Psych 215L: Language Acquisition

Lecture 13
Poverty of the Stimulus: Anaphoric One

An induction problem by any other name...

One of the most controversial claims in linguistics is that children face an induction problem:

“Plato’s Problem” (Chomsky 1980, Dresher 2003)

Basic claim:
The data encountered are compatible with multiple hypotheses.

The induction problem

Extended claim:
Given this, the data are insufficient for identifying the correct hypothesis as quickly as children do (Legate & Yang 2002) – or at all.

Big question: How do children do it, then?

One answer: Children come prepared

- Children are not unbiased learners.
- But if children come equipped with helpful learning biases, then what is the nature of these necessary biases?
  - Are they innate or derived from the input somehow?
  - Are they domain-specific or domain-general?
  - Are they about what’s being learned or about how to learn?

The Universal Grammar (UG) hypothesis (Chomsky 1965, Chomsky 1993):
These biases are innate and domain-specific.
Induction problems, UG, and informative data

The direct evidence assumption

If you want to learn linguistic knowledge $L$, you learn it by observing examples of $L$ in your input (and possibly by also being sensitive to indirect negative evidence about what examples are missing from the input.)

Learning complex yes/no questions

Direct evidence $L$:
"Is the boy who is in the corner $t$, happy?"

Possible indirect negative evidence:
"Is the boy who $t$ is in the corner is happy?"

The direct evidence assumption

Traditional Idea
Induction problems $\rightarrow$ Universal Grammar (UG)

The direct evidence assumption

Learning the representation of English anaphoric one

Direct evidence $L$:
"I see a red bottle... but there isn’t another one around."

Possible indirect negative evidence:
"I like the student of linguistics and he likes the one of computer science."

The direct evidence assumption

Learning syntactic islands

Direct evidence $L$:
"What did the teacher think $t_{sw}$ inspired the students?"
"Who did the teacher think the letter from the soldier inspired $t_{sw}$?"
"Who $t_{sw}$ thought the letter from the soldier inspired the students?"

Possible indirect negative evidence:
"Who did the teacher think [the letter from $t_{sw}$] inspired the students?"
A broader set of informative data

**Indirect evidence:** other kinds of data that may also be relevant, thereby broadening the set of informative data

Recent computational models have been exploring this:
- Complex yes/no questions (Perfors, Tenenbaum, & Regier 2006, 2011)
- Anaphoric one (Regier & Gold 2004, Pearl & Lids 2009, Forsaier et al. 2009)

Mapping out UG & the acquisition process

**Big questions:**
- When induction problems exist, what does it take to solve them?
  - What indirect evidence is available? How might a child leverage this evidence?
  - What learning biases can get the job done, given the available data? Are they necessarily innate and domain-specific (UG)?
- How can the necessary learning biases inform us about how the acquisition process works?

**Anaphoric One**

**Look - a red bottle!**

Do you see another **one**?

**Anaphoric One**

**Look - a red bottle!**

Do you see another **one**?

Process: First determine the antecedent of **one** (what string **one** is referring to). \( \rightarrow \) "red bottle"
Look - a red bottle!

Do you see 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

Anaphoric One

Two steps:
1. Identify syntactic antecedent (based on syntactic category of one)
2. Identify semantic referent (based on syntactic antecedent)

Anaphoric One: Syntactic Category

Standard linguistic theory claims that one in these kind of utterances is a syntactic category smaller than an entire noun phrase, but larger than just a noun (N⁰). This category is sometimes called N’. This category includes strings like “bottle” and “red bottle”.

NP
  det
  another
  N⁰

NP
  det
  another
  adj
  N⁰

N’s

bottle

NP
  det
  another
  adj
  red
  N⁰

N’s

bottle

NP
  det
  one
  N⁰

N’s

bottle

NP
  det
  another
  adj
  red
  N⁰

N’s

bottle
Importantly, one is not N0. If it was, it could only have strings like "bottle" as its antecedent, and could never have strings like "red bottle" as its antecedent.

If one was N0, we would have a different interpretation of "Look – a red bottle! Do you see another one?"

Because one’s antecedent could only be "bottle", we would have to interpret the second part as "Do you see another bottle?" and the purple bottle would be a fine referent for one.

Since one's antecedent is "red bottle", and "red bottle" cannot be N0, one must not be N0.

Lidz, Waxman, & Freedman (2003) [LWF] found that 18-month-olds have a preference for the red bottle in the same situation.

"Look – a red bottle! Do you see another one?"

LWF interpretation & conclusion: Preference for the RED BOTTLE means the preferred syntactic antecedent is "red bottle".

LWF conclude that 18-month-old knowledge = syntactic category of one = N syntactic antecedent when modifier is present includes modifier (e.g., red) = referent has modifier property

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 data children encounter are ambiguous.

Syntactically (SYN) ambiguous data:

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

One's referent = BOTTLE
One's antecedent = \([n\_bottle]\) or \([w\_bottle]\)?

Previous proposals for learning about *one*


How then? Children have innate, domain-specific knowledge restricting the hypotheses about *one*: cannot be syntactic category N°.

What about when there are multiple N° antecedents? \([n\_red\_bottle]\) or \([w\_bottle]\)?

(No specific proposal for this.)
Previous proposals for learning about one referent have included using innate domain-general statistical learning abilities to track how often one's referent has the mentioned property (e.g., red). If the referent often has the property (RED BOTTLE), this is a suspicious coincidence unless the antecedent really does include the modifier ("red bottle") and one's category is N.\[N N\]

Pearl & Lidz 2009 \[P&L\]: Ambiguous data cannot be leveraged, even if they are used. Look—another one!

Why? These data cause an "equal-opportunity" probabilistic learner to think one's category is N0.\[N0\]

Foraker et al. 2009 \[F&al\]: Leverage the syntactic distribution of one with innate domain-general statistical learning, by using subtle domain-specific semantic distinctions that indicate syntactic category. "ball with stripes", "side of the road", "one with dots", "one of the river"\[modifiers\]\[complements = conceptually evoked by head noun\]

How? Indirect negative evidence (never seeing one with a complement, even though other nouns take complements) indicates one is not N0.

A new proposal: Broadening the data set

Pearl & Mis, submitted \[P&M\]: Other pronouns in the language can also be used anaphorically: him, her, it, … Look at the cute penguin. I want to hug him/her/it.\[NP him/her/it\]\[NP the \[N cute \[N N0 penguin\]\]\]

Look! A cute penguin. I want one.\[NP one\]\[NP a \[N cute \[N N0 penguin\]\]\]

Note: The issue of one's category only occurs when one is used in a syntactic environment that indicates it is smaller than an NP (<NP).
A new proposal: Broadening the data set

Pearl & Mis, submitted [P&M]: Track how often the referent of the anaphoric element (one, him, her, it, etc.) has the property mentioned in the potential antecedent, using innate domain-general statistical learning abilities.

Important: This applies, even when the syntactic category is known.

Look at the cute penguin. I want to hug him/her/it. Look! A cute penguin. I want one.

Is the referent cute? Yes! So it's important that the antecedent include the modifier "cute".

Data set comparisons:

Learners using syntactic and semantic information

Unamb <NP
"Look at a red bottle! Hmm - there doesn't seem to be another one here, though."
Learners: Baker, R & G, P & L's EO, P & M

Sem-Syn Amb
"Look at a red bottle! Oh, look - another one!"
Learners: R & G, P & L's EO, P & M

Syn Amb
"Look at a red bottle! Oh, look - another one!"
Learners: P & L's EO, P & M

Unamb NP
"Look at a red bottle! I want one/it."
Learners: P & M

Information in the data

previous context = ex: "...a red bottle..."

current usage = pronoun ex: "...another one..."

REFERENTIAL INTENT
Property mentioned?

Property important?

Antecedent string includes property?

Antecedent string includes modifier?

SYNTACTIC USAGE
Pronoun used

Syntactic category of pronoun

Syntactic environment

Actual antecedent string

Object referred to

Observed

Latent
Information in the data

**previous context** = ex: "...a red bottle..."

**current usage = pronoun**

ex: "...another one..."

**REFERENTIAL INTENT**
- Property mentioned?
- Property important?
- Antecedent string includes property?

**SYNTACTIC USAGE**
- Pronoun used
- Syntactic category of pronoun
- Antecedent string includes modifier?
- Syntactic environment

**Actual antecedent string**

**Observed**

**Latent**

Object referred to

"red bottle",

"bottle"
Information in the data

previous context =
ex: "...a red bottle..."

current usage = pronoun
ex: "...another one..."

PROPERTY MENTIONED?

Syntactic category of pronoun

Antecedent string includes property?

 Antarecedent string
includes modifier?

Syntactic environment

Actual antecedent string

Object referred to

Information in the data: Unamb <NP

previous context =
ex: "...a red bottle..."

current usage = pronoun
ex: "...another one..."

PROPERTY MENTIONED?

Syntactic category of pronoun

Antecedent string includes property?

 Antarecedent string
includes modifier?

Syntactic environment

Actual antecedent string

Object referred to

Information in the data: Sem-Syn Amb

previous context =
ex: "...a red bottle..."

current usage = pronoun
ex: "...another one..."

PROPERTY MENTIONED?

Syntactic category of pronoun

Antecedent string includes property?

 Antarecedent string
includes modifier?

Syntactic environment

Actual antecedent string

Object referred to

Information in the data: Syn Amb

previous context =
ex: "...a bottle..."

current usage = pronoun
ex: "...another one..."

PROPERTY MENTIONED?

Syntactic category of pronoun

Antecedent string includes property?

 Antarecedent string
includes modifier?

Syntactic environment

Actual antecedent string

Object referred to
The online probabilistic framework

### General form of update equations for $p_I$ (adapted from Chew 1971):

$$ p_I = \frac{\alpha + \text{data}_x}{\alpha + \beta + \text{totaldata}_x} $$

Where:
- $\alpha = 1$ (A very weak prior)
- $\beta$ is the total informative data seen w.r.t $x$

After every informative data point encountered:
- $\text{data}_x = \text{data}_x + \text{data}$ (Incremented by probability that data point suggests $x$ is true)
- $\text{totaldata}_x = \text{totaldata}_x + 1$ (One informative data point seen)

### The online probabilistic framework:

#### Updating $p_I$

<table>
<thead>
<tr>
<th>Property mentioned</th>
<th>Property important?</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td>$\frac{\rho}{\rho + p_1 + p_2}$ Probability property is important</td>
</tr>
<tr>
<td>yes</td>
<td>yes</td>
<td>$p_I$</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>$\frac{1}{\rho + p_1 + p_2}$ Probability property is not important</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>$\frac{p_1}{\rho + p_1 + p_2}$ Category = N', choose N' with modifier, property is important</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>$\frac{1}{\rho + p_1 + p_2}$ Category = N', property is not important, choose object with property by chance</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>$\frac{1}{\rho + p_1 + p_2}$ Category = N, property is not important, choose object with property by chance</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>$\frac{1}{\rho + p_1 + p_2}$ Category = N, property is not important, choose object with property by chance</td>
</tr>
</tbody>
</table>

### The online probabilistic framework:

**Information in the data: Unamb NP**

**previous context =**

"...a red bottle..."

**current usage = pronoun**

"...want one..."

**REFERENTIAL INTENT**

Property mentioned? Yes

Property important? Yes

Antecedent string includes property? Yes

**SYNTACTIC USAGE**

Pronoun used: one

Antecedent category of pronoun: NP

Antecedent string includes modifier? Yes

**Syntactic environment**

Property important? Yes

Syntactic category of pronoun: NP

**Object referred to:** a red bottle

**Observed Latent**

Actual antecedent string: a red bottle

**The online probabilistic framework**

Tracking the probability that a property mentioned in the potential antecedent is important: $p_I$

#### $p_I$ Property mentioned = yes

Property important? Yes

Tracking the probability that the syntactic category is N' when it is smaller than NP: $p_N'$

#### $p_N'$ Syntactic category of pronoun = N'

Syntactic environment = <NP
The online probabilistic framework:

<table>
<thead>
<tr>
<th>Explanation</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unamb &lt;NP</td>
<td>1</td>
</tr>
<tr>
<td>Unamb NP</td>
<td>N/A</td>
</tr>
<tr>
<td>Syn Amb</td>
<td>( \frac{\rho}{m+n} )</td>
</tr>
<tr>
<td>Sem-Syn Amb</td>
<td>( \frac{\rho^2 + \rho}{m+n} )</td>
</tr>
</tbody>
</table>

\( \rho = \frac{m}{m+n} \): Category = N’, choose N’ with modifier, property is important

\( \rho = \frac{n}{m+n} \): Category = N’, choose N’ without modifier, property is not important, choose object with property by chance

\( \rho = \frac{(1-p^2)(1-t^2)}{f} \): Category = N’, property is not important, choose object with property by chance

Example updates

Start with \( p_{N'} = p_I = 0.50 \)

One Unamb <NP data point: \( p_{N'} = 0.67, p_I = 0.67 \)

One Unamb NP data point: \( p_{N'} = 0.50, p_I = 0.67 \)

One Sem-Syn Amb data point: \( p_{N'} = 0.56, p_I = 0.47 \)

One Syn Amb data point: \( p_{N'} = 0.58, p_I = 0.62 \)

One Syn Amb data point: \( p_{N'} = 0.48, p_I = 0.50 \)

Corpus Analysis & Learner Input

Brown/Eve corpus (CHILDES; MacWhinney 2000): starting at 18 months

17,521 utterances of child-directed speech, 2,874 referential pronoun utterances

<table>
<thead>
<tr>
<th>N’</th>
<th>0.66%</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>7.52%</td>
</tr>
<tr>
<td>Unamb NP</td>
<td>8.42%</td>
</tr>
<tr>
<td>Uninformative</td>
<td>83.4%</td>
</tr>
</tbody>
</table>

Pearl & Lida (2009): Children learn one’s representation between 14 and 18 months.

Based on estimates of the number of utterances children hear from birth until 18 months (Akhtar et al., 2004), we can calculate the data distribution in their input (36,500 referential pronoun utterances total).
Corpus Analysis & Learner Input

Learner Input based on Brown/Eve corpus distributions

<table>
<thead>
<tr>
<th></th>
<th>Baker</th>
<th>R&amp;G</th>
<th>P&amp;L’s EO</th>
<th>P&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unamb &lt;NP</td>
<td>0.00%</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sem-Syn Amb</td>
<td>0.66%</td>
<td>0</td>
<td>242</td>
<td>242</td>
</tr>
<tr>
<td>Syn Amb</td>
<td>7.52%</td>
<td>0</td>
<td>0</td>
<td>2743</td>
</tr>
<tr>
<td>Unamb NP</td>
<td>8.42%</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Uninformative</td>
<td>88.3%</td>
<td>36500</td>
<td>36258</td>
<td>33515</td>
</tr>
</tbody>
</table>

Pearl & Lidz (2009): Children learn one’s representation between 14 and 18 months.

Based on estimates of the number of utterances children hear from birth until 18 months (Akhtar et al., 2004), we can calculate the data distribution in their input (36,500 referential pronoun utterances total).

Measures of Success: LWF children’s behavior

In addition to directly assessing \( p_I \) and \( p_{N'} \), we can measure how often a learner would reproduce the behavior in the LWF experiment.

1. Look – a red bottle!
2. Do you see another one?

Testing LWF’s assumption about what behavior means

In addition to directly assessing the learner’s behavior, we can assess LWF’s assumption that correct behavior indicates the children have the correct representation for one.

Is it possible to get correct behavior in the LWF experiment without having the correct representation for one in general (as measured by \( p_I \) and \( p_{N'} \))? Is it possible to get correct behavior in the LWF experiment without having the correct representation for one at the time the behavior is being produced?

\[
\begin{align*}
\rho_1 & = p_N * \frac{n_1}{m_1} + p_I * \frac{n_1}{m_1} & \text{Category} = N', \text{antecedent} = \text{red bottle} \\
\rho_2 & = p_N * \frac{n_1}{m_1} + (1 - p_N) * \frac{n_1}{m_1} & \text{Category} = N', \text{antecedent} = \text{bottle} \\
\rho_3 & = (1 - p_N) * (1 - p_I) * \frac{n_1}{m_1} & \text{Category} = N', \text{antecedent} = \text{bottle} \\
\end{align*}
\]

\[
\begin{align*}
\rho_1 + \rho_2 + \rho_3 \geq \rho_I \text{ given the learner looks at the red bottle}
\end{align*}
\]
### Learner Results

Averages over 1000 simulations, standard deviations in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Baker</th>
<th>R&amp;G</th>
<th>P&amp;L’s EO</th>
<th>P&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_i )</td>
<td>0.50 (&lt;.01)</td>
<td>0.95 (&lt;.01)</td>
<td>0.02 (&lt;.01)</td>
<td>&gt;0.99 (&lt;.01)</td>
</tr>
<tr>
<td>( p_{ NI' } )</td>
<td>0.50 (&lt;.01)</td>
<td>0.97 (&lt;.01)</td>
<td>0.17 (&lt;.02)</td>
<td>0.37 (.04)</td>
</tr>
<tr>
<td>( p(\text{LWF behavior}) )</td>
<td>( p(\text{correct representation when producing LWF behavior}) )</td>
<td>0.53 (&lt;.01)</td>
<td>0.93 (&lt;.01)</td>
<td>0.50 (&lt;.01)</td>
</tr>
</tbody>
</table>

As previous studies found:
- Traditional unambiguous data alone fails (Baker).
- Leveraging Sem-Syn ambiguous data succeeds (R&G, P&L).
- Leveraging Syn ambiguous data in addition fails (P&L’s EO).

New result: Leveraging Unamb NP data (P&M) does not yield the correct representation in general \( p_{ NI' } \) is low, but...

#### What this means:
LWF’s assumption that correct behavior indicates the child has the correct representation does not seem to hold.

### Learner Results

Averages over 1000 simulations, standard deviations in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Baker</th>
<th>R&amp;G</th>
<th>P&amp;L’s EO</th>
<th>P&amp;M</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_i )</td>
<td>0.50 (&lt;.01)</td>
<td>0.95 (&lt;.01)</td>
<td>0.02 (&lt;.01)</td>
<td>&gt;0.99 (&lt;.01)</td>
</tr>
<tr>
<td>( p_{ NI' } )</td>
<td>0.50 (&lt;.01)</td>
<td>0.97 (&lt;.01)</td>
<td>0.17 (&lt;.02)</td>
<td>0.37 (.04)</td>
</tr>
<tr>
<td>( p(\text{LWF behavior}) )</td>
<td>( p(\text{correct representation when producing LWF behavior}) )</td>
<td>0.53 (&lt;.01)</td>
<td>0.93 (&lt;.01)</td>
<td>0.50 (&lt;.01)</td>
</tr>
</tbody>
</table>

New result: The probability of producing the LWF behavior with this incorrect representation is high.

#### How does this work?
If \( p_i \) is high, then when a property is mentioned (like "red"), the learner believes that property is relevant – which means the referent must include that property (RED BOTTLE).

#### Or does it?
When the child produces the correct behavior in the LWF experiment, the probability that the child has the correct representation when making that interpretation is very high, even if the probability for the correct representation in general (e.g., when there is no modifier present) is very low.

#### Upshot:
LWF were not wrong about children’s representation when interpreting utterances like those in their experiment.
Learner Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Baker</th>
<th>R&amp;G</th>
<th>P&amp;L or E &amp; M</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_j$</td>
<td>0.50</td>
<td>0.95</td>
<td>0.02</td>
</tr>
<tr>
<td>$p_{j'}$</td>
<td>0.50</td>
<td>0.97</td>
<td>0.17</td>
</tr>
<tr>
<td>$p(LWF \ behavior)$</td>
<td>0.53</td>
<td>0.93</td>
<td>0.50</td>
</tr>
<tr>
<td>$p(\text{correct representation when producing LWF behavior})$</td>
<td>0.22</td>
<td>0.92</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Also, the other learners behave as LWF expect:

When they show the correct behavior, they have the correct representation.
When they show incorrect behavior, they have the incorrect representation.

Recap & Implications

• Children may be able to learn the correct interpretation for one in certain situations (such as the LWF experiment) by broadening the set of data they consider relevant.

• Just because children demonstrate that they have the correct interpretation some of the time does not mean they have the correct representation all of the time.

Recap & Implications

• While children must eventually learn the correct representation of one, they do not necessarily need to do so by 18 months.

• Instead, they may realize that one’s category is N’ (rather than N0) at some later point.

One possibility:

[F & al] (2014) proposed to use domain-general statistical learning, by using sub-lexical domain-specific semantic distinctions that indicate syntactic category.

The Acquisition Trajectory

Before 18 months:

- Need domain-specific knowledge
- Recognize that one is similar to other anaphoric elements (it, him, etc.)

- How to get it?
  - Derive it by using innate domain-general statistical learning abilities to observe the distribution of one compared to these other elements.
The Acquisition Trajectory

**Before 18 months:**
Track how often a mentioned property is important for a referent to have.

How to get it?
Use innate domain-general statistical learning abilities to track this.

**18 months:**
Be able to assign the correct interpretation to utterances like those in the LWF experiment. (Know that one is \( N \) in these cases.)

"Look – a red bottle! Do you see another one?"

**After 18 months:**
Need domain-specific knowledge about subtle semantic distinctions that indicate syntactic category in order to leverage the syntactic distribution of one with innate domain-general statistical learning.

How?
May come from innate domain-specific knowledge (UG) about language.

---

**Back to the bigger questions**

- When induction problems exist, what does it take to solve them?
  - What indirect evidence is available? How might a child leverage this evidence?
  - Broader data sets that are identifiable via innate domain-general learning abilities may be additional sources of useful information.
Back to the bigger questions

- When induction problems exist, what does it take to solve them?
  • What learning biases can get the job done, given the available data?
    Are they necessarily innate and domain-specific (UG)?
  → In this case study, the first step may not involve this kind of knowledge, although achieving the final adult knowledge state may.

<table>
<thead>
<tr>
<th>Stage I</th>
<th>18-month-old behavior</th>
<th>Stage II</th>
</tr>
</thead>
<tbody>
<tr>
<td>derived domain-specific knowledge</td>
<td>derived domain-specific knowledge?</td>
<td></td>
</tr>
<tr>
<td>innate domain-general statistical learning</td>
<td>innate domain-general statistical learning</td>
<td></td>
</tr>
</tbody>
</table>

Back to the bigger questions

- How can the necessary learning biases inform us about how the acquisition process works?
  → identify learning biases needed to achieve 18-month-old behavior
  → identify knowledge state those biases suggest
  → suggest a two stage acquisition process for learning anaphoric one

<table>
<thead>
<tr>
<th>Stage I</th>
<th>18-month-old behavior</th>
<th>Stage II</th>
</tr>
</thead>
<tbody>
<tr>
<td>derived domain-specific knowledge</td>
<td>derived domain-specific knowledge?</td>
<td></td>
</tr>
<tr>
<td>innate domain-general statistical learning</td>
<td>innate domain-general statistical learning</td>
<td></td>
</tr>
</tbody>
</table>

The big picture

- Indirect evidence does not necessarily mean indirect negative evidence – it can come from considering a broader pool of informative data
- Indirect evidence does not necessarily negate the need for learning biases (of whatever kind)
- Considering indirect evidence and its impact on acquisition can help define concrete proposals about what is necessarily innate and domain-specific, and thus what is in Universal Grammar
- Knowing the impact of the necessary learning biases on acquisition may also inform us about the acquisition trajectory