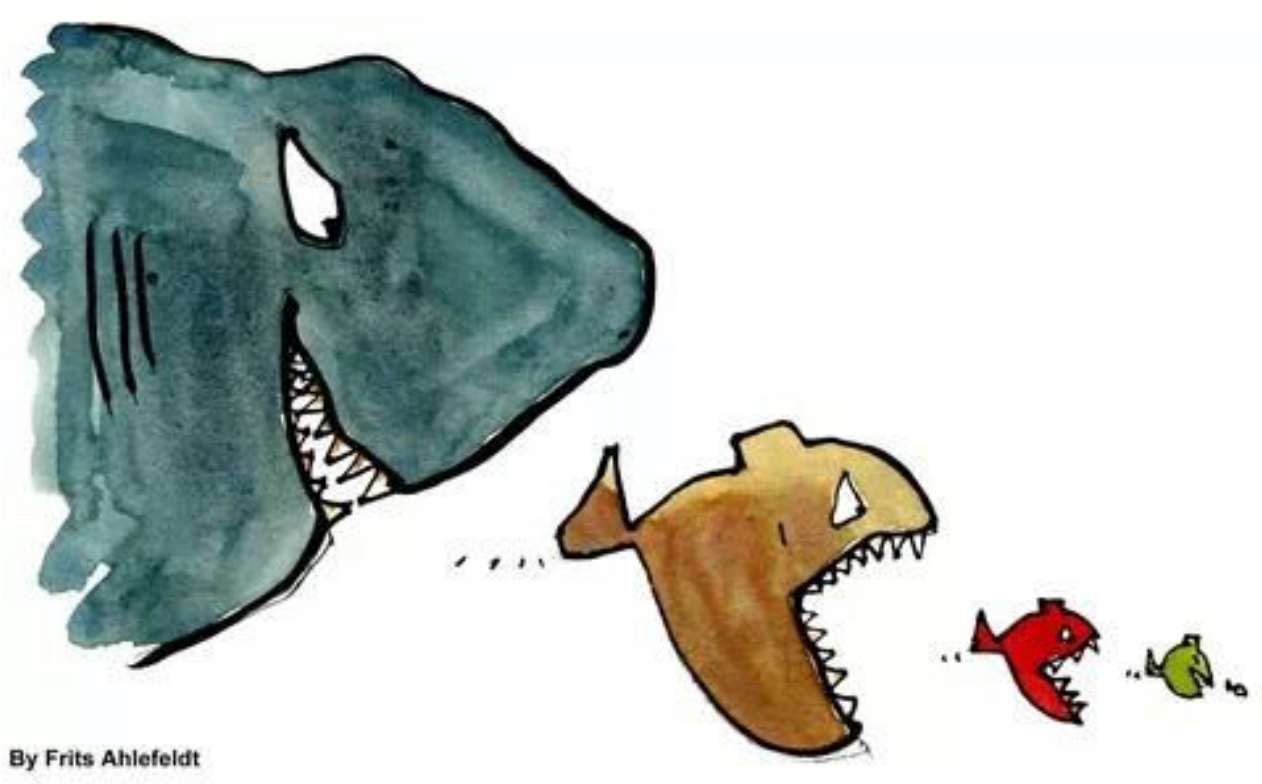
Surprisal and Hallucinations in Garden Path Sentences

Alankrita Bhatt
Sharbatanu Chatterjee

Ambiguous Sentences



Fish fish fish eat eat eat is a **valid sentence.**
So is *Fish fish fish fish fish eat eat eat eat eat!*

Garden-Path Sentences – What are they?

- Sentences that lead the human sentence processor (HSP) to construct an **initial syntactic structure**, which turns out to be **incorrect,** and thus requires syntactic (and semantic) **reanalysis.**

Examples –

- "Time flies like an arrow, fruit flies like a banana"
- "The old man the boat"

Previous Work – Theories of Sentence Processing

- **Incremental Sentence Processing Theory** : Hypotheses about syntactic structures and semantic roles are made as soon as each word is encountered.

- This theory states that **Input Recognition** and **Syntactic Analysis** are distinct.

Newer Theories: Statistical Framework

- Computing the probability of a sentence using PCFG rules

$$P(\text{Sentence}) = \prod (\text{Probability of all rules used to generate it})$$

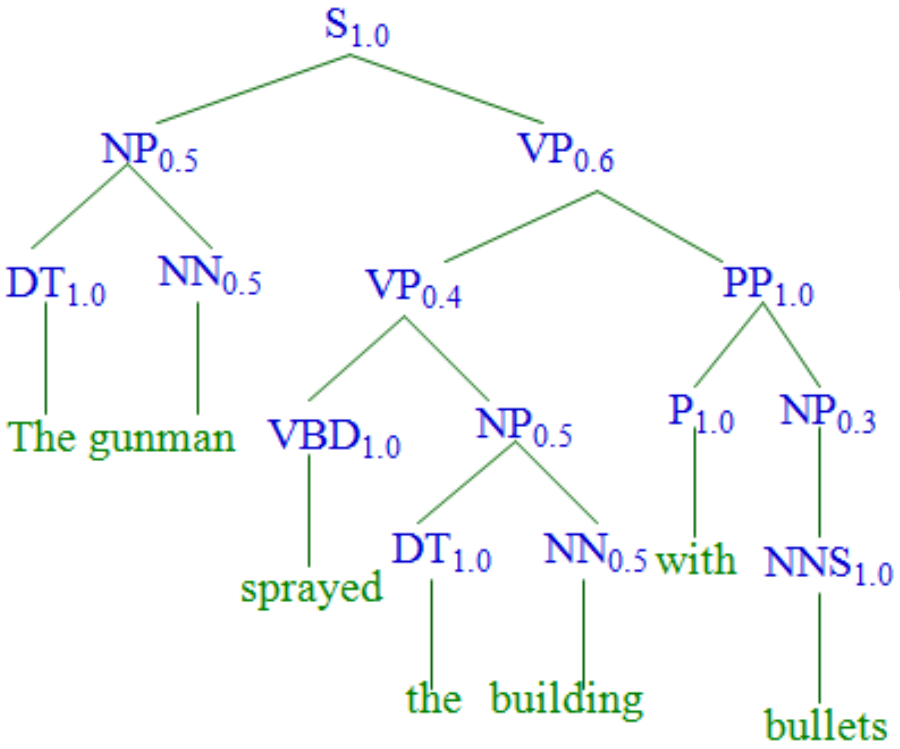
TOP	→ S .	1.000000
S	→ INVERTED NP	0.003257
S	→ SBAR S	0.012289
S	→ SBAR , S	0.041753
S	→ NP VP	0.942701
INVERTED	→ PP VBD	1.000000
SBAR	→ INSBAR S	1.000000
VP	→ VBD RB	0.002149
VP	→ VBD PP	0.202024
VP	→ VBD NP	0.393660
VP	→ VBD PP PP	0.028029
VP	→ VBD RP	0.005731
VP	→ VBD	0.222441
VP	→ VBD JJ	0.145966
PP	→ IN NP	1.000000

Example of PCFG Probabilities taken from the parsed Brown corpus.

Example

- "The gunman sprayed the building with bullets"

One way of parsing

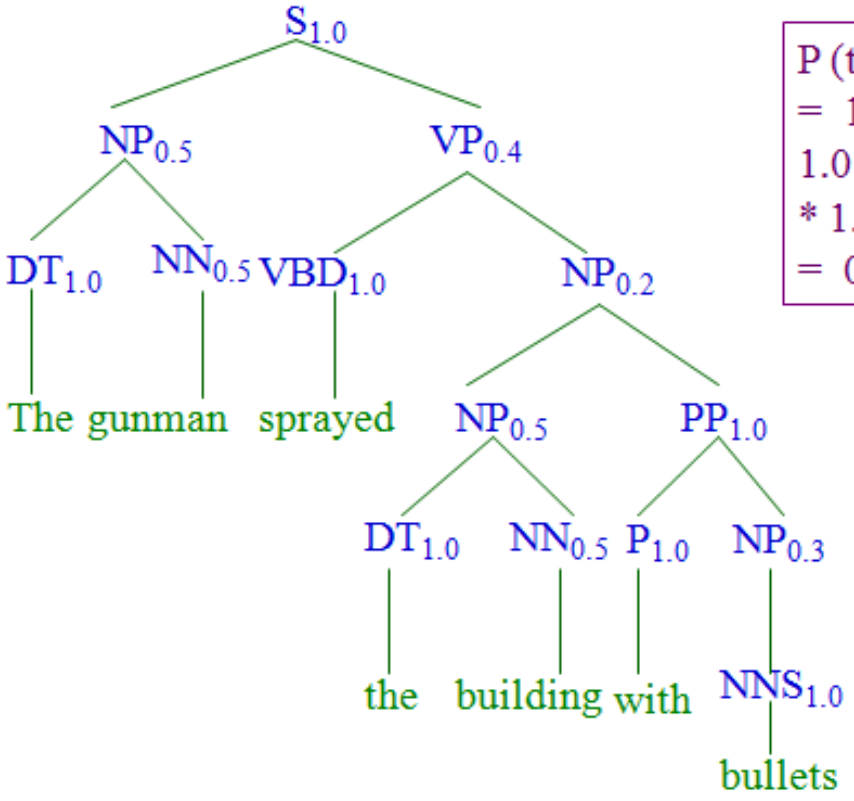


$P(t_1)$
$= 1.0 * 0.5 * 1.0 * 0.5 * 0.6 *$
$0.4 * 1.0 * 0.5 * 1.0 * 0.5 * 1.0$
$* 1.0 * 0.3 * 1.0$
$= 0.00225$

Example

"The gunman sprayed the building with bullets"

Another way of parsing



$P(t_2)$
$= 1.0 * 0.5 * 1.0 * 0.5 * 0.4 *$
$1.0 * 0.2 * 0.5 * 1.0 * 0.5 * 1.0$
$* 1.0 * 0.3 * 1.0$
$= 0.0015$

Hence total probability of the sentence
$= 0.0015 + 0.00225 = 0.00375$

Surprisal

- The Cognitive "effort" required to process a particular word in a sentence can be quantified in terms of an information theoretic measure defined as the "Surprisal"

$$\text{Effort}(w_i) \propto \log \frac{1}{P(w_i|w_{1\ldots i-1}, Ctxt)}$$

Example: The horse raced past the barn fell.

$$P(\textit{fell}|w_{1\ldots 6}) = \sum_j P(\textit{fell}|T_j, w_{1\ldots 6}) P(T_j|w_{1\ldots 6})$$

Here T_j represents pre-disambiguation analyses.

Our Hypotheses

1. Linguistic information is used both proactively and retroactively (This is optimal!) –

   "Hallucinations" may occur in some cases and people might process a distorted input that has a high prior probability.

2. The surprisal is linearly related to the reading time.

3. The "Good Enough" theory of sentence processing is the most probable.

Hallucinations

- "As the soldiers marched, toward the tank lurched an injured enemy combatant"

- When asked *Did the soldiers march toward the tank?* , many subjects gave a positive reply

- Many also believed that there wasn't a comma in the original sentence.

Methodology

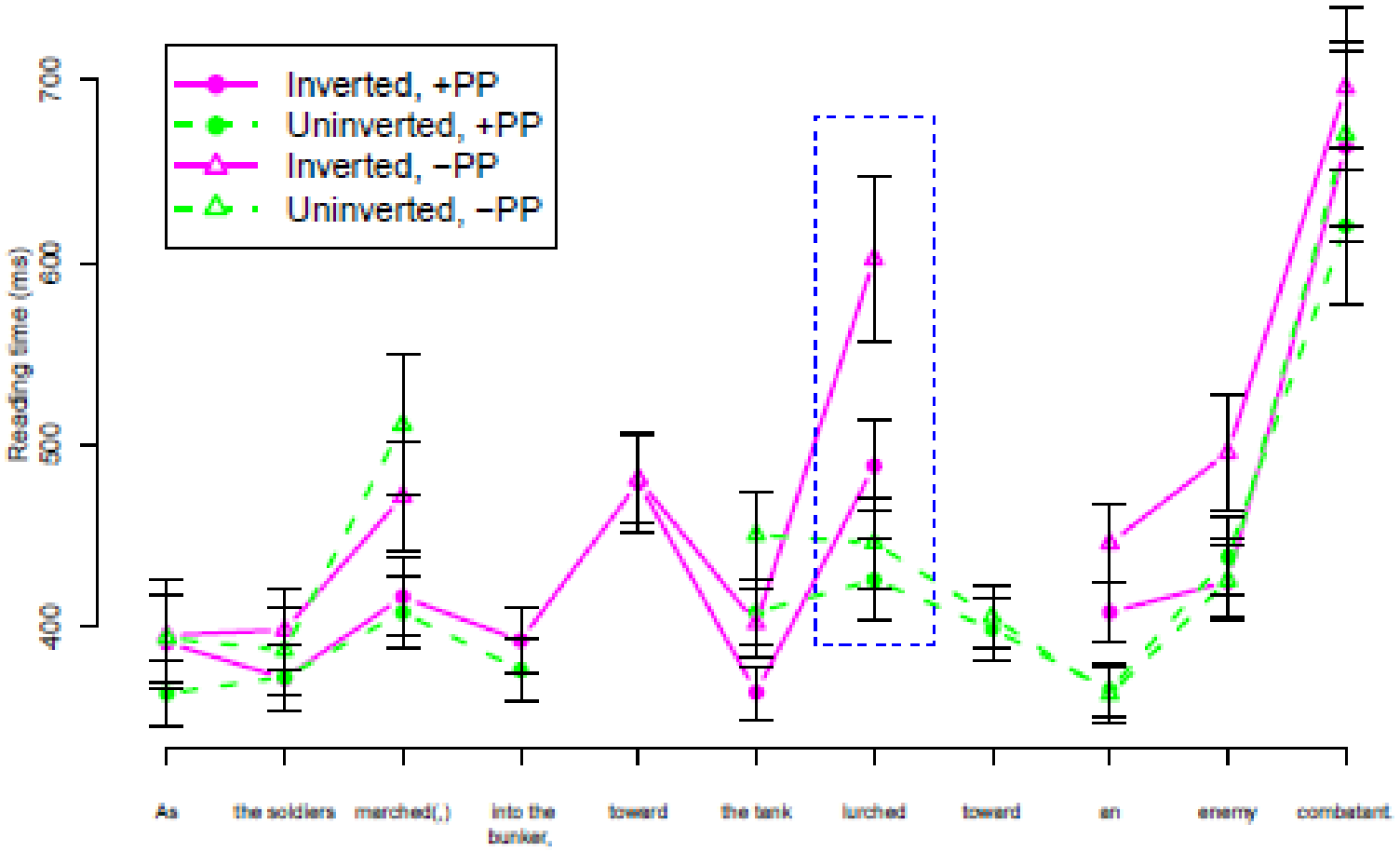
- Self paced reading study.
- Participants press a button on a keyboard to reveal the successive word in a sentence.
- The times between subsequent button presses are taken as a measure of incremental processing difficulty.
- Each sentence is followed by a yes/no question.
- Experimental items interspersed with “filler” sentences.

Gaze Tracking

- By observing where the gaze fixates, we hope to receive further affirmation of our hypothesis.

- “Lingering misinterpretations” are also expected.

What we expect



Ambiguous Sentences in other Languages

- এসেছি যা ফেলে তার চিহ্ন তবু ছাড়ে না অভ্যাস
  সমস্ত চিতকারশব্দে শুনতে পাই স্তব্ধতার শ্বাস।

- কাদম্বিনী মরিয়া প্রমাণ করিল যে সে মরে নাই।

Possible future work

- Do Language 2 speakers parse Language 2 the same way as Language 1 speakers do?

- Misinterpretations can give us an idea about the mental representations of structures which can help in pedagogical techniques.

References

- Levy, Roger. 2011. “Integrating surprisal and uncertain-input models in online sentence comprehension: Formal techniques and empirical results”. In Proceedings of the 49th annual meeting of the Association for Computational Linguistics, Stroudsburg, PA: Association for Computational Linguistics.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics, pages 159-166.
- Levy, R. (2008). Expectation based syntactic comprehension. Cognition, 106:1126-1177.
- Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In Proceedings of the 13th Conference on Empirical Methods in Natural Language Processing, pages 234-243.
- Slattery, Timothy J. et al (2013) Lingering misinterpretations of garden path sentences arise from competing syntactic representations, Journal of Memory and Language, pages 104-120.

Thank you!
Questions?