involving the accuracy factor in lateral electrodes. Midline analysis did not show main effects of either ambiguity or accuracy, nor was there interaction between ambiguity and accuracy.

These results indicated that the disambiguating verb (e.g., *ran*) triggered more syntactic reanalysis in ambiguous than unambiguous sentences, but that P600 amplitude was not associated with question accuracy, which is again contrary to one of the predictions of the good-enough processing account.

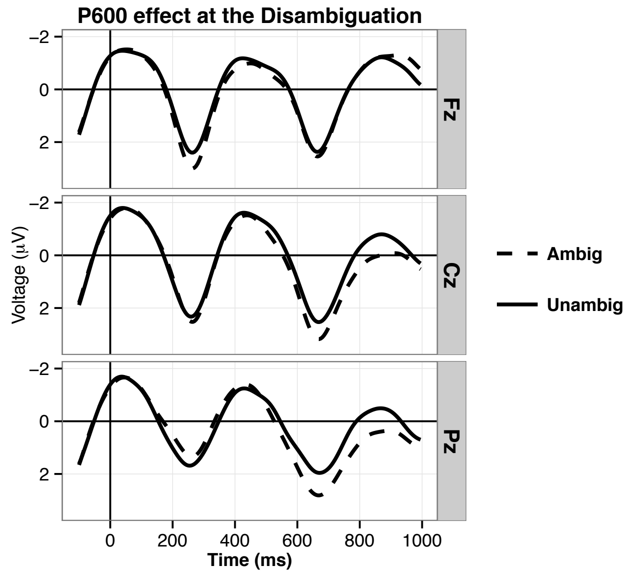

Figure 3: Grand average ERPs of the disambiguating verb at midline electrodes in Ambiguous and Unambiguous conditions in Experiment 2, baselined on 100 ms before the onset of the disambiguating verb.

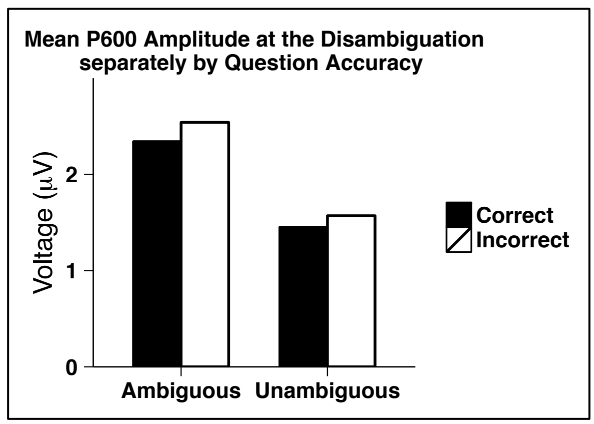

Figure 4: Mean amplitude of the P600 component averaging across all centro-parietal electrodes, separately by response accuracy.

Experiment 3

Experiment 3 aimed to examine whether or not properties of the events described in the sentences and comprehension questions (*Did the man hunt the deer?*) might affect how likely participants were to incorrectly respond "yes" to them.

Method

Twenty-six native speakers of English (14 males; mean age 19.9; range 18-22) were asked to provide a percentage rating to questions like *How likely is it that the man hunted the deer?*, after reading sentences like *While the man hunted(,) the deer that was brown and graceful ran into the woods*. Materials were ambiguous and unambiguous versions of the forty items used in Experiments 1 and 2, and were distributed over two lists. Twenty-four fillers, which were ambiguous and unambiguous versions of sentences like *While Anna dressed(,) the baby who was cute and small spit up on the bed*, were added to each list. Lists were randomized for each participant. Item-by-item plausibility rating was obtained by averaging across participants.

Results

The average likelihood rating across all items was 71%. Questions following ambiguous sentences were rated significantly more likely than those following unambiguous sentences (mean ambiguous: 78%; mean unambiguous: 65%; t=3.9, p<.001).

After entering likelihood information into the multi-level models as an independent variable to examine its effect on comprehension accuracy, results of both Experiment 1 and Experiment 2 showed a main effect of likelihood (Experiment 1: β=0.02, z=3.0, p<.01; Experiment 2: β=0.02, z=3.7, p<.001), suggesting that the likelihood of the events described in the sentences and comprehension questions reliably predicted question responses. Questions describing more likely events led to more incorrect "yes" responses. The main effect of ambiguity remained significant in Experiment 1 (β=0.39, z=3.2, p<.01) and marginally significant in Experiment 2 (β=0.34, z=1.8, p=.07), showing that likelihood had separable effects on question accuracy.

Taken together, these findings suggest that response accuracy to comprehension questions following garden-path sentences such as (1) in Experiments 1 and 2 was determined in part by the likelihood of the events described in those questions. Therefore, responses to such questions does not seem to provide good evidence about whether readers reanalyzed initial misinterpretation in garden-path sentences.

Conclusion

Experiments 1 and 2 converged on showing that there was no evidence of more reanalysis during trials with correct question responses than during those with incorrect question responses. In Experiment 1, readers did not spend more time reading the disambiguating verb and in Experiment 2 P600 amplitude was not larger at the disambiguating verb on trials with correctly answered questions. These results are not consistent with one prediction of the good-enough processing account, i.e., lingering misinterpretation is resulted from incomplete reanalysis and therefore there should be more evidence of reanalysis for trials that were answered correctly. However, our results do not rule out the Good-Enough Processing Account, as we did not test the second possibility: reanalysis is completed, but interpretations from both initial analysis and reanalysis linger.

The likelihood judgments obtained in Experiment 3 and the results of including them in re-analyses of Experiments 1 and 2 suggest that accuracy to comprehension questions following garden-path sentences may not provide a good indicator of whether or not people fully revise an initial misinterpretation of a garden-path sentence. Instead, responses to such questions seem to be strongly influenced by the likelihood of the events described in the sentence and its question.

References

Barton, S.B., & Sanford, A.J. (1993). A case study of anomaly detection: Shallow semantic processing and cohesion establishment. *Memory & Cognition, 21*, 477-487.

Christianson, K. (2010). Effects of Plausibility on Structural Priming. *Journal of experimental psychology. Learning, memory, and cognition, 36*, 538-544.

Christianson, K., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368-407.

Christianson, K., Williams, C.C., Zacks, R.T., & Ferreira, F. (2006). Younger and older adults' "good-enough" interpretations of garden-path sentences. *Discourse Processes, 42*, 205-205-238. doi: 10.1207/s15326950dp4202_6

Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*, 107-126. doi: doi:10.1017/S0142716406060206

Ferreira, F., Bailey, K.G.D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*, 11-11-15.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research, 30*, 3-3-20.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*, 348-368.

Ferreira, F., & Henderson, J.M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language, 30*, 725-745.

Ferreira, F., & Henderson, J.M. (1998). Syntactic reanalysis, thematic processing, and sentence comprehension. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 73-100). Dordrecht: The Netherlands: Kluwer Academic.

Fodor, J.D., & Inoue, A. (1998). Attach anyway. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (Vol. 21, pp. 101-141). Kluwer: Dordrecht.

Frisson, S. (2009). Semantic underspecification in language processing. *Language and Linguistics Compass, 3*, 111-127.

Greenhouse, Samuel W, & Geisser, Seymour. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95-112.

Kaschak, M.P., & Glenberg, A.M. (2004). This construction needs learned. *Journal of Experimental Psychology: General, 133*, 450-467.

Lau, E.F., & Ferreira, F. (2005). Lingering effects of disfluent material on comprehension of garden path sentences. *Language and Cognitive Processes, 20*, 633-666.

Oldfield, Richard C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97-113.

Osterhout, L., Holcomb, P.J., & Swinney, D.A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 786-803.

Sanford, A.J., & Sturt, P. (2002). Depth of processing in language comprehension: not noticing the evidence. *Trends in Cognitive Sciences, 6*, 382-386.

Sturt, P., Pickering, M.J., & Crocker, M.W. (1999). Structural change and reanalysis difficulty in language comprehension. *Journal of Memory and Language, 40*, 136-150.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition, 36*, 201-216.

Trueswell, J.C., Tanenhaus, M.K., & Garnsey, S.M. (1994). Semantic Influences On Parsing: Use of Thematic Role Information in Syntactic Ambiguity Resolution. *Journal of Memory and Language, 33*, 285-318.