**Question:**

What determines / predicts best how people will produce something that’s given information?

**Hypotheses:**

1. Whether it matters if it’s interpolated or separate data for speakers (across all)
2. Recency of mention (explicit, non explicit, both)
3. Frequency of mention (since last explicit mention, last non explicit, and both)
4. How many other things have been mentioned (explicit, non-explicit, and both)
5. Part of speech:
   1. Mention part of speech: pronoun, proper noun
   2. Previous mention part of speech: pronoun, proper noun (previous explicit, non-explicit, both)
6. Syntax
   1. Current mention syntactic function: direct object, indirect object, subject
   2. Prior mention syntactic function: ditto (previous explicit, non-explicit, both)
   3. Syntactic complexity (will discuss how to measure more next week)

Speaker information: test all both with coreference in Stanford CoreNLP turned on across speakers and turned off across speakers

Referential NP

**Tables Versions to Generate:**

* Speaker coreference off
* Speaker coreference on

**Columns to Add:**

* Coreference chain column **DONE**
* How recently explicitly mentioned (hypothesis 1)
* How recently non-explicitly mentioned (1)
* How recently explicitly or non-explicitly mentioned (1)
* How many times explicitly mentioned (2)
* How many times non-explicitly mentioned (2)
* How many times explicitly or non-explicitly mentioned (2)
* How many things have been explicitly mentioned in current session (3) **DONE**
* Create intonational phrases - (give unique ID number)
* Part of speech of previous explicit reference (4b)
* Part of speech of previous non-explicit reference (4b)
* Part of speech of previous reference, whether explicit or non-explicit (4b)
* Syntactic function of current word (5a)
* Syntactic function of previous explicit mention (5b)
* Syntactic function of previous non-explicit mention (5b)
* Syntactic function of previous explicit or non-explicit mention (5b)
* Rose: Syntactic complexity measure of current mention (5c)
  + Constituent height/depth
  + Minimum spanning tree if two coreferences within a sentence