Ling151/ Psych156A
Winter 2018
Review Questions: Structure

(1) Terms/concepts to know: Navajo Code Talker Paradox, linguistic parameter, Universal Grammar, expletive subject, head directionality, overhypothesis, verb second, variational learning, sampling, unambiguous data, metrical stress, violable constraints, constraint ranking, multiple stress contours, acquisitional intake, Tolerance Principle, data filter

(2) 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?

(3) 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?

(4) 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?

(5) According to Greenberg, do languages with the same structural patterns need to have a shared history?

(6) What is an example of a linguistic parameter that was proposed to account for a variety of observed structural patterns? (Hint: Think about English vs. Italian, or English vs. Japanese)

(7) When thinking about the structure of human languages, are principles or parameters the invariant component? Which component is the part that varies in a constrained way?

(8) Why would it be easier to learn about the structure of your language if linguistic parameters were involved? (Hint: Think about things that are hard-to-learn vs. easy-to-learn. What if a parameter connected to all these things? How would that help?)

(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: q1 or q2.

> (a) Suppose we observe that structures A, B, and C often exhibit behavior characteristic of value q1, 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 q1. Which structures are likely to also have values q1 (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 q_1, while D shows behavior characteristic of value z_1. Given your answer to (b), how might a child learn the value for structure C? What value would this be? What kind of evidence would the child have used to learn about structure C?

(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) Using the idea of overhypotheses, how might an English child learn that complex yes/no questions in English rely on structured representations? (Hint: Think about other data in the language that might signal structured representations are preferred over other kinds of representations.) What evidence do we have that this is possible, given realistic English child-directed speech?

(13) 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? (Hint: Think about issues of interacting parameters. Is there more than one possible way to generate this word order?)

(14) 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?)

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

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

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

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

(19) 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.)

(20) 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?

(21) 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.

(22) Why is English a difficult language to learn the metrical stress system for? (Hint: Think about the issue of irregularities. What happens when there's more than one stress contour for a word form? Is it possible for any one grammar to account for all the data?) Do multiple stress contours appear a lot in English child-directed speech?

(23) Is there also a poverty of the stimulus problem for English metricl phonology? Why? (Hint: Think about what it means if data are ambiguous.)

(24) Why would it be helpful to focus only on productive data when trying to figure out the stress system of English? (Hint: Think about what it means to know a system – you know the rules of how to generate a stress contour. Which data seem to follow a regular, predictable rule?)

(25) Do the three metrical stress grammars considered do better when they focus only on capturing productive data?

(26) How does the Tolerance Principle relate to the idea of a data intake filter for the child's acquisitional intake? (Hint: What does it do when there's more than one stress contour for a word form?)

(27) Why might it be hard for English children to learn the currently proposed English grammars in each of the three systems discussed (HV, Hayes, and OT)? (Hint: Think about whether those grammars are the ones that account for the most productive data.) Do any of the systems have minor variants of the proposed English grammars that are much more learnable from English child-directed speech data?