

Psych156A/ Ling150
Spring 2016
Review Questions: Structure

- (1) Terms/concepts to know: Navajo Code Talker Paradox, Universal Grammar, parameter, parameter values, Greenberg's generalizations, overhypothesis, variational learning, unambiguous data, sampling
- (2) What is an example of a movement rule in syntax?
- (3) Suppose a child encounters a sentence with the word order "Subject Verb Object". (For example: "Sarah likes Hoggle.") Why can't the child be certain about how this word order was produced?
- (4) How does language learning work under Chomsky's Universal Grammar? Do children require input from the native language to learn it? Why or why not?
- (5) If there are 3 language structure parameters with 2 values each, how many different languages could be represented? What about if there are 2 language structure parameters with 3 values each? What if there are 7 language structure parameters with 2 values each? What if there are 2 language structure parameters with 10 values each?
- (6) What kinds of generalizations can be made about the structure of languages which have "Subject Verb Object" as the basic word order, according to Greenberg? What about languages which have "Subject Object Verb" as the basic word order?
- (7) According to Greenberg, do languages with the same structural patterns need to have a shared history?
- (8) What is the value of parameters for learning language structure? That is, why are they useful? How can they make acquisition easier?
- (9) Suppose we have a parameter Q. We are not sure exactly which structures in the observable language it connects with, but we think it might connect to structures A, B, C, and D. Suppose that parameter Q can take one of two values: x1 or x2.
 - (a) Suppose we observe that structures A, B, and C often exhibit behavior characteristic of value x1, but structure D exhibits behavior characteristic of a different value z1. Which structures are likely to be connected together under parameter Q?
 - (b) Suppose we then find that parameter Q should have value x1. Which structures are likely to also have values x1 (A, B, C, and/or D)?
 - (c) Suppose then that a child rarely sees structure C, but often sees structures A, B, and D. A and B show behavior characteristic of value x1, while D shows

behavior characteristic of value z1. Given your answer to (b), how might a child learn the value for structure C? What value would this be?

(10) How is the idea of an overhypothesis similar to the idea of a linguistic parameter? (Hint: Think about what both allow you to do for data you haven't seen examples of before.)

(11) In the Dewar & Xu (2010) experiment, how was the control condition for the "unexpected outcome" implemented? Why is this a reasonable control? (Hint: Think about the different process that happens in order to generate each, even though the same objects appear at the end.) Why do infants' responses in this study suggest that they are able to form and use overhypotheses to make predictions?

(12) Though parameters can make language acquisition easier, why might acquisition still be difficult even if all children have to do is learn the appropriate parameter value for their linguistic parameters? (Hint: Are observable data ever ambiguous with respect to what parameter value they indicate?)

(13) Which grammar (native or non-native) is supposed to be the most successful in variational learning? Why?

(14) If there are 8 grammars available, what probability will a variational learner initially assign each one?

(15) Which grammar(s) will be able to successfully analyze unambiguous data for a language? Why?

(16) Is it easier to find unambiguous data for entire grammars or for individual parameters within a grammar?

(17) Suppose a grammar is comprised of two structural language parameters: P1 and P2. Suppose each of these parameters has two possible values, and these are the current probabilities the learner has associated with each of those values:

(P1) Value A: 0.2 Value B: 0.8
(P2) Value C: 0.4 Value D: 0.6

(i) What is the probability of a variational learner selecting the following grammars to test out on a data point?

- (a) P1 = A, P2 = C
- (b) P1 = B, P2 = D
- (c) P1 = A, P2 = D
- (d) P1 = B, P2 = C

(ii) Which grammar is most likely to be selected? Which grammar is least likely to be selected? (Hint: Your answers to part (i) will be helpful.)

(18) What does the quantity of unambiguous data that a child encounters have to do with when a child learns a particular structural property (usually signaled by a parameter value) of the language?

(19) Suppose we discover that Guin has the following structural properties: +wh-fronting, -verb-raising, -verb-second, -subject-drop, -topic-drop, and +intermediate-wh. Suppose we also determine how much unambiguous data a Guin child would likely encounter for each of these structural properties:

+wh-fronting: 5% of input
-verb-raising: 7.5% of input
-verb-second: 2.5% of input
-subject-drop: 20% of input
-topic-drop: 10% of the input
+intermediate-wh: 0.5% of input

(a) Which structural property would Yang (2004) predict a variational learner would learn first? Why?

(b) Which structural property would Yang (2004) predict a variational learner would learn last? Why?

(c) Give the order in which Yang (2004)'s variational learner would learn these structural properties, starting from the one acquired earliest and ending with the one acquired latest.

Extra Material (not required for this class):

(E1) What did artificial intelligence researchers discover about the relative difficulty of playing chess compared to the relative difficulty of translating language? How did they do this? Are there aspects of language use that still seem to be difficult for machines?