A new way to find developmentally-meaningful variation in children’s input: A look at syntactic knowledge across socio-economic status

Lisa Pearl
University of California, Irvine

Computation of Language Laboratory
UC Irvine

Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lppearl@uci.edu

July 7, 2020
ForMA, Institute of Language Studies
Unicamp, Brazil
There’s lots of variation in children’s input
Developmentally-meaningful variation impacts language development
Meaningful input deficits can lead to language delays.
If there’s an input-based language delay and we know what the crucial input deficit is, we can intervene and fix that deficit.
Input-based language delays appear across socio-economic status (SES). Lower-SES children are often behind their higher-SES peers.
Low-SES language input can differ from high-SES input in both overall quantity of speech and the quality of that speech (Hart & Risley 1995, Huttenlocher et al. 2010, Rowe 2012, Schwab & Lew-Williams 2016, Rowe et al. 2017).
Quality can be measured by different aspects of the input, like diversity of vocabulary …
Quality can be measured by different aspects of the input, like diversity of vocabulary, diversity of syntactic constructions …
Quality can be measured by different aspects of the input, like diversity of vocabulary, diversity of syntactic constructions, and frequency of decontextualized speech.

The kitty wasn't there
Because we're going tomorrow

The penguins should be at the zoo
Because the penguins were being fed.

We'll see the kitty on Friday
Because we're going tomorrow

We saw her yesterday, didn't we?
How can we tell if any particular input difference is developmentally-meaningful (that is, it impacts language development)?
One (standard) way:
• Notice that there’s a difference
One (standard) way:
• Notice that there’s a **difference**
• Measure language acquisition **outcomes**
One (standard) way:
• Notice that there’s a **difference**
• Measure language acquisition **outcomes**
• See if that input difference **correlates** with any **outcome differences**
One (standard) way:

- Notice that there’s a **difference**
- Measure language acquisition **outcomes**
- See if that input difference **correlates** with any **outcome differences**

If so, then the input difference **might cause** the outcome difference and so be **meaningful**.
One (standard) way

≠

A new (complementary) way uses developmental computational modeling.
One (standard) way

A developmental computational model implements a specific learning theory …
One (standard) way

A developmental computational model implements a specific learning theory about how children use their input …
One (standard) way

A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.
One (standard) way

A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.

Important: the learning theory implemented by the model specifies what aspect of the input matters.
One (standard) way

A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.

If we know what input part matters, we can target that part for intervention if needed.
One (standard) way

So, a developmental computational model can predict the language outcome on the basis of the input.
One (standard) way

If the predicted outcomes differ, then it's because the input difference caused that outcome difference. So, the input difference is predicted to be meaningful.
One (standard) way

If the predicted outcomes differ, then it’s because the input difference caused that outcome difference. So, the input difference is predicted to be meaningful.

These outcome predictions will need to be verified, though.
One (standard) way

Bonus: Because the learning theory in the model is causal, we can predict if the input should cause similar outcomes, too.

In that case, the input difference isn’t meaningful.
Detecting meaningful input differences

One (standard) way

A new (complementary) way
Detecting meaningful input differences

One (standard) way

Today’s focus

A new (complementary) way
Detecting meaningful input differences

One (standard) way

Today’s focus

A new (complementary) way

Case study:
Syntactic island acquisition

Who does…
One (standard) way

Today’s focus

A new (complementary) way

Case study: Syntactic island acquisition

Why? It’s higher-order syntactic knowledge where we don’t know much about developmentally-meaningful input differences across SES.
Syntactic island constraints involve *wh*-dependencies.

*This kitty was bought as a present for someone.*

*Lily thinks this kitty is pretty.*

*Who does Lily think the kitty for is pretty?*

*What does Lily think is pretty, and who does she think it’s for?*
Syntactic island constraints involve *wh*-dependencies.

**What’s going on here?**

There’s a *dependency* between the *wh*-word *who* and where it’s understood (*the gap*)

*Who does Lily think the kitty for *who* is pretty?*

This dependency is *not allowed* in English.

One explanation: The dependency crosses a “syntactic island” (Ross 1967)
Syntactic island constraints involve *wh*-dependencies.

What’s going on here? syntactic island (Ross 1967)

*Who does Lily think the kitty for __who is pretty?* Subject island
Syntactic island constraints involve *wh*-dependencies.

*Who does Lily think the kitty for* __who is pretty?__  *Subject island*

*What’s going on here?*

Jack is somewhat tricksy.

He claimed he bought something.

*What did Jack make the claim that he bought* __what?__
Syntactic island constraints involve *wh*-dependencies.

What’s going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for who is pretty? Subject island

What did Jack make the claim that he bought what? Complex NP island

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth wondered if he actually did and what it was.

What did Elizabeth wonder whether Jack bought what?
Syntactic island constraints involve *wh*-dependencies.

Who does Lily think the kitty for *who* is pretty?  
What did Jack make the claim that he bought *what*?  
What did Elizabeth wonder whether Jack bought *what*?

---

Jack is somewhat tricksy.  
He claimed he bought something.  
Elizabeth worried it was something dangerous.

What did Elizabeth worry if Jack bought *what*?
Syntactic island constraints involve *wh*-dependencies.

What’s going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for _who is pretty? Subject island
What did Jack make the claim that he bought _what? Complex NP island
What did Elizabeth wonder whether Jack bought _what? Whether island
What did Elizabeth worry if Jack bought _what? Adjunct island

Important: It’s not about the length of the dependency.

(Chomsky 1965, Ross 1967)
Syntactic island constraints involve *wh*-dependencies.

Who does Lily think the kitty for __who is pretty?  [Subject island]
What did Jack make the claim that he bought __what?  [Complex NP island]
What did Elizabeth wonder whether Jack bought __what?  [Whether island]
What did Elizabeth worry if Jack bought __what?  [Adjunct island]

Important: It’s not about the length of the dependency.

Elizabeth
Syntactic island constraints involve *wh*-dependencies.

**Who does Lily think the kitty for** __who__ is pretty?  
**What did Jack make the claim that he bought** __what__?  
**What did Elizabeth wonder whether Jack bought** __what__?  
**What did Elizabeth worry if Jack bought** __what__?

Important: It’s not about the length of the dependency.

**What did Elizabeth think Jack said** __what__?
Syntactic island constraints involve *wh*-dependencies.

Who does Lily think the kitty for *who* is pretty?  
What did Jack make the claim that he bought *what*?  
What did Elizabeth wonder whether Jack bought *what*?  
What did Elizabeth worry if Jack bought *what*?  

What did Elizabeth think Jack said Lily saw *what*?

Important: It’s not about the length of the dependency.
Syntactic island constraints involve *wh*-dependencies.

Who does Lily think the kitty for **who** is pretty?  **Subject island**

What did Jack make the claim that he bought **what**?  **Complex NP island**

What did Elizabeth wonder whether Jack bought **what**?  **Whether island**

What did Elizabeth worry if Jack bought **what**?  **Adjunct island**

High-SES adults judge these dependencies to be far worse than many others, including others that are very similar except that they don’t cross syntactic islands (Sprouse et al. 2012).

These judgments are an observable behavior signaling that acquisition of syntactic island knowledge has occurred.
Syntactic island constraints involve *wh*-dependencies.

Who does Lily think the kitty for *who* is pretty? **Subject island**

What did Jack make the claim that he bought *what*? **Complex NP island**

What did Elizabeth wonder whether Jack bought *what*? **Whether island**

What did Elizabeth worry if Jack bought *what*? **Adjunct island**

High-SES adults **judge** these dependencies to be far worse than many others, including others that are very similar except that they don’t cross syntactic islands (Sprouse et al. 2012).

So, these judgments can serve as a target for successful acquisition — an outcome we can measure.
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. 2012: magnitude estimation judgments
• factorial definition controlling for two salient properties of island-crossing dependencies

length of dependency
(matrix vs. embedded)

Who

[CP... __who]?

Who

[non-island]

presence of an island structure
(non-island vs. island)

Who

[non-island]

Who

[island]?
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

length of dependency (matrix vs. embedded) \times presence of an island structure (non-island vs. island)

Complex NP island stimuli

Who __ claimed [that Lily forgot the necklace]? matrix | non-island

What did the teacher claim [that Lily forgot ___]? embedded | non-island

Who __ made [the claim that Lily forgot the necklace]? matrix | island

*What did the teacher make [the claim that Lily forgot ___ ]? embedded | island

Sprouse et al. 2012
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

length of dependency \( \times \) presence of an island structure

(matrix vs. embedded) \( \times \) (non-island vs. island)

Subject island stimuli

<table>
<thead>
<tr>
<th>Who __ thinks [the necklace is expensive]?</th>
<th>matrix</th>
<th>non-island</th>
</tr>
</thead>
<tbody>
<tr>
<td>What does Jack think [__ is expensive]?:</td>
<td>embedded</td>
<td>non-island</td>
</tr>
<tr>
<td>Who __ thinks [the necklace for Lily] is expensive?</td>
<td>matrix</td>
<td>island</td>
</tr>
<tr>
<td>*Who does Jack think [the necklace for ___] is expensive?</td>
<td>embedded</td>
<td>island</td>
</tr>
</tbody>
</table>

Sprouse et al. 2012
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

- length of dependency
- presence of an island structure

(matrix vs. embedded) \( \times \) (non-island vs. island)

Whether island stimuli

Who __ thinks [that Jack stole the necklace]? 
What does the teacher think [that Jack stole __]? 
Who __ wonders [whether Jack stole the necklace]? 
*What does the teacher wonder [whether Jack stole __]?

Sprouse et al. 2012
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

length of dependency (matrix vs. embedded) \( \times \) presence of an island structure (non-island vs. island)

Adjunct island stimuli

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Matrix</th>
<th>Non-island</th>
<th>Embedded</th>
<th>Non-island</th>
<th>Matrix</th>
<th>Island</th>
<th>Embedded</th>
<th>Island</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who __ thinks [that Lily forgot the necklace]?</td>
<td>matrix</td>
<td>non-island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What does the teacher think [that Lily forgot __ ]?</td>
<td>embedded</td>
<td>non-island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who __ worries [if Lily forgot the necklace]?</td>
<td>matrix</td>
<td>island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*What does the teacher worry [if Lily forgot __ ]?</td>
<td>embedded</td>
<td>island</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sprouse et al. 2012
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

length of dependency
(matrix vs. embedded) \( \times \) presence of an island structure
(non-island vs. island)

Syntactic island = superadditive interaction of the two factors
(additional unacceptability that arises when the two factors — length & presence of an island structure — are combined, above and beyond the independent contribution of each factor).

Sprouse et al. 2012
Syntactic island constraints

High-SES adult judgments = behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

- length of dependency (matrix vs. embedded) × presence of an island structure (non-island vs. island)

Syntactic island = superadditive interaction of the two factors

Who does...?

Who [non-island]?

Who [island]?

Additional unacceptability

Who [CP... who]?
Syntactic island constraints
High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior
length of dependency (matrix vs. embedded) X presence of an island structure (non-island vs. island)

Syntactic island = superadditive interaction of the two factors

Sprouse et al. 2012
Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

- length of dependency (matrix vs. embedded) × presence of an island structure (non-island vs. island)

Syntactic island = superadditive interaction of the two factors

Sprouse et al. (2012): acceptability judgments from 173 adult subjects

Superadditivity for all four island types
Syntactic island constraints
High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior
length of dependency (matrix vs. embedded) \(\times\) presence of an island structure (non-island vs. island)

Syntactic island = **superadditive** interaction of the two factors

Sprouse et al. (2012): acceptability judgments from 173 adult subjects

Superadditivity for all four island types = knowledge that dependencies can’t cross these island structures.
Okay, so what’s the relevant input for learning this target knowledge?
That depends on **how** we think children learn it.
That depends on how we think children learn it.

Pearl & Sprouse 2013 intuition:
- Learn what you can from the dependencies you do actually observe in the input
- Apply it to make a judgment about the dependencies you haven’t seen before, like syntactic islands.
A concrete learning strategy (Pearl & Sprouse 2013): View \textit{wh}-dependencies in terms of their \textbf{building blocks} and \textbf{track} those building blocks in the input.
A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

What phrases contain the gap (but not the wh-word)?

Pearl & Sprouse 2013
A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

What phrases contain the gap (but not the wh-word)?

What did you see __?
= What did [IP you [VP see __]]?
= start-IP-VP-end
A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

What phrases contain the gap (but not the *wh*-word)?

- What did you see ___?
  - What did [IP you [VP see ___]]?
  - *start-IP-VP-end*

- What ___ happened?
  - What [IP ___ happened]?
  - *start-IP-end*
A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

What phrases contain the gap (but not the wh-word)?

What did you see ___?
= What did [\textit{IP} you [\textit{VP} see ___]]?  
= start-IP-VP-end

What ___ happened?
= What [\textit{IP} ___ happened]?  
= start-IP-end

What did she want to do ___?
= What did [\textit{IP} she [\textit{VP} want [\textit{IP} to [\textit{VP} do ___]]]]?  
= start-IP-VP-IP-VP-end
A strategy for learning syntactic islands

What did you see ___?
= What did [IP you [VP see ___]]?
= start-IP-VP-end

What ___ happened?
= What [IP ___ happened]?
= start-IP-end

What did she want to do ___?
= What did [IP she [VP want [IP to [VP do ___]]]]?
= start-IP-VP-IP-VP-end

Ungrammatical dependencies have low probability segments

[CP Who did [IP Lily [VP think [CP [IP [NP the kitty [PP for ___]] was pretty ?]]]]]

start-IP-VP-CP-IP-NP-PP-end

Pearl & Sprouse 2013
A strategy for learning syntactic islands

So if children break these dependencies into smaller building blocks, they can identify if a dependency has a bad segment (made up of one or more low probability building blocks).

Pearl & Sprouse 2013
A strategy for learning syntactic islands

The building blocks: trigrams of container nodes

start-IP-VP-end
start-IP-VP-CP-IP-NP-PP-end
start-IP-VP-IP-VP-end

syntactic trigrams
A strategy for learning syntactic islands

The building blocks: **trigrams of container nodes**

- start-IP-VP-end
- start-IP-VP-CP-IP-NP-PP-end
- start-IP-end
- start-IP-VP-IP-VP-end
- syntactic trigrams

Pearl & Sprouse 2013
A strategy for learning syntactic islands

The building blocks: trigrams of container nodes

start-IP-VP-end      start-IP-VP-CP-IP-NP-PP-end
start-IP-end

start-IP-VP-VP-end
start-IP-VP
IP-VP-IP
VP-IP-VP
IP-VP-end

Pearl & Sprouse 2013
A strategy for learning syntactic islands

The building blocks: trigrams of container nodes

start-IP-VP
IP-VP-IP
VP-IP-VP
IP-VP-end

start-IP-VP-end
start-IP-VP-IP-VP-end
start-IP-VP-CP-IP-NP-PP-end

syntactic trigrams

start-IP-end
start-IP-end

Pearl & Sprouse 2013
A strategy for learning syntactic islands

The building blocks: trigrams of container nodes

- start-IP-VP-end
- start-IP-end
- start-IP-VP-IP-VP-end
- start-IP-VP-CP-IP-NP-PP-end
- start-IP-VP
- IP-VP-CP
- VP-CP-IP
- CP-IP-NP
- IP-NP-PP
- NP-PP-end

VP-CP-IP
CP-IP-NP
NP-PP-

start-IP-VP-end
start-IP-end
start-IP-VP-IP-VP-end

syntactic trigrams

Pearl & Sprouse 2013
A strategy for learning syntactic islands

The strategy: Track the relative frequency of the syntactic trigrams in your input

Pearl & Sprouse 2013
A strategy for learning syntactic islands

Who does…

Some of them are common and some of them aren’t.

start-IP-VP-end
start-IP-VP-CP-IP-NP-PP-end
start-IP-end
start-IP-VP-IP-VP-end

start-IP-VP
start-IP-end
IP-VP-end

IP-VP-CP
IP-VP-IP
VP-PP-end

IP-VP-PP
VP-CP-IP
VP-IP-VP
NP-PP-end

Pearl & Sprouse 2013
A strategy for learning syntactic islands

Some of them are common and some of them aren’t.

(And some never occur at all.)
A strategy for learning syntactic islands

Relative syntactic trigram frequency:

\[ p(t) \approx \frac{\# \text{ trigram}}{\text{total} \# \text{ trigrams}} \]
A strategy for learning syntactic islands

Any *wh*-dependency can then be constructed from its syntactic trigram building blocks.
A strategy for learning syntactic islands

\[ \prod_{t \in \text{trigrams}} p(t) \]

Who does...

Pearl & Sprouse 2013
A strategy for learning syntactic islands

Who does…

\[ \prod_{t \in \text{trigrams}} p(t) \]

Pearl & Sprouse 2013
A strategy for learning syntactic islands

$\prod_{t \in \text{trigrams}} p(t)$
A strategy for learning syntactic islands

\[ \prod_{t \in \text{trigrams}} p(t) \]

A wh-dependency’s probability can stand in for its judged acceptability.
A strategy for learning syntactic islands

\[ \prod_{t \in \text{trigrams}} p(t) \]

Lower probability dependencies are dispreferred, compared to higher probability dependencies.
A strategy for learning syntactic islands

Each set of island stimuli from Sprouse et al. 2012…

Complex NP island stimuli

Who __ claimed [that Lily forgot the necklace]?  
What did the teacher claim [that Lily forgot __]?  
Who __ made [the claim that Lily forgot the necklace]?  
*What did the teacher make [the claim that Lily forgot __]?  

\[ \Pi_{t \in \text{trigrams}} p(t) \]
A strategy for learning syntactic islands

Each *wh*-dependency from the island stimuli of Sprouse et al. 2012
- can be transformed into container node sequences

<table>
<thead>
<tr>
<th>Complex NP island stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>start-IP-end</td>
</tr>
<tr>
<td>start-IP-VP-CP_that-IP-VP-end</td>
</tr>
<tr>
<td>start-IP-end</td>
</tr>
<tr>
<td>start-IP-VP-NP-CP_that-IP-VP-end</td>
</tr>
</tbody>
</table>

\[ \prod_{t \in \text{trigrams}} p(t) \]
A strategy for learning syntactic islands

Each *wh*-dependency from the island stimuli of Sprouse et al. 2012
• can be transformed into container node sequences
• can be broken into **syntactic trigram building blocks** and have its probability calculated

**Complex NP island stimuli**

- `start-IP-end` | matrix | non-island
- `start-IP-VP-CP_{that}IP-VP-end` | embedded | non-island
- `start-IP-end` | matrix | island
- `start-IP-VP-NP-CP_{that}IP-VP-end` | embedded | island

\[ \prod_{t \in \text{trigrams}} p(t) \]
A strategy for learning syntactic islands

These probabilities can then be plotted to see if superadditivity is present.

**Complex NP island stimuli**

- $\text{start-IP-end}$
- $\text{start-IP-VP-CP_{that}-IP-VP-end}$
- $\text{start-IP-end}$
- $\text{start-IP-VP-NP-CP_{that}-IP-VP-end}$
A strategy for learning syntactic islands

If so, then the child would have syntactic island knowledge that allows the same judgment pattern as adults, learned from the building blocks in children's input.

Complex NP island stimuli

- start-IP-end
- start-IP-VP-CP\textsubscript{that}IP-VP-end
- start-IP-end
- start-IP-VP-NP-CP\textsubscript{that}IP-VP-end

\[ \prod_{t \in \text{trigrams}} p(t) \]
This strategy works for high-SES children’s input

Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data.
This strategy works for high-SES children’s input. Judges from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data.

Superadditivity for all four islands.
This strategy works for high-SES children’s input

Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data.

Implication: High-SES child input can support the acquisition of syntactic islands, using this learning strategy that depends on a certain part of the input.
This strategy works for high-SES children’s input

Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data.

That input part is the *wh*-dependencies, and their building blocks (the syntactic trigrams).
Who does…

\[ \prod_{t \in \text{trigrams}} p(t) \]

Are there meaningful differences across SES in this part of the input (the \textit{wh}-dependencies and syntactic trigrams)?

\textit{Bates & Pearl 2019, in prep.}
Are there meaningful differences across SES in this part of the input (the \emph{wh}-dependencies and syntactic trigrams)?

Let’s use developmental modeling to find out.
But first… how different does this input look across SES?

Let’s look at the distribution of the relevant parts: the *wh*-dependencies and the syntactic trigrams.
One way to measure differences in distribution:
the **Jensen-Shannon divergence (JSDiv)** (Endres & Schindelin 2003).

$$0 \leq \text{JSDiv} \leq 1$$

- identical distributions
  
- dissimilar distributions
Who does…

\[ \prod_{t \in \text{trigrams}} p(t) \]

\[ I \neq \neq \]

\[ 0 \leq \text{JSDiv} \leq 1 \]

Measurable input differences

The input samples

High-SES child-directed

102K utterances (21K wh-dependencies) from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 high-SES children between the ages of 1 and 5 years old.

Bates & Pearl 2019, in prep.
Who does…

\[ \prod_{t \in \text{trigrams}} p(t) \]

\[ 0 \leq JSDiv \leq 1 \]

Measurable input differences

31.8K utterances (3.9K wh-dependencies) from a subpart of the HSLLD corpus (Dickinson & Tabors 2001) in the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 78 low-SES children between the ages of 3 and 5.
Measurable input differences

\[ \Pi_{t \in \text{trigrams}} p(t) \]

\[ = \not\quad \neq \quad 0 \leq JSDiv \leq 1 \]

High-SES child-directed
21K \textit{wh}-dependencies

Low-SES child-directed
3.9K \textit{wh}-dependencies

Note: SES was defined by the creators of the HSLLD corpus according to \textit{maternal education} (6 years to some post-high school education) and \textit{annual income} (70\% reported < $20K/year).
Who does…

\[ \prod_{t \in \text{trigrams}} p(t) \]

\[
0 \leq JSDiv \leq 1
\]

Measurable input differences

74.6K utterances (8.5K wh-dependencies) from the Switchboard corpus (Marcus et al. 1999) of adults speaking to each other over the phone.

The input samples

- High-SES adult-directed
- Low-SES child-directed
- High-SES child-directed
- 21K wh-dependencies
- 3.9K wh-dependencies

Bates & Pearl 2019, in prep.
Measurable input differences

\[ \Pi_{t \in \text{trigrams}} p(t) \]

0 \leq JSDiv \leq 1

So what do we find?

In particular, is high-SES child-directed speech more like low-SES child-directed speech or more like high-SES adult-directed speech?

High-SES adult-directed

High-SES child-directed

Low-SES child-directed

8.5K wh-dependencies

21K wh-dependencies

3.9K wh-dependencies

Bates & Pearl 2019, in prep.
Measurable input differences

\[ \prod_{t \in \text{trigrams}} p(t) = 0 \leq JSDiv \leq 1 \]

If high-SES child-directed speech is more like low-SES child-directed speech, then SES differences matter less than who the speech is directed at.
Measurable input differences

\[ \prod_{t \in \text{trigrams}} p(t) = \neq \]

\[ 0 \leq \text{JSDiv} \leq 1 \]

If high-SES child-directed speech is more like high-SES adult-directed speech, then SES differences matter more than who the speech is directed at.
Whether we look at \emph{wh}-dependencies or syntactic trigrams, we find the same pattern: high-SES and low-SES child-directed speech are more similar than high-SES child-directed and high-SES adult-directed speech.

\[ \prod_{t \in \text{trigrams}} p(t) \]

\[ 0 \leq JSDiv \leq 1 \]
For *wh*-dependencies, high-SES child-directed speech is twice as similar to low-SES child-directed speech as it is to high-SES adult-directed speech.
\prod_{t \in \text{trigrams}} p(t) = 0 \leq \text{JSDiv} \leq 1

For syntactic trigrams, high-SES child-directed speech is twice as similar to low-SES child-directed speech as it is to high-SES adult-directed speech.

- High-SES adult-directed: 8.5K wh-dependencies
- High-SES child-directed: 21K wh-dependencies
- Low-SES child-directed: 3.9K wh-dependencies

\[ \text{JSDiv} = 0.01825 \]

\[ \text{JSDiv} = 0.00850 \]
Measurable input differences

\[ \prod_{t \in \text{trigrams}} p(t) = 0 \leq \text{JSDiv} \leq 1 \]

Takeaway: This part of the input looks pretty similar across SES — more similar than child-directed vs. adult-directed speech within SES.
But does this part of the input act differently? That is, are any differences (even if they’re small) meaningful?

They might be — small differences in the input distribution might snowball into **learning outcome differences**.
Meaningful input differences

But does this part of the input act differently? That is, are any differences (even if they’re small) meaningful?

They might be — small differences in the input distribution might snowball into learning outcome differences.

\[
\Pi_{t \in \text{trigrams}} p(t)
\]

<table>
<thead>
<tr>
<th>76.7%</th>
<th>start-IP-VP-end</th>
<th>75.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>What did Lily read ___what?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10.3%</th>
<th>start-IP-end</th>
<th>12.8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>What ___what happened?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
But does this part of the input act differently? That is, are any differences (even if they’re small) meaningful?

They might be — small differences in the input distribution might snowball into learning outcome differences.
Meaningful input differences

Let's use developmental computational modeling to find out.
Meaningful input differences

Judgments from a modeled child learning from the same amount of data as low-SES children seem to, with those data having the same composition as low-SES child-directed speech data.

Looking for superadditivity as the sign of syntactic islands knowledge
Meaningful input differences

Judgments from a modeled child learning from the same amount of data as low-SES children seem to, with those data having the same composition as low-SES child-directed speech data.

Superadditivity for all four islands!
Judgments from a modeled child learning from the same amount of data as low-SES children seem to, with those data having the same composition as low-SES child-directed speech data.

This means low-SES input is predicted to support the same learning outcome knowledge (of these four syntactic islands).
No meaningful input differences predicted

So, our developmental computational model predicts no meaningful input differences across SES when it comes to learning this syntactic island knowledge from this part of the input.
No meaningful input differences predicted

Useful: Because we know how the input is predicted to cause the knowledge to develop, we know which building blocks are particularly important.

$\prod_{t \in \text{trigrams}} p(t)$
No meaningful input differences predicted

Key building blocks for success involve complementizer *that* (CP\text{that}) - this is because two of the islands (whether and adjunct) only differ from grammatical dependencies by the complementizer used.

- What does the teacher think *that* Lily forgot __?  
  - embedded | non-island

- *What does the teacher wonder* whether Lily forgot __?  
  - embedded | island

- *What does the teacher worry* if Lily forgot __?  
  - embedded | island

*Bates & Pearl 2019, in prep.*
No meaningful input differences predicted

Key building blocks for success involve complementizer *that* (CP_{that}) - this is because two of the islands (whether and adjunct) only differ from grammatical dependencies by the complementizer used.

\[ \prod_{t \in \text{trigrams}} p(t) \]

- **start-IP-VP-CP_{\text{that}}**  \leftrightarrow  IP-VP-end

  - embedded  |  non-island

- * **start-IP-VP-CP_{\text{whether}}**  \leftrightarrow  IP-VP-end

  - embedded  |  island

- * **start-IP-VP-CP_{\text{adjunct}}**  \leftrightarrow  IP-VP-end

  - embedded  |  island

*Bates & Pearl 2019, in prep.*
No meaningful input differences predicted

So, children need to encounter grammatical *wh*-dependencies that involve *CP_that*. These are actually *pretty rare* in child-directed speech.

**Low-SES**
- child-directed
  - 2 instances of 3.9K (=.05%)

**High-SES**
- child-directed
  - 2 instances of 21K (<.01%)

What do you think that ___what happens?  
What do you think that Jack read ___what?

---

Bates & Pearl 2019, in prep.
No meaningful input differences predicted

\[ \prod_{t \in \text{trigrams}} p(t) \]

But with enough input (over several years), even these rare cases are predicted to support learning.

Low-SES child-directed
- 2 instances of 3.9K (=.05%)

High-SES child-directed
- 2 instances of 21K (<.01%)

What do you think that __what happens?

What do you think that Jack read __what?

Bates & Pearl 2019, in prep.
No meaningful input differences predicted

\[ \prod_{t \in \text{trigrams}} p(t) \]

And in fact, if the samples are reasonably accurate, low-SES children actually see this building block more often.

Low-SES child-directed

2 instances of 3.9K (=.05%)

What do you think that __what happens?

High-SES child-directed

2 instances of 21K (<.01%)

What do you think that Jack read __what?

Bates & Pearl 2019, in prep.
No meaningful input differences predicted

\[ \prod_{t \in \text{trigrams}} p(t) \]

Interesting: The \textit{wh}-dependency with this building block is typically judged to be ungrammatical in the high-SES dialect (a \textit{that}-trace violation).

Low-SES child-directed

2 instances of 3.9K (=.05%)

What do you think that \underline{what} happens?

High-SES child-directed

2 instances of 21K (<.01%)

What do you think that Jack read \underline{what}?
No meaningful input differences predicted

\[ \prod_{t \in \text{trigrams}} p(t) \]

Upshot: Low-SES children are predicted to achieve the same learning outcome as high-SES children by leveraging crucial building blocks from sources a high-SES child wouldn’t hear (because they’re ungrammatical for high-SES speakers).

Low-SES child-directed

High-SES child-directed

What do you think that what happens?
No meaningful input differences predicted

Takeaway: This is one reason why differences in the input might not be meaningful differences. The building blocks may show up in different places, but they’re still present in the input.

What do you think that __what happens?

Bates & Pearl 2019, in prep.
Who does $\prod_{t \in \text{trigrams}} p(t)$

So now what?

No meaningful input differences predicted

Bates & Pearl 2019, in prep.
We should measure the learning outcomes in children across SES to see if in fact there are any learning outcome differences.
One caveat: If there are in fact differences, it could be due to other factors besides input differences.

Example factor: Language processing ability is known to differ across SES, with low-SES children sometimes slower compared to their high-SES counterparts (Fernald et al. 2013, Weisleder & Fernald 2013). If low-SES children are less able to harness the information in their input (even if it’s there), they might be delayed in acquiring syntactic island knowledge.
But, if there aren’t outcome differences (perhaps after any language processing ability differences have resolved), then this supports syntactic island input quality being the same across SES.
Building block origins

\[ \prod_{t \in \text{trigrams}} p(t) \]

Low-SES child-directed

What do you think that \_what happens?

Remember that key building blocks involving CP\_ that are predicted to come from a particular wh-dependency in low-SES child-directed speech that’s ungrammatical in the high-SES dialect.

Bates & Pearl 2019, in prep.
Building block origins

Who does…

\[ \prod_{t \in \text{trigrams}} p(t) \]

Low-SES child-directed

What do you think that \_\_ what happens?

This means low-SES adults are predicted to view this *wh*-dependency as grammatical if we expect low-SES children to hear it and harness those crucial CP that building blocks from it.

We can test this.
The big picture

One (standard) way

Developmental computational modeling complements existing techniques for assessing developmentally-meaningful input differences.

A new (complementary) way
The big picture

One (standard) way

≠

✓

Developmental computational modeling complements existing techniques for assessing developmentally-meaningful input differences.

A new (complementary) way

≠

✓

Who does...

We demonstrated this for syntactic island knowledge, and predicted no meaningful input differences across SES.
The big picture

One (standard) way

≠

A new (complementary) way

Developmental computational modeling complements existing techniques for assessing developmentally-meaningful input differences.

This means we predict that no input-based interventions would be impactful if there actually are any differences in the acquisition of these syntactic islands across SES.
Something useful: This technique can provide a causal explanation for how input differences could affect learning outcomes.
Something useful: This technique can provide a causal explanation for how input differences could affect learning outcomes.

For syntactic islands, the building blocks needed for this knowledge don’t seem to differ enough to matter.
The big picture

One (standard) way

Something else useful: This technique can make predictions about differences we expect in both child outcomes and eventual adult knowledge.

A new (complementary) way

What do you think that __ what happens?
Something important: Any predicted differences still need to be measured. But at least we know what to look for.
The big picture

One (standard) way

Bonus: Modeling is often faster (and cheaper to do) than behavioral work. So it can be very useful as a first-pass input-quality assessment.

A new (complementary) way

Extra bonus: Possible to do in pandemic times.

What do you think that __what happens?
So let's use developmental computational modeling when we want to identify and understand developmentally-meaningful input variation!
Thank you!

Alandi Bates

BUCLD 2018
UCI Institute for Mathematical Behavioral Sciences 2019
UCSD Linguistics 2020
UCI QuantLang Collective

Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lpearl@uci.edu