The Role of Animacy Information in Expectation-based Sentence Comprehension
Zhong Chen (Rochester Institute of Technology, USA)
z.chen@rit.edu

1. Introduction

- Sentence processing is predictive. Each incoming word introduces information that helps us to shape expectations about the rest of the sentence (Hale, 2001; Levy, 2008).
- **Structural Information**, e.g., grammatical category, phrase structure hierarchy, or syntactic movement, often revises existing parses. In contrast, new words also render **Non-structural Information** like thematic relation or information structure.
- But how do structural and non-structural expectations interact during comprehension? This work investigates the role of noun phrase **Animacy** in parsing by modeling two eye-tracking experiments using **Entropy Reduction**.

2. Entropy Reduction

- Put **Entropy**, the information-theoretic notion, in a language-processing scenario. The random variable $X$ may take values that are derivations on a probabilistic grammar $G$.

$$H(X) = - \sum_{x \in X} p(x) \log_2 p(x)$$

- Extend the entropy notation to express conditioning events. If $w_1 \ldots w_t$ is an initial substring of a sentence generated by $G$, the conditional entropy $H_s$ will be the uncertainty about derivations that have a $w_1 \ldots w_t$ prefix.

**Entropy Reduction** (ER) is a complexity metric of sentence comprehension that quantifies the amount of information a word contributes towards reducing uncertainty (Hale, 2003, 2006, Frank et al., 2015).

$$ER_t = \begin{cases} H_{t-1} - H_t & \text{if this difference > 0} \\ 0 & \text{otherwise} \end{cases}$$

- The uncertainty level depends on weighted, predictive syntactic analyses that are “still in play” at a given point. This work takes one step further by considering readers’ expectation on the non-structural factor animacy.

3. Animacy Information

- This study adopts a binary classification of noun phrase animacy, namely +ANIM vs -ANIM, similar to previous works (MacWhinney et al., 1984, Traxler et al., 2002).
- The frequency distributions of NP animacy are obtained from the annotated hand-parsed Switchboard corpus of conversational American English (Zaenen et al., 2004).

4. Modeling Incremental Comprehension (Cornell CCPC)

5. Example 1: Traxler et al. (2002, 2005)

- Animacy facilitates the assignment of thematic roles and affects how ambiguities are resolved in relative clauses. The easiness of SRs (SUBJ ADV) is less prominent when inanimate heads are involved, especially among readers with high working memory capacity (WMC).

- ER predictions mirror the reversed SUBJ ADV observed within the high WMC group, confirming a recent finding such that ER is a stronger predictor of processing difficulty when a larger amount of syntactic lookahead is considered (Linzen & Jaeger, in press).

- SIRs with inanimate heads are in fact harder than their OR counterparts because (1) they are less frequent and (2) the parser is highly uncertain about choosing SR or OR as the continuation given an inanimate head. More disambiguation effort will be made later within the RC region.


- Animacy also interacts with the depth of structural relations. The difficulty of integrating a subject with an action verb is significantly reduced when the subject is inanimate and the verb occurs across the clausal boundary in an embedded clause, e.g. SR.

- ER predictions are similar such that the magnitude of clausal boundary effect is amplified with an inanimate subject. This is because the inanimate subject allows a variety of continuations while more disambiguation effort is done at the main verb than at the embedded verb.

7. Conclusions

- This work models the animacy effect observed in comprehension experiments and provides linguistically plausible interpretations.
- It visualizes alternative derivations that are “still in-play” by using probabilistic grammars and therefore allows us to describe the interaction between structural and non-structural expectations.

Selected References


