

## Word-order uncertainty induces alternative, non-veridical structures in online comprehension

Gabriel Doyle & Roger Levy (University of California, San Diego)

gdoyle@ling.ucsd.edu

Parsing; Uncertainty; Sentence comprehension; Hallucinated garden-paths; Self-paced reading; English

Human sentence processing takes place in noisy conditions. As a result, a rational approach to sentence processing should include and use uncertainty about word identities and word order during online comprehension [1]. However, the traditional approach to human sentence comprehension has an implicit modularity assumption that the words of the sentence are read without residual uncertainty, and without the possibility of later revision.

Previous work suggests that this modularity assumption is too strong. Humans do maintain uncertainty that is affected by later information [2,3], and revising this uncertainty may explain online comprehension effects, such as regressive eye-movements during reading [4,5] and garden-pathing onto syntactic analyses inconsistent with veridical input [6]. This previous work investigated structural uncertainty stemming from word identity uncertainty.

In this experiment, we look for evidence of word order uncertainty. We consider a set of sentences where a sentence-initial relative clause may be re-analyzed as a complement clause (CC) if the semantic roles of the embedded and extracted NPs were swapped. If readers maintain uncertainty about the order of the words they have previously read (motivated by the possibility of, e.g., speaker error or faulty memory), they may assign some probability to this non-veridical swap-CC structure.

To look for this structure, we designed a 2x2 moving-window self-paced reading study using sentences like (1). These start with an ORC-modified animate noun, with an inanimate NP as the RC subject. The relative pronoun varies between *that*, which allows swapping the NPs to get a CC partial-parse (2), or *who*, which blocks both the NP swap and CC partial-parse. We also vary whether the embedded NP is able and biased to take a CC or unable to take one. The relative clause verb is chosen to be plausible but unlikely with the inanimate embedded subject, encouraging the swap-CC partial-parse. This swap-CC alternative is removed when the main verb appears, terminating the CC without an object.

If readers maintain word order uncertainty and update their distribution over previous words based on later information, we predict an interaction effect at the relative clause verb (RCV) and main verb (MV), as the likelihood of the swap-CC parse increases at RCV and decreases at MV. The *that*-biasCC condition is predicted to exhibit a stronger slowdown at these verbs than conditions with only *that* or only a CC-biased NP. The prediction is borne out in self-paced reading, where the *that*-biasCC condition has superadditively high reading times at both RCV and MV ( $p < .001$ ). It is not obvious how this superadditive effect would be explained under existing retrieval/interference frameworks.

This gives evidence that readers maintain uncertainty about word order as part of a rational response to input over a noisy channel. Furthermore, this uncertainty influences online parsing to the point that comprehenders can pursue an analysis of the sentence that is inconsistent with the surface input.

### Stimuli

- |     |   |               |
|-----|---|---------------|
| (1) | a. The journalist that the fact surprised came to the press conference late.    | [that-biasCC] |
|     | b. The journalist who the fact surprised came to the press conference late.     | [who-biasCC]  |
|     | c. The journalist that the article surprised came to the press conference late. | [that-noCC]   |
|     | d. The journalist who the article surprised came to the press conference late.  | [who-noCC]    |
| (2) | The fact that the journalist surprised (someone) came... [swap-CC]              |               |

### References

- [1] Levy (2008) A noisy-channel model of rational human sentence comprehension under uncertain input. EMNLP  
 [2] Marslen-Wilson (1987) Functional parallelism in spoken word recognition. Cognition  
 [3] Connine et al (1991) Effects of subsequent sentence context in auditory word recognition. JML  
 [4] Levy et al (2009) Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. PNAS  
 [5] Bicknell (2011) Eye movements in reading as rational behavior. PhD dissertation.  
 [6] Levy (2011) Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results. ACL.