

## Empirical re-assessment of stimulus poverty arguments

Legate & Yang

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

- The Case
- The challenge
- The rebuttal
- The message

2

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Case

- The logic of APS is simple:  
if you know  $X$ , and  $X$  is underdetermined by learning experience,  
then the knowledge of  $X$  must be innate.

Given language data  $D$ , and a simple but incorrect hypothesis of  $D$ ,  $H$ ,

☞ the child behaves as though he does not entertain  $H$

☞ the evidence necessary to rule out  $H$  is not available to the child

- The child possesses innate knowledge excluding  $H$  from the hypothesis space

3

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Case

- Forming a question in English involves inversion of the main clause auxiliary verb and the subject:

#### *The data*

☞ Is Alex  $e$  singing a song? (Alex is singing a song)

☞ Has Robin  $e$  finished reading? (Robin has finished reading)

☞ Are the boys  $e$  back yet? (The boys are back)

- Many possible hypotheses:
  - Linear: front the first/last auxiliary
  - Linear + hierarchical: front the first auxiliary following the first NP
  - Creative ones
  - ...

4

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ■ The Case

- Many possible hypotheses:
  - Linear: front the first/last auxiliary
  - Linear + hierarchical: front the first auxiliary following the first NP
  - Creative ones
  - ...
- The point: Some of these hypotheses are less plausible, but with DDL the child must eliminate all competing hypotheses.

5

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ■ The Case

- Question formation is structure dependent and involves parsing the sentence:  
fronting the auxiliary that follows the subject NP:

👂 Is [the woman who is singing] e happy?  
👂 Has [the man that is reading a book] e eaten supper?

- Simpler: The first auxiliary hypothesis:

👂 Is [the woman who e singing] is happy?  
👂 Is [the man that e reading a book] has eaten supper?

- APS: if enough evidence is available surely children would reject the first auxiliary hypothesis. BUT, such evidence is virtually absent from the linguistic data.

6

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ■ The Challenge

- P&S challenge to empirical content (not logic – snark!) of APS: children do encounter disconfirming evidence which serves to rule out the incorrect, structure dependent, hypothesis.
- There is only one alternative hypothesis to be ruled out: The first auxiliary hypothesis.

L&Y: This is incorrect. Remember all the possible hypotheses from earlier?  
But, since this would be an even harder task, they accept the challenge!

7

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ■ The Challenge

- It is not only auxiliary inversion in yes/no questions such as 6) that rules out the first auxiliary hypothesis, wh-questions with an inverted auxiliary over a complex NP are also informative (7)

6) Is [the boy who is in the corner]<sub>NP</sub> e smiling?

7) How could [anyone that was awake]<sub>NP</sub> e not hear that?

8

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Challenge

- Corpus search results in 1% of sentences – or 5 out of 500 - of the 6) and 7) type
- Estimates that no less than 0.1 to 1% of all input sentences are of the forms 6) and 7)
- P&S: Point is: the critical evidence does exist!

L&Y: An extra logical step is missing:

The existence of disconfirming evidence says nothing about its *sufficiency* to rule out the first auxiliary hypothesis.

9

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- How much data is sufficient? (we will return to this question)
- Suppose: two problems of acquisition,  $P_1$  and  $P_2$
- Frequency of data that can settle problem,  $F_1$  and  $F_2$
- Suppose:  $P_1$  and  $P_2$  acquired at same developmental stage
- Then expect  $F_1$  and  $F_2$  to be roughly the same
- Conversely:  $F_1$  and  $F_2$  significantly different
- Then  $P_1$  and  $P_2$  must represent different learning problems

10

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- Samples from submitted questions:
  - This assumes that all language rules are independent of all other language rules; that learning of one rule cannot help you learn another rule.
  - If we take prior knowledge into consideration, and A is learned before B, it seems reasonable that knowledge of A can assist in learning of B.

11

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- Take  $F_1$  to be the frequency of sentences like 6) and 7)
- Need a comparative setting of language acquisition (more on the problem with this in the next slide, so hold that thought):
- The subject-drop phenomenon (acquired at same age)  
In some languages (like Spanish) it is optional to drop the subject but this is not done in English
  - Count the frequency of 'there expletive sentences' to get  $F_2$   
1.2%
- V2 (verb second movement) sentences in German and Dutch  
A Lisa style example: Sarah *must* the labyrinth solve
  - Count the frequency of V2  
1.2%

12

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- Samples from submitted questions:
  - These are totally different linguistic phenomena.
- If we take prior knowledge into consideration, and prior knowledge must be equivalent across acquisition problems, then comparison can not be made across languages.

13

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- First auxiliary hypothesis double check:
  - Frequency of 6) and 7) that rule out the first-auxiliary hypothesis should also be approx. 1.2%
- Now we get a little snippy!
  - 'It is rather odd that...P&S cite anecdotes...What's more [the data is] curiously selective.'
- So, they do the 'real' count for all 56 files in the Nina corpus:
  - 46,499 sentences of which 20,651 are questions
  - none were yes/no type in 6) and 14 were wh-questions of the type in 7).
  - $F_7$  approx. 0.068%

14

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- Also, check count from Adam corpus:
  - 20,372 sentences of which 8,889 are questions
  - none were yes/no type in 6) and 4 were wh-questions of the type in 7)
  - $F_7$  approx. 0.045%
- **APS stands unchallenged: the knowledge of structure dependence in syntax is available to the child in the absence of experience.**

15

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Rebuttal

- Samples from submitted questions:
  - A low count is not a non-existence proof. Maybe 1.2% was too high and the true sufficient amount is what ever non-zero proportion of the data that the child encounters.
- Can a calculation of frequency on a corpus be used to determine the critical acquisition threshold?

16

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Message

- The challenge of the APS fails

- Some quotes:

"...DDL, proudly touted as innateness free, turns out to have many hidden assumptions – otherwise known as innate knowledge"

"...a more serious problem with DDL ... has to do with the wild statistical disparities between what is presented to children and how children actually learn"

"...it must be concluded that the innate knowledge of UG provides important learning priors to skew the distributional relations between adult and child languages."

"...innately primed learning is, still, 'the only game in town' "

17

## Empirical re-assessment of stimulus poverty arguments Legate & Yang

### ▪ The Message

- Samples from submitted questions:

- Their argument didn't establish that innate knowledge must be specifically grammatical (UG).

- How specific are the UG biases we are arguing about?

- Differences between adult and child output could be the result of maturation unfolding in the DDL

- Why not develop an artificial language, sprinkle in some bad rules with a higher frequency than normally encountered, and see what happens?

18