Surprisal and Hallucinations in Garden Path Sentences

Alankrita Bhatt
Sharbatanu Chatterjee
Fish fish fish is a valid sentence.
So is Fish 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”
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.

Input Recognition

Syntactic Analysis

Prediction of upcoming input
Newer Theories: Statistical Framework

- Computing the probability of a sentence using PCFG rules

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

Example of PCFG Probabilities taken from the parsed Brown corpus.
“The gunman sprayed the building with bullets”

One way of parsing

\[
P(t_1) = 1.0 \times 0.5 \times 1.0 \times 0.5 \times 0.6 \times
0.4 \times 1.0 \times 0.5 \times 1.0 \times 0.5 \times 1.0
\times 1.0 \times 0.3 \times 1.0
= 0.00225
\]
Example

“The gunman sprayed the building with bullets”

Another way of parsing

Hence total probability of the sentence

\[ \text{P} (t_2) = 1.0 \times 0.5 \times 1.0 \times 0.5 \times 0.4 \times 1.0 \times 0.2 \times 0.5 \times 1.0 \times 0.5 \times 1.0 \times 1.0 \times 0.3 \times 1.0 = 0.0015 \]

\[ \text{Hence total probability of the sentence } = 0.0015 + 0.00225 = 0.00375 \]
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”

\[
\text{Effort}(w_i) \propto \log \frac{1}{P(w_i | w_1...i-1, Ctxt)}
\]

Example: The horse raced past the barn fell.

\[
P(\text{fell} | w_1...6) = \sum_j P(\text{fell} | T_j, w_1...6) P(T_j | w_1...6)
\]

Here \(T_j\) represents pre-disambiguation analyses.
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.
“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.
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.
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 চিতকারণ শুনতে পাই শ্বশুরতার স্নাস।

- 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


Thank you!
Questions?