

Psych156A/ Ling150

Spring 2012

Review Questions: Grammatical Categorization

(1) Terms/concepts to know: grammatical category, frequent frame

(2) One way to think about a word's grammatical category (like noun or verb) is that it is really just a description of the way that word can be used in the language. Given the following contexts, say whether you think the novel word in each example (indicated by CAPITALS) is a noun (like "goblin"), a verb (like "sing"), an adjective (like "hot"), or an adverb (like "dreamily"). Be sure to briefly explain why you think so. (You might find it helpful to substitute words you know in place of the novel words, and see which ones fit best.)

- (a) That's a very BOFT bog.
- (b) Is Sir Didymus REKKing the Bog of Eternal Stench?
- (c) Ludo shouldn't have STROOPed his paw in the bog.
- (d) They had to beware the BREER's tricks.
- (e) Hoggle couldn't believe how FREEMILY he was able to do it.

(3) What is the Semantic Bootstrapping Hypothesis? What is one problem with it?

(4) Here an example utterance from the imaginary Guin language.

"felgo bofty mu az berg mu merk berg felgo zu mu var berg az porto mu freggo berg felgo seech mu set berg draz pino trem felgo trem mu peri berg lootem bleeter."

- (a) What is the most frequent frame in this utterance?
- (b) Which words would this frame cluster together?
- (c) What is the second most frequent frame in this utterance?
- (d) What words would this second most frequent frame cluster together?

(5) By using the frequent frames learning strategy on a large set of Guin data, suppose you were able to identify 120 words that seem to behave like nouns. When Guin speakers are asked about them, you find out that 65 of the 120 words you identified did belong to the same category (nouns). However, from the data set you had, you should have found 35 more nouns than you actually did correctly identify. Calculate the precision and recall for the frequent frames approach on this set of Guin data. Be sure to show how you calculated each one.

(6) What does it mean if a grammatical categorization model has a high precision score, but a low recall score? That is, is this model likely to correctly identify a word as belonging to a particular grammatical category? And is this model likely to correctly group all the words of one category together?

(7) What experimental evidence do we have that very young children are sensitive to frequent frames in the input?

(8) Do frequent frames always operate best over words? (Hint: Think about languages like Turkish and German.) How well do frequent frames (using any unit) operate cross-linguistically?

(9) What considerations about children's memory did the model used by Wang & Mintz (2008) make? (Think about what they were trying to improve on from Mintz (2003)'s study.) How did the model by Wang & Mintz (2008) perform, compared to the study done by Mintz (2003)?