How to succeed at syntactic island acquisition without really trying: Learning the right building blocks

Lisa Pearl
University of California, Irvine

February 23, 2024
Pomona Acquisition Workshop (PAW)
Pomona College
What does it mean to succeed at syntactic island acquisition?
One answer: To develop the target behavior we observe about syntactic islands...
One answer: To develop the target behavior we observe about syntactic islands...

Some example behavior: judgment patterns and (dis)preferences for certain utterances related to syntactic islands...

- What?
- What?
- [What [what]]?
- [What [what]]?
One answer: To develop the target behavior we observe about syntactic islands, given the input children get and the time they have to learn.
Acquisition success for syntactic islands
“…without really trying…”

What does it mean to try?

\[
\text{What } [\text{BN1 } [\text{BN2 } \_\text{what}]]? 
\]
“without really trying”

Learn about syntactic islands indirectly by learning about *wh*-dependencies more generally.
“Learning the right building blocks”

Proposal: The child is trying to learn about the building blocks that combine into wh-dependencies.
Proposal: Learn about syntactic islands indirectly by learning about the probabilities of the building blocks for wh-dependencies.
Proposal: Learn simultaneously from the input
(i) what the building blocks are, and
(ii) their probabilities

This turns out to work really well.

Dickson, Pearl & Futrell 2022, Dickson, Futrell, & Pearl 2024, Dickson, Futrell, & Pearl in prep
But first, let’s briefly review some relevant information about the acquisition of syntactic islands.
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*).

*Who does Lily think the kitty for whom is pretty?*
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 strongly dispreferred in English.

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

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

Subject island  

(Ross 1967)
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*?

Jack is somewhat tricksy.

He claimed he bought something.
Syntactic islands involve *wh*-dependencies.

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.

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.
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 worry if Jack bought *[what]?**  **Whether island**

What did Elizabeth wonder whether Jack bought *[what]?**  **Adjunct island**

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

What did Elizabeth think Jack said *[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]

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?  

*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*? 

English adults *judge* these island-crossing dependencies to be *far less acceptable* 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.

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 island-crossing dependencies compared to others (de Villiers et al. 2008).
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*

Additional *wh*-dependency knowledge:
The *frequency* of a lexical item can also affect adult acceptability judgments of potential syntactic islands.

What did Elizabeth say that Jack saw *what*?  

What did Elizabeth whine that Jack 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**

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.

*Syntactic island* (Ross 1967)

So, these judgments and *(dis)preference* can serve as a target for successful acquisition — an outcome we can measure.
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

[CP... __who]? 

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

Who

[non-island ]?

Who

[island ]?
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

Who ___ thinks [the necklace is expensive]?

What does Jack think [___ is expensive]?

Who ___ thinks [the necklace for Lily] is expensive?

*Who does Jack think [the necklace for ___] is expensive?

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)

**Whether island stimuli**

- Who ___ thinks [that Jack stole the necklace]?
  - matrix | non-island
- What does the teacher think [that Jack stole ___]?  
  - embedded | non-island
- Who ___ wonders [whether Jack stole the necklace]?
  - matrix | island
- *What does the teacher wonder [whether Jack stole ___]?  
  - 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)

Adjunct island stimuli

<table>
<thead>
<tr>
<th>Question</th>
<th>Type</th>
<th></th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who ___ thinks [that Lily forgot the necklace]?</td>
<td>matrix</td>
<td>non-island</td>
<td></td>
</tr>
<tr>
<td>What does the teacher think [that Lily forgot ___]?</td>
<td>embedded</td>
<td>non-island</td>
<td></td>
</tr>
<tr>
<td>Who ___ worries [if Lily forgot the necklace]?</td>
<td>matrix</td>
<td>island</td>
<td></td>
</tr>
<tr>
<td>*What does the teacher worry [if Lily forgot ___]?</td>
<td>embedded</td>
<td>island</td>
<td></td>
</tr>
</tbody>
</table>

Sprouse et al. 2012
Syntactic islands
Adult knowledge as measured by acceptability judgment behavior

- Length of dependency: (matrix vs. embedded)
- 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
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.
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 [non-island]? Who [island]?

Who [non-island]? Who [island]?

Who [CP... _who]? Who [non-island]?
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
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

Who [non-island]?
Who [island]?

matrix
embedded

non-island structure
island structure

Who __who? __who?

[CP... __who]?
Syntactic islands

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

Who [non-island]?

Who [island]?

Who [CP... who]?
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
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

Who [non-island]
Who [island]

Who [CP... __who]?

Who __who?

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

Who [non-island]? Who [island]?

Who [CP... who]?
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

"island difference": extra difference in acceptability

Who [non-island]?

Who [island]?

Who

__who__?

Who [CP... __who__]?
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)

Syntactic island = superadditive interaction of the two factors

“island difference”: superadditivity = positive difference (positive slope)
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

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

(non-parallel lines)

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
Adult knowledge as measured by \textit{acceptability judgment behavior}.

- \textit{Length of dependency} (matrix vs. embedded) \times \textit{Presence of an island structure} (non-island vs. island).

**Syntactic islands** = \textit{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

Syntactic islands
Child judgments
= behavioral target outcome

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?

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

How do children prefer to interpret potentially ambiguous wh-questions?
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__?
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 needle and thread*
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]]]]]??]

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?

children strongly prefer this interpretation
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.
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.
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]]]]]]]]?
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 [NP the cat [that [was [lying [on [the table [with [___what]]]]]]]]?*

+ other wh-dependencies

*What did the mother say she bought ___?*

*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 ___?*

De Villiers et al. 2008
Adult knowledge as measured by acceptability judgment behavior


What did she think [that he saw ___]?  
What did she say [that he saw ___]?  
What did she whine [that he saw ___]?  
What did she mumble [that he saw ___]?
Syntactic islands
More adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior


<table>
<thead>
<tr>
<th>Verb</th>
<th>Frequency of Main Verb with CP (log-transformed)</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>hope</td>
<td>-6</td>
<td>0.9</td>
</tr>
<tr>
<td>think</td>
<td>-5</td>
<td>0.8</td>
</tr>
<tr>
<td>say</td>
<td>-4</td>
<td>0.8</td>
</tr>
<tr>
<td>learn</td>
<td>-3</td>
<td>0.8</td>
</tr>
<tr>
<td>forget</td>
<td>-2</td>
<td>0.8</td>
</tr>
<tr>
<td>know</td>
<td>-1</td>
<td>0.8</td>
</tr>
<tr>
<td>whine</td>
<td>0</td>
<td>0.8</td>
</tr>
<tr>
<td>murmur</td>
<td>1</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Liu et al. 2019, 2022
Adult knowledge as measured by acceptability judgment behavior


Important pattern: Positive correlation between main verb with CP frequency and judged acceptability.
Syntactic islands
Adult & child judgments
= behavioral target outcome

Subject island
Complex NP island
Whether island
Adjunct island

Sprouse et al. 2012

What [__what]?

De Villiers et al. 2008

Liu et al. 2019, 2022
What [[ what ]]?  

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.

Complex NP + other wh-dependencies

Child Preference

0.00 0.25 0.50 0.75 1.00
Syntactic islands

How long do children have to learn?

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

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

This is the acquisition problem

What [__what]?
Syntactic islands

...which is where a theory of acquisition comes in.
Syntactic islands

...which is where a theory of acquisition comes in.

Learn the right building blocks
We can **evaluate a theory** by implementing it concretely in a computational cognitive model.

Learn the right building blocks
The model generates predictions that can be compared with available empirical data.
Syntactic islands

And then we can look inside it to see what makes it work (or not work).

Learn the right building blocks
Syntactic islands

So let’s do this for our theory.

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

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, 2024, in prep.
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 (or other longer *wh*-dependencies).

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, 2024, in prep.
View *wh*-dependencies in terms of their **building blocks** and **track** those building blocks in the input.
What phrases contain the gap (but not the *wh*-word)?

Dependencies represented as a sequence of container nodes.

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, 2024, in prep.
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*
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*
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
What __ happened?
= What [IP __ happened]?
= start-IP-end

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

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

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, 2024, in prep.
What __ happened?
= What [IP __ happened]?
= start-IP-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

What did you see __?
= What did [IP you [VP see ___]]?
= start-IP-VP-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).

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, 2024, in prep.
Theory: The child tries to learn what the “best” building blocks are at the same time she learns about their distributions in the input.
the best building blocks

(1) Look for the best-sized units
(2) Sometimes include the lexical item with the phrasal node (XP)

Dickson et al. 2022, 2024, in prep.
Learn the right 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

Dickson et al. 2022, 2024, in prep.
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.

Dickson et al. 2022, 2024, in prep.
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.
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, 2024, in prep.
Learn the right 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

learning efficient building blocks

Dickson et al. 2022, 2024, in prep.
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 __ ?

Dickson et al. 2022, 2024, in prep.
Learn the right 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}} \textit{say-CP}_{\text{IP}} \textit{past-VP}_{\text{see-end}}

\textit{Dickson et al. 2022, 2024, in prep.}
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

Pieces can be *small*, so that many of them make up a *wh*-dependency
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.

Dickson et al. 2022, 2024, in prep.
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

\[ \text{slower because many} \]

But these pieces may get reused, so that makes them faster to access.

Dickson et al. 2022, 2024, in prep.
Learn the right 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

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

But these pieces may get reused, so that makes them faster to access.

Dickson et al. 2022, 2024, in prep.
Learn the right 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{verbatim}
start-IP_{past}-VP_{say}-CP_{that}-IP_{past}-VP_{see}-end
\end{verbatim}

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

Dickson et al. 2022, 2024, in prep.
a balance between
(1) how big they are
(2) how fast they are to put together to make a \textit{wh-}dependency

\begin{verbatim}
start-IP_{past}\text{-}VP_{say}\text{-}CP_{that}\text{-}IP_{past}\text{-}VP_{see}\text{-}end
\end{verbatim}

It may be faster to put together one big piece.

Dickson et al. 2022, 2024, in prep.
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, 2024, in prep.
Learning efficient building blocks involves finding a balance between:

1. How big they are
2. How fast they are to put together to make a *wh*-dependency.

```
start-IP_{past}-VP_{say}-CP_{that}-IP_{past}-VP_{see}-end
```

Many smaller reused components are needed, but one big fast component because one.

It may be slower if the piece is used rarely.

Dickson et al. 2022, 2024, in prep.
What did she say that he saw...

...start past say that past see end

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

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

Dickson et al. 2022, 2024, in prep.
Learn the right 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

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

many smaller
many reused

Dickson et al. 2022, 2024, in prep.
How can children find the best balance?

Dickson et al. 2022, 2024, in prep.
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.

Dickson et al. 2022, 2024, in prep.
Learn the right building blocks

So that’s what the modeled child will do

Dickson et al. 2022, 2024, in prep.
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?

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.
12.7K *wh*-dependencies from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 children between the ages of 1 and 5 years old.

Evaluating the theory

This lets us estimate which wh-dependencies children hear and how often they hear them (the 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 2022
Dickson et al. 2022, 2024, in prep.
Evaluating the theory

Children begin to represent the full structure of *wh*-dependencies (e.g., *wh*-questions and relative clauses) around **18 months**: Perkins & Lidz 2021.

Evaluating the theory

Educated guess: This is when children can start processing *wh*-dependencies reliably from their input.

Evaluating the theory

wh-dependency distribution

(18 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).

Evaluating the theory

wh-dependency distribution

(18 months ≤ age < 60 months)

How many minutes: 
≈10,442,258

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.
Evaluating the theory

How many wh-dependencies is this? (≈10,442,258 minutes)

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.
Evaluating the theory

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

Evaluating the theory

wh-dependency distribution

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.

≈ 2.15 million wh-dependencies

(≈10,442,258 minutes)
Evaluating the theory

An additional wrinkle: Children’s memory isn’t perfect

Dickson et al. 2024
Evaluating the theory

Memory is an important part of processing dependencies (McElree et al. 2003).

\[
\text{wh-dependency distribution}
\]

≈ 2.15 million \textit{wh}-dependencies

What \([\text{__what]}\)?
Evaluating the theory


≈ 2.15 million \(wh\)-dependencies
Evaluating the theory

\[ \text{wh-dependency distribution} \approx 2.15 \text{ million wh-dependencies} \]

Upshot: Children might **not take in** all the *wh*-dependency distribution information in their input.

Dickson et al. 2024
Evaluating the theory

Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

Dickson et al. 2024
Learner intake:
Some parts of any particular wh-dependency may be forgotten in the moment.
Evaluating the theory

Learner intake:
Some parts of any particular wh-dependency may be forgotten in the moment.
Evaluating the theory

imperfect intake (a little forgetting)

Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.
Evaluating the theory

imperfect intake
(a little more forgetting)

Learner intake:
Some parts of any particular wh-dependency may be forgotten in the moment.

Dickson et al. 2024
Evaluating the theory

imperfect intake
(a lot of forgetting)

Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

Dickson et al. 2024
Evaluating the theory

We can investigate memory-impacted learners with a **recency effect** (Anderson & Milson 1989), where more recent information is more likely to be remembered.

Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

Dickson et al. 2024
Evaluating the theory

Anderson & Milson 1989

\[
\frac{1}{\text{position} + 1}^\alpha
\]

Learner intake:
Some parts of any particular wh-dependency may be forgotten in the moment.

Dickson et al. 2024
Evaluating the theory

Can the modeled learner produce the appropriate observable behavior?

\[ \approx 2.15 \text{ million } \text{wh-dependencies} \]

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.
Evaluating the theory

Reminder: **Target behavior**

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

**Evaluating the theory**

Sprouse et al. 2012

De Villiers et al. 2008

Liu et al. 2019, 2022

What did she VERB [that he saw __ ]?

- **Complex NP island**
- + other *wh*-dependencies

**Child Preference**

- Frequency of main verb with CP (log-transformed)
- What
  - _what?_
Evaluating the theory

Sprouse et al. 2012

Dickson et al. 2022, 2024, in prep.
Evaluating the theory

The modeled learners can predict the observed pattern (positive slope = “island difference”).

Sprouse et al. 2012

Dickson et al. 2022, 2024, in prep.
Evaluating the theory

Sprouse et al. 2012

Subject island
Complex NP island
Whether island
Adjunct island

z-score rating

Island Difference Score

Position from end

Remember probability

Complex NP | Subject

Whether | Adjunct

modeled learner (log) probability

% Forgotten

Dickson et al. 2022, 2024, in prep.

This happens even if they forget a whole lot.
Evaluating the theory

Remember probability

Liu et al. 2019, 2022

Dickson et al. 2022, 2024, in prep.

What did she VERB [that he saw _]?
Most modeled learners can predict the observed pattern (positive correlation with frequency).
Evaluating the theory

What did she VERB [that he saw __ ]?

This happens, except when there’s an extraordinary amount of forgetting.

Dickson et al. 2022, 2024, in prep.
Evaluating the theory

Dickson et al. 2022, 2024, in prep.

De Villiers et al. 2008

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

What did she VERB [that he saw __]?

One way to think about these is as qualitative categories.

De Villiers et al. 2008

Dickson et al. 2022, 2024, in prep.
Evaluating the theory

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

Then, we want to try to connect these observed qualitative categories to modeled learner predicted judgments.

Dickson et al. 2022, 2024, in prep.

De Villiers et al. 2008
Evaluating the theory

The different quadrants correspond to child vs. modeled learner qualitative predictions.

De Villiers et al. 2008

Dickson et al. 2022, 2024, in prep.
Evaluating the theory

We'd like the modeled learner predictions to qualitatively match the observed child behavior.

Dickson et al. 2022, 2024, in prep.

De Villiers et al. 2008
Evaluating the theory

What did she VERB [that he saw __ ]?

Complex NP + other wh-dependencies

We’d like the modeled learner predictions to qualitatively match the observed child behavior.

Dickson et al. 2022, 2024, in prep.
Evaluating the theory

Dickson et al. 2022, 2024, in prep.

De Villiers et al. 2008

They mostly do, but there are some exceptions.

This is true, no matter how much forgetting there is.
Takeaway: Modeled learners implementing this learning theory can generate most of the observed target behavior patterns, even with human(-like) memory limitations.
Learn the right building blocks

Bigger takeaway:
This theory can work (pretty well) for learning knowledge about syntactic islands.

Dickson et al. 2022, 2024, in prep.
Learn the right building blocks

Key idea: Learning about the building blocks of *wh*-dependencies leads to knowledge about syntactic islands.

Dickson et al. 2022, 2024, in prep.
Learn the right building blocks

Key idea: This strategy works when the child’s goal is finding efficient building blocks.
One way to succeed at learning about constraints on wh-dependencies (syntactic islands) is to learn them indirectly.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022, Dickson et al. 2022, Dickson et al. 2024, in prep.
In particular, learn how to represent \textit{wh}-dependencies efficiently using pieces that can be used to represent any \textit{wh}-dependency.
The efficient pieces that are learned, with their associated probabilities, allow the constraints to emerge automatically.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022, Dickson et al. 2022, Dickson et al. 2024, in prep.
Thank you!

Niels Dickson
Jon Sprouse
Richard Futrell
Alandi Bates

BUCLD 2018    UCSD Linguistics 2020    ForMA Group 2020
UMD Linguistics 2020    BUCLD 2021    SCiL 2022
UArizona Linguistics 2022    UChicago LEAP 2022
UPenn 2023    UCI QuantLang Collective

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