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

Lisa Pearl

University of California, Irvine

Computation of Language Laboratory

UC Irvine

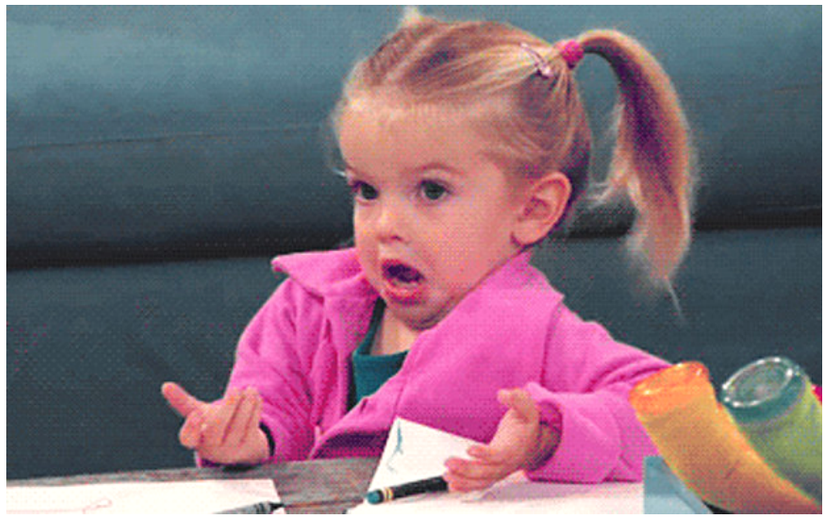
Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lpearl@uci.edu

November 4, 2022

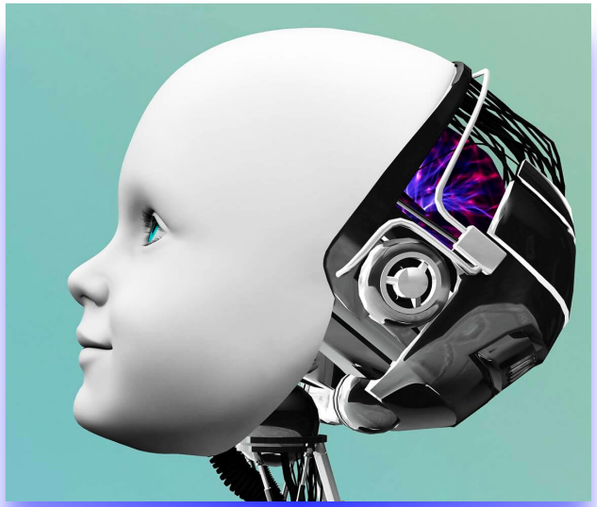
Linguistics Colloquium

University of Arizona

What's a computational model of acquisition?

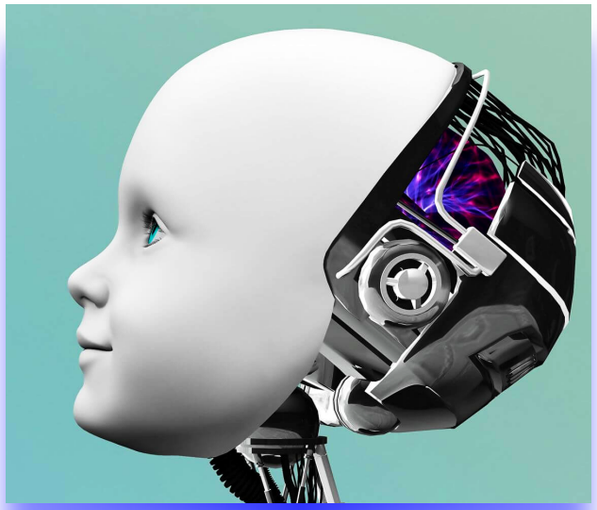


What's a computational model of acquisition?



"the embodiment of a specific theory about acquisition"

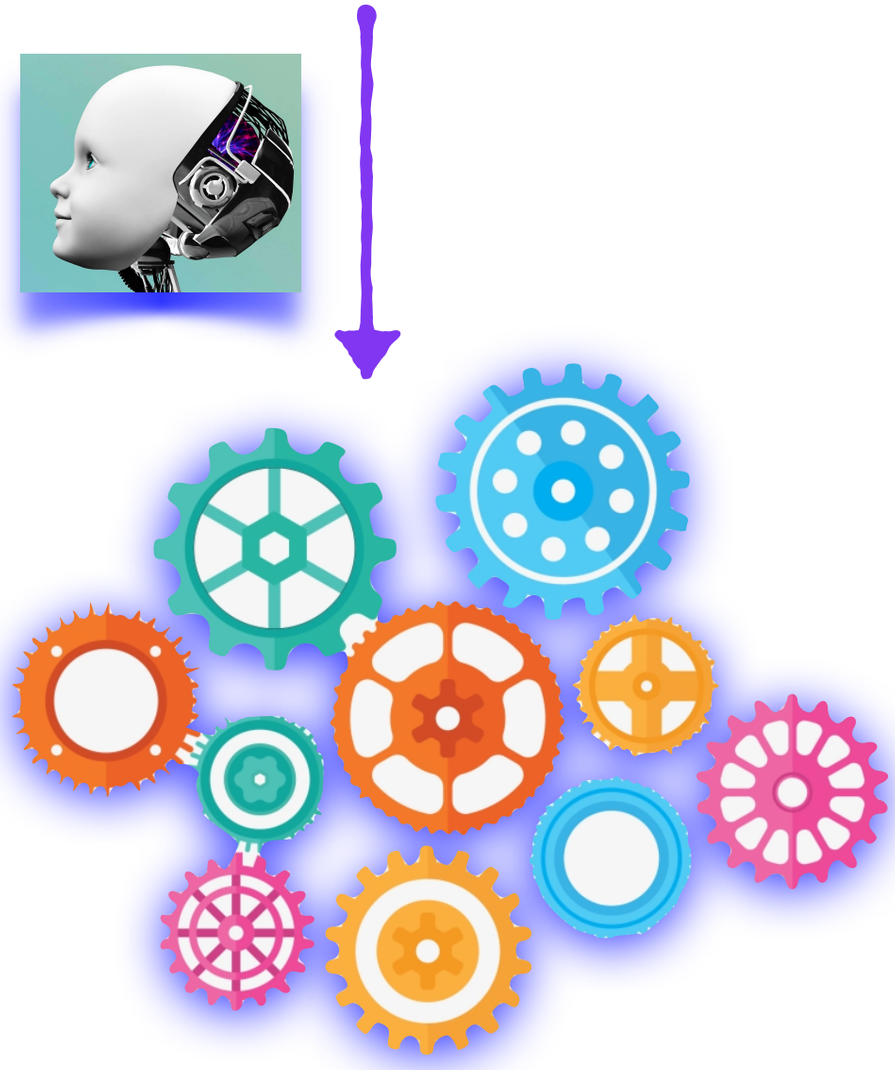
- Pearl in press



computational model

"…you need to first have a theory about how acquisition works." - Pearl in press

theory of acquisition



“Then, the model can be used to

(1) make all the components of that acquisition theory explicit…” - Pearl in press

theory of acquisition

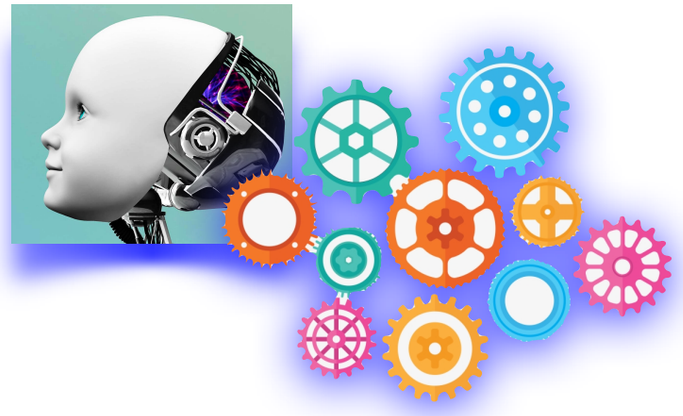
“Then, the model can be used to

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

(2) evaluate whether it actually works…”

- Pearl in press

How?



theory of acquisition

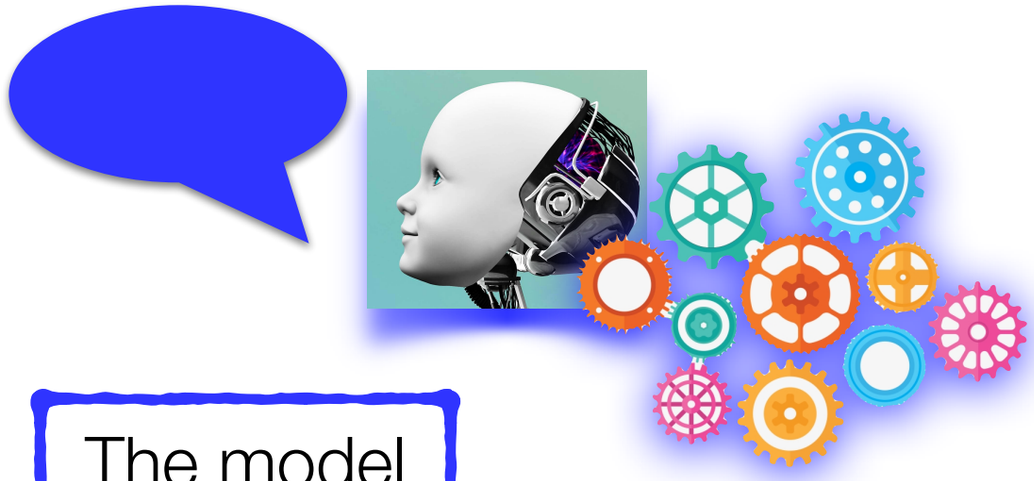
“Then, the model can be used to

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

(2) evaluate whether it actually works…”

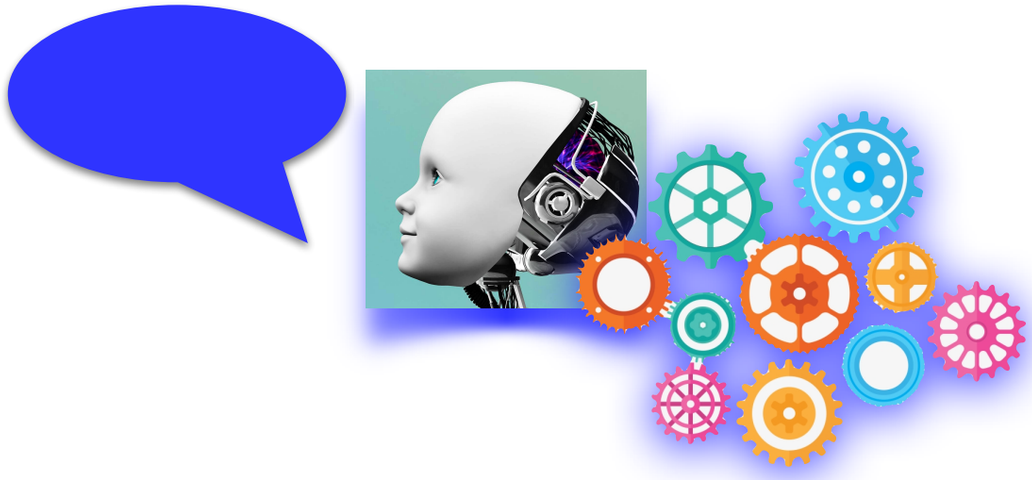
- Pearl in press

How?



The model generates predictions

theory of acquisition



“Then, the model can be used to

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

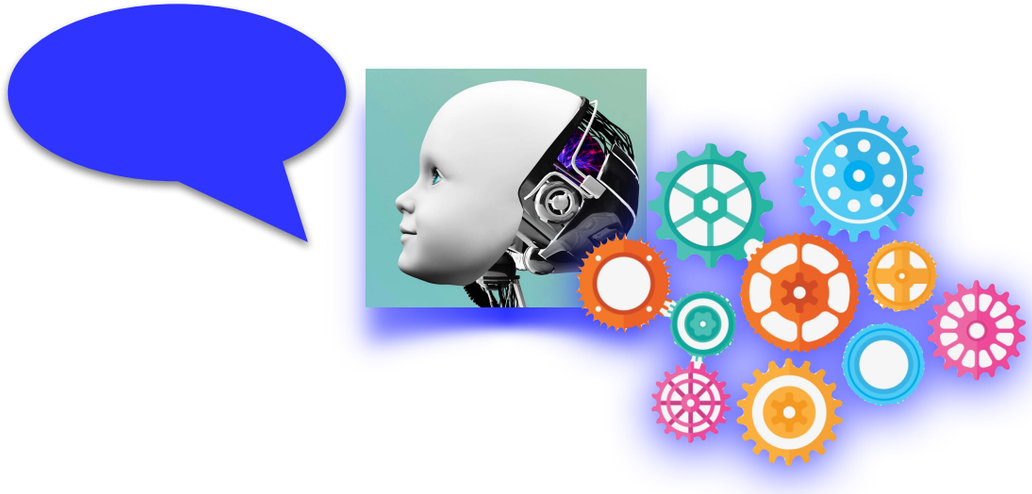
(2) evaluate whether it actually works…”

- Pearl in press

How?

Compare the model’s predictions against empirical data from children

theory of acquisition



“Then, the model can be used to

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

(2) evaluate whether it actually works…”

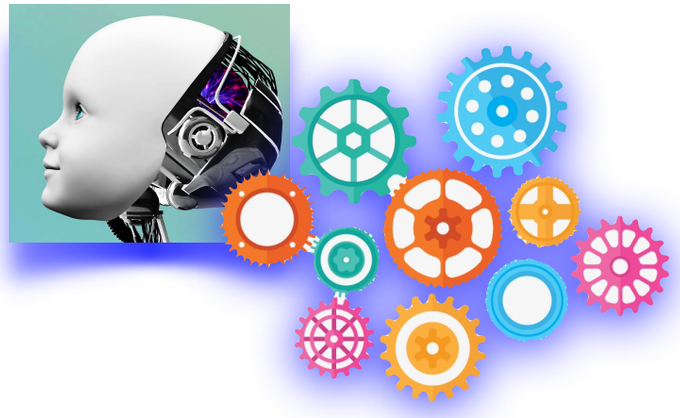
- Pearl in press

How?

✓ ✗

Either they match or they don’t.

theory of acquisition



“Then, the model can be used to

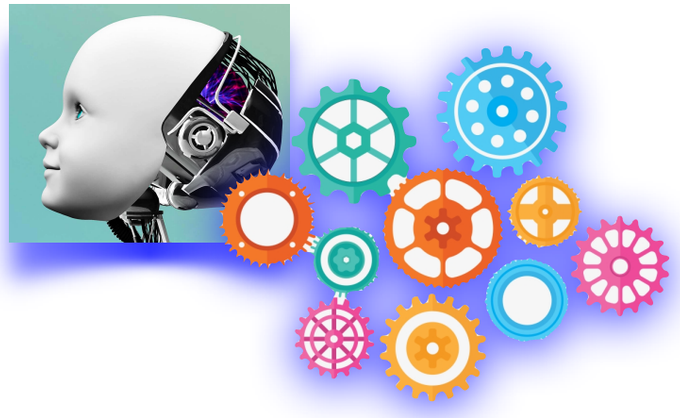
(1) make all the components of that acquisition theory explicit,

(2) evaluate whether it actually works, and

(3) determine precisely what makes it work (or not work).”

- Pearl in press

theory of acquisition



“Then, the model can be used to

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

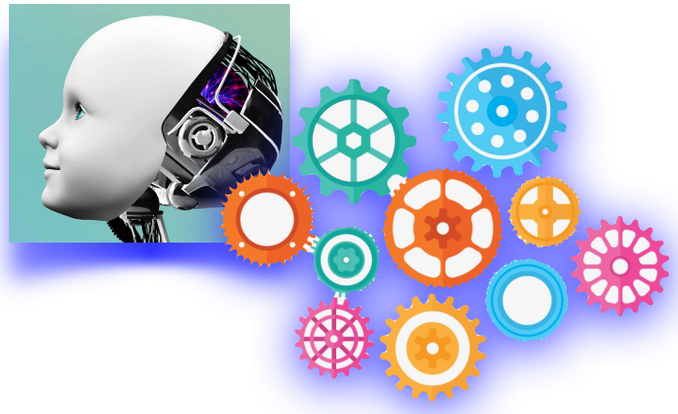
(2) evaluate whether it actually works, and

(3) determine precisely what makes it work (or not work).”

- Pearl in press

Looking inside a model that implements a theory is easier than looking inside a real child’s mind and making sense of it.

theory of acquisition



“Then, the model can be used to

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

(2) evaluate whether it actually works, and

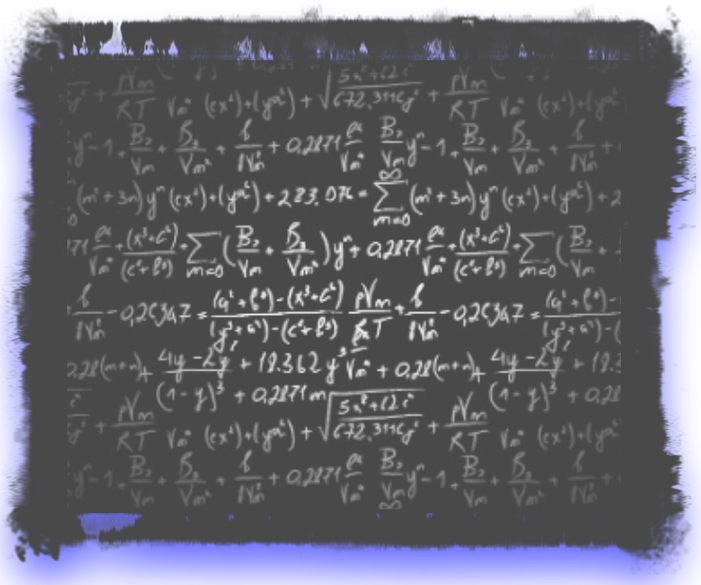
(3) determine precisely what makes it work (or not work).”

- Pearl in press

…but an informative model ideally allows us to understand what’s going on inside a child’s mind.

computational model

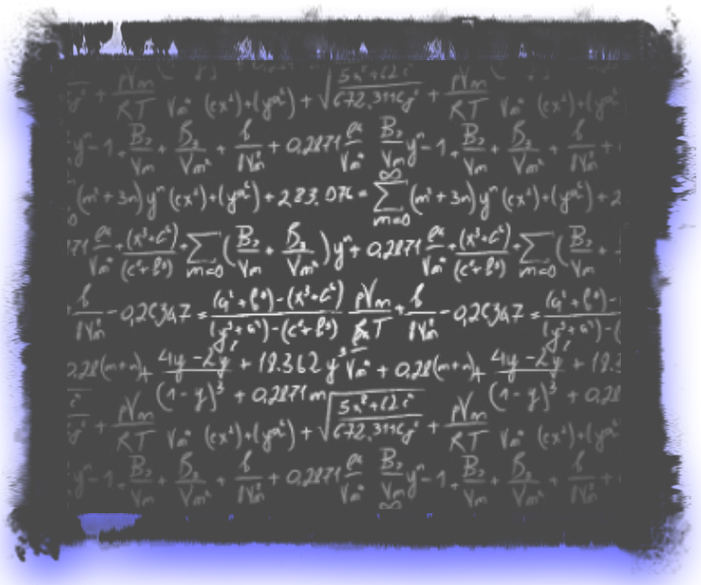
theory of acquisition



We can use quantitative techniques to implement a computational model.

computational model

theory of acquisition

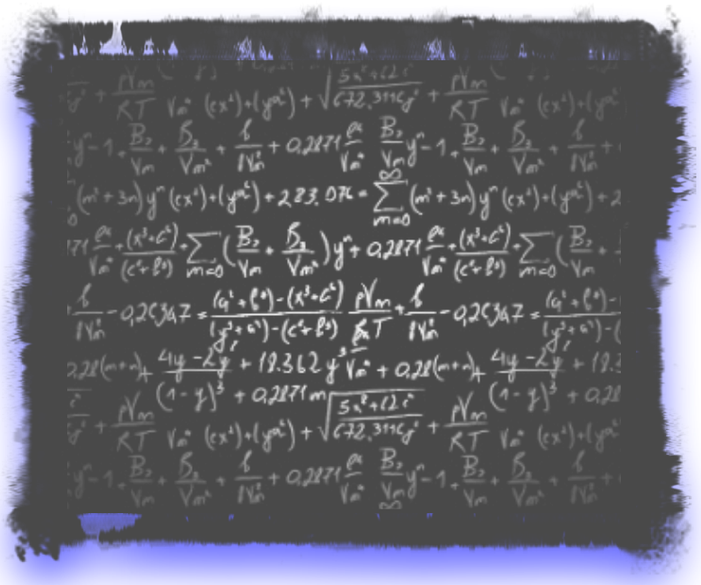


quantitative

One main part: Counting things

computational model

theory of acquisition



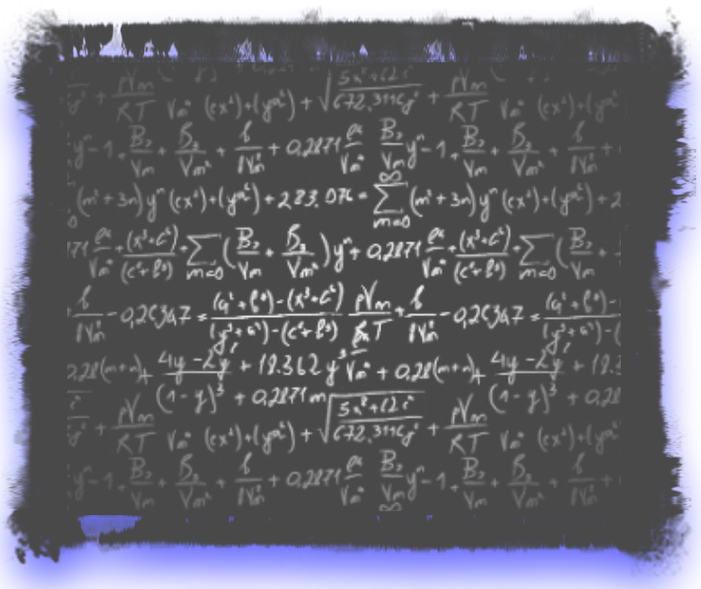
quantitative

One main part: Counting things

(sometimes we count a lot of things)

computational model

theory of acquisition



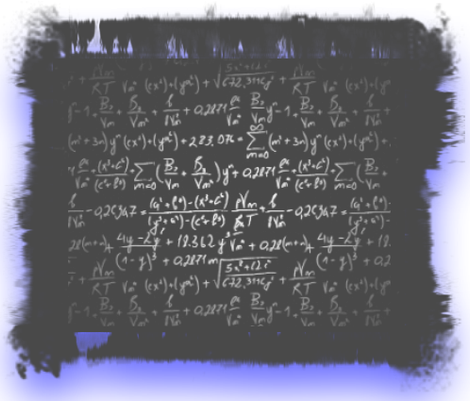
quantitative

counting

Another part: principled reasoning
based on those counts

Bayesian inference

$$p(Hypothesis \mid Data) \propto p(Hypothesis) \cdot p(Data \mid Hypothesis)$$

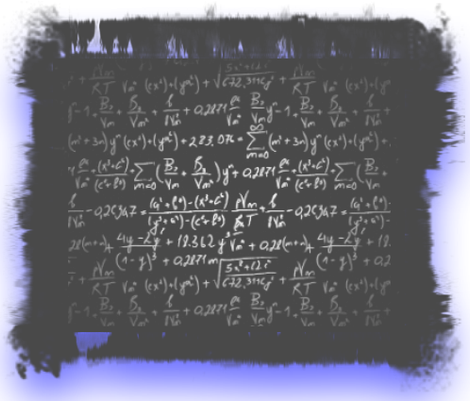


quantitative

computational model

theory of acquisition

We think the child is learning by counting different parts of her input and reasoning over those counts in a sensible way.

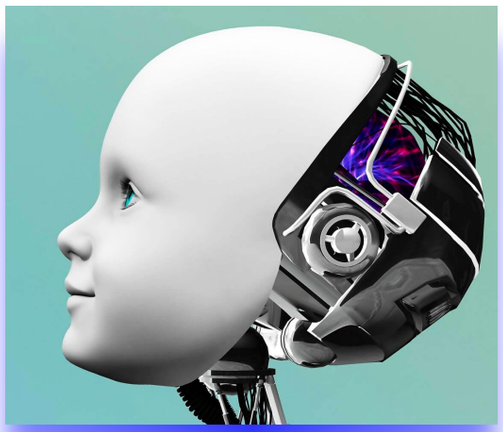


quantitative

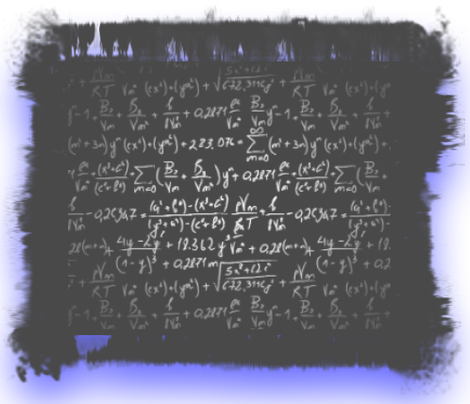
computational model

theory of acquisition

We think the child is learning by counting different parts of her input and reasoning over those counts in a sensible way.

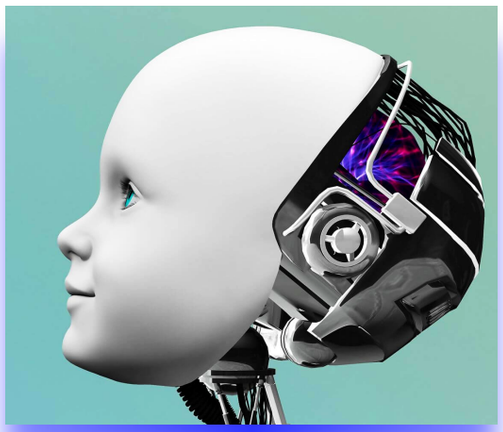


So, the model will count those same things and learn about language by doing principled reasoning over those counts.

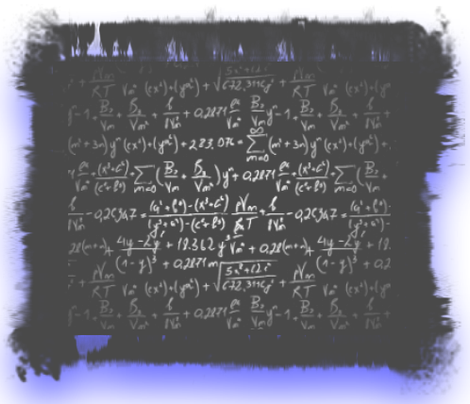


quantitative

computational model

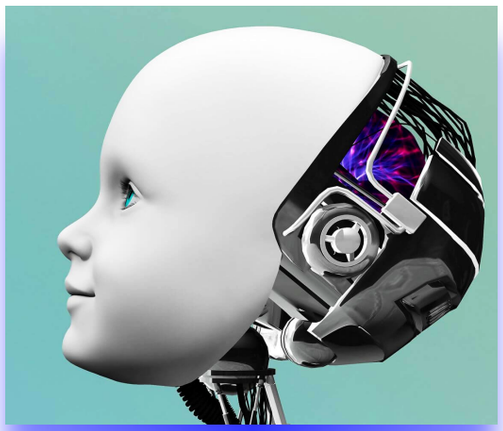


This counting and reasoning is what the theory of acquisition is about.

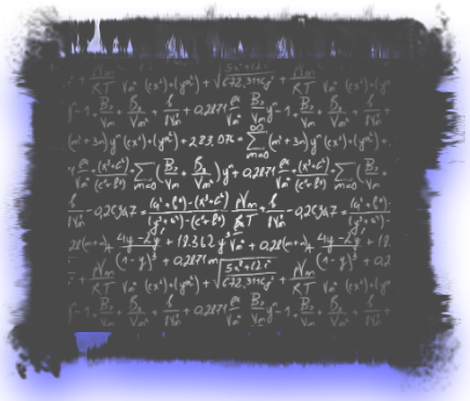


quantitative

computational model



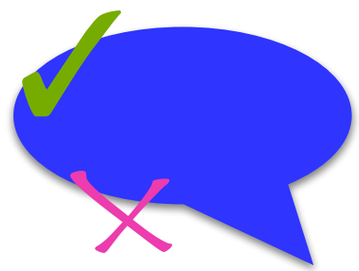
In particular: counting what? And reasoning how?



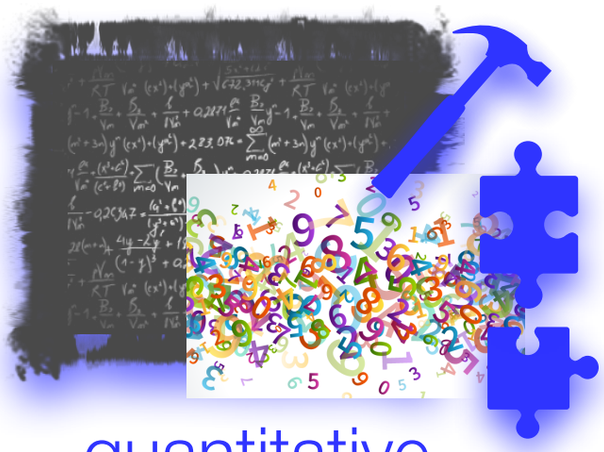
quantitative

computational model

theory of acquisition



We build the model according to what the theory specifies in order to evaluate it. Modeling results can help us refine our theory.



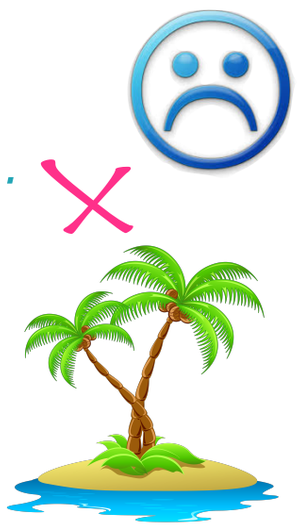
quantitative

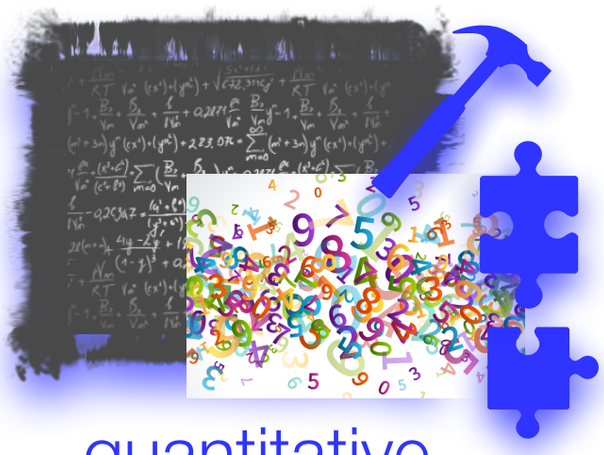
computational model

theory of acquisition

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

Who does…





quantitative

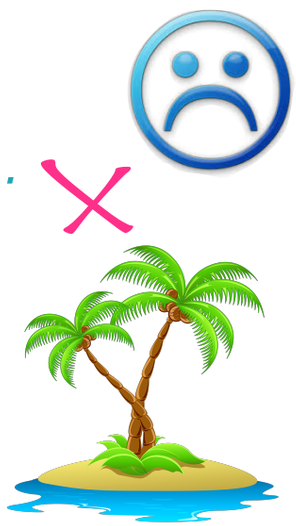
computational model

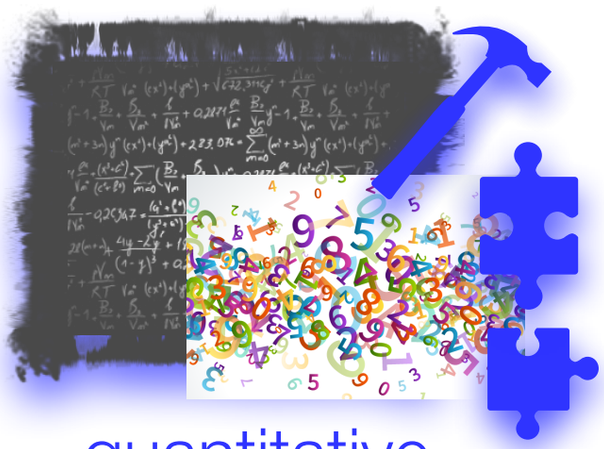
theory of acquisition

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

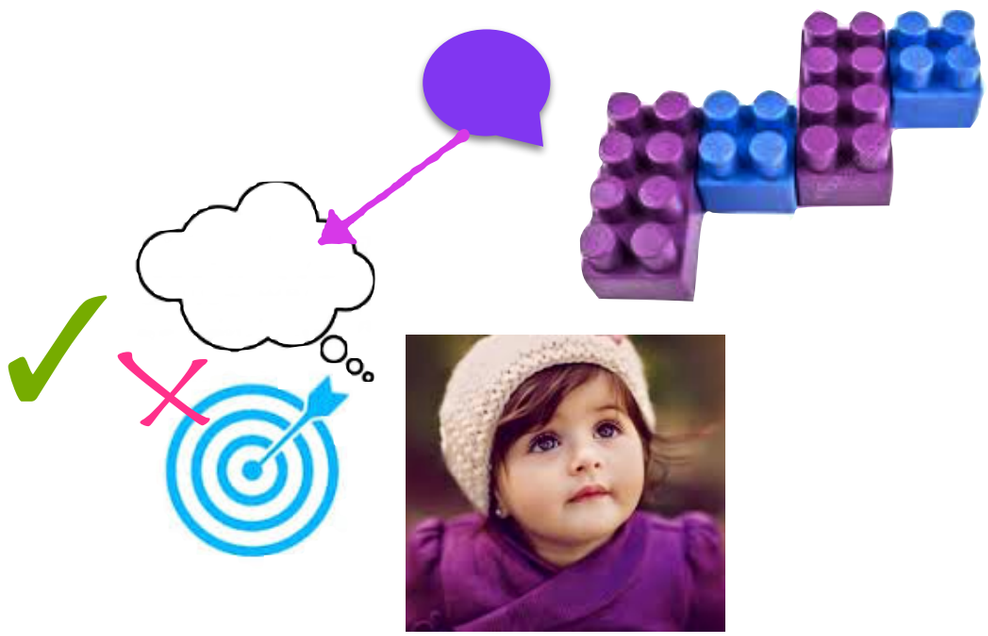
computational model

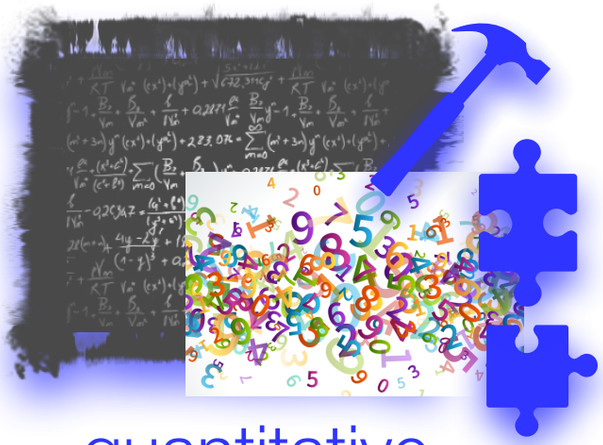
theory of acquisition

Who does…

1: syntactic islands acquisition

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





quantitative

computational model

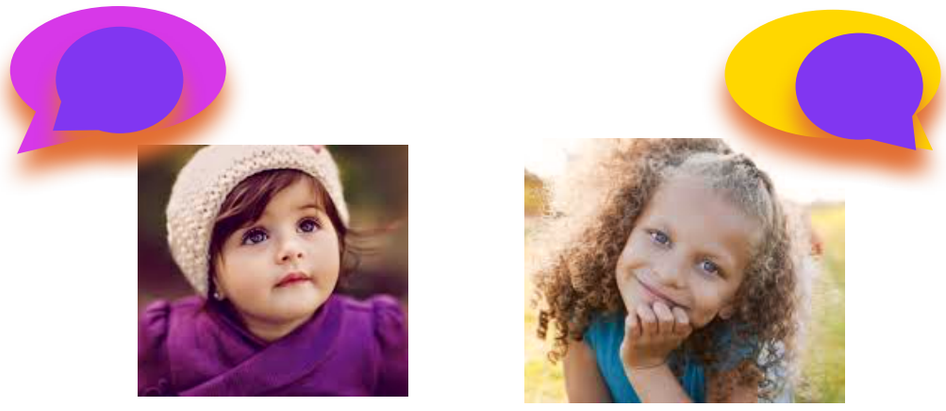
theory of acquisition

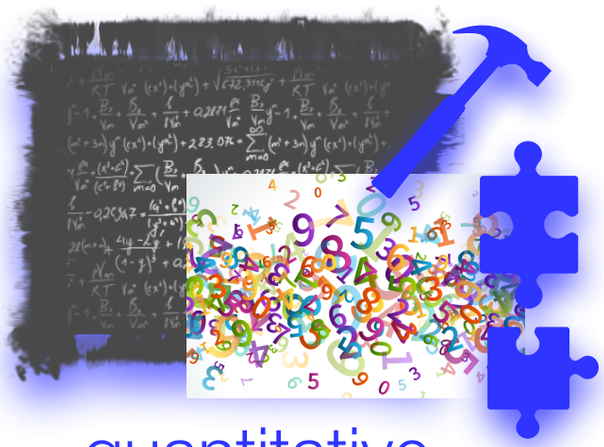
Who does…

1: syntactic islands acquisition

2: Evaluating a theory

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





quantitative

computational model

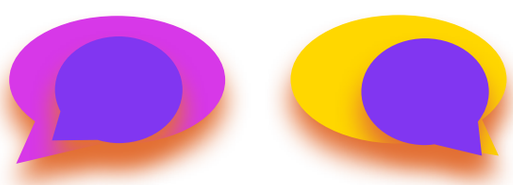
theory of acquisition

Who does…

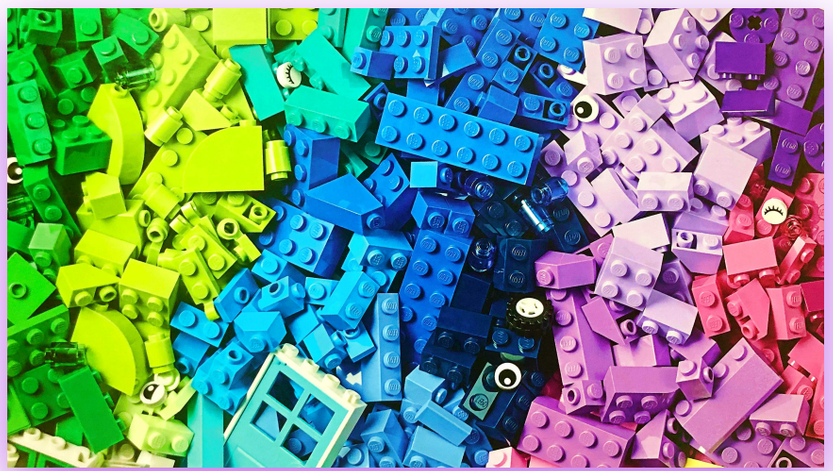
1: syntactic islands acquisition

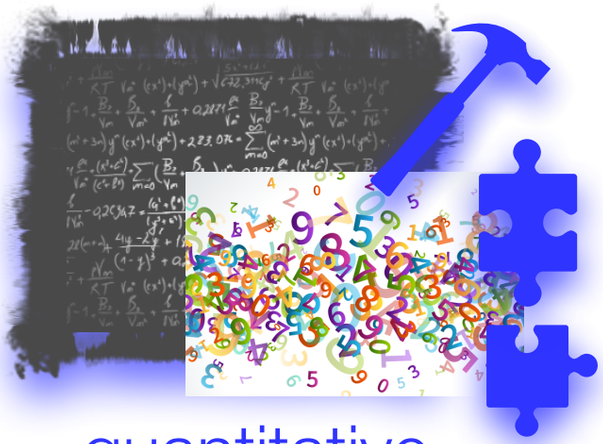
2: Evaluating a theory

3: dialectal variation



Part 4: Evaluating a variant of this theory where the child learns what the building blocks are first





quantitative

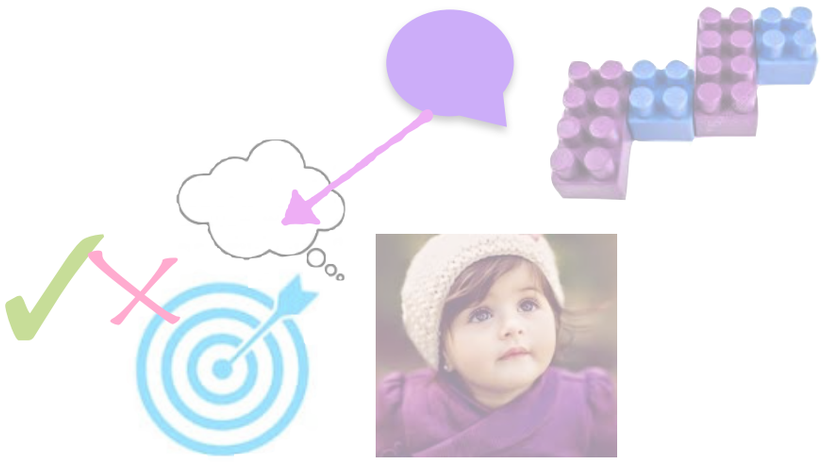
computational model

theory of acquisition

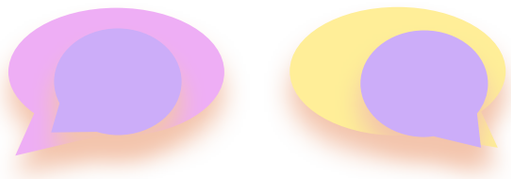
Who does…

1: syntactic islands acquisition

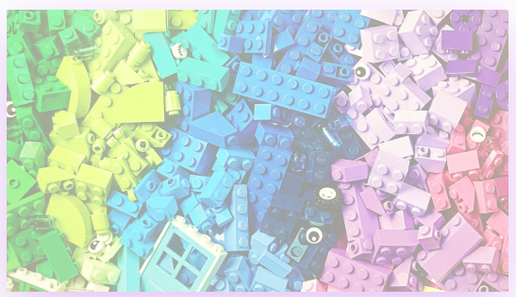
2: Evaluating a theory



3: dialectal variation



4: learning the building blocks



Syntactic islands

involve *wh*-dependencies.

This kitty was bought as a present for someone.

Lily thinks this kitty is pretty.

What's going on here?

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__who 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 not allowed in English.

One explanation: The dependency crosses a "syntactic island" (Ross 1967)

Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

Jack is somewhat tricksy.

He claimed he bought something.

What did Jack make the claim that he 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

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.

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

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.

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.

What did Elizabeth think __what?

Elizabeth

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

Elizabeth

Jack

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

What did Elizabeth think Jack said __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

Elizabeth

Jack

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

What did Elizabeth think Jack said Lily saw __what?

Lily

Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

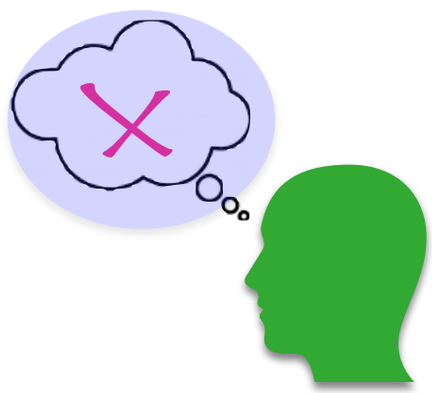
Who does Lily think the kitty for __who is pretty? Subject island

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 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.

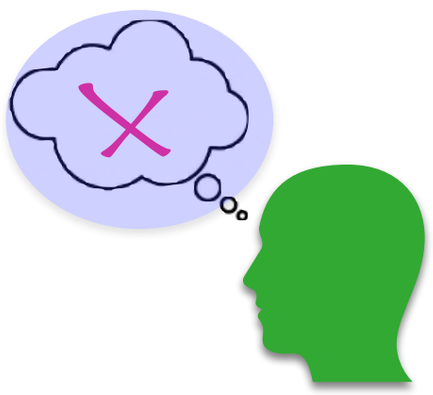
syntactic island (Ross 1967)

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.

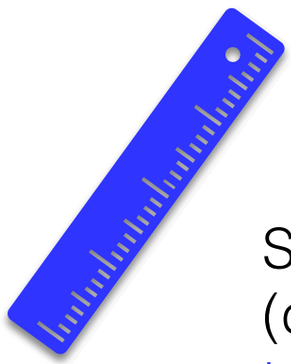
syntactic island (Ross 1967)

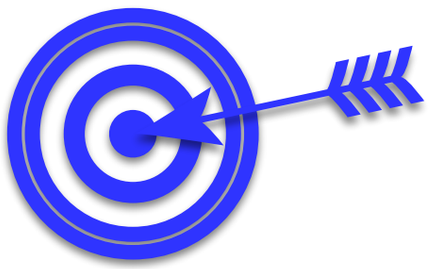
Who does Lily think the kitty for __who is pretty? Subject island

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



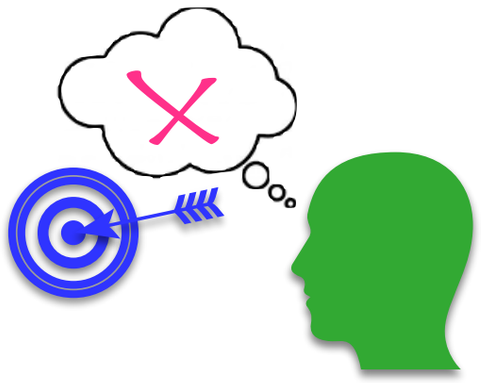


So, these judgments and (dis)preferences can serve as a target for successful acquisition — an outcome we can measure.

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 *__who?*

Who [CP… *__who*]?

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

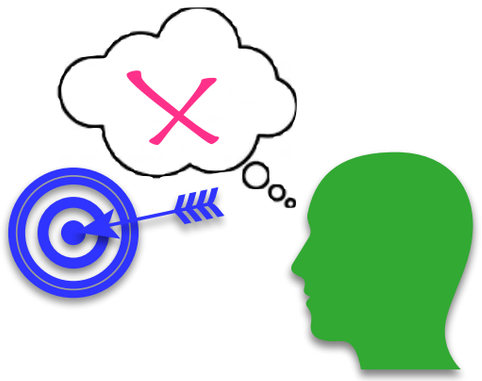
Who [non-island]?

Who [island]?

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**)

Subject island stimuli

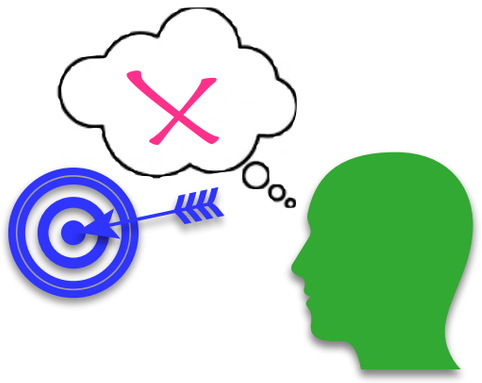
Who __ thinks [the necklace is expensive]?	matrix | non-island
What does Jack think [__ is expensive]?	embedded | non-island
Who __ thinks [the necklace for Lily] is expensive?	matrix | island
*Who does Jack think [the necklace for __] is expensive?	embedded | island

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

Whether island stimuli

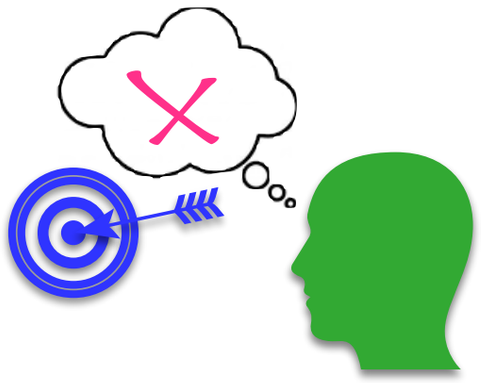
Who __ thinks [that Jack stole the necklace]?	matrix | non-island
What does the teacher think [that Jack stole __]?	embedded | non-island
Who __ wonders [whether Jack stole the necklace]?	matrix | island
*What does the teacher wonder [whether Jack stole __]?	embedded | island

Sprouse et al. 2012

Syntactic islands

Adult judgments

= behavioral target outcome



Adult knowledge as measured by acceptability judgment behavior

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

Adjunct island stimuli

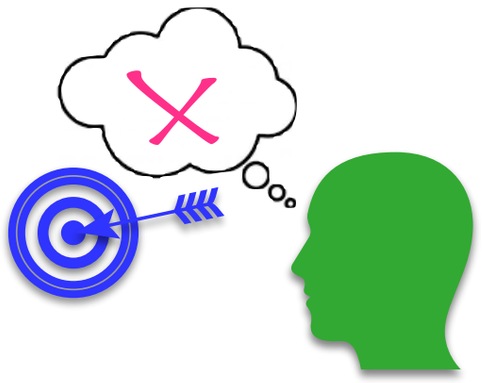
Who __ thinks [that Lily forgot the necklace]?	matrix | non-island
What does the teacher think [that Lily forgot __]?	embedded | non-island
Who __ worries [if Lily forgot the necklace]?	matrix | island
*What does the teacher worry [if Lily forgot __]?	embedded | island

Sprouse et al. 2012

Syntactic islands

Adult judgments

= behavioral target outcome



Adult knowledge as measured by acceptability judgment behavior

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

Complex NP island stimuli

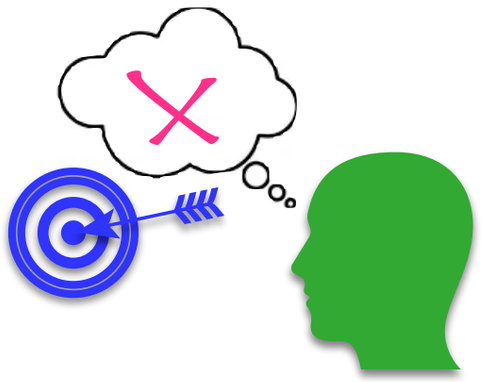
Who __ claimed [that Lily forgot the necklace]?	matrix | non-island
What did the teacher claim [that Lily forgot __]?	embedded | non-island
Who __ made [the claim that Lily forgot the necklace]?	matrix | island
*What did the teacher make [the claim that Lily forgot __]?	embedded | island

Sprouse et al. 2012

Syntactic 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**)

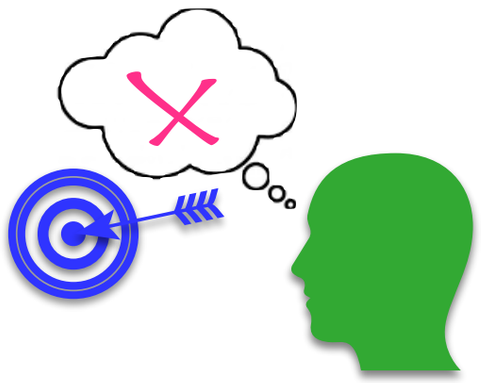
Syntactic island = **superadditive** interaction of the two factors.
This is additional unacceptability that arises when the two factors — length & presence of an island structure — are combined, above and beyond the independent contribution of each factor.

Sprouse et al. 2012

Syntactic islands

Adult judgments

= behavioral target outcome



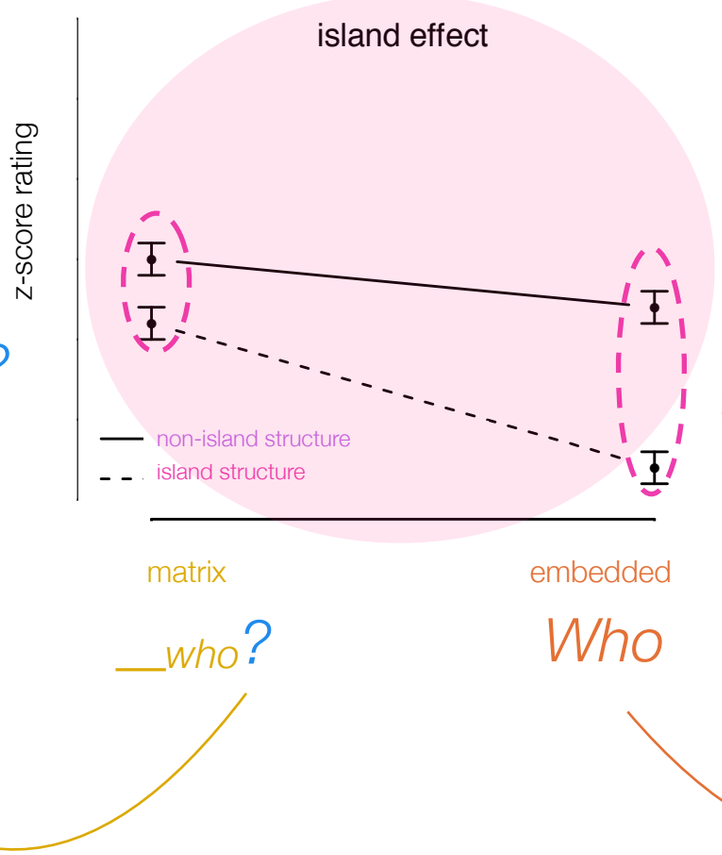
Adult knowledge as measured by acceptability judgment behavior

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

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

Who [non-island]?

Who [island]?



additional unacceptability

Who *__who*?

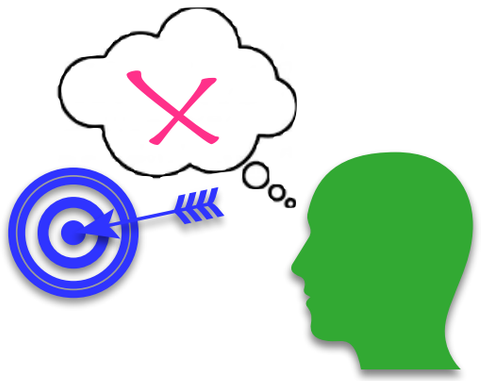
Who [CP… *__who*]?

Sprouse et al. 2012

Syntactic islands

Adult judgments

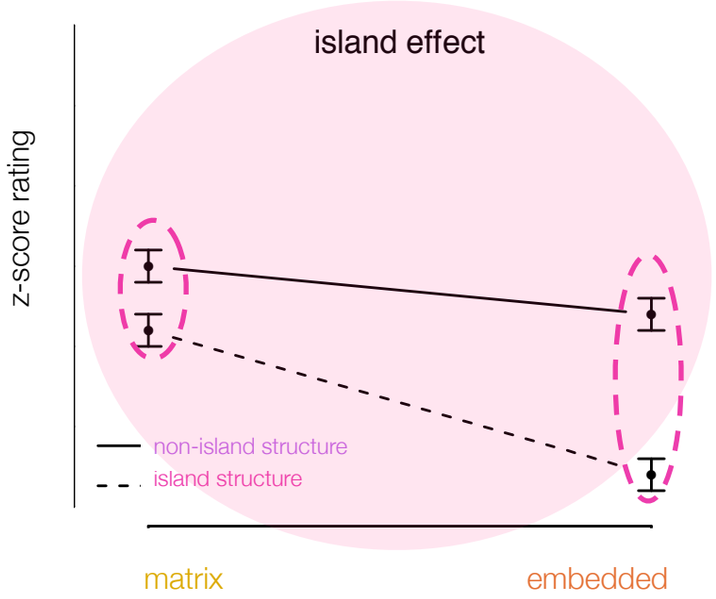
= behavioral target outcome

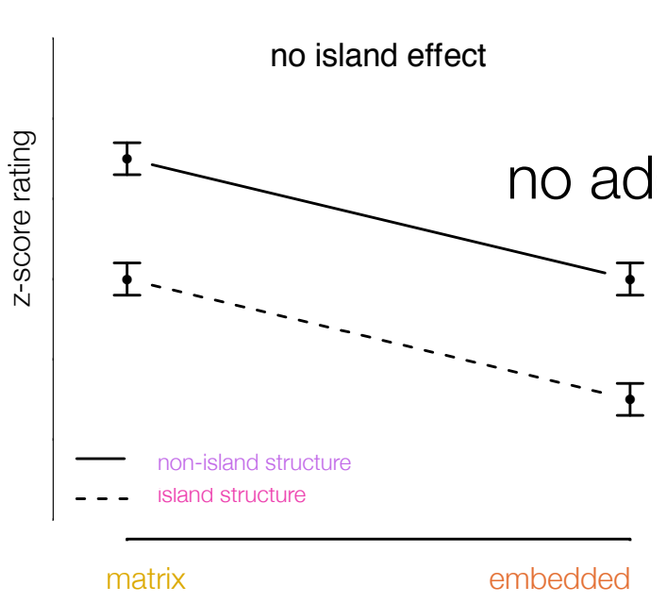


Adult knowledge as measured by acceptability judgment behavior

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

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





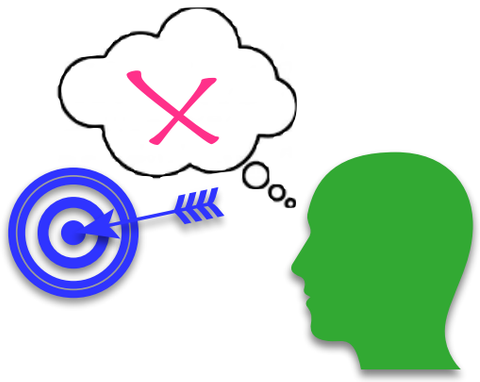
no additional unacceptability

Sprouse et al. 2012

Syntactic islands

Adult judgments

= behavioral target outcome

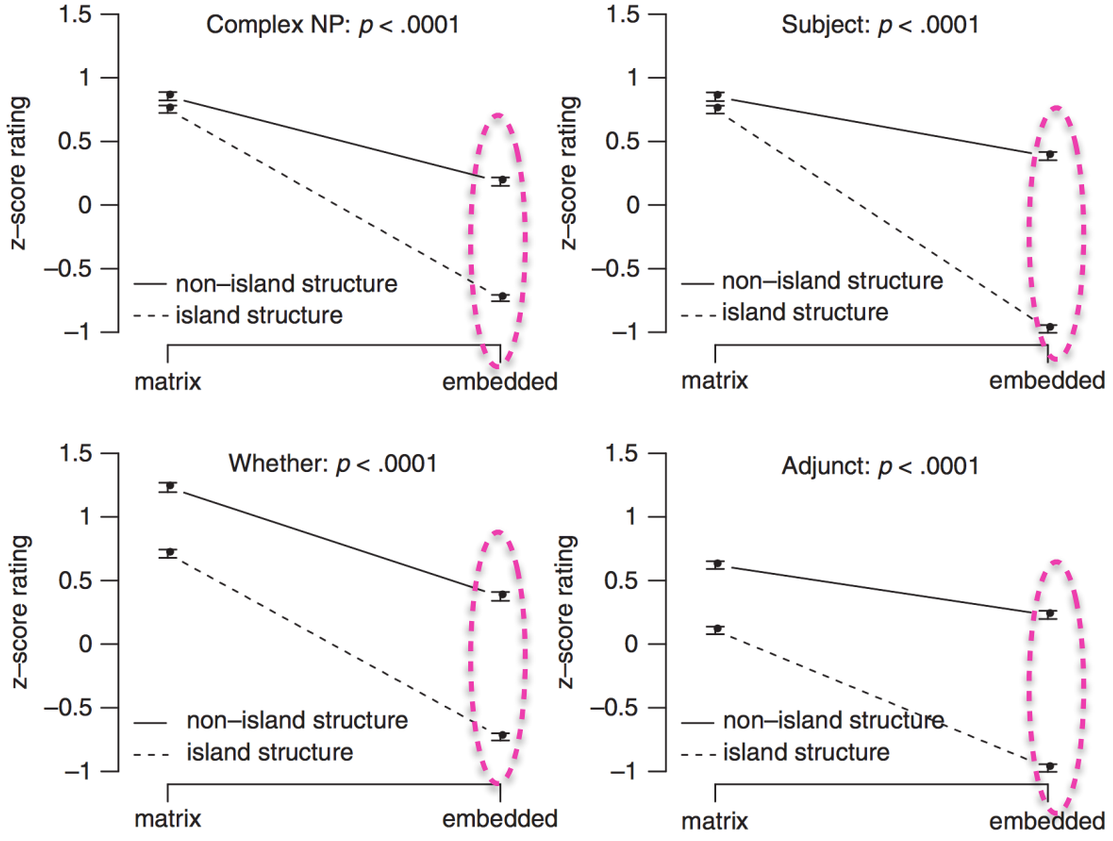


Adult knowledge as measured by acceptability judgment behavior

length of dependency **(matrix vs. embedded)** X 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

Sprouse et al. 2012

Syntactic islands

Adult judgments

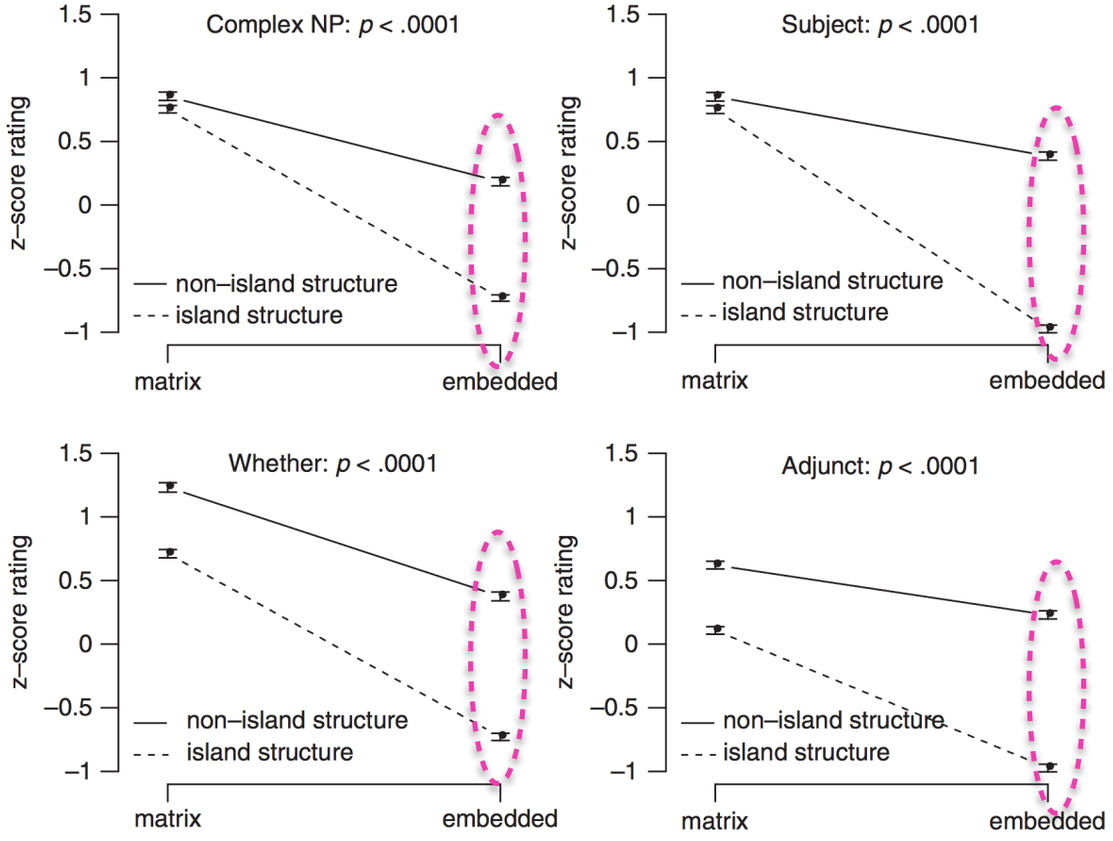
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

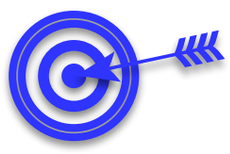
Syntactic island = **superadditive** interaction of the two factors

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.

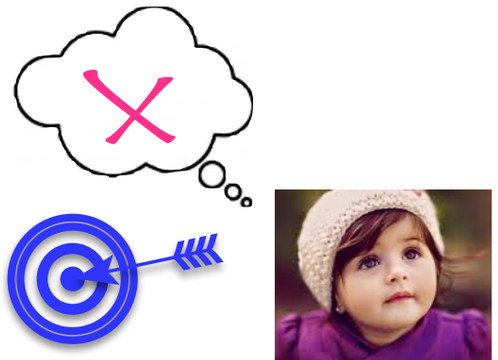


Sprouse et al. 2012

Syntactic islands

Child judgments

= behavioral target outcome



Child knowledge as measured by preferred interpretation behavior

De Villiers et al. 2008:

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

Syntactic islands

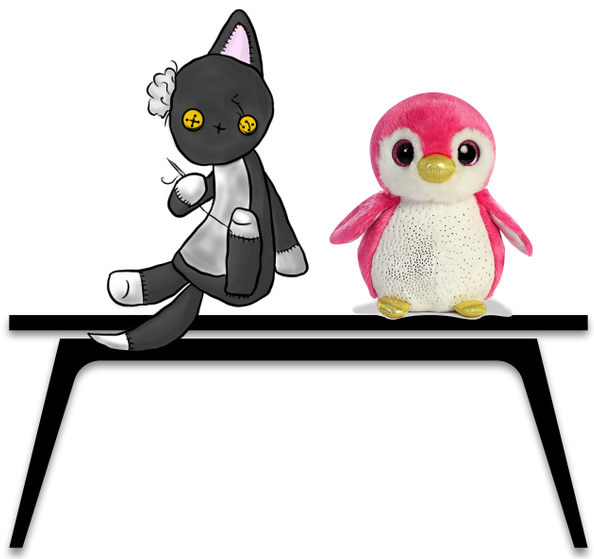
Child judgments

= 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*?



De Villiers et al. 2008

Syntactic islands

Child judgments

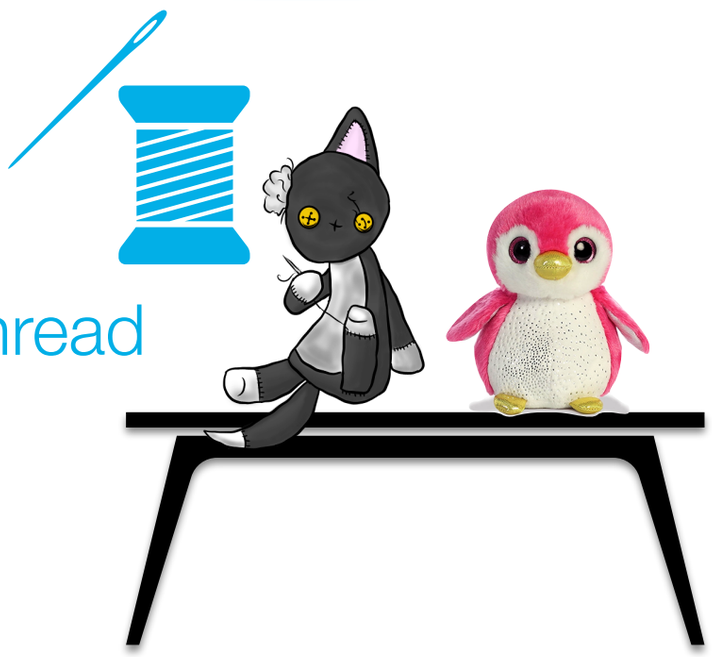
= behavioral target outcome

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 needle and thread



De Villiers et al. 2008

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

De Villiers et al. 2008

Syntactic islands

Child judgments

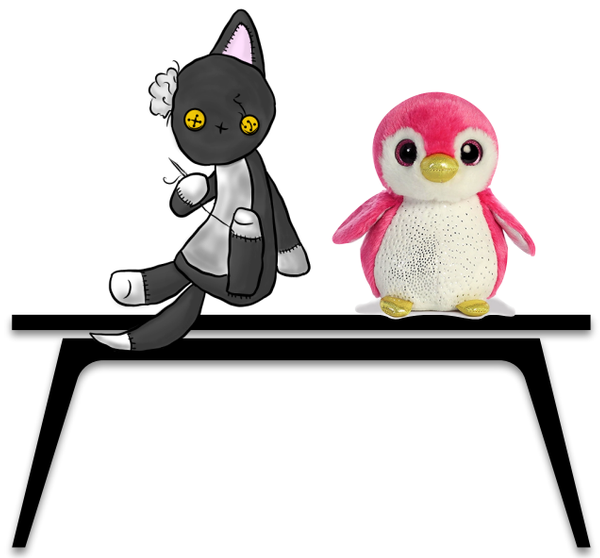
= behavioral target outcome

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*?

children strongly prefer this interpretation



De Villiers et al. 2008

Syntactic islands

Child judgments

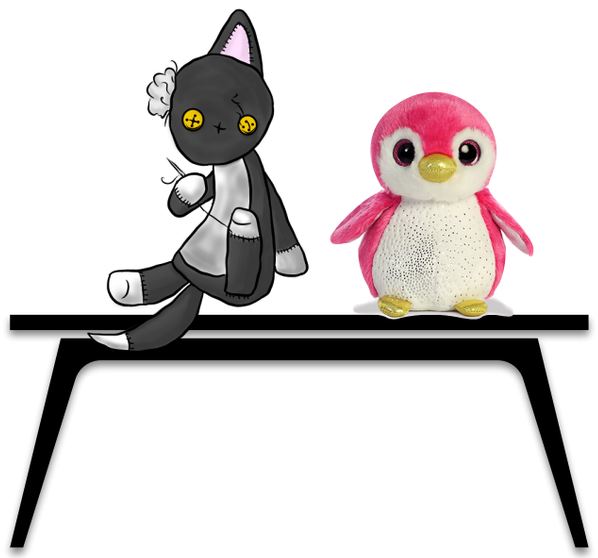
= behavioral target outcome

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

Syntactic islands

Child judgments

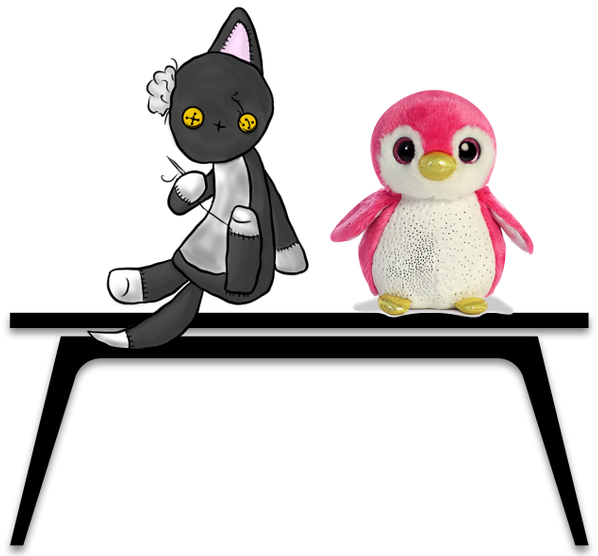
= behavioral target outcome

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*]]]]]]]]?

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



De Villiers et al. 2008

Syntactic islands

Child judgments

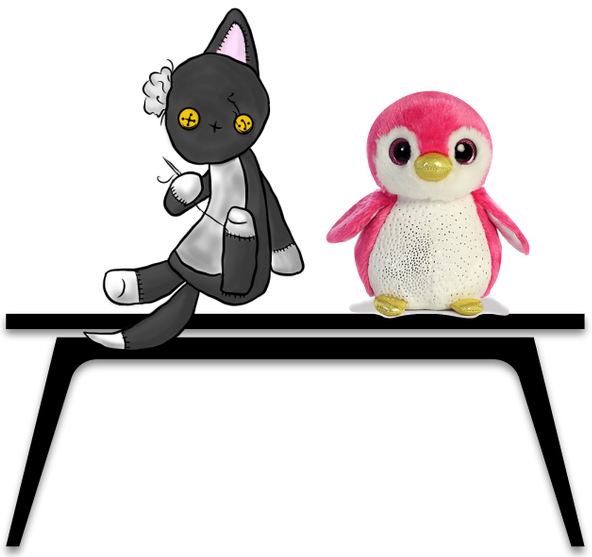
= behavioral target outcome

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*]]]]]]]]?

…which is a dependency that crosses a Complex NP.

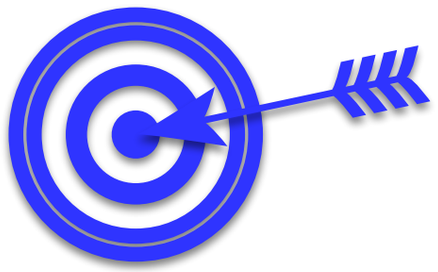


De Villiers et al. 2008

Syntactic islands

Adult & child judgments

= behavioral target outcome

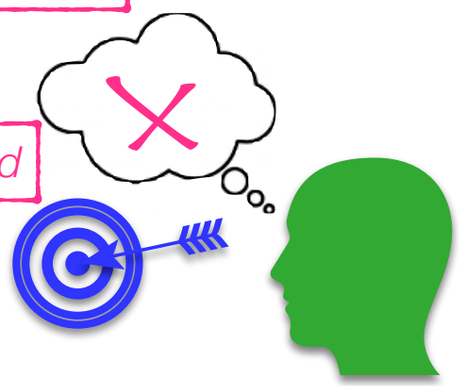


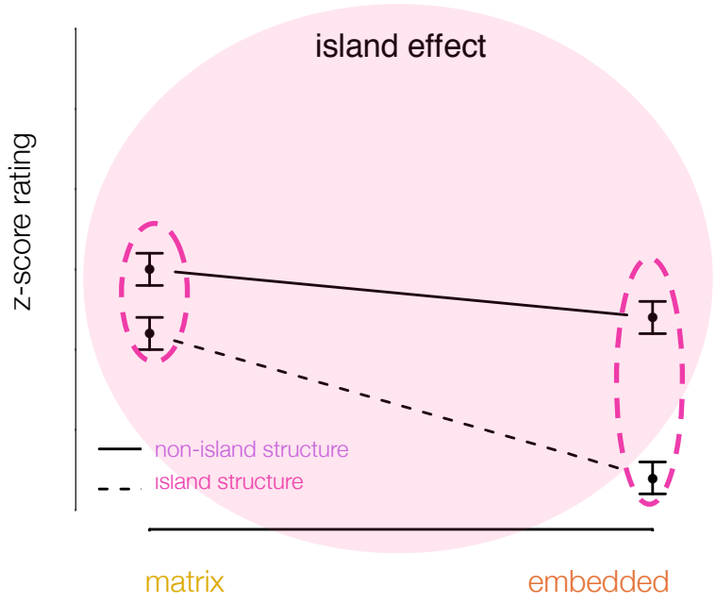
Subject island

Complex NP island

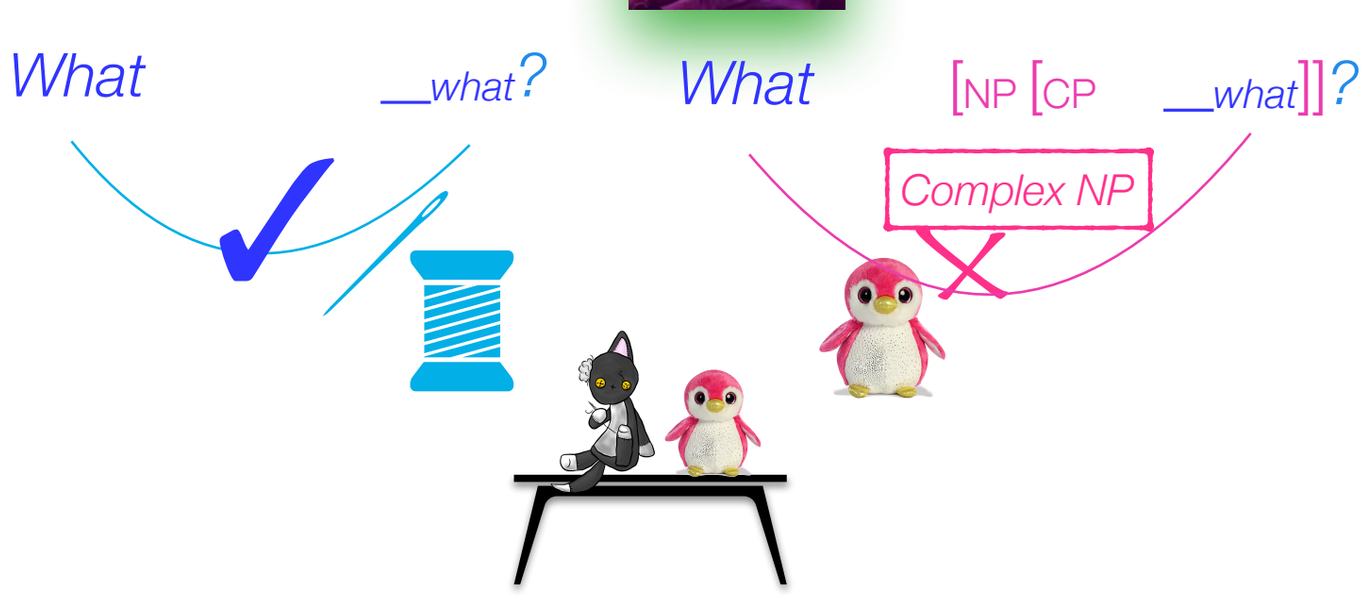
Whether island

Adjunct island

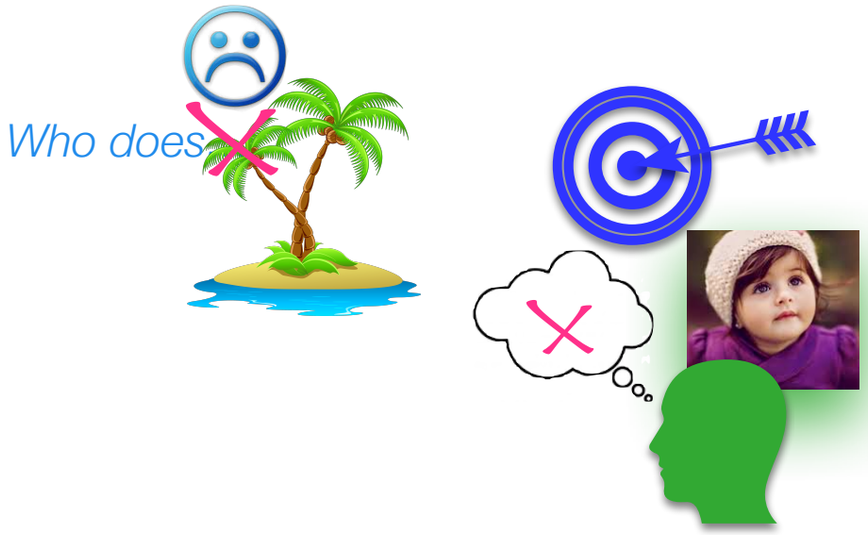




Sprouse et al. 2012



De Villiers et al. 2008



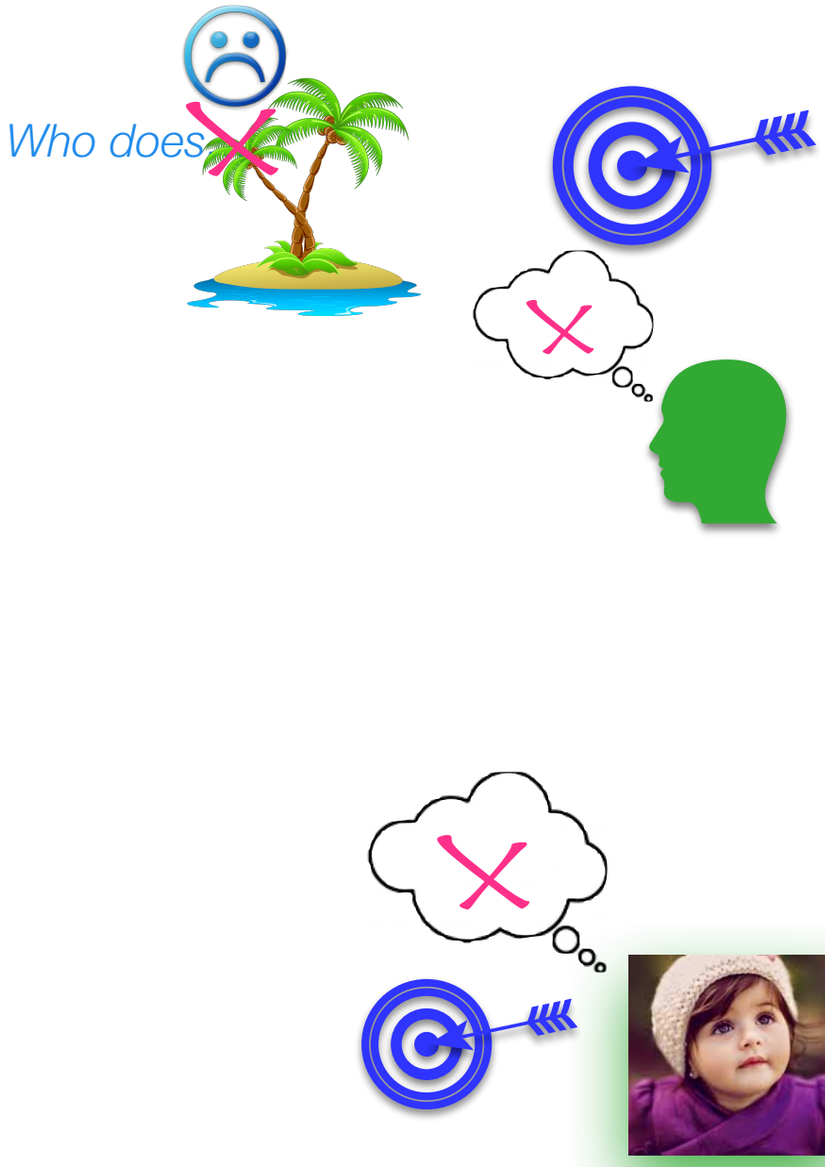
Syntactic islands

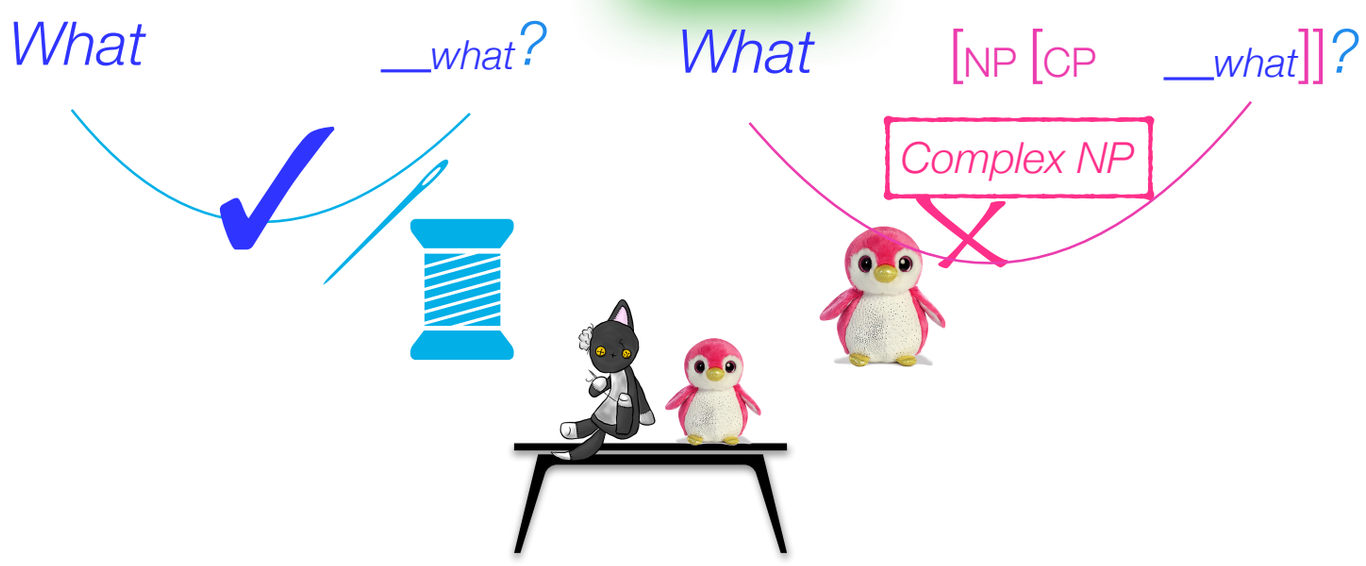
How long do children have to learn?

Syntactic islands

How long do children have to learn?

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



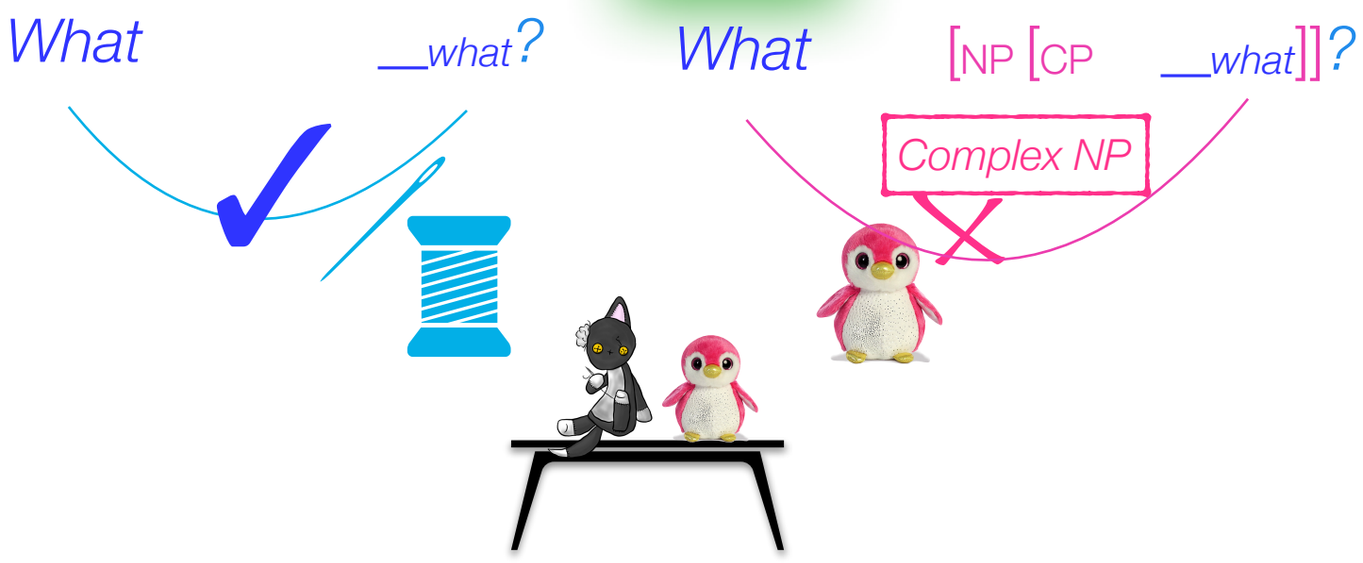


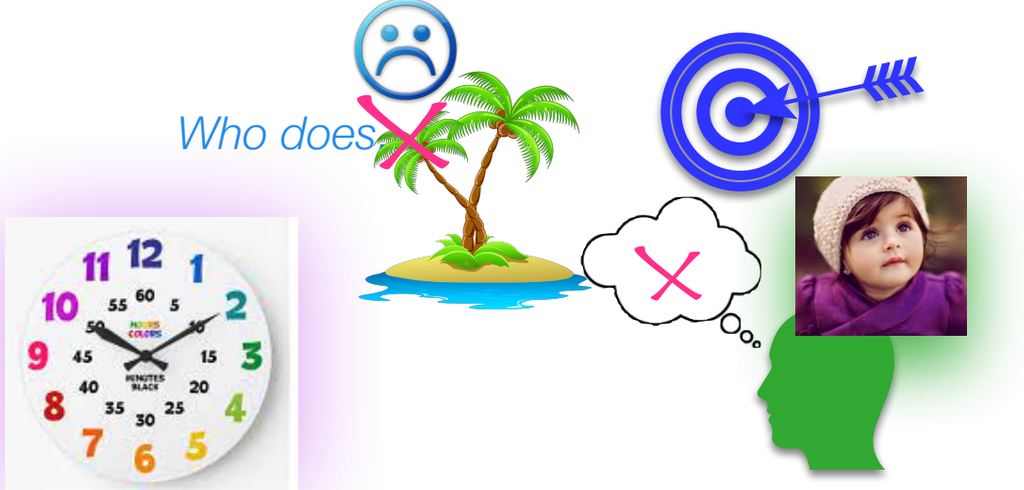
Syntactic islands

How long do children have to learn?

So input through age four.

(<60 months)



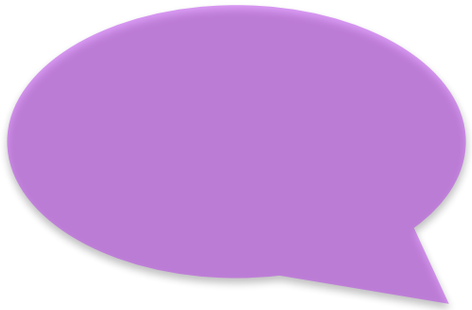


Syntactic islands

What input do children get?

Syntactic islands

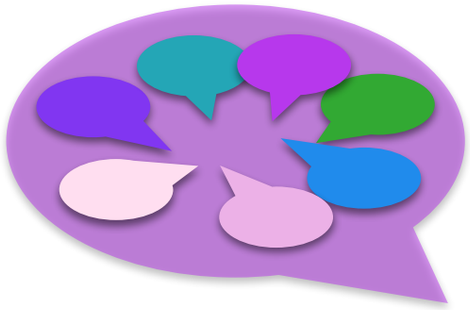
What input do children get?



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

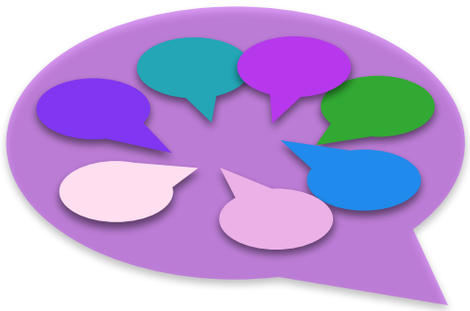
Syntactic islands

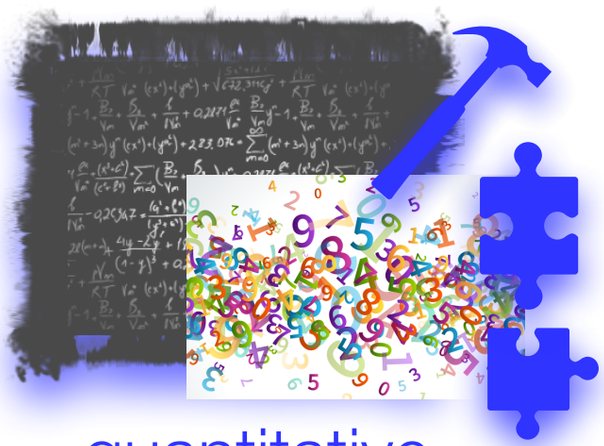
This is the acquisition problem



Syntactic islands

…which is where the theory of acquisition comes in.





quantitative

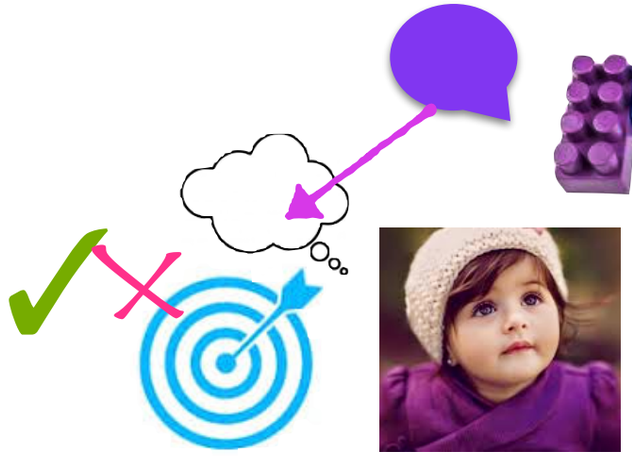
computational model

theory of acquisition

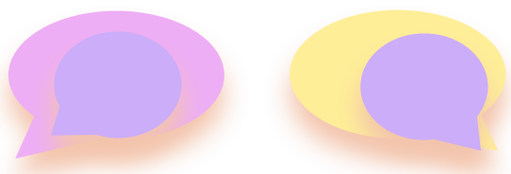
2: Evaluating a theory

Who does…

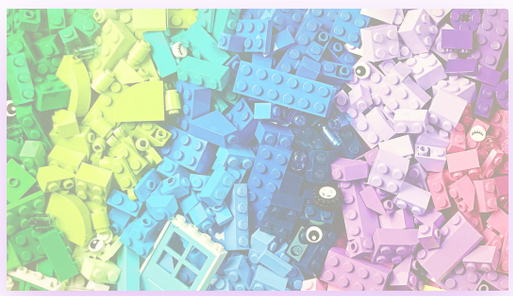
1: syntactic islands acquisition



3: dialectal variation



4: learning the building blocks



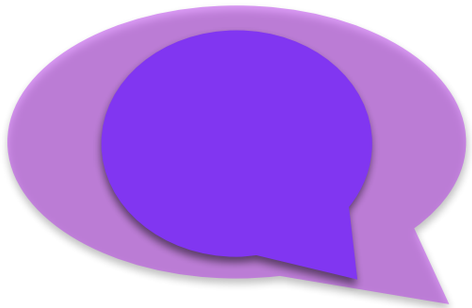
Syntactic islands:
A theory of acquisition

Intuition:

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

Pearl & Sprouse 2013

Syntactic islands:
A theory of acquisition



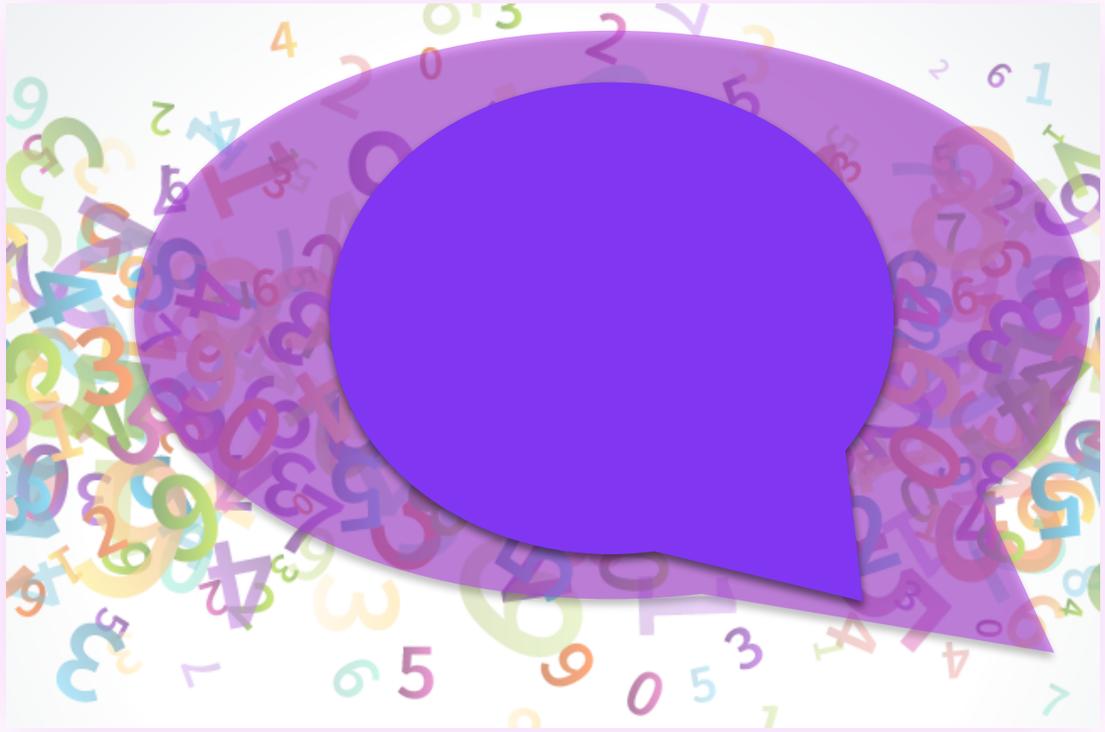
Intuition:

- 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.

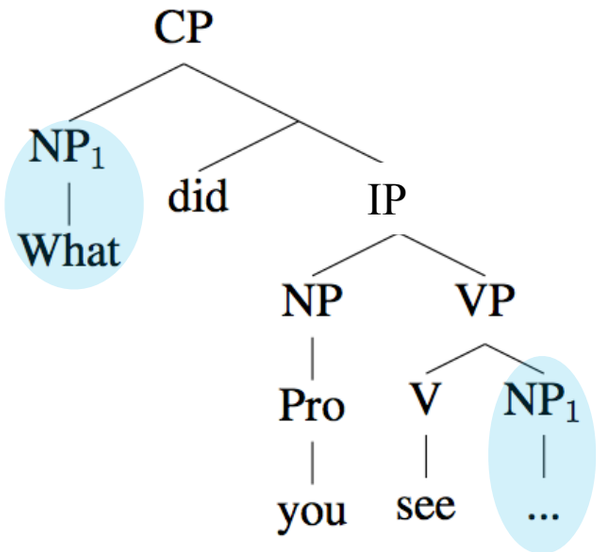


Pearl & Sprouse 2013

Syntactic islands:
A theory of acquisition

Dependencies represented as a sequence of container nodes

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



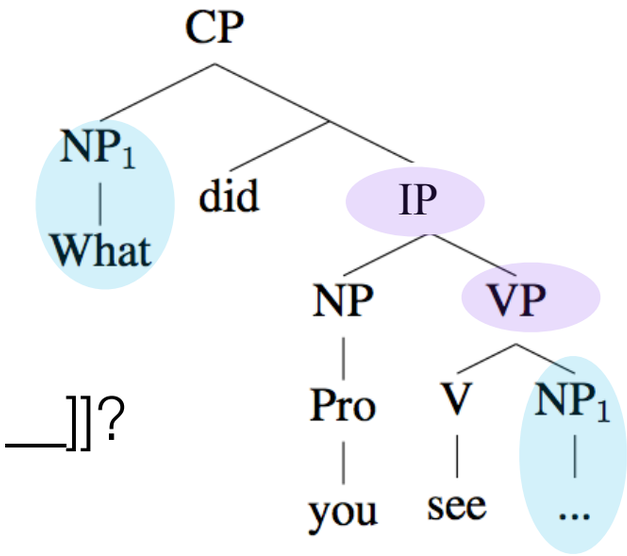
Pearl & Sprouse 2013

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*



Pearl & Sprouse 2013

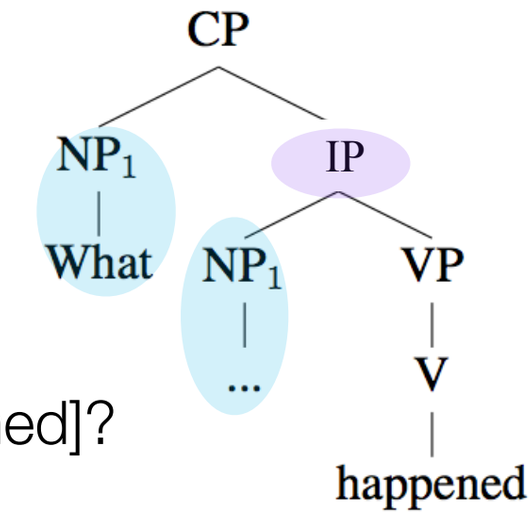
Syntactic islands:
A theory of acquisition

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*



Pearl & Sprouse 2013

Syntactic islands:
A theory of acquisition

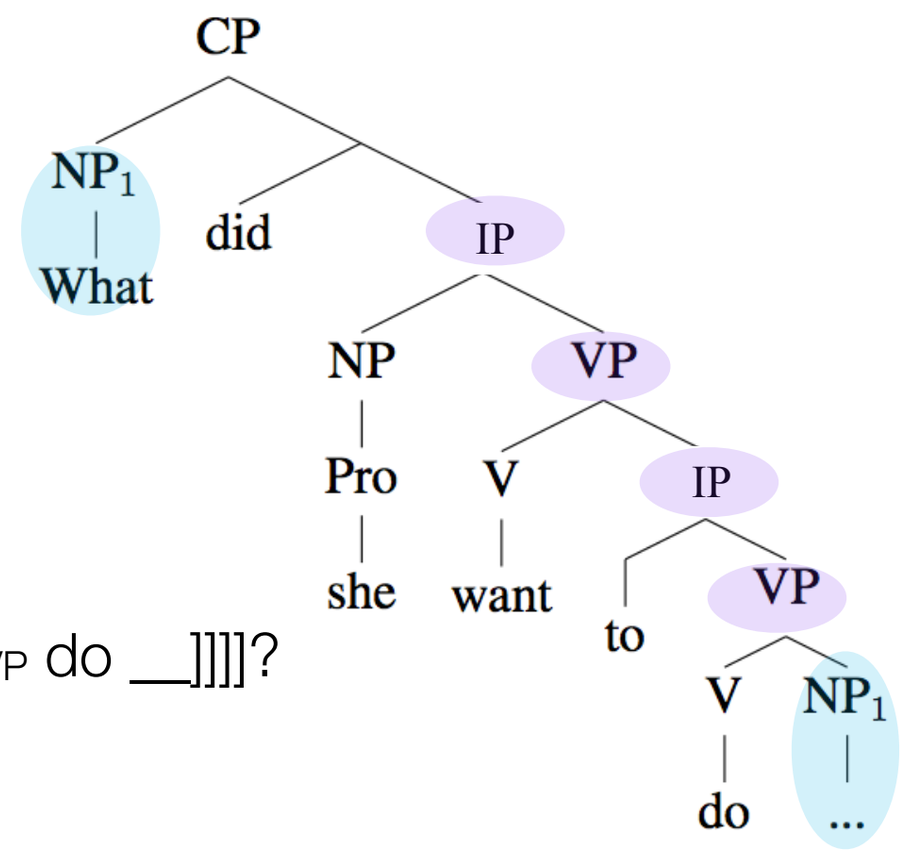
Dependencies represented as a sequence of container nodes

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 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*

(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

Syntactic islands: A theory of acquisition

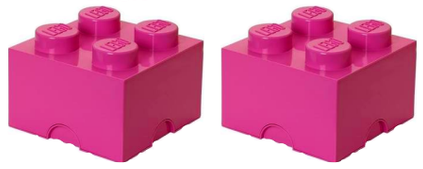
What did you see __?
= What did [IP you [VP see __]]?
= *start*-IP-VP-*end*

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*

[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*



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

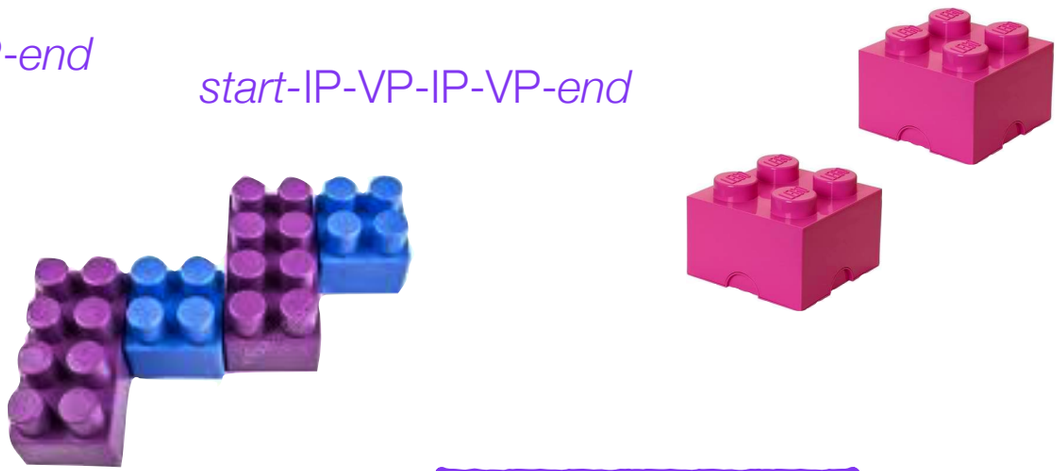
Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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



syntactic trigrams

The building blocks: trigrams of container nodes

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

syntactic trigrams

The building blocks: trigrams of container nodes

start-IP-VP-*end*

start-IP-VP

IP-VP-*end*

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

syntactic trigrams

The building blocks: trigrams of container nodes

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

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-*end*

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

syntactic trigrams

The building blocks: trigrams of container nodes

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-*end*

start-IP-*end*

start-IP-*end*

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

syntactic trigrams

The building blocks: trigrams of container nodes

IP-VP-IP

VP-IP-VP

IP-VP-*end*

start-IP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-VP

IP-VP-CP$_{that}$

VP-CP$_{that}$-IP

CP$_{that}$-IP-NP

IP-NP-PP

NP-PP-*end*

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

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

start-IP-VP

IP-VP-*end*

IP-VP-IP

VP-IP-VP

start-IP-*end*

IP-VP-CP$_{that}$

IP-VP-PP

VP-PP-*end*

VP-CP$_{that}$-IP

NP-PP-*end*

Pearl & Sprouse 2013

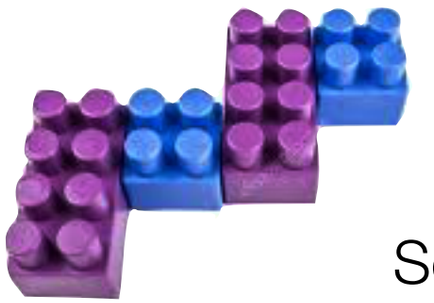
Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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



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

start-IP-VP

start-IP-*end*

IP-VP-*end*

IP-VP-CP$_{that}$

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-IP-VP

VP-CP$_{that}$-IP

NP-PP-*end*

Pearl & Sprouse 2013

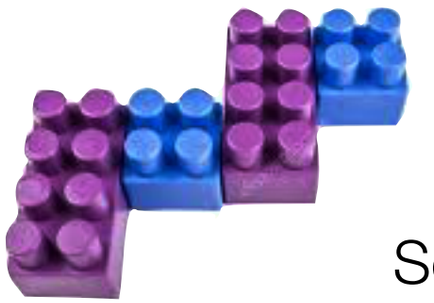
Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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



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

(And some never occur at all.)

start-IP-VP

start-IP-*end*

IP-VP-*end*

IP-VP-CP$_{that}$

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-IP-VP

NP-PP-*end*

VP-CP$_{that}$-IP

CP$_{that}$-IP-NP

IP-NP-PP

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

Relative syntactic trigram frequency:

$$= p(t) \approx \frac{\#\ trigram}{total\ \#\ trigrams}$$

start-IP-VP

start-IP-*end*

IP-VP-*end*

IP-VP-CP$_{that}$

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-IP-VP

VP-CP$_{that}$-IP

NP-PP-*end*

CP$_{that}$-IP-NP

IP-NP-PP

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

Any *wh*-dependency can then be constructed from its syntactic trigram building blocks

start-IP-VP

start-IP-*end*

IP-VP-*end*

IP-VP-CP$_{that}$

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-IP-VP

NP-PP-*end*

VP-CP$_{that}$-IP

CP$_{that}$-IP-NP

IP-NP-PP

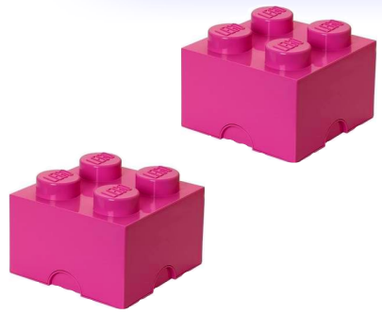
Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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



start-IP-VP-*end*

start-IP-VP

IP-VP-*end*

$$\Pi_{t \in trigrams}\, p(t)$$

start-IP-*end*

IP-VP-CP$_{that}$

IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

NP-PP-*end*

VP-CP$_{that}$-IP

CP$_{that}$-IP-NP

IP-NP-PP

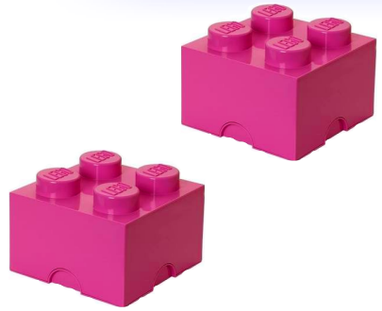
Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*



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

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-*end*

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

start-IP-*end*

IP-VP-CP$_{that}$

IP-VP-PP

VP-PP-*end*

NP-PP-*end*

VP-CP$_{that}$-IP

CP$_{that}$-IP-NP

IP-NP-PP

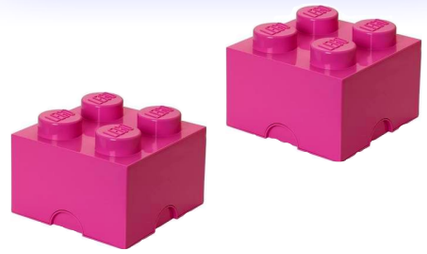
Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

start-IP-*end*

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



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

start-IP-VP

IP-VP-CP_{that}

VP-CP_{that}-IP

CP_{that}-IP-NP

IP-NP-PP

NP-PP-*end*

$$\Pi_{t \in trigrams}\, p(t)$$

start-IP-*end*

IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

Pearl & Sprouse 2013

Syntactic islands: A theory of acquisition

start-IP-VP-*end*

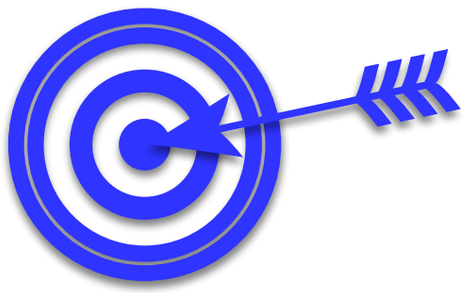
start-IP-VP-CP_{that}-IP-NP-PP-*end*

start-IP-*end*

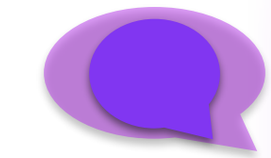
start-IP-VP-IP-VP-*end*

$$\Pi_{t \in trigrams}\, p(t)$$

A *wh*-dependency's probability can stand in for its predicted acceptability or preference.



Pearl & Sprouse 2013



Syntactic islands: A theory of acquisition

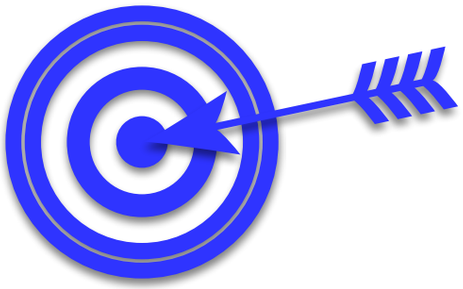
start-IP-VP-*end*

start-IP-VP-CP$_{that}$-IP-NP-PP-*end*

start-IP-*end*

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

$$\Pi_{t \in trigrams}\, p(t)$$

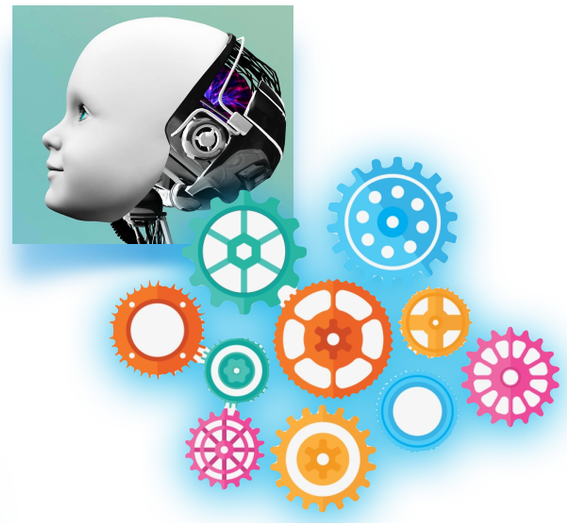


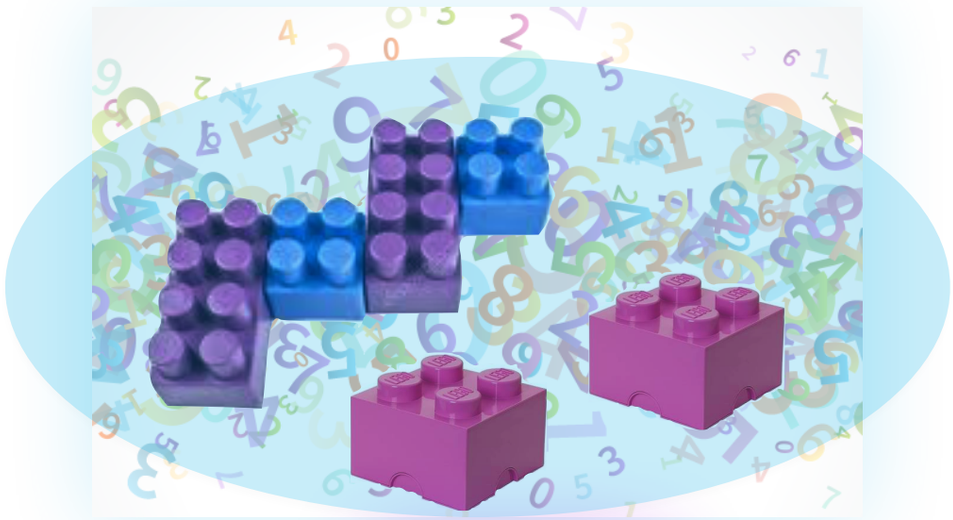
Lower probability dependencies are predicted to be less acceptable (dispreferred), compared to higher probability dependencies.

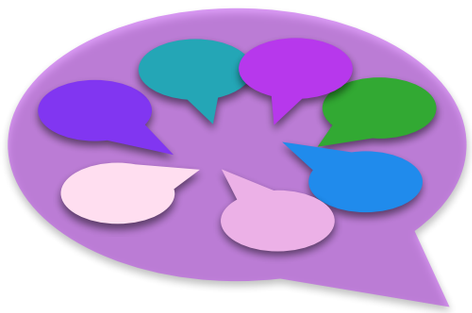
Pearl & Sprouse 2013

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?

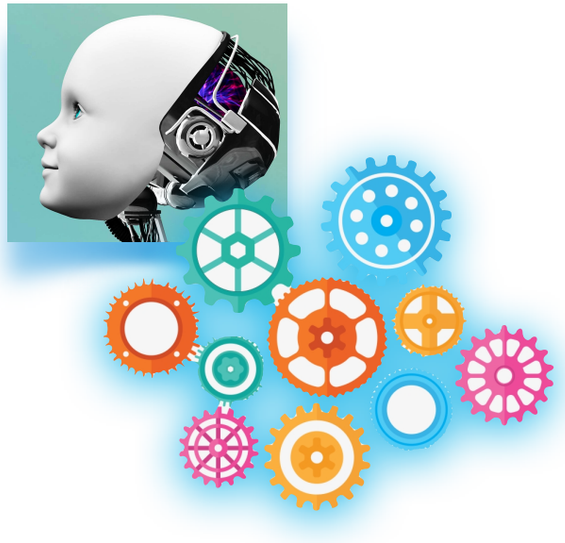


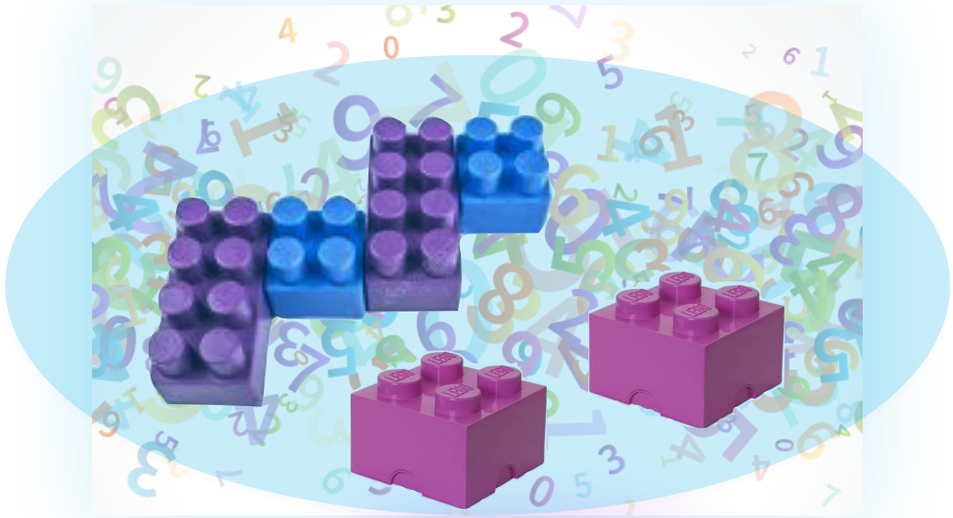


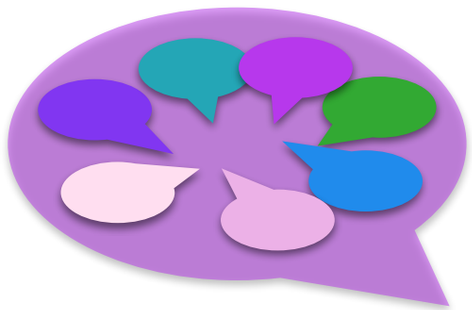


Pearl & Sprouse 2013

Evaluating the theory



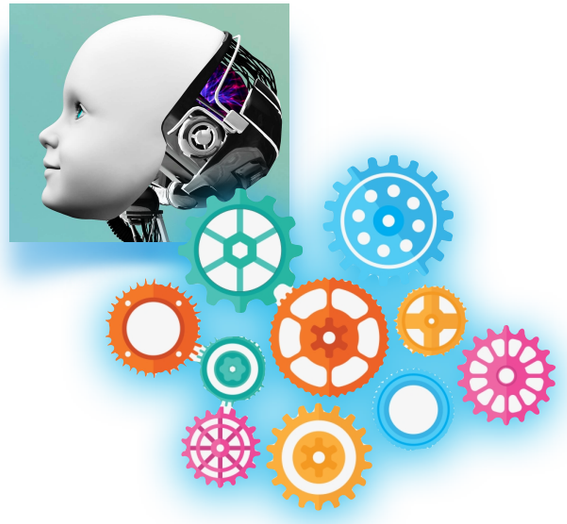


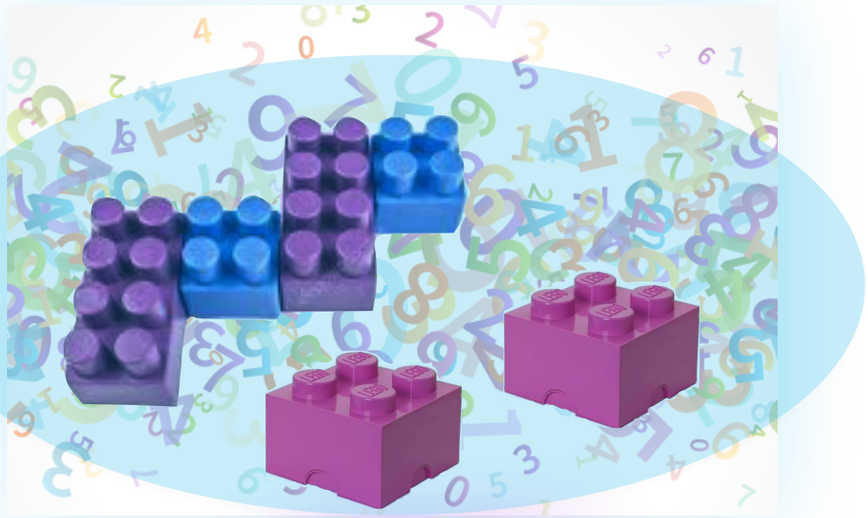


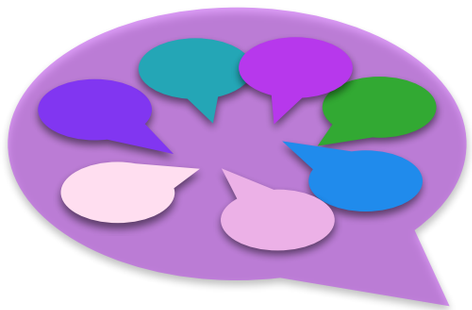
What's the input look like?

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



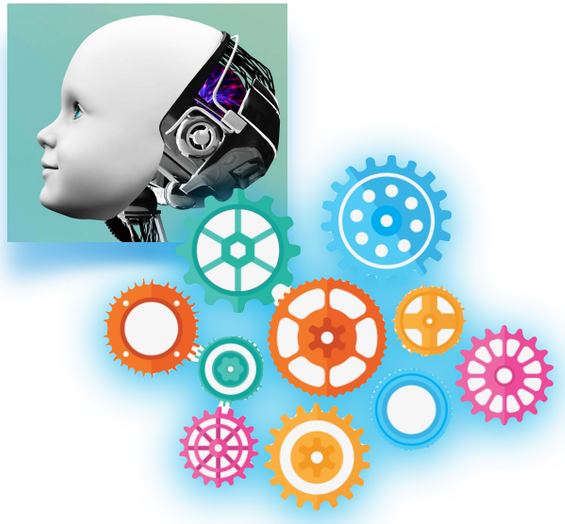


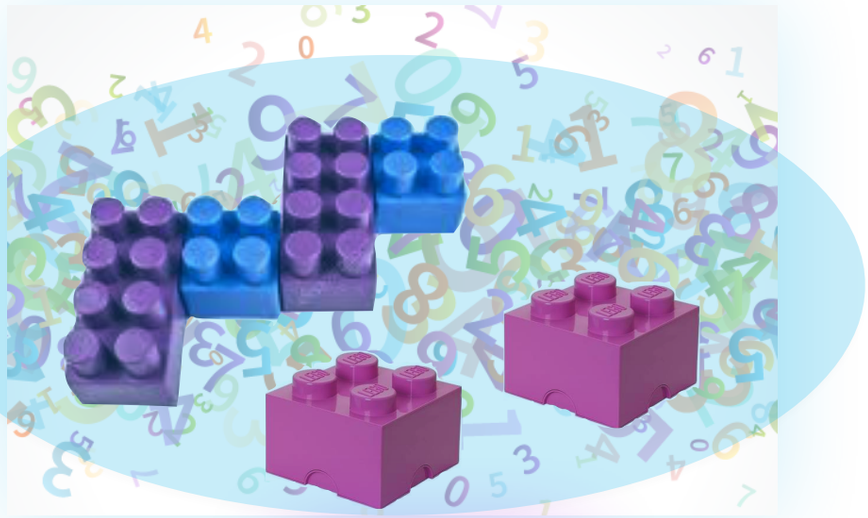


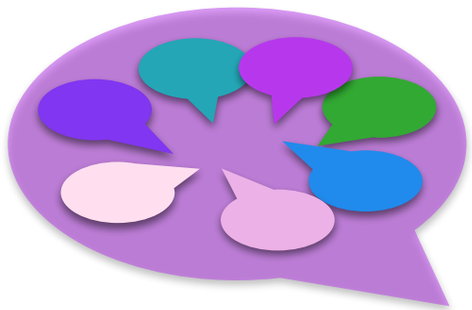
102K utterances (≈21K *wh*-dependencies) from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 children between the ages of 1 and 5 years old.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

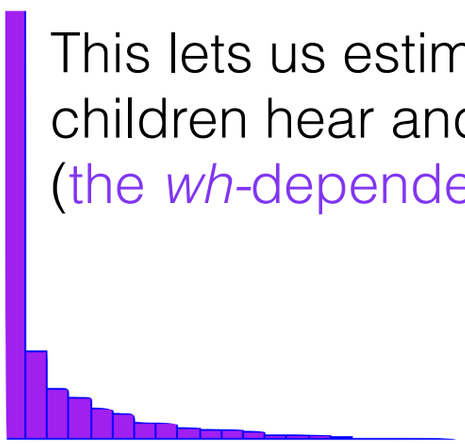
Evaluating the theory





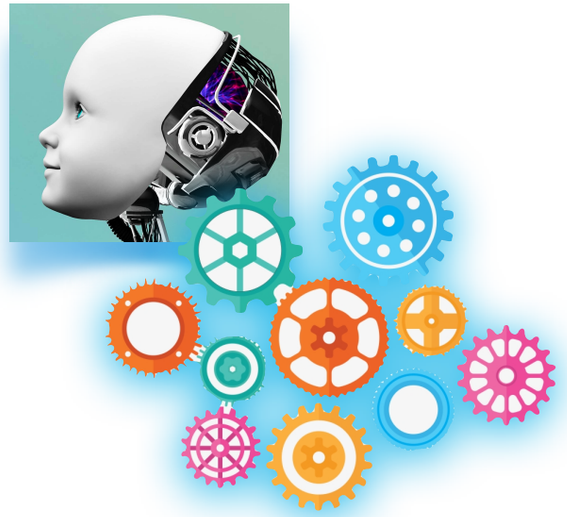


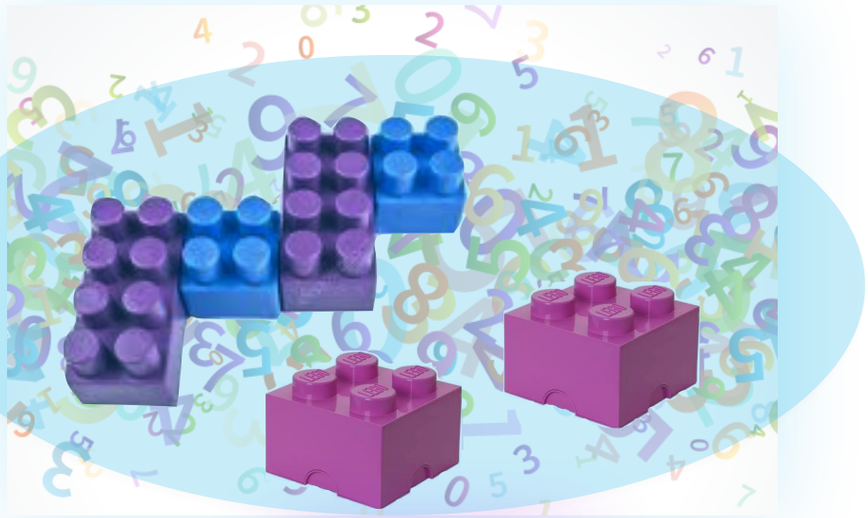
This lets us estimate which *wh*-dependencies children hear and how often they hear them (the *wh*-dependency distribution).

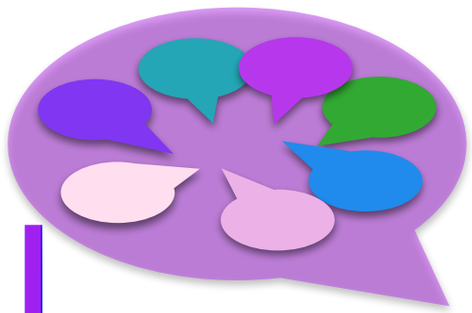


Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

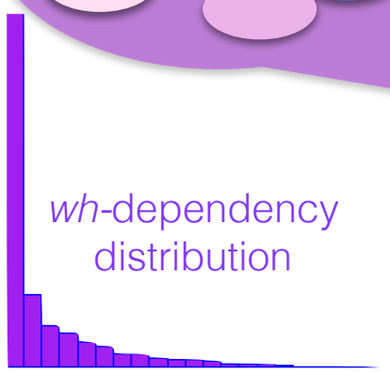
Evaluating the theory







wh-dependency distribution



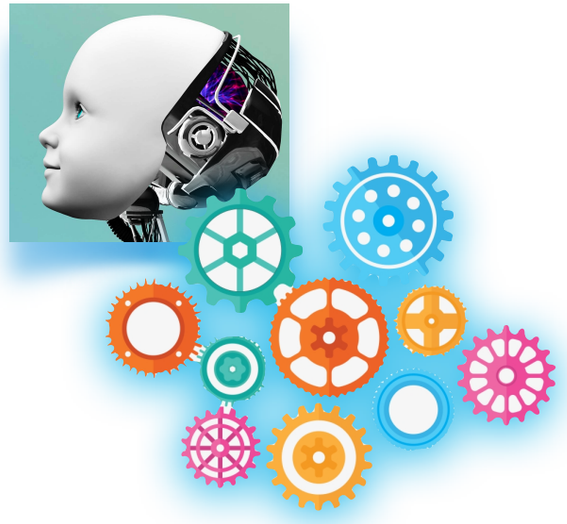
#

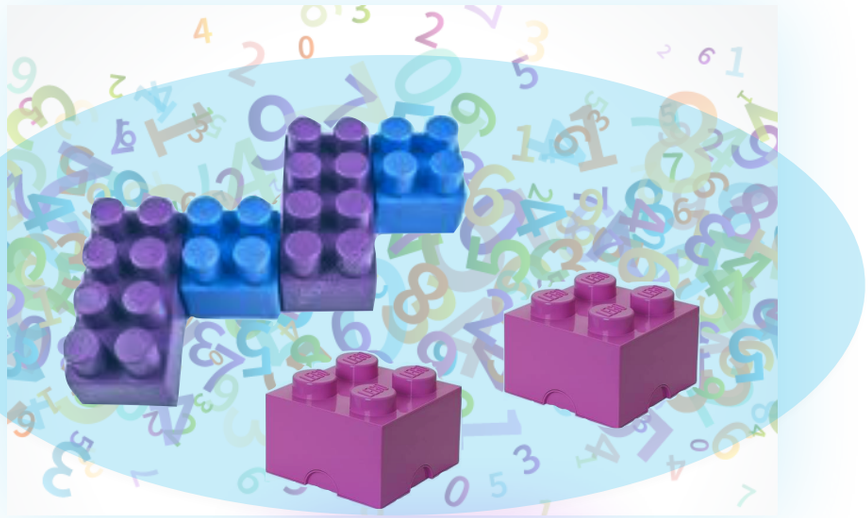
We can then estimate how many *wh*-dependencies children hear during the learning period.

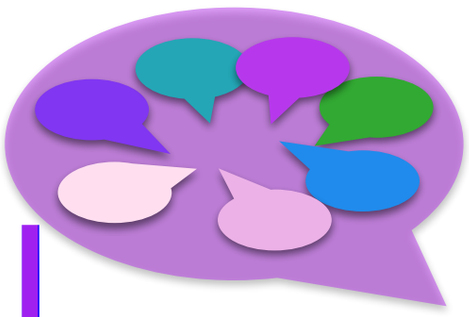
(<60 months)

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory

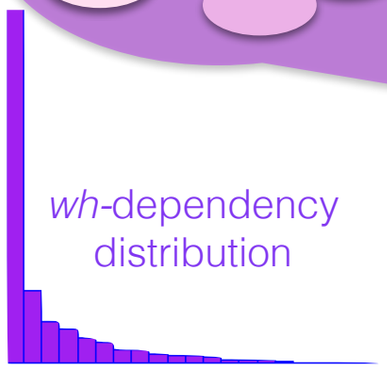






#

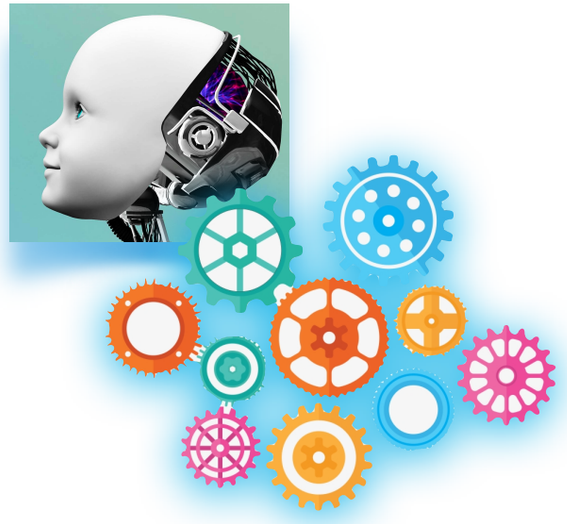
(<60 months)

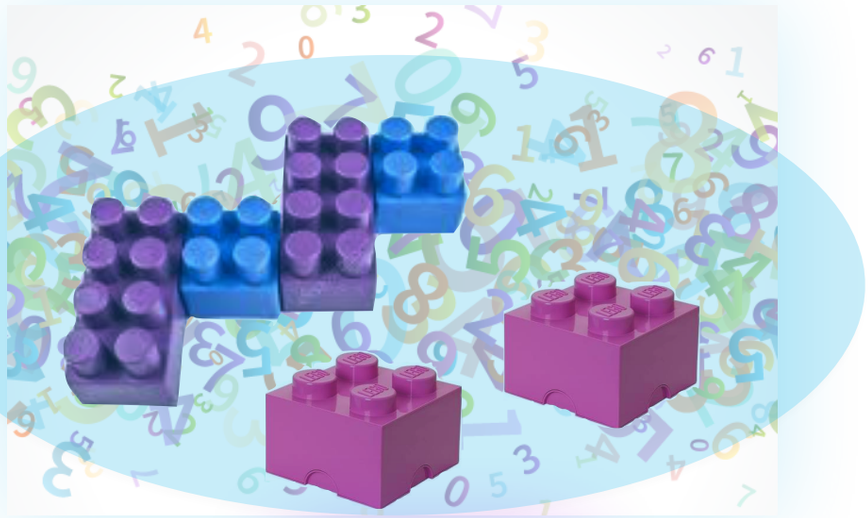


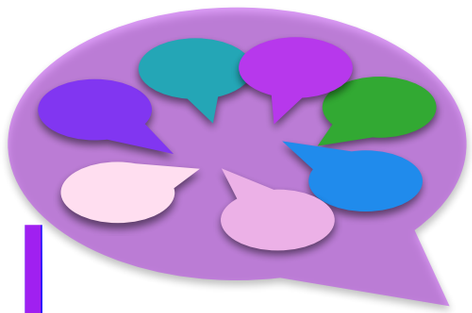
Children begin to represent the full structure of *wh*-dependencies (e.g., *wh*-questions and relative clauses) around 20 months: Seidl et al. 2003, Gagliardi et al. 2016, Perkins & Lidz 2020.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory

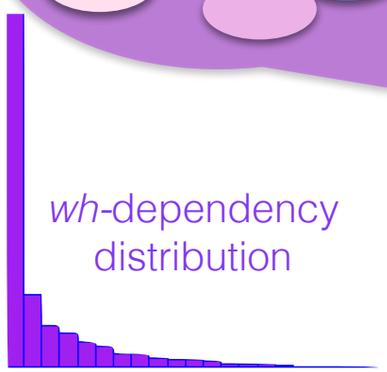




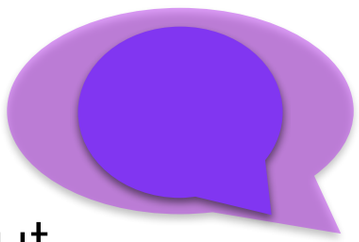


#

wh-dependency distribution



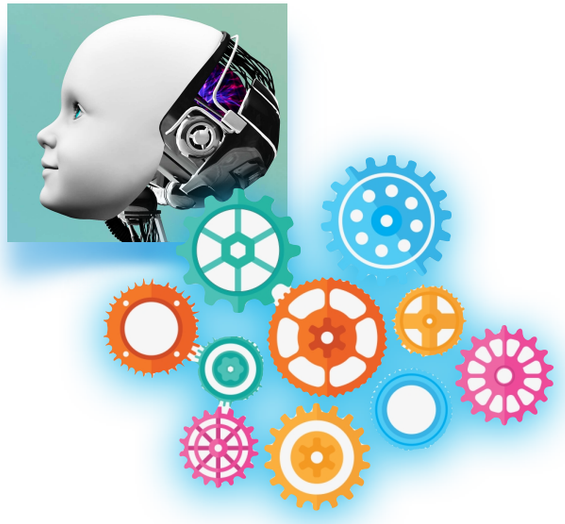
(20 months ≤ age < 60 months)

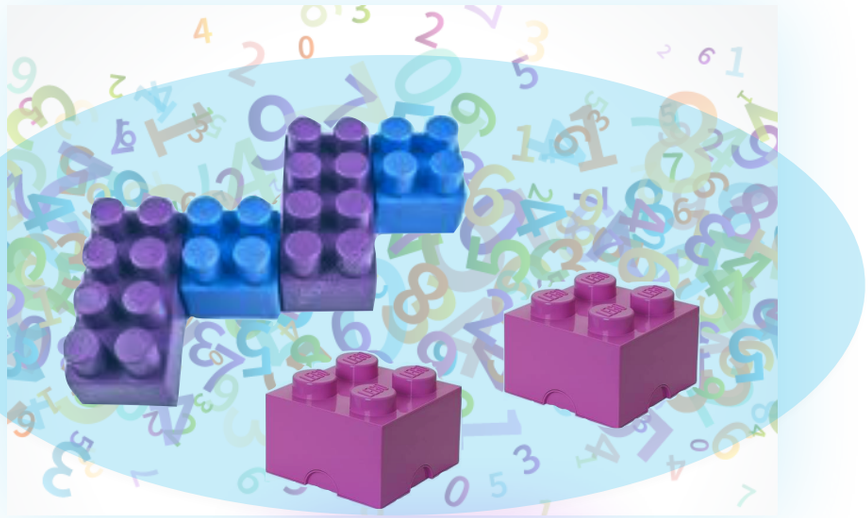


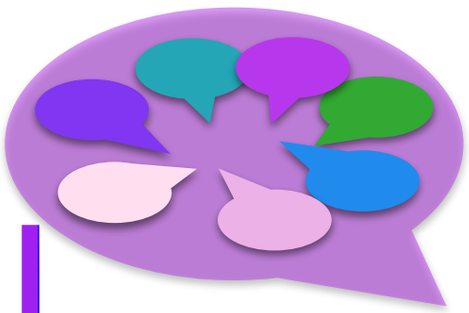
Educated guess: This is when children can start processing *wh*-dependencies reliably from their input.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory

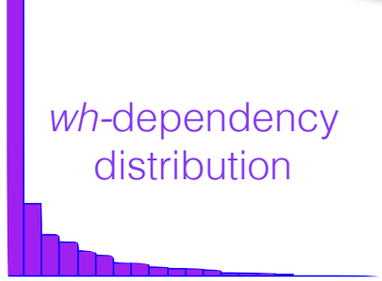


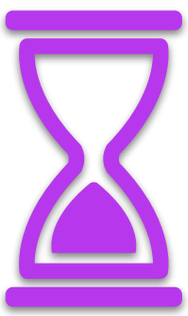




#

(20 months ≤ age < 60 months)

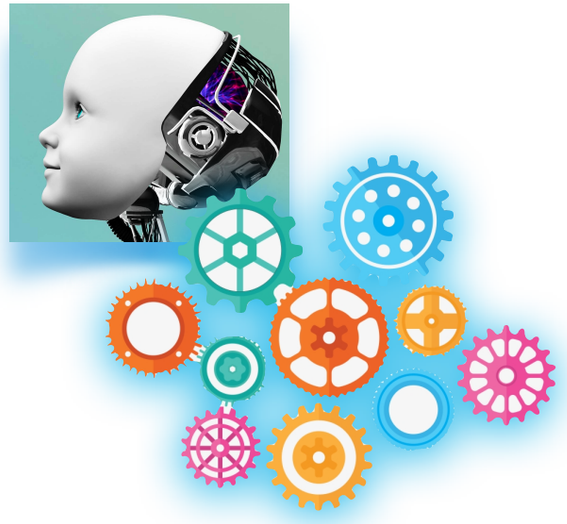


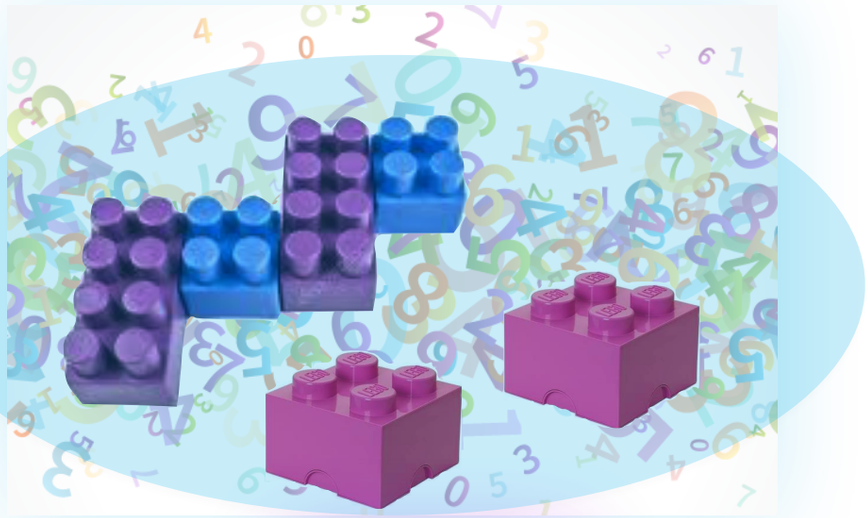


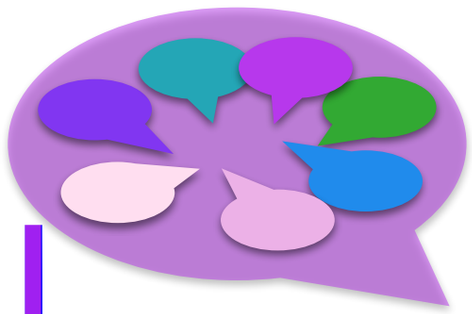
How many minutes is this? In particular, children are awake for only a certain portion of the day at different ages (Davis et al. 2004).

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory







wh-dependency distribution

#

(20 months ≤ age < 60 months)

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

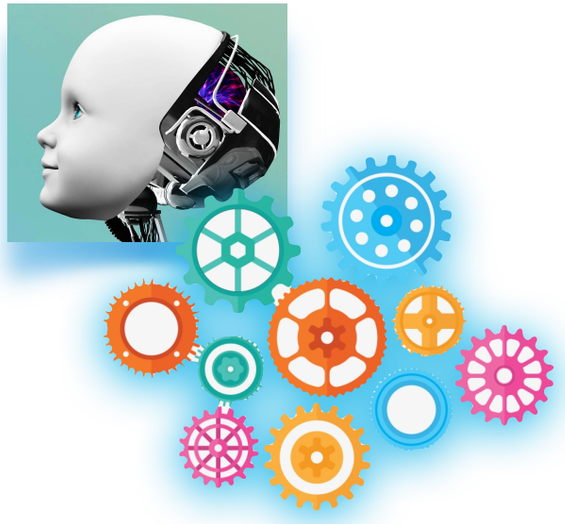
cumulative waking mins

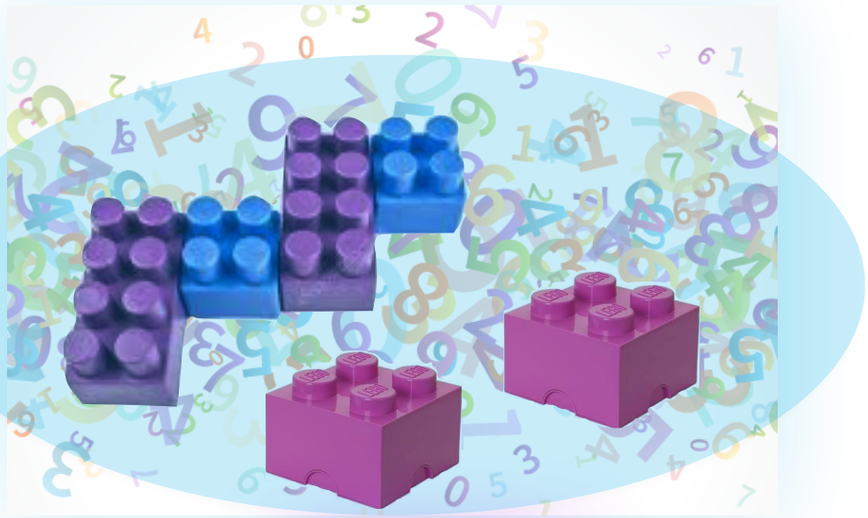
14174.17 * 60 min/hour

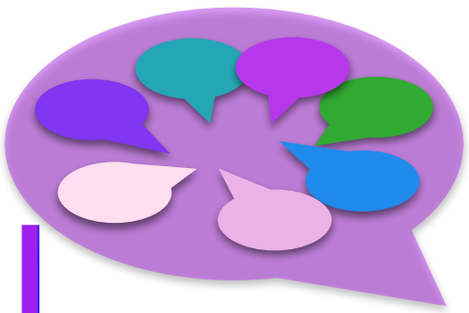
How many minutes 850450.2

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory







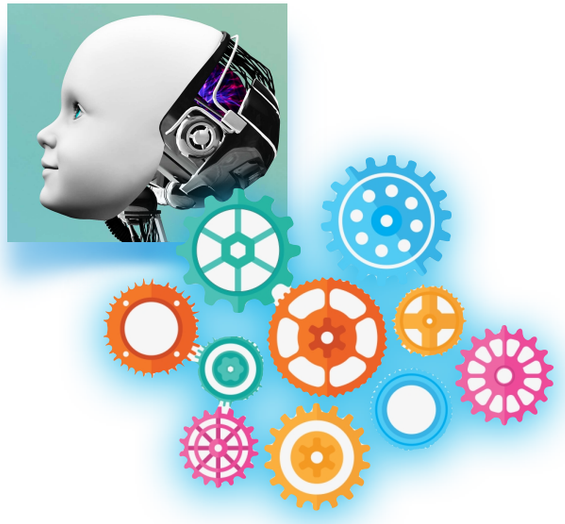
wh-dependency distribution

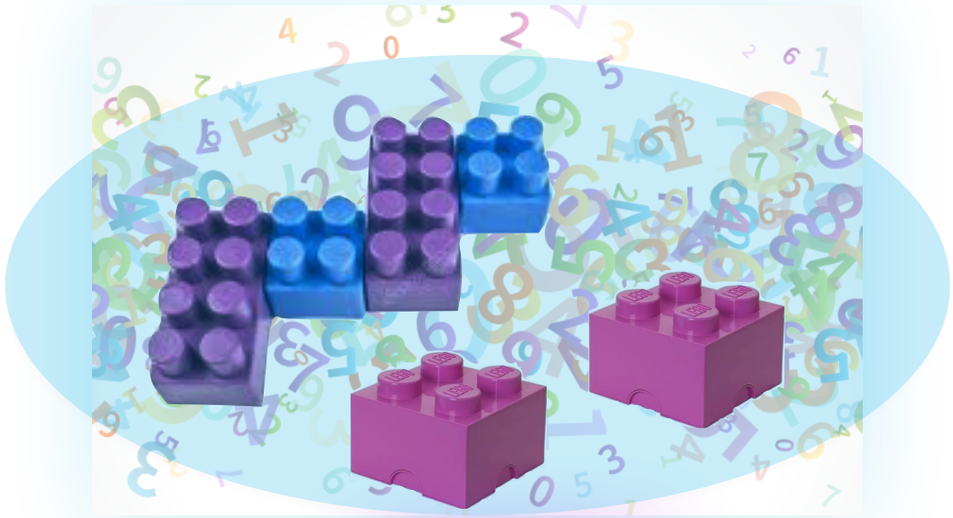
(≈850450 minutes)

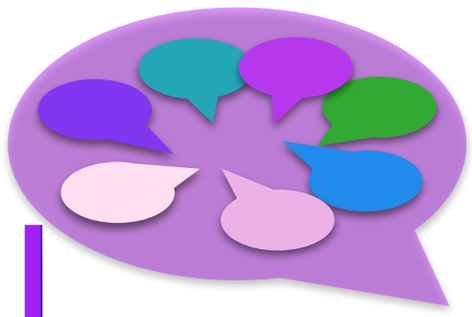
How many *wh*-dependencies is this? #

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory







wh-dependency distribution

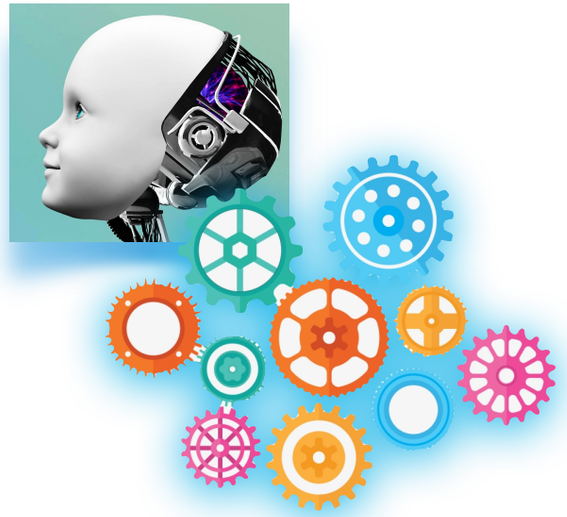
(≈850450 minutes)

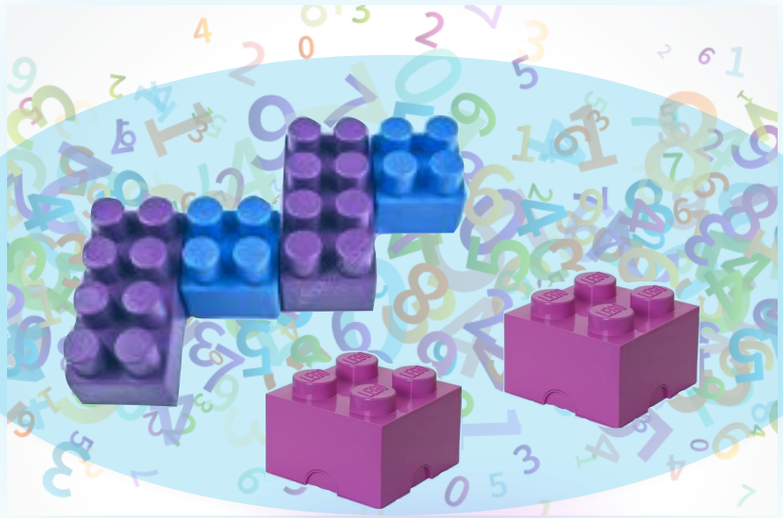
Hoff-Ginsberg (1998) and Rowe (2012): Estimates of utterances per minute in speech directed at children from different backgrounds.

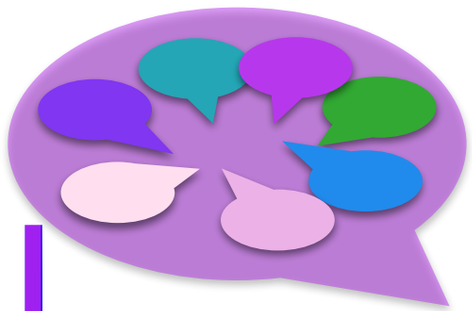
#

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory







wh-dependency distribution

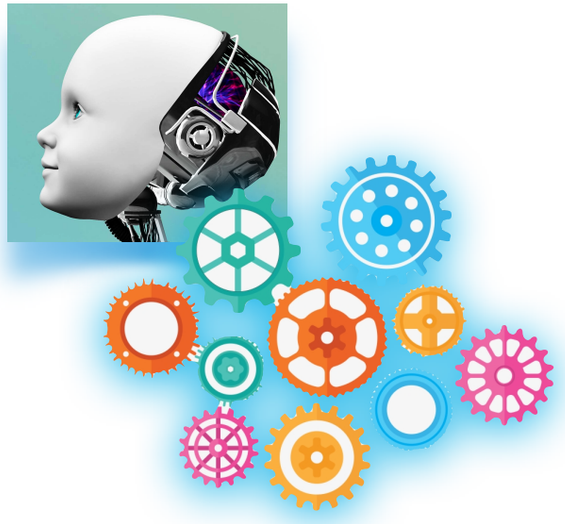
(≈850450 minutes)

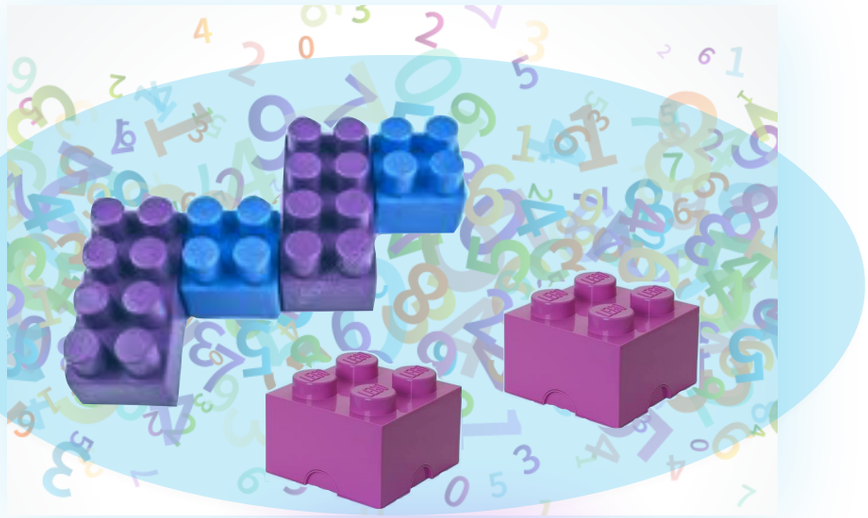
#

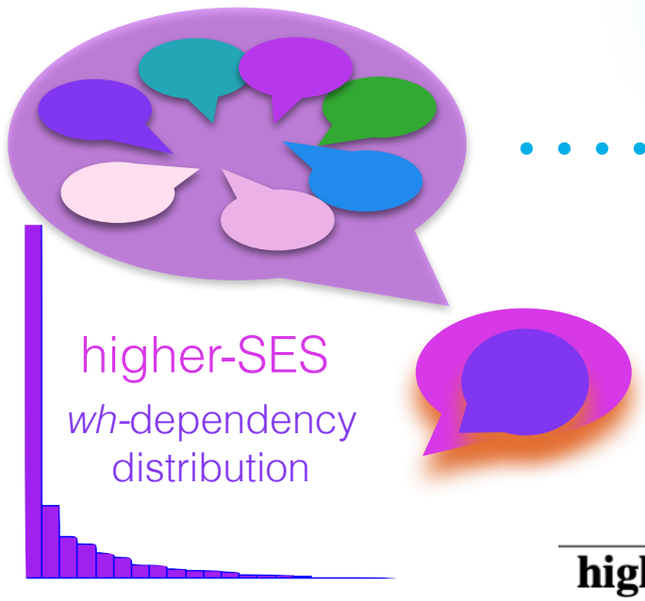
Data from children from higher socio-economic status (SES) backgrounds.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory





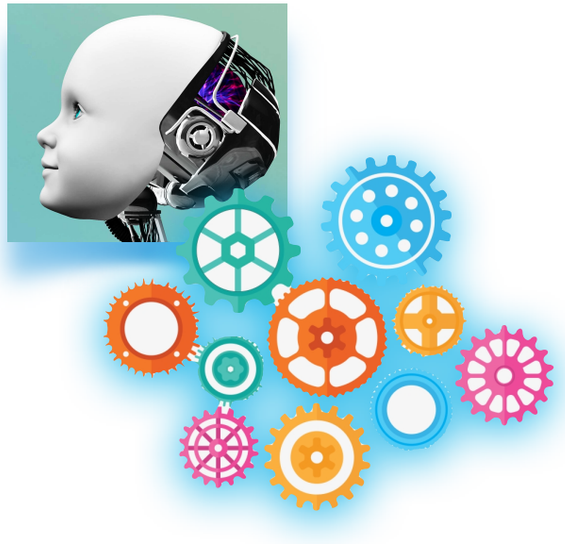


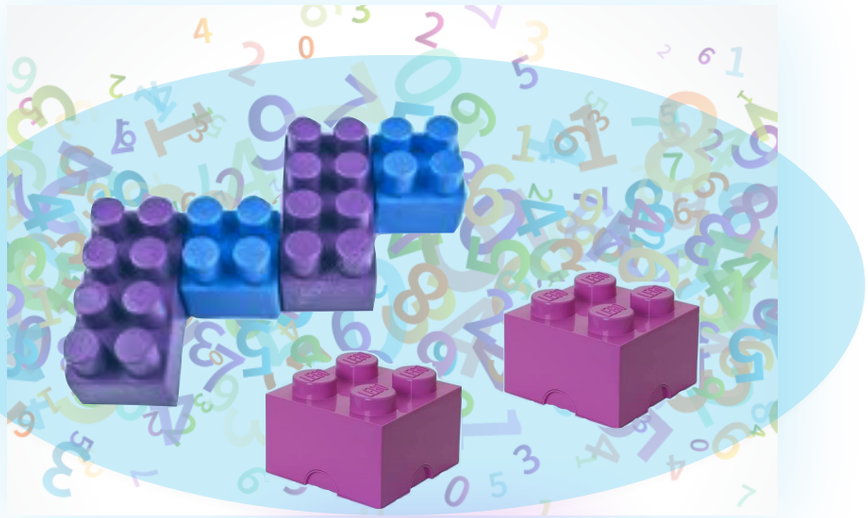
from our own corpus samples: rate of wh-dependencies/ utterance

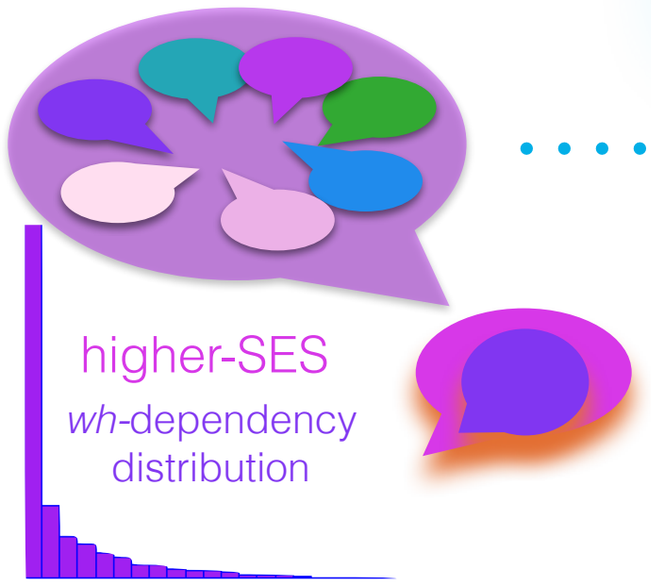
	utt/min	*	min	*	*wh*-dep/utt	=	total *wh*-dep
higher-SES		*	850,450.2	*	20,932/101,838		
- 2 s.d.	7.4					=	1,293,545
- 1 s.d.	11.6		#			=	2,027,719
average	15.8					=	2,761,893
+ 1 s.d.	20.0					=	3,496,067
+ 2 s.d.	24.2					=	4,230,241

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory







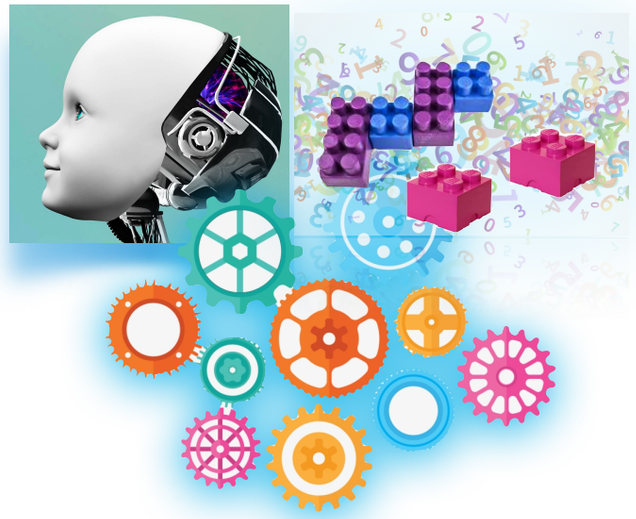
#

≈1.3 million - 4.2 million
wh-dependencies

Can the modeled learner produce the appropriate observable behavior?

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



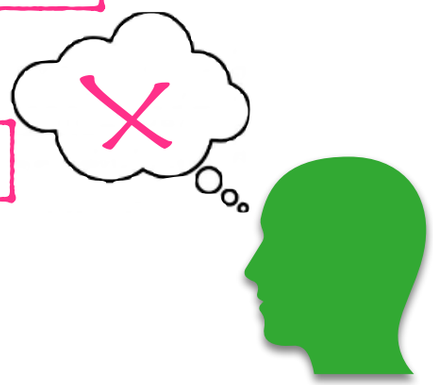
Reminder: Target behavior

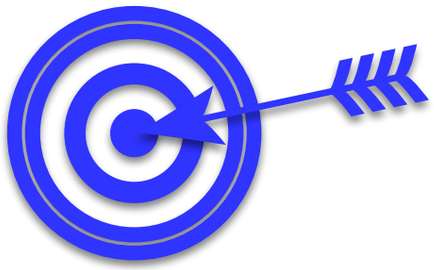
Subject island

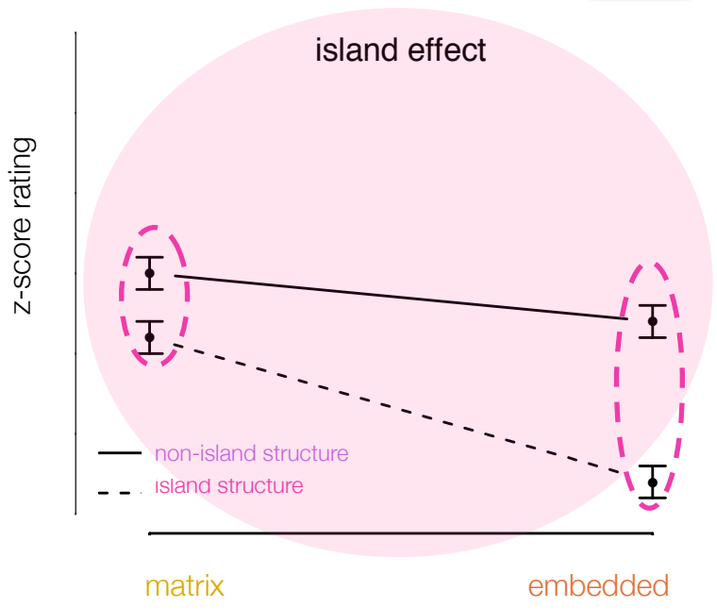
Complex NP island

Whether island

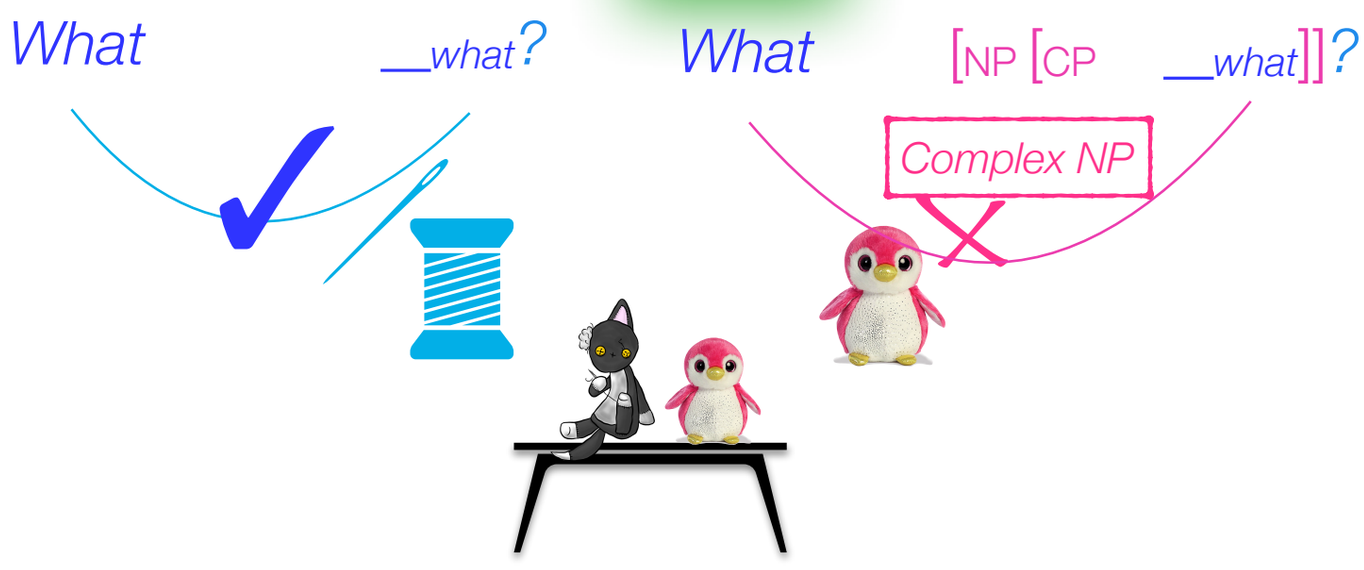
Adjunct island







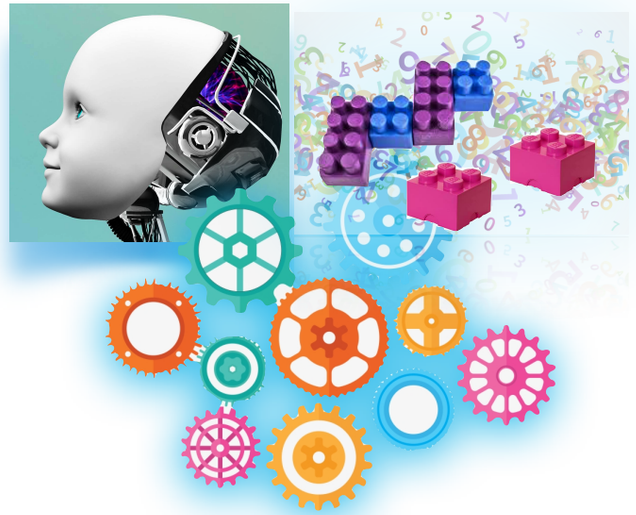
Sprouse et al. 2012



De Villiers et al. 2008

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



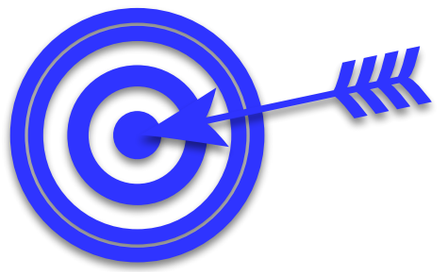
Reminder: Target behavior

Subject island

Complex NP island

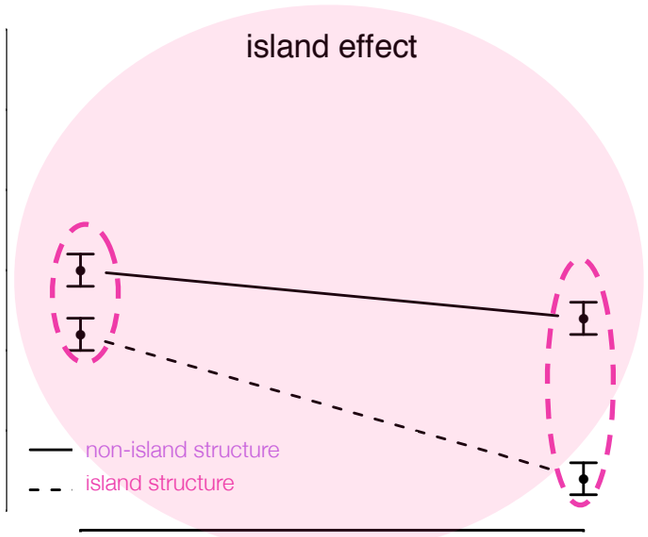
Whether island

Adjunct island



Who [non-island]?

Who [island]?



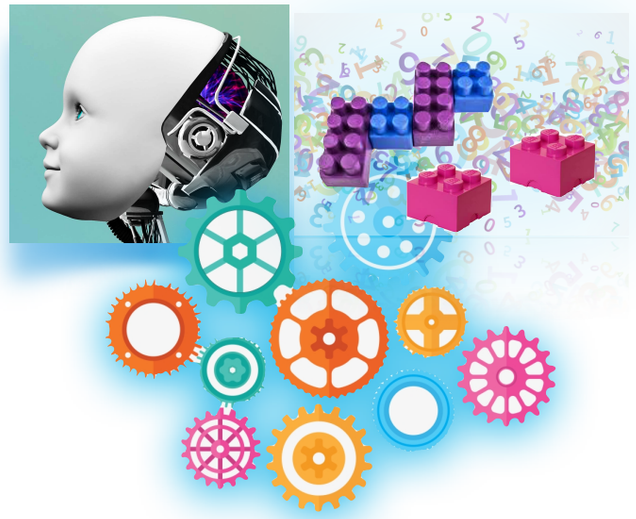
Who *__who*?

Who [CP… *__who*]?

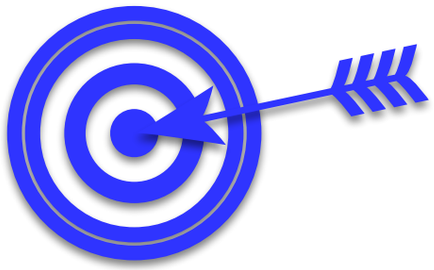
Looking for superadditivity in selected judgments as the sign of syntactic islands knowledge

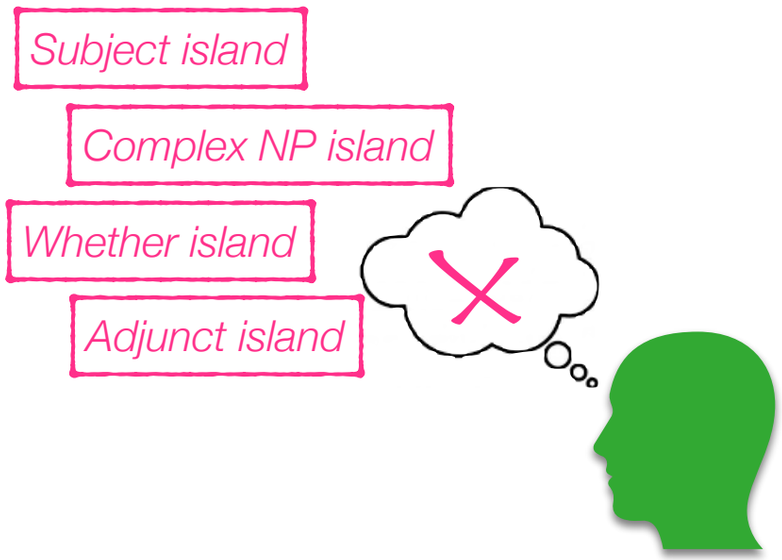
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



Reminder: Target behavior





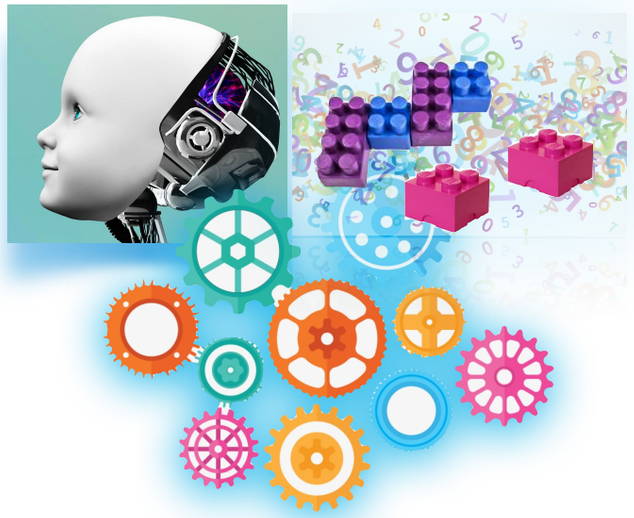
Each set of island stimuli from Sprouse et al. 2012…

Complex NP island stimuli

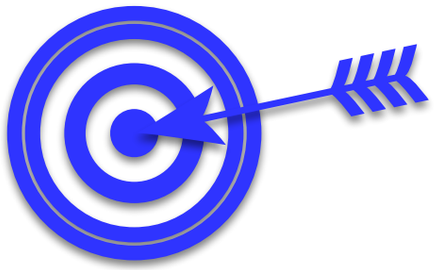
Who __ claimed [that Lily forgot the necklace]?	matrix | non-island
What did the teacher claim [that Lily forgot __]?	embedded | non-island
Who __ made [the claim that Lily forgot the necklace]?	matrix | island
*What did the teacher make [the claim that Lily forgot __]?	embedded | island

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



Reminder: Target behavior

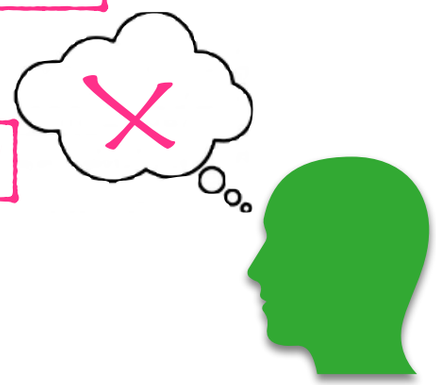


Subject island

Complex NP island

Whether island

Adjunct island



Each *wh*-dependency from the island stimuli of Sprouse et al. 2012

- can be transformed into container node sequences

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

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory

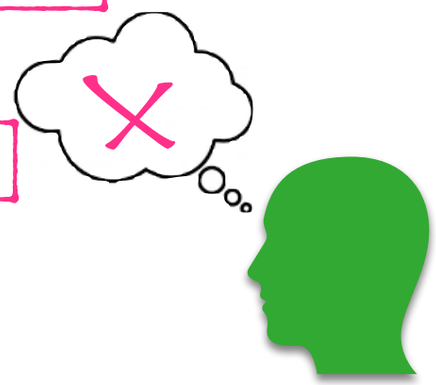
Reminder: Target behavior

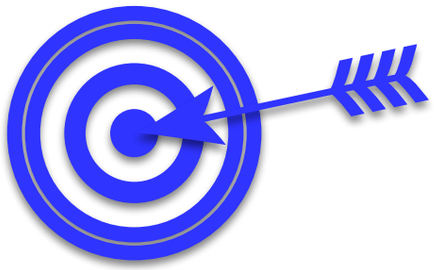
Subject island

Complex NP island

Whether island

Adjunct island





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

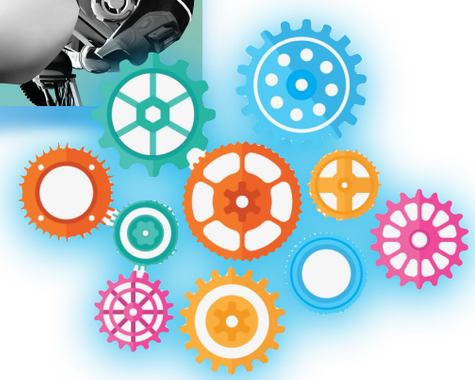
$$\prod_{t \in trigrams} p(t)$$

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

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



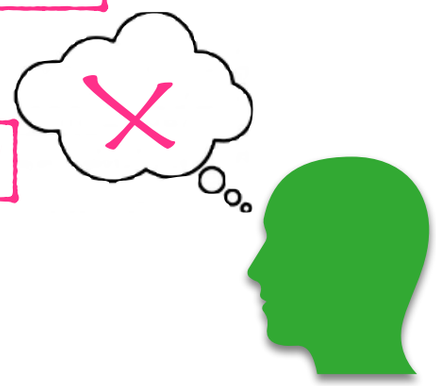
Reminder: Target behavior

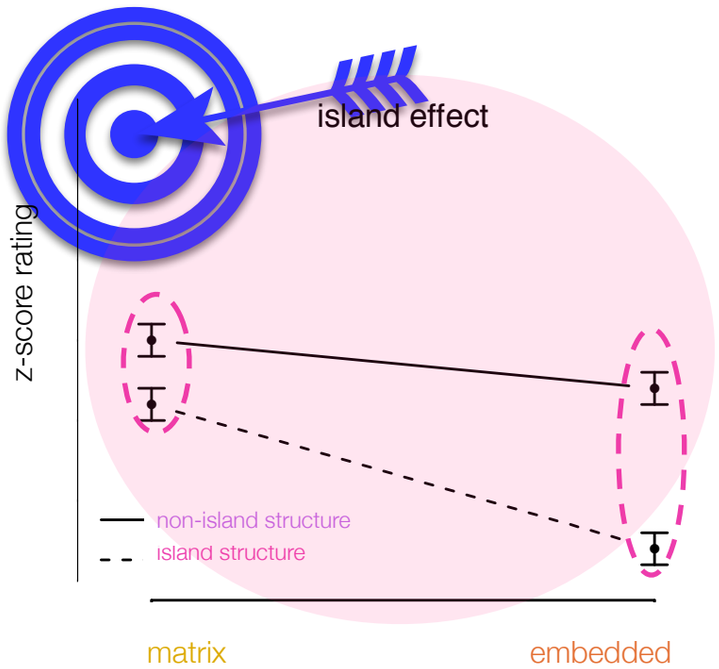
Subject island

Complex NP island

Whether island

Adjunct island





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

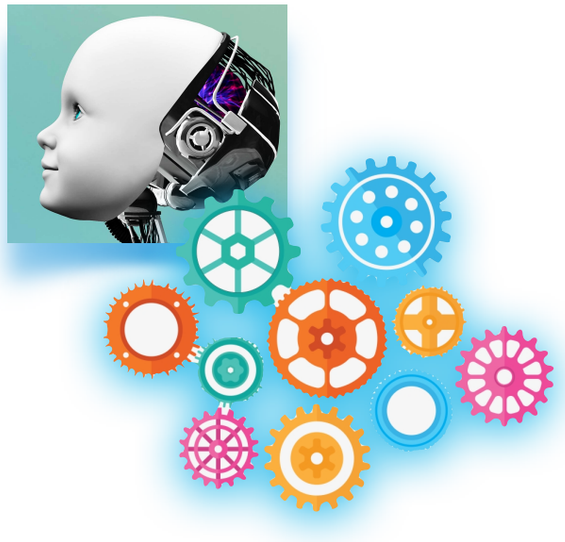
Complex NP island stimuli

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

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

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory



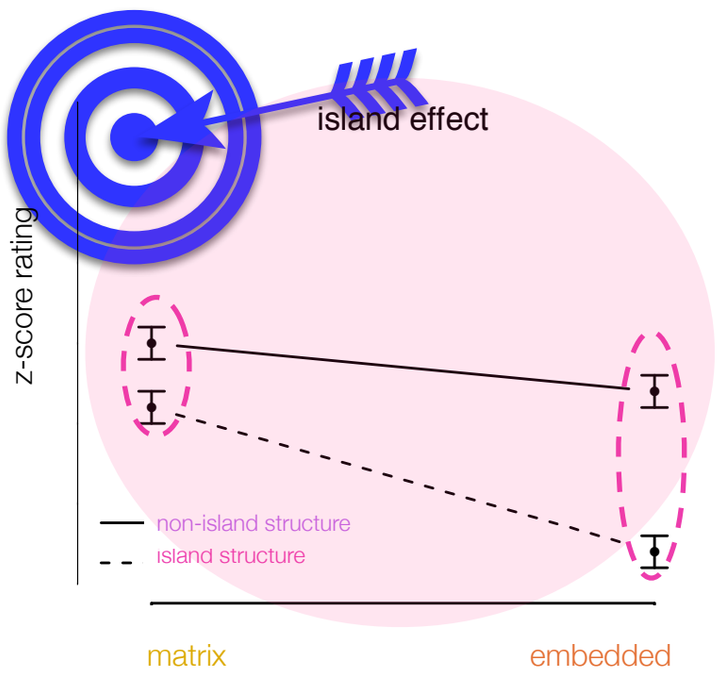
Reminder: Target behavior

Subject island

Complex NP island

Whether island

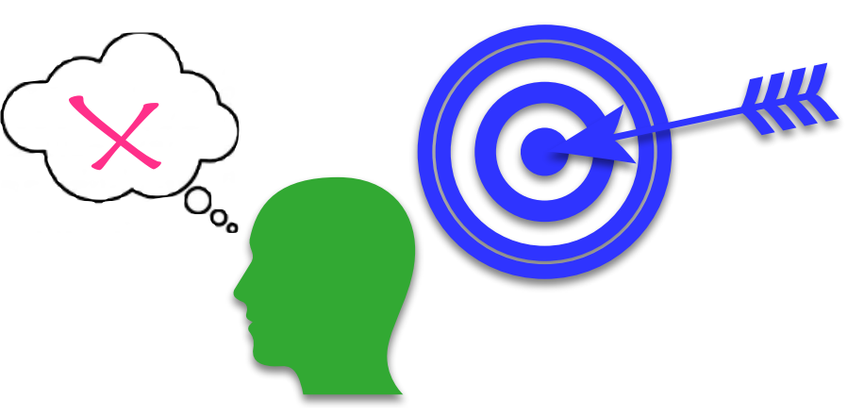
Adjunct island



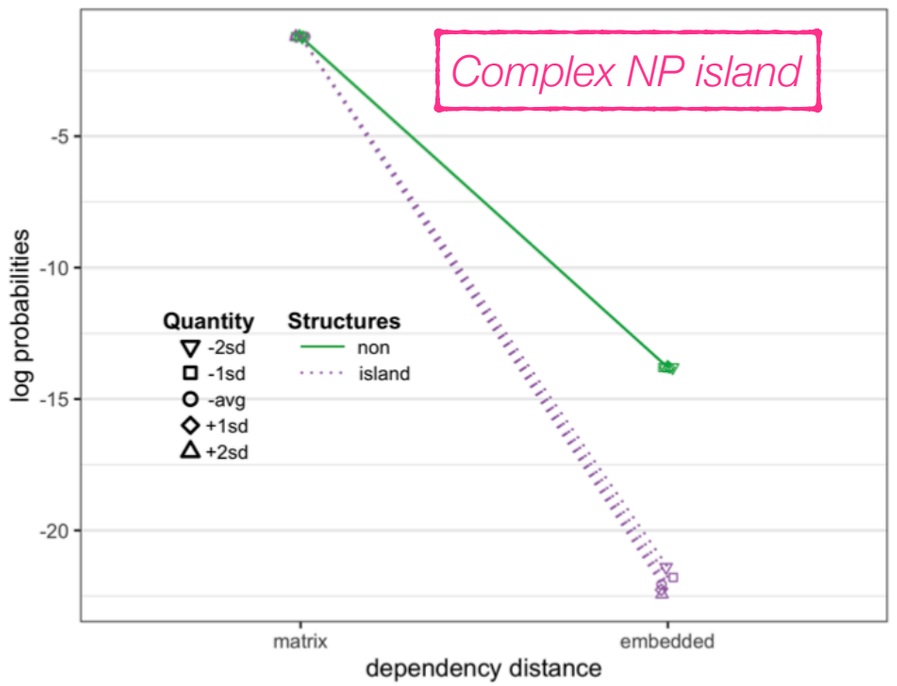
If so, then we predict the modeled child has syntactic island knowledge that allows the same judgment pattern as adults, learned from the building blocks in children's input.

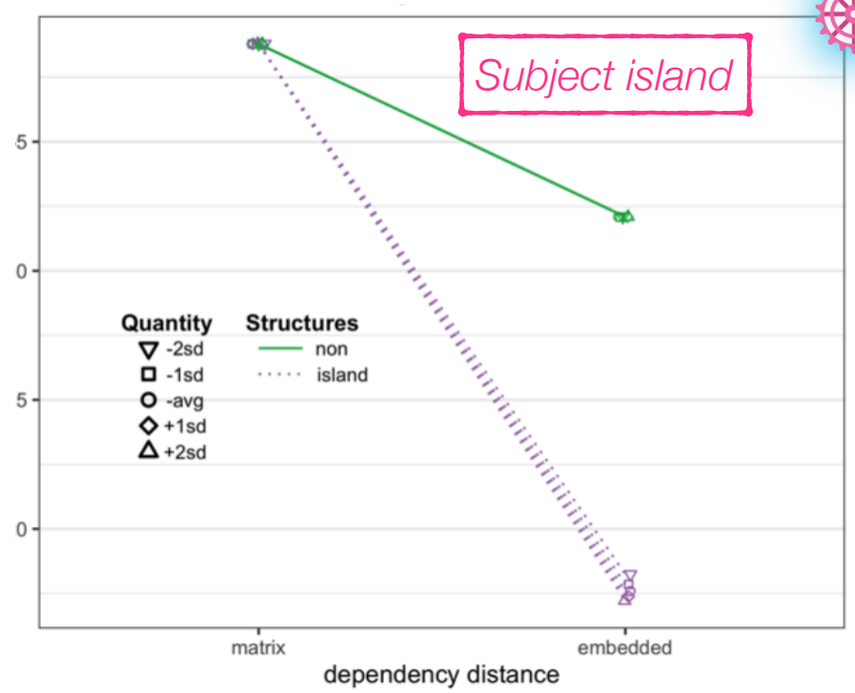
$$\Pi_{t \in trigrams}\, p(t)$$

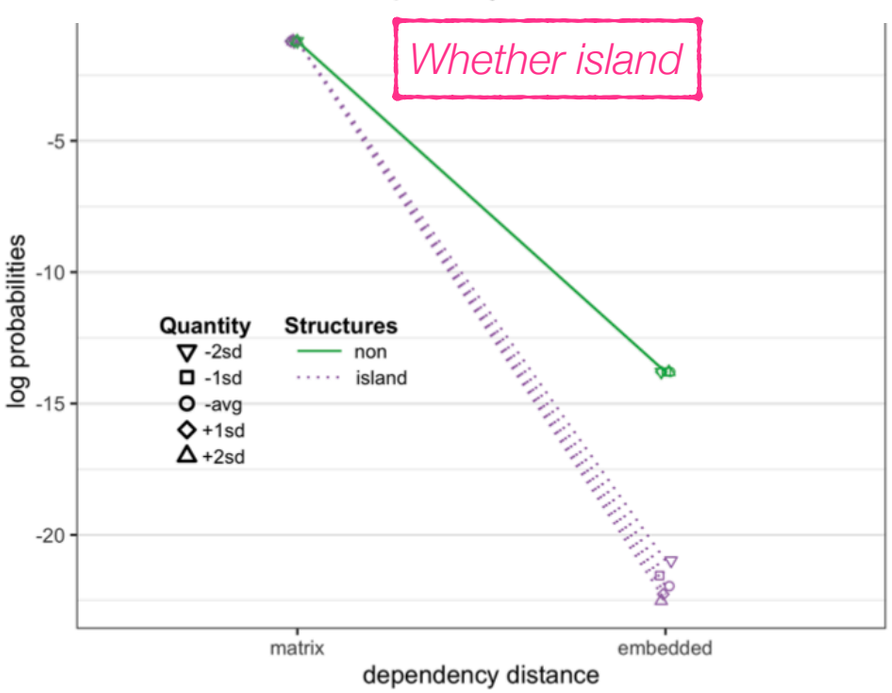
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

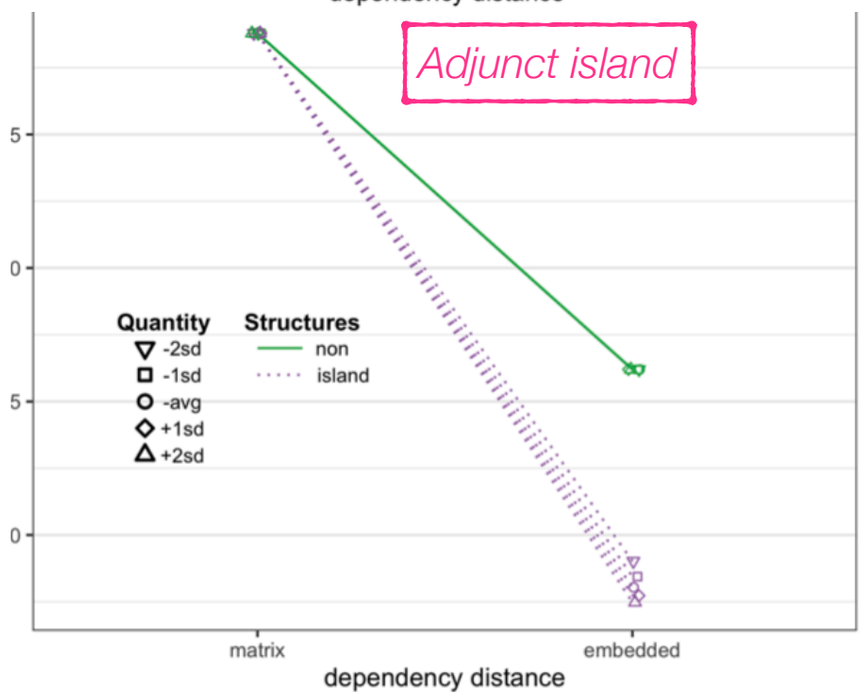


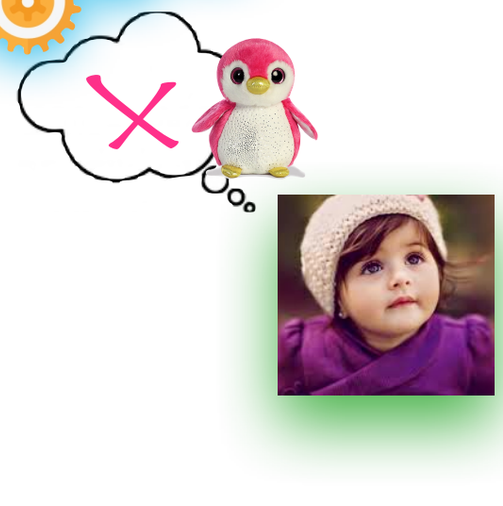
Evaluating the theory



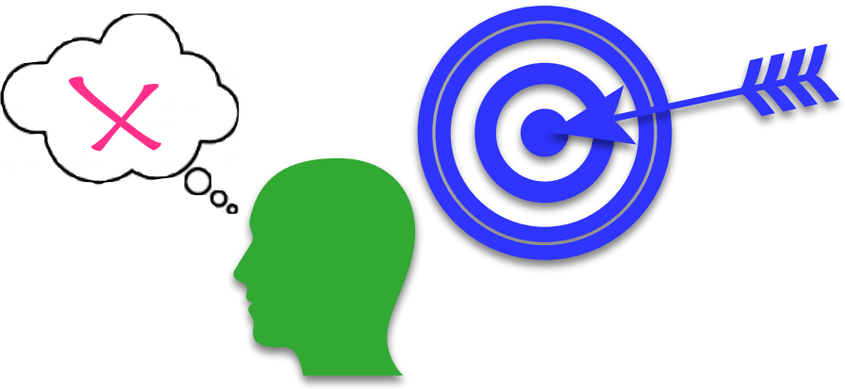




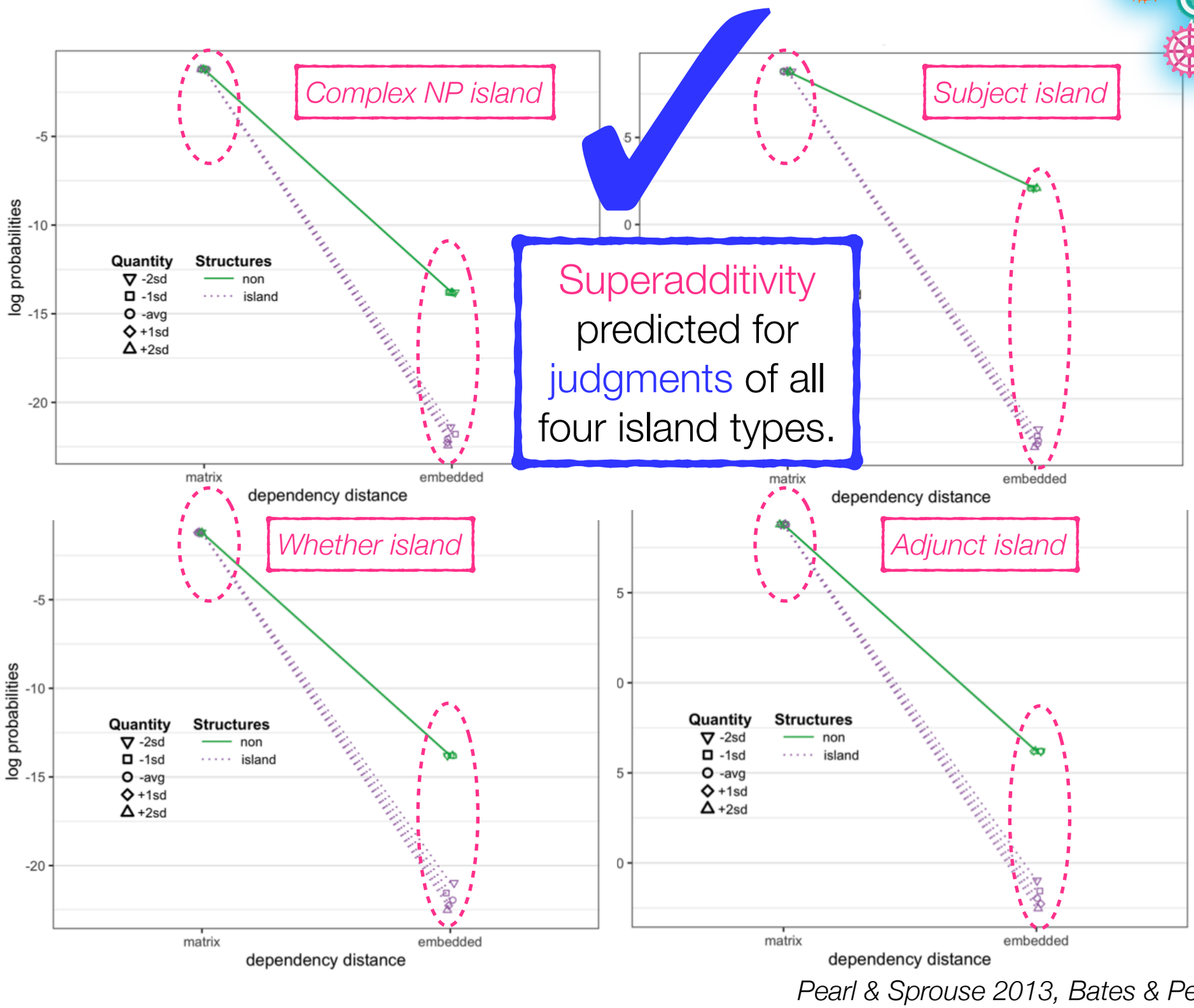




Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

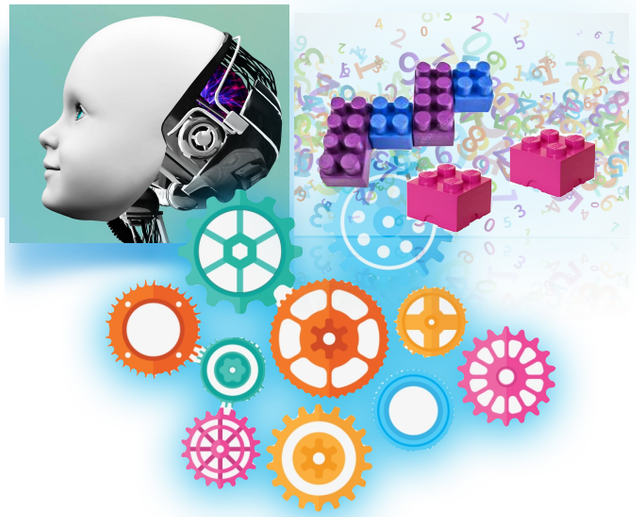


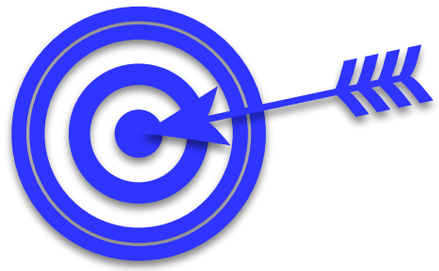
Evaluating the theory



Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory





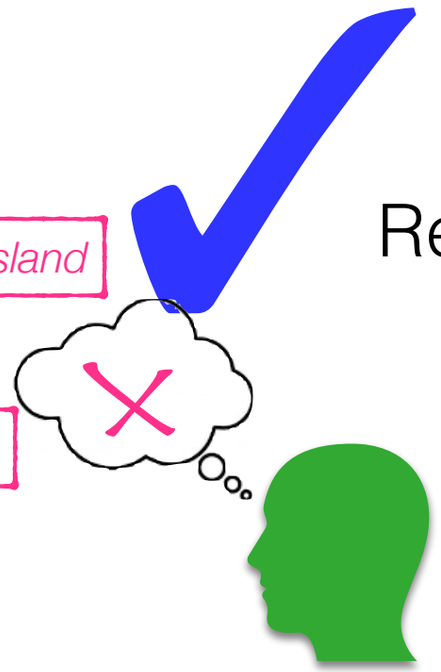
Reminder: Target behavior

Subject island

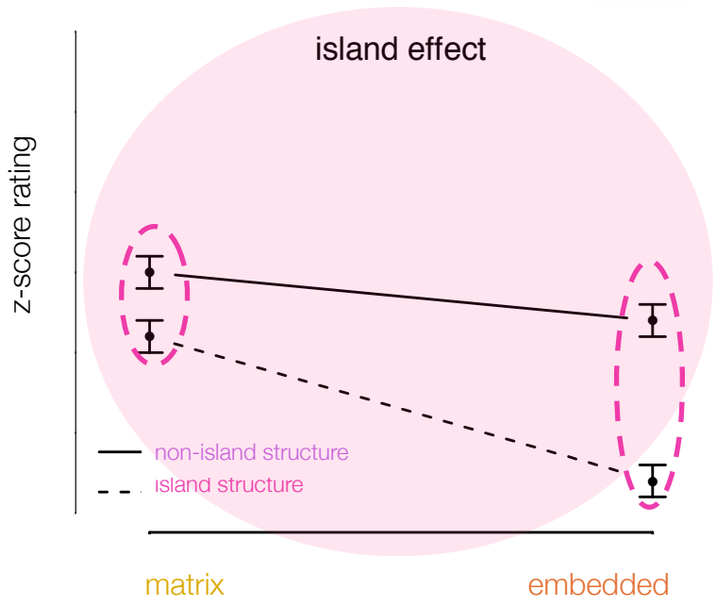
Complex NP island

Whether island

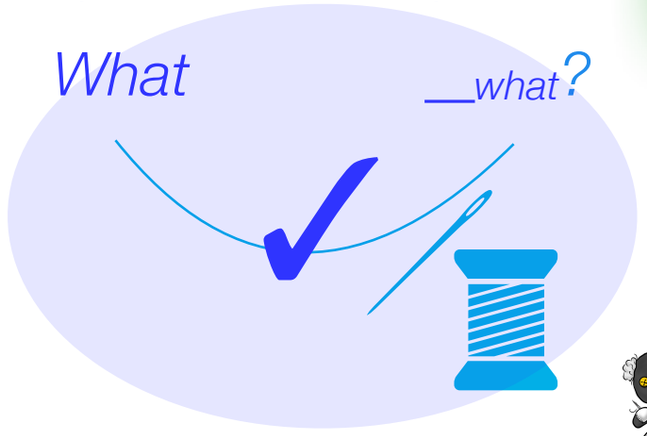
Adjunct island

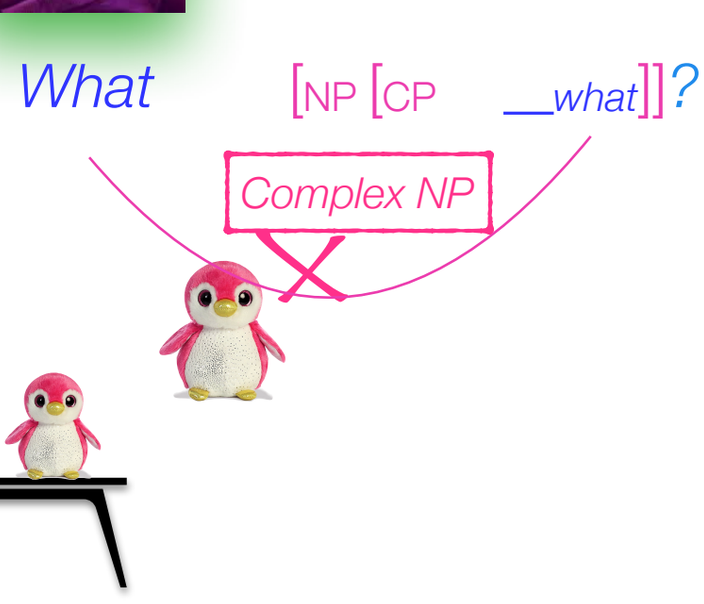


Children prefer this interpretation.



Sprouse et al. 2012

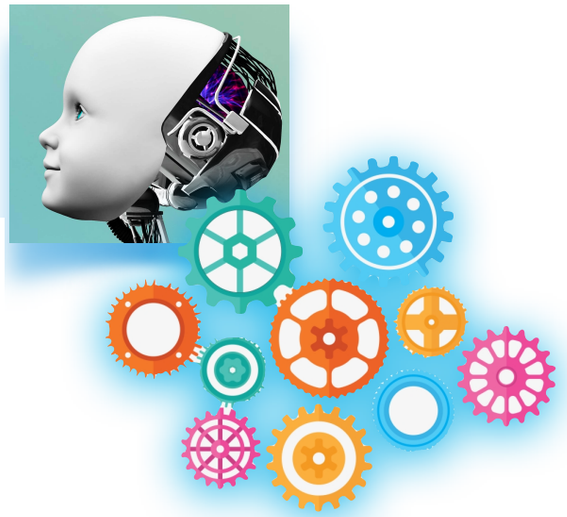


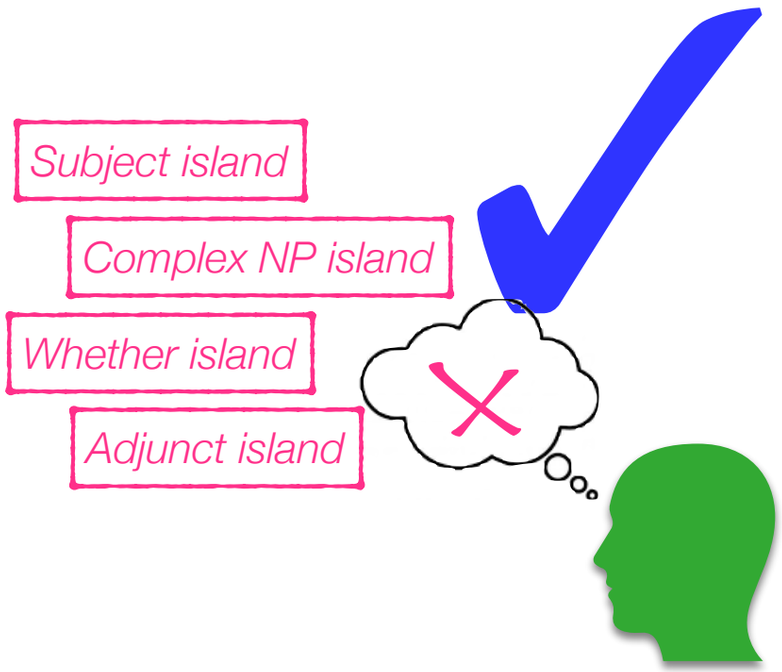


De Villiers et al. 2008

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

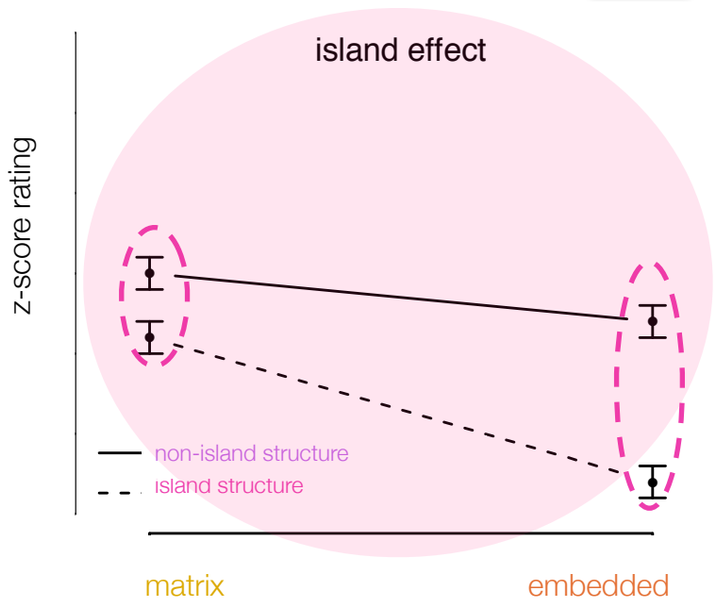
Evaluating the theory



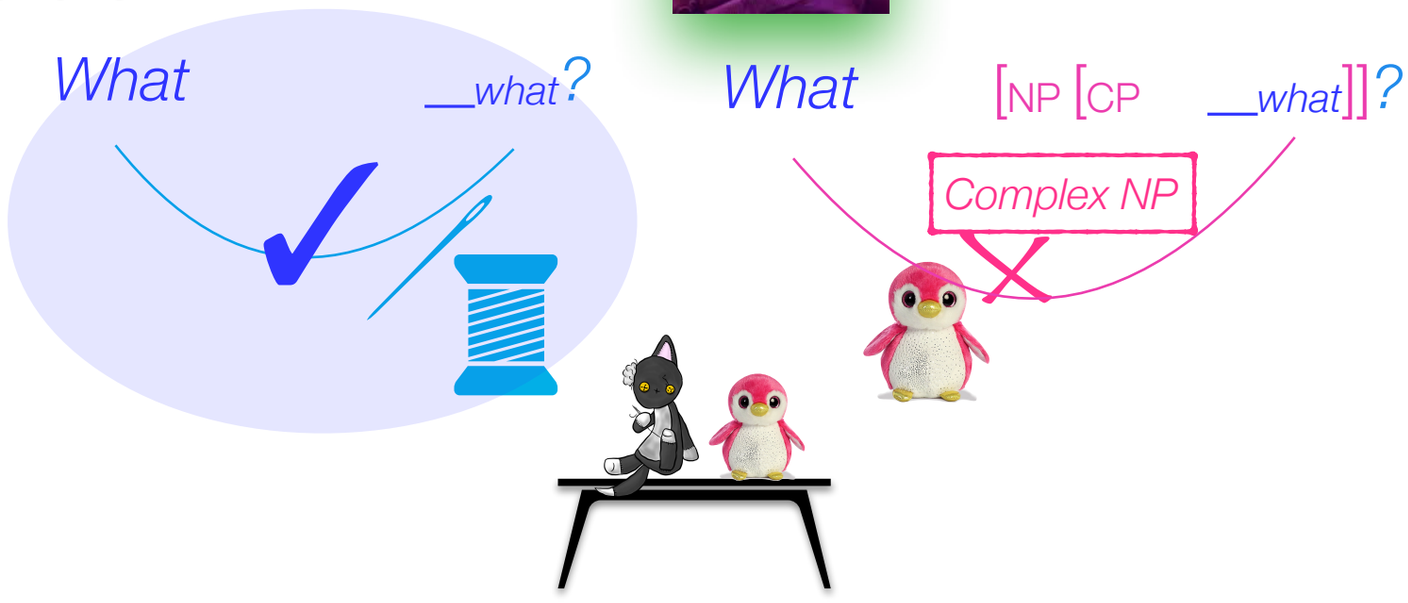


$$\Pi_{t \in trigrams}\, p(t)$$

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



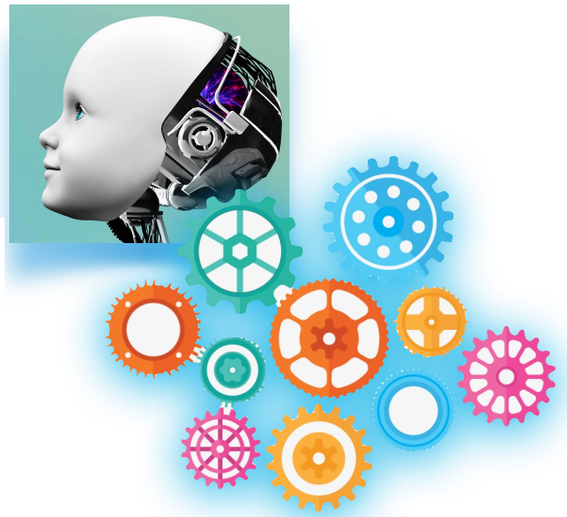
Sprouse et al. 2012

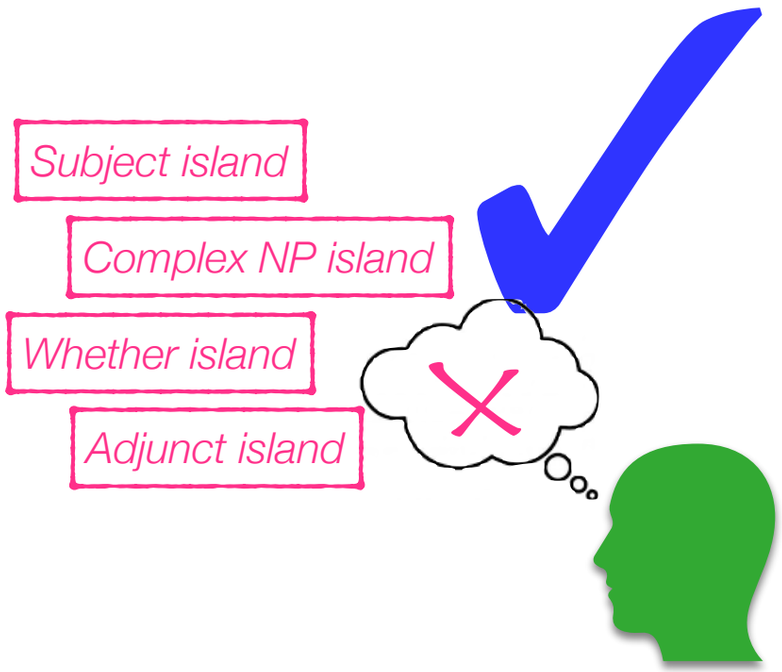


De Villiers et al. 2008

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

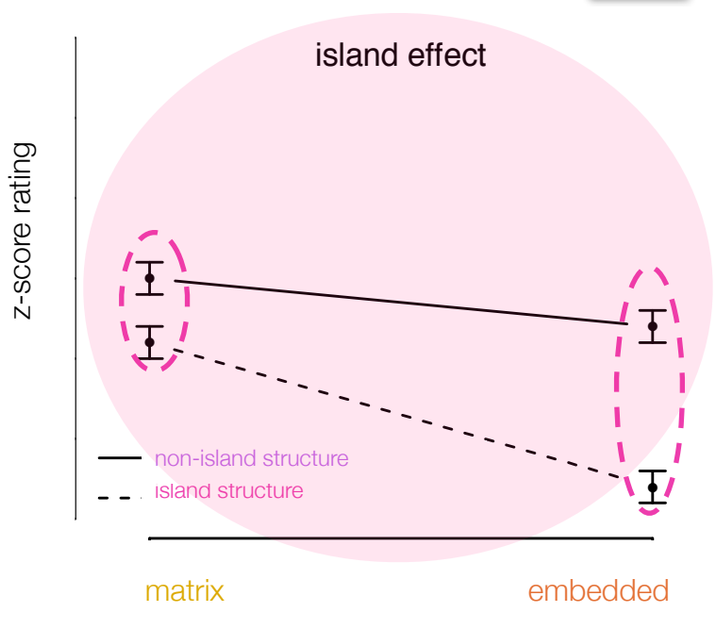
Evaluating the theory



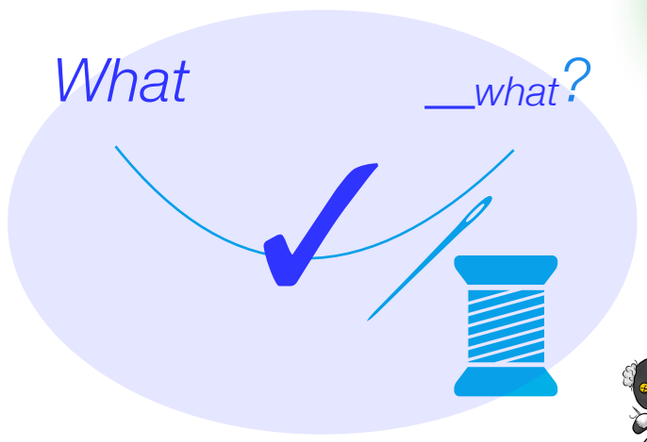


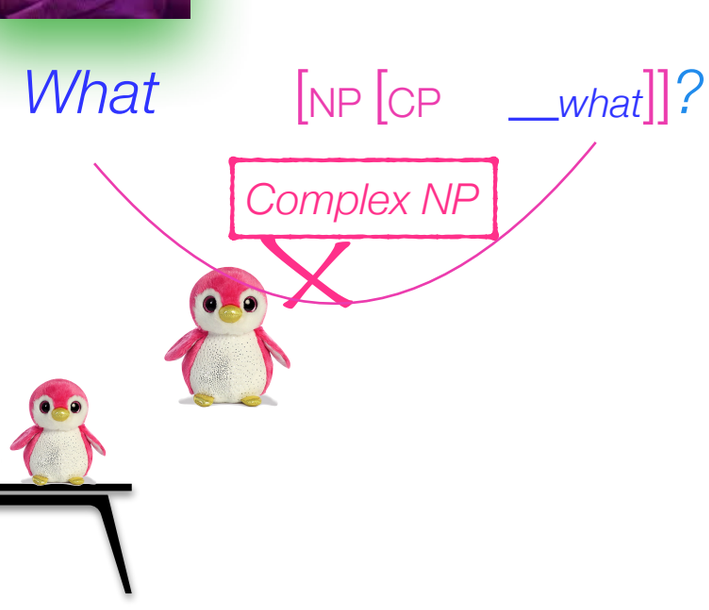
$$\Pi_{t \in trigrams}\, p(t)$$

So, the modeled child prefers it.



Sprouse et al. 2012

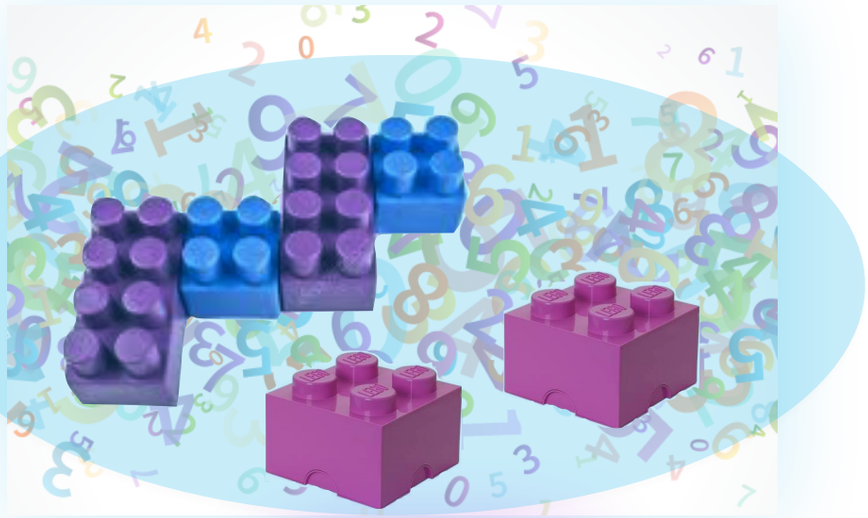


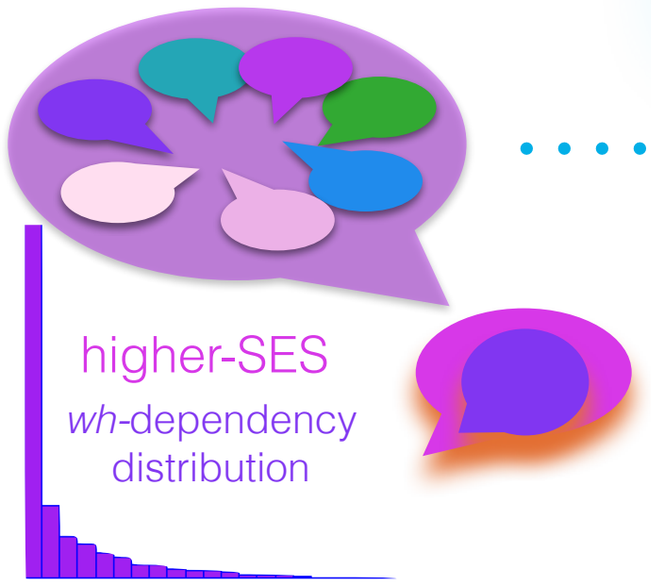


De Villiers et al. 2008

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Takeaway:
This theory can work for learning
knowledge about syntactic islands.

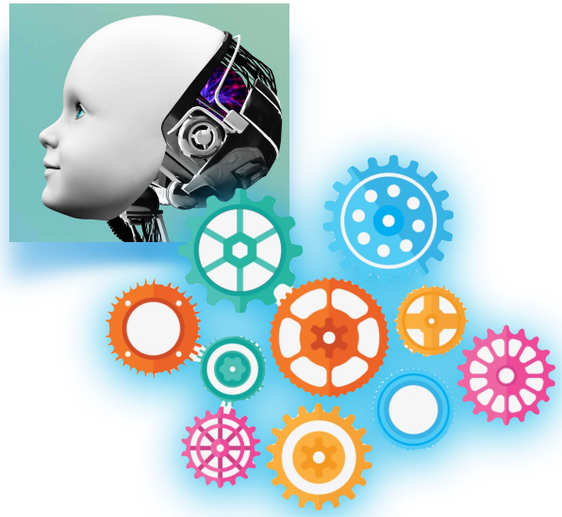


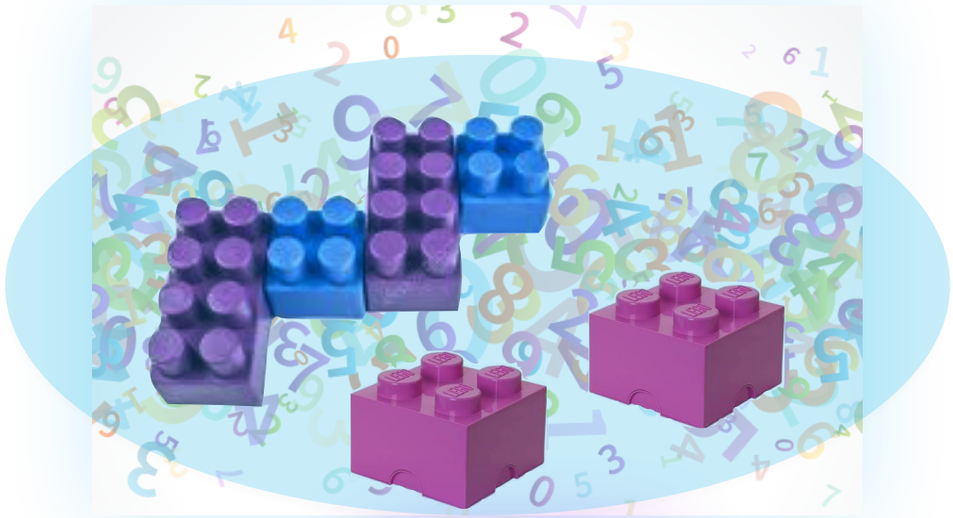


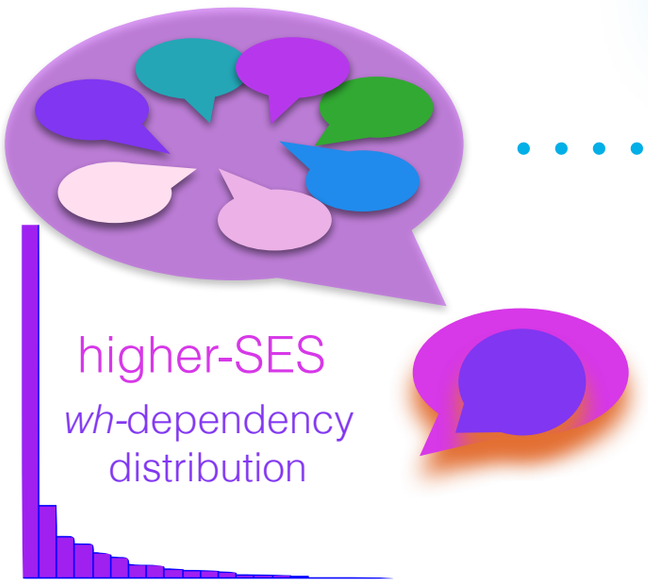
#

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

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

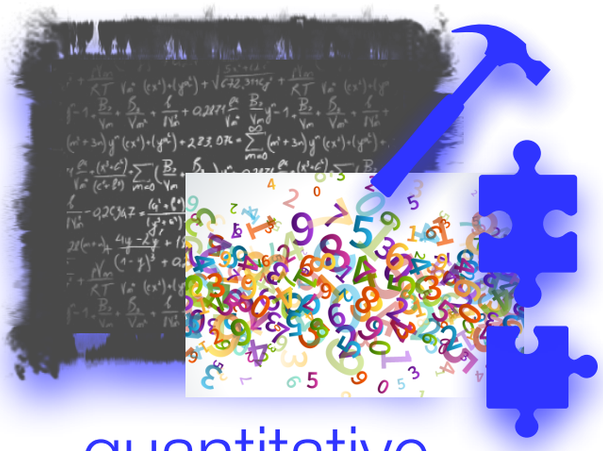






#

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press



quantitative

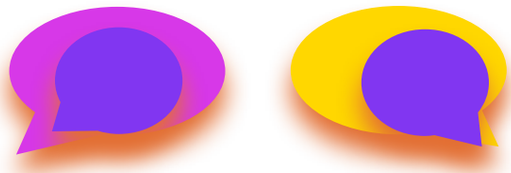
computational model

theory of acquisition

3: dialectal variation

Who does…

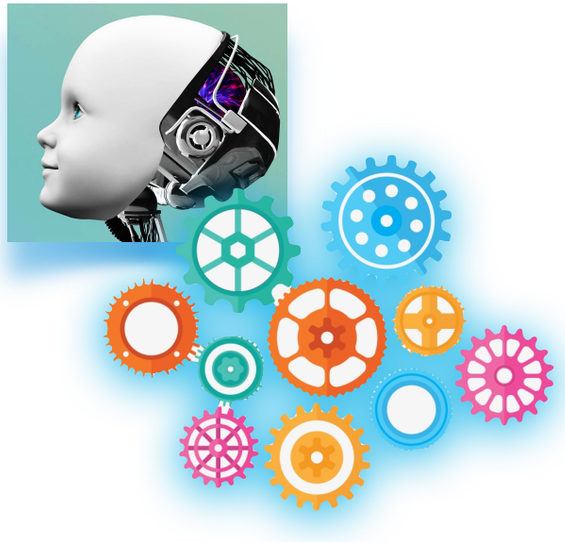
1: syntactic islands acquisition

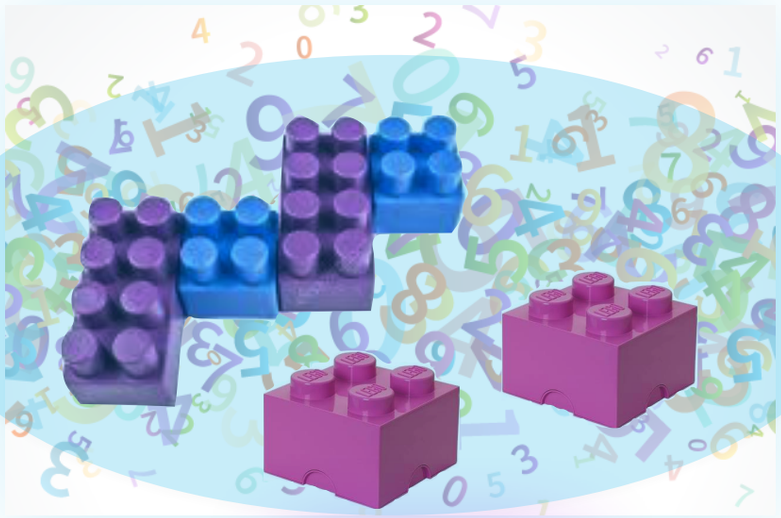


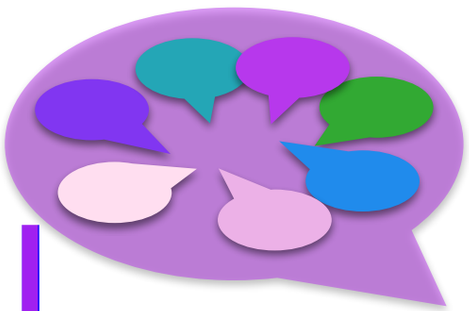
2: Evaluating a theory

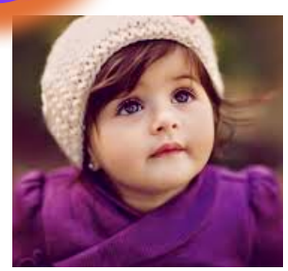
4: learning the building blocks

Evaluating the theory across dialects









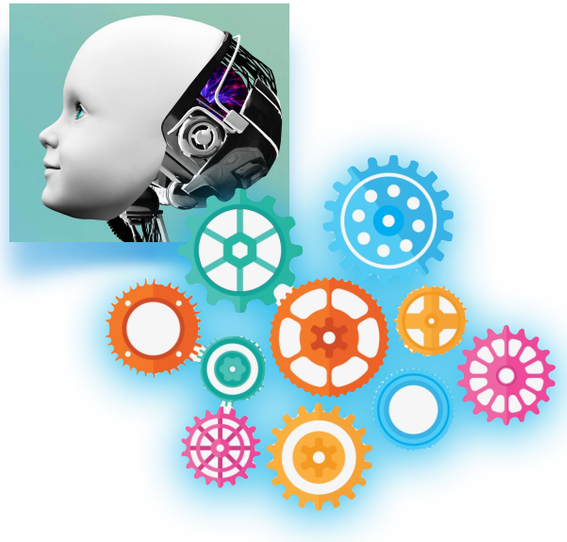
wh-dependency distribution

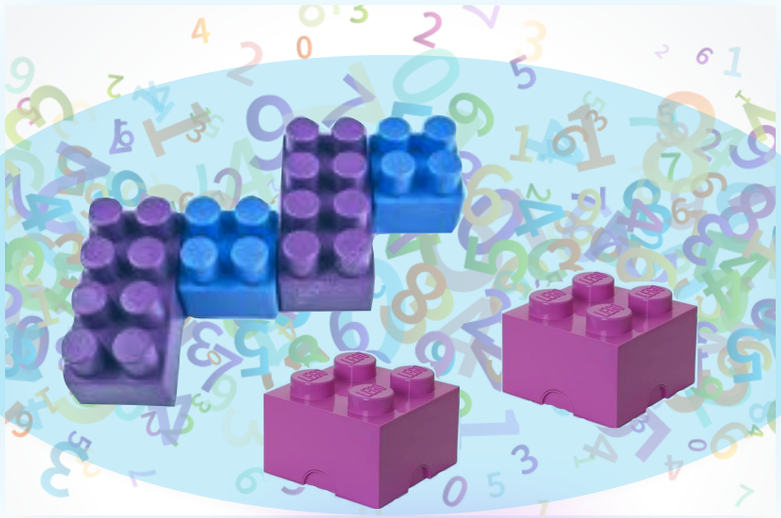
(≈850450 minutes)

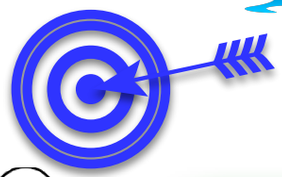
Our previous input sample came from speech directed at children from higher-SES backgrounds.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects

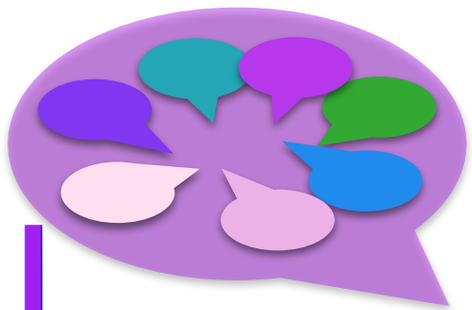






(≈850450 minutes)

#



wh-dependency distribution

What about children from lower-SES backgrounds?

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Input differences across SES

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).

Input differences across SES

Quality can be measured by different aspects of the input, like diversity of vocabulary …

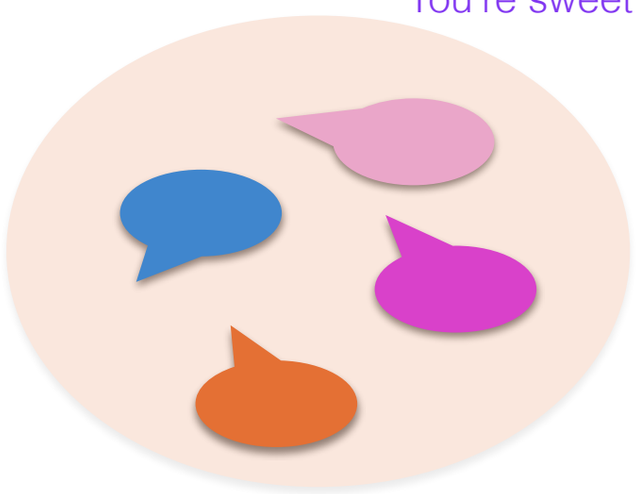
penguin penguin
kitty penguin penguin kitty
kitty penguin kitty penguin
kitty

whale seal penguin
kitty birdie monkey
kitty puppy monkey penguin
cat

Input differences across SES

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

Go play
Go eat
I love you
Come here
Go to bed
He's nice
I see you
You're cute
Come with me
Eat dinner
You're sweet



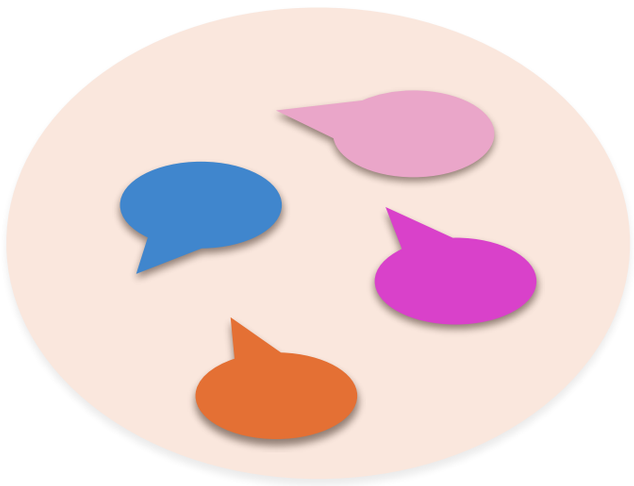
Come here
Is that a kitty?
What would you like to do today?
You're cute
Come play
What do you see?
I love you
What would you like?
Eat dinner
You like that, don't you?
What a cutie!

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



We saw her yesterday, didn't we?

The penguins should be at the zoo

Because the penguins were being fed.

The kitty wasn't there

Because we're going tomorrow

We'll see the kitty on Friday

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?

What happened?

What did you do?

Who did you think that saw the kitty?

What happened?

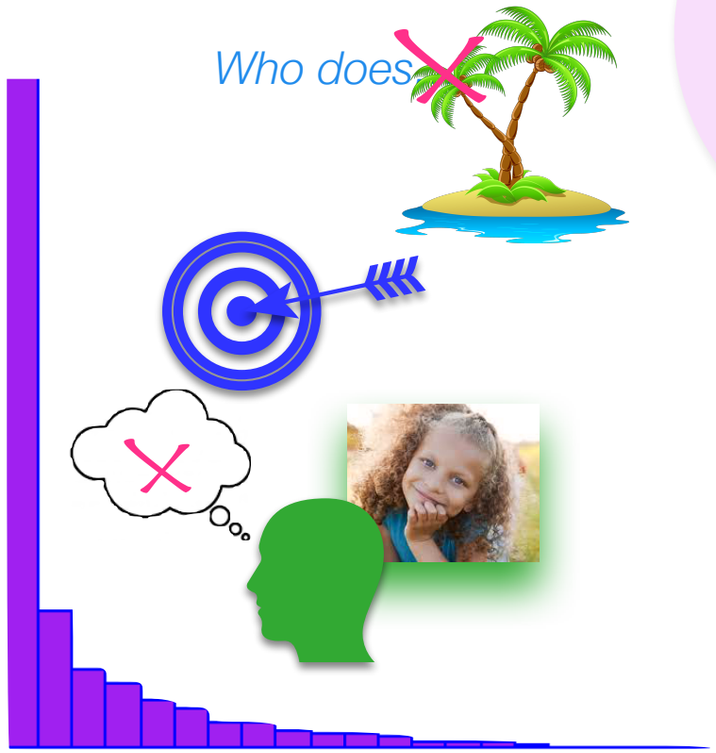
What happened?

What did you see?

Who did you think that you saw?

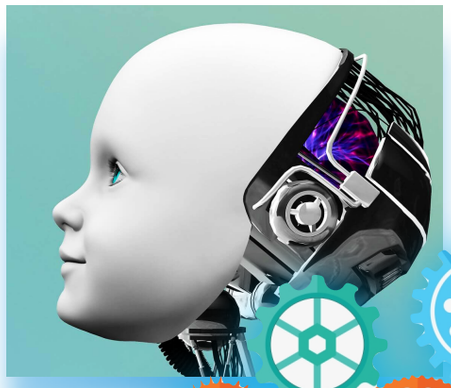
What happened?

Who did you see?



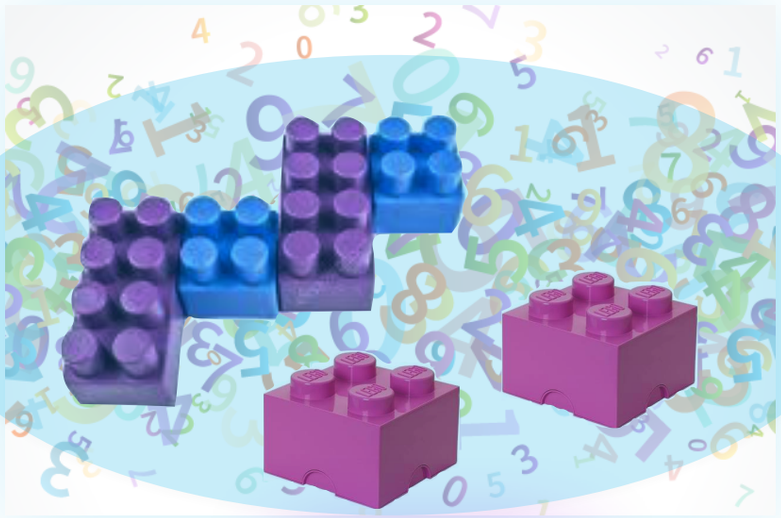
Input differences across SES for this theory

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?



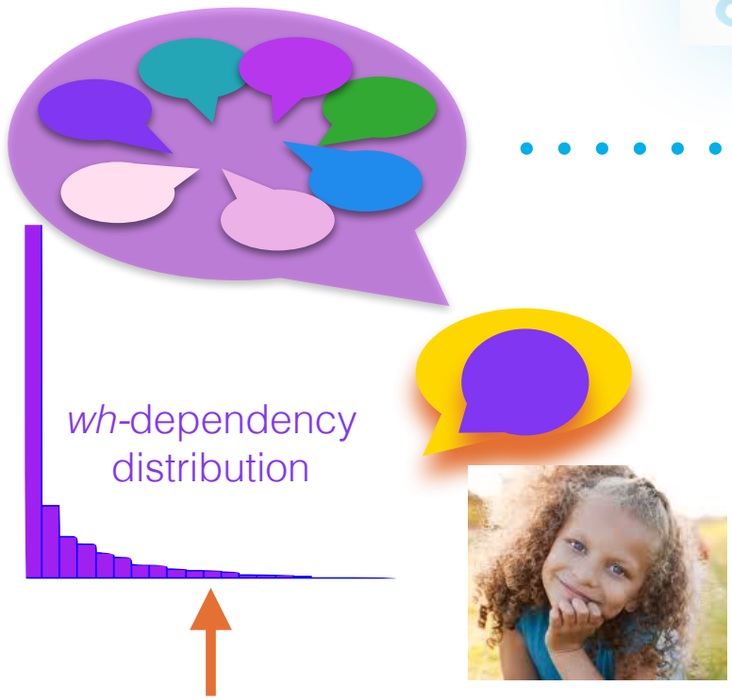
Let's find out!

Evaluating the theory across dialects



(≈850450 minutes)

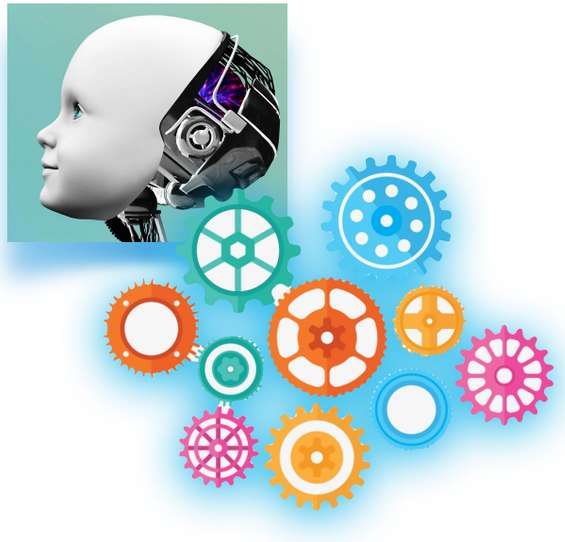
#

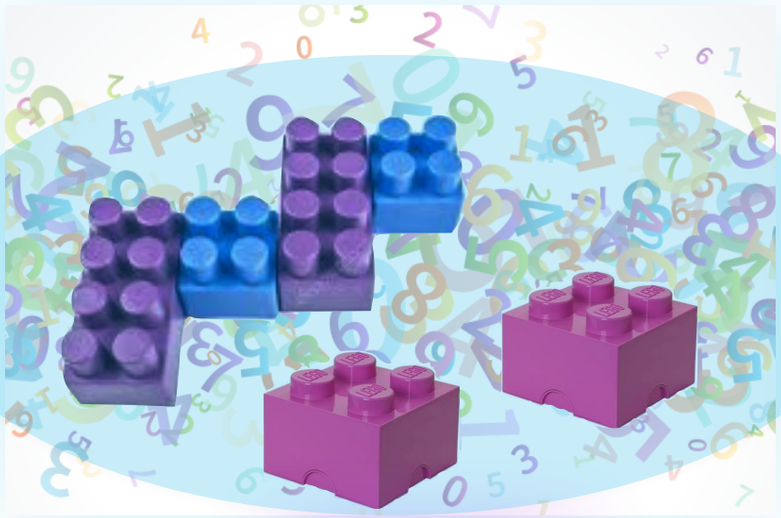


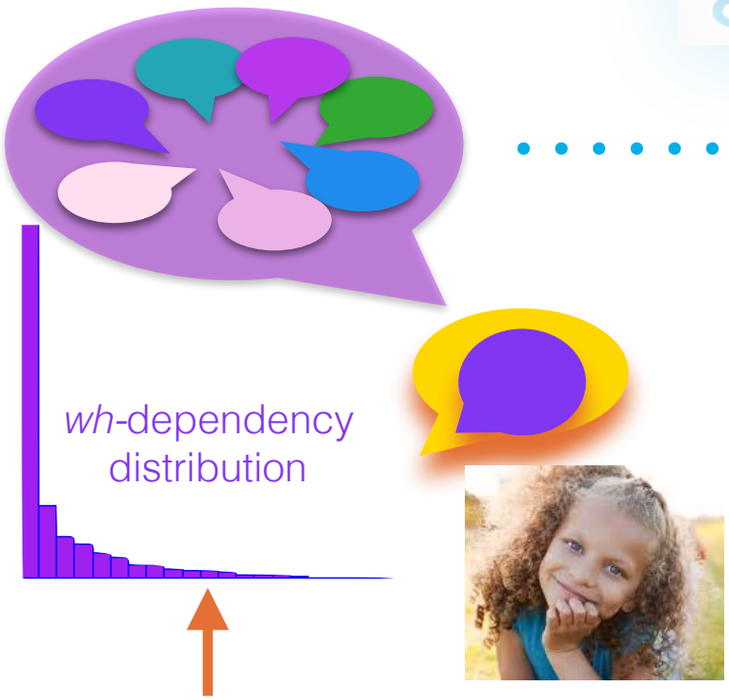
31.8K utterances (3.9K *wh*-dependencies) from a subpart of the HSLLD corpus (Dickinson & Tabors 2001) in the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 78 lower-SES children between the ages of 3 and 5.

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

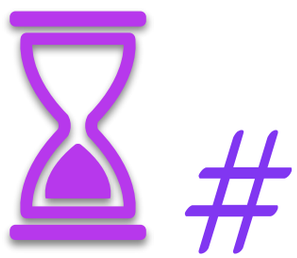
Evaluating the theory across dialects







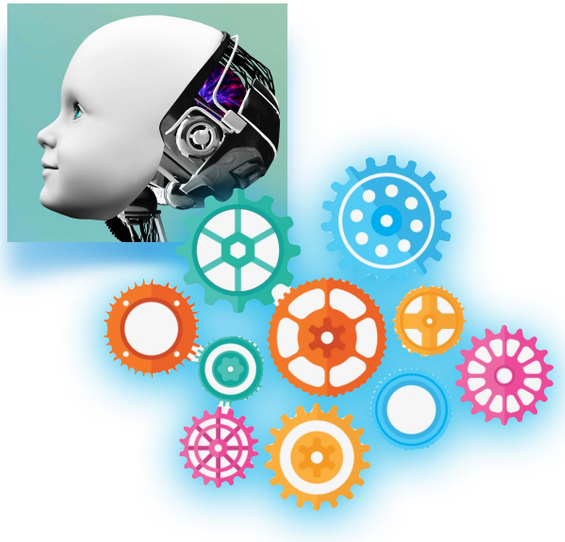
(≈850450 minutes)

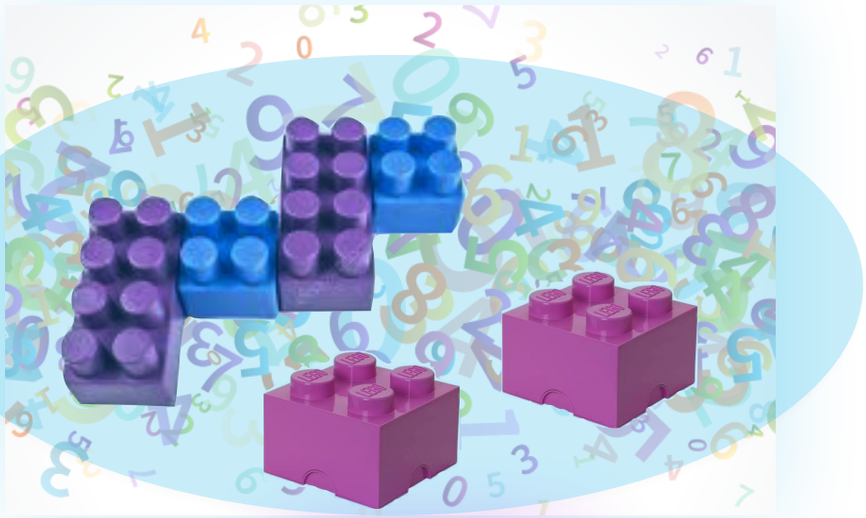


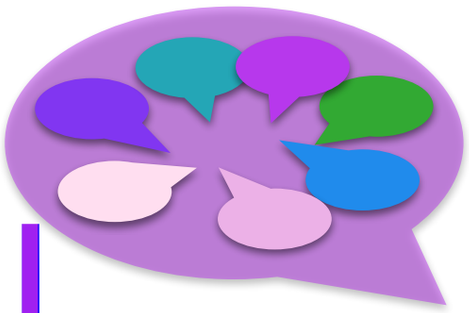
Note: Lower-SES was defined by the creators of the HSLLD corpus according to maternal education (6 years to some post-high school education) and annual income (70% reported < $20K/year).

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

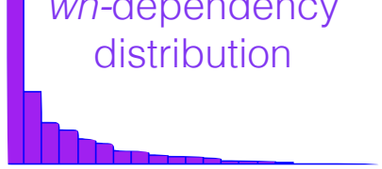
Evaluating the theory across dialects







(≈850450 minutes)

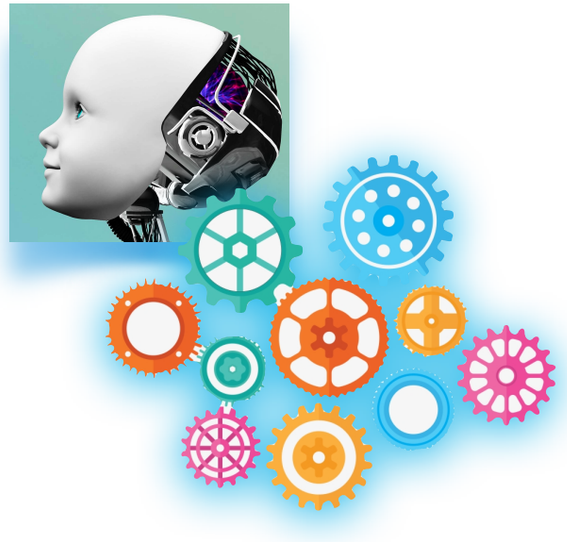


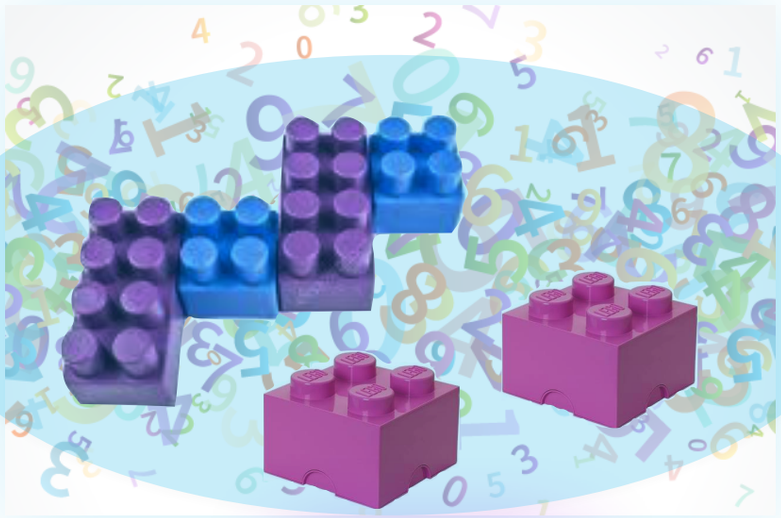
Hoff-Ginsberg (1998) and Rowe (2012):
Estimates of utterances per minute in speech directed at children from different backgrounds.

#

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects





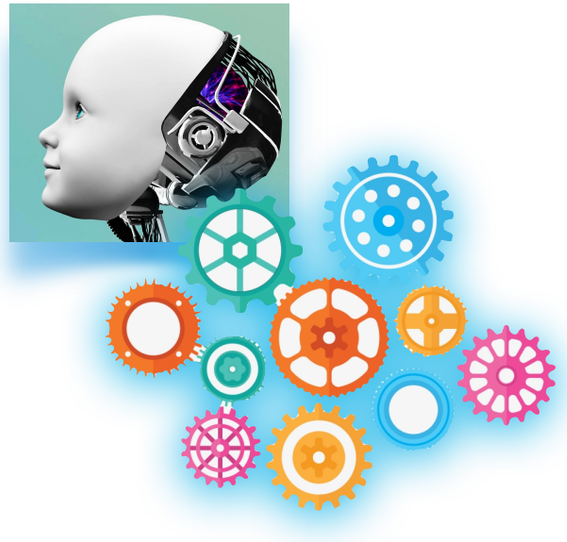
(≈850450 minutes)

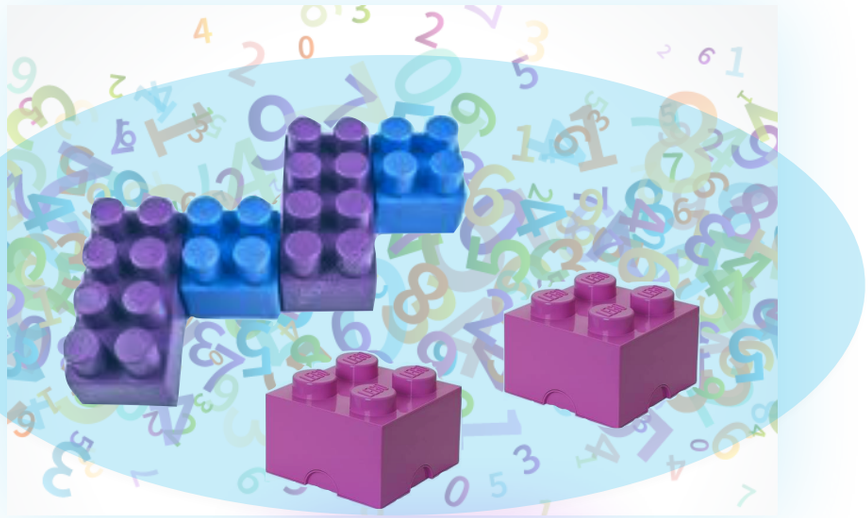
#

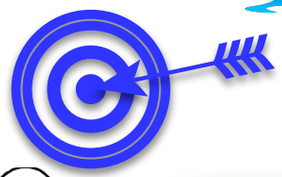
Data from children from lower socio-economic status (SES) backgrounds.

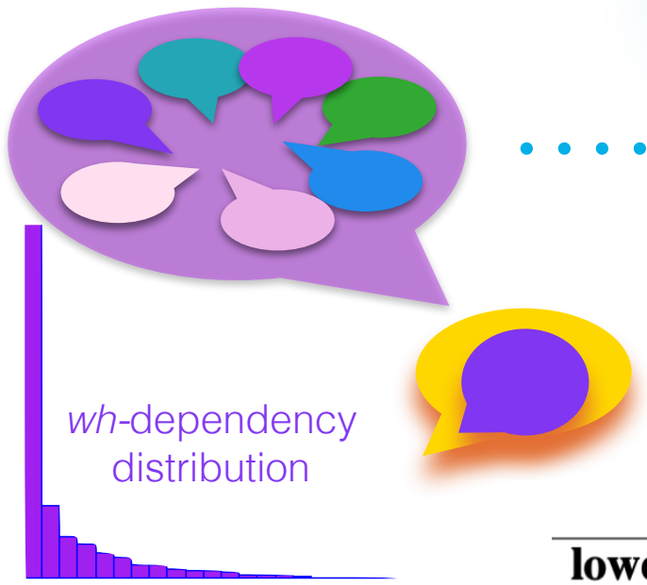
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects







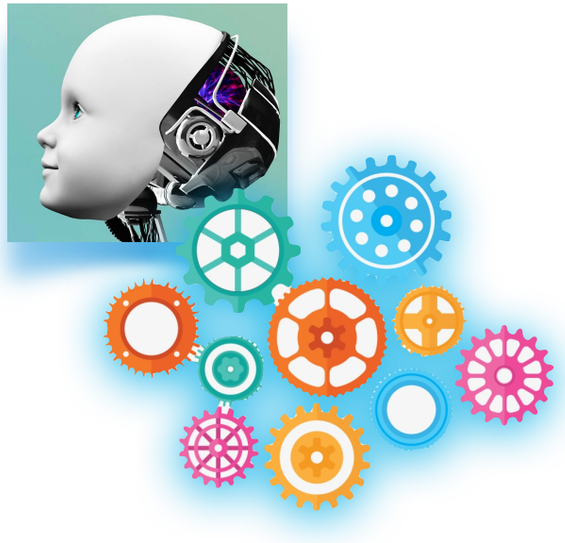


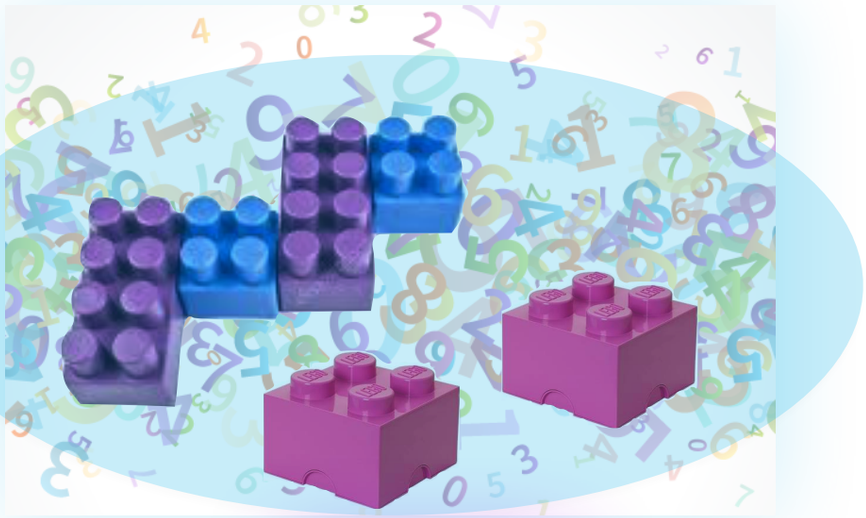
from our own corpus samples: rate of wh-dependencies/ utterance

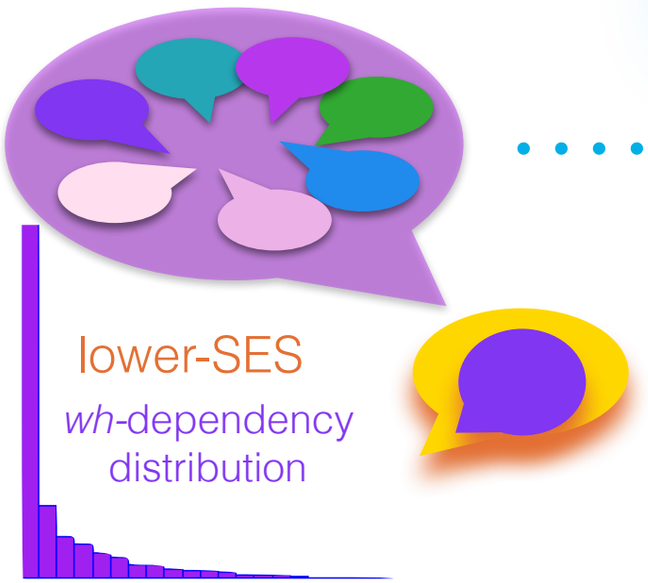
	utt/min	*	min	*	*wh*-dep/utt	=	total *wh*-dep
lower-SES		*	850,450.2	*	3,904/31,875		
- 2 s.d.	4.6					=	479,144
- 1 s.d.	8.8					=	916,624
average	13.0		#			=	1,354,103
+ 1 s.d.	17.2					=	1,791,583
+ 2 s.d.	21.4					=	2,229,063

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects







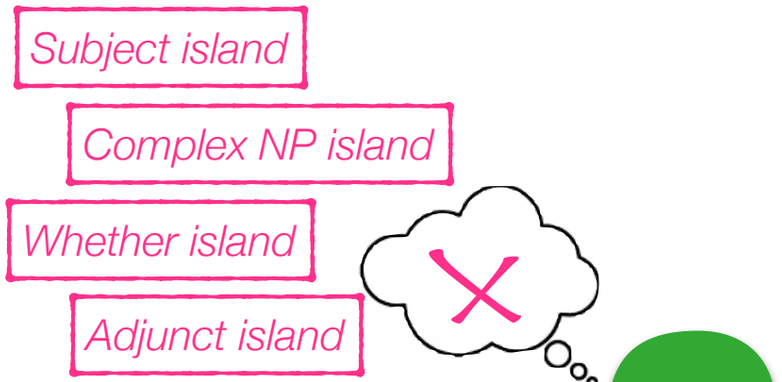
≈ 479K- 2.2 million *wh*-dependencies

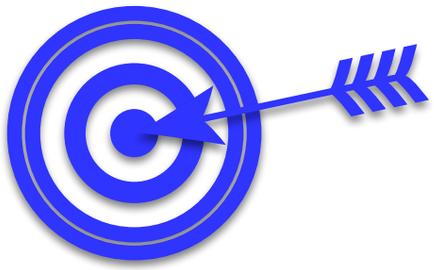
Can the modeled child produce the appropriate observable behavior?

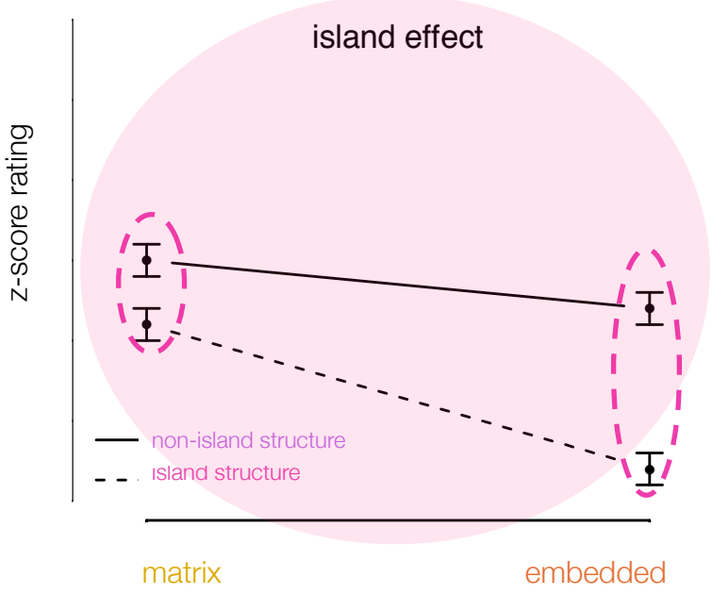
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects

Reminder: Target behavior

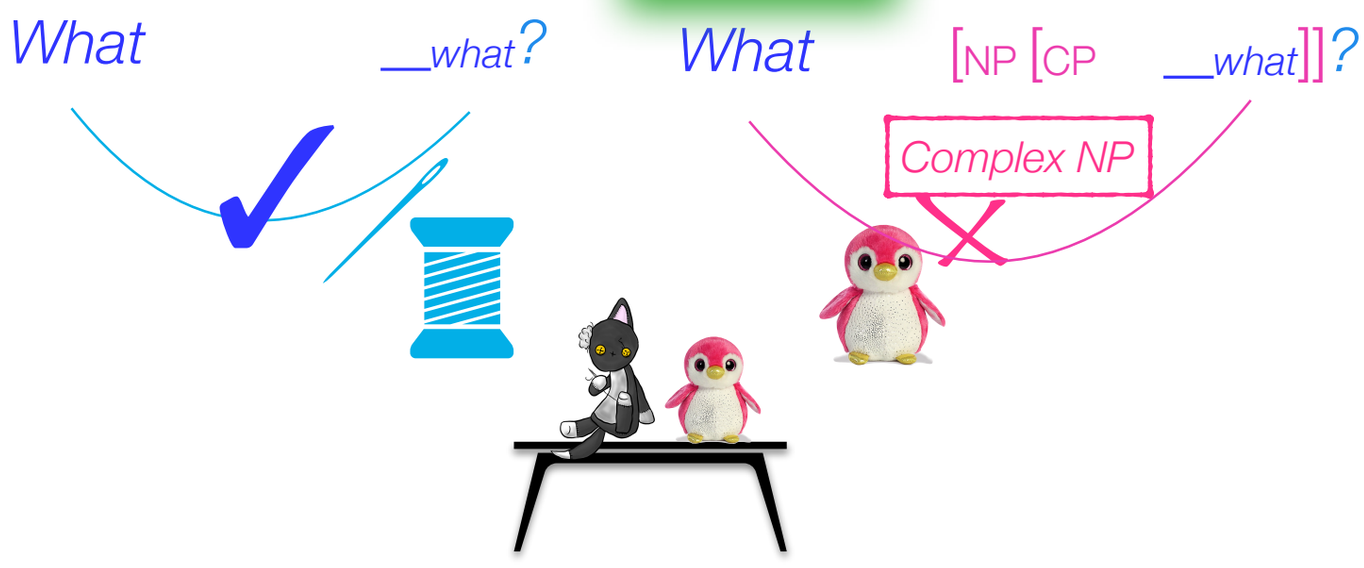






higher-SES *Sprouse et al. 2012*

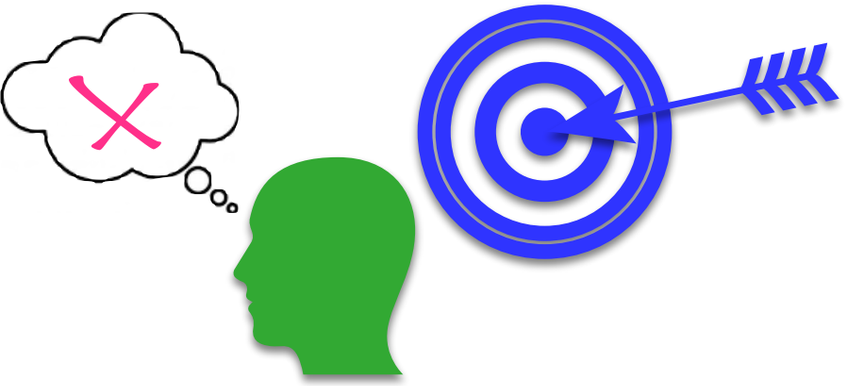
lower-SES *Bates et al. in prep.*



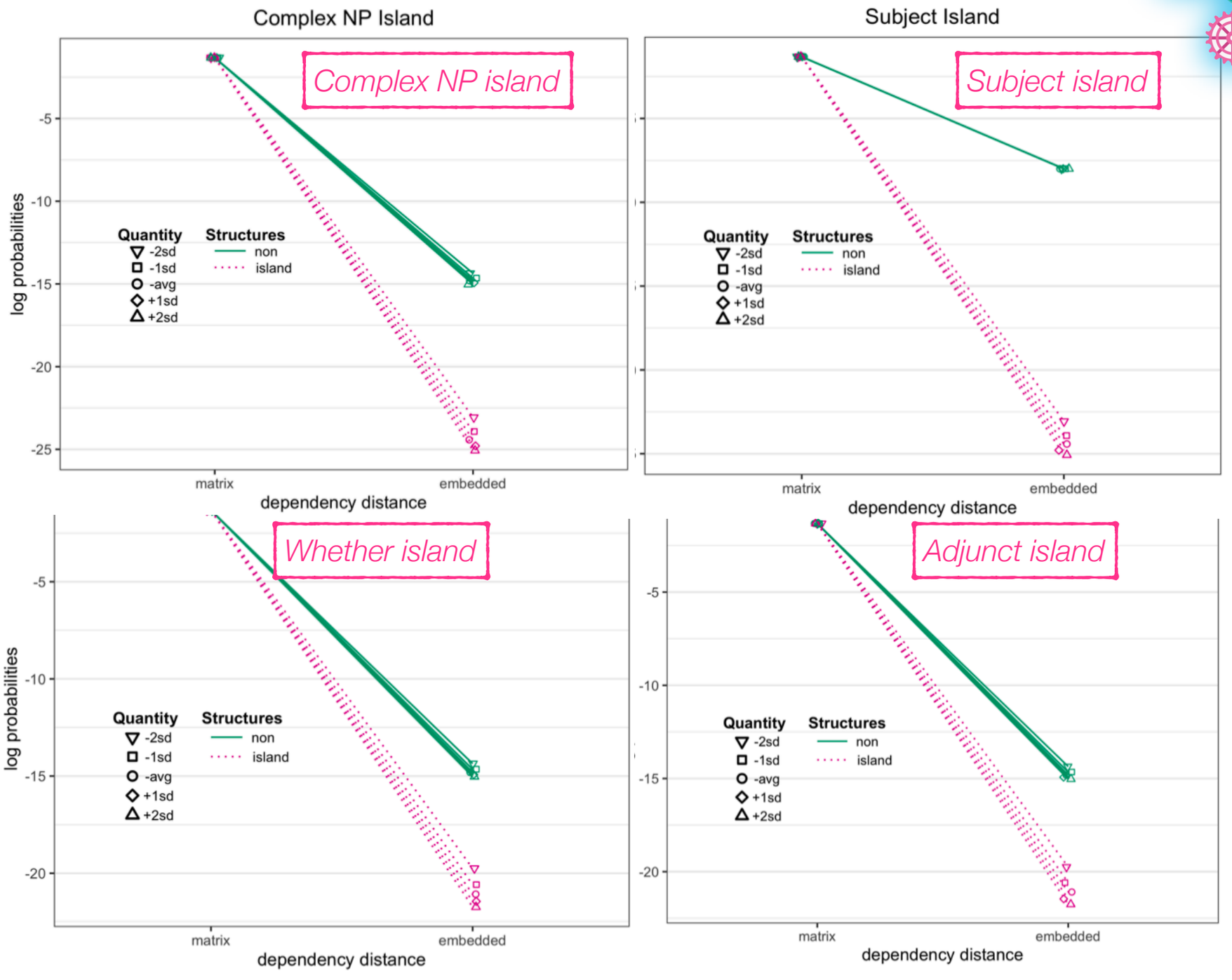
higher-SES *De Villiers et al. 2008*

lower-SES

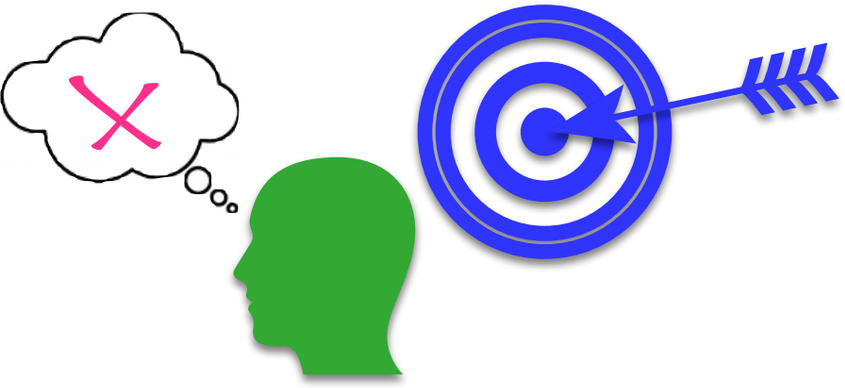
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press



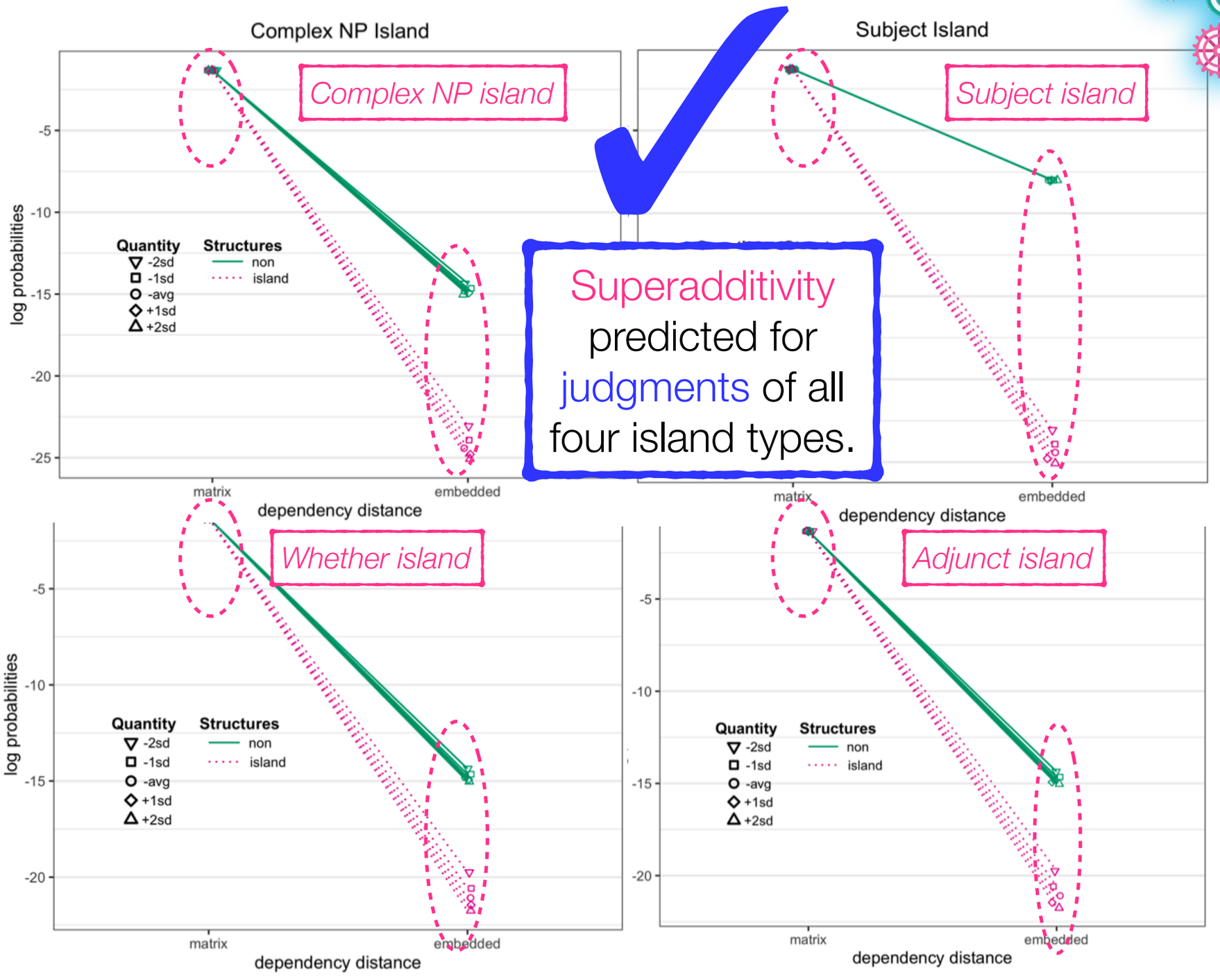
Evaluating the theory across dialects



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

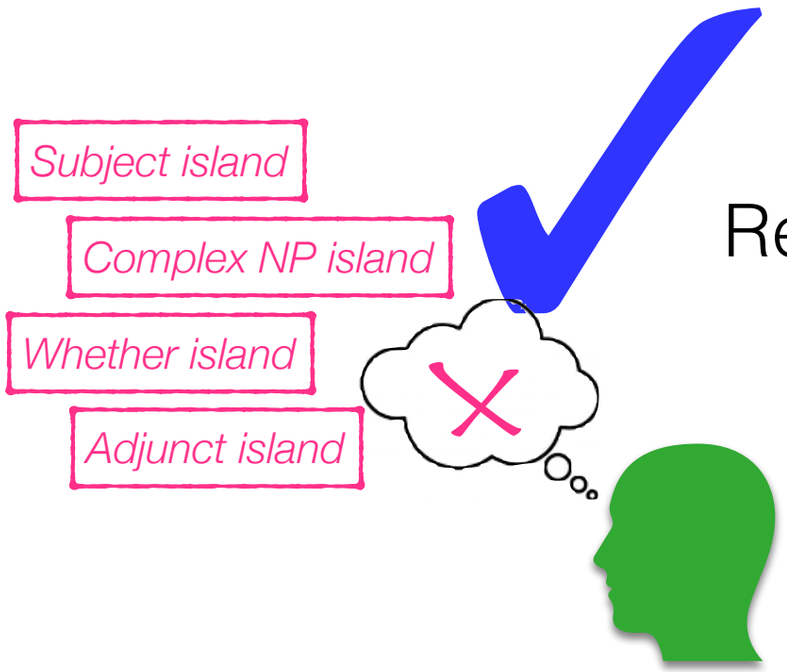
Reminder: Target behavior

Subject island

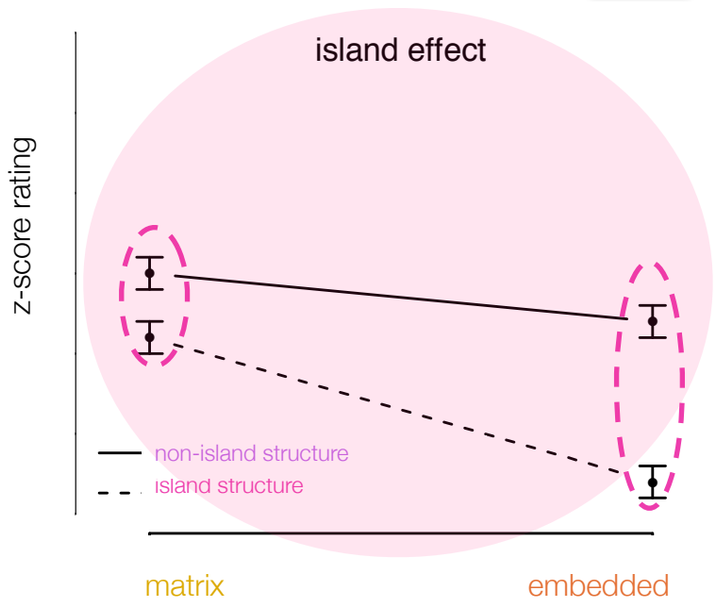
Complex NP island

Whether island

Adjunct island

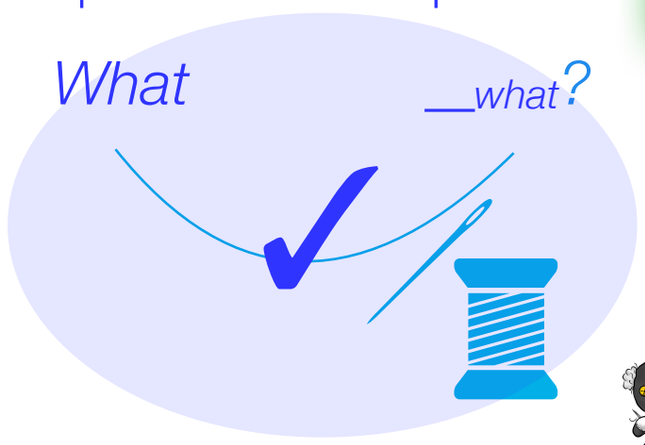


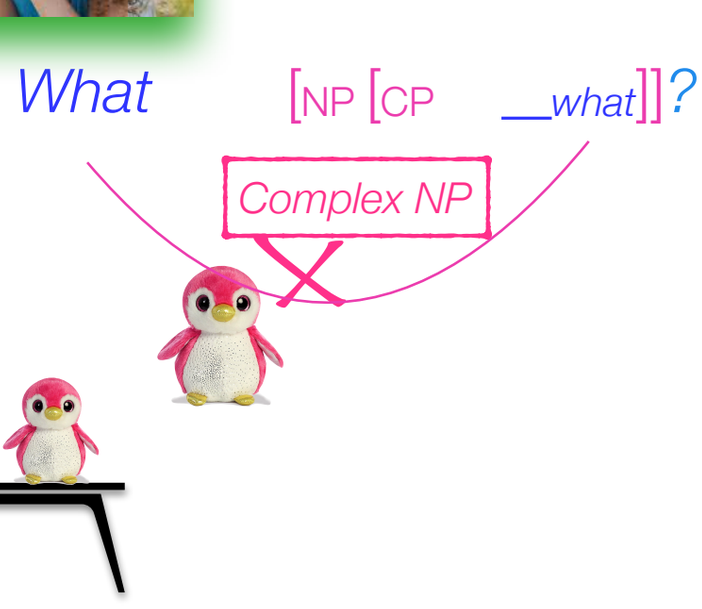
Children prefer this interpretation.



higher-SES *Sprouse et al. 2012*

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



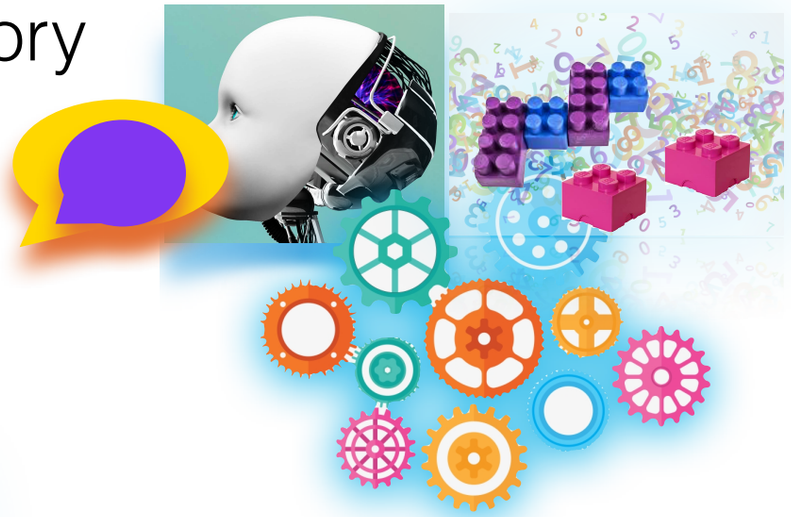


higher-SES *De Villiers et al. 2008*

lower-SES

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects

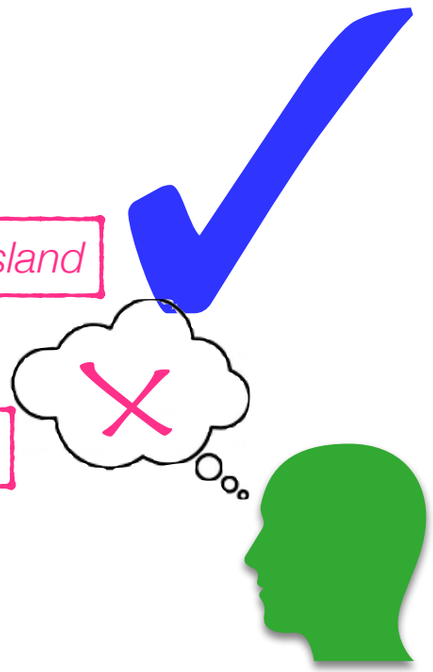


Subject island

Complex NP island

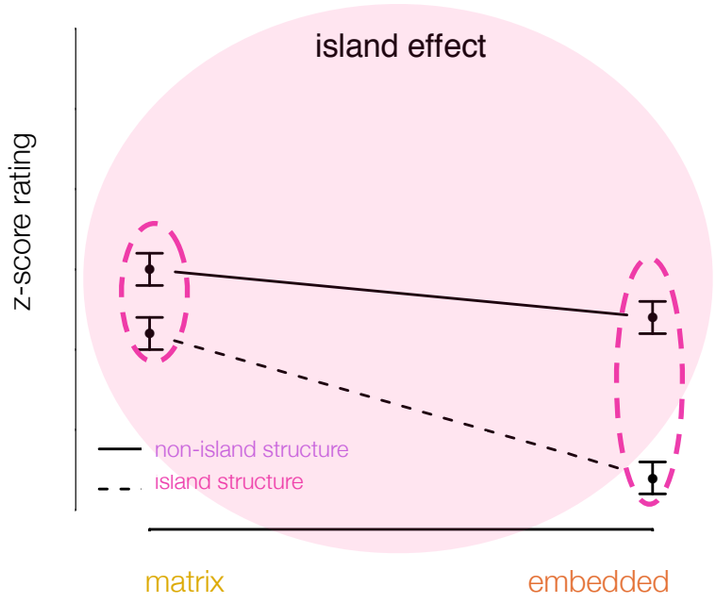
Whether island

Adjunct island



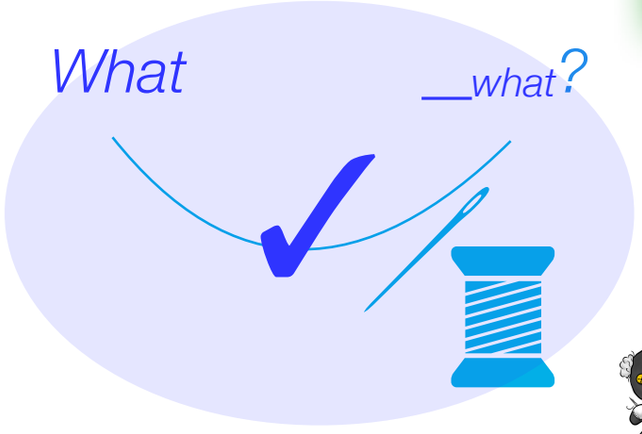
$\Pi_{t \in trigrams}\, p(t)$

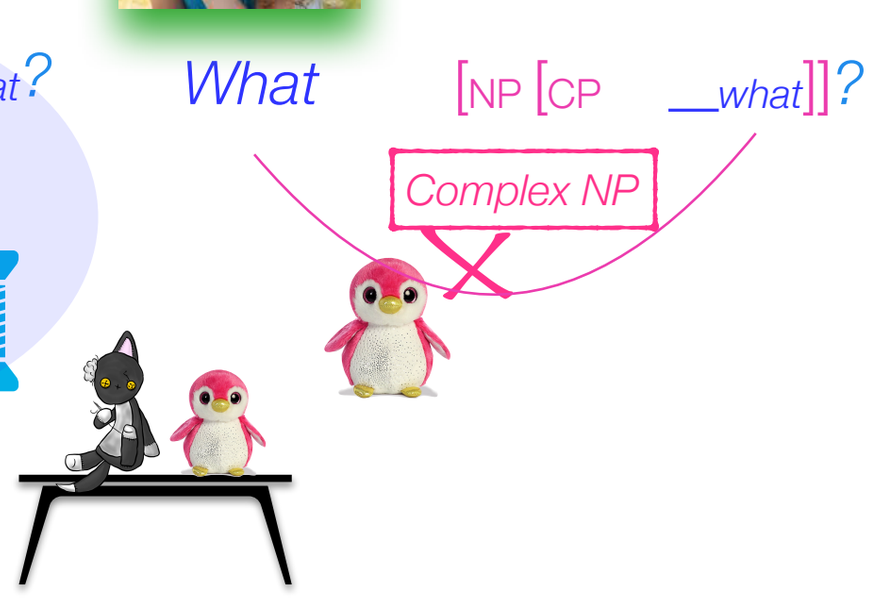
The *wh*-dependency this interpretation relies on is 10^{21} times more probable than the other one.



higher-SES *Sprouse et al. 2012*

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



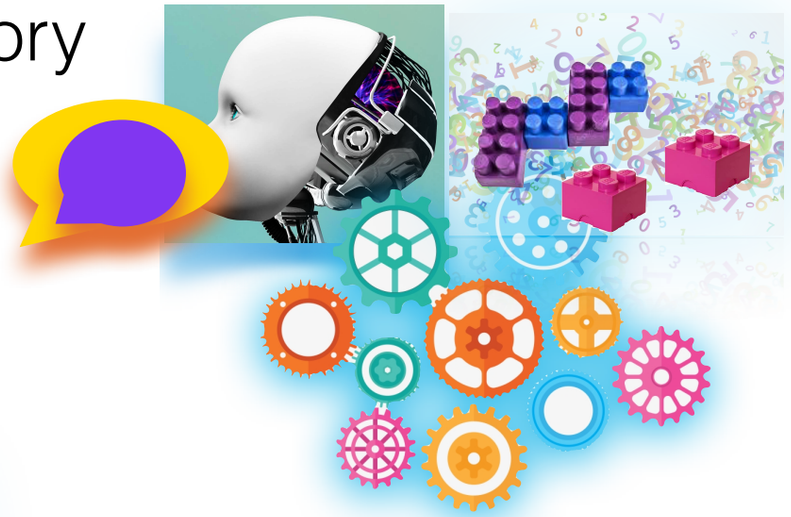


higher-SES *De Villiers et al. 2008*

lower-SES

Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Evaluating the theory across dialects

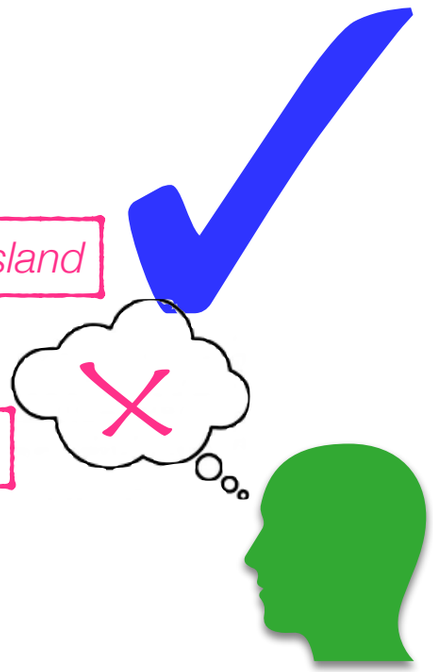


Subject island

Complex NP island

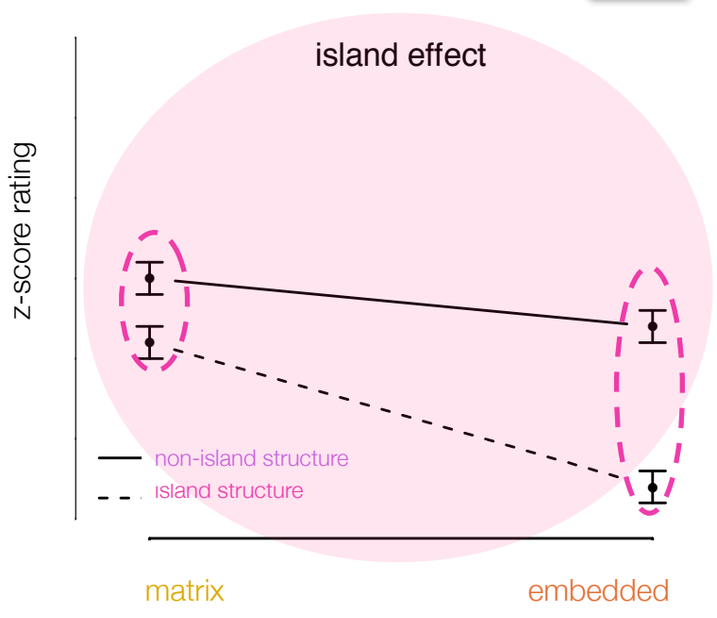
Whether island

Adjunct island



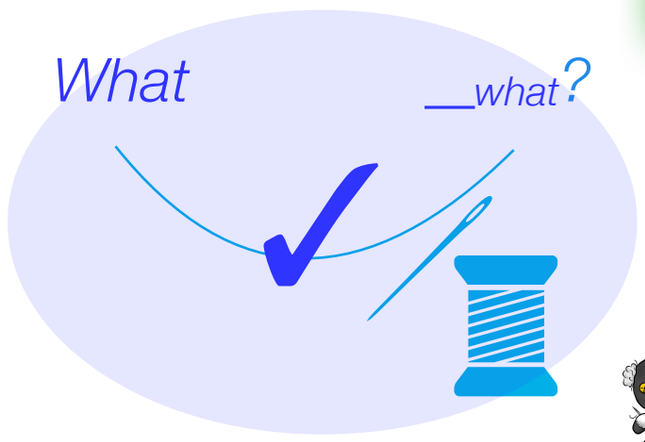
$\Pi_{t \in trigrams}\, p(t)$

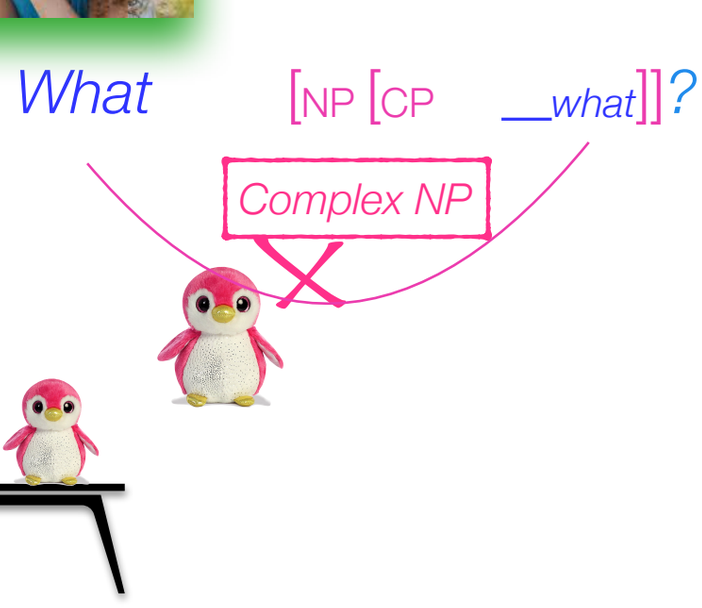
So, the modeled child prefers it.



higher-SES *Sprouse et al. 2012*

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





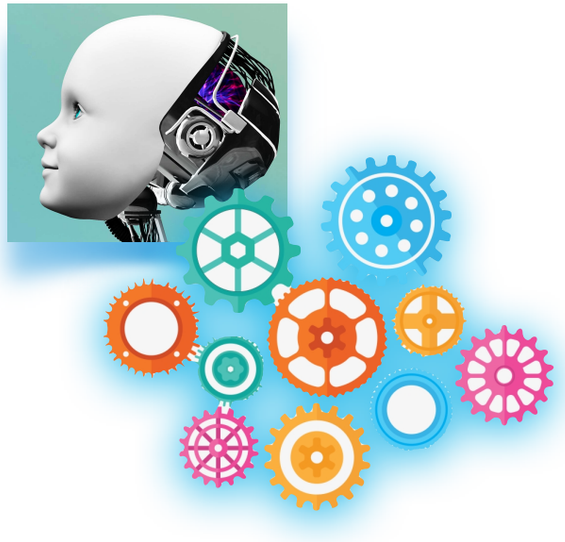
higher-SES *De Villiers et al. 2008*

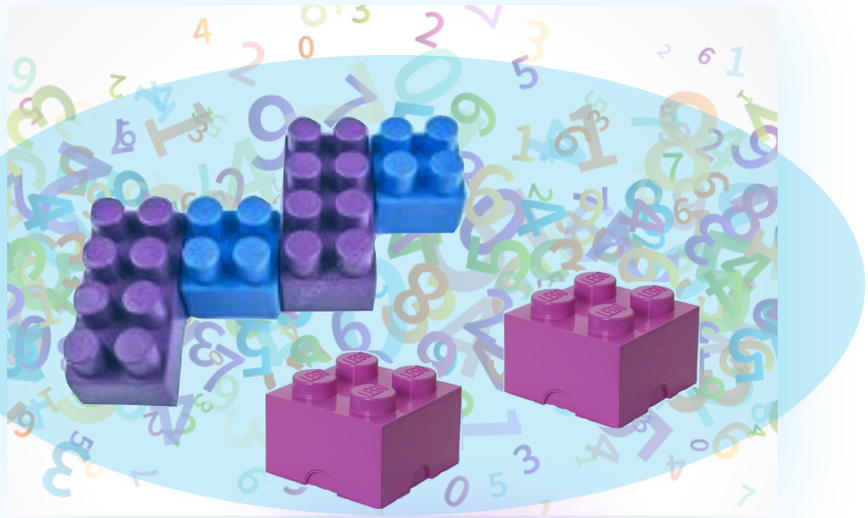
lower-SES

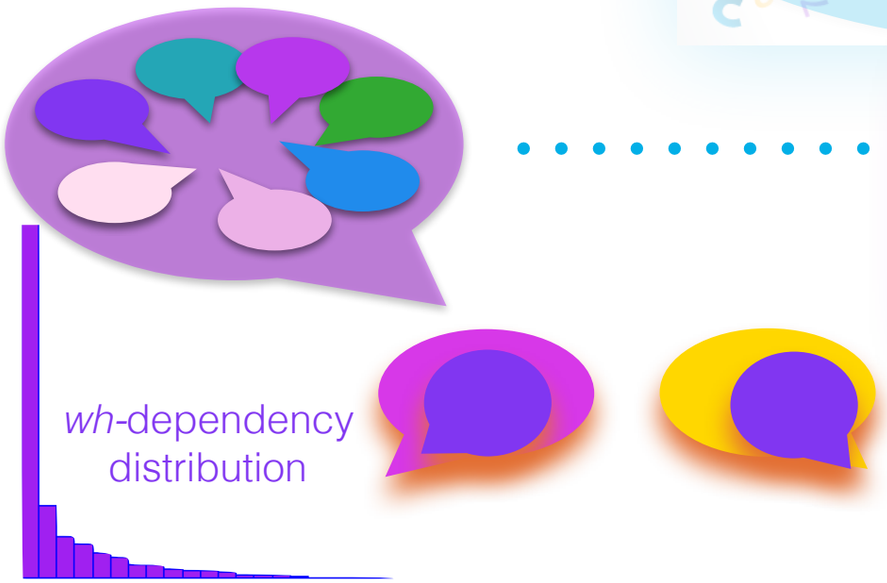
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Takeaway:

This theory can work for learning knowledge about syntactic islands across dialects.

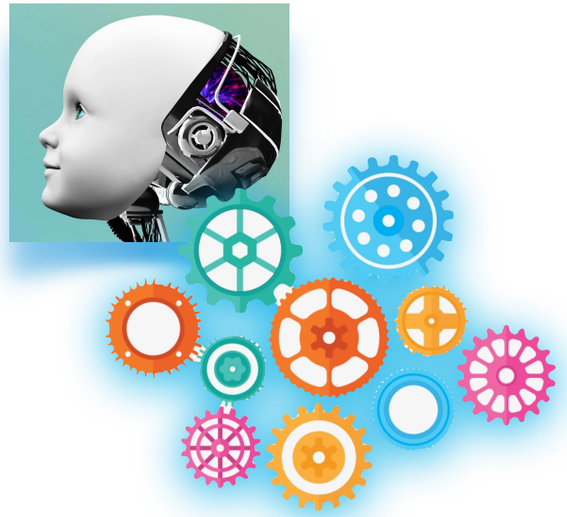


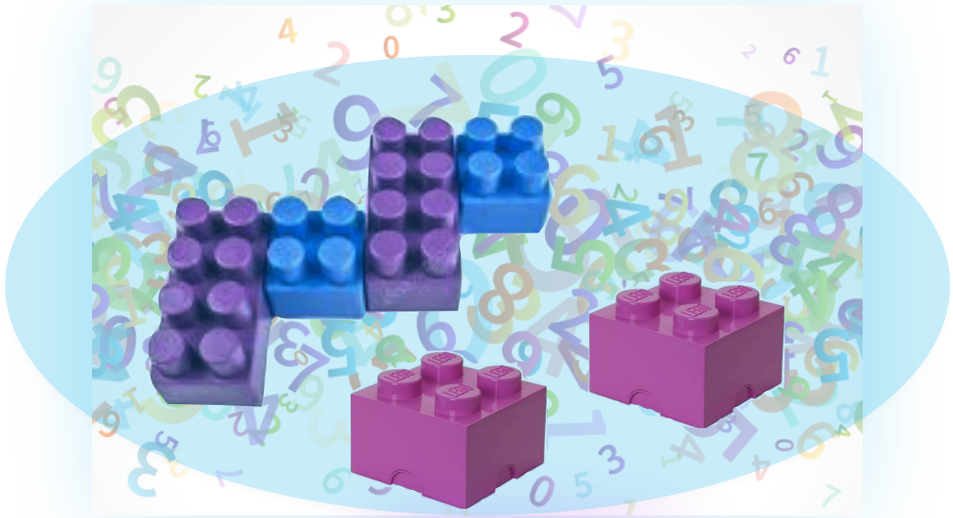


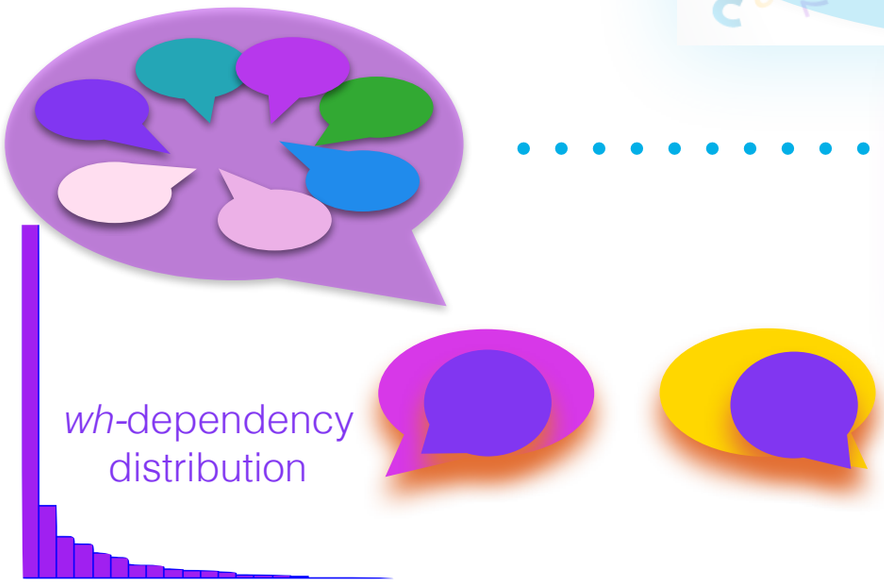


Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

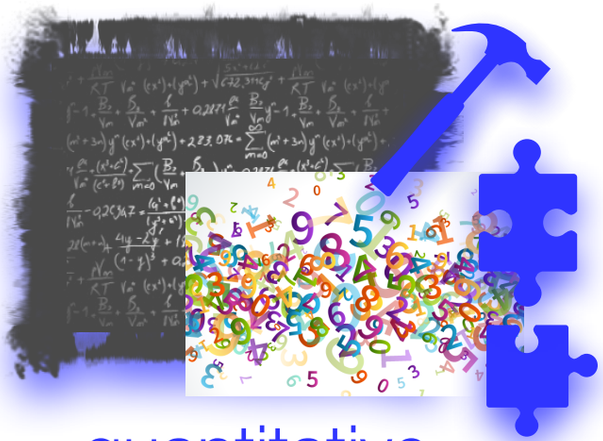
Key idea (again): Learning about the building blocks of *wh*-dependencies leads to knowledge about syntactic islands even when there's input variation.







Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press



quantitative

computational model

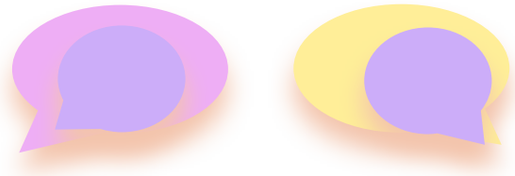
theory of acquisition

Who does…

1: syntactic islands acquisition

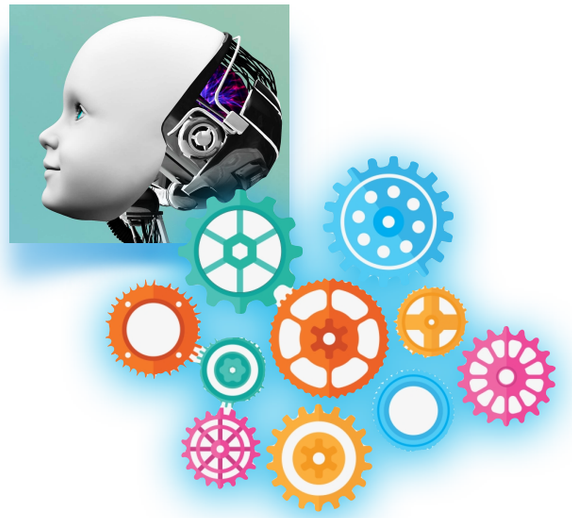
2: Evaluating a theory

3: dialectal variation

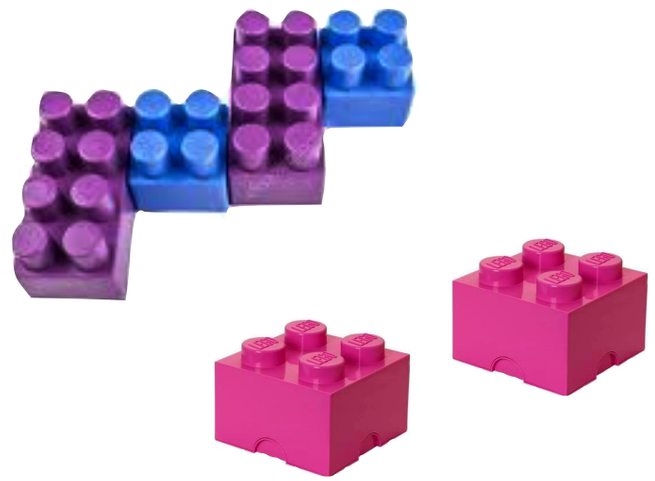


4: learning the building blocks

Learning the building blocks



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



start-IP-VP

IP-VP-*end*

IP-VP-IP

VP-IP-VP

start-IP-*end*

IP-VP-CP$_{that}$

VP-PP-*end*

IP-VP-PP

VP-CP$_{that}$-IP

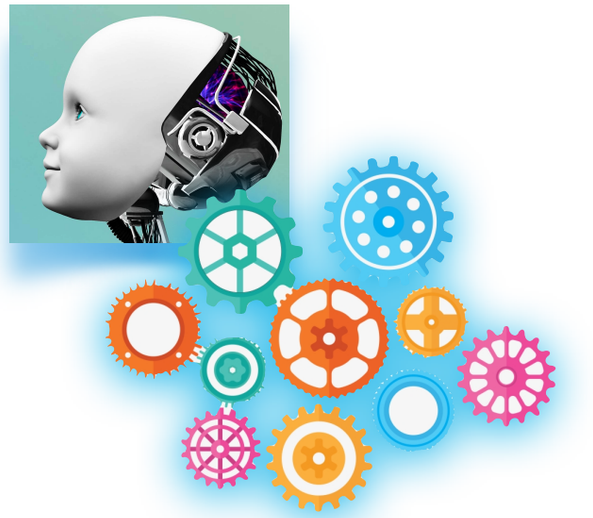
NP-PP-*end*

CP$_{that}$-IP-NP

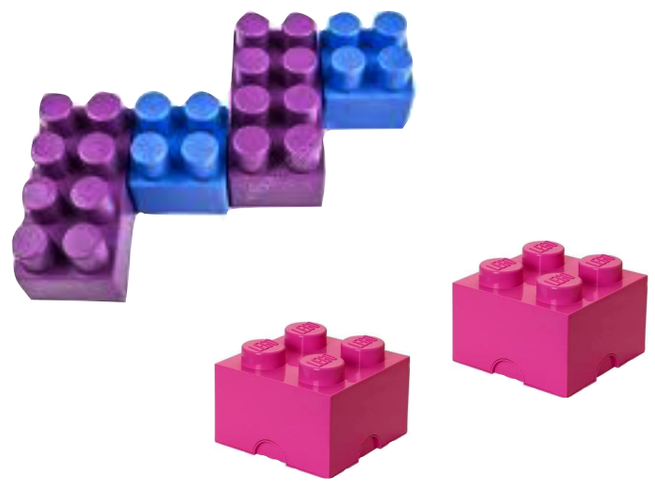
IP-NP-PP

Dickson et al. 2022, in prep.

Learning the building blocks



In particular:
(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-IP

VP-IP-VP

start-IP-*end*

IP-VP-CP$_{that}$

IP-VP-PP

VP-PP-*end*

VP-CP$_{that}$-IP

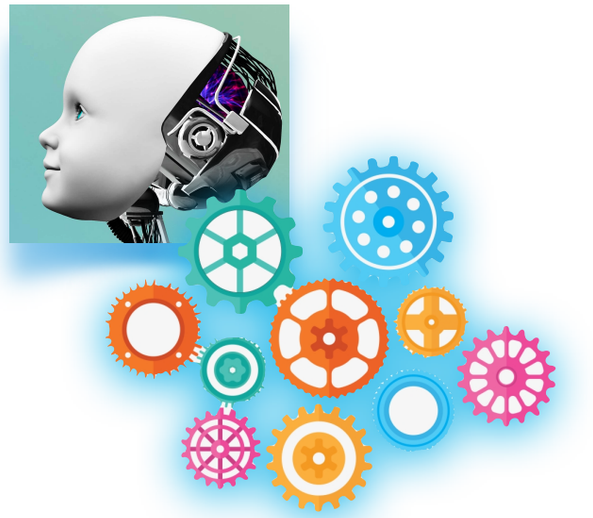
NP-PP-*end*

CP$_{that}$-IP-NP

IP-NP-PP

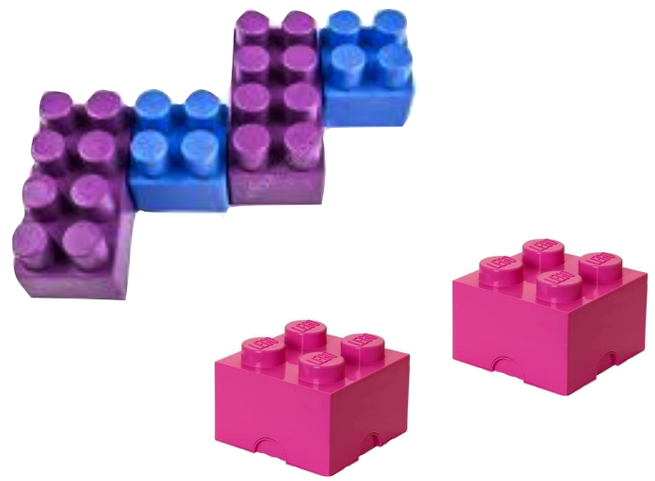
Dickson et al. 2022, in prep.

Learning the building blocks



…CP_{that}…

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



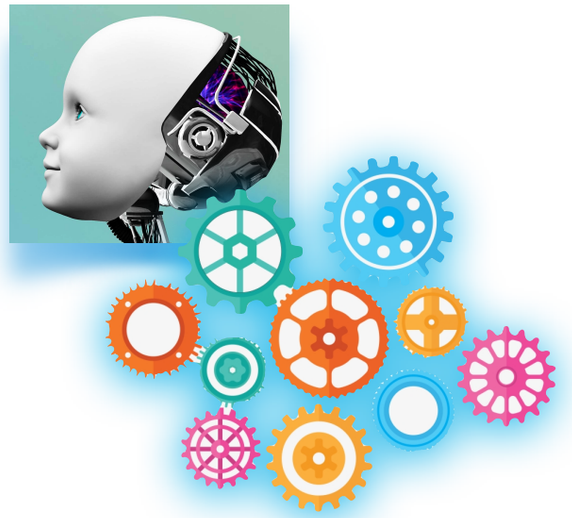
What does the teacher think [that Lily forgot __]? embedded | non-island

Whether *What does the teacher wonder [whether Lily forgot __]? embedded | island

Adjunct *What does the teacher worry [if Lily forgot __]? embedded | island

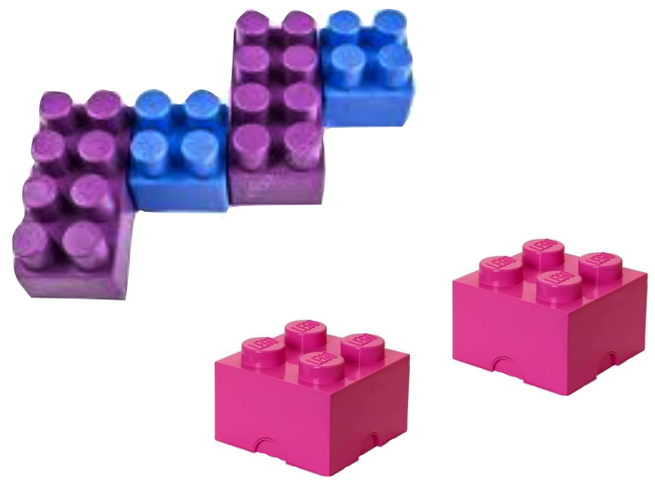
Dickson et al. 2022, in prep.

Learning the building blocks



…CP_{that}…

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

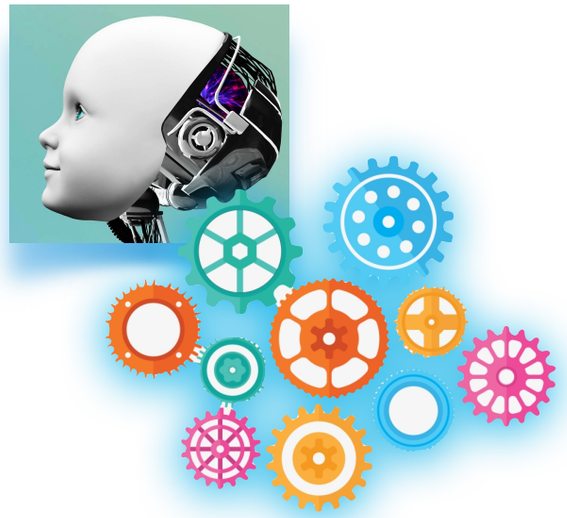


	start-IP-VP-CP_{that}- IP-VP-*end*	embedded | non-island
Whether	* *start*-IP-VP-$CP_{whether}$-IP-VP-*end*	embedded | island
Adjunct	* *start*-IP-VP-CP_{if}- IP-VP-*end*	embedded | island

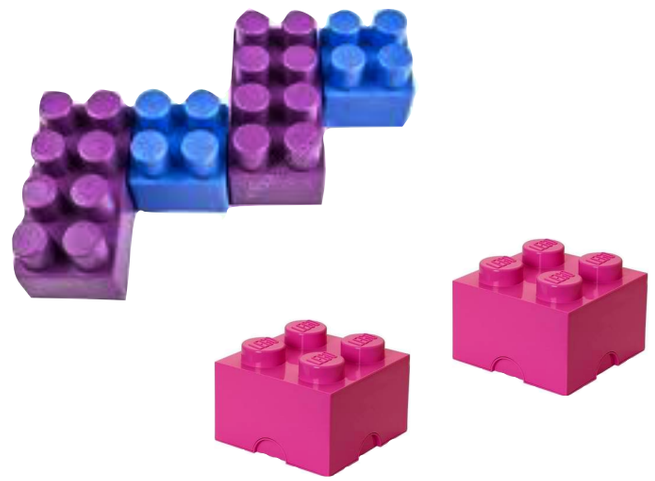
Dickson et al. 2022, in prep.

Learning the building blocks

…CP_{that}…



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

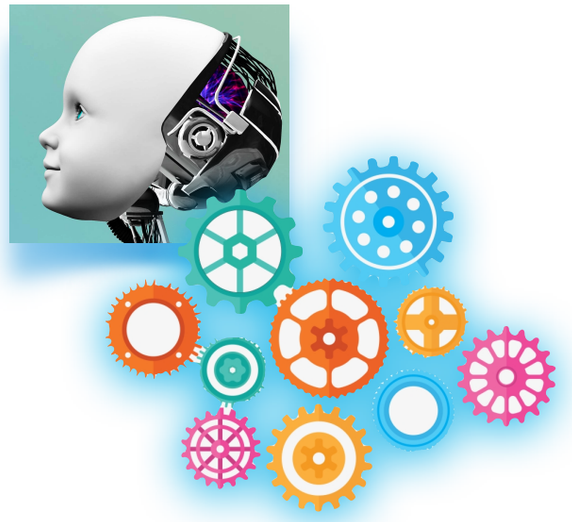


	start-IP-VP-CP_{that}- IP-VP-*end*	embedded | non-island
Whether	* *start*-IP-VP-$CP_{whether}$-IP-VP-*end*	embedded | island
Adjunct	* *start*-IP-VP-CP_{if}- IP-VP-*end*	embedded | island

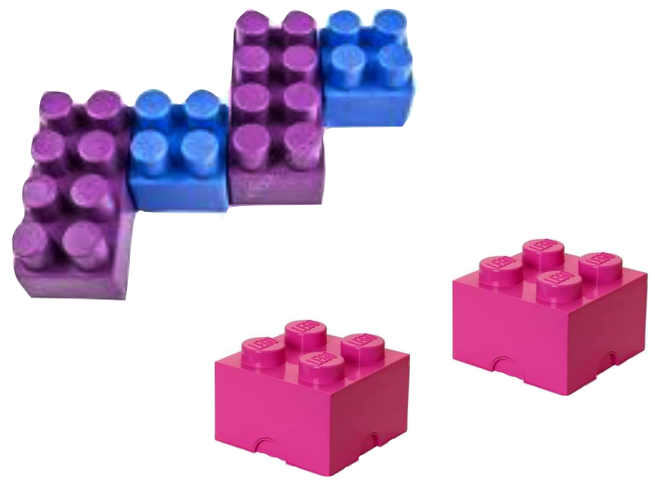
Dickson et al. 2022, in prep.

Learning the building blocks

…CP_{that}…



Is this the only one needed?

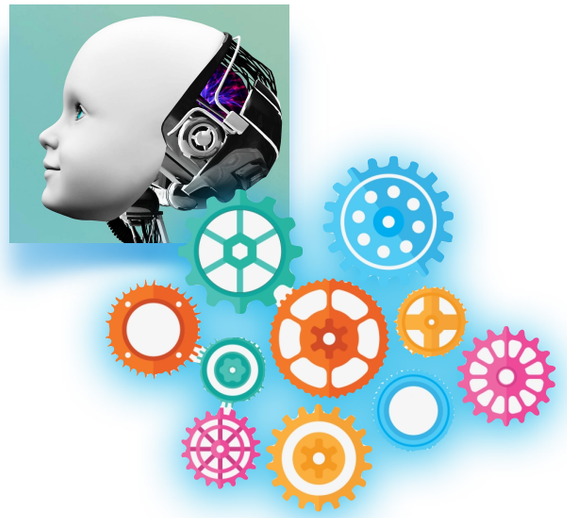


	start-IP-VP-CP_{that}- IP-VP-*end*	embedded | non-island
Whether	* *start*-IP-VP-$CP_{whether}$-IP-VP-*end*	embedded | island
Adjunct	* *start*-IP-VP-CP_{if}- IP-VP-*end*	embedded | island

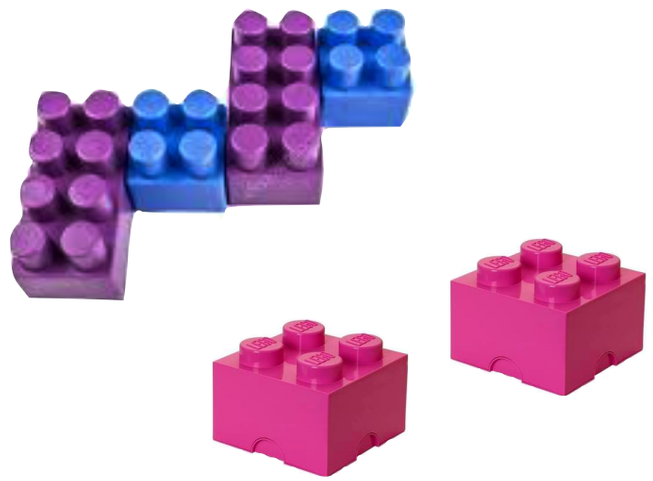
Dickson et al. 2022, in prep.

Learning the building blocks

…CP$_{that}$…



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



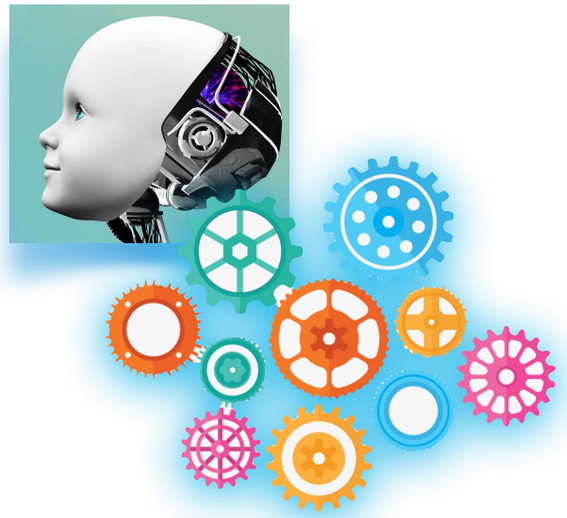
What did she think [that he saw __]?
What did she say [that he saw __]?

What did she whine [that he saw __]?
What did she mumble [that he saw __]?

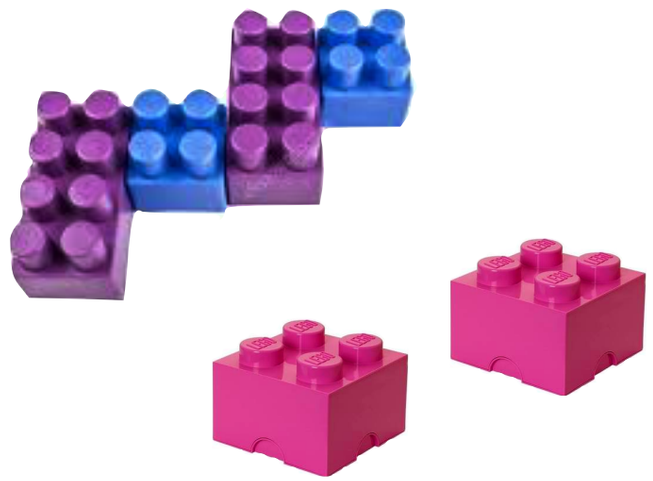
Dickson et al. 2022, in prep.

Learning the building blocks

…CP_{that}…



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



start-IP-VP_{think} -CP_{that}-IP-VP-*end*

start-IP-VP_{say} -CP_{that}-IP-VP-*end*

start-IP-VP_{whine} -CP_{that}-IP-VP-*end*

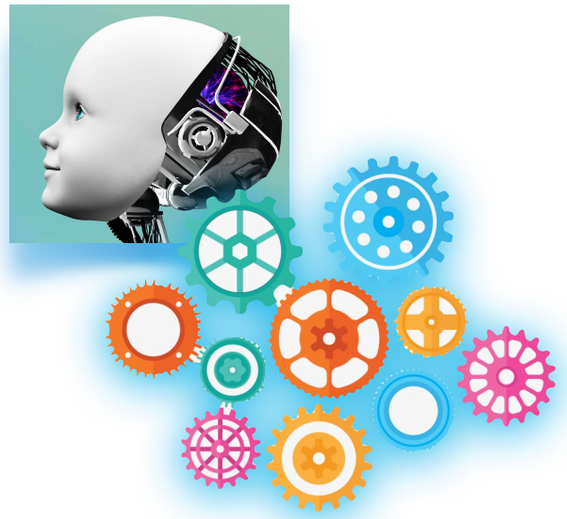
start-IP-VP_{mumble} -CP_{that}-IP-VP-*end*

Dickson et al. 2022, in prep.

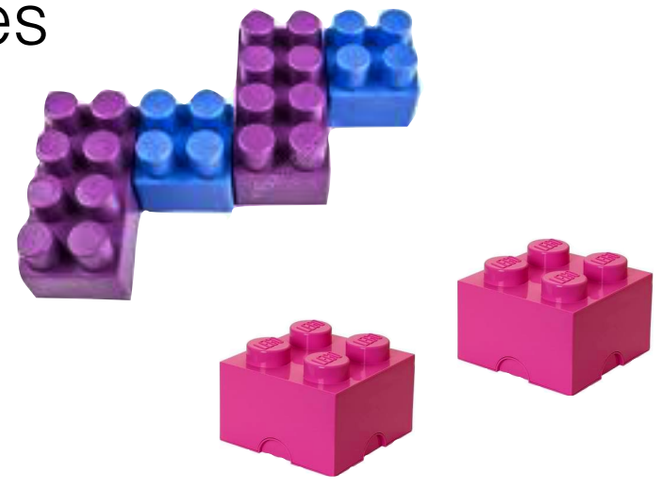
Learning the building blocks

…CP_{that}…

…VP_{think}…

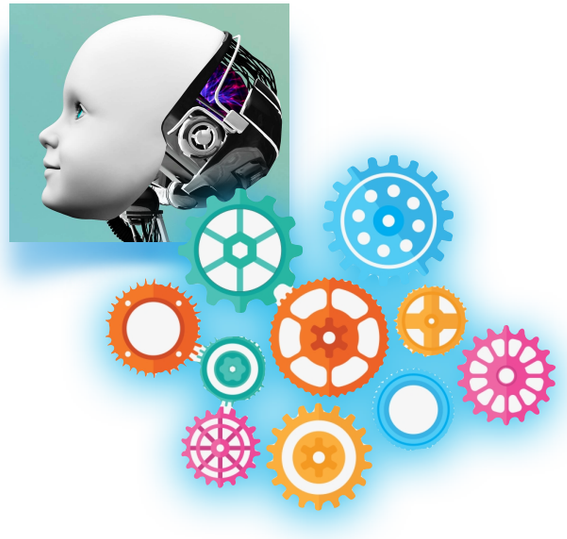


Bigger question: Are there other lexical item types the building blocks need to include?

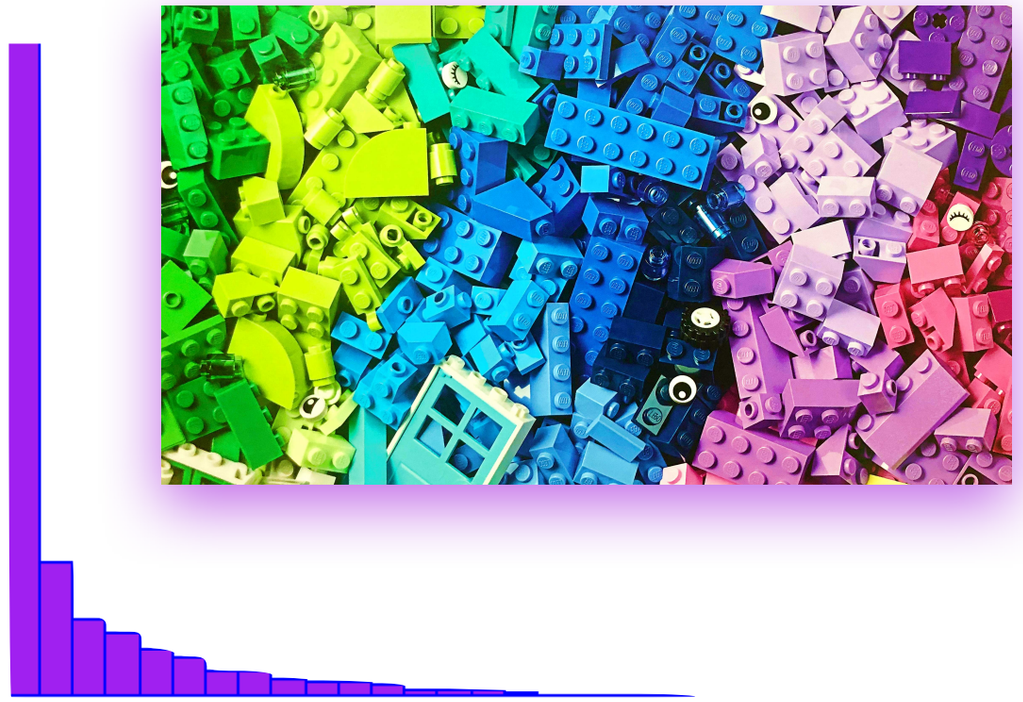


Dickson et al. 2022, in prep.

Learning the building blocks

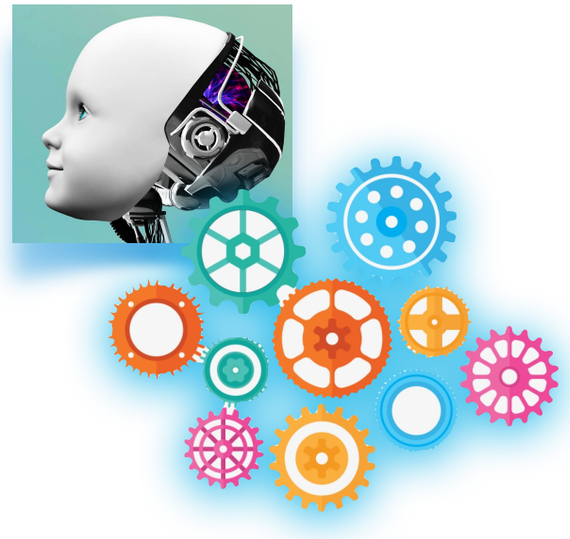


Idea: What if the child is trying to learn
what the best building blocks are
at the same time she learns about their
distributions in the input?



Dickson et al. 2022, in prep.

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

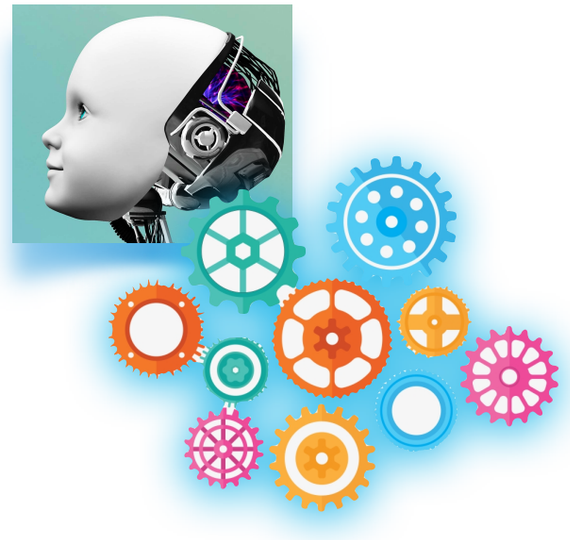
IP-VP-*end*

IP-VP-CP_{that}

VP-CP_{that}-IP

Dickson et al. 2022, in prep.

Learning the building blocks



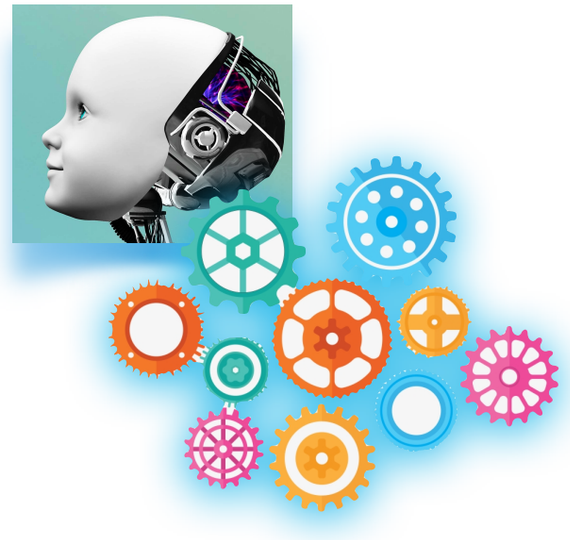
the best building blocks

(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.

Dickson et al. 2022, in prep.

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

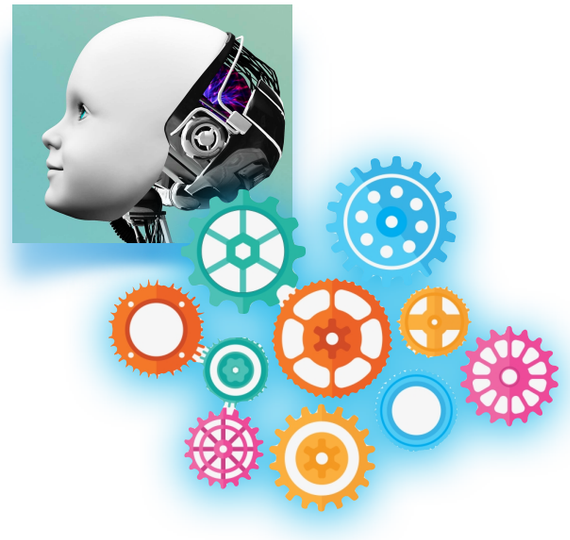
IP-VP-*end*

IP-VP-CP_{that}

VP-CP_{that}-IP

Dickson et al. 2022, in prep.

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}

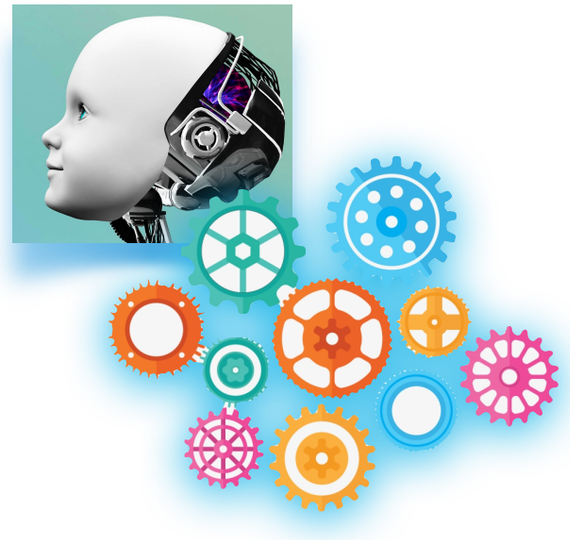
start-IP

IP-VP-*end*

CP_{that}-IP

Dickson et al. 2022, in prep.

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.

start-IP-VP_{think}-CP_{that}

start-IP

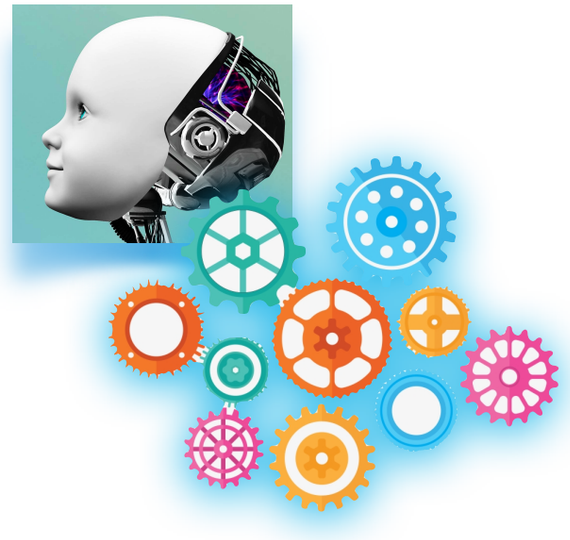
IP-VP-*end*

CP_{that}-IP

start-IP-VP_{say}-CP_{that}

Dickson et al. 2022, in prep.

Learning the building blocks



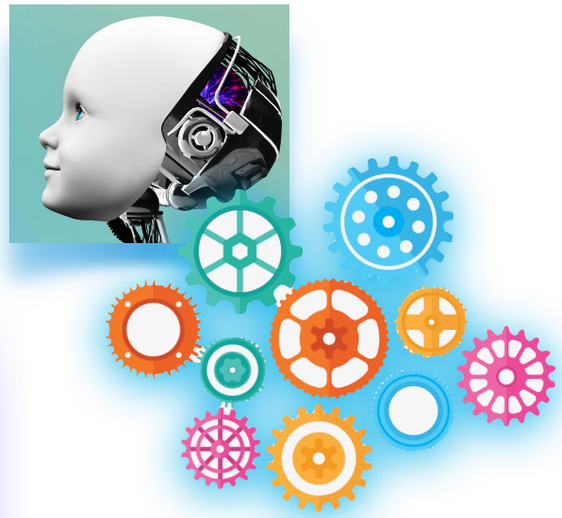
the best building blocks

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

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

Dickson et al. 2022, in prep.

Learning the building blocks

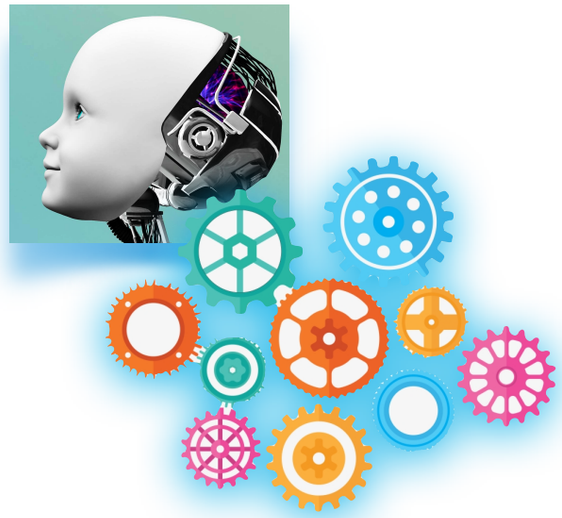


How can the child learn what the best building blocks are?

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

Dickson et al. 2022, in prep.

Learning the building blocks



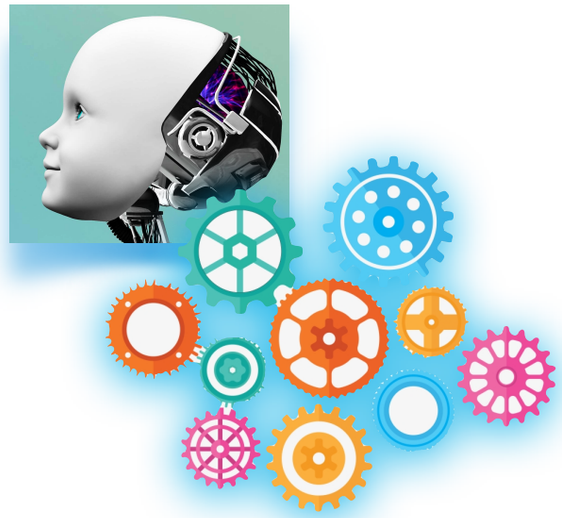
How can the child learn what the best building blocks are?

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

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

Dickson et al. 2022, in prep.

Learning the building blocks



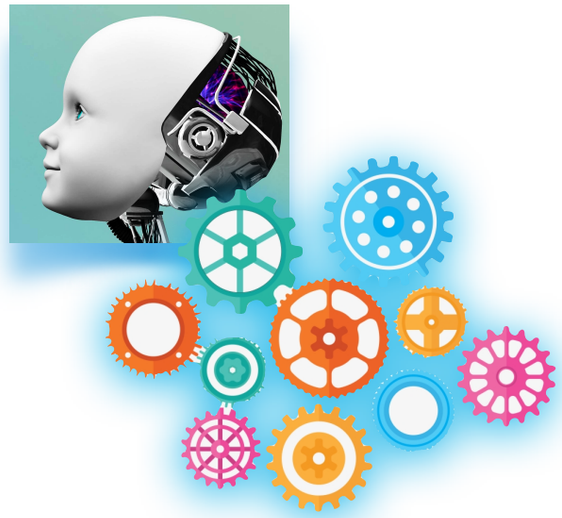
How can the child learn what the best building blocks are?

(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.

Dickson et al. 2022, in prep.

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

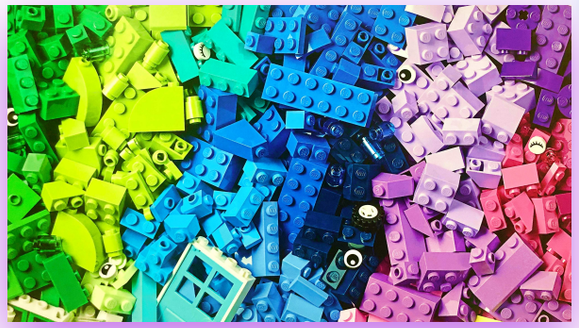
Efficient building blocks allow the representation of current and future *wh*-dependencies to be more probable.

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

Dickson et al. 2022, in prep.

Learning the building blocks

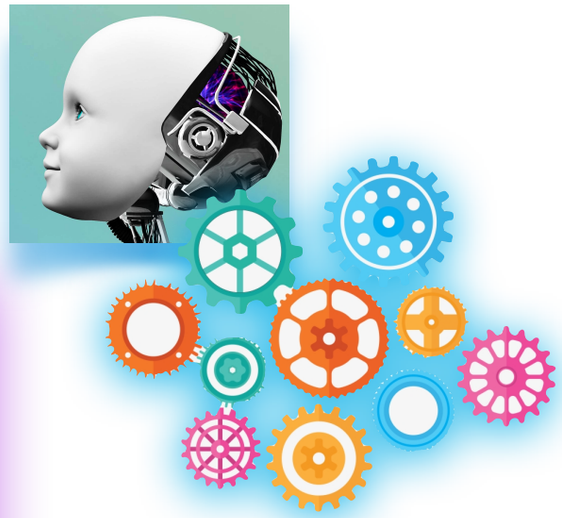
learning efficient building blocks



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.

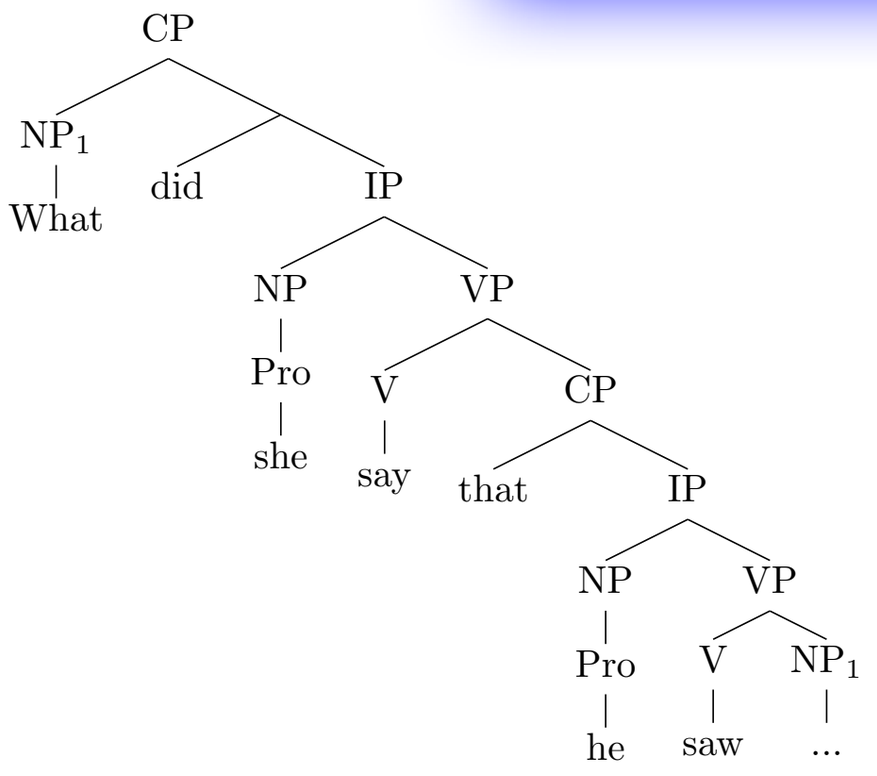
Learning the building blocks



learning efficient building blocks

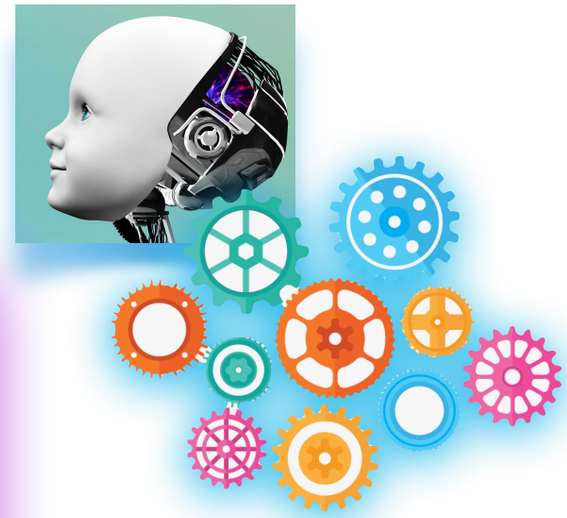
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

What did she say that he saw __ ?



Dickson et al. 2022, in prep.

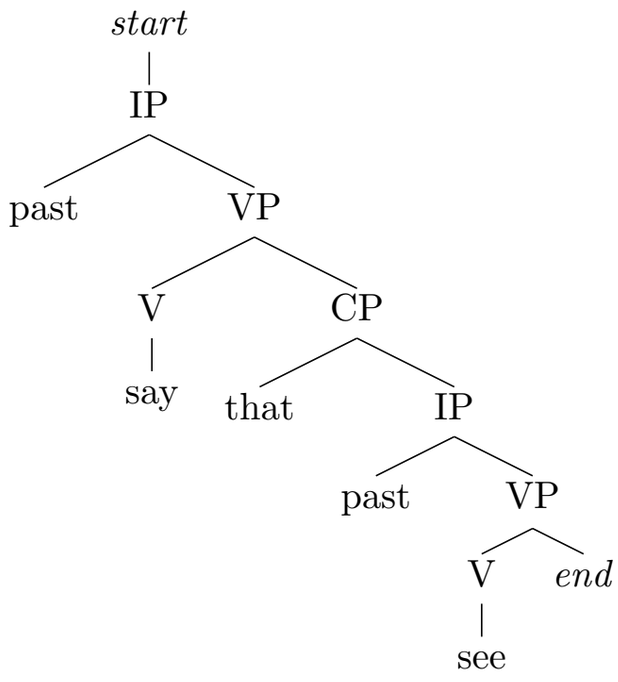
Learning the building blocks



learning efficient building blocks

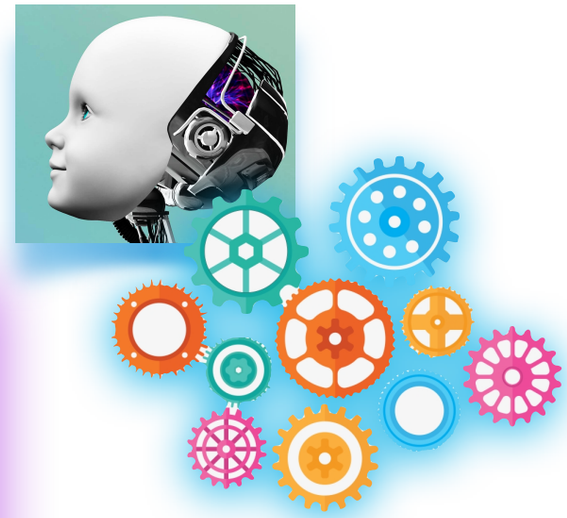
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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



Dickson et al. 2022, in prep.

Learning the building blocks

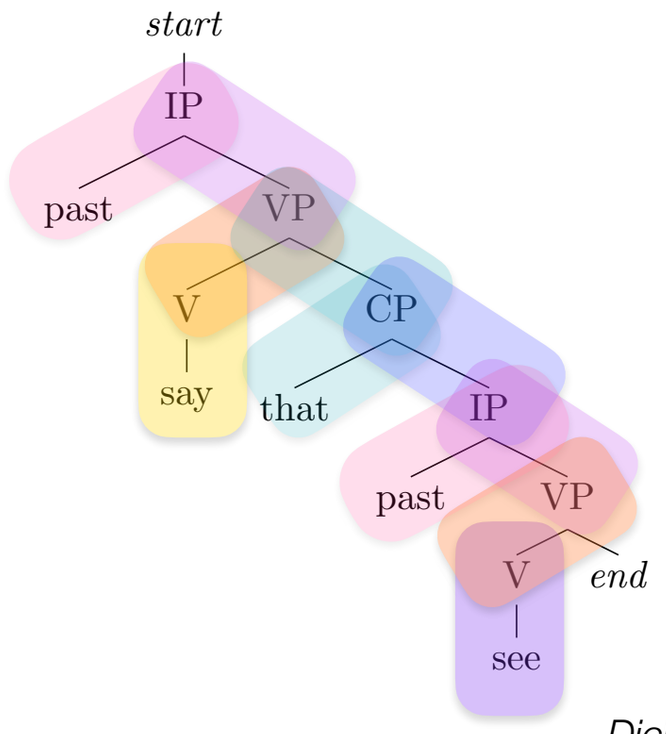


learning efficient building blocks

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

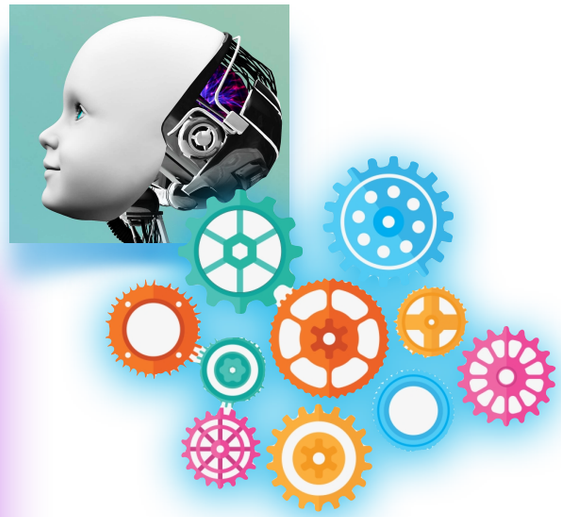
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

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



Dickson et al. 2022, in prep.

Learning the building blocks

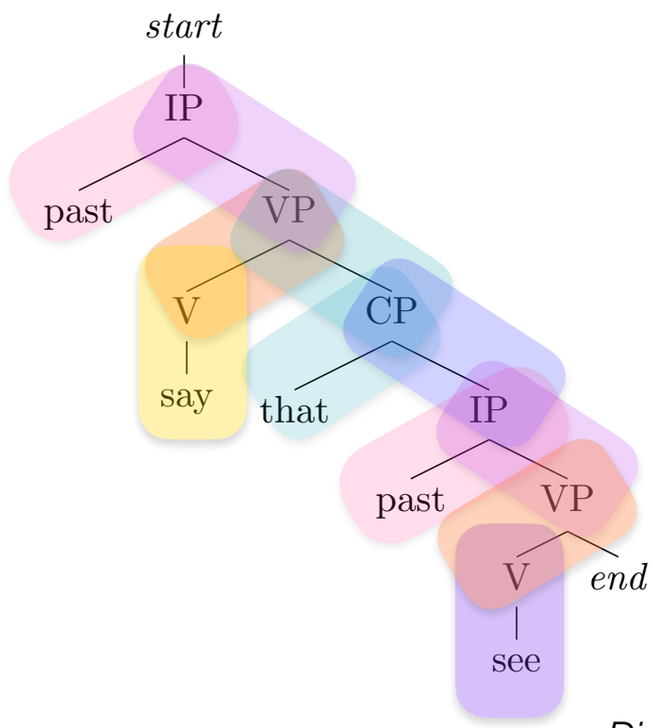


learning efficient building blocks

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

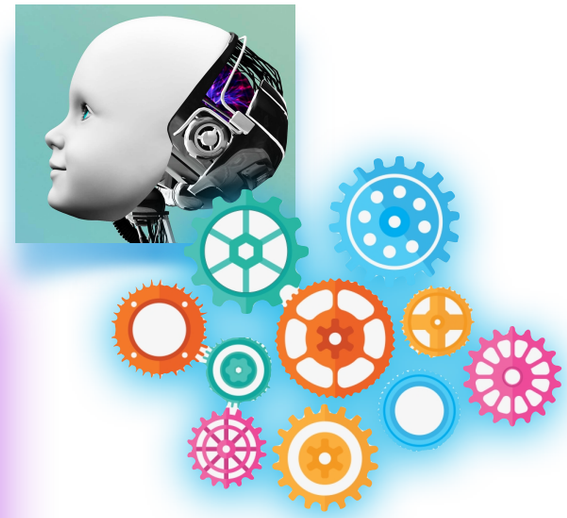
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

It may be slower to put together many small pieces.



Dickson et al. 2022, in prep.

Learning the building blocks



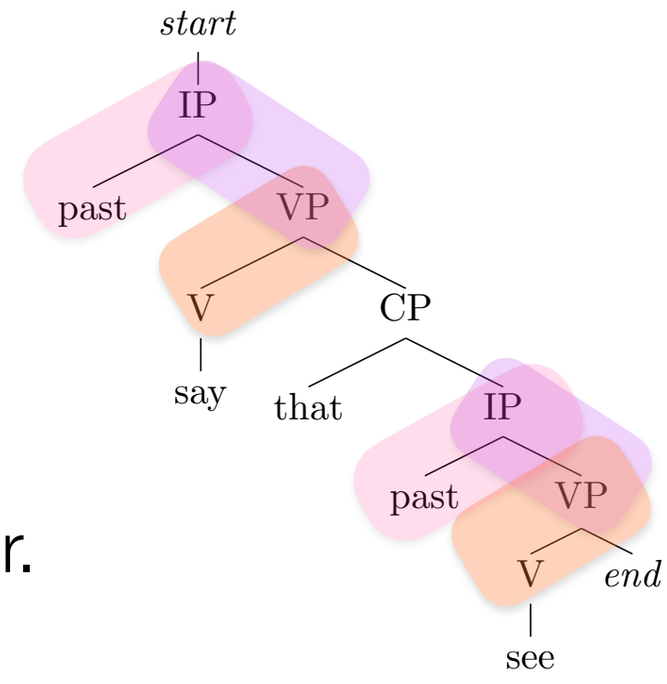
learning efficient building blocks

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

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

many smaller

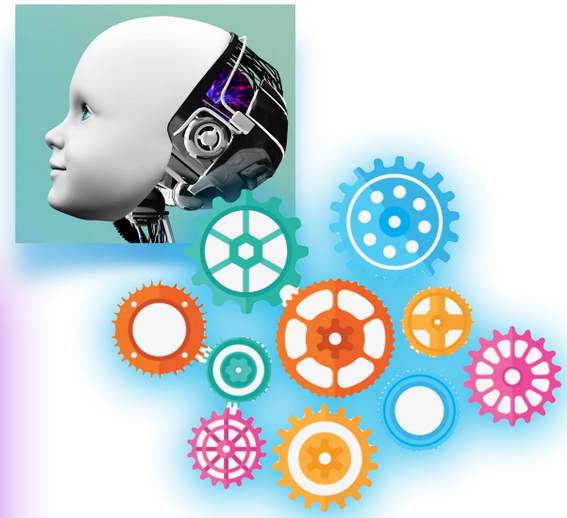
slower because many



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

Dickson et al. 2022, in prep.

Learning the building blocks



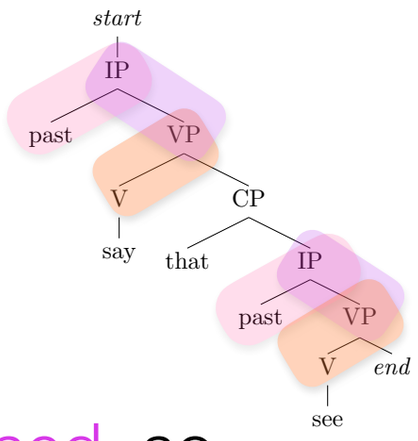
learning efficient building blocks

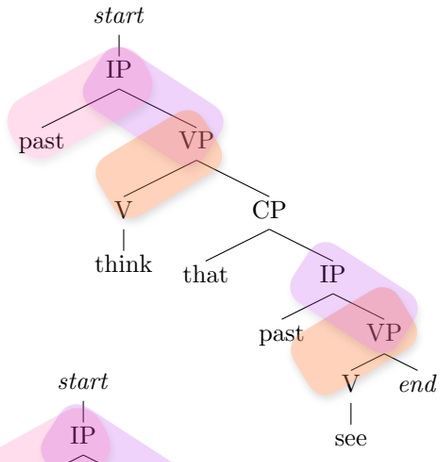
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

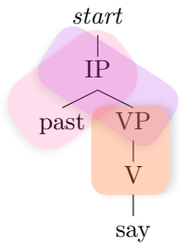
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

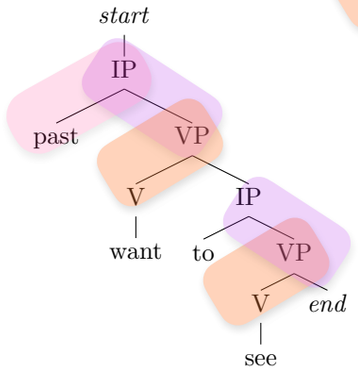
many smaller

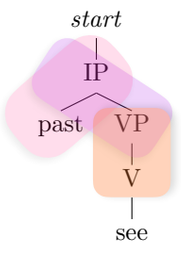
slower because many







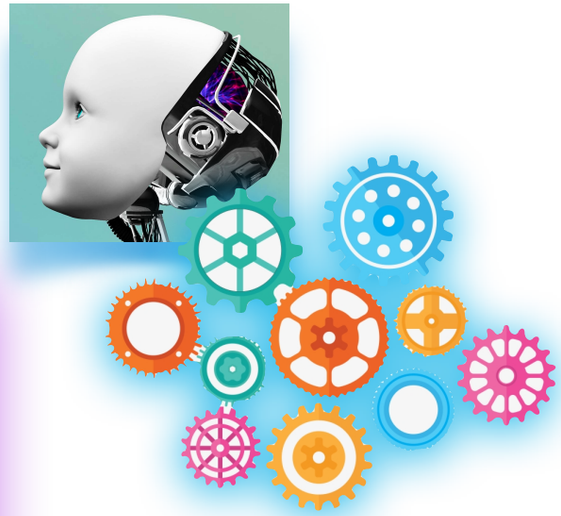




But these pieces may get reused, so that makes them faster to put together.

Dickson et al. 2022, in prep.

Learning the building blocks

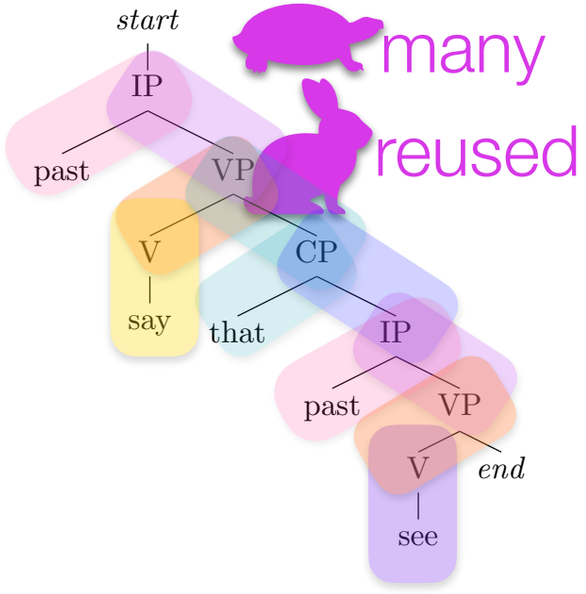


learning efficient building blocks

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

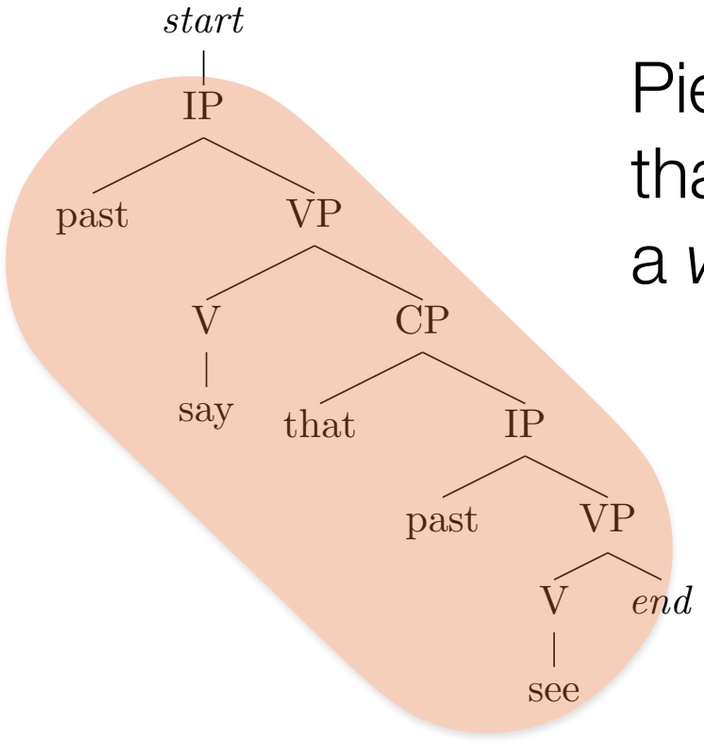
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller



many

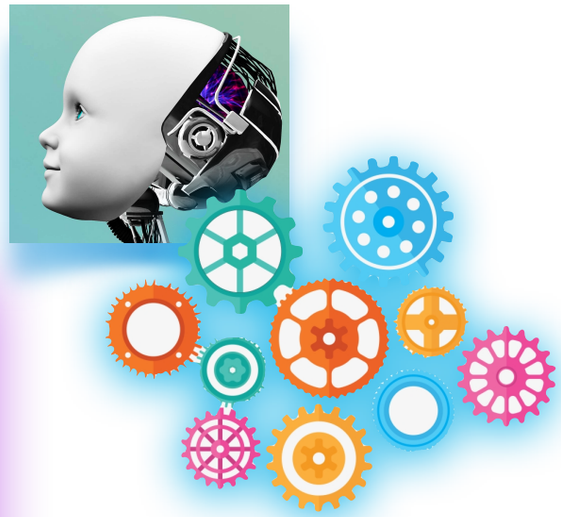
reused



Pieces can be big, so that only one makes up a *wh*-depdendency

Dickson et al. 2022, in prep.

Learning the building blocks

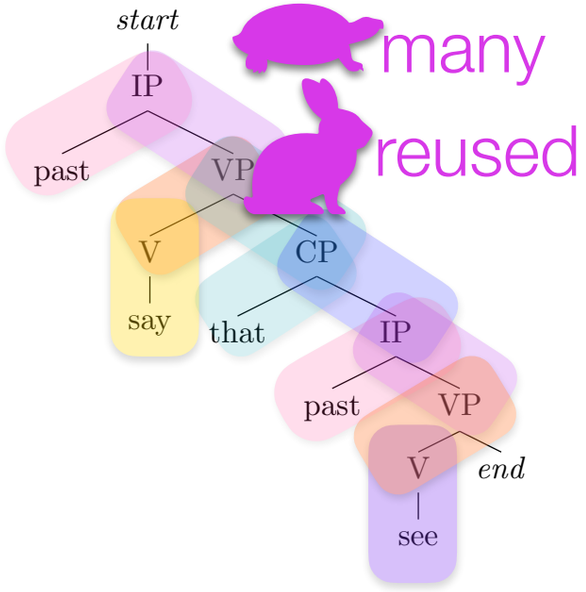


learning efficient building blocks

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

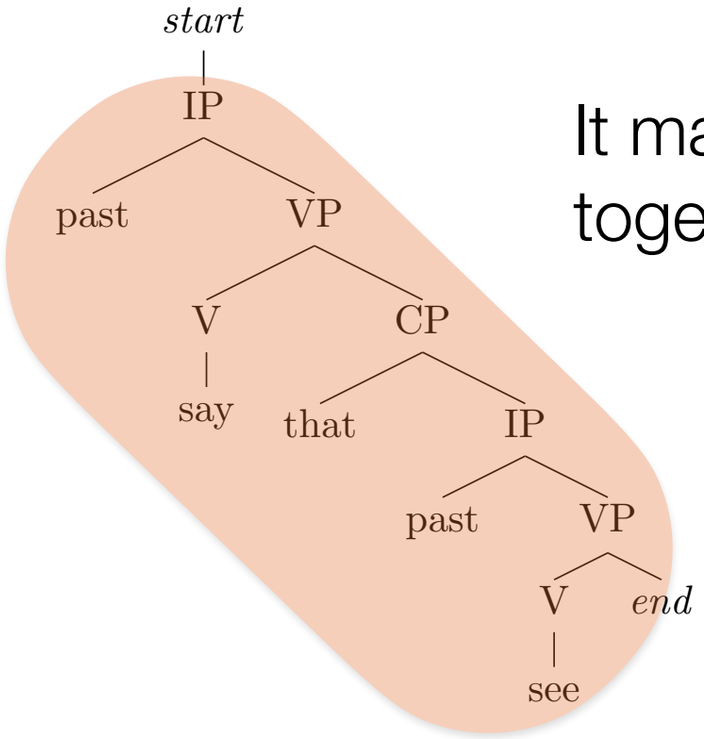
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller



many

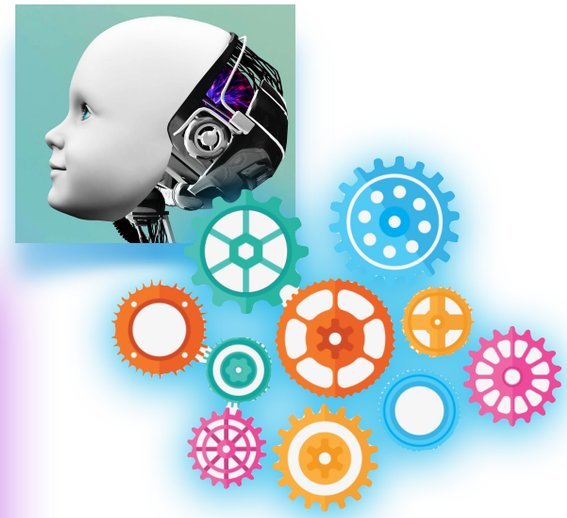
reused



It may be faster to put together one big piece.

Dickson et al. 2022, in prep.

Learning the building blocks

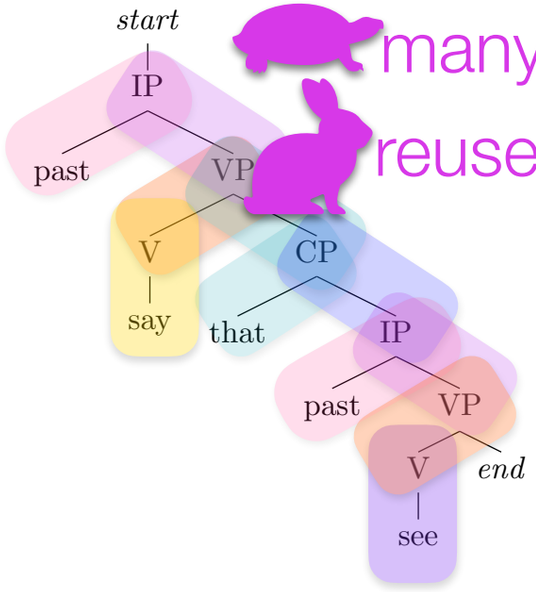


learning efficient building blocks

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

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

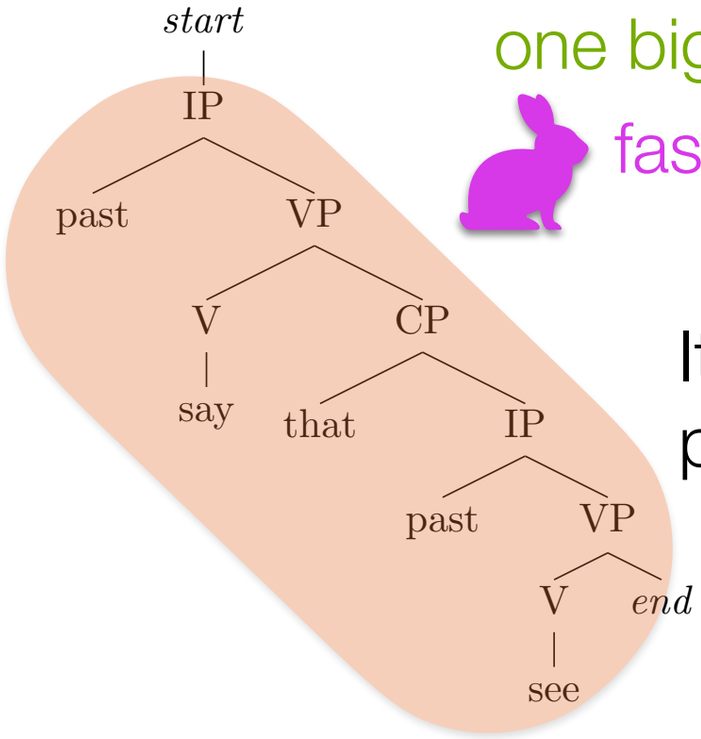
many smaller



many

reused

one big

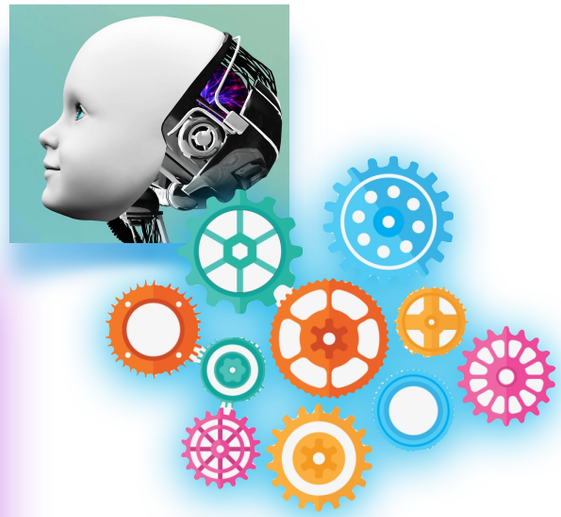


faster because one

It may be slower if the piece is used rarely.

Dickson et al. 2022, in prep.

Learning the building blocks

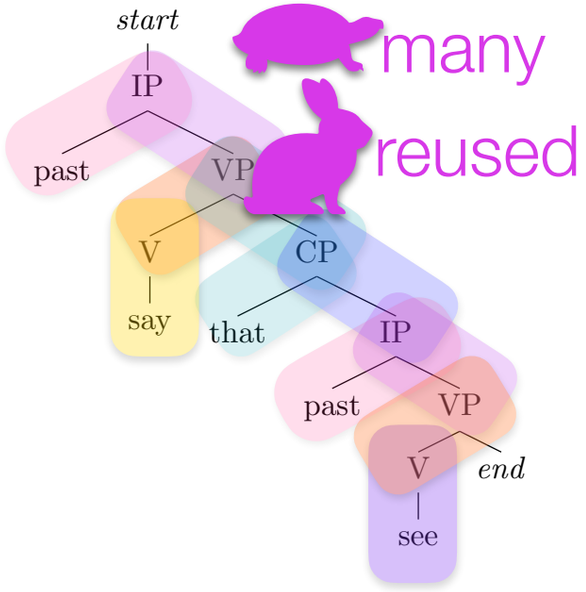


learning efficient building blocks

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

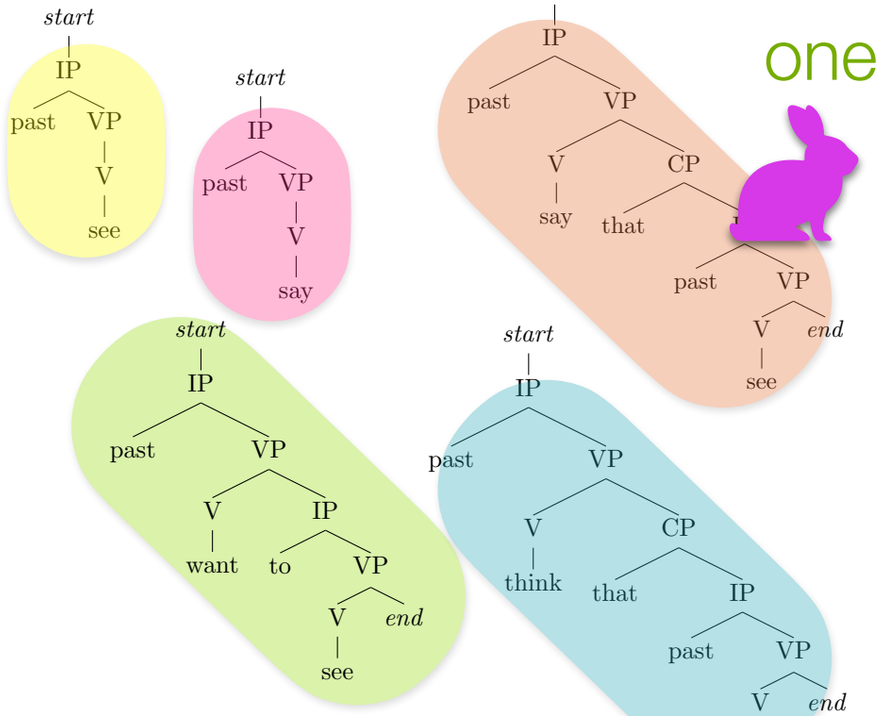
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller



many

reused



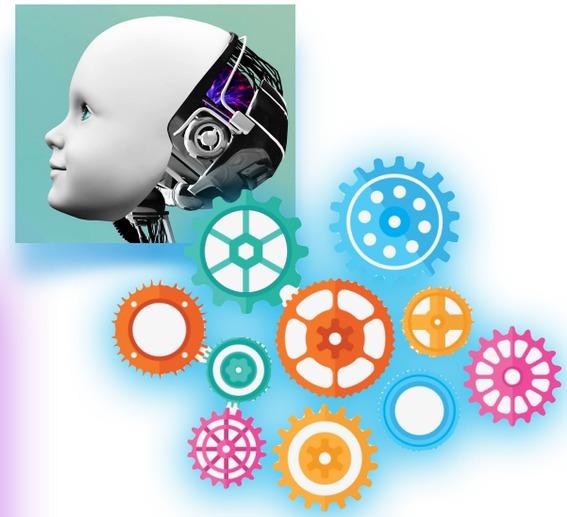
one big

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

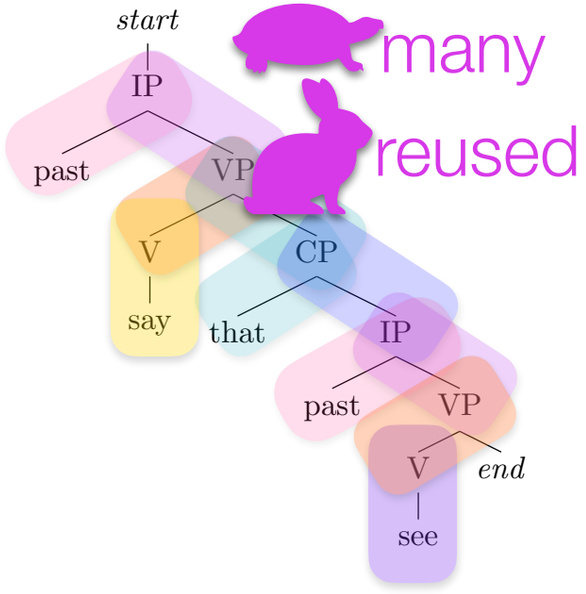
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

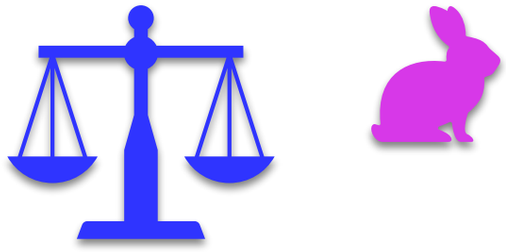
many smaller

many

reused



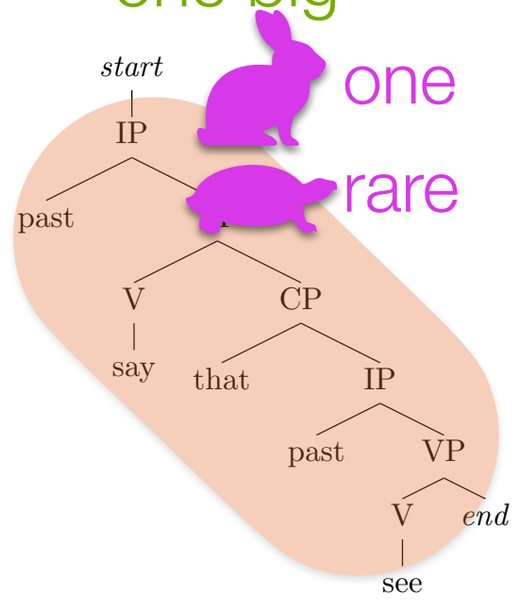
The most efficient option is probably a balance of bigger and smaller blocks that collectively are faster to put together.



one big

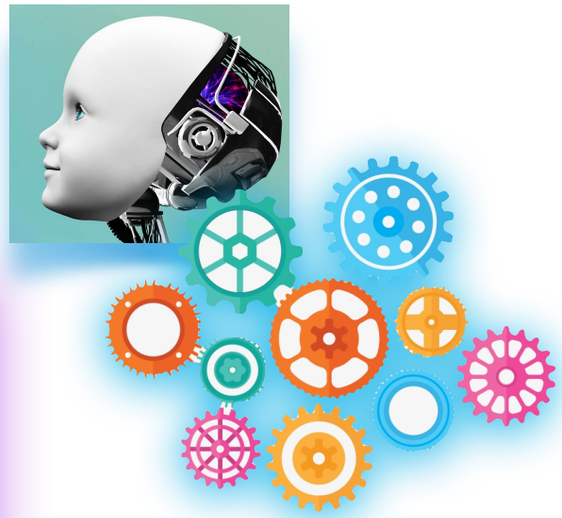
one

rare



Dickson et al. 2022, in prep.

Learning the building blocks

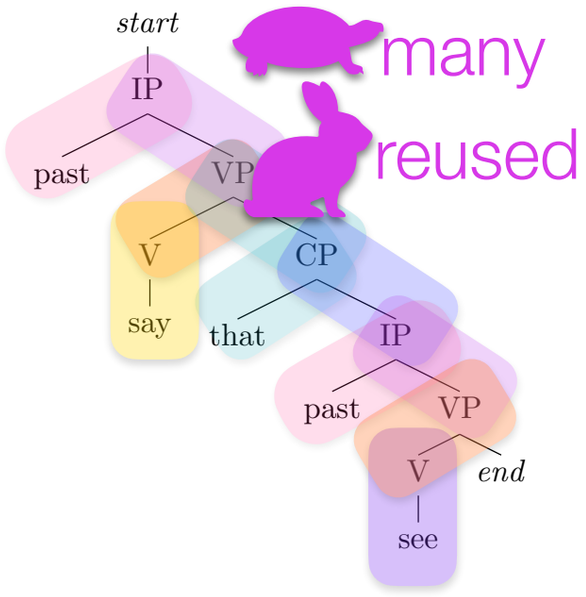


learning efficient building blocks

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

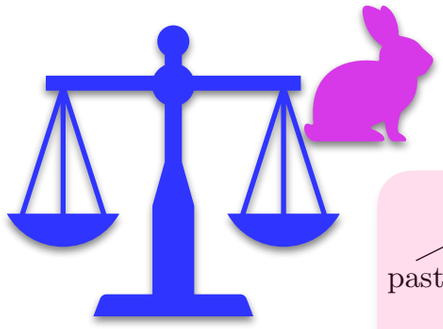
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

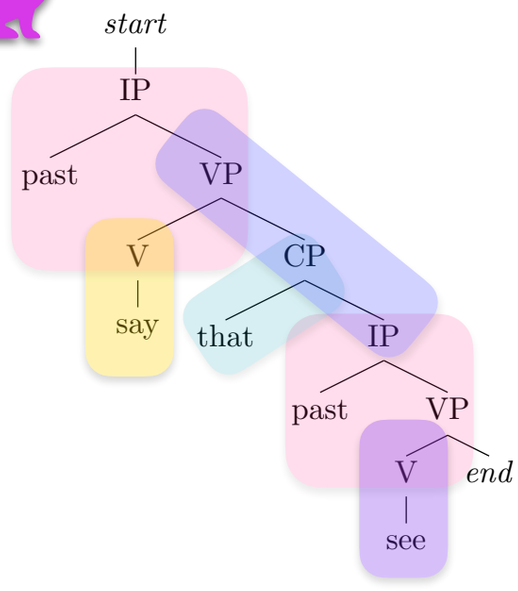
many smaller



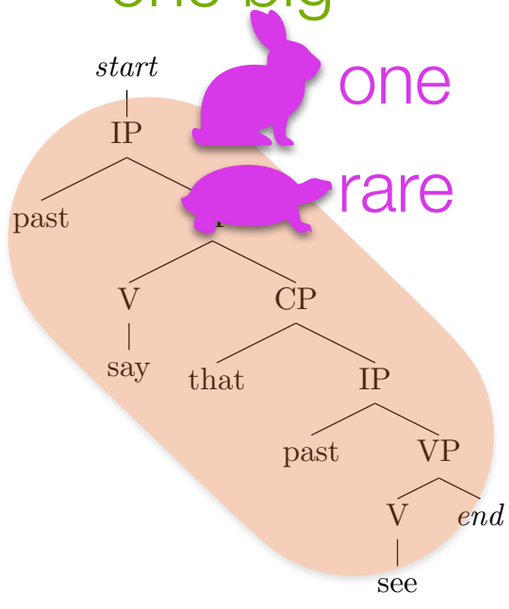
many

reused





one big



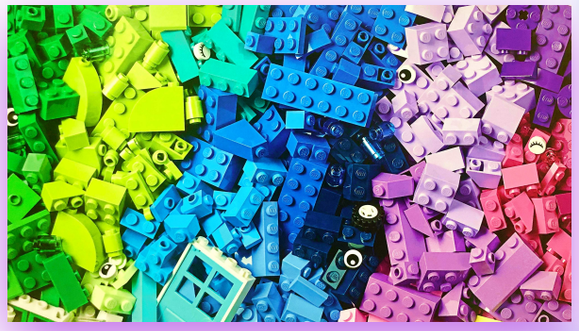
one

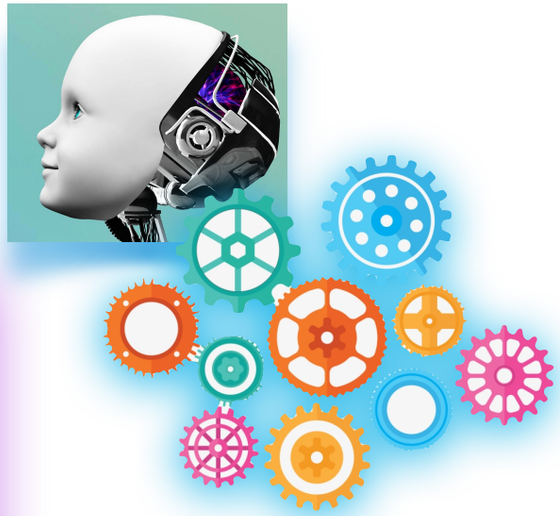
rare

Dickson et al. 2022, in prep.

Learning the building blocks

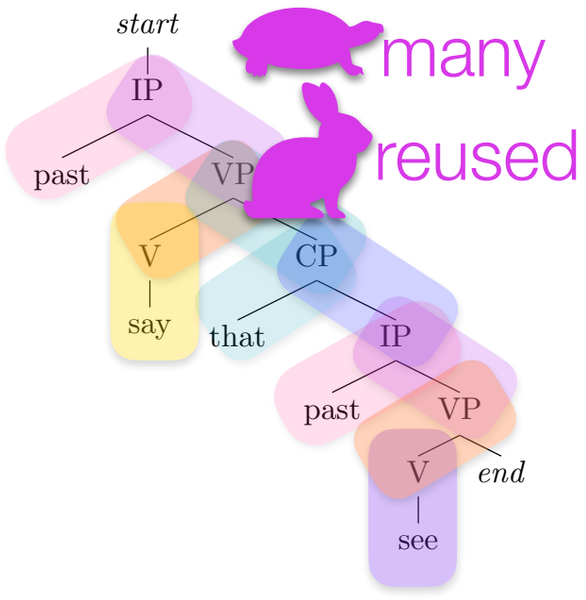
learning efficient building blocks





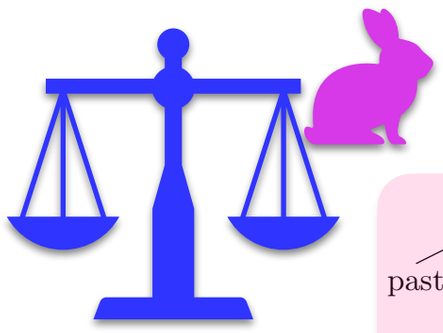
How can children find the best balance?

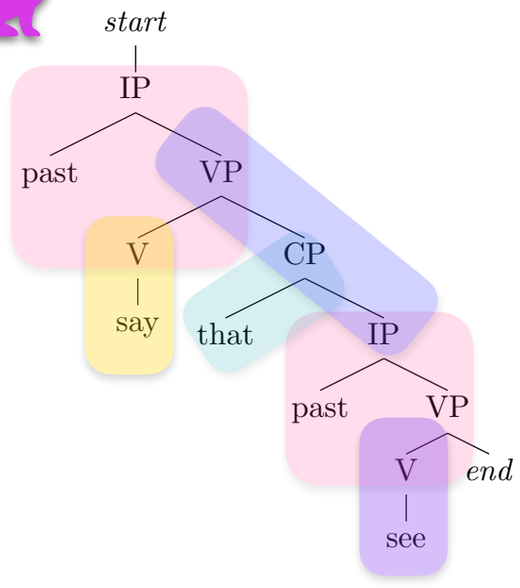
many smaller



many

reused

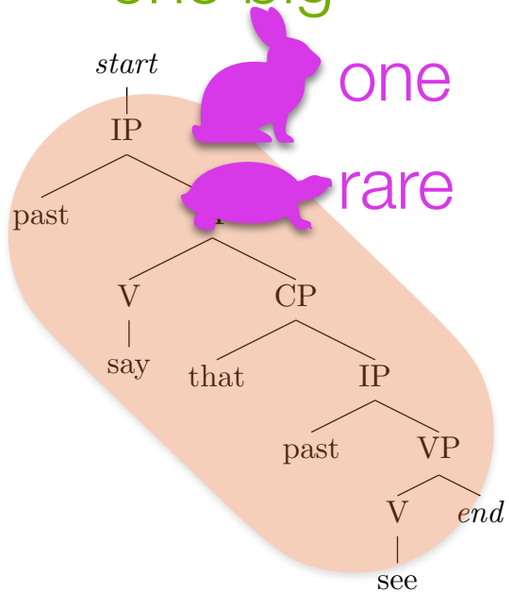




one big

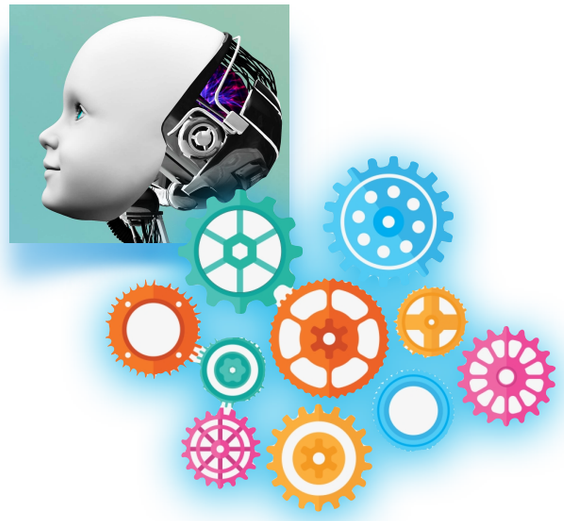
one

rare



Dickson et al. 2022, in prep.

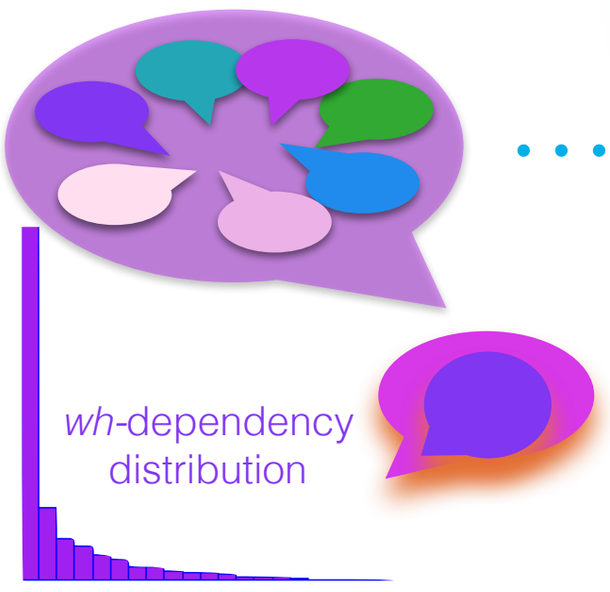
Learning the building blocks



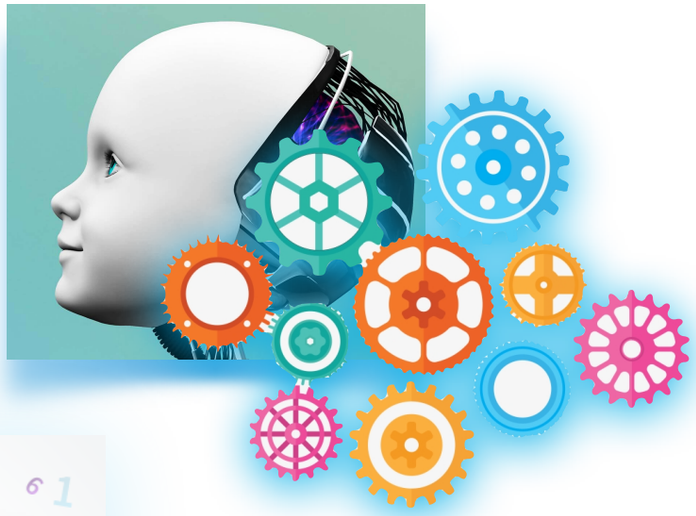
Use Bayesian inference to search through the hypothesis space of all possible building blocks (O'Donnell 2015) and find an efficient set for children's input.

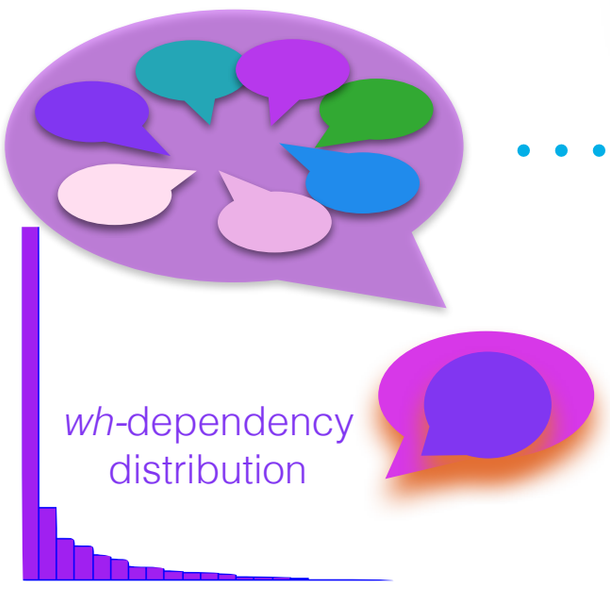
Dickson et al. 2022, in prep.

So that's what the modeled child will do



Dickson et al. 2022, in prep.

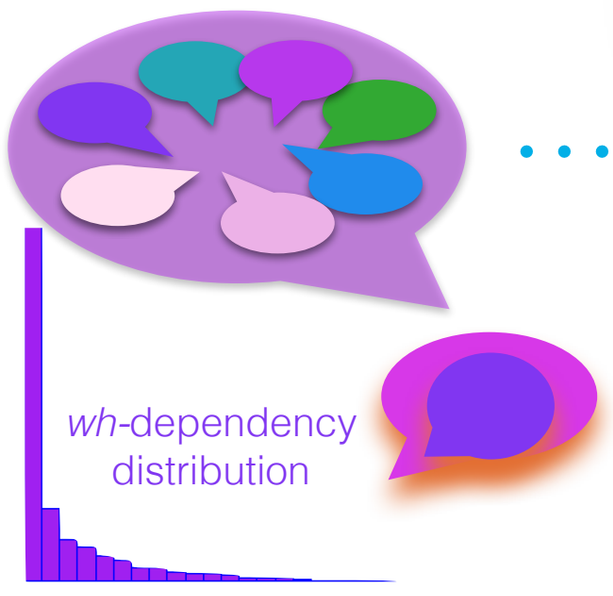


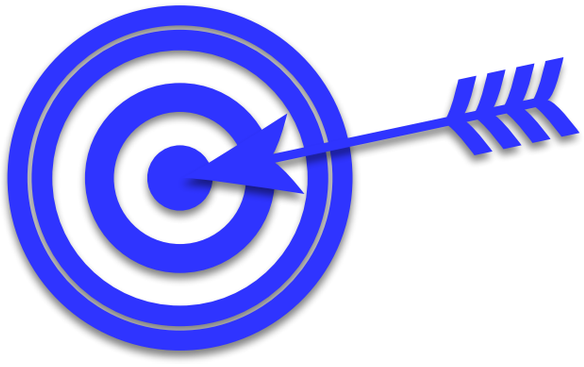


There's additional target behavior about *wh*-dependencies we'd like to capture.

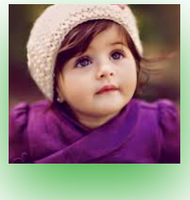
Dickson et al. 2022, in prep.

Who does





Before:
Adult judgments + child preferences of
certain *wh*-dependencies



Dickson et al. 2022, in prep.

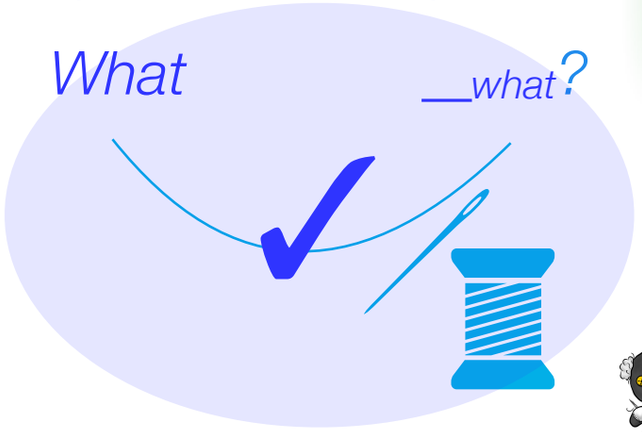
Before: Target behavior

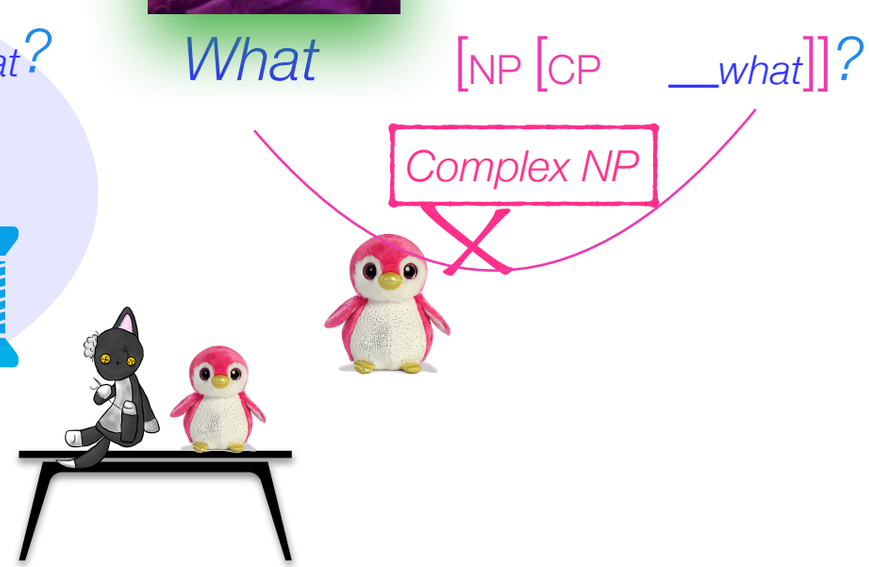
certain *wh*-dependencies

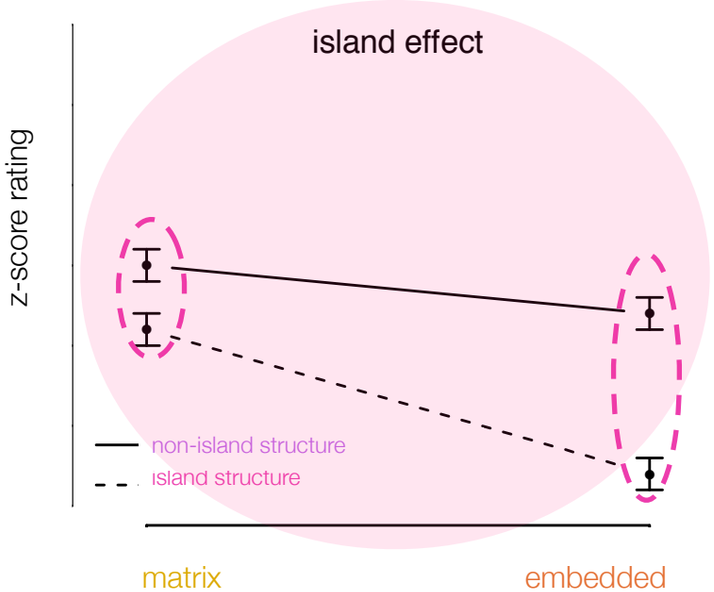
- *Subject island*
- *Complex NP island*
- *Whether island*
- *Adjunct island*

certain *wh*-dependencies

Children prefer this interpretation.







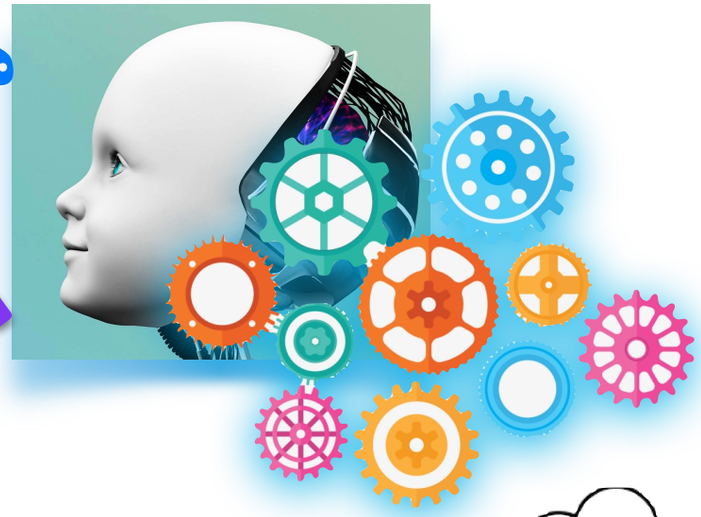
higher-SES *Sprouse et al. 2012*

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

higher-SES *De Villiers et al. 2008*

lower-SES

Dickson et al. 2022, in prep.



certain *wh*-dependencies

Subject island

Complex NP island

Whether island

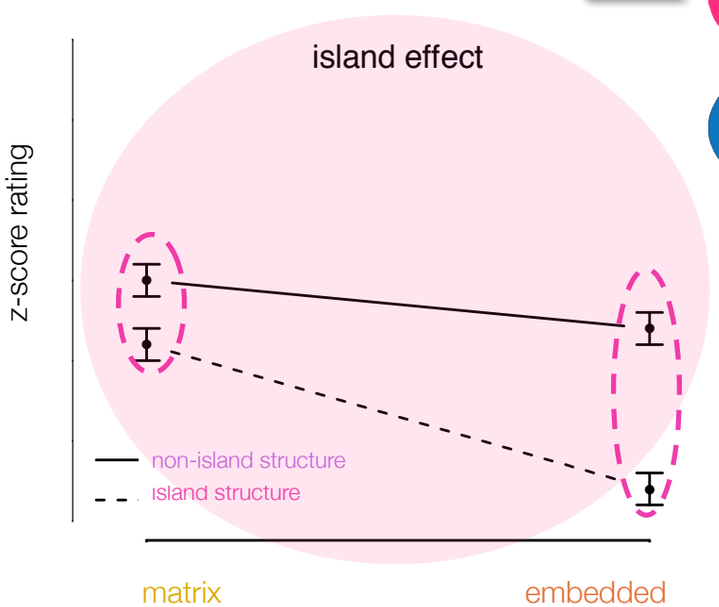
Adjunct island

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

What did she 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 __]?

Liu et al. 2019



Dickson et al. 2022, in prep.

certain *wh*-dependencies

Subject island

Complex NP island

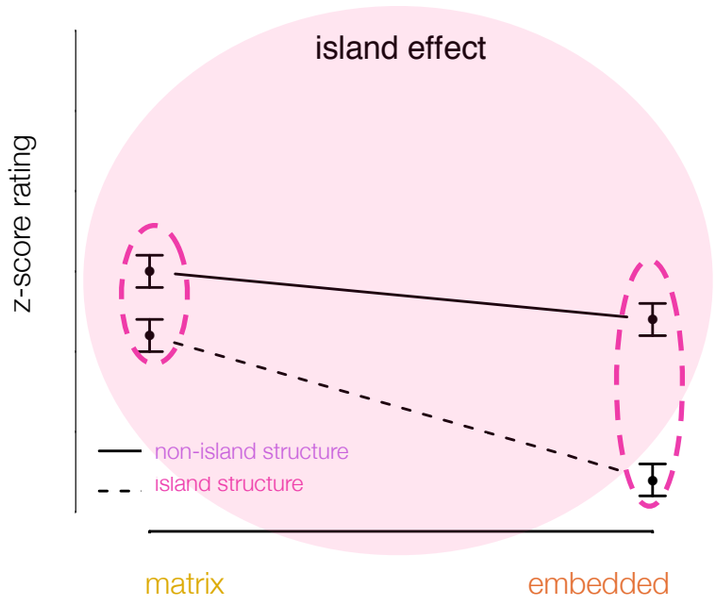
Whether island

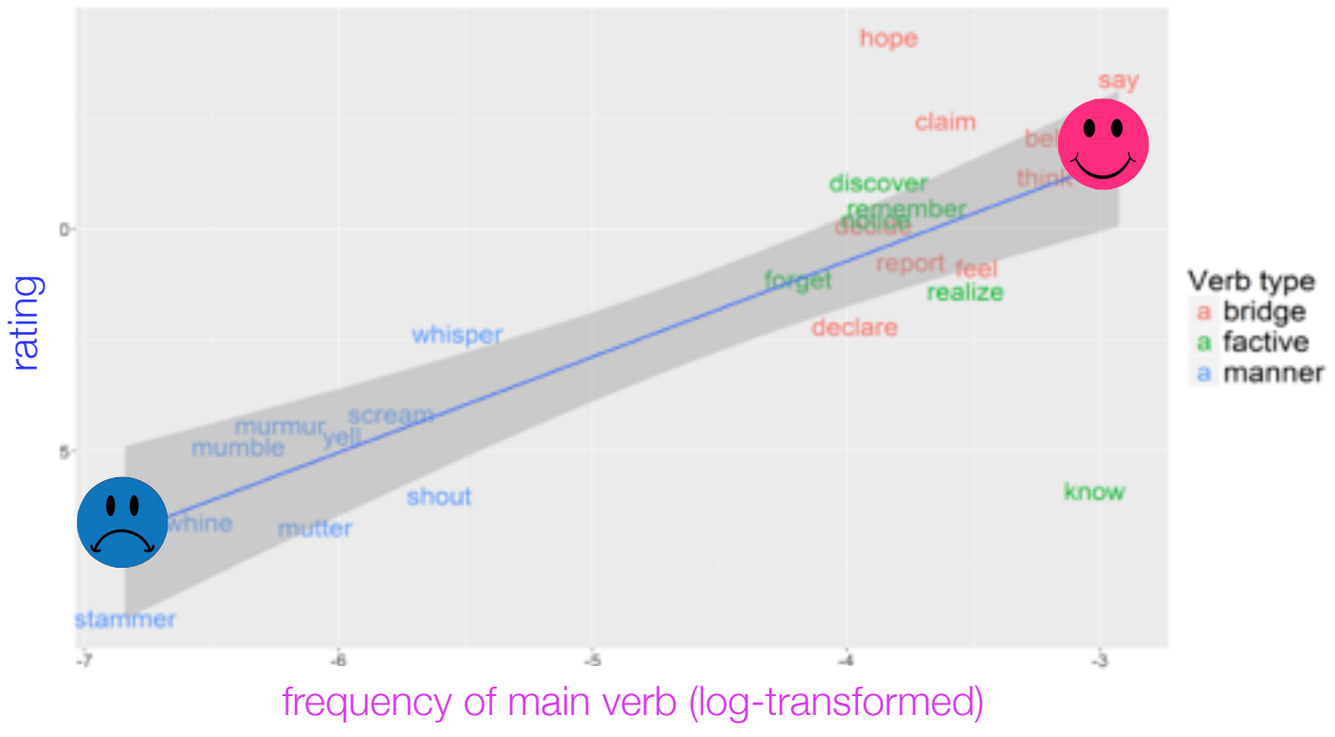
Adjunct island

+ 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

Subject island

Complex NP island

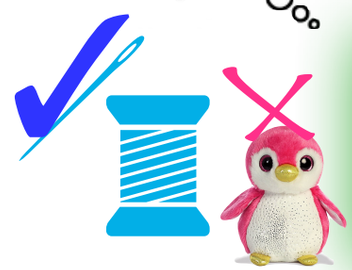
Whether island

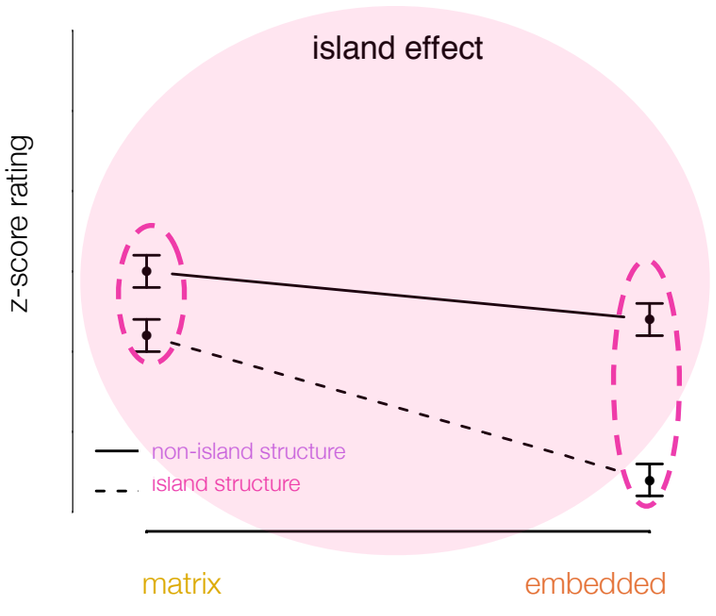
Adjunct island

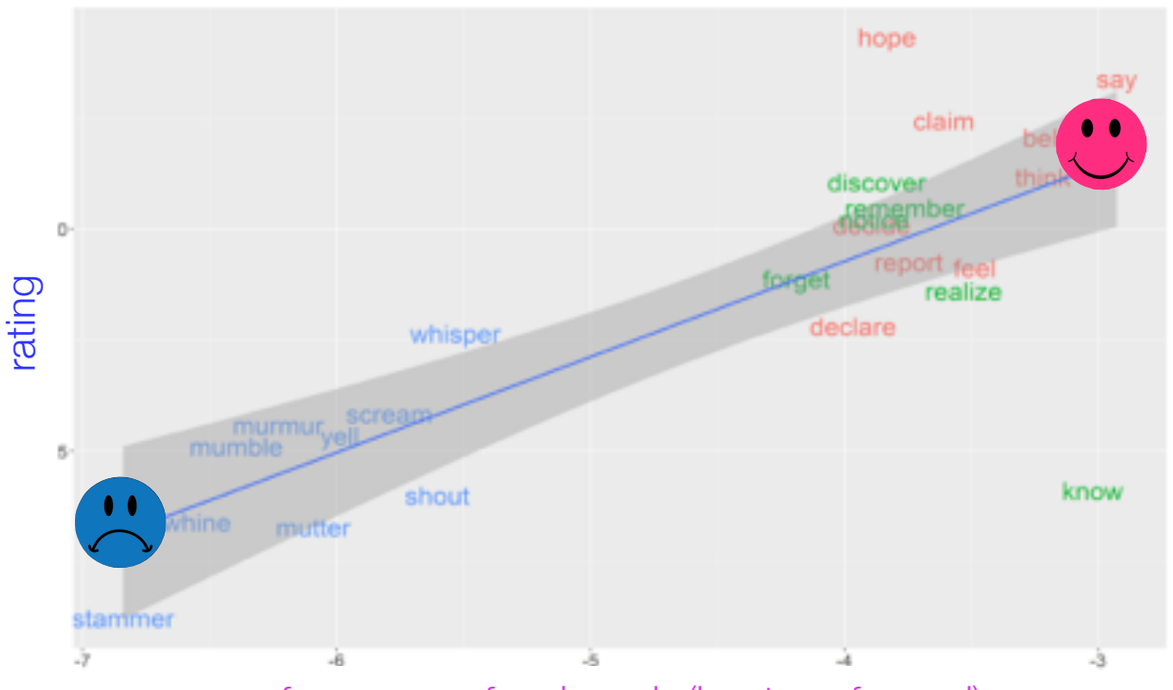
+ additional
target behavior
with *wh*-dependencies that
vary main verb frequency

What did she VERB [that he saw __]?

Liu et al. 2019



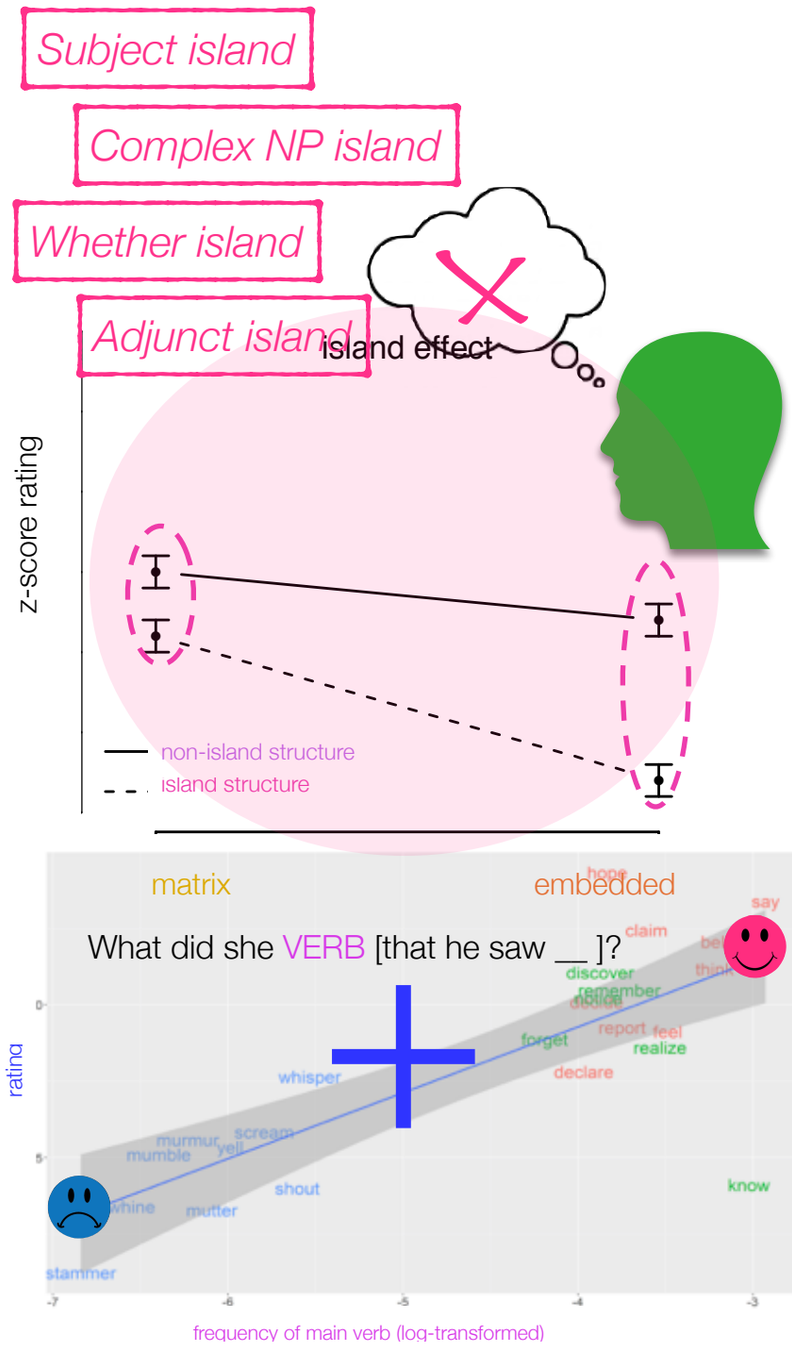


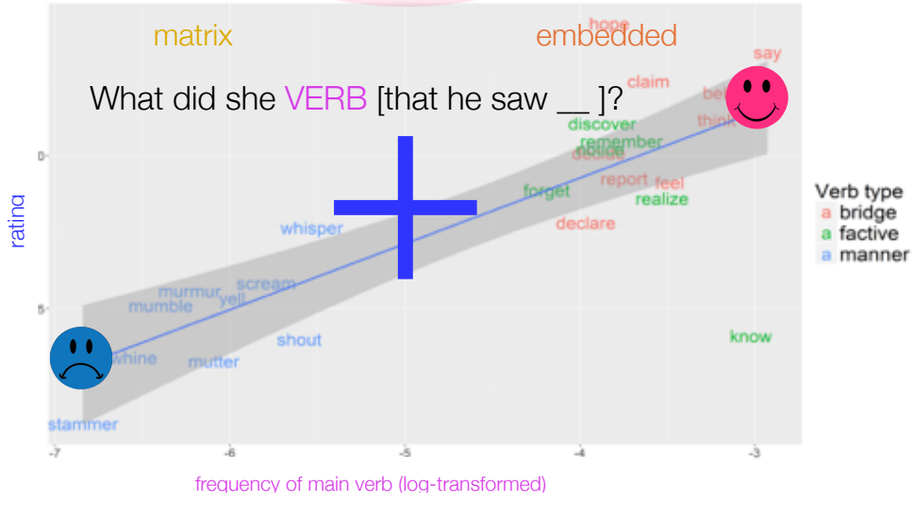


Important pattern:
Positive correlation between main verb frequency and judged acceptability.

Dickson et al. 2022, in prep.

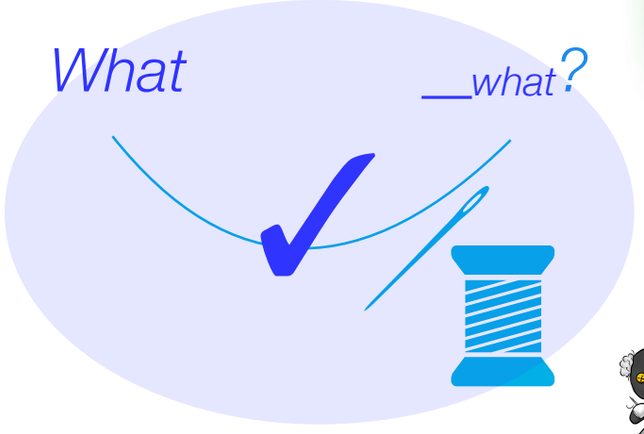
Before: Target behavior

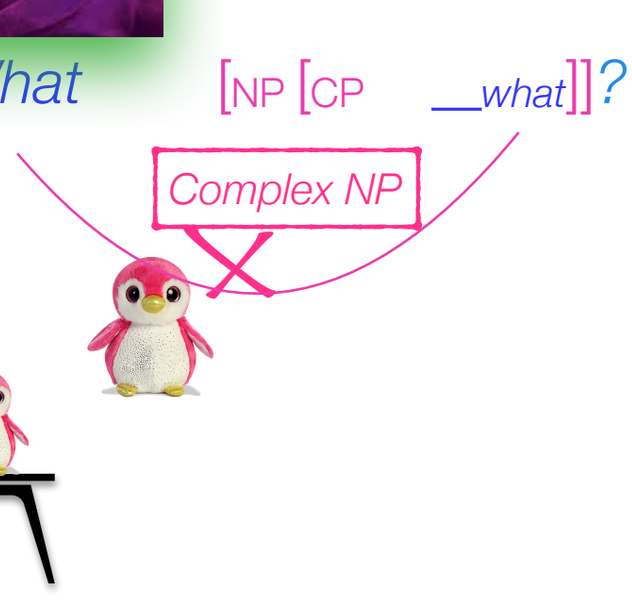




certain *wh*-dependencies

Children prefer this interpretation.





higher-SES

lower-SES

De Villiers et al. 2008

Dickson et al. 2022, in prep.

+ additional
target behavior
with other *wh*-dependencies

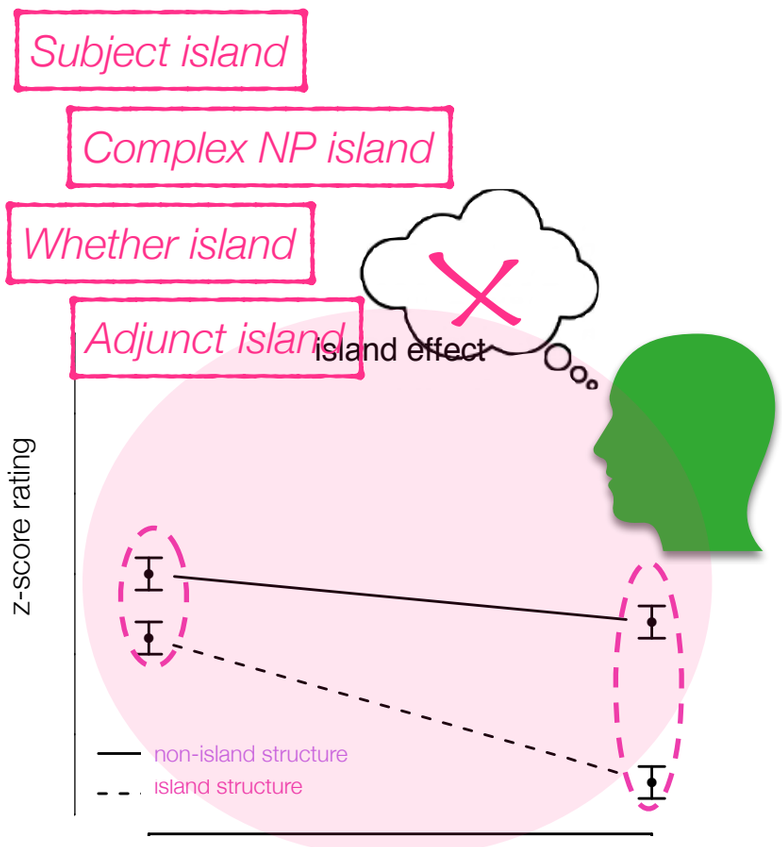
Who did the little sister ask how to see?

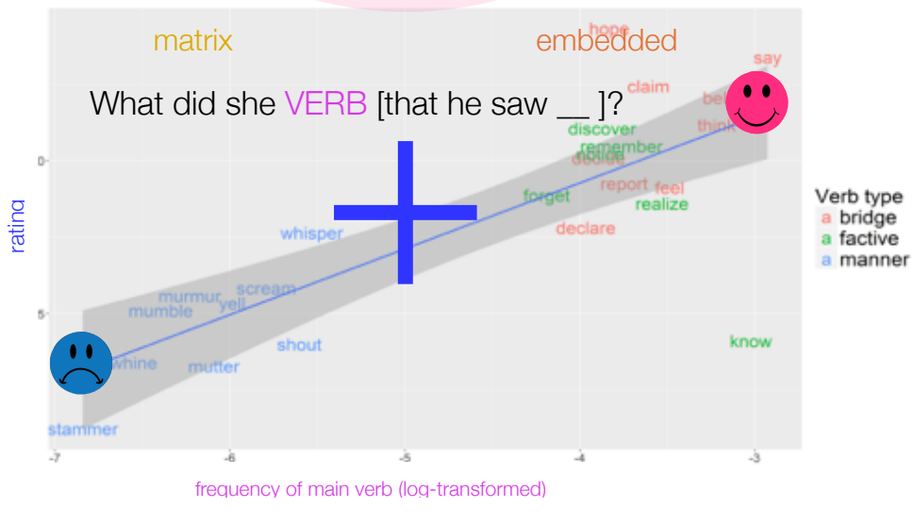
Who did the boy ask what to bring?

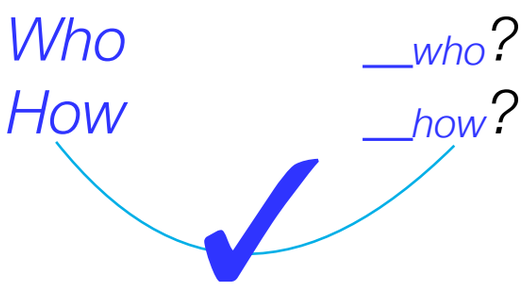
How did the mom learn what to bake?

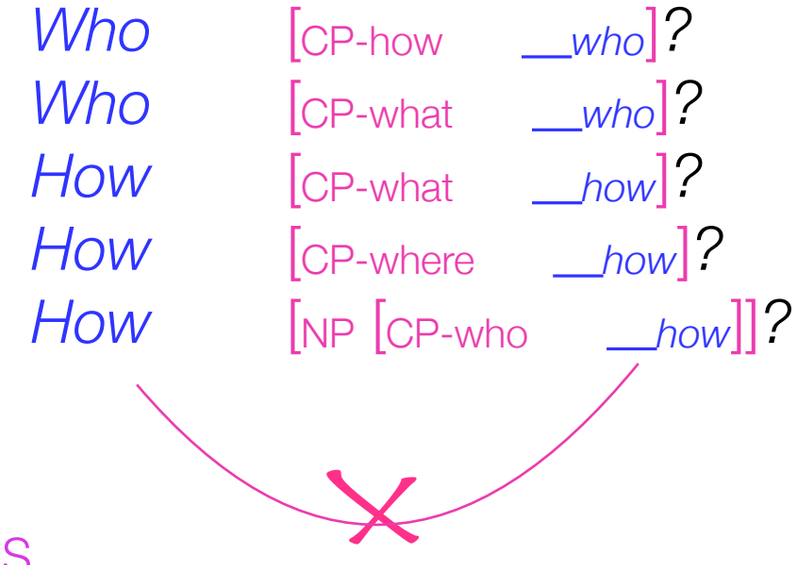
How did the girl ask where to ride?

How did the boy who sneezed drink the milk?







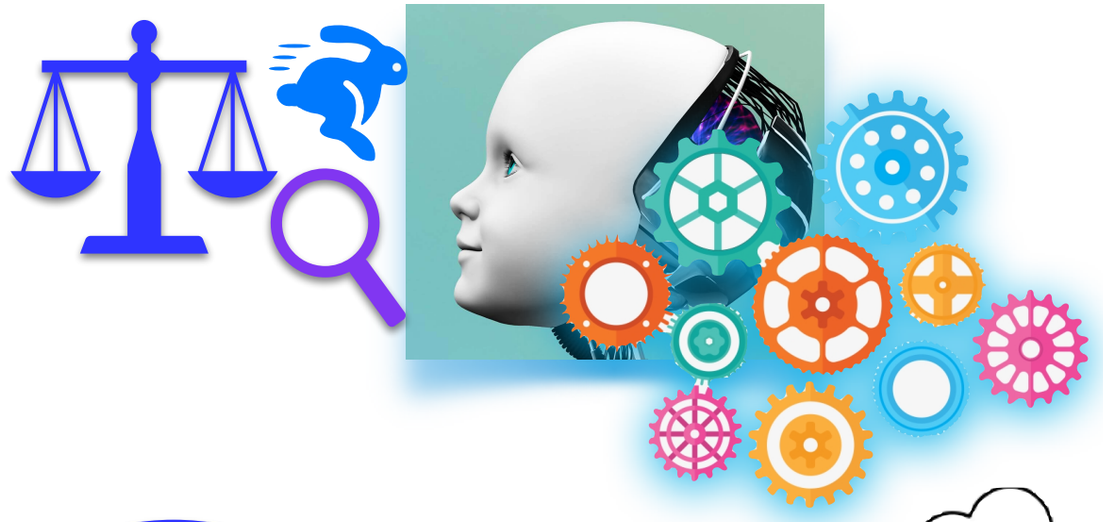


higher-SES

lower-SES

De Villiers et al. 2008

Dickson et al. 2022, in prep.

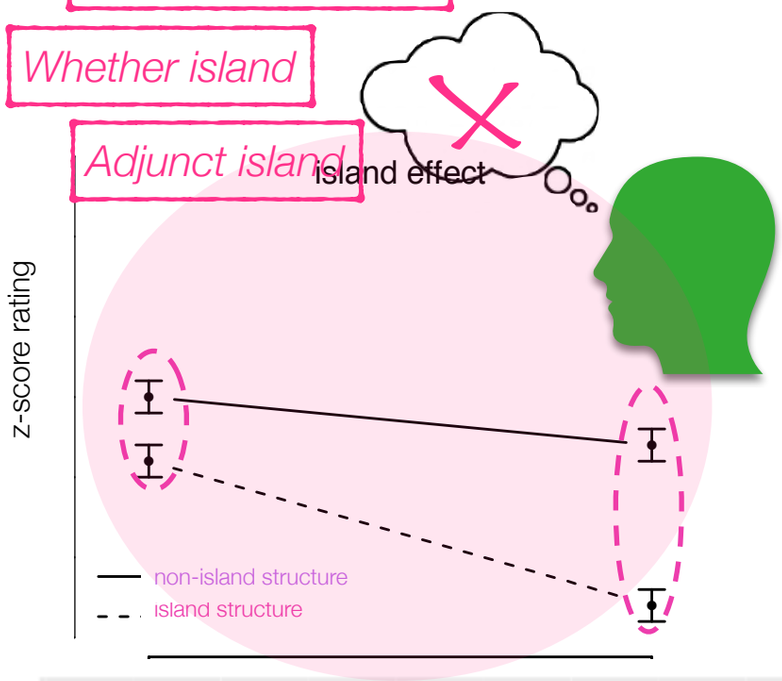


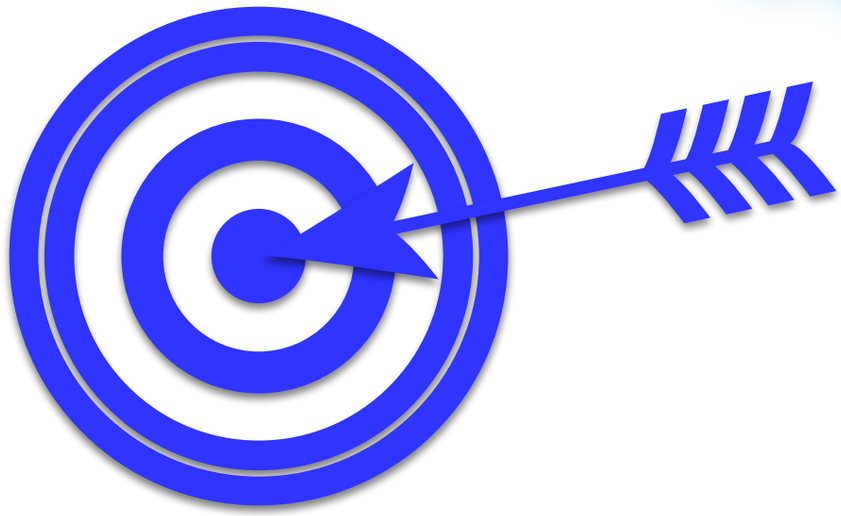
Subject island

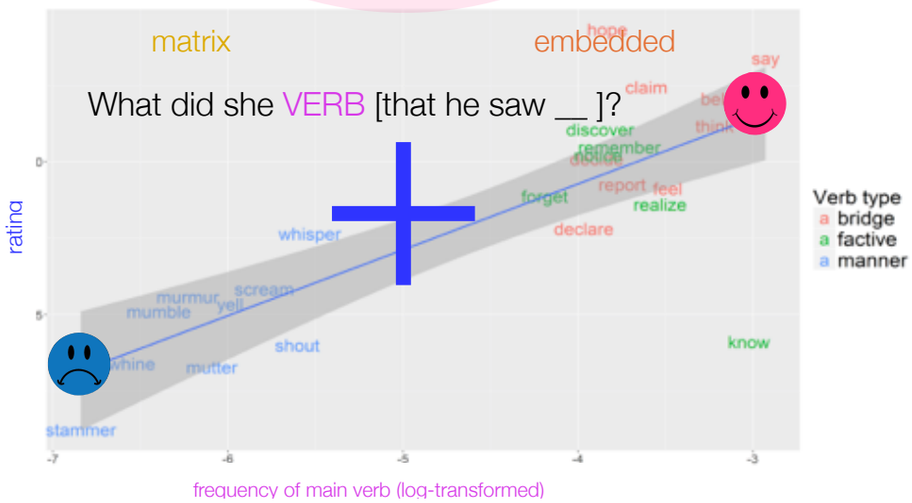
Complex NP island

Whether island

Adjunct island







Who *__who*?
How *__how*?

Who [CP-how *__who*]?
Who [CP-what *__who*]?
How [CP-what *__how*]?
How [CP-where *__how*]?
How [NP [CP-who *__how*]]?

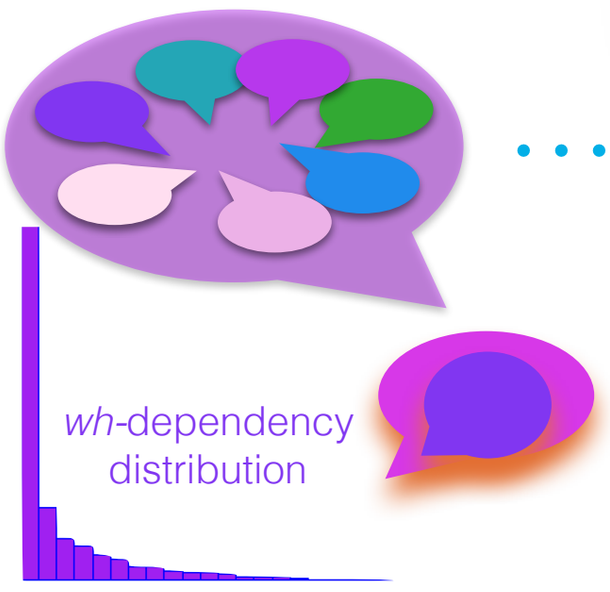
higher-SES

lower-SES

De Villiers et al. 2008

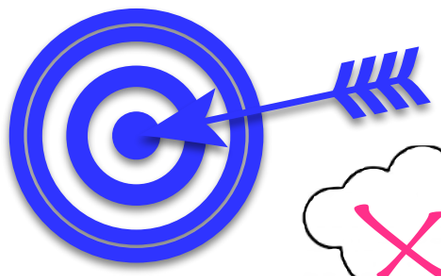
Dickson et al. 2022, in prep.

So what does the modeled child do?

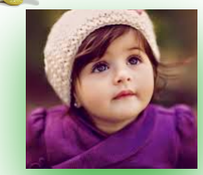


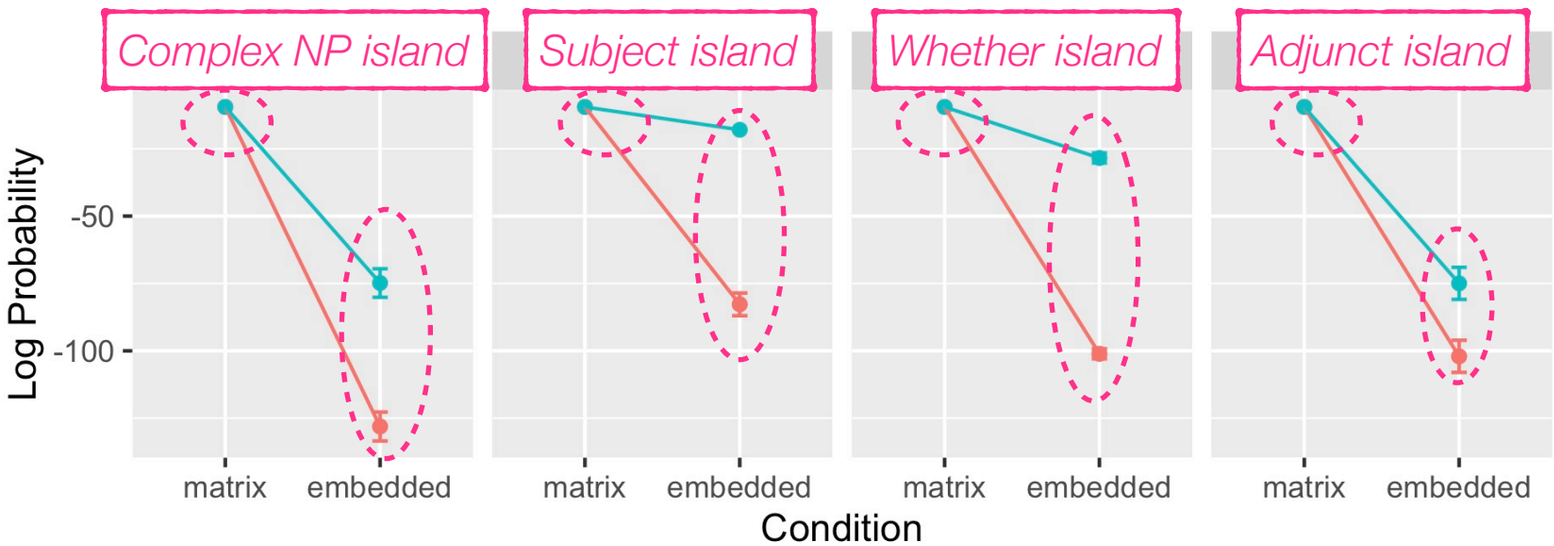
Dickson et al. 2022, in prep.

Learning the building blocks



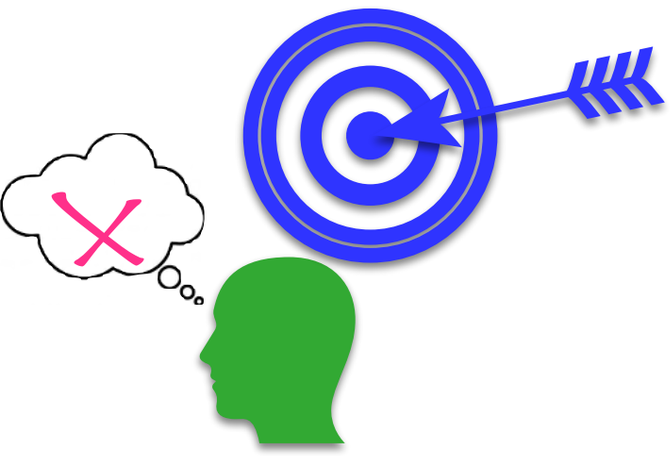
Superadditivity predicted for judgments of all four island types.

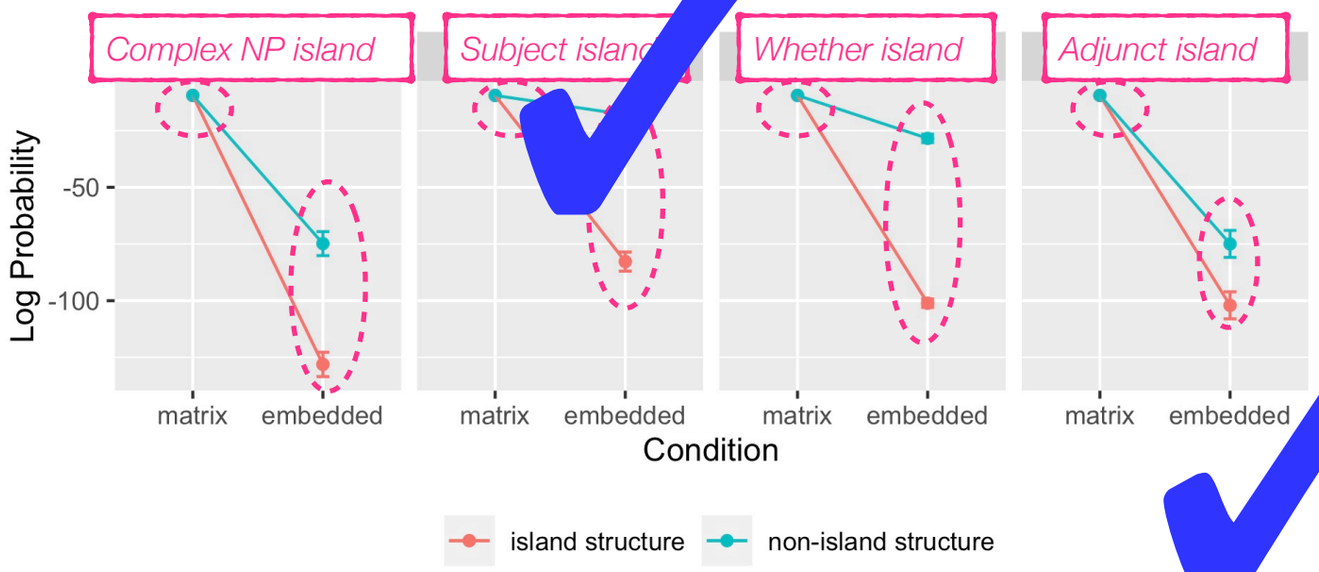




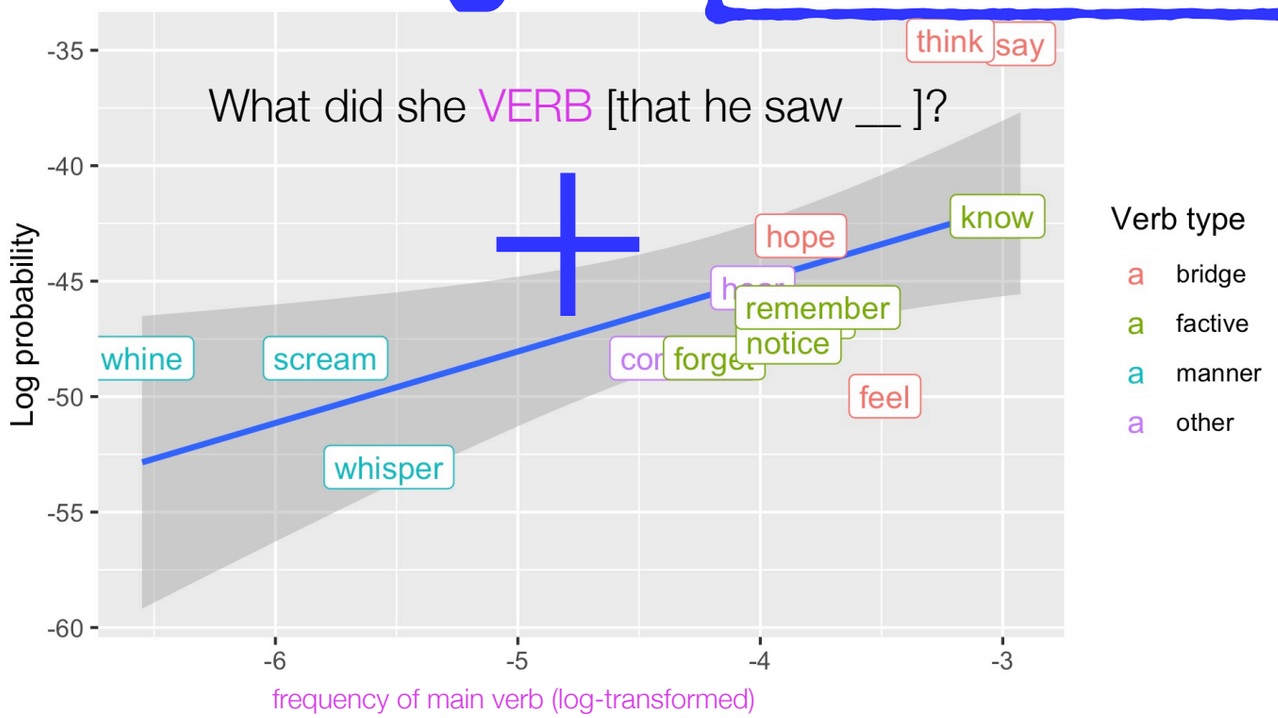
Dickson et al. 2022, in prep.

Learning the building blocks



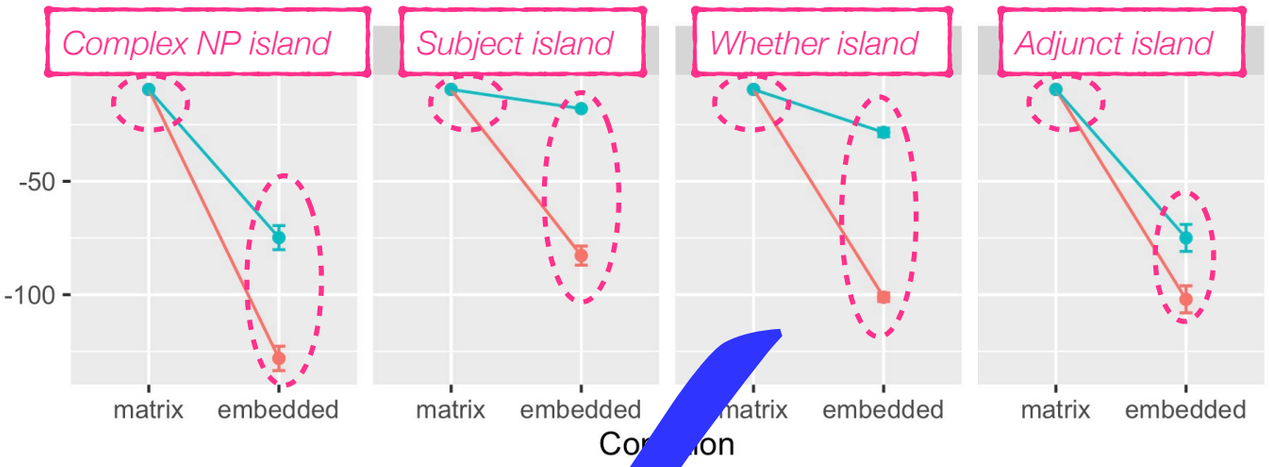


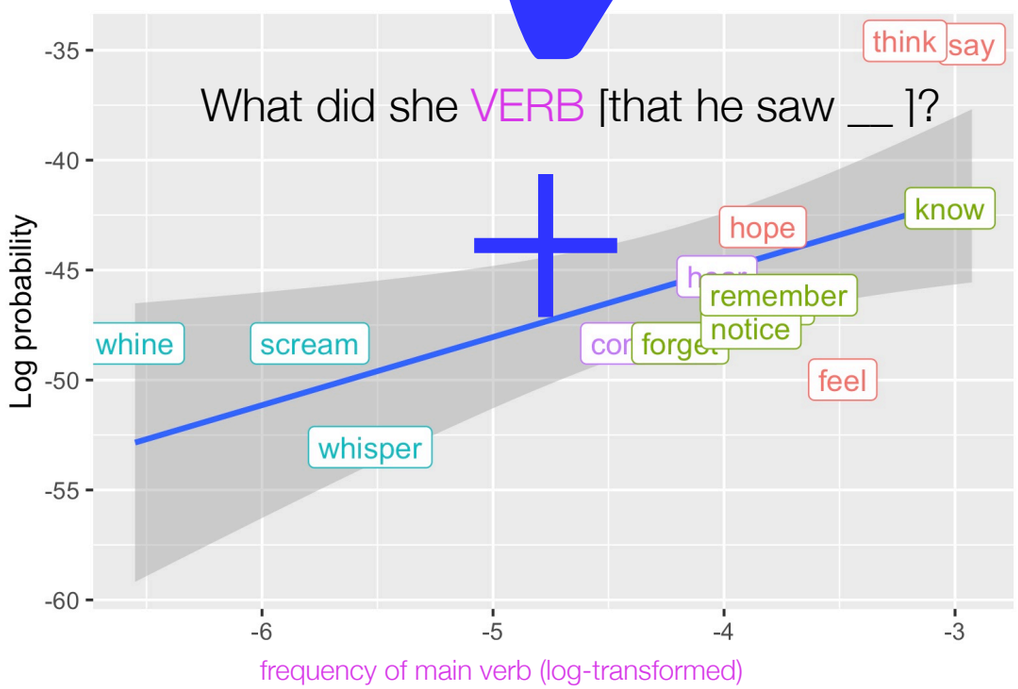
Positive correlation predicted with verb frequency for judgments of this *wh*-dependency.

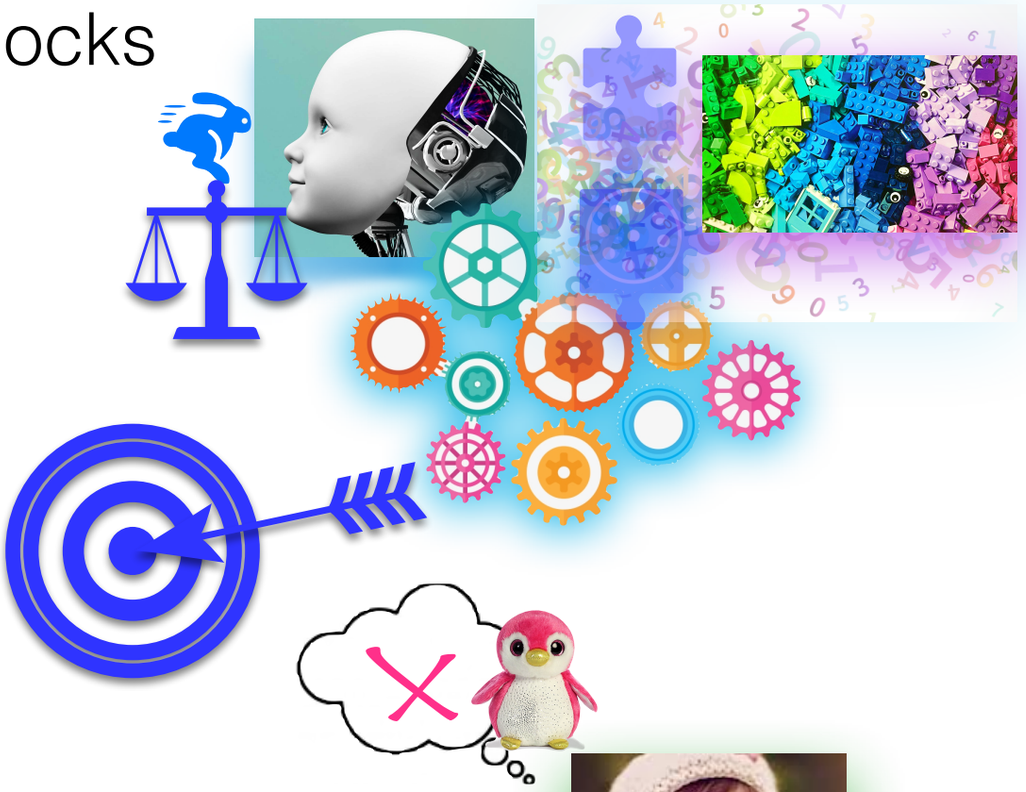


Dickson et al. 2022, in prep.

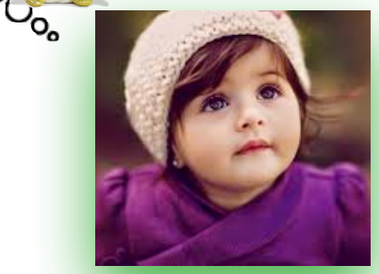
Learning the building blocks

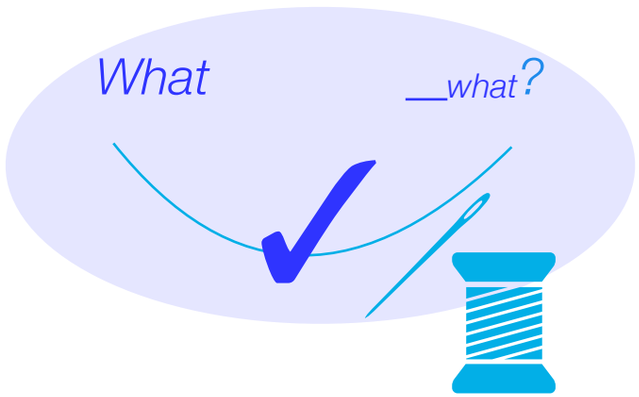


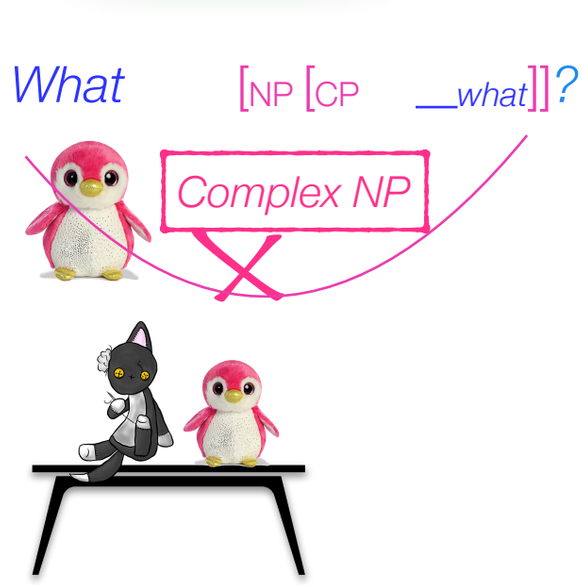




Children prefer this interpretation.

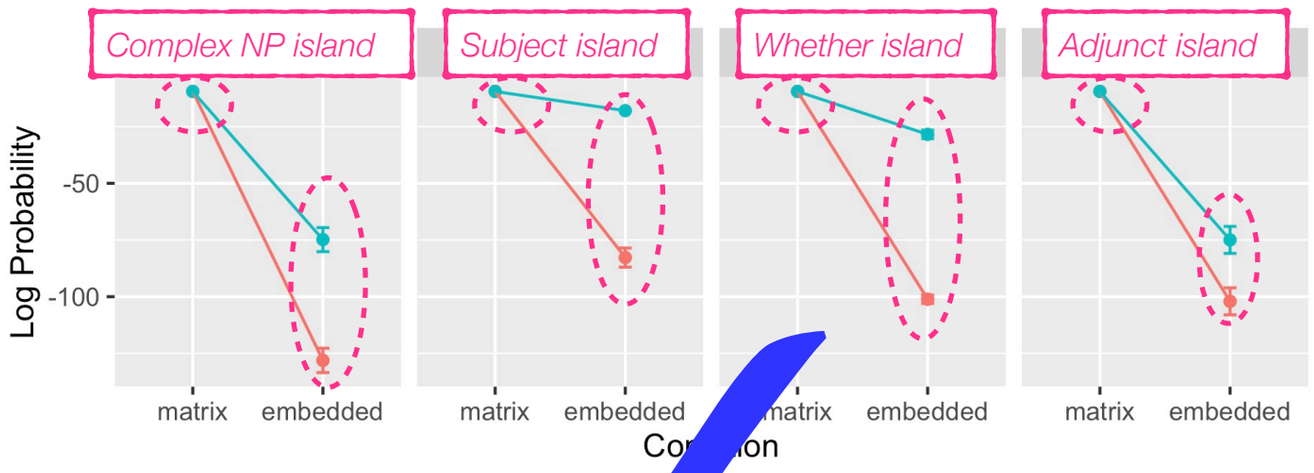


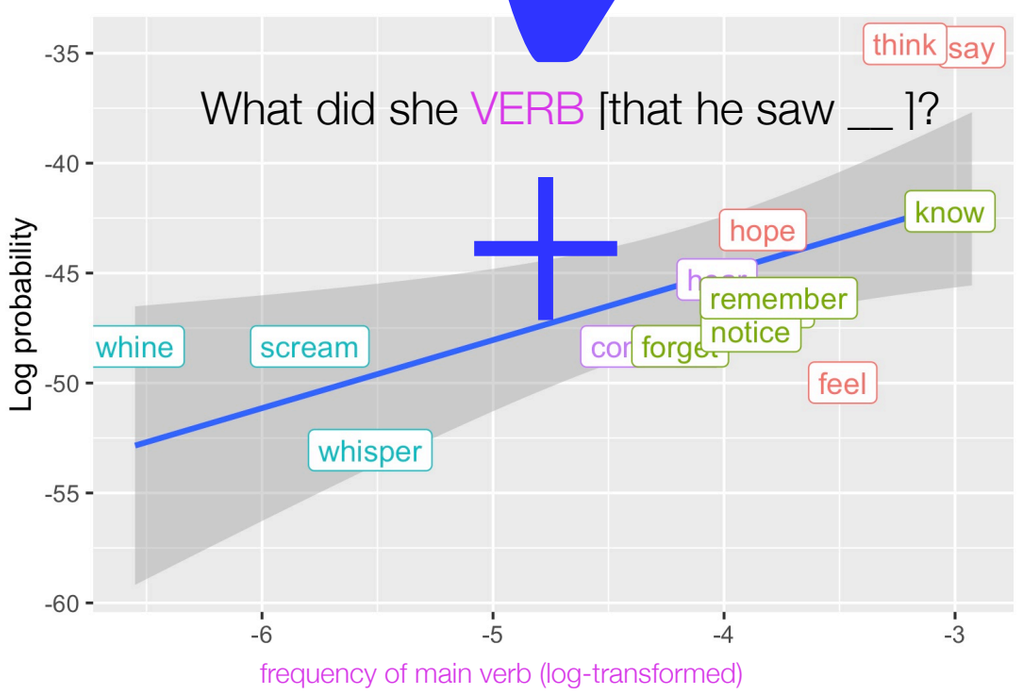




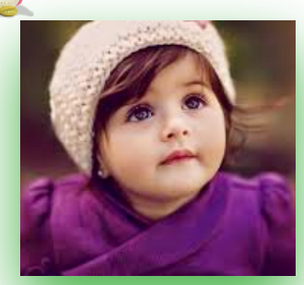
Dickson et al. 2022, in prep.

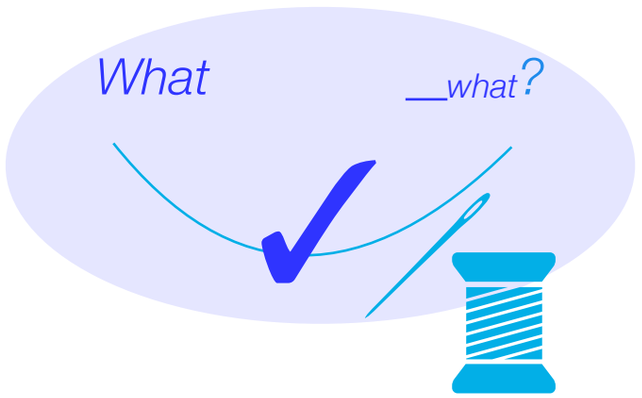
Learning the building blocks

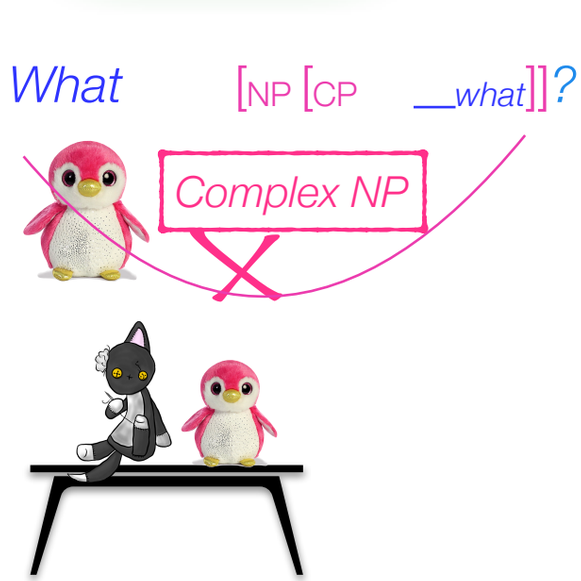




The *wh*-dependency this interpretation relies on is 10^5 times more probable than the other one.

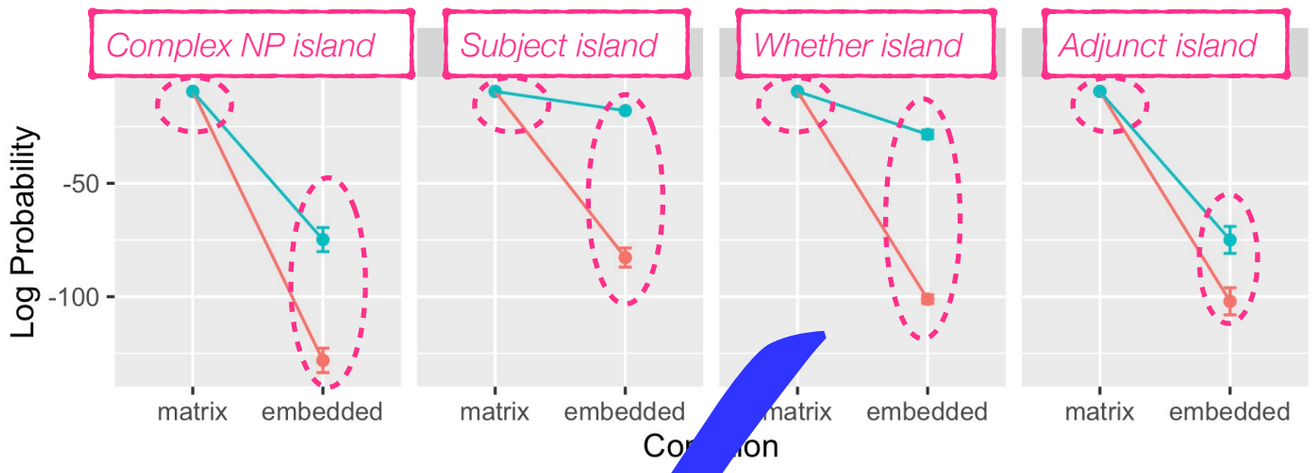


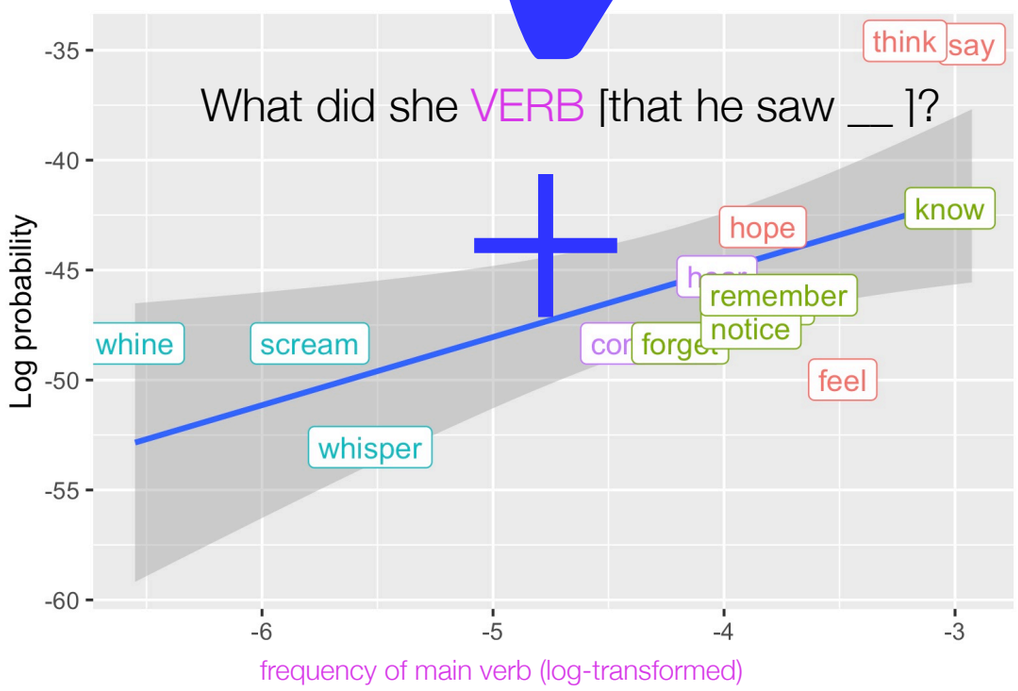


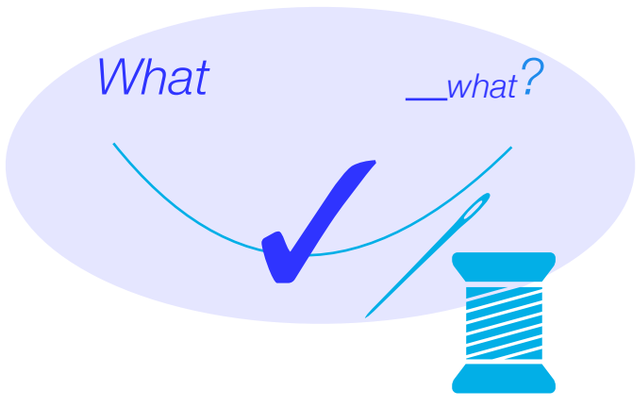


Dickson et al. 2022, in prep.

Learning the building blocks



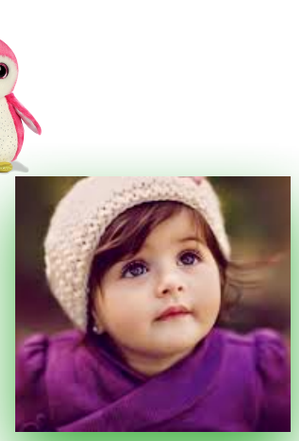


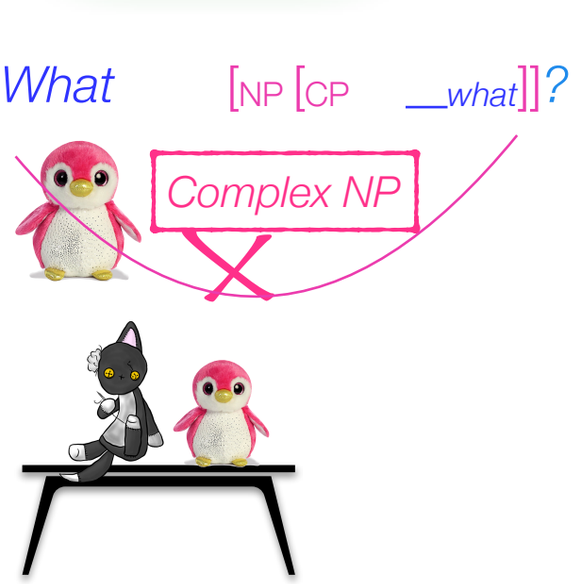


Verb type

- a bridge
- a factive
- a manner
- a other

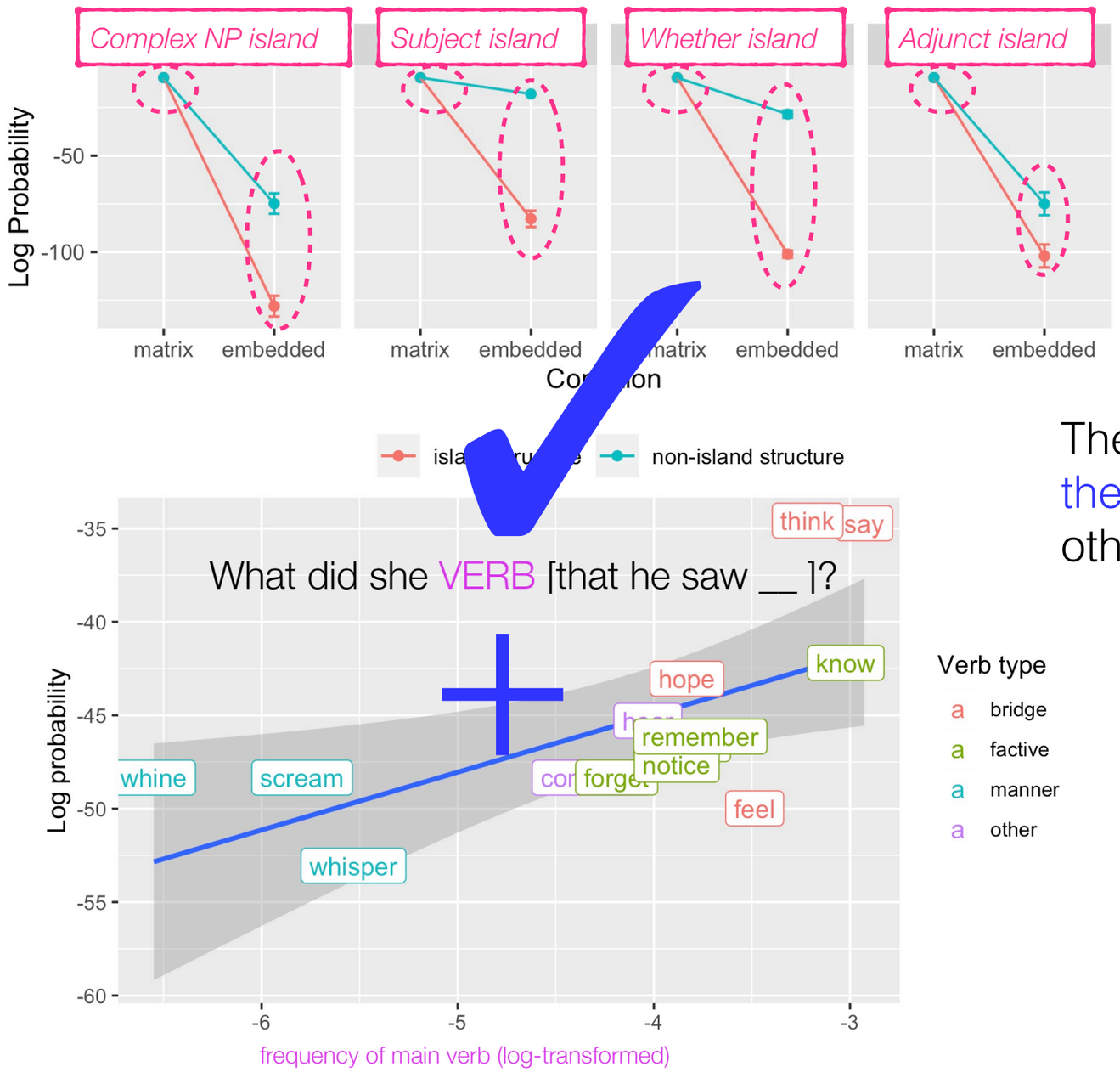
So, the modeled child prefers it.

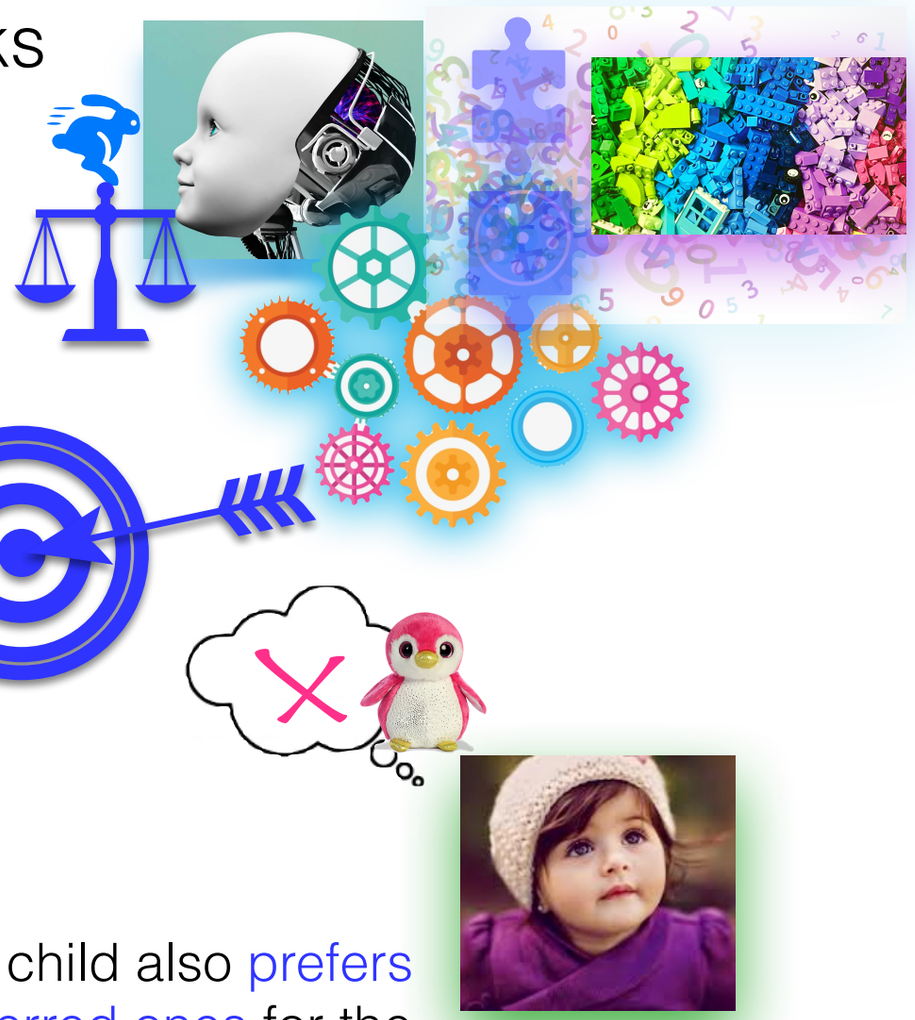




Dickson et al. 2022, in prep.

Learning the building blocks





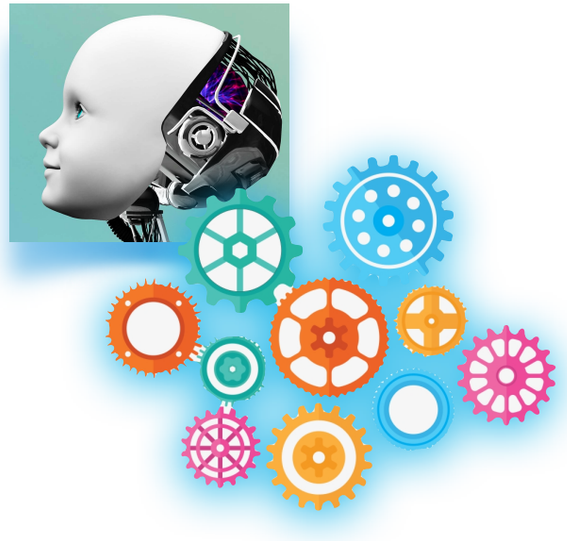
The modeled child also prefers the child-preferred ones for the other *wh*-dependencies.

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

Who ___*who*?
How ___*how*?

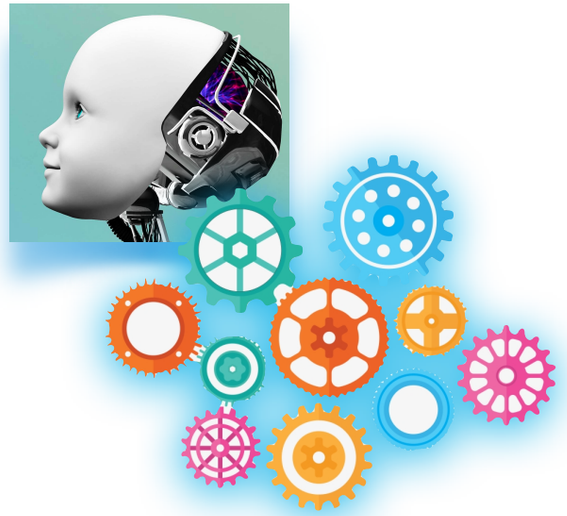
Who [CP-how ___*who*]?
Who [CP-what ___*who*]?
How [CP-what ___*how*]?
How [CP-where ___*how*]?
How [NP [CP-who ___*how*]]?

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.



learned building blocks

Different sizes that the modeled child learned

V_{say}

V_{go}

$VP_{V:know}$

VP_{V}

$IP_{I:present}$-$VP_{V:be}$

$IP_{I:past}$-VP

$CP_{C:null}$-IP

NP_{N}-$PP_{P:of}$

$IP_{I:-finite}$-VP_{V}

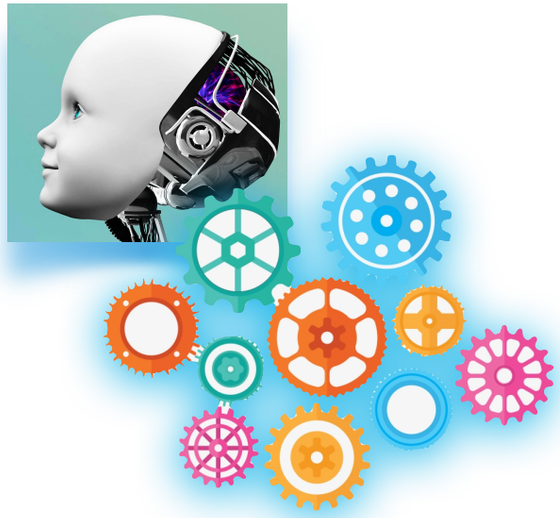
$VP_{V:go}$-VP

$IP_{I:present}$-VP_{V}-$IP_{I:-finite}$-$VP_{V:talk}$-$PP_{P:to}$

$VP_{V:think}$-$CP_{C:null}$-IP

$IP_{I:past}$-$VP_{V:say}$-CP

Dickson et al. 2022, in prep.



learned building blocks

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.).

V_{say}

V_{go}

$VP_{V:know}$

VP_{V}

$IP_{I:present}\text{-}VP_{V:be}$

$IP_{I:past}\text{-}VP$

$CP_{C:null}\text{-}IP$

$NP_{N}\text{-}PP_{P:of}$

$IP_{I:-finite}\text{-}VP_{V}$

$VP_{V:go}\text{-}VP$

$IP_{I:present}\text{-}VP_{V}\text{-}IP_{I:-finite}\text{-}VP_{V:talk}\text{-}PP_{P:to}$

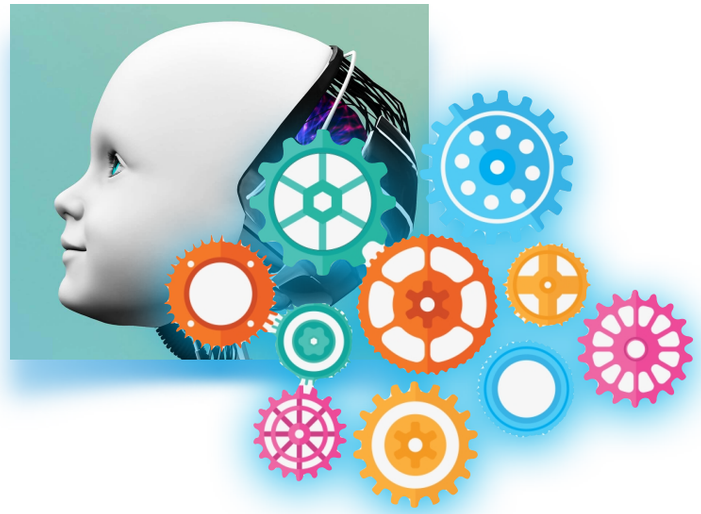
$VP_{V:think}\text{-}CP_{C:null}\text{-}IP$

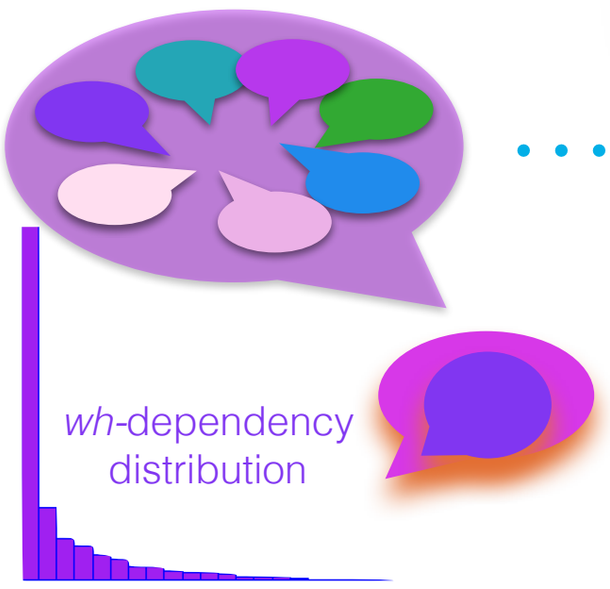
$IP_{I:past}\text{-}VP_{V:say}\text{-}CP$

Dickson et al. 2022, in prep.

Takeaway:

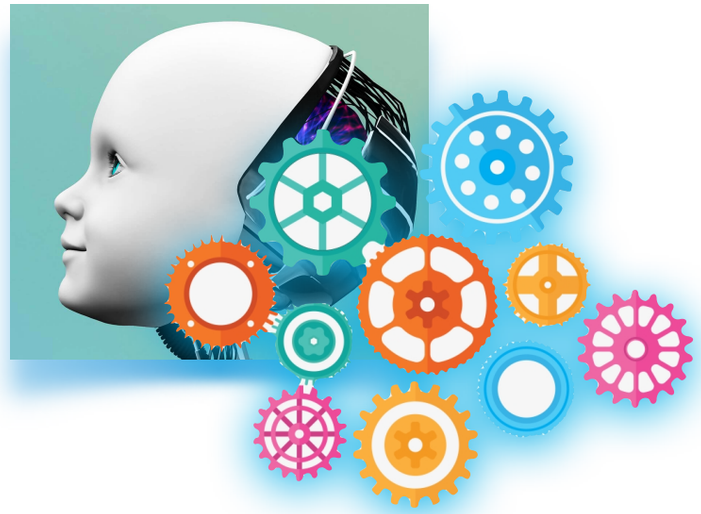
This theory — which has to learn the building blocks — can work for learning knowledge about syntactic islands.

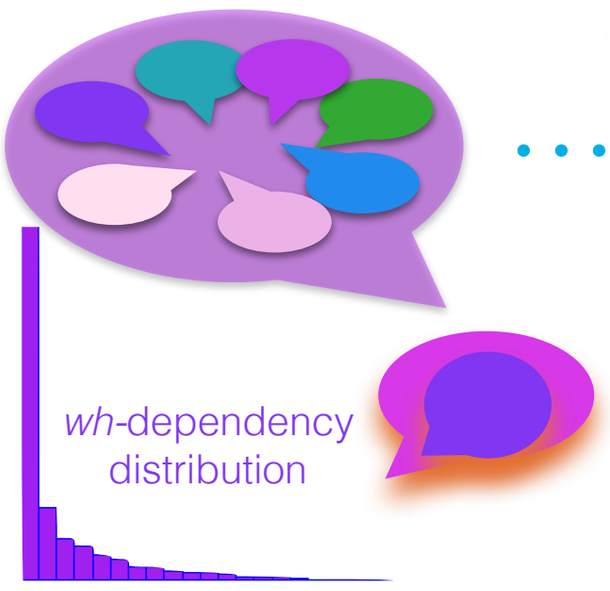




Dickson et al. 2022, in prep.

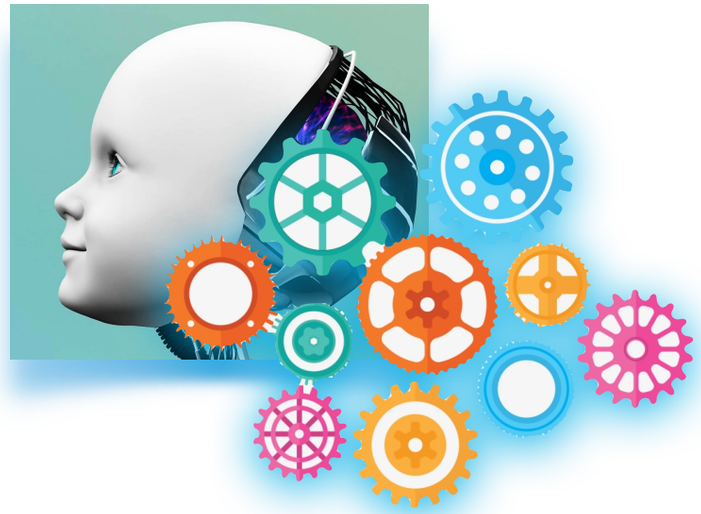
Key idea (again): Learning about the building blocks of *wh*-dependencies leads to knowledge about syntactic islands, even when there's less knowledge built in.

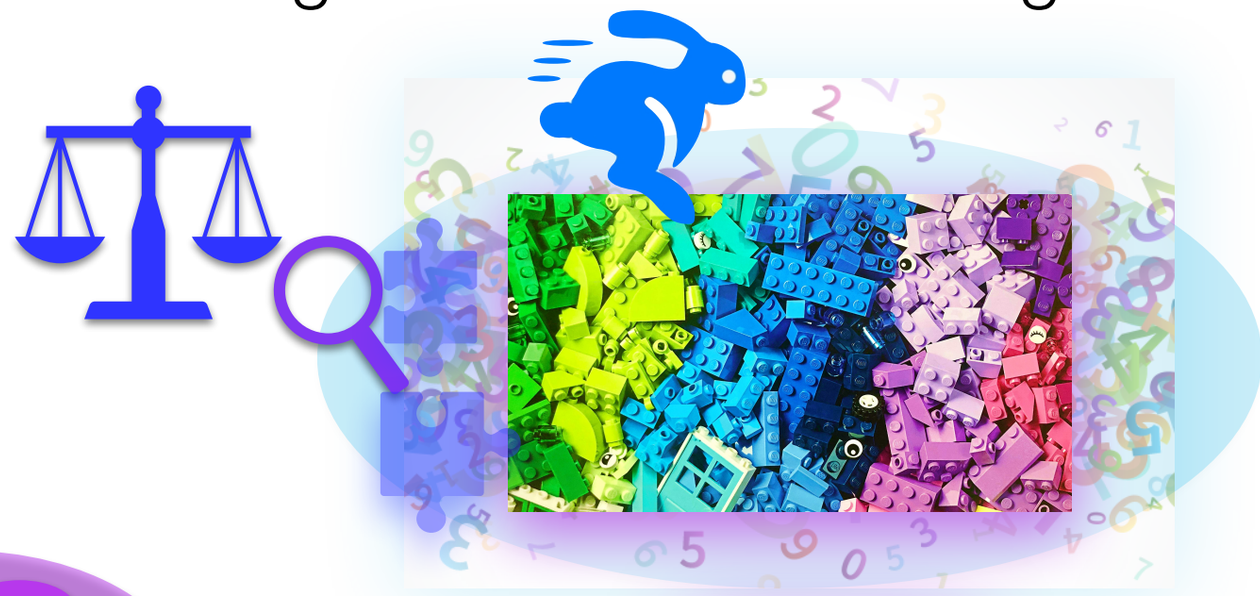


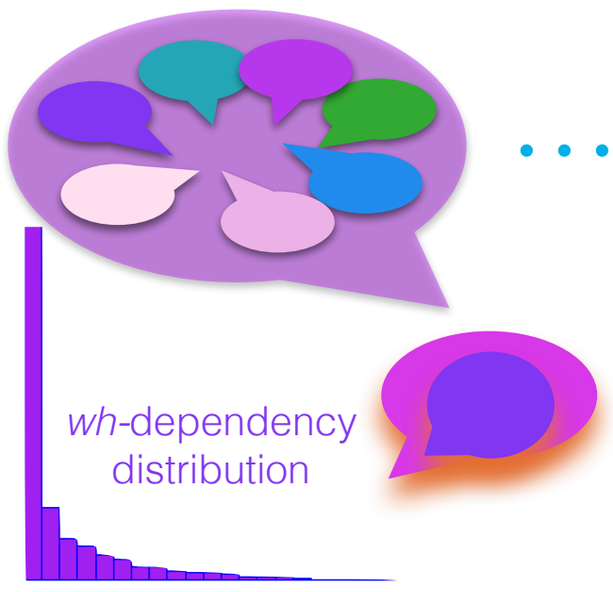


Dickson et al. 2022, in prep.

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.

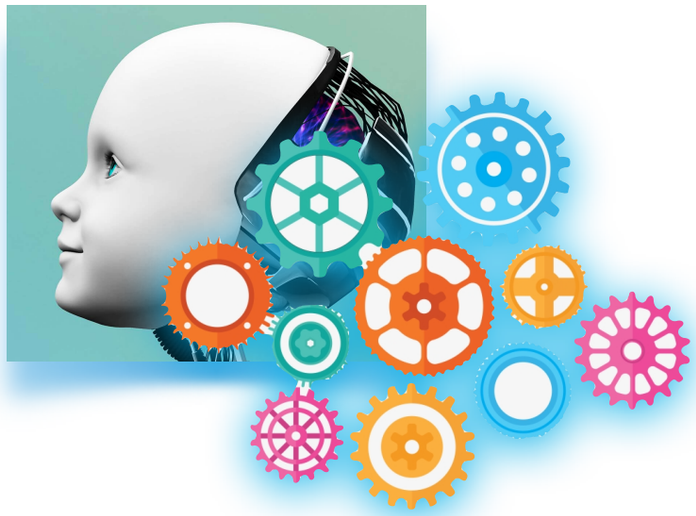




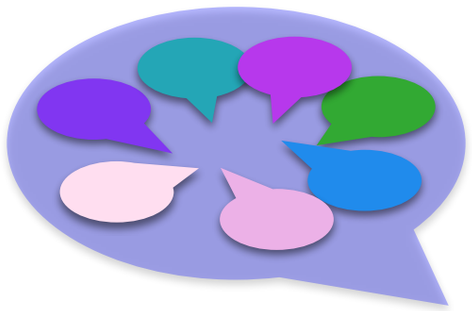


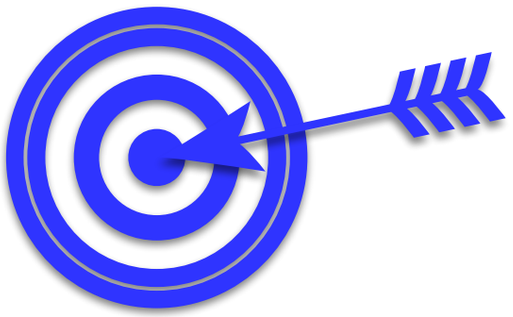
Dickson et al. 2022, in prep.

The big picture

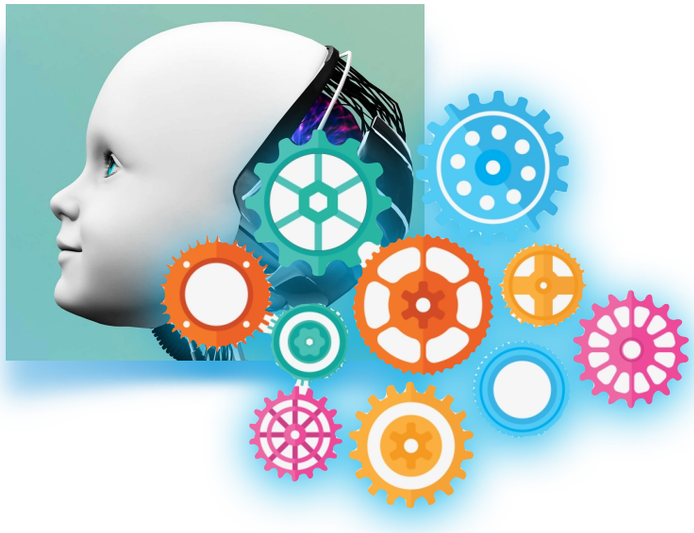


This case study demonstrates how we can use computational cognitive modeling, grounded in empirical data, to evaluate our theories about how children could learn what they do about language.





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

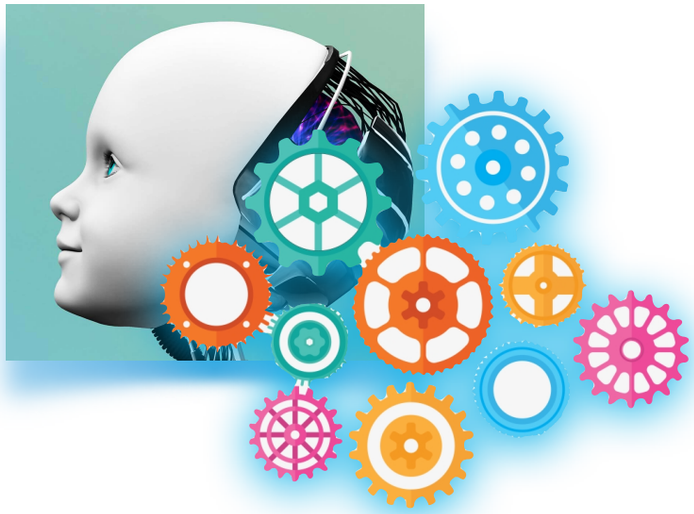


The big picture

Takeaway: Theories that view specific linguistic knowledge (like syntactic islands and *wh*-dependencies more generally) as built up of smaller building blocks can work well.

Who does

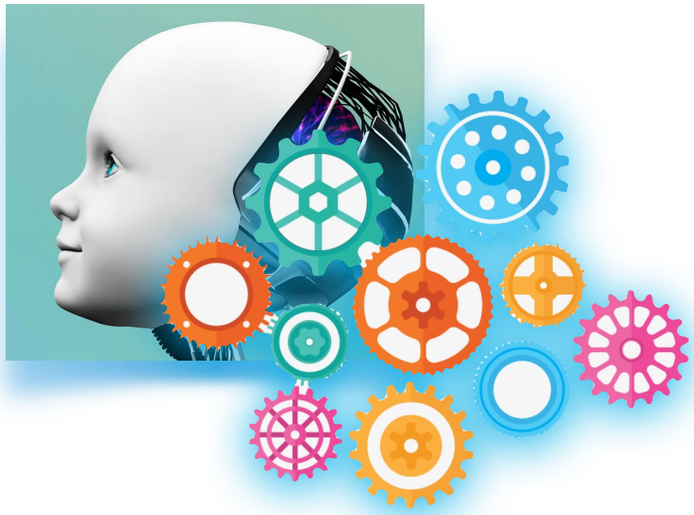
Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.



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

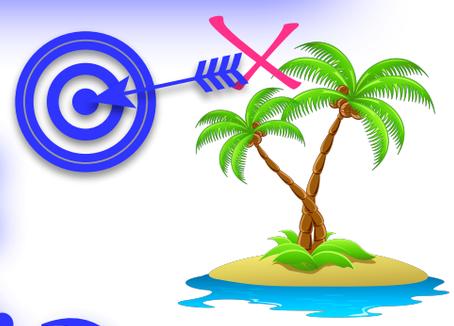
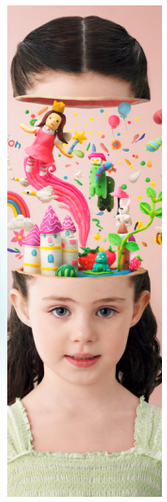
Niels Dickson

BUCLD 2018 UCSD Linguistics 2020 ForMA Group 2020

UMD Linguistics 2020 BUCLD 2021 SCiL 2022

UCI QuantLang Collective

Computation of Language Laboratory

UC Irvine

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