**Prospectus Defense**

**Big Question:** What are the discourse-functional correlates of lexical flexibility?

**Lexical flexibility** – go over examples

* (traditional, imprecise) – The ability for a lexeme to function in more than one part of speech
* (constructional, operationalized) – The ability for a lexeme to occur in multiple pragmatic functions with zero coding

**Specific Question:** Does information status correlate to choice of lexical category for flexible items?

**Motivation:** To understand how discourse-functional pressures give rise to parts of speech (in the sense of typological markedness patterns)

* Builds on Hopper & Thompson (1984), Thompson (1989), Croft (2001), Nakayama (2003)

**Implications:**

* Languages described as having rigid word classes show strongly grammaticized pragmatic functions
* Languages described as having flexible word classes show weakly grammaticized pragmatic functions
* Languages may become more or less flexible with time

**Structure of the Dissertation**

* Main question: How discourse factors affect categorical choice
* Other questions are meant to control for / rule out other possible factors in choice of lexical category
* Methodological questions:
  + corpus size
  + token frequency
  + corpus dispersion
* Semantic factors – So these can be accounted for
* Discourse factors
  + referent tracking (via grammatical role rather than anaphora)
    - Could potentially look at anaphora too though
  + information status (given, new, activated)
* Conclusion
  + Overview of the factors and their effects on categorical choice
  + Using the same lexemes for each chapter may allow me to used a mixed-effects model

**Data & Methods**

* Make the raw quantitative summaries available in both an Appendix and online
* Method for quantifying lexical flexibility (review sample data)
  + Will plan to also make various qualitative observations as I code the data
* Conclusion: Mixed-effects model
* Database: The coding process will result in a set of corpora tagged for flexible cases, allowing for future studies

**Issues**

* Synchronic vs. diachronic sense of “flexible”
  + Is something flexible because it *can* change functions, or because it *has* changed functions?
  + Empirically we can only know the one (using corpora at least), but the wording is important
* What counts as a lexeme?
  + Gries. 2006. Corpus-based methods and cognitive semantics: The many meanings of *to run*
* flexibility = conversion
  + Conversion is a proxy – a useful correlate of flexibility
  + I think flexibility should be thought of as something cognitive, and broader than just conversion
* semantic shifts
  + I can probably also code for semantic shift as I code for other variables
  + Which shifts/conversions actually happened and are attested, and which did not?
  + How do these compare to non-zero-marked shifts?
  + What determines whether a derivation will be marked or unmarked?
  + Coding for semantic shift will likely provide strong empirical evidence in support of Croft’s universal (direction of semantic shift towards the prototype). This would suggest that lexical items are always prespecified for category, although perhaps to differing degrees.
* terminology
  + *prototypical* vs. *prototypal*
* size of corpora
  + Chitimacha: 32,000
  + SBC: 249,000