Lexical Variety in Sentence Disambiguation

George Kachergis and Jamie Olson

Background
Ambiguity resolution in natural languages is arguably the most difficult problem concerning natural language comprehension. Despite much research, little progress has been made towards a comprehensive, robust model of human ambiguity resolution. Linguistic ambiguity occurs in situations where a single sentence can have multiple different interpretations. Ambiguity in language arises mainly out of phonological, lexical and syntactic ambiguity. We investigate possible ways of resolving syntactic ambiguities in English sentences.

Design
We used custom software to present participants with a series of 62 randomly ordered sentences. They were asked to read each sentence and press a key to validate it as a syntactic English sentence, or press a different key to reject it. 20 of the trials were non-sentences to check accuracy, 20 were unambiguous sentences similar in form and content to the ambiguous ones, and 22 were ambiguous sentences of the following forms:

- **prepositional** - "The girl saw the man with the telescope."
- **possessive** - "They fed her rat poison."
- **gerund** - "Visiting relatives can be boring."
- **adjective** - "The old men and women stayed home."

After validating ambiguous sentences, participants were asked to indicate how they resolved the ambiguity (ex: Who had the telescope? Man or Girl).

Results
All sentences within each category had identical possible structural interpretations. Most participants had different structural interpretation within sub-categories at least once. That is, they interpreted structurally identical sentences differently depending on the lexical variety and semantic context. This is shown in the table (left), where the percentage of participants that disambiguated syntactically identical sentences differently is displayed. The number of sentences in each category is displayed because it is reasonable to expect that the more sentences there are in a category, the more lexical variety there is and it is more likely that participants will respond differently.

Syntactic ambiguity involves two different competing structural interpretations of a sentence. Sentence structure is often represented using phrase structure diagrams. For example, the sentence "Mary kissed the children in the kitchen," can be interpreted as either Mary kissing the children who happened to be in the kitchen or as Mary kissing the children while she (Mary) was in the kitchen:

```
S
  | NP  VP
  |   |
  | Mary  kissed  NP
  |   PP
  | the  children
  | in  det  noun
  | the  kitchen
```

```
S
  | NP  VP
  |   |
  | Mary  kissed  NP
  |   PP
  | the  children
  | in  det  noun
  | the  kitchen
```

The fact that participants disambiguated sentences within the same ambiguous sub-category differently from sentence to sentence suggests that no participant is following a structural rule that dictates how they resolve ambiguities. Instead, it indicates that lexical variety and context are important when humans are resolving syntactic ambiguities.

There remain many questions to be answered about how humans resolve structural ambiguities in language. While it has been shown that people use lexical and contextual information, little research has been done to determine precisely what information is used and how it affects disambiguation.