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 $w_1 \ldots w_i$ is an initial substring of a sentence generated by G, the conditional entropy H_i will be the uncertainty about derivations that have a $w_1 \ldots w_i$ 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).

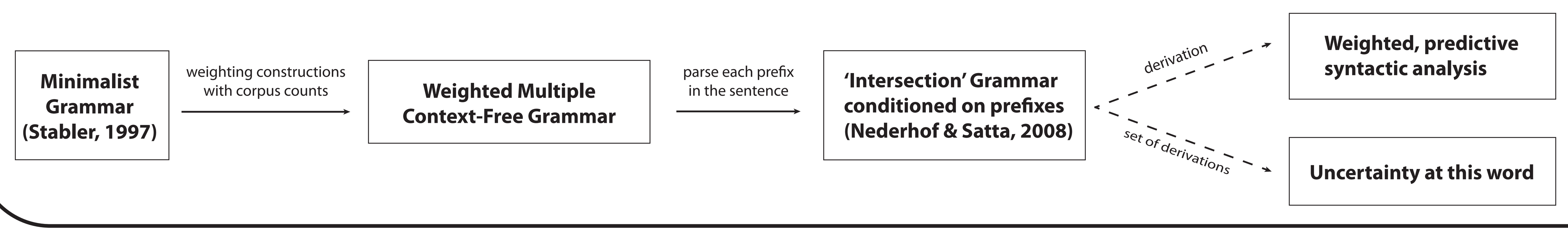
$$ER_i = \begin{cases} H_{i-1} - H_i & \text{if this difference} > 0 \\ 0 & \text{otherwise} \end{cases}$$

- 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).

Type	Examples	Head	Embedded	**SUBJ ADV** within the three-word region after the head		
				Traxler et al (2002)	Traxler et al (2005) High WMC	This study (ER)
SR	The director that watched the movie…	+anim	−anim	294 ms	239 ms	1.01 bit
OR	The director that the movie pleased…	+anim	−anim			
SR	The movie that pleased the director…	−anim	+anim	39 ms	**-49 ms**	**-0.73 bit**
OR	The movie that the director watched…	−anim	+anim			

- 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).
- SRs 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.

Prob	Remainder
0.106	***Det AniN who*** *Vt Det InaniN Vi*
0.078	***Det AniN who*** *Vt Pronoun Vi*
0.056	***Det AniN who*** *Vt Det AniN Vi*
0.055	***Det AniN who*** *Vi Vi*
0.026	***Det AniN who*** *Vt Det InaniN Vt Det InaniN*
0.020	***Det AniN who*** *Vt Pronoun Vt Det InaniN*
0.019	***Det AniN who*** *Det AniN Vt Vi*
0.018	***Det AniN who*** *Pronoun Vt Vi*
0.017	***Det AniN who*** *Vt Det InaniN Vdi Det AniN to Det InaniN*
0.015	***Det AniN who*** *Vt Det InaniN Vdi Det InaniN to Det InaniN*
…	…
entropy = 7.929	

(a) Animate prefix

Prob	Remainder
0.047	***Det InaniN who*** *Vt Det InaniN Vi*
0.047	***Det InaniN who*** *Det AniN Vt Vi*
0.046	***Det InaniN who*** *Pronoun Vt Vi*
0.035	***Det InaniN who*** *Vt Pronoun Vi*
0.034	***Det InaniN who*** *Vi Vi*
0.025	***Det InaniN who*** *Vt Det AniN Vi*
0.016	***Det InaniN who*** *Vt Det InaniN Vt Det InaniN*
0.016	***Det InaniN who*** *Det AniN Vt Vt Det InaniN*
0.016	***Det InaniN who*** *Pronoun Vt Vt Det InaniN*
0.012	***Det InaniN who*** *Vt Pronoun Vt Det InaniN*
…	…
entropy = 9.092	

(b) Inanimate prefix

6. Example 2: Lowder & Gordon (2012)

- 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 occur across the clausal boundary in an embedded clause, e.g. SR.

Subject NP	Verb Location	Examples	The clausal-boundary effect at the verb			
			Lowder & Gordon (2012)		This Study (ER)	
+anim	matrix	The cowboy concealed the pistol…	379 ms	36 ms	1.12 bit	0.08 bit
	embedded	The cowboy **that** concealed the pistol…	343 ms		1.04 bit	
−anim	matrix	The pistol injured the cowboy…	415 ms	93 ms	2.32 bit	0.45 bit
	embedded	The pistol **that** injured the cowboy…	322 ms		1.86 bit	

- 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 is done at the main verb than at the embedded verb.

Prob	Remainder
0.304	***Det InaniN*** *Vi*
0.105	***Det InaniN*** *Vt Det Inaninoun*
0.067	***Det InaniN*** *Vdi Det Aninoun to Det Inaninoun*
0.061	***Det InaniN*** *Vdi Det Inaninoun to Det Inaninoun*
0.055	***Det InaniN*** *Vt Det Aninoun*
0.054	***Det InaniN*** *Vi in Det Inaninoun*
0.037	***Det InaniN*** *Vt Pronoun*
0.019	***Det InaniN*** *Vi in Det Aninoun*
0.012	***Det InaniN*** *Vdi Det Aninoun to Det Aninoun*
0.011	***Det InaniN*** *Vdi Det Inaninoun to Det Aninoun*
…	…
entropy = 5.374	

ER=2.32 →

Prob	Remainder
0.424	***Det Inaninoun Vt*** *Det Inaninoun*
0.223	***Det Inaninoun Vt*** *Det Aninoun*
0.148	***Det Inaninoun Vt*** *Pronoun*
0.043	***Det Inaninoun Vt*** *Det Inaninoun in Det Inaninoun*
0.023	***Det Inaninoun Vt*** *Det Aninoun in Det Inaninoun*
0.015	***Det Inaninoun Vt*** *Det Inaninoun in Det Aninoun*
0.015	***Det Inaninoun Vt*** *Pronoun in Det Inaninoun*
0.008	***Det Inaninoun Vt*** *Det Aninoun in Det Aninoun*
0.006	***Det Inaninoun Vt*** *Det Inaninoun who Vt Det Inaninoun*
0.006	***Det Inaninoun Vt*** *Det Inaninoun who Det Aninoun Vt*
…	…
entropy = 3.058	

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

Frank, S., Otten, L., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and language*, 140, 1-11.
Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, 30(4).
Linzen, T. & Jaeger T.F. (In press) Uncertainty and Expectation in Sentence Processing: Evidence From Subcategorization Distributions. *Cognitive Science*.
Lowder, M., & Gordon, P. (2012). The pistol that injured the cowboy: Difficulty with inanimate subject-verb integration is reduced by structural separation. *JML*, 66(4), 819–832.
Traxler, M., Morris, R., & Seely, R. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, 47, 69–90.
Traxler, M., Williams, R., Blozis, S., & Morris, R. (2005). Working memory, animacy, and verb class in the processing of relative clauses. *Journal of Memory and Language*, 53, 204–224.