

Lingering misinterpretation of garden-path sentences and structural representation

Yujing Huang (Boston College) & Fernanda Ferreira (UC Davis)

yujing.huang@bc.edu

A key question in parsing research concerns how sentences that have been misinterpreted are reanalyzed, and to what extent the parser's attempts at reanalysis are successful. Past work demonstrating that misinterpretations associated with misparses linger (Christianson et al., 2001) suggests that the parser fails in at least some of the steps required for successful reinterpretation. What is not yet established is the level of representation responsible for misinterpretations: Is the syntax of the sentence incorrect, or do intermediate meanings associated with the initial misparse linger in memory? To attempt to answer this question, Slattery et al. (2013) reported that fixation durations on a downstream reflexive were longer when the gender of the reflexive mismatched that of its antecedent (See Table 1). They argued that this gender mismatch effect showed that comprehenders constructed the binding relationship between the reflexive and its antecedent and therefore had built the correct structural representation.

A problem with the logic of their argument, however, is that they did not probe the subjects' interpretations, and therefore it is possible that the gender mismatch effect they found was driven by trials on which the sentences were interpreted correctly, and that the mismatch effect would not have been found on trials on which the subjects failed to interpret the sentence correctly. To address this question concerning the level of representation responsible for misinterpretations and so-called "good enough" effects in sentence comprehension, we conducted two studies where we combined the reading/reaction time measures with an offline comprehension task which enabled us to test whether there is a reading/reaction time difference on correct versus incorrect trials.

Our first study used eye-tracking while reading paradigm and the second study was a self-paced reading study. Stimuli were similar to those in Slattery et al. (2013) and remained unaltered between our two studies. We manipulated two factors: ambiguity (comma/no comma) and gender match versus mismatch (see Table 1). Four lists were created with a counterbalanced design so that each participant only saw the same item once. Each list consisted of 80 test items and 160 fillers. All test items were followed by a comprehension question probing the possible misinterpretation (Table 1). Comprehension questions also followed half of the fillers. Participants ($n=120$; $n=144$ native English speakers) answered these questions after reading the sentence on a separate screen.

For the eye-tracking study, we analyzed the same measures as those in Slattery et al., i.e. first-pass, go-past and total reading times. For the self-paced reading study, we analyzed the reaction time on the critical word, and in two spill-over regions which are the two words after the critical word. Our linear mixed-effect models reveal a significant effect of ambiguity on the disambiguating verb in the reading/reaction time measures ($p's < 0.01$), which suggested that participants were garden-pathed when the sentence was temporarily ambiguous. We also replicated Christianson et al. (2001) and found that the response accuracy to the ambiguous sentences was 48% and 46% in our two studies. In addition, we found an overall gender mismatch effect for reflexives, replicating Slattery et al. (2013). When we added the response accuracy as a predictor to the model, there was a significant gender mismatch effect but no interaction between gender matching and response accuracy (Table 2). We also found a negative correlation between response accuracy and reading/reaction time measures in the disambiguation region ($p's < 0.05$), suggesting that more difficulty at the point of reanalysis leads to lower response accuracy.

Our results suggest that comprehenders can succeed at syntactic reanalysis to a degree that permits them to detect a downstream gender mismatch based on a binding relationship. Using comprehension questions, we confirmed that reanalysis does not always result in the correct interpretation. We conclude from this pattern of results that the misinterpretation of garden-path sentences does not arise from the illicit syntactic representation but instead is a result of the lingering in memory of meanings associated with the initial misparse.

References

- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive psychology*, 42(4), 368-407.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive psychology*, 14(2), 178-210.
- Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, 69(2), 104-120.

Table 1 Example of stimuli

Condition	Gender match	Gender mismatch
Garden-path	While Mary bathed the <i>baby boy</i> that was cute and cuddly fell off the bed and hurt <u>himself</u> very badly.	While Mary bathed the <i>baby girl</i> that was cute and cuddly fell off the bed and hurt <u>himself</u> very badly.
Non-garden-path	While Mary bathed, the <i>baby boy</i> that was cute and cuddly fell off the bed and hurt <u>himself</u> very badly.	While Mary bathed, the <i>baby girl</i> that was cute and cuddly fell off the bed and hurt <u>himself</u> very badly.
Comprehension question	Did Mary bathe the baby?	

Table 2 Summary of results from mixed-effect models in the reflexive region

	Predictor	β	p-value
Experiment 1	Matching	-11.27	<0.001 ***
	Accuracy	-4.47	0.35
	Interaction	0.17	0.97
Experiment 2	Matching	-13.35	<0.001 ***
	Accuracy	-8.11	0.07
	Interaction	-1.90	0.75

Note: The dependent variable in Experiment 1 is the first-pass reading time on the reflexive and the in Experiment 2 is the reaction time on the first spill-over word.

Figure 1 Reading time/reaction time on the reflexive/spill-over region (GP=garden-path)

