Adventures in computational modeling for syntactic acquisition: A look at syntactic islands

Lisa Pearl
University of California, Irvine

November 4, 2022
Linguistics Colloquium
University of Arizona
What’s a computational model of acquisition?
What’s a computational model of acquisition?

“the embodiment of a specific theory about acquisition”
- Pearl in press
“…you need to first have a theory about how acquisition works.” — Pearl in press
“Then, the model can be used to

(1) make all the components of that acquisition theory explicit…”  - Pearl in press
“Then, the model can be used to

(1) make all the components of that acquisition theory explicit,

(2) evaluate whether it actually works…”

- Pearl in press

How?
“Then, the model can be used to
(1) make all the components of that acquisition theory explicit,
(2) evaluate whether it actually works…”
- Pearl in press

The model generates predictions
How?
“Then, the model can be used to

(1) make all the components of that acquisition theory explicit,

(2) evaluate whether it actually works…”

- Pearl in press
“Then, the model can be used to
(1) make all the components of that acquisition theory explicit,
(2) evaluate whether it actually works…”

- Pearl in press
“Then, the model can be used to

1. make all the components of that acquisition theory explicit,
2. evaluate whether it actually works, and
3. determine precisely what makes it work (or not work).”

- Pearl in press
“Then, the model can be used to

1. **make all the components** of that acquisition theory **explicit**,
2. **evaluate whether it actually works**, and
3. **determine precisely what makes it work (or not work)**.

- Pearl in press
“Then, the model can be used to
(1) make all the components of that acquisition theory explicit,
(2) evaluate whether it actually works, and
(3) determine precisely what makes it work (or not work).”
- Pearl in press

…but an informative model ideally allows us to understand what’s going on inside a child’s mind.
We can use quantitative techniques to implement a computational model.
theory of acquisition

computational model

quantitative

One main part: Counting things
One main part: **Counting** things

(sometimes we count a lot of things)
theory of acquisition

computational model

quantitative

Bayesian inference

Another part: principled reasoning based on those counts

\[ p(\text{Hypothesis} | \text{Data}) \propto p(\text{Hypothesis}) \cdot p(\text{Data} | \text{Hypothesis}) \]
We think the child is learning by counting different parts of her input and reasoning over those counts in a sensible way.
We think the child is learning by counting different parts of her input and reasoning over those counts in a sensible way.

So, the model will count those same things and learn about language by doing principled reasoning over those counts.
This counting and reasoning is what the theory of acquisition is about.
In particular: **counting what?** And **reasoning how?**
We build the model according to what the theory specifies in order to evaluate it. Modeling results can help us refine our theory.
Today: A case study with the acquisition of complex syntactic knowledge known as syntactic islands.

Who does...
Today: A case study with the acquisition of complex syntactic knowledge known as syntactic islands.

Part 1:
About syntactic islands and their acquisition

Who does…
quantitative

Part 2: Evaluating a theory of acquisition for English that assumes certain building blocks
quantitative

Who does...

1: syntactic islands acquisition

2: Evaluating a theory

Part 3: Evaluating this theory when there's dialectal variation

Part 3: Evaluating this theory

theory of acquisition

computational model
quantitative

theory of acquisition

computational model

1: syntactic islands acquisition

Who does…

2: Evaluating a theory

3: dialectal variation

Part 4: Evaluating a variant of this theory where the child learns what the building blocks are first
Syntactic islands 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 islands involve *wh*-dependencies.

What’s going on here?

There’s a **dependency** between the *wh*-word *who* and where it’s understood (**the gap**)
Syntactic islands 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 islands involve *wh*-dependencies.

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

(Subject island)
Syntactic islands involve *wh*-dependencies.

Who does Lily think the kitty for whom is pretty? [Subject island]

Jack is somewhat tricksy.

He claimed he bought something.

*What did Jack make the claim that he bought what?*
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 islands 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**

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

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

**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 islands 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.

What did Elizabeth think ___ what?
Syntactic islands involve *wh*-dependencies.

**Who does Lily think the kitty for** [Subject island]

**What did Jack make the claim that he bought** [Complex NP island]

**What did Elizabeth wonder whether Jack bought** [Whether island]

**What did Elizabeth worry if Jack bought** [Adjunct island]

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

**What did Elizabeth think Jack said**
Syntactic islands 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.

```
What did Elizabeth think Jack said Lily saw \_\_what?
```
Syntactic islands 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**

English 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).
Syntactic islands involve *wh*-dependencies.

*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

English-learning children strongly *disprefer* one of these dependencies compared to others (*de Villiers* et al. 2008).
Syntactic islands involve *wh*-dependencies. 

**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]

These judgments and (dis)preferences are a measurable observable behavior that can signal the successful acquisition of syntactic island knowledge.
Syntactic islands 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?*
Syntactic islands

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

Who [CP... __who]?

Who [non-island]

Who [island]

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

Who __who?

Who [non-island]?

Who [island]?
**Syntactic islands**

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)

<table>
<thead>
<tr>
<th>Subject island stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who ___ thinks [the necklace is expensive]?</td>
</tr>
<tr>
<td>What does Jack think [ ___ is expensive]?</td>
</tr>
<tr>
<td>Who ___ thinks [the necklace for Lily] is expensive?</td>
</tr>
<tr>
<td>*Who does Jack think [the necklace for ___ ] is expensive?</td>
</tr>
</tbody>
</table>

Sprouse et al. 2012
Syntactic islands

Adult judgments  
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

Whether island stimuli

<table>
<thead>
<tr>
<th>Who __ thinks [that Jack stole the necklace]?</th>
<th>matrix</th>
<th>non-island</th>
</tr>
</thead>
<tbody>
<tr>
<td>What does the teacher think [that Jack stole __ ]?</td>
<td>embedded</td>
<td>non-island</td>
</tr>
<tr>
<td>Who __ wonders [whether Jack stole the necklace]?</td>
<td>matrix</td>
<td>island</td>
</tr>
<tr>
<td>*What does the teacher wonder [whether Jack stole __ ]?</td>
<td>embedded</td>
<td>island</td>
</tr>
</tbody>
</table>

Sprouse et al. 2012
Syntactic islands

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)

Adjunct island stimuli

Who ___ thinks [that Lily forgot the necklace]? matrix | non-island
What does the teacher think [that Lily forgot ___ ]? embedded | non-island
Who ___ worries [if Lily forgot the necklace]? matrix | island
*What does the teacher worry [if Lily forgot ___ ]? embedded | island

Sprouse et al. 2012
Syntactic islands

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)

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 islands

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. This is **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.
Syntactic islands

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

Who does...? Who does...?
Who [non-island] Who [island]?
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

Syntactic islands

Adult judgments

= behavioral target outcome
Syntactic islands

Adult knowledge as measured by acceptability judgment behavior

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

Syntactic island = \text{superadditive interaction of the two factors}

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

Superadditivity for all four island types

Sprouse et al. 2012
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

= knowledge that dependencies crossing these island structures are dispreferred.
Child knowledge as measured by preferred interpretation behavior

De Villiers et al. 2008:
How do children prefer to interpret potentially ambiguous wh-questions?
Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous *wh*-questions?

What did the boy fix the cat that was lying on the table with ___what?
Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous \textit{wh}-questions?

\textbf{What} did the boy \textbf{fix the cat} \textit{that was lying on the table} \textbf{with} \underline{\textit{what}}?

\textit{a needle and thread}
Syntactic islands
Child judgments = behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous *wh*-questions?

What did the boy [fix [the cat [that [was [lying [on [the table [with ___what]]]]]]]]? a penguin
Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous wh-questions?

**What** did the boy fix the cat that was lying on the table with __**what**?

Children strongly prefer this interpretation.
Syntactic islands
Child judgments
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous wh-questions?

What did the boy fix the cat that was lying on the table with ___what?

…and strongly disprefer this interpretation

De Villiers et al. 2008
Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous *wh*-questions?

What did the boy [fix [the cat [that [was [lying [on [the table [with ___what]]]]]]]]

This means they strongly disprefer the *wh*-dependency this interpretation relies on.

De Villiers et al. 2008
Child knowledge as measured by preferred interpretation behavior

How do children prefer to interpret potentially ambiguous *wh*-questions?

What did the boy [fix \[NP the cat [that [was [lying [on [the table [with ___what]]]]]]]]? 

…which is a dependency that crosses a Complex NP.
Syntactic islands
Adult & child judgments
= behavioral target outcome

Subject island
Complex NP island
Whether island
Adjunct island

z-score rating

matrix
eMBEDDED

Sprouse et al. 2012

De Villiers et al. 2008
Syntactic islands

How long do children have to learn?
Syntactic islands

How long do children have to learn?

De Villiers et al. 2008: Data from four-year-olds.

What [NP [CP __what]]?
Syntactic islands

How long do children have to learn?

So input through age four.
(<60 months)
What input do children get?

Syntactic islands
Syntactic islands

What input do children get?

We can estimate this from samples of child-directed speech.
Syntactic islands

This is the acquisition problem
Syntactic islands

...which is where the theory of acquisition comes in.
quantitative

computational model

2: Evaluating a theory

Who does…

1: syntactic islands
acquisition

3: dialectal variation

4: learning the building blocks

theory of acquisition
Syntactic islands: A theory of acquisition

Intuition:
• Learn what you can from the *wh*-dependencies you observe in the input over time
Syntactic islands:
A theory of acquisition

Intuition:
• Learn what you can from the *wh*-dependencies you observe in the input over time

• Apply it to generate behavior for *wh*-dependencies you haven’t seen before, like those crossing syntactic islands (and maybe other longer *wh*-dependencies, too).

Pearl & Sprouse 2013
Syntactic islands: A theory of acquisition

View *wh*-dependencies in terms of their **building blocks** and **track** those building blocks in the input.
Syntactic islands: A theory of acquisition

Dependencies represented as a sequence of container nodes

What phrases contain the gap (but not the wh-word)?
Syntactic islands: A theory of acquisition

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
Syntactic islands: A theory of acquisition

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
Syntactic islands: A theory of acquisition

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

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

Pearl & Sprouse 2013
Syntactic islands: A theory of acquisition

What ___ happened?
= What [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

(Much) less acceptable dependencies have low probability segments

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

start-IP-VP-CP-that-IP-NP-PP-end
Syntactic islands:
A theory of acquisition

What did you see __?  
= What did \[IP \text{ you} [VP \text{ see } __]\]?  
= \(start-IP-VP-end\)

What __ happened?  
= What \[IP \text{ } __ \text{ happened}\]?  
= \(start-IP-end\)

What did she want to do __?  
= What did \[IP \text{ she} [VP \text{ want } [IP \text{ to} [VP \text{ do } __]]]\]?  
= \(start-IP-VP-IP-VP-end\)

\[[[CP \text{ Who} \text{ did} \text{ } [IP \text{ Lily} [VP \text{ think}[CP\text{-that} \text{ [IP [NP \text{ the kitty} [PP \text{ for } __] \text{ was pretty ?}]]]]]]]]

\(start-IP-VP-CP_{\text{that-IP-NP-PP-end}}\)

So if children break these dependencies into smaller building blocks, they can identify if a dependency has bad segments (made up of one or more low probability building blocks).
Syntactic islands: A theory of acquisition

The building blocks: trigrams of container nodes
Syntactic islands: A theory of acquisition

The building blocks: trigrams of container nodes

start-IP-VP-end
start-IP-CP_{that}-IP-NP-PP-end
start-IP-end
start-IP-VP-IP-VP-end
start-IP-VP
Syntactic islands: A theory of acquisition

The building blocks: trigrams of container nodes
Syntactic islands: A theory of acquisition

The building blocks: syntactic trigrams

trigrams of container nodes

start-IP-VP-end

start-IP-VP-CP_{that}-IP-NP-PP-end

start-IP-end

start-IP-VP-IP-VP-end

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-end
Syntactic islands: A theory of acquisition

The building blocks: trigrams of container nodes

- start-IP-VP-end
- start-IP-CP_{\text{that}}-IP-NP-PP-end
- start-IP-end
- start-IP-VP-IP-VP-end
- IP-VP-IP
- VP-IP-VP
- IP-VP-end
- start-IP-VP
- IP-VP-CP_{\text{that}}
- VP-CP_{\text{that}}-IP
- CP_{\text{that}}-IP-NP
- IP-NP-PP
- NP-PP-end

syntactic trigrams

Pearl & Sprouse 2013
Syntactic islands: A theory of acquisition

Learning: Track the relative frequency of the syntactic trigrams in the input

start-IP-VP-end  start-IP-VP-CP_that-IP-NP-PP-end
start-IP-end      start-IP-VP-IP-VP-end
start-IP-VP       start-IP-end
IP-VP-end         start-IP-end
IP-VP-IP          IP-VP-end
IP-VP-CP_that     VP-IP-VP
VP-CP_that-IP     start-IP-end
IP-VP-PP          VP-PP-end
NP-PP-end         Pearl & Sprouse 2013
Syntactic islands: A theory of acquisition

Some of them are common and some of them aren’t.
Some of them are common and some of them aren’t. (And some never occur at all.)
Syntactic islands: A theory of acquisition

Relative syntactic trigram frequency:

\[ p(t) \approx \frac{\# \text{trigram}}{\text{total} \# \text{trigrams}} \]
Syntactic islands: A theory of acquisition

Any wh-dependency can then be constructed from its syntactic trigram building blocks.
Syntactic islands: A theory of acquisition

\[
\prod_{t \in \text{trigrams}} p(t)
\]
Syntactic islands: A theory of acquisition

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

\text{start-IP-VP-end}

\text{start-IP-VP-CP}_{\text{that}}-\text{IP-NP-PP-end}

\text{start-IP-VP-IP-VP-end}

\text{start-IP-VP}

\text{IP-VP-IP}

\text{VP-IP-VP}

\text{IP-VP-end}

\text{IP-VP-CP}_{\text{that}}-\text{IP}

\text{IP-VP-PP}

\text{VP-PP-end}

\text{NP-PP-end}

\text{CP}_{\text{that}}-\text{IP-NP}

\text{IP-NP-PP}

Pearl & Sprouse 2013
Syntactic islands: A theory of acquisition

\[ \Pi_{t \in \text{trigrams}} p(t) \]
A wh-dependency’s probability can stand in for its predicted acceptability or preference.
Syntactic islands: A theory of acquisition

Lower probability dependencies are predicted to be less acceptable (dispreferred), compared to higher probability dependencies.
Evaluating the theory

If we learn from the input children get the way this theory specifies, can this theory output the behavior children (should) produce?
Evaluating the theory

What's the input look like?
Evaluating the theory

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

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

This lets us estimate which \textit{wh}-dependencies children hear and how often they hear them (the \textit{wh}-dependency distribution).
Evaluating the theory

We can then estimate how many wh-dependencies children hear during the learning period. (<60 months)

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

Children begin to represent the full structure of *wh*-dependencies (e.g., *wh*-questions and relative clauses) around **20 months**: Seidl et al. 2003, Gagliardi et al. 2016, Perkins & Lidz 2020.
Educated guess: This is when children can start processing *wh*-dependencies reliably from their input.

(20 months ≤ age < 60 months)
Evaluating the theory

(20 months ≤ age < 60 months)

How many minutes is this? In particular, children are awake for only a certain portion of the day at different ages (Davis et al. 2004).

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

<table>
<thead>
<tr>
<th>age</th>
<th>age range</th>
<th>waking</th>
<th>total waking hours</th>
<th>cumulative waking hrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>20-23 months</td>
<td>10</td>
<td>11 hrs/day * 365 days/yr * 4/12 = 1216.67</td>
<td>1216.67</td>
</tr>
<tr>
<td>two</td>
<td>24-35 months</td>
<td>11</td>
<td>11 hrs/day * 365 days/yr = 4015</td>
<td>5231.67</td>
</tr>
<tr>
<td>three</td>
<td>36-47 months</td>
<td>12</td>
<td>12 hrs/day * 365 days/yr = 4380</td>
<td>9611.67</td>
</tr>
<tr>
<td>four</td>
<td>48-59 months</td>
<td>12.5</td>
<td>12.5 hrs/day * 365 days/yr = 4562.5</td>
<td>14174.17</td>
</tr>
</tbody>
</table>

cumulative waking mins
14174.17 * 60 min/hour
850450.2

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

How many wh-dependencies is this? 

(≈850450 minutes)
Hoff-Ginsberg (1998) and Rowe (2012): Estimates of *utterances per minute* in speech directed at children from different backgrounds.
Evaluating the theory

Data from children from higher socio-economic status (SES) backgrounds.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

higher-SES
wh-dependency
distribution

from our own
corpus samples:
rate of wh-
dependencies/
utterance

<table>
<thead>
<tr>
<th>higher-SES</th>
<th>utt/min</th>
<th>* 850,450.2</th>
<th>* wh-dep/utt</th>
<th>total wh-dep</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 2 s.d.</td>
<td>7.4</td>
<td></td>
<td></td>
<td>= 1,293,545</td>
</tr>
<tr>
<td>- 1 s.d.</td>
<td>11.6</td>
<td></td>
<td></td>
<td>= 2,027,719</td>
</tr>
<tr>
<td>average</td>
<td>15.8</td>
<td></td>
<td></td>
<td>= 2,761,893</td>
</tr>
<tr>
<td>+ 1 s.d.</td>
<td>20.0</td>
<td></td>
<td></td>
<td>= 3,496,067</td>
</tr>
<tr>
<td>+ 2 s.d.</td>
<td>24.2</td>
<td></td>
<td></td>
<td>= 4,230,241</td>
</tr>
</tbody>
</table>

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

Can the modeled learner produce the appropriate observable behavior?

≈1.3 million - 4.2 million wh-dependencies

higher-SES wh-dependency distribution
Evaluating the theory

Reminder: Target behavior

- Subject island
- Complex NP island
- Whether island
- Adjunct island

\[
\begin{align*}
\text{z-score rating} & \quad \text{island effect} \\
\text{matrix} & \quad \text{embedded} \\
\text{non-island structure} & \quad \text{island structure}
\end{align*}
\]

Sprouse et al. 2012

De Villiers et al. 2008

What \[
[\text{NP} \quad \text{[CP \quad \_what]}?]
\]

Complex NP

What \[
\_what?_?
\]

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory

Reminder: **Target behavior**

Subject island
Complex NP island
Whether island
Adjunct island

Looking for **superadditivity** in selected judgments as the sign of syntactic islands knowledge

Who
Who
Who
Who

Who
Who

Who
Who

matrix
embedded
non-island structure
island structure
Evaluating the theory

Reminder: Target behavior

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 __]?
Evaluating the theory

Reminder: **Target behavior**

- **Subject island**
- **Complex NP island**
- **Whether island**
- **Adjunct island**

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

**Complex NP island stimuli**

<table>
<thead>
<tr>
<th>Structure</th>
<th>Type</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>start-IP-end</code></td>
<td>matrix</td>
<td>non-island</td>
</tr>
<tr>
<td><code>start-IP-VP-CP_{that}IP-VP-end</code></td>
<td>embedded</td>
<td>non-island</td>
</tr>
<tr>
<td><code>start-IP-end</code></td>
<td>matrix</td>
<td>island</td>
</tr>
<tr>
<td><code>start-IP-VP-NP-CP_{that}IP-VP-end</code></td>
<td>embedded</td>
<td>island</td>
</tr>
</tbody>
</table>

*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press*
Evaluating the theory

Reminder: **Target behavior**

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**

<table>
<thead>
<tr>
<th>start-IP-end</th>
<th>matrix</th>
<th>non-island</th>
</tr>
</thead>
<tbody>
<tr>
<td>start-IP-VP-CP<em>that</em>IP-VP-end</td>
<td>embedded</td>
<td>non-island</td>
</tr>
<tr>
<td>start-IP-end</td>
<td>matrix</td>
<td>island</td>
</tr>
<tr>
<td>start-IP-VP-NP-CP<em>that</em>IP-VP-end</td>
<td>embedded</td>
<td>island</td>
</tr>
</tbody>
</table>
Evaluating the theory

Reminder: Target behavior

- Subject island
- Complex NP island
- Whether island
- Adjunct island

These probabilities can then be plotted to see if superadditivity is present in the predicted acceptability judgments.

### 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

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

Reminder: **Target behavior**

Subject island
Complex NP island
Whether island
Adjunct island

If so, then we predict the modeled child has **syntactic island knowledge** that allows the **same judgment pattern** as adults, learned from the **building blocks** in children’s input.
Evaluating the theory

Complex NP island

Subject island

Whether island

Adjunct island
Evaluating the theory

Superadditivity predicted for judgments of all four island types.
Evaluating the theory

Subject island
Complex NP island
Whether island
Adjunct island

Reminder: Target behavior

Children prefer this interpretation.

What ___what?

What [NP [CP ___what]]?

Complex NP

De Villiers et al. 2008

Sprouse et al. 2012

Student

De Villiers et al. 2008

Sprouse et al. 2012

Complex NP island
Whether island
Adjunct island

z-score rating

matrix
embedded

island effect

non-island structure
island structure
Evaluating the theory

The *wh*-dependency this interpretation relies on is $10^{18}$ times more probable than the other one.
Evaluating the theory

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

So, the modeled child prefers it.

\[ [\text{NP [CP } \_\text{what}]]? \]
Takeaway:
This theory can work for learning knowledge about syntactic islands.
Key idea: Learning about the building blocks of *wh*-dependencies leads to knowledge about syntactic islands.
quantitative

computational model

3: dialectal variation

1: syntactic islands
acquisition

2: Evaluating a theory

4: learning the building blocks

theory of acquisition
Evaluating the theory across dialects

Our previous input sample came from speech directed at children from higher-SES backgrounds.
What about children from lower-SES backgrounds?

Evaluation the theory across dialects

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Lower-SES language input can differ from higher-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 ...
Input differences across SES

Quality can be measured by different aspects of the input, like diversity of vocabulary, diversity of syntactic constructions …
Input differences across SES

**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.

The kitty wasn’t there

We’ll see the kitty on Friday

Because we’re going tomorrow

We saw her yesterday, didn’t we?
Input differences across SES

What we don’t know: Do input differences impact the *wh*-dependency distribution in a way that could affect the acquisition of **syntactic island knowledge**?
Input differences across SES for this theory

What we don’t know: Do input differences impact the *wh*-dependency building block distribution in a way that could affect the acquisition of *syntactic island knowledge*?

Let’s find out!
Evaluating the theory across dialects

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 lower-SES children between the ages of 3 and 5.
Evaluating the theory across dialects

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

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory across dialects

Hoff-Ginsberg (1998) and Rowe (2012): Estimates of utterances per minute in speech directed at children from different backgrounds.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory across dialects

Data from children from lower socio-economic status (SES) backgrounds.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory across dialects

from our own corpus samples: rate of wh-dependencies/utterance

<table>
<thead>
<tr>
<th>lower-SES</th>
<th>utt/min</th>
<th>min</th>
<th>wh-dep/utt</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 2 s.d.</td>
<td>4.6</td>
<td>850,450.2</td>
<td>3,904/31,875</td>
</tr>
<tr>
<td>- 1 s.d.</td>
<td>8.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>13.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 1 s.d.</td>
<td>17.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ 2 s.d.</td>
<td>21.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory across dialects

Can the modeled child produce the appropriate observable behavior?

Who does…

≈ 479K- 2.2 million wh-dependencies

lower-SES wh-dependency distribution
Reminder: Target behavior

- Subject island
- Complex NP island
- Whether island
- Adjunct island

Evaluating the theory across dialects

Whether island

\[\text{What} \ [\text{NP } [\text{CP } \ldots]]?\]

\[\text{Complex NP}\]

higher-SES \ Sprouse et al. 2012
lower-SES \ Bates et al. in prep.

higher-SES \ De Villiers et al. 2008
lower-SES
Evaluating the theory across dialects

Complex NP island

Subject island

Whether island

Adjunct island

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory across dialects

Superadditivity predicted for judgments of all four island types.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press
Evaluating the theory across dialects

Reminder: Target behavior

Children prefer this interpretation.

**Subject island**

**Complex NP island**

**Whether island**

**Adjunct island**

**What [NP [CP _what]]?**

[NP [CP _what]]?

**Complex NP**

**What _what?**

**non-island structure**

**island structure**

**matrix**

**embedded**

- **non-island structure**
- **island structure**

**higher-SES Sprouse et al. 2012**

**lower-SES Bates et al. in prep.**

**higher-SES De Villiers et al. 2008**

**lower-SES**
Evaluating the theory across dialects

The *wh*-dependency this interpretation relies on is $10^{21}$ times more probable than the other one.

Evaluating the theory across dialects

**Subject island**

**Complex NP island**

**Whether island**

**Adjunct island**

higher-SES  Sprouse et al. 2012

lower-SES  Bates et al. in prep.

higher-SES  De Villiers et al. 2008

lower-SES
Evaluating the theory across dialects

Subject island
Complex NP island
Whether island
Adjunct island

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

So, the modeled child prefers it.

\begin{align*}
\text{higher-SES} & \quad \text{Sprouse et al. 2012} \\
\text{lower-SES} & \quad \text{Bates et al. in prep.}
\end{align*}

\begin{align*}
\text{higher-SES} & \quad \text{De Villiers et al. 2008} \\
\text{lower-SES}
\end{align*}
Takeaway:
This theory can work for learning knowledge about syntactic islands across dialects.
Key idea (again): Learning about the building blocks of *wh*-dependencies leads to knowledge about syntactic islands even when there’s input variation.
quantitative

computational model

1: syntactic islands acquisition

2: Evaluating a theory

3: dialectal variation

4: learning the building blocks

Who does…

theory of acquisition
The building blocks from Pearl & Sprouse (2013) were pre-specified. The modeled child already knew to look for syntactic trigrams of a certain kind.
Learning the building blocks

In particular:
(1) Look for groups of three units
(2) If the unit is a CP, include the *lexical item*

Dickson et al. 2022, in prep.
Learning the building blocks

Empirical motivation for the CP lexical item: Two of the islands (Whether and Adjunct) only differ from more acceptable \textit{wh}-dependencies by the \textit{complementizer} used.

What does the teacher think \textbf{[that Lily forgot …]}? \hspace{1cm} embedded | non-island

```
Whether
```
```
Adjunct
```

*What does the teacher wonder \textbf{[whether Lily forgot …]}? \hspace{1cm} embedded | island

```
Whether
```
```
```

*What does the teacher worry \textbf{[if Lily forgot …]}? \hspace{1cm} embedded | island

```
Whether
```
```
```

Dickson et al. 2022, in prep.
Learning the building blocks

Empirical motivation for the CP lexical item: Two of the islands (Whether and Adjunct) only differ from more acceptable \textit{wh}-dependencies by the \textit{complementizer} used.

\begin{itemize}
  \item \textit{Non-island:} \texttt{start-IP-VP-CP_{that-}IP-VP-end}
  \item \textit{Island:} \texttt{* start-IP-VP-CP_{whether-}IP-VP-end}
  \item \texttt{* start-IP-VP-CP_{if-}IP-VP-end}
\end{itemize}

\scriptsize{Dickson et al. 2022, \textit{in prep.}}
Learning the building blocks

So the building blocks need to include this lexical item type.

- start-IP-VP-CP\textsubscript{that}-IP-VP-end (embedded | non-island)
- * start-IP-VP-CP\textsubscript{whether}-IP-VP-end (embedded | island)
- * start-IP-VP-CP\textsubscript{if}-IP-VP-end (embedded | island)

Dickson et al. 2022, in prep.
Learning the building blocks

Is this the only one needed?

\[ \text{start-IP-VP-CP}_{\text{that}} \text{-IP-VP-end} \]

\[ \text{embedded|non-island} \]

\[ \text{start-IP-VP-CP}_{\text{whether}} \text{-IP-VP-end} \]

\[ \text{embedded|island} \]

\[ \text{start-IP-VP-CP}_{\text{If}} \text{-IP-VP-end} \]

\[ \text{embedded|island} \]

Dickson et al. 2022, in prep.
Liu et al 2019: Acceptability of *wh*-dependencies can depend on the *lexical item* in the main verb.

Dickson et al. 2022, *in prep.*
Liu et al 2019: Acceptability of *wh*-dependencies can depend on the **lexical item in the main verb**.
Learning the building blocks

Bigger question: Are there other lexical item types the building blocks need to include?

Dickson et al. 2022, in prep.
Learning the building blocks

Idea: What if the child is trying to learn what the best building blocks are at the same time she learns about their distributions in the input?
Learning the building blocks

the best building blocks

Before:
(1) Look for groups of three units
(2) If the unit is a CP, include the lexical item

- start-IP-VP
- IP-VP-end
- IP-VP-\text{CP}_{\text{that}}
- VP-\text{CP}_{\text{that}}-\text{IP}
Learning the building blocks

the best building blocks

(1) Look for groups of three units
(2) If the unit is a CP, include the lexical item

Maybe the best size is sometimes bigger than three and sometimes smaller.
Learning the building blocks

the best building blocks

(1) Look for groups of three units
(2) If the unit is a CP, include the lexical item
Learning the building blocks

the best building blocks

(1) Look for groups of three units
(2) If the unit is a CP, include the **lexical item**

```
start-IP-VP-CP_{that}
start-IP
IP-VP-end
CP_{that}-IP
```
Learning the building blocks

the best building blocks

(1) Look for the **best-sized units**
(2) If the unit is a CP, include the **lexical item**

Maybe the lexical item is needed sometimes…but sometimes **not**.

\[
\text{start-IP-VP}_{\text{think-CP}_{\text{that}}}
\]

\[
\text{start-IP}
\]

\[
\text{IP-VP-end}
\]

\[
\text{CP}_{\text{that-IP}}
\]

\[
\text{start-IP-VP}_{\text{say-CP}_{\text{that}}}
\]
Learning the building blocks

the best building blocks

(1) Look for the best-sized units
(2) Sometimes include the lexical item

Maybe the lexical item is needed sometimes…but sometimes not.
Learning the building blocks

How can the child learn what the best building blocks are?

(1) Look for the best-sized units
(2) Sometimes include the lexical item
Learning the building blocks

How can the child learn what the best building blocks are?

(1) Look for the best-sized units
(2) Sometimes include the lexical item

Theory: Look for an “efficient” set of building blocks.
Learning the building blocks

How can the child learn what the best building blocks are?

(1) Look for the best-sized units
(2) Sometimes include the lexical item

Efficient building blocks allow the representation of current and future *wh*-dependencies to be more probable.
Learning the building blocks

How can the child learn what the best building blocks are?

(1) Look for the best-sized units
(2) Sometimes include the lexical item

Efficient building blocks allow the representation of current and future *wh*-dependencies to be more probable.

Why? One idea: Higher probability *wh*-dependencies are faster to process (comprehending or producing).

Dickson et al. 2022, in prep.
Learning the building blocks

Learning efficient building blocks

How? Look for building blocks that are a balance between
(1) how big they are
(2) how fast they are to put together to make a wh-dependency

Dickson et al. 2022, in prep.
Learning the building blocks

learning **efficient building blocks**

a **balance** between
(1) **how big** they are
(2) **how fast** they are to put together to
make a *wh*-dependency

What did she say that he saw __ ?
Learning the building blocks

learning efficient building blocks

a balance between
- how big they are
- how fast they are to put together to make a *wh*-dependency

\[ \text{start-IP}_{\text{past-VP}} \text{say-CP that-IP}_{\text{past-VP}} \text{see-end} \]
Learning the building blocks

learning efficient building blocks

a balance between

1. how big they are
2. how fast they are to put together to make a \textit{wh}-dependency

\begin{align*}
\textit{start}-&\text{IP}_{\text{past}}-\text{VP}_{\text{say}}-\text{CP}_{\text{that}}-\text{IP}_{\text{past}}-\text{VP}_{\text{see}}-\text{end}
\end{align*}

Pieces can be small, so that many of them make up a \textit{wh}-dependency

Dickson et al. 2022, in prep.
Learning the building blocks

learning efficient building blocks

a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

\[ \text{start-IP}_{\text{past-VP}}\text{say-CP that-IP}_{\text{past-VP}}\text{see-end} \]

It may be slower to put together many small pieces.
Learning the building blocks

Learning efficient building blocks

a balance between
(1) how big they are
(2) how fast they are to put together to make a wh-dependency

\[ \text{start-IP_{\text{past}}-VP_{\text{say}}-CP_{\text{that}}-IP_{\text{past}}-VP_{\text{see}}-end} \]

many smaller

slower because many

But these pieces may get reused, so that makes them faster to put together.
Learning the building blocks

Learning efficient building blocks

a balance between
(1) how big they are
(2) how fast they are to put together to make a wh-dependency

\[ \text{start-IP}_{\text{past}}-\text{VP}_{\text{say}}-\text{CP}_{\text{that}}-\text{IP}_{\text{past}}-\text{VP}_{\text{see}}-\text{end} \]

many smaller

\[ \text{start} \rightarrow \text{IP}_{\text{past}} \rightarrow \text{VP}_{\text{say}} \rightarrow \text{CP}_{\text{that}} \rightarrow \text{IP}_{\text{past}} \rightarrow \text{VP}_{\text{see}} \rightarrow \text{end} \]

But these pieces may get reused, so that makes them faster to put together.
Learning the building blocks

Learning efficient building blocks

a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

\[ \text{start-IP}_{\text{past}} \text{-VP}_{\text{say}} \text{-CP}_{\text{that}} \text{-IP}_{\text{past}} \text{-VP}_{\text{see-\text{end}}} \]

Pieces can be \textbf{big}, so that only one makes up a *wh*-dependency

Dickson et al. 2022, in prep.
Learning the building blocks

Learning efficient building blocks

a balance between
(1) how big they are
(2) how fast they are to put together to make a wh-dependency

\[ \text{start-IP}_{\text{past}}-\text{VP}_{\text{say}}-\text{CP}_{\text{that}}-\text{IP}_{\text{past}}-\text{VP}_{\text{see}}-\text{end} \]

many smaller

many reused

It may be faster to put together one big piece.

Dickson et al. 2022, in prep.
Learning the building blocks

Learning **efficient building blocks**

a **balance** between

(1) **how big** they are
(2) **how fast** they are to put together to make a *wh*-dependency

\[ \text{start-IP}_{\text{past}}-\text{VP}_{\text{say}-\text{CP}_{\text{that}}-\text{IP}_{\text{past}}-\text{VP}_{\text{see}}-\text{end}} \]

many smaller

many

reused

one big

faster because one

It may be **slower** if the piece is used *rarely.*

Dickson et al. 2022, *in prep.*
Learning the building blocks

Learning efficient building blocks

A balance between
(1) how big they are
(2) how fast they are to put together to make a $wh$-dependency

$\text{start-IP}_{\text{past}}\text{-VP}_{\text{say}}\text{-CP}_{\text{that}}\text{-IP}_{\text{past}}\text{-VP}_{\text{see}}\text{-end}$

Many smaller

Many

Many reused

One big

Faster because one

It may be slower if the piece is used rarely.

Dickson et al. 2022, in prep.
Learning the building blocks

learning efficient building blocks

a balance between
(1) how big they are
(2) how fast they are to put together to make a \textit{wh}-dependency

\[\textit{start-IP}_\text{past-VP}_\text{say-CP}_\text{that-IP}_\text{past-VP}_\text{see-end}\]

many smaller
many reused

The most \textbf{efficient} option is probably a balance of \textbf{bigger} and \textbf{smaller} blocks that collectively are faster to put together.

\textit{Dickson et al. 2022, in prep.}
Learning the building blocks

learning efficient building blocks

a balance between
(1) **how big** they are
(2) **how fast** they are to put together to make a *wh*-dependency

\[ \text{start-IP}_{\text{past-VP}} \text{say-CP}_{\text{that-IP}} \text{past-VP}_{\text{see-end}} \]

many smaller
many reused

one big
one rare

Dickson et al. 2022, in prep.
Learning the building blocks

How can children find the best balance?

many smaller
many reused

one big
one rare

Dickson et al. 2022, in prep.
Learning the building blocks

Use Bayesian inference to search through the hypothesis space of all possible building blocks (O’Donnell 2015) and find an efficient set for children’s input.
So that’s what the modeled child will do
There’s additional target behavior about *wh*-dependencies we’d like to capture.
Before: Adult judgments + child preferences of certain *wh*-dependencies
Before: Target behavior

Children prefer this interpretation.

What \[ [NP [CP \_what]] \]?

- Subject island
- Complex NP island
- Whether island
- Adjunct island

whether island structure
non-island structure

higher-SES: Sprouse et al. 2012
lower-SES: Bates et al. in prep.

higher-SES: De Villiers et al. 2008
lower-SES

certain wh-dependencies

z-score rating

matrix embedded
certain \textit{wh}-dependencies

\textbf{Subject island}

\textbf{Complex NP island}

\textbf{Whether island}

\textbf{Adjunct island}

\underline{+ additional target behavior with \textit{wh}-dependencies that vary main verb frequency}

\begin{itemize}
  \item \textbf{What did she think} \[ \text{[that he saw \underline{__}]}? \]
  \item \textbf{What did she say} \[ \text{[that he saw \underline{__}]}? \]
  \item \textbf{What did she whine} \[ \text{[that he saw \underline{__}]}? \]
  \item \textbf{What did she mutter} \[ \text{[that he saw \underline{__}]}? \]
\end{itemize}

\textit{Liu et al. 2019}
certain *wh*-dependencies

Subject island
Complex NP island
Whether island
Adjunct island

+ additional target behavior with *wh*-dependencies that vary main verb frequency

What did she VERB [that he saw __ ]?

*Liu et al. 2019*

Dickson et al. 2022, in prep.
Certain wh-dependencies

+ additional target behavior with wh-dependencies that vary main verb frequency

What did she VERB [that he saw __ ]?

Liu et al. 2019

Important pattern: Positive correlation between main verb frequency and judged acceptability.

Dickson et al. 2022, in prep.
Before: Target behavior

Children prefer this interpretation.

What __what__?  

What [NP [CP __what__]]?

Complex NP

higher-SES  

lower-SES

De Villiers et al. 2008

Dickson et al. 2022, in prep.
+ additional target behavior with other wh-dependencies

Who did the little sister ask how to see?
Who did the boy ask what to bring?
How did the mom learn what to bake?
How did the girl ask where to ride?
How did the boy who sneezed drink the milk?

Who [CP-how ___who]?
Who [CP-what ___who]?
How [CP-what ___how]?
How [CP-where ___how]?
How [NP [CP-who ___how]]?

higher-SES lower-SES

De Villiers et al. 2008
Dickson et al. 2022, in prep.
Whether island

Subject island
Complex NP island

Adjunct island

z-score rating

What did she VERB [that he saw __]? [CP-how _who]?
Who [CP-what _who]?
How [NP [CP-who _how]]? [CP-where _how]?
higher-SES lower-SES

De Villiers et al. 2008

Dickson et al. 2022, in prep.
So what does the **modeled child** do?
Learning the building blocks

Superadditivity predicted for judgments of all four island types.

Dickson et al. 2022, in prep.
Learning the building blocks

- Complex NP island
- Subject island
- Whether island
- Adjunct island

✓ frequency of main verb (log-transformed)

What did she **VERB** [that he saw __ ]?

Positive correlation predicted with verb frequency for judgments of this *wh*-dependency.

Dickson et al. 2022, in prep.
Learning the building blocks

Complex NP island
Subject island
Whether island
Adjunct island

Log Probability

What did she VERB [that he saw __ ]?

Verb type
- bridge
- factive
- manner
- other

whine
scream
remember
notice

Dickson et al. 2022, in prep.
Learning the building blocks

The wh-dependency this interpretation relies on is $10^5$ times more probable than the other one.

Dickson et al. 2022, in prep.
Learning the building blocks

Complex NP island
Subject island
Whether island
Adjunct island

Log Probability

What did she VERB [that he saw __ ]?

Log probability

Verb type
- a: bridge
- a: factive
- a: manner
- a: other

Dickson et al. 2022, in prep.
Learning the building blocks

The modeled child also prefers the child-preferred ones for the other wh-dependencies.

Dickson et al. 2022, in prep.
What do the learned building blocks that lead to this behavior look like?

Dickson et al. 2022, in prep.
Different sizes that the modeled child learned
Some lexicalization based on frequency: more frequent lexical items are used. The frequency threshold is learned by the modeled child per node type (IP, VP, CP, etc.).
Takeaway:
This theory — which has to learn the building blocks — can work for learning knowledge about syntactic islands.
Key idea (again): Learning about the **building blocks of** \(wh\)-dependencies leads to **knowledge about syntactic islands**, even when there’s **less knowledge built in**.

\(wh\)-dependency distribution

Dickson et al. 2022, in prep.
Key idea (again): Learning about the building blocks of wh-dependencies leads to knowledge about syntactic islands when the child’s goal is efficient building blocks.
The big picture

This case study demonstrates how we can use computational cognitive modeling, grounded in empirical data, to evaluate our theories about how children could learn what they do about language.

Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.
Takeaway: Theories that view specific linguistic knowledge (like syntactic islands and \textit{wh}-dependencies more generally) as built up of smaller building blocks can work well.

Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.
The big picture

Takeaway: Children can learn about the building blocks, which allows them to implicitly learn sophisticated knowledge (such as syntactic islands).

Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.
The big picture

Takeaway: Learning good building blocks is possible if children’s goal is an efficient set of building blocks that allows future language processing to be easier (faster).

Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.
Thank you!

Jon Sprouse  Alandi Bates  Richard Futrell  Niels Dickson

BUCLD 2018  UCSD Linguistics 2020  ForMA Group 2020
UMD Linguistics 2020  BUCLD 2021  SCiL 2022
UCI QuantLang Collective

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