Psych156A/ Ling150
Spring 2012
Homework 3: Phrases & Poverty of the Stimulus
Or "The Language Adventures of Sigmund von Hacklestein, part 3"

Remember to write your full name and University ID number on your assignment. If you collaborate with other students in the class, please make sure to indicate who you worked with.

(91 points total)

(1) Sigmund has been studying yet more details of the Guin language. This time, he's trying to figure out what words make up the phrases of Guin. Here are some data that he's gathered:

Known categories: A, B, C, D

The transitional probability of the category sequence AB is 0.20.
The transitional probability of category sequence AC is 0.10.
The transitional probability of category sequence AD is 0.70.
The transitional probability of category sequence BC is 0.80.
The transitional probability of category sequence BD is 0.10.
The transitional probability of category sequence CD is 0.90.

Assume that the transitional probability learner that Thompson & Newport (2007) believe in will only posit a phrase boundary if the transitional probability between grammatical categories is less than 0.50. For each sequence below, indicate whether this transitional probability learner would be likely to think that sequence is a phrase by stating **Yes** or **No**, and then **briefly explain why you think so**. [3 pts each]

Ex: AB
No.
The transitional probability from A to B is less than 0.50, so this learner will put a phrase boundary between A and B, which makes them part of separate phrases.

(a) ABC
(b) BCD
(c) ABCD
(d) ABD
(e) ACD

(2) Sigmund has been investigating some of the syntactic rules in Guin, and has been particularly interested how Guin children learn these rules. He noticed the following pattern in Guin children's input, having to do with the order of subjects, verbs, and objects:

(i) Subject Verb Object (English translation: "He likes her")
(ii) Verb Subject Object (English translation: "Likes he her")

(iii) Subject Object Verb (English translation: “He her likes”)
(iv) Object Subject Verb (English translation: “Her he likes”)

When Sigmund tested 3-year-old Guin children on which orders they allowed, he found they allowed these four orders and additionally allowed the following two:

(v) Object Verb Subject (English translation: “Her likes he”)
(vi) Verb Object Subject (English translation: “Likes her he”)

(a) Given the the three elements (Subject, Verb, Object), how many possible sentence orderings of these elements are there? [1pt]

(b) Do Guin children observe all the possible sentence orderings in their input? (Yes or No) [1pt]

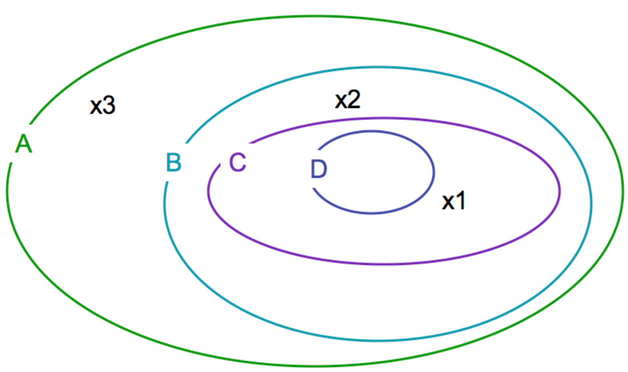
(c) Based on Sigmund’s observations, do **3-year-old** Guin children generalize beyond their input data? (Yes or No) [1pt]

(d) Based on Sigmund’s observations, do **3-year-old** Guin children show constrained generalization? (Yes or No) [1pt]

(e) Sigmund found to his surprise that adult Guin speakers disallow order (vi) [Verb Object Subject], and so only allow orders (i)-(v) in their language. When he tested **5-year-old** Guin children, he found that they too disallowed order (vi). Is this an example of constrained generalization for the Guin **adults and 5-year-olds**? (Yes or No) [1pt]

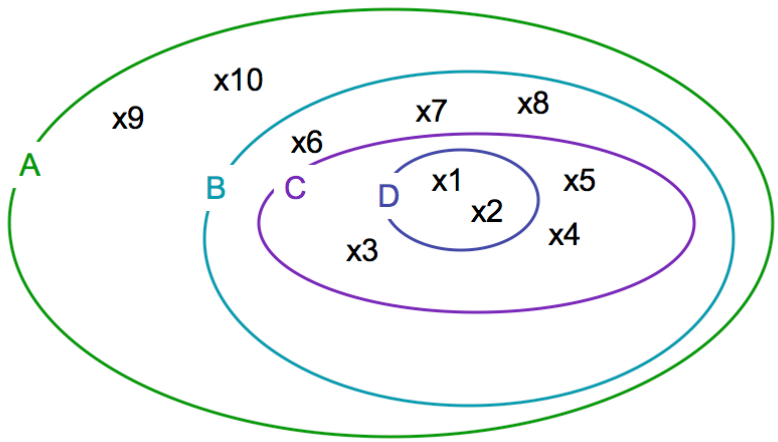
(f) Would Sigmund be able to successfully argue that Guin acquisition behavior of this word order rule must be due to innate knowledge about language? (Yes or No) (Hint: Think about what behavior is expected if children possess innate knowledge. Do *all* Guin children show this behavior?) [1pt]

(3) Below is a schematized picture representing four different generalizations children might make: A, B, C, and D. Sample data points children might encounter are represented by x1, x2, and x3.



(a) Is generalization A a subset of generalization B? (Yes or No) [1pt]
(b) Is generalization B a superset of generalization D? (Yes or No) [1pt]
(c) Is generalization C a subset of generalization A? (Yes or No) [1pt]
(d) Is generalization D a superset of generalization C? (Yes or No) [1pt]
(e) Would initially choosing generalization B when generalization C is actually correct be an example of the Subset problem? (Yes or No) [1pt]
(f) Would initially choosing generalization A when generalization D is actually correct be an example of the Subset problem? (Yes or No) [1pt]
(g) Would initially choosing generalization C when generalization B is actually correct be an example of the Subset problem? (Yes or No) [1pt]
(h) Would initially choosing generalization D when generalization C is actually correct be an example of the Subset problem? (Yes or No) [1pt]
(i) Would the Subset principle help a learner solve the problem of having chosen generalization B when generalization A was actually correct? (Yes or No) [1pt]
(j) Would the Subset principle help a learner solve the problem of having chosen generalization D when generalization C was actually correct? (Yes or No) [1pt]
(k) Would the Subset principle help a learner solve the problem of having chosen generalization C when generalization D was actually correct? (Yes or No) [1pt]
(l) Would the Subset principle help a learner solve the problem of having chosen generalization A when generalization B was actually correct? (Yes or No) [1pt]
(m) Would a Bayesian learner that uses the Size Principle and is sensitive to counterexamples eventually be able to learn that generalization C was correct for a language even if that learner began by thinking generalization B was correct? (Yes or No) [1pt]
(n) Would a Bayesian learner that uses the Size Principle and is sensitive to counterexamples eventually be able to learn that generalization B was correct for a language even if that learner began by thinking generalization C was correct? (Yes or No) [1pt]
(o) Would a Bayesian learner that uses the Size Principle and is sensitive to counterexamples eventually be able to learn that generalization A was correct for a language even if that learner began by thinking generalization D was correct? (Yes or No) [1pt]
(p) Would a Bayesian learner that uses the Size Principle and is sensitive to counterexamples eventually be able to learn that generalization C was correct for a language even if that learner began by thinking generalization A was correct? (Yes or No) [1pt]

(4) Suppose there are four hypotheses that a learner is considering for a particular language, and they account for the examples x1-x10, as shown in the diagram below. Suppose we also assume that x1-x10 are the only potential items that could be in the language.



Suppose the prior probability (before the child has encountered any input data) of each hypothesis is equal, so this is the prior that will be used to calculate the probabilities in (a) and (b) below.

(a) Suppose the child first encounters data point x2.

(i) What is the likelihood of each of the generalizations? (Hint: You need to calculate p(x2 | A), p(x2 | B), p(x2 | C), and p(x2 | D).) [8pts]

(ii) What is the posterior probability of each of the generalizations? (Hint: Use Bayes' Rule to calculate p(A | x2), p(B | x2), p(C | x2), and p(D | x2).) You may round your answers to three decimal places. [8pts]

(b) Suppose the child encounters the following five examples first: x2, x4, x3, x1, x6. In other words, the data *d* that the child needs to account for is this set of data points.

(i) What is the likelihood of each of the generalizations? (Hint: You need to calculate p(*d* | A), p(*d* | B), p(*d* | C), and p(*d* | D).) [8pts]

(ii) What is the posterior probability of each of the generalizations? (Hint: Use Bayes' Rule to calculate p(A | *d*), p(B | *d*), p(C | *d*), and p(D | *d*). Remember that the likelihood of producing five data points is the product of the likelihood of each data point.) You may round your answer to five decimal places. [8pts]

(c) Suppose the child has already encountered data point x2 (and so has reassessed the probability of each generalization already). The posterior probability from (a) now becomes the prior probability for each generalization. Suppose the child then encounters data point x8. In other words, the data *d* that the child needs to account for is x2 and x8.

(i) What is the likelihood of each of the generalizations? (Hint: You need to calculate p(*d* | A), p(*d* | B), p(*d* | C), and p(*d* | D). Remember that the likelihood of producing two data points is the product of the likelihood of each data point.) You may round your answers to four decimal places. [8pts]

(ii) What is the posterior probability of each of the generalizations? (Hint: Use Bayes' Rule to calculate p(A | *d*), p(B | *d*), p(C | *d*), and p(D | *d*), and remember to use your answer from (a-ii) as the prior probability for each generalization). You may round your answers to three decimal places. [8pts]

(5) Sigmund was very intrigued by the idea that computational modeling studies can help answer questions about what kind of knowledge children need to solve different poverty of the stimulus problems. He has been investigating some phenomena in the language of Guin, and has discovered a poverty of the stimulus problem. When discussing this with other language science researchers who have studied Guin, he found out that a current proposal for learning the correct rules pertaining to those phenomena is that children must **innately** know

(i) that a certain kind of **linguistic** structure is important, and that no more than two of these structures are allowed in the same sentence

(ii) that this **linguistic** structure exists

(a) Is the current proposal that Sigmund heard about compatible with a linguistic nativist view? (Yes or No) Is the current proposal that Sigmund heard about compatible with a nativist view that doesn't believe in domain-specific innate knowledge? (Yes or No)
(Hint: Think about where the current proposal falls on the dimensions of domain-specific vs. domain-general, and innate vs. derived.) [2pts]

(b) Sigmund designed a computational modeling study, and discovered that the necessary linguistic structure is derivable from previous experience with the Guin language if the child has innate statistical learning abilities. The knowledge that a child needs is

(i) to **innately** know that a certain kind of **linguistic** structure is important, and that no more than two of these structures are allowed in the same sentence

(ii) how to **derive** this **linguistic** structure from the input data by using **innate** statistical learning abilities

Is this new idea about how to solve the poverty of the stimulus problem compatible with a linguistic nativist view? (Yes or No) Is this new idea compatible with a nativist view that doesn't believe in domain-specific innate knowledge? (Hint: Think about where the current proposal falls on the dimensions of domain-specific vs. domain-general, and innate vs. derived.) [2pts]

(c) Sigmund designed a second computational modeling study, and again discovered that the necessary linguistic structure is derivable from previous experience with the Guin language if the child has innate statistical learning abilities. In addition, Sigmund's new model suggests that the child can also derive the knowledge that no more than two of these structures are allowed in the same sentence through previous experience with other sentences in the Guin language if the child has innate statistical learning abilities. The knowledge that a child needs is

(i) how to **derive** the knowledge that no more than two of these structures are allowed in the same sentence by using **innate** statistical learning

(ii) how to **derive** this **linguistic** structure from the input data by using **innate** statistical learning abilities

Is this new idea about how to solve the poverty of the stimulus problem compatible with a linguistic nativist view? (Yes or No) Is this new idea compatible with a nativist view that doesn't believe in domain-specific innate knowledge? (Hint: Think about where the current proposal falls on the dimensions of domain-specific vs. domain-general, and innate vs. derived.) [2pts]