# Psych 156A/ Ling 151: Acquisition of Language II

Lecture 15

Poverty of the stimulus III

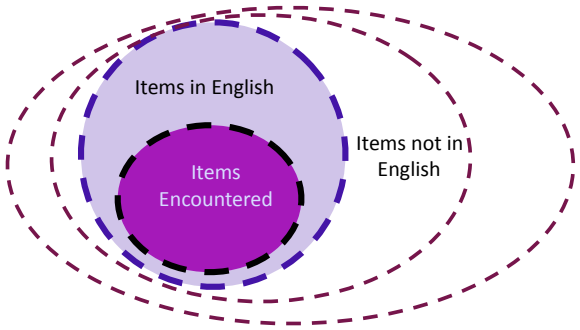
## Announcements

Be working on HW3 (due: 5/26/16)

Please fill out course evaluations :)

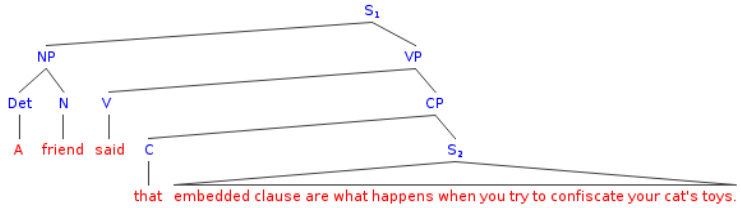
## Poverty of the stimulus

The data encountered are compatible with both the correct hypothesis and other incorrect hypotheses about the rules and patterns of the language.




## Reasonable questions

- What are some examples of linguistic knowledge that seem to present a poverty of the stimulus situation?
  - Structure-dependent rules



## English yes/no questions

Crain & Nakayama (1987) showed that children as young as 3 years old have restrictions on they form complex yes/no questions. They use some kind of structure-dependent rule.

Rule: Move main clause auxiliary

Is the girl who can solve the labyrinth happy?

## English yes/no questions

The problem is that simple yes/no questions are compatible with a lot of different rules, both structure-independent and structure-dependent.

Rule: Move first auxiliary?

Jareth can alter time.
Can Jareth alter time?

Rule: Move last auxiliary?

Rule: Move main clause auxiliary?

Rule: Move auxiliary in even-numbered position in sentence?

Rule: Move auxiliary closest to a noun?

## English yes/no questions

The correct rule is a structure-dependent rule (it requires the child to know that sentences can be divided into main and embedded clauses).

~~Rule: Move first auxiliary?~~

Jareth can alter time.
Can Jareth alter time?

~~Rule: Move last auxiliary?~~

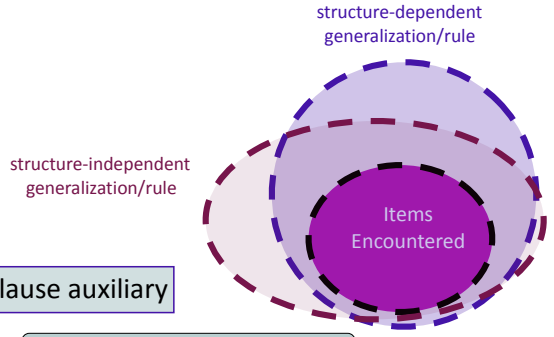
Rule: Move main clause auxiliary?

~~Rule: Move auxiliary in even-numbered position in sentence?~~

~~Rule: Move auxiliary closest to a noun?~~

## English yes/no questions

How do children choose the right rule from all the possible rules that are compatible? That is, how do they generalize the right way from the subset of the data they encounter?

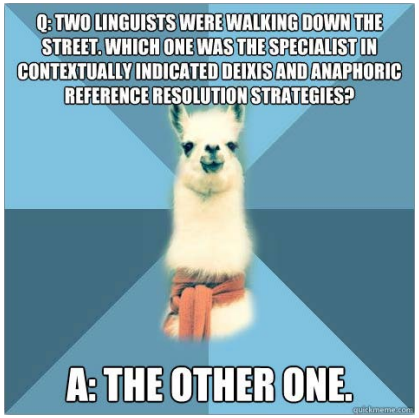



Rule: Move main clause auxiliary

Is the girl who can solve the labyrinth happy?

## Reasonable questions

- What are some examples of linguistic knowledge that seem to present a poverty of the stimulus situation?
  - Structure-dependent rules
  - Anaphoric *one* in English




## English anaphoric *one*

Look - a red bottle!

## English anaphoric *one*

Look - a red bottle!

Look – another *one*!

## English anaphoric *one*

Look - a red bottle!

red bottle

Look – another *one*!

Process: First determine the antecedent of *one* (what expression *one* is referring to).
→"red bottle"

## English anaphoric *one*

Look - a red bottle!

red bottle

Look – another *one*!

Process: Because the antecedent ("red bottle") includes the modifier "red", the property RED is important for the referent of *one* to have.

→ referent of *one* = RED BOTTLE

## English anaphoric *one*

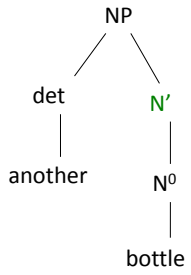
Look - a red bottle!

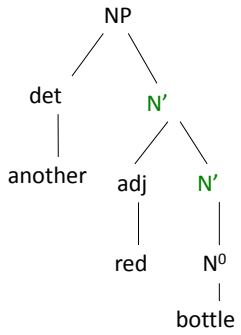
Look – another *one*!

Two steps:
(1) Identify linguistic antecedent
(2) Identify referent (based on linguistic antecedent)

## Anaphoric *one*: Syntactic category

Standard linguistic theory (Chomsky 1970, Jackendoff 1977) posits that *one* in these kinds of utterances is a syntactic category smaller than an entire noun phrase (NP), but larger than just a noun ($N^0$). This category is N'. This category includes strings like "bottle" and "red bottle".



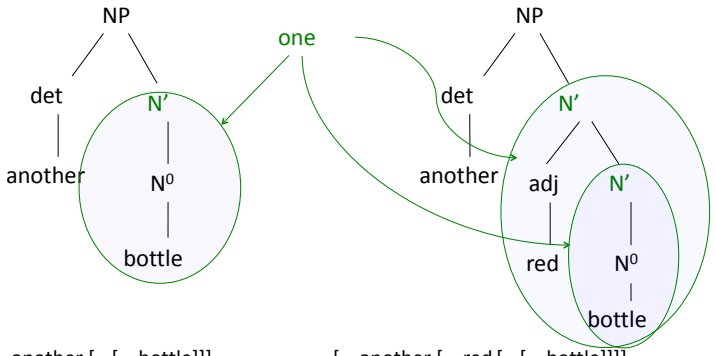




[NP another [N' [N0 bottle]]]

[NP another [N' red [N' [N0 bottle]]]]

## Anaphoric *one*: Syntactic category

Standard linguistic theory (Chomsky 1970, Jackendoff 1977) posits that *one* in these kinds of utterances is a syntactic category smaller than an entire noun phrase (NP), but larger than just a noun ($N^0$). This category is N'. This category includes strings like "bottle" and "red bottle".

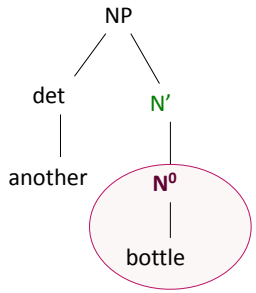



[NP another [N' [N0 bottle]]]

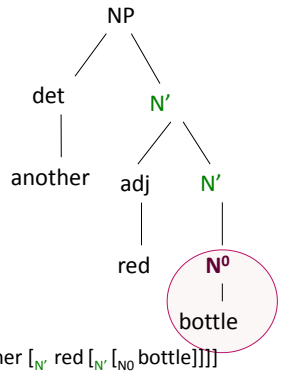
[NP another [N' red [N' [N0 bottle]]]]

## Anaphoric *one*: Syntactic category

Importantly, *one* is not $N^0$. If it was, it could only have strings like "bottle" as its antecedent, and could never have strings like "red bottle" as its antecedent.




[$_{NP}$ another [$_{N'}$ [$_{N0}$ bottle]]]




[$_{NP}$ another [$_{N'}$ red [$_{N'}$ [$_{N0}$ bottle]]]]

## Anaphoric *one*: Interpretations based on syntactic category

If *one* was $N^0$, we would not be able to have the "red bottle" interpretation:

"Look – a red bottle!  Look – another one!"

Because *one*'s antecedent could only be "bottle", we would have to interpret the second part as "Look - another *bottle*!"

Since *one*'s antecedent can be "red bottle", and "red bottle" cannot be $N^0$, *one* must not be $N^0$ (in this context at least).

## Anaphoric *one*: Adult knowledge

"Look – a red bottle! Look – another one!"

≈ "Look – a red bottle! Look – another red bottle!"

Target knowledge state for acquisition:

Syntactic knowledge: category N'

Referential knowledge: mentioned property ("red") is included in the linguistic antecedent (antecedent = "red bottle"), so referent has property.

## Anaphoric *one*: Adult knowledge

"Look – a red bottle! Look – another one!"

≈ "Look – a red bottle! Look – another red bottle!"

Target behavior state (based on target knowledge state):

In this scenario, adults expect to see another red bottle – not just another bottle. So, they will look for a second red bottle.

## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

"Look – a red bottle!" 

"Now look…"

Control:
"What do you see now?"

Anaphoric:
"Do you see another one?"

## Anaphoric *one*: Children's knowledge

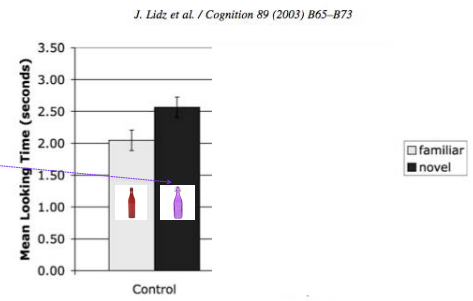
Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

"Look – a red bottle!" 

"Now look…"

Control:
"What do you see now?"
Baseline novelty preference:
[~2.0s  vs. ~2.5s  ]

Anaphoric:
"Do you see another one?"




## Anaphoric *one*: Children's knowledge

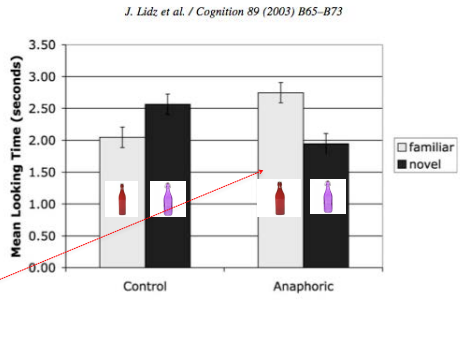
Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

"Look – a red bottle!" 

"Now look…"

Control:
"What do you see now?"
Baseline novelty preference

Anaphoric:
"Do you see another one?"
[~2.75s  vs. ~1.95s  ]
Adjusted familiarity preference




## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

"Look – a red bottle!" 

"Now look…"

Noun:
"Do you see another bottle?"

Adjective-noun:
"Do you see another red bottle?"

## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

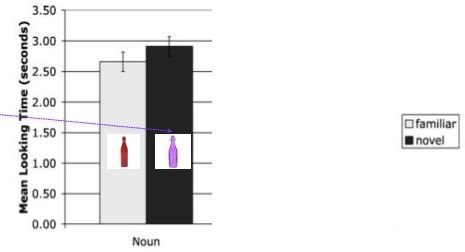
"Look – a red bottle!" 

"Now look…"

Noun:
"Do you see another bottle?"
Baseline novelty preference:
[~2.65s  vs. ~2.95s  ]

Adjective-noun:
"Do you see another red bottle?"

J. Lidz et al. / Cognition 89 (2003) B65–B73




## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

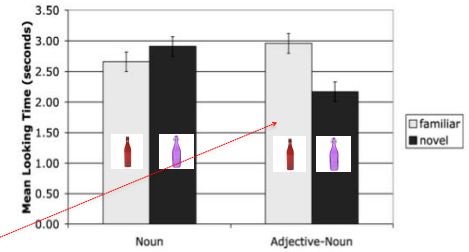
"Look – a red bottle!" 

"Now look…"

Noun:
"Do you see another bottle?"
Baseline novelty preference

Adjective-noun:
"Do you see another red bottle?"
[~3.0s  vs. ~2.1s  ]
Adjusted familiarity preference

J. Lidz et al. / Cognition 89 (2003) B65–B73




## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

"Look – a red bottle!" 

"Now look…"

Control/Noun:
"What do you see now?"
"Do you see another bottle?"
Baseline novelty preference
Average probability of looking to familiar bottle: 0.459

Anaphoric/Adjective-Noun:
"Do you see another one?"
"Do you see another red bottle?"
Adjusted familiarity preference
Average probability of looking to familiar bottle: 0.587

## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

LWF interpretation:
Given 18-month-olds' baseline novelty preference and adjusted familiarity preference, preference for RED BOTTLE means the preferred antecedent is "red bottle".

## Anaphoric *one*: Children's knowledge

Lidz, Waxman, & Freedman (2003) [**LWF**] investigated 18-month-old behavior in this scenario.

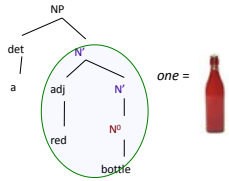
LWF interpretation:

Given 18-month-olds' baseline novelty preference and adjusted familiarity preference, preference for RED BOTTLE means the preferred antecedent is "red bottle".

LWF conclusion about 18-month-old knowledge state:

(1) syntactic category of *one* = N'

(2) linguistic antecedent when modifier is present (i.e., property is mentioned) includes modifier (e.g., "red") = referent has modifier property




## Anaphoric *One*: The induction problem

Acquisition: Children must learn the right syntactic category for *one*, and the right interpretation preference for *one* in situations with more than one option.

## Anaphoric *One*: The induction problem

Acquisition: Children must learn the right syntactic category for *one*, and the right interpretation preference for *one* in situations with more than one option.

Problem: Most direct evidence children encounter is ambiguous.

Syntactically (SYN) ambiguous data

(92% according to corpus study by Pearl & Mis 2011, 2016):

"Look – a bottle! Oh, look – another one."

*one*'s referent = BOTTLE

*one*'s antecedent = $[_{N'}[_{N0}$ bottle]] or $[_{N0}$ bottle]?

## Anaphoric *One*: The induction problem

Acquisition: Children must learn the right syntactic category for *one*, and the right interpretation preference for *one* in situations with more than one option.

Problem: Most data children encounter are ambiguous.

Referentially and syntactically (REF-SYN) ambiguous

(8% according to corpus study by Pearl & Mis 2011, 2016)

"Look – a red bottle! Oh, look – another one."

*one*'s referent = RED BOTTLE or BOTTLE?

*one*'s antecedent = $[_{N'}$ red$[_{N'}[_{N0}$ bottle]]] or $[_{N'}[_{N0}$ bottle]] or $[_{N0}$ bottle]?

## Anaphoric *One*: The induction problem

Acquisition: Children must learn the right syntactic category for *one*, and the right interpretation preference for *one* in situations with more than one option.

Problem: Unambiguous data are extremely rare

Unambiguous (UNAMB) data

(0% according to corpus study by Pearl & Mis 2011, 2016)

"Look – a red bottle! Hmmm - there doesn't seem to be another one here, though."

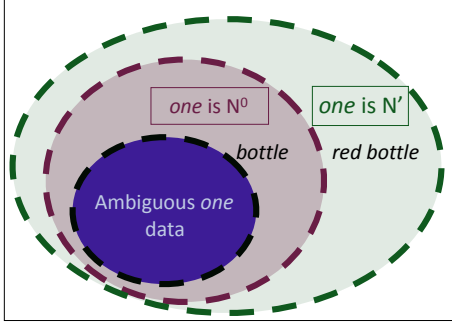
*one*'s referent = BOTTLE? If so, *one*'s antecedent = "bottle".
But it's strange to claim there's not another *bottle* here.
So, *one*'s referent must be RED BOTTLE, and *one*'s antecedent = $[_{N'}$ red$[_{N'}[_{N0}$ bottle$]]]$.

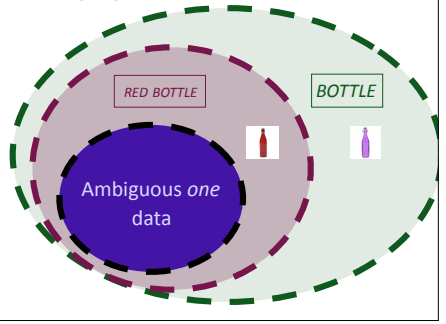
## Anaphoric *One*: The induction problem

Acquisition: Children must learn the right syntactic category for *one*, and the right interpretation preference for *one* in situations with more than one option.

Problem: If children don't encounter unambiguous data often enough (or at all), they are left with data that are compatible with both hypotheses – that *one* is $N^0$ and that *one* is N' for the syntax, that *one*'s antecedent doesn't include the modifier (BOTTLE) and that it does (RED BOTTLE) for the meaning.

How do children know which is the right generalization?




## Anaphoric *One*: The induction problem

Acquisition: Children must learn the right syntactic category for *one*, and the right interpretation preference for *one* in situations with more than one option.

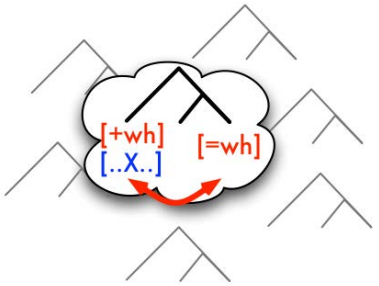
Problem: If children don't encounter unambiguous data often enough (or at all), they are left with data that are compatible with both hypotheses – that *one* is $N^0$ and that *one* is N' for the syntax, that *one*'s antecedent doesn't include the modifier (BOTTLE) and that it does (RED BOTTLE) for the meaning.

How do children know which is the right generalization?

See Pearl & Mis (2016) for an answer:
"…these results suggest that children may be leveraging broader sets of data to make the syntactic generalizations leading to their observed behavior…"

## Reasonable questions

- What are some examples of linguistic knowledge that seem to present a poverty of the stimulus situation?
  - Structure-dependent rules
  - Anaphoric *one* in English
  - Syntactic islands



## Syntactic islands

Dependencies between a wh-word and where it's understood (its gap) can exist when these two items are not adjacent, and these dependencies do not appear to be constrained by length (Chomsky 1965, Ross 1967).

What does Jack think __?
What does Jack think that Lily said __?
What does Jack think that Lily said that Sarah heard __?
What does Jack think that Lily said that Sarah heard that Jareth stole __?

## Syntactic islands

However, if the gap position appears inside certain structures (called "syntactic islands" by Ross 1967), the dependency seems to be ungrammatical.

*What did you make [the claim that Jack bought __]?
*What do you think [the joke about __] offended Jack?
*What do you wonder [whether Jack bought __]?
*What do you worry [if Jack buys __]?
*What did you meet [the scientist who invented __]?
*What did [that Jack wrote __] offend the editor?
*What did Jack buy [a book and __]?
*Which did Jack borrow [__ book]?

## The input: Induction problems

Data from five corpora of child-directed speech (Brown-Adam, Brown-Eve, Valian, Suppes, Valian) from CHILDES (MacWhinney 2000): speech to 25 children between the ages of one and five years old.

Total utterances: 813,036
Utterances containing a *wh*-dependency: 31,247

Pearl & Sprouse (2013, 2015) discovered that more complex dependencies were fairly rare in general (<0.01% of the input).

## The input: Induction problems

What kind of *wh*-dependencies are in the input?

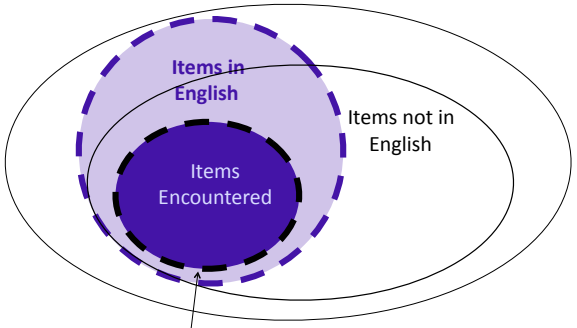
| | |
|---|---|
| 76.7% | *What did you see __?* |
| 12.8% | *What __ happened?* |
| 5.6% | *What did she want to do __?* |
| 2.5% | *What did she read from __?* |
| 1.1% | *What did she think he said __?* |

…

## The input: Induction problems

Important: Some grammatical utterances never appeared at all. This means that only a subset of grammatical utterances appeared, and the child has to generalize appropriately from this subset.

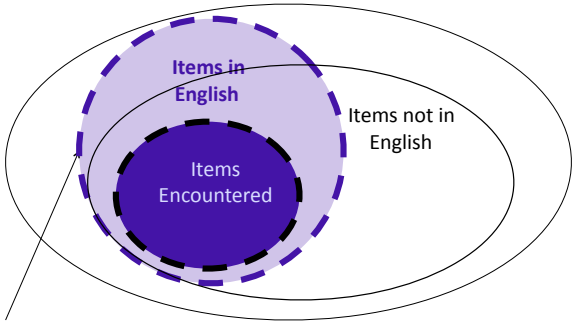
## Syntactic Islands: Induction Problem




*wh*-questions in input (usually fairly simple)

What did you see?
What happened?
…

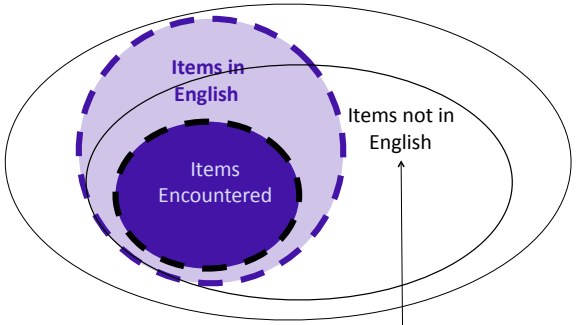
## Syntactic Islands: Induction Problem




Grammatical *wh*-questions

What did you see?
What happened?
Who did Jack think that Lily saw?
What did Jack think happened?

## Syntactic Islands: Induction Problem




Ungrammatical *wh*-questions: Syntactic islands

*What did you make [the claim that Jack bought __]?
*What do you think [the joke about __] offended Jack?
*What do you wonder [whether Jack bought __]?
*What do you worry [if Jack buys __]?

## Syntactic islands:
## One answer for some of the islands

*What did you make [the claim that Jack bought __]?
*What do you think [the joke about __] offended Jack?
*What do you wonder [whether Jack bought __]?
*What do you worry [if Jack buys __]?
*What did you meet [the scientist who invented __]?
*What did [that Jack wrote __] offend the editor?
*What did Jack buy [a book and __]?
*Which did Jack borrow [__ book]?

Pearl & Sprouse 2013, 2015: Learn what you can from the dependencies you do actually observe in the data and apply it to make a judgment about the dependencies you haven't seen before, like these syntactic islands.

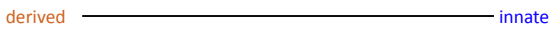
That is, leverage a broader set of data to make syntactic generalizations.

## Reasonable questions

- When there is a poverty of the stimulus situation, what kind of "knowledge" do children need in order to end up with the right answer?
  - Knowledge kinds (at least three dimensions to consider):

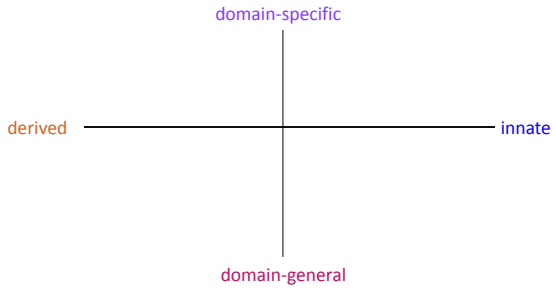
## Reasonable questions

- When there is a poverty of the stimulus situation, what kind of "knowledge" do children need in order to end up with the right answer?
  - Knowledge kinds (at least three dimensions to consider):
    innate vs. derived from prior (language) experience



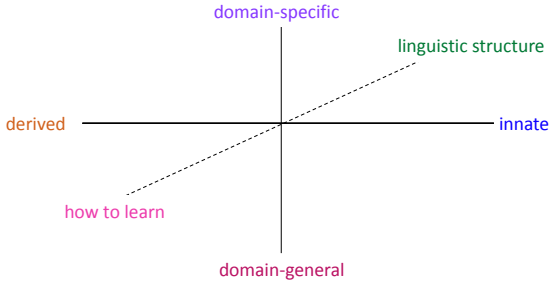

## Reasonable questions

- When there is a poverty of the stimulus situation, what kind of "knowledge" do children need in order to end up with the right answer?
  - Knowledge kinds (at least three dimensions to consider):
    innate vs. derived from prior (language) experience
    domain-specific vs. domain-general



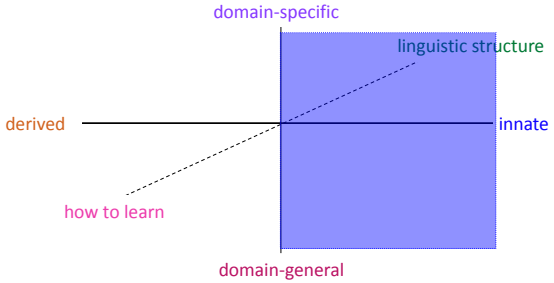
## Reasonable questions

- When there is a poverty of the stimulus situation, what kind of "knowledge" do children need in order to end up with the right answer?
  - Knowledge kinds (at least three dimensions to consider):
    innate vs. derived from prior (language) experience
    domain-specific vs. domain-general
    about linguistic structure vs. about how to learn



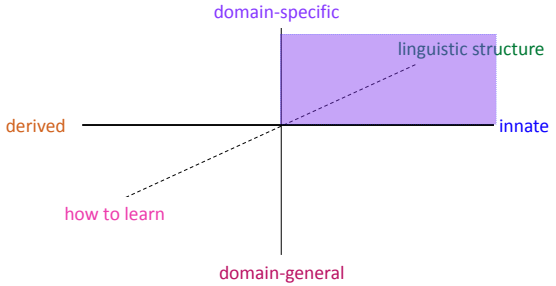

## Reasonable questions

Nativists believe that the necessary knowledge is innate, but may be either domain-specific or domain-general.



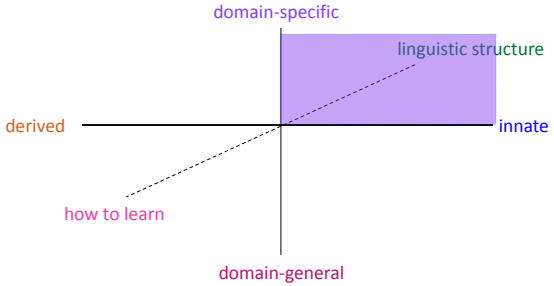

## Reasonable questions

Linguistic nativists believe that the necessary knowledge is both innate and domain-specific. This is sometimes called the Universal Grammar (UG) hypothesis. Linguistics nativists believe that because children have Universal Grammar, they can solve these poverty of the stimulus problems.




## Universal Grammar

Ideas for what could be in Universal Grammar often come from examining specific language acquisition problems, and figuring out what would be needed to solve those problems.



## Reasonable questions

- How can we test different ideas about what the necessary knowledge might be?
  - Computational modeling studies can help us identify the necessary knowledge

We can construct a computational model where we have precise control over these important components of the language acquisition process:

- The hypotheses the child is considering at any given point
  [hypothesis space]
- How the child represents the data & which data the child uses
  [data intake]
- How the child changes belief based on those data
  [update procedure]

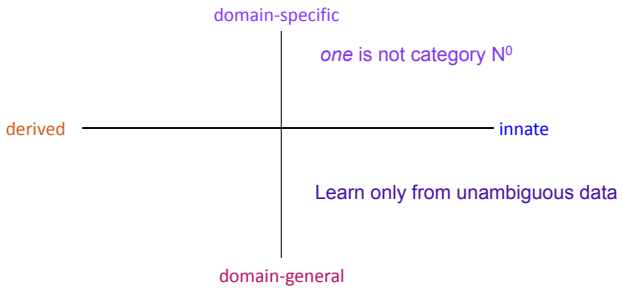
## Computational modeling studies

Several recent computational models have attempted to address poverty of the stimulus questions, and rely on probabilistic learning (often Bayesian inference) as the main method of learning. By modeling the acquisition process for these linguistic phenomena, these models hope to pinpoint the kind of knowledge required for language acquisition.

- Anaphoric *one*: Regier & Gahl (2004), Foraker et al. (2009), Pearl & Lidz (2009), Pearl & Mis (2011, 2013, 2016)
- Syntactic islands: Pearl & Sprouse (2013a, 2013b, 2015)

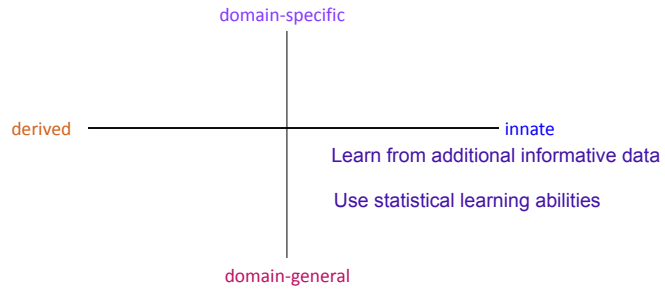
## English anaphoric *one*

Baker (1978) assumed only unambiguous data are informative, and these data are rare (or non-existent). So, he proposed that children needed to know that *one* could not be syntactic category $N^0$.



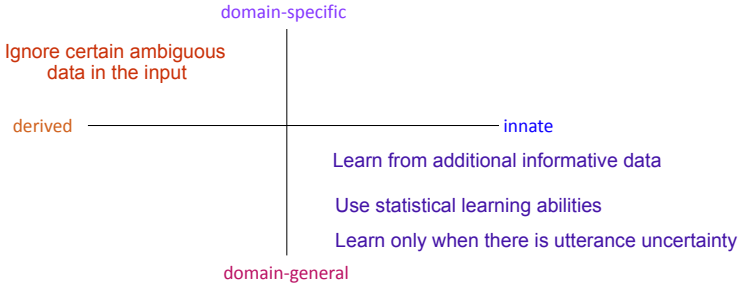

## English anaphoric *one*

Regier & Gahl (2004) used a Bayesian learner computational model to show that children could learn *one* is category N' if they learned from some of the available ambiguous data and used their statistical learning abilities to track suspicious coincidences in the input.



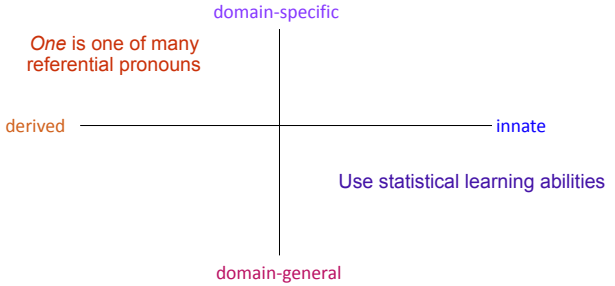
## English anaphoric *one*

Pearl & Lidz (2009) discovered that a Bayesian learner must ignore certain ambiguous data (even if they're informative) in order to learn that *one* is category N'. This can be derived from an innate, domain-general preference for learning when there is uncertainty about the meaning of the utterance heard.




## English anaphoric *one*

Pearl & Mis (2011, 2016) discovered that a Bayesian learner can learn from all ambiguous *one* data and still learn to interpret *one* appropriately in experiments like Lidz, Waxman, & Freedman (2003) if the learner also learns from data containing other referential pronouns like *it*.

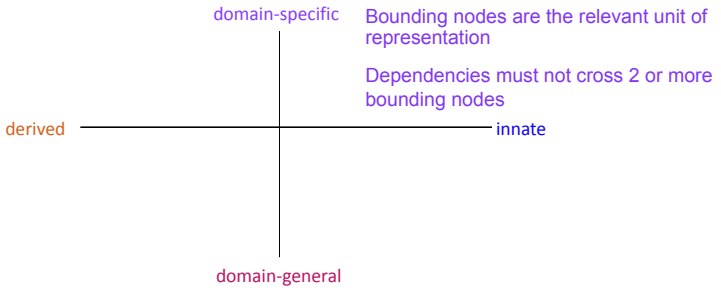



## Syntactic islands

Chomsky (1973), Huang (1982), and Lasnik & Saito (1984) proposed that children must know that dependencies cannot cross 2 or more bounding nodes (a language-specific representation that refers to certain types of phrases).


## Syntactic islands

Chomsky (1973), Huang (1982), and Lasnik & Saito (1984) proposed that children must know that dependencies cannot cross 2 or more bounding nodes (a language-specific representation that refers to certain types of phrases).



# Syntactic islands

Pearl & Sprouse (2013a, 2013b, 2015) discovered that a probabilistic learner that tracks sequences of a different linguistic abstract representation (container nodes) can learn some of the syntactic islands.

[$_{CP}$ What did [$_{IP}$ you [$_{VP}$ see __]]]?
IP VP ← Container nodes

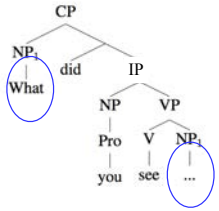
Container node sequence: IP-VP

# Syntactic islands

Pearl & Sprouse (2013a, 2013b, 2015) discovered that a probabilistic learner that tracks sequences of a different linguistic abstract representation (container nodes) can learn some of the syntactic islands.

[$_{CP}$ What did [$_{IP}$ you [$_{VP}$ see __]]]?
IP VP

Container node sequence: IP-VP



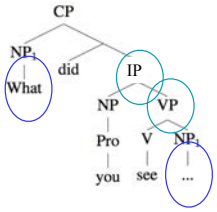

How to describe this dependency:
What phrases is the gap inside but the wh-word isn't inside?

# Syntactic islands

Pearl & Sprouse (2013a, 2013b, 2015) discovered that a probabilistic learner that tracks sequences of a different linguistic abstract representation (container nodes) can learn some of the syntactic islands.

[$_{CP}$ What did [$_{IP}$ you [$_{VP}$ see __]]]?
IP VP

Container node sequence: IP-VP

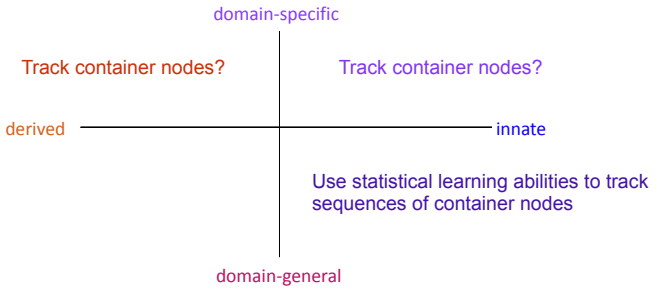



How to describe this dependency:
What phrases is the gap inside but the wh-word isn't inside?

What did you see __?
= What did [$_{IP}$ you [$_{VP}$ see __]]?
= IP-VP

# Syntactic islands

Pearl & Sprouse (2013a, 2013b, 2015) discovered that a probabilistic learner that tracks sequences of a different linguistic abstract representation (container nodes) can learn some of the syntactic islands. It's unclear whether container nodes are innate or derived knowledge.



## Recap

- There are several examples of poverty of the stimulus problems in language learning, such as the representation of English anaphoric *one*, and the existence of syntactic islands.
- Children require some knowledge to help them solve these problems, but there are different kinds of knowledge they could have.
- Nativists believe at least some of the knowledge is innate. Linguistic nativists believe that at least some of the knowledge is both innate and specific to language.
- Computational modeling studies can help us determine what knowledge is necessary for successful acquisition to occur.

## Questions?

You should be able to do all the questions on HW3 and the review questions for poverty of the stimulus. Use the rest of the time to work on these and ask us questions about them.