Psych 215L: Language Acquisition

Lecture 11

Poverty of the Stimulus

Reminder: Poverty of the Stimulus

Language

Can be thought of as the set of legal items in the language (sentences, strings, etc.). The child's job: figure out the rules that generate that legal set and don't generate illegal items.

Reminder: Poverty of the Stimulus

The Logic of Poverty of the Stimulus (The Logical Problem of Language Acquisition)

1) Suppose there are some data.
2) Suppose there is an incorrect hypothesis compatible with the data.
3) Suppose children behave as if they never entertain the incorrect hypothesis.

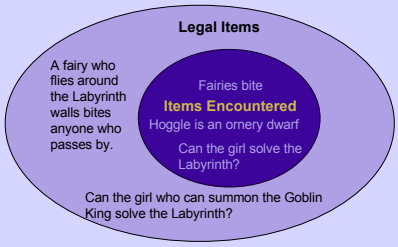
   Addendum (interpretation): Or children converge on the correct hypothesis much earlier than expected (Legate & Yang 2002).

Conclusion: Children possess prior knowledge ruling out the incorrect hypothesis from the hypothesis space considered.

Addendum (Interpretation): The initial hypothesis space does not include all hypotheses. Specifically, the incorrect ones of a particular kind are not in the child's hypothesis space.

Reminder: Poverty of the Stimulus

The argument for having innate biases to guide language acquisition



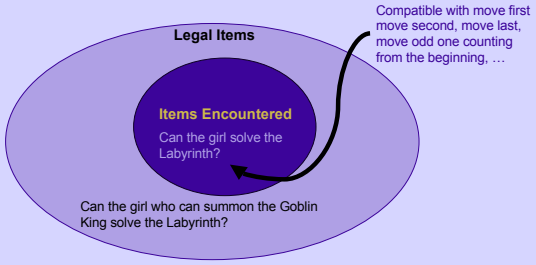
Idea: The data available to the child are compatible with a number of generalizations. However, children only seem to pick the right ones. Therefore, they must have some other constraints guiding their language learning.

The innate part: The guiding information must be available prior to learning.

Reminder: LPLA

Induction Problem: Logical Problem of Language Acquisition (Standard Theory)

Children don't get access to all the data in the language by the time they have the correct generalization. They learn from a subset of the legal items in the language. And *still* they seem to converge on the right generalizations…without trying out many (or all) of the wrong ones.



Pullum & Scholz (2002)

"…linguistic nativism is the view…that human infants have at least some linguistically specific innate knowledge"

"…issue is whether a full description of that predisposition incorporates anything that entails specific contingent facts about natural languages"

[poverty of the stimulus]

"…argument…turns on the claim that during the language acquisition process, children often come to know things about the language they are acquiring despite not having access to the crucial evidence that shows these things to be true of the language."

Pullum & Scholz (2002): Frustration with PoS Proponents

"Instead of clarifying the reasoning, each successive writer on this topic shakes together an idiosyncratic cocktail of claims about children's learning of language, and concludes nativism is thereby supported. Most of the frequently encountered claims are about children's observable accomplishments or aspects of the child's environment."

Pullum & Scholz (2002): Frustration with PoS Proponents

Children's observable accomplishments

Speed: Children learn so fast.

Reliability: Children always succeed.

Productivity: Children learn a system.

Selectivity: Children pick the correct option from a bunch of incorrect (and "seductive") alternatives.

Underdetermination: Children arrive at systems of knowledge underdetermined by the data.

Convergence: Children end up with the right system.

Universality: The system acquired has a lot of properties in common with other language systems of the world.

Pullum & Scholz (2002): Frustration with PoS Proponents

Children's observable accomplishments

Speed: Children learn so fast.

Reliability: Children always succeed.

Productivity: Children learn a system.

Selectivity: Children pick the correct option from a bunch of incorrect (and "seductive") alternatives.

Underdetermination: Children arrive at systems of knowledge underdetermined by the data.

Convergence: Children end up with the right system.

Universality: The system acquired has a lot of properties in common with other language systems of the world.

Pullum & Scholz (2002): Frustration with PoS Proponents

Children's observable accomplishments

Speed: Children learn so fast.

Relevant Interpretation: Faster than expected, given available data.

Selectivity: Children pick the correct option from a bunch of incorrect (and "seductive") alternatives.

Relevant Interpretation: Seductive because also compatible with data.

Underdetermination: Children arrive at systems of knowledge underdetermined by the data.

Relevant Interpretation: Alternative hypotheses also compatible with data.

Pullum & Scholz (2002): Frustration with PoS Proponents

Aspects of Child's Environment

Ingratitude: No explicit payoff for correct language usage.

Finiteness: Children don't get infinite data to learn from.

Idiosyncracy: The subset of data children encounter varies from child to child.

Incompleteness: Children don't hear everything in the language.

Positivity: No explicit instruction of what isn't in the language.

Degeneracy: Input to children has noise.

Pullum & Scholz (2002): Frustration with PoS Proponents

Aspects of Child's Environment

Ingratitude: No explicit payoff for correct language usage.

Finiteness: Children don't get infinite data to learn from.

Idiosyncracy: The subset of data children encounter varies from child to child.

Incompleteness: Children don't hear everything in the language.

Positivity: No explicit instruction of what isn't in the language.

Degeneracy: Input to children has noise.

Pullum & Scholz (2002): Frustration with PoS Proponents

Aspects of Child's Environment

Finiteness: Children don't get infinite data to learn from.
- Relevant Interpretation: Make generalization from incomplete data set.

Idiosyncracy: The subset of data children encounter varies from child to child.
- Relevant Interpretation: Make generalization from incomplete data set.

Incompleteness: Children don't hear everything in the language.
- Relevant Interpretation: Make generalization from incomplete data set.

Positivity: No explicit instruction of what isn't in the language.
- Relevant Interpretation: Make generalization from incomplete data set.

Pullum & Scholz (2002): The Version Chosen To Attack

"People attain knowledge of the structure of their language for which *no* evidence is available in the data to which they were exposed as children." - Hornstein & Lightfoot (1981)

"We replace total lack of evidence by *lack of evidence that is adequate to the task*…would not emerge in conversational data near often enough to guarantee that any particular child would ever encounter it." - Pullum & Scholz

"…'the APS' to stand for 'the Argument Selected by Pullum & Scholz' "

Pullum & Scholz (2002): How to Support APS

Step 1: Describe in detail what is known.

Step 2a: Identify the crucial data that would lead a data-driven learner to that knowledge.

Step 2b: Give reason to believe that's the crucial data.

Step 3: Show learners don't have access to that crucial data.

Step 4: Show that learners nonetheless acquire the right knowledge.

Pullum & Scholz (2002): How to Support APS

Step 1: Describe in detail what is known.

Step 2a: Identify the crucial data that would lead a data-driven learner to that knowledge.

Step 2b: Give reason to believe that's the crucial data.

Step 3: Show learners don't have access to that crucial data.
- One way: Look for really rare data types. These are likely to be close enough to absent.

Step 4: Show that learners nonetheless acquire the right knowledge.

Pullum & Scholz (2002): Case Studies

Case 1: Plurals in noun-noun compounds

3-6 yr olds behavior:

Irregular plural pattern (plural marker on first noun okay)

1 tooth-eater or 1 *teeth-eater*

1 mouse-eater or 1 *mice-eater*

Regular plural pattern (plural marker on first noun not okay)

1 toy-eater (but not 1 *toys-eater*)

1 rat-eater (but not 1 *rats-eater*)

Pullum & Scholz (2002): Case Studies

Case 1: Plurals in noun-noun compounds

Knowledge of incomplete paradigm:

tooth-eater

teeth-eater

toy-eater

**toys-eater*

Important point: No generalization to regular plural nouns.

Pullum & Scholz (2002): Case Studies

Case 1: Plurals in noun-noun compounds

Gordon (1986):

Brown corpus (1,000,000 words):

irregular sg compounds [*tooth-eater*] (153 tokens)

irregular pl compounds [*teeth-eater*] *(3 tokens)*

regular sg compounds [toy-eater] (…more…?)

regular pl compounds [**toys-eater*] *(0 tokens)*

Argument: Irregular pl compounds appear so rarely, they are similar in frequency to regular pl compounds (which never appear because they're ungrammatical.) But children still produce the irregular pl compounds and do not produce the irregular sg compounds. This is hard to explain if they're data-driven. (Though see Foraker et al. (2009) for an example where an ideal learner can make use of even slight differences in data distribution, and then Lidz, Waxman, & Freedman (2003) and Pearl & Lidz (2009) for discussion about how "informative" data that have very low frequency may not be helpful to real learners…)

Pullum & Scholz (2002): Case Studies

Case 1: Plurals in noun-noun compounds

P&S rebuttals:

Not clear 3-6 yr old behavior was really true outside the experimental setup (method flaws).

Point: Not children's behavior.

Not clear that regular pl compounds are ungrammatical. Examples: "rules committee", "chemicals-maker", "citizens-sponsored" appear in Wall Street Journal corpus.

Point: Not adult's behavior either.

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

It rains, It may rain, It may have rained, It may be raining, It has rained, It has been raining, It is raining, It may have been raining,…

Rule: (…Aux Verb: {rains, may rain, may have rained, …}

Aux --> Tense (Modal) (have +en) (be +ing)

{present} {may, might} {have VERBen} {be VERBing}

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

It rains, It may rain, It may have rained, It may be raining, It has rained, It has been raining, It is raining, It may have been raining,…

Rule: (…Aux Verb: {rains, may rain, may have rained, …}

Aux --> Tense (Modal) (have +en) (be +ing)

{present} {may, might} {have VERBen} {be VERBing}

{present} + rain = rains

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

It rains, It may rain, It may have rained, It may be raining, It has rained, It has been raining, It is raining, It may have been raining,…

Rule: (…Aux Verb: {rains, may rain, may have rained, …}

Aux --> Tense (Modal) (have +en) (be +ing)

{present} {may, might} {have VERBen} {be VERBing}

{present} + {Modal} = may rain

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

It rains, It may rain, It may have rained, It may be raining, It has rained, It has been raining, It is raining, It may have been raining,…

Rule: (…Aux Verb: {rains, may rain, may have rained, …}

Aux --> Tense (Modal) (have +en) (be +ing)

{present} {may, might} {have VERBen} {be VERBing}

{past} + {Modal} + {have + en} = may have rained

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

It rains, It may rain, It may have rained, It may be raining, It has rained, It has been raining, It is raining, It may have been raining,…

Rule: (…Aux Verb: {rains, may rain, may have rained, …}

Aux --> Tense (Modal) (have +en) (be +ing)

{present} {may, might} {have VERBen} {be VERBing}

{past} + {have + en} = had rained

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

It rains, It may rain, It may have rained, It may be raining, It has rained, It has been raining, It is raining, It may have been raining,…

Rule: (…Aux Verb: {rains, may rain, may have rained, …}

Aux --> Tense (Modal) (have +en) (be +ing)

{present} {may, might} {have VERBen} {be VERBing}

{past} + {Modal} + {be+ing} = may have been raining

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

Kimball 1973:

Aux --> Tense (Modal) (have +en) (be +ing)

Crucial data to get proper rule sequence have all three optional components:

"It may have been raining"

No examples in 1,000,000 word corpus, vanishingly rare in conversation…

Pullum & Scholz (2002): Case Studies

Case 2: Auxiliary sequences

P&S rebuttal:

Is that rule really what children are acquiring? Instead, children may be able to abstract the necessary sequence from other sequences of not exactly that type.

Also, data not so vanishingly rare: hundreds of examples in adult literature (Moby Dick, Wuthering Heights) and many in children's literature (Peter Pan, Alice in Wonderland, The Wizard of Oz)

Estimate: 1 approximately every 3000-4000 sentences

The real question: How many is enough? Need a quantitative claim from the linguists (see Legate & Yang 2002, Hsu & Chater 2010 for some attempts at this).

Pullum & Scholz (2002): Case Studies

Case 3: Anaphoric One

Originally, Baker 1978

Recently described accessibly in Foraker et al. (2007, 2009) and Pearl & Mis (2011)

"I liked the debate about acquisition. You liked the one about modeling."

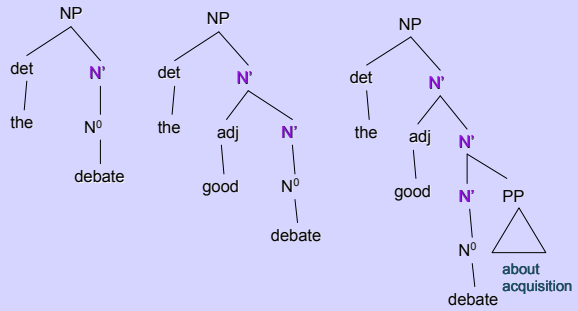
* I'll walk by the side of the road and you can walk by the one of the river."

Syntactic distribution distinction: Difference between complement-taking nouns (*side*) and modifier-taking nouns (*debate*).

Syntactic distinction: complement-taking nouns = Category N^0, modifier-taking nouns = Category N' (larger than N)

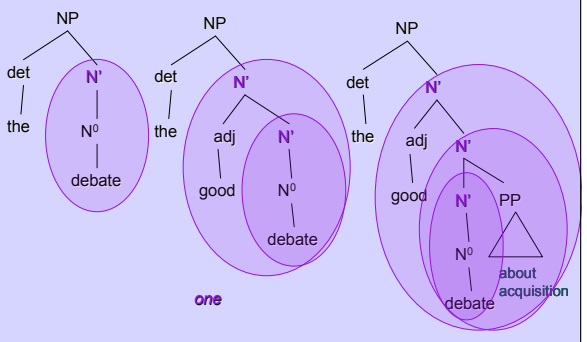
Semantic distinction: Complement-taking nouns are conceptually different from modifier-taking nouns. (*side = side of what?*, *debate* can stand by itself)

Anaphoric *One* structure

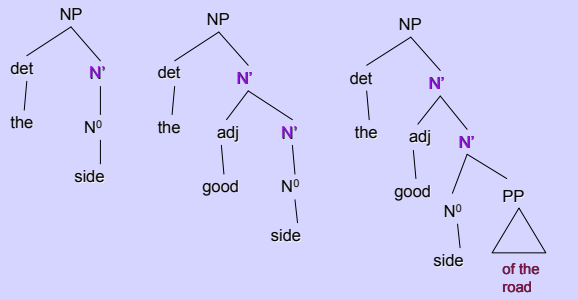


[modifier]

Anaphoric *One* structure

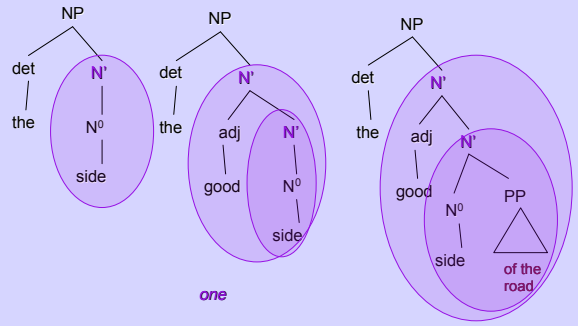


Anaphoric *One* structure



[complement]

Anaphoric *One* structure



Pullum & Scholz (2002): Case Studies

Case 3: Anaphoric One

Originally, Baker 1978

Necessary Data:

Baker 1978: rule out *one* = N^0

Need specific utterance & world situation:

Utterance:

"Look – a red bottle! We want another *one* and there doesn't seem to be *one* here."

Situation

Reasoning:
one ≠ "bottle", since another bottle is present.
Therefore, one = "red bottle", which can only be N' (not N^0)

Pullum & Scholz (2002): Case Studies

Case 3: Anaphoric One

18-month olds behave as if they have the right interpretations
(Lidz, Waxman, & Freedman 2003)

Unambiguous data are pretty rare in child-directed speech:
(Lidz, Waxman, & Freedman 2003; Pearl & Lidz 2009)

~0.25% of anaphoric one utterances are unambiguous for
one != category N^0, but instead one = something larger like N'

Similar P&S rebuttal as before: How rare is too rare?
(see Legate & Yang 2002, Hsu & Chater 2010 for some suggestions for how to quantify "too rare")

Pullum & Scholz (2002): Case Studies

Case 3: Anaphoric One

Rebuttal of another kind: this is not the crucial evidence

"Can be learned from other available data"
Regier & Gahl 2004, Foraker et al. 2009

"…but not without some other learning constraints/knowledge, too"
Foraker et al. 2009, Pearl & Lidz 2009

"…unless you broaden your idea of what counts as informative data…and even then, you may need still need some additional knowledge"
Pearl & Mis (2011)

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

Chomsky 1971: Adult Knowledge

The girl is easily fooled.
Is the girl easily fooled?

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?

Someone who is not easily fooled could trick someone who is.
Could someone who is not easily fooled trick someone who is?

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

Chomsky 1971: Child Behavior (Crain & Nakayama (1987))

The girl is easily fooled.
Is the girl easily fooled?

Rule: Move first auxiliary?

Rule: Move main-clause auxiliary?

Rule: Move odd-numbered auxiliary?

Rule: Move auxiliary next to female noun?

…

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Rule: Move main-clause auxiliary

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

Chomsky 1971: Child Data

The girl is easily fooled.
Is the girl easily fooled?

Very frequent

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Very infrequent

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

P&S rebuttal:

The girl is easily fooled.
Is the girl easily fooled?

Very frequent

I could borrow your pencil when you're done.
When you're done, could I borrow your pencil?

Rules out "front first aux" hypothesis:
Should be very frequent

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Very infrequent

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

P&S rebuttal:

The girl is easily fooled.
Is the girl easily fooled?

Very frequent

The changes these events portend are how fundamental.
How fundamental are the changes these events portend?

Rules out "front first aux" hypothesis, though not in yes/no questions:
15th sentence in WSJ corpus

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Very infrequent

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

P&S rebuttal:

The girl is easily fooled.
Is the girl easily fooled?

Very frequent

What I'm doing is in the shareholders' best interest.
Is what I'm doing in the shareholders' best interest?

Rules out "front first aux" hypothesis:
180th sentence in WSJ corpus

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Very infrequent

Legate & Yang (2002): Rebuttal

Data set = 500 sentences of the Wall Street Journal
"How fundamental are the changes these events portend?"
"Is what I'm doing in the shareholders' best interest?"

Not really a good sample of child-directed speech, unless you have a very unique child.

Also, Legate & Yang (2002) found that 1% are of this data type (5 of the 500 sentences).

Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

P&S rebuttal:

The girl is easily fooled.
Is the girl easily fooled?

Very frequent

The other dolly that was in here is where.
Where's the other dolly that was in here?

Rules out "front first aux" hypothesis, though not yes/no question:
Child-directed speech

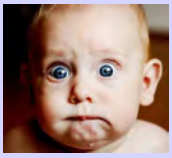
The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Very infrequent

Legate & Yang (2002): Rebuttal

Looking at the Nina corpus:
- 46,499 sentences
- 20,651 questions
- 14 unambiguous data examples (all of wh-question type)

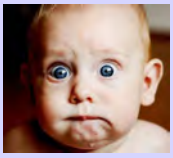
Frequency of unambiguous data: 0.068%



Legate & Yang (2002): Rebuttal

Looking at the Brown-Adam corpus:
- 20,372 sentences
- 8,889 questions
- 4 unambiguous data examples (all of wh-question type)

Frequency of unambiguous data: 0.045%



Pullum & Scholz (2002): Case Studies

Case 4: Auxiliary Fronting

P&S rebuttal:

Main point (similar to previous ones): How much data is enough?

Legate & Yang (2002) and Hsu & Chater (2010) offer some ideas for how to assess this.

The girl who can solve the labyrinth is easily fooled.
Is the girl who can solve the labyrinth easily fooled?
* Can the girl who solve the labyrinth is eaily fooled?

Very infrequent

Pullum & Scholz (2002): Summary

Linguists should be careful about what knowledge they think children are acquiring.

It's not that there is no evidence for the child to learn from in most cases. It's just that it's rare. It's an open question about how rare is too rare.

Additional Note: Larger point about PoS (from Crain & Pietroski (2002))

"…it's not enough to mention ways in which children could learn some things without Universal Grammar. To rebut poverty-of-the-stimulus arguments, one has to show how children could learn [everything] adults actually know; and as close investigation reveals, adults know a lot more than casual inspection suggests. *That* is the nativist's main point."

Examples of linguistic knowledge that some researchers believe is hard:

- restrictions on meaning interpretation (Crain & Pietroski 2002)
- restrictions on syntactic case (Valian 2009, ch2, section 2.4)
- restrictions on syntactic islands (Pearl & Sprouse 2011)

Innate Bias = Domain-Specific?

Poverty of the Stimulus (the existence of an induction problem) is usually used as the motivation for Universal Grammar. But just because an induction problem exists doesn't mean innate domain-specific knowledge like UG is required to solve it. The knowledge required could be derived from prior knowledge (domain-specific or domain-general) or simply be domain-general to begin with.

Exploring the Nature of the Necessary Bias(es): Computational Modeling Work

Domain-general biases explored:

- -prefer subset hypothesis: Regier & Gahl 2004
- -prefer simplicity: Perfors, Tenenbaum, & Regier 2006, 2011
- -use only maximally informative data: Pearl & Weinberg 2007, Pearl 2008, Pearl & Lidz 2009, Pearl & Mis 2011
- -prefer highly probable sequences: Reali & Christiansen 2005, Kam et al. 2008, Pearl & Sprouse 2011

Domain-specific biases explored:

- -ignore certain kinds of ambiguous data that are identified using domain-specific (linguistic) knowledge: Regier & Gahl 2004, Pearl & Lidz 2009, Pearl & Mis 2011
- -ignore embedded clause data: Pearl & Weinberg 2007
- -prefer syntactic information over semantic information: Foraker et al. 2009