How to succeed at syntactic island acquisition without really trying: Learning about local structure

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

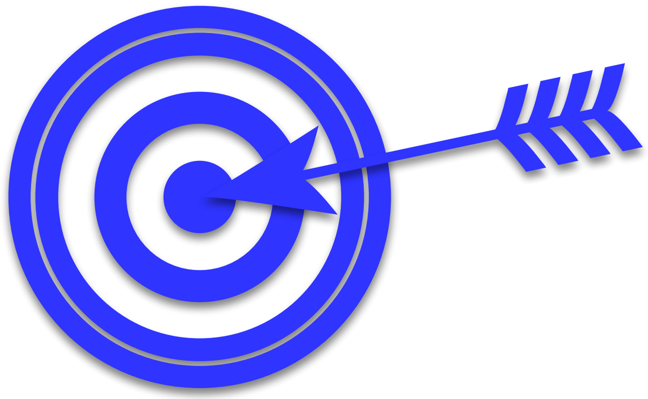
lpearl@uci.edu

March 31, 2023

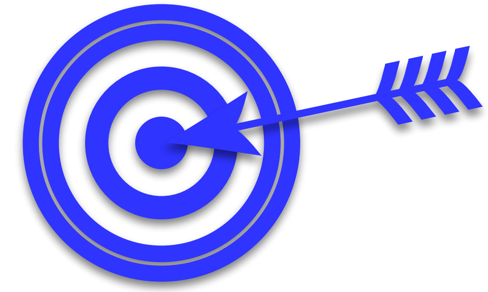
Workshop on Locality in Theory, Processing, and Acquisition

University of Pennsylvania

What does it mean to succeed
at syntactic island acquisition?

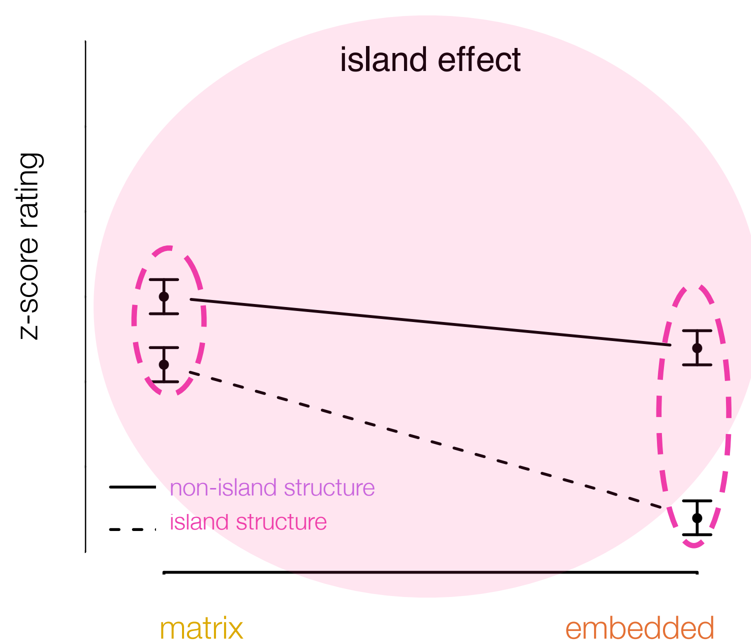


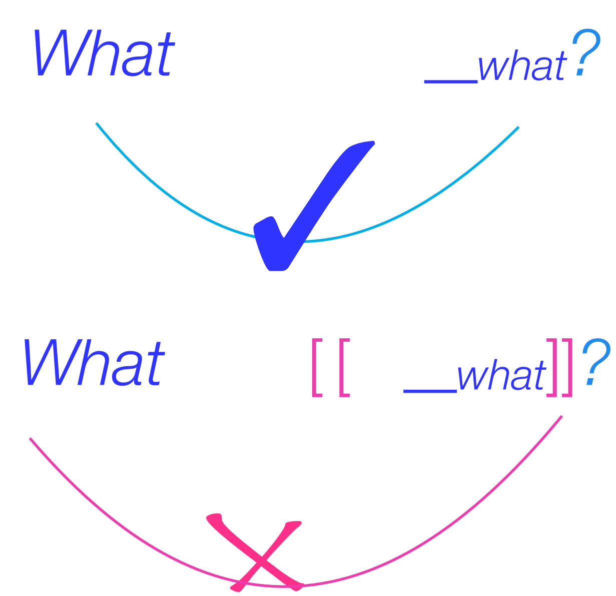
One answer: To develop the target behavior
we observe about syntactic islands…



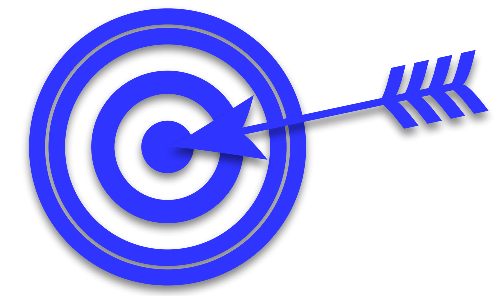
One answer: To develop the target behavior we observe about syntactic islands…

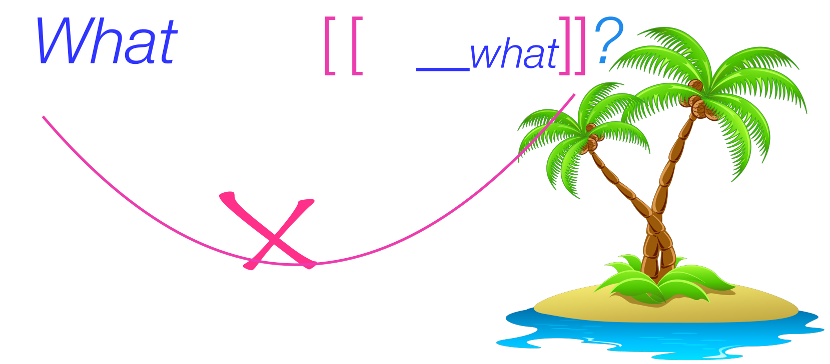
Some example behavior:
judgment patterns and (dis)preferences for certain utterances related to syntactic islands



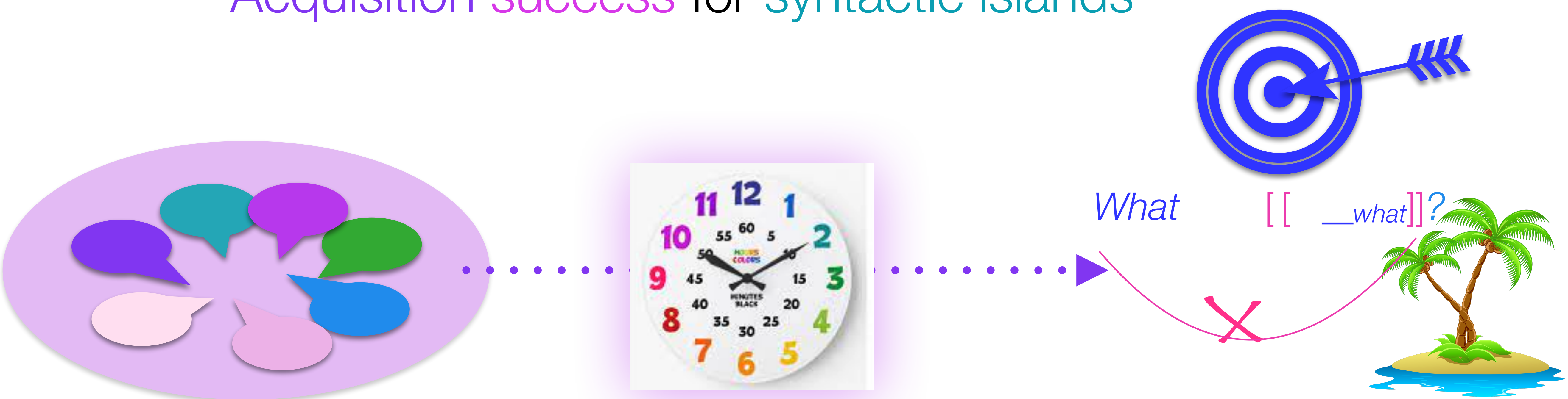


One answer: To develop the target behavior
we observe about syntactic islands,
given the input children get
and the time they have to learn.





Acquisition success for syntactic islands



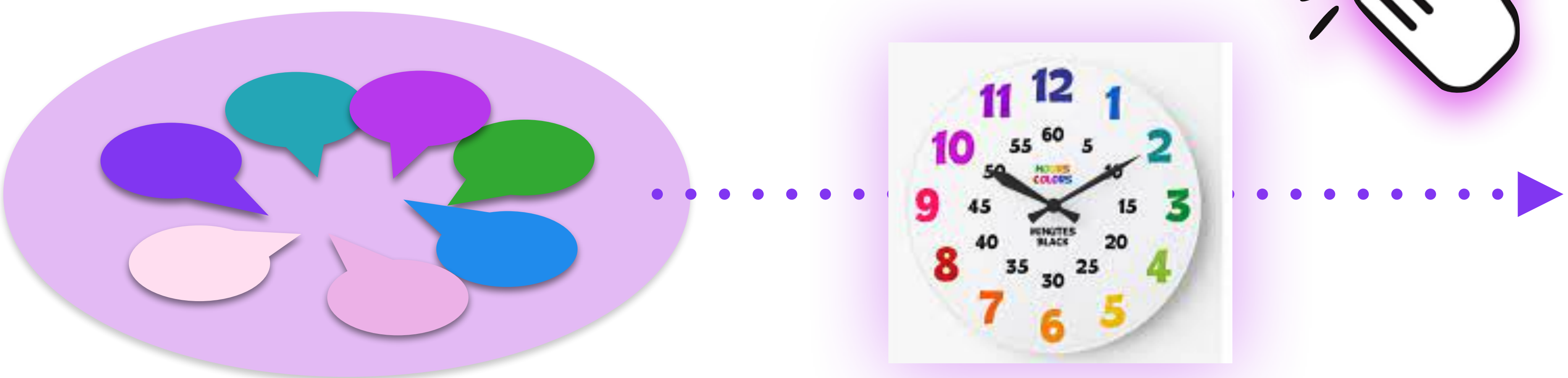
“…without really trying…”

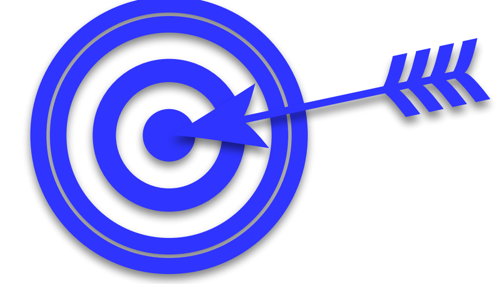
What does it mean to try?

“trying”

One answer: Learn about syntactic islands directly. For instance, look for language-specific “bounding nodes” (Subjacency: Chomsky 1973, Huang 1982, Lasnik & Saito 1984) that signal syntactic island structure in *wh*-dependencies.

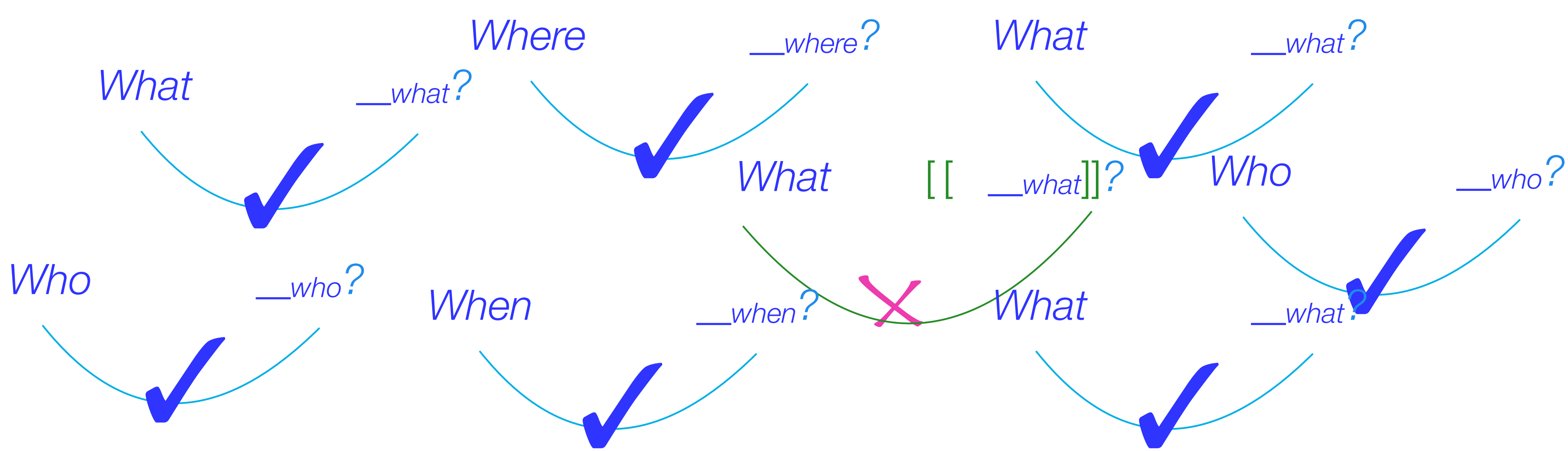
What [BN1 [BN2 *__what*]]?

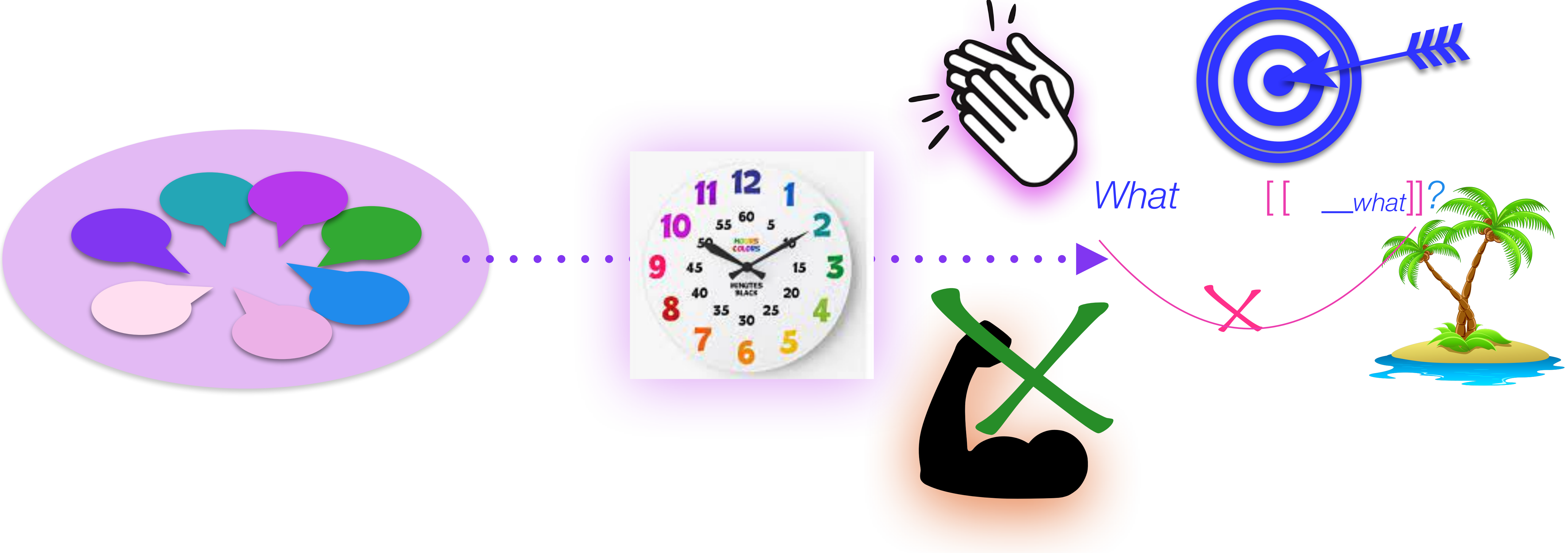




“without really trying”

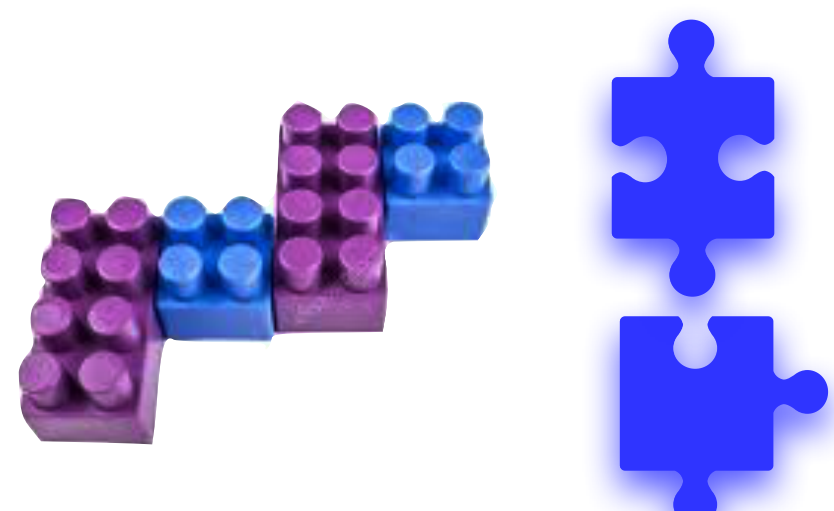
Learn about syntactic islands indirectly by learning about *wh*-dependencies more generally.

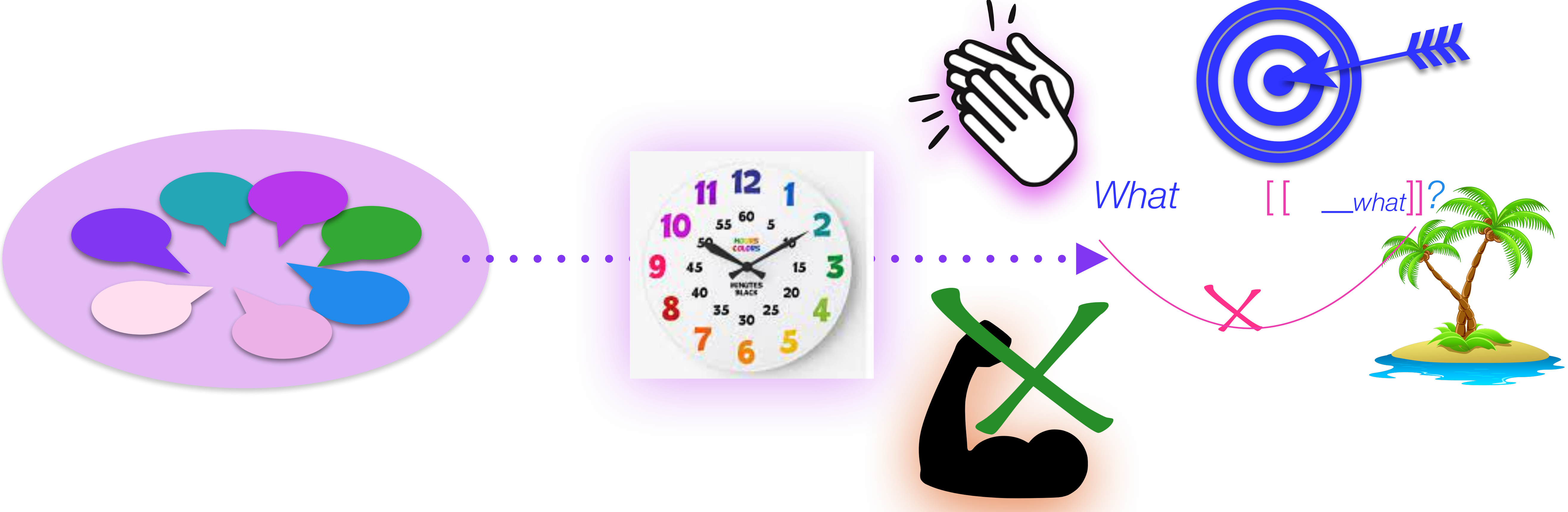




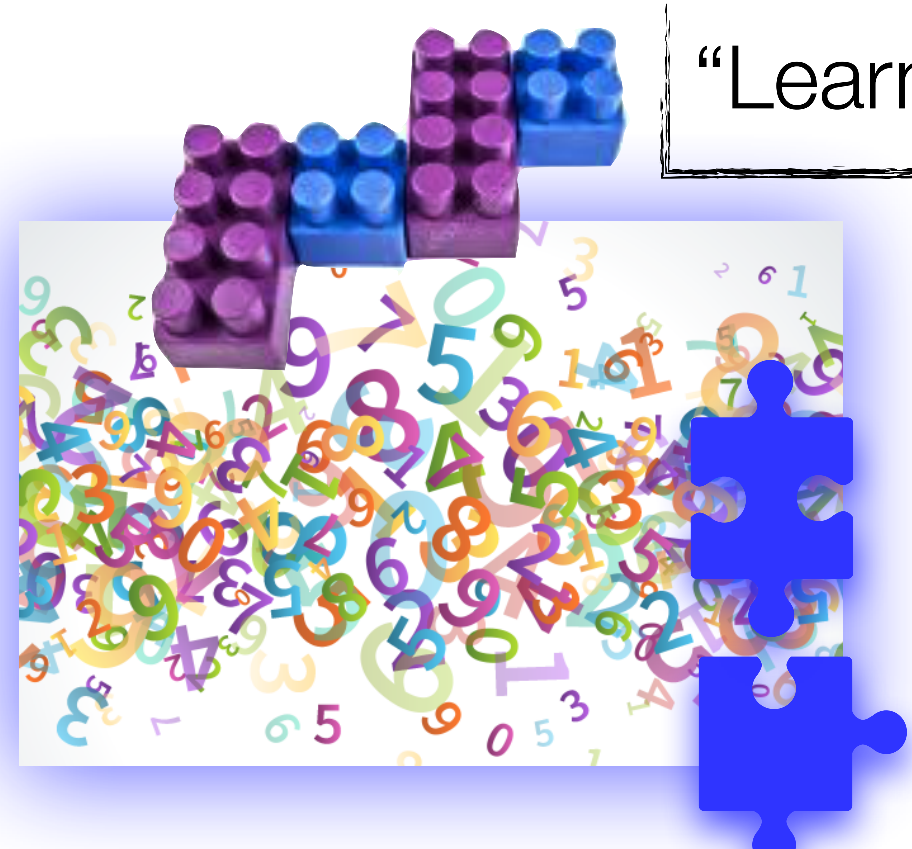
“Learning about local structure”

Proposal: The relevant local structure is pieces that combine to build *wh*-dependencies.





“Learning about local structure”



Proposal: Learn about syntactic islands indirectly by learning about the probabilities of the pieces that build *wh*-dependencies.

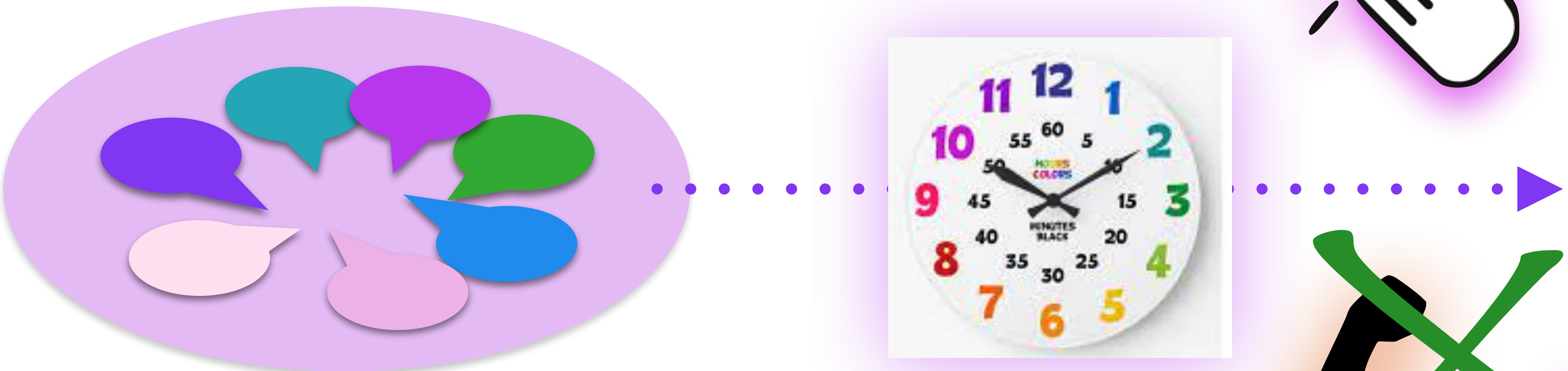
Part 1: Learning the probabilities of pre-specified pieces from the input (Pearl & Sprouse 2013).

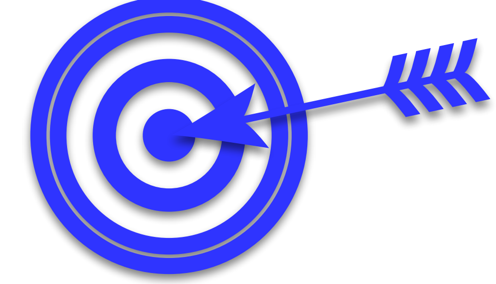
This turns out to work pretty well.

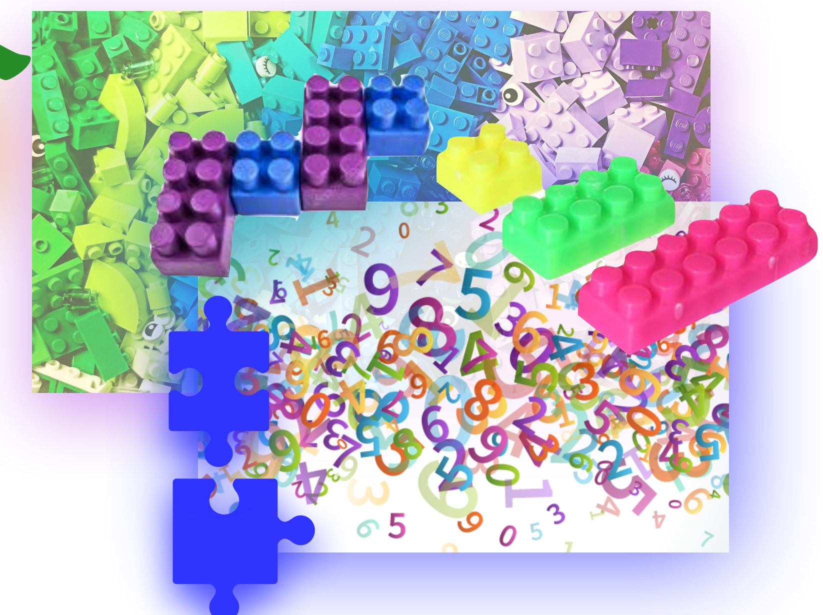
1: probabilities of pre-specified pieces

Part 2: Learning what the pieces are and their probabilities from the input (Dickson, Pearl & Futrell 2022, in prep).

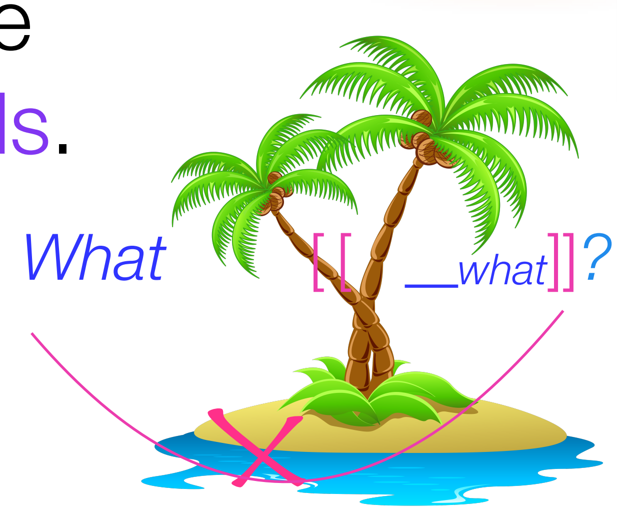
This turns out to work even better.





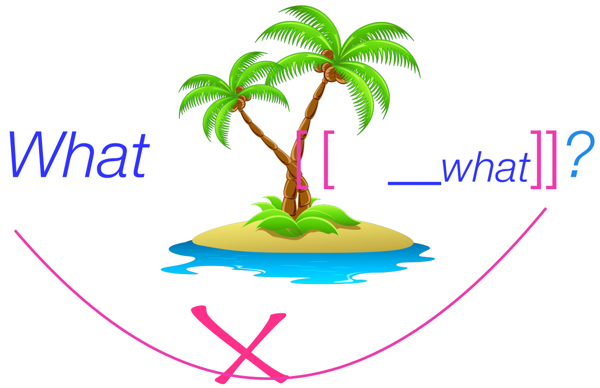


But first, let’s briefly review some relevant information about the acquisition of syntactic islands.



1: probabilities of pre-specified pieces

2: what the pieces are and their probabilities



Syntactic islands

involve *wh*-dependencies.

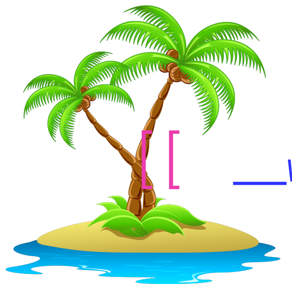
This kitty was bought as a present for someone.

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?



What [[__*what*]]?

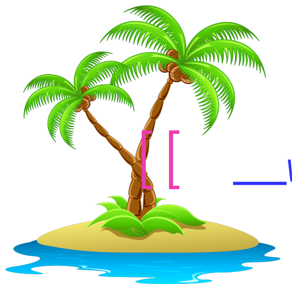
Syntactic islands

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?

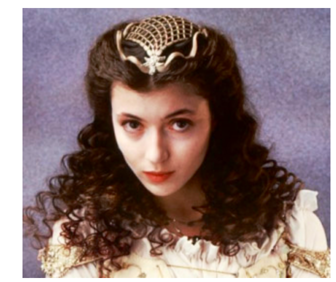
What  [[___*what*]]?

Syntactic islands

involve *wh*-dependencies.

What's going on here?

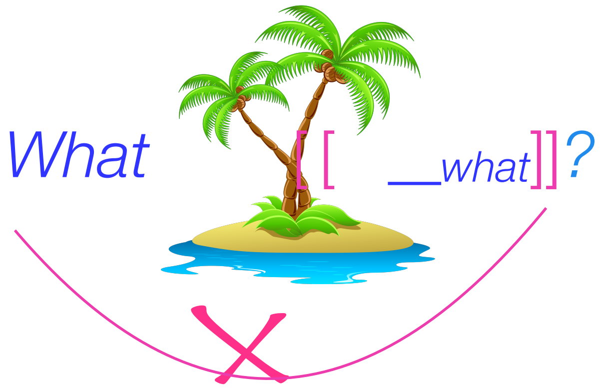
There's a dependency between the *wh*-word *who* and where it's understood (the gap)



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)



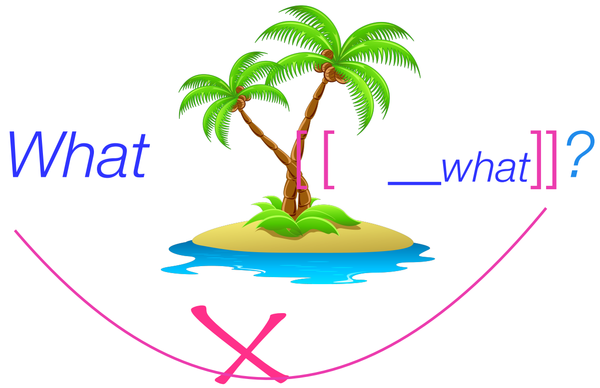
Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

Who does Lily think the kitty for __who is pretty?

Subject island



Syntactic islands

involve *wh*-dependencies.

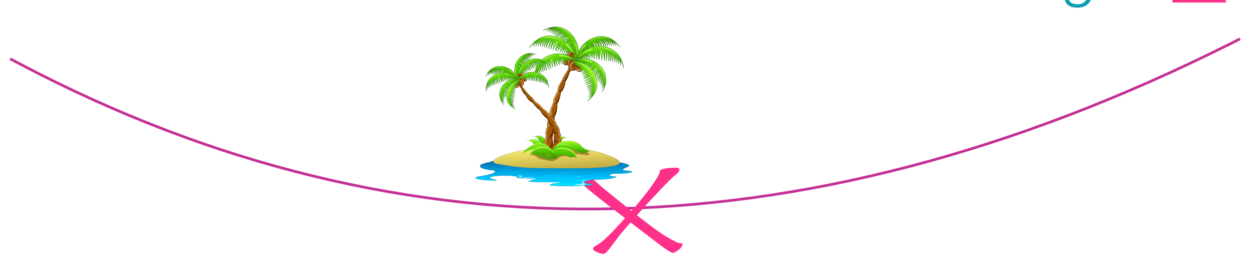
syntactic island (Ross 1967)

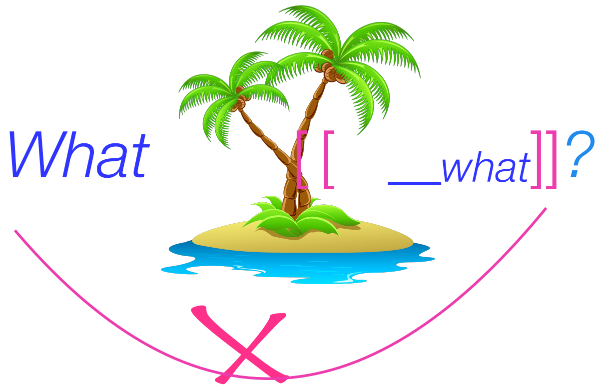
Who does Lily think the kitty for __who is pretty? $\boxed{\textit{Subject island}}$

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

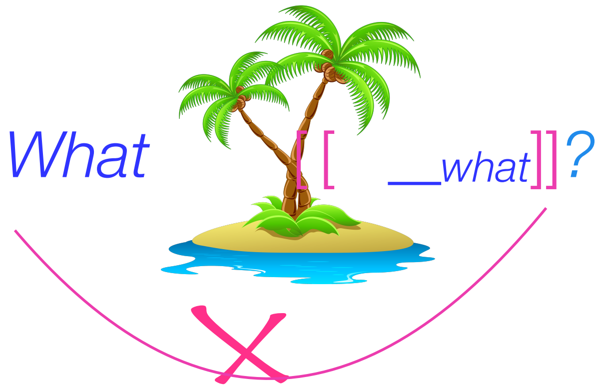
What did Jack make the claim that he bought __what? Complex NP island

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

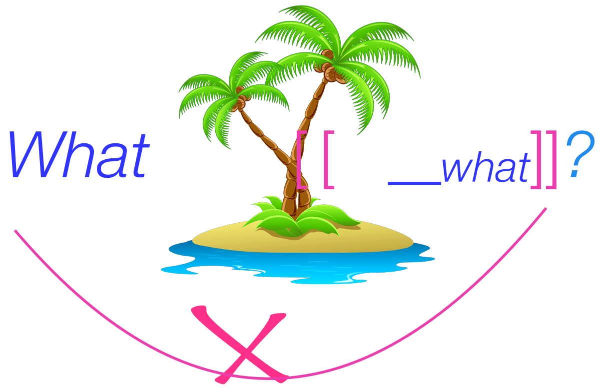
What did Elizabeth wonder whether Jack bought __what? Whether island

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

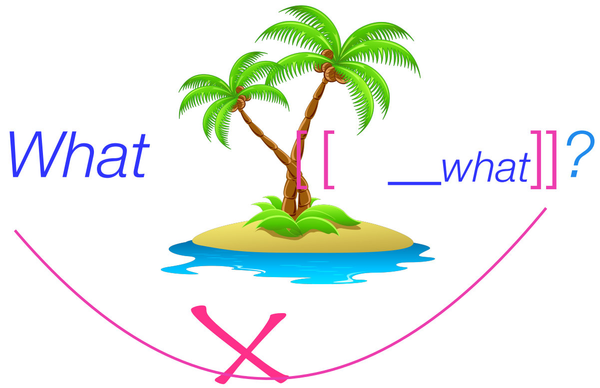
What did Jack make the claim that he bought __what? Complex NP island

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

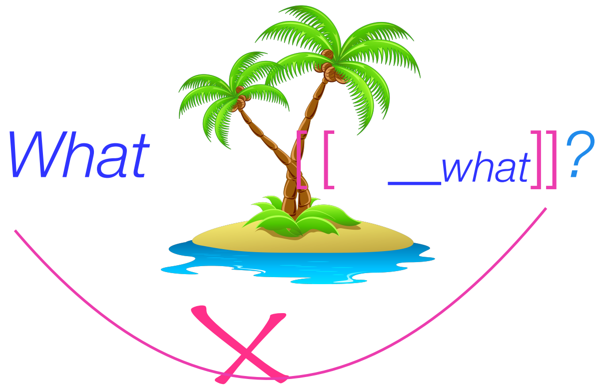
What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

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

Elizabeth

What did Elizabeth think __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

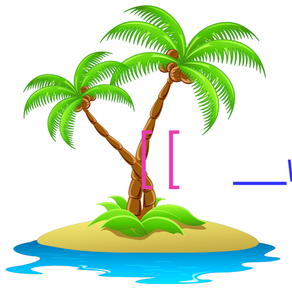
What did Elizabeth worry if Jack bought __what? Adjunct island

Elizabeth

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

Jack

What did Elizabeth think Jack said __what?

What  [[__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

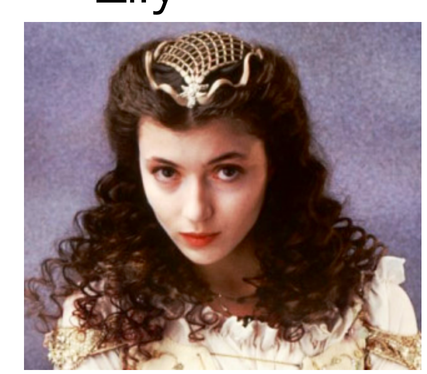
Elizabeth

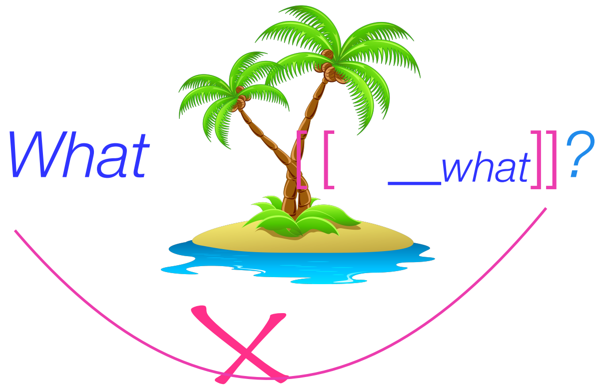
Jack

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)

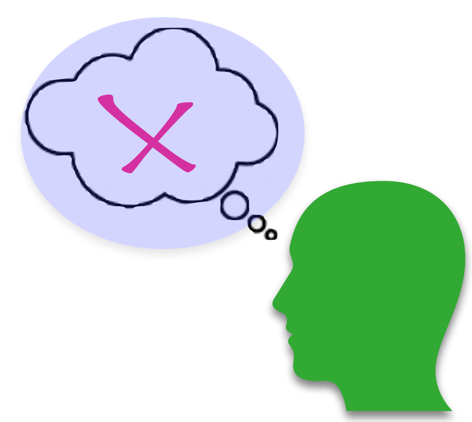
Who does Lily think the kitty for __who is pretty? Subject island

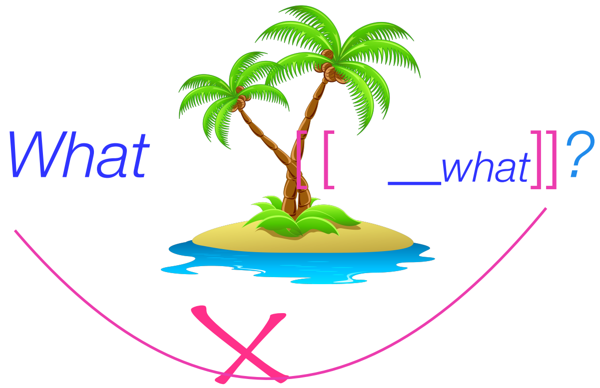
What did Jack make the claim that he bought __what? Complex NP island

What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

English adults judge these island-crossing 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.

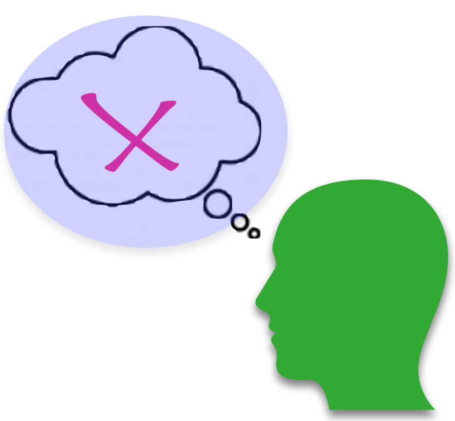
syntactic island (Ross 1967)

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

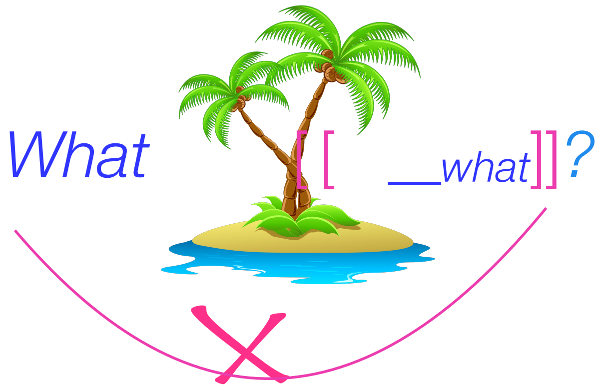
What did Jack make the claim that he bought __what? Complex NP island

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



Syntactic islands

involve *wh*-dependencies.

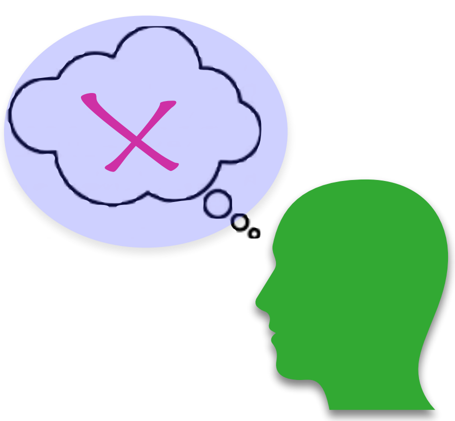
syntactic island (Ross 1967)

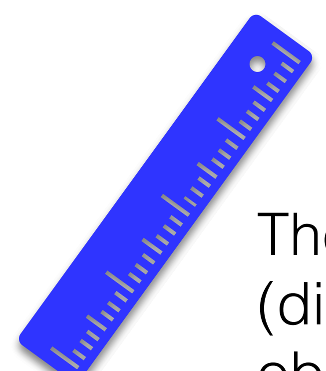
Who does Lily think the kitty for __who is pretty? Subject island

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

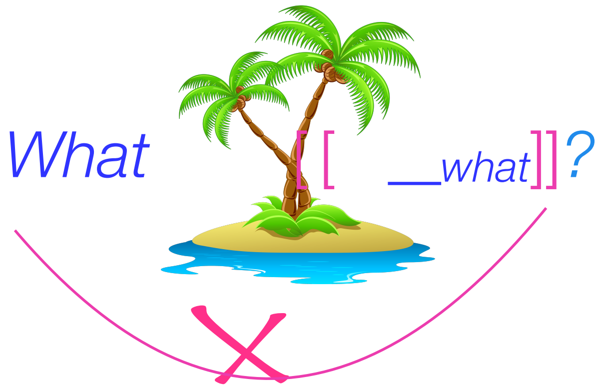
What did Elizabeth wonder whether Jack bought __what? Whether island

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.

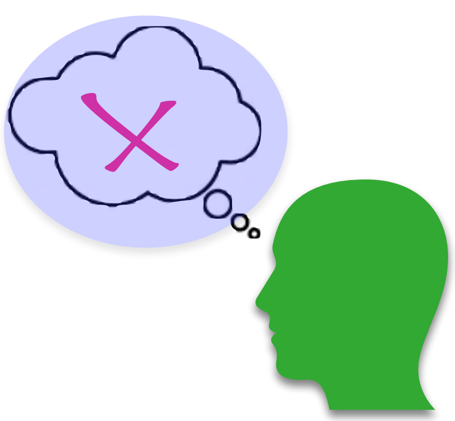
syntactic island (Ross 1967)

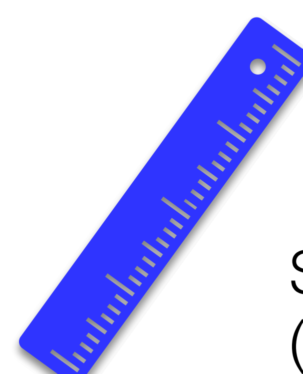
Who does Lily think the kitty for __who is pretty? Subject island

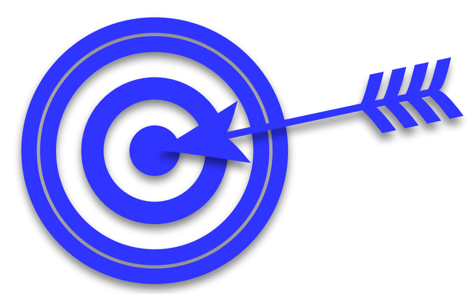
What did Jack make the claim that he bought __what? Complex NP island

What did Elizabeth wonder whether Jack bought __what? Whether island

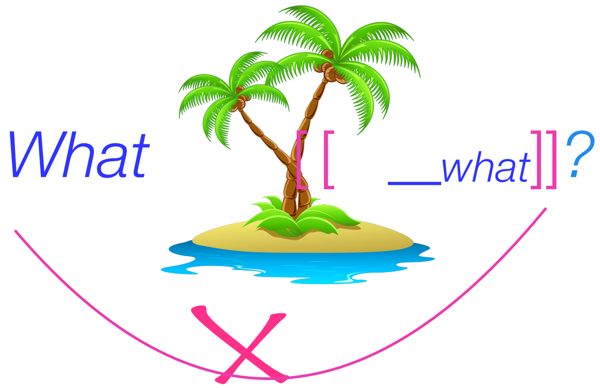
What did Elizabeth worry if Jack bought __what? Adjunct island







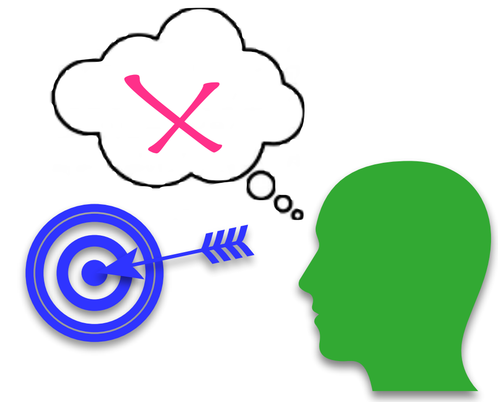
So, these judgments and (dis)preferences can serve as a target for successful acquisition — an outcome we can measure.



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)

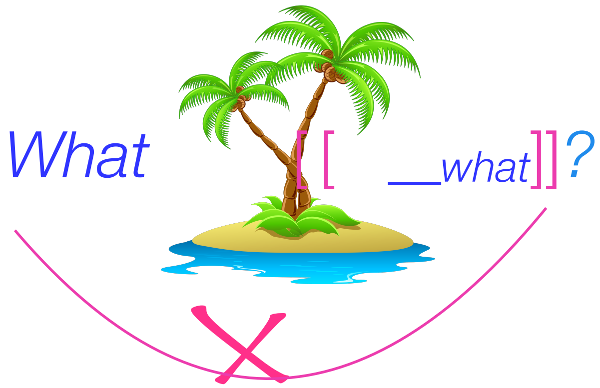
Who __*who*?

Who [CP… __*who*]?

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

Who [non-island]?

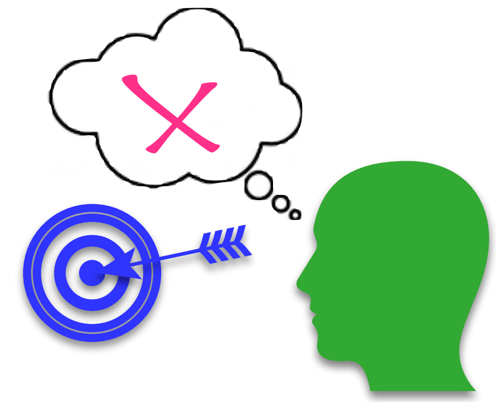
Who [island]?



Syntactic islands

Adult judgments

= behavioral target outcome



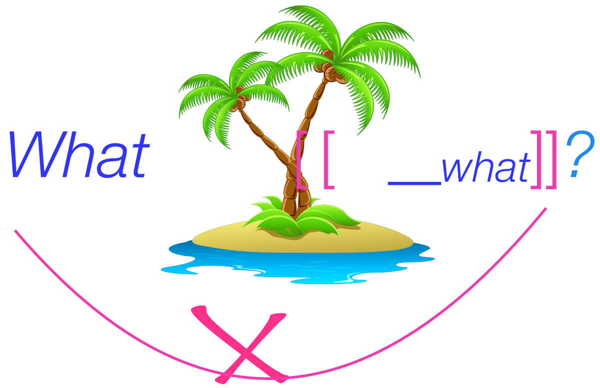
Adult knowledge as measured by acceptability judgment behavior

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

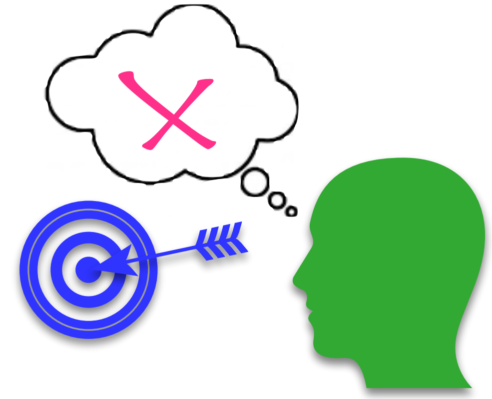
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



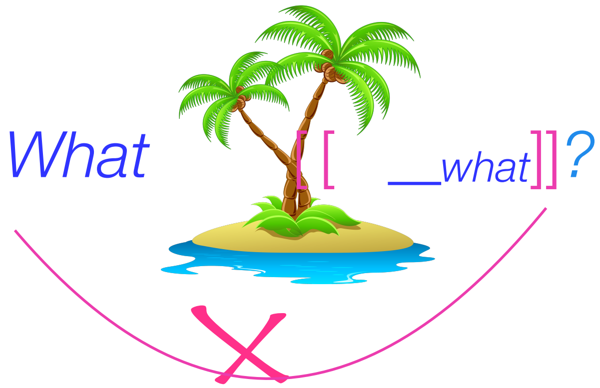
Adult knowledge as measured by acceptability judgment behavior

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

Whether island stimuli

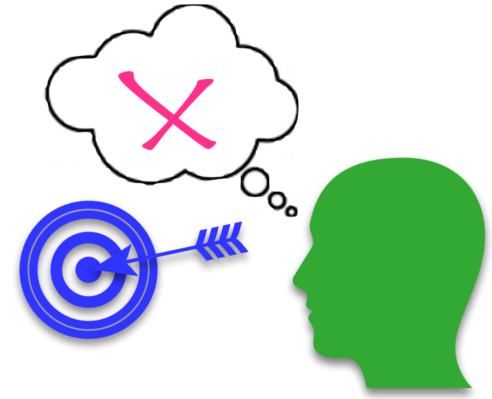
Who __ thinks [that Jack stole the necklace]?	matrix | non-island
What does the teacher think [that Jack stole __]?	embedded | non-island
Who __ wonders [whether Jack stole the necklace]?	matrix | island
*What does the teacher wonder [whether Jack stole __]?	embedded | island

Sprouse et al. 2012



Syntactic islands

Adult judgments
= behavioral target outcome



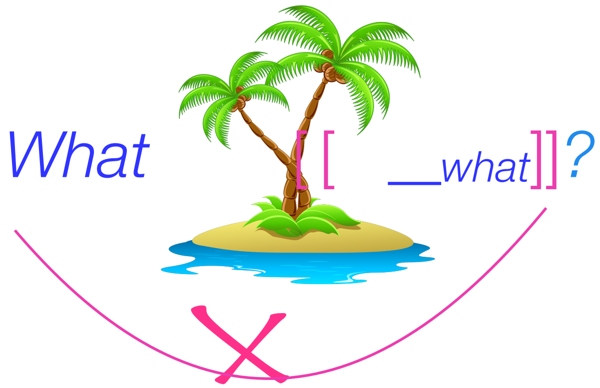
Adult knowledge as measured by acceptability judgment behavior

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

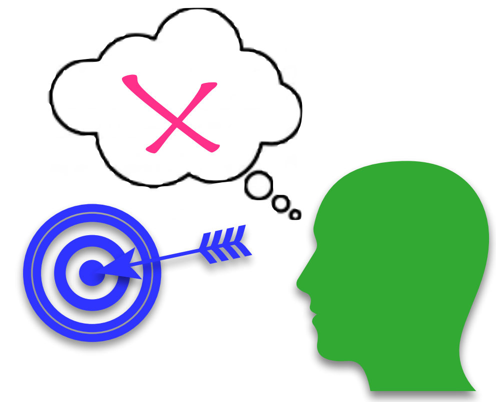
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



Adult knowledge as measured by acceptability judgment behavior

length of dependency
(matrix vs. embedded)

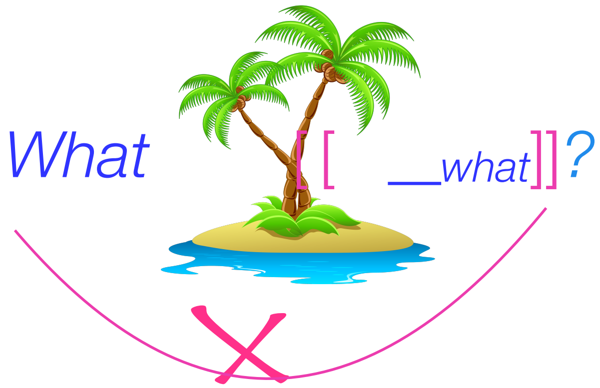
X

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

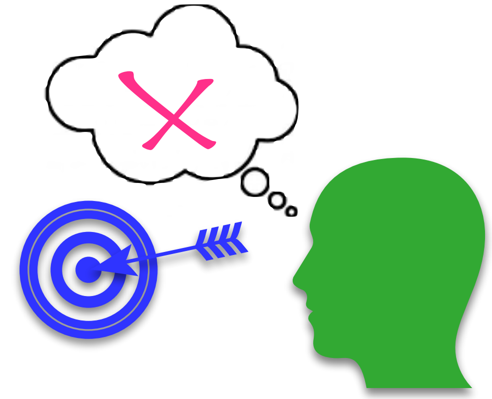
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



Adult knowledge as measured by acceptability judgment behavior

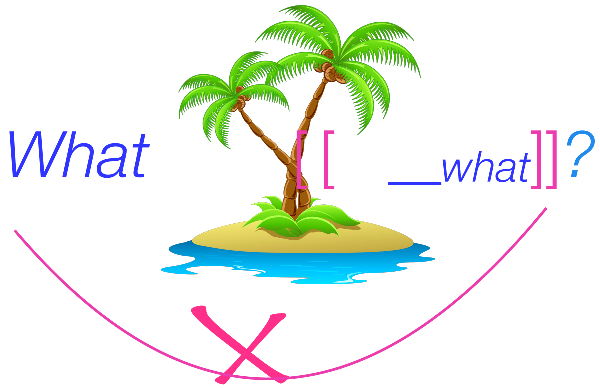
length of dependency
(**matrix** **vs.** **embedded**)

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.

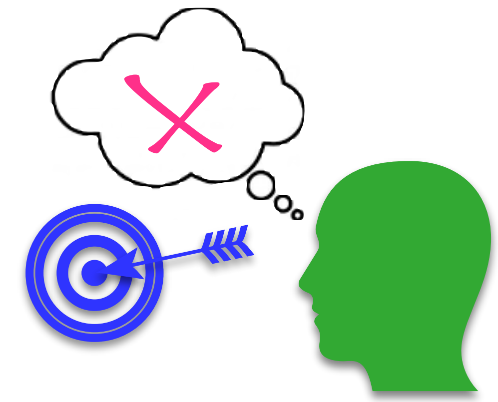
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



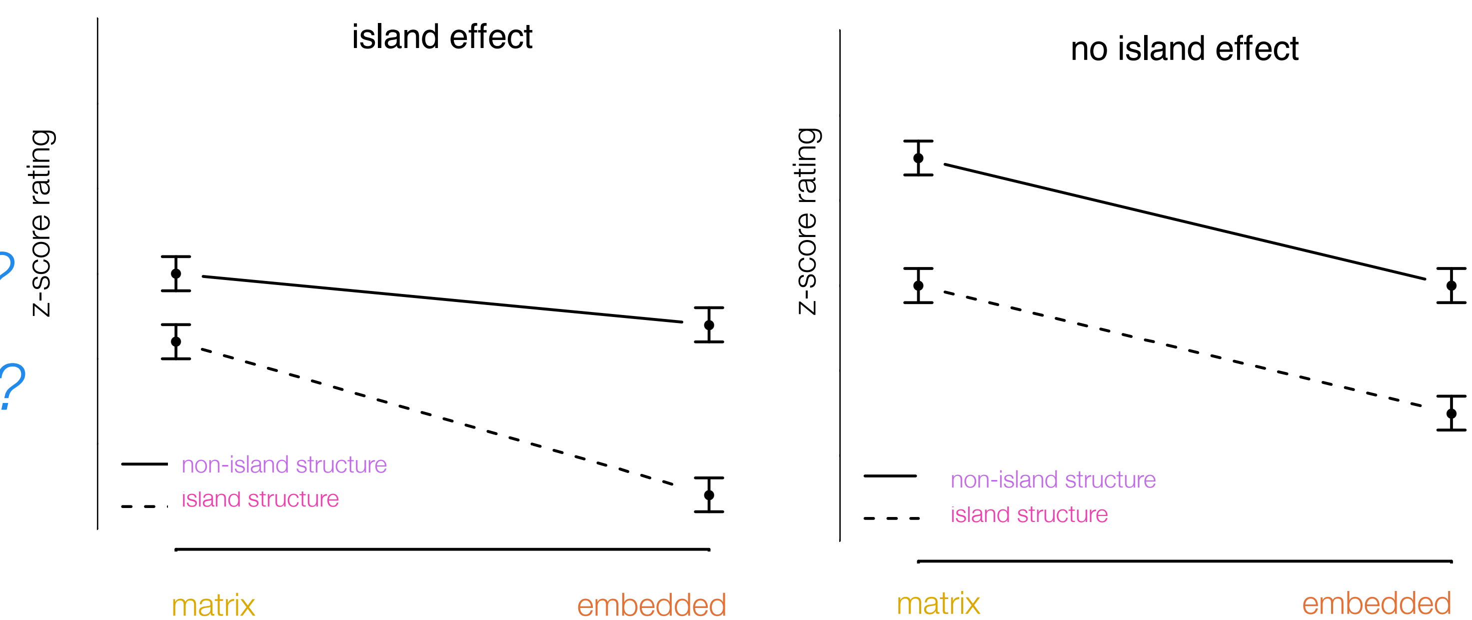
Adult knowledge as measured by acceptability judgment behavior

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]?

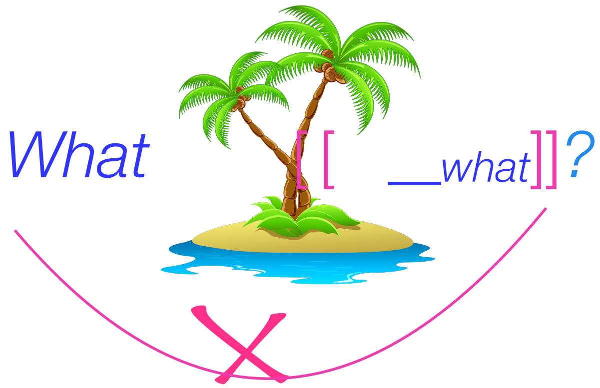
Who [island]?



Who __who?

Who [CP… __who]?

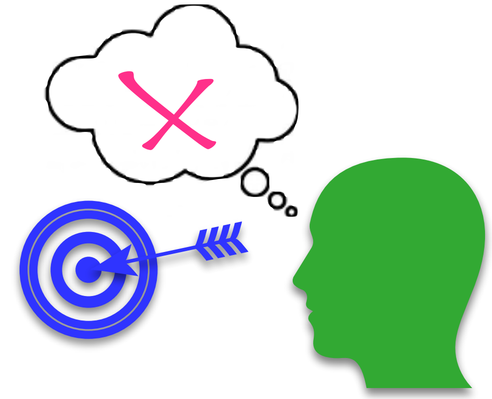
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



Adult knowledge as measured by acceptability judgment behavior

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

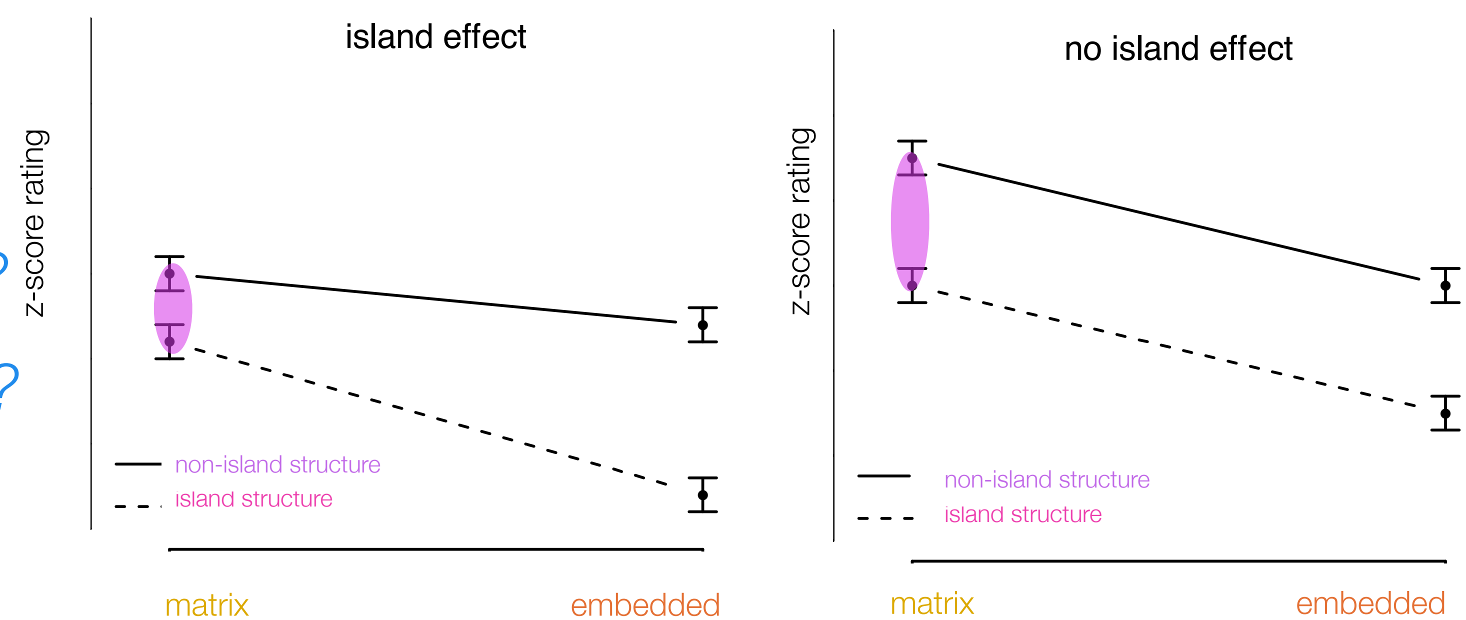
X

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

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

Who [non-island]?

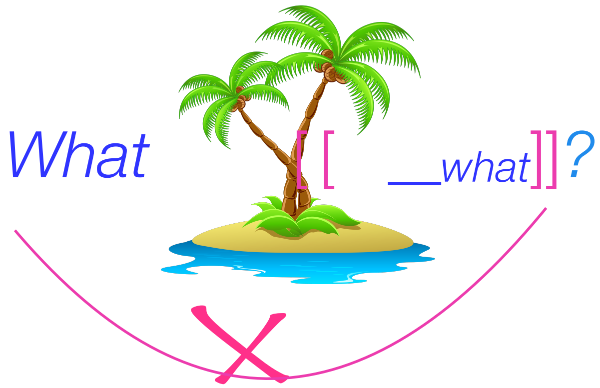
Who [island]?



Who *__who*?

Who [CP… *__who*]?

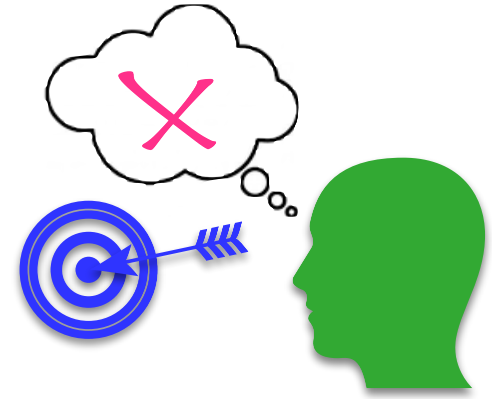
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



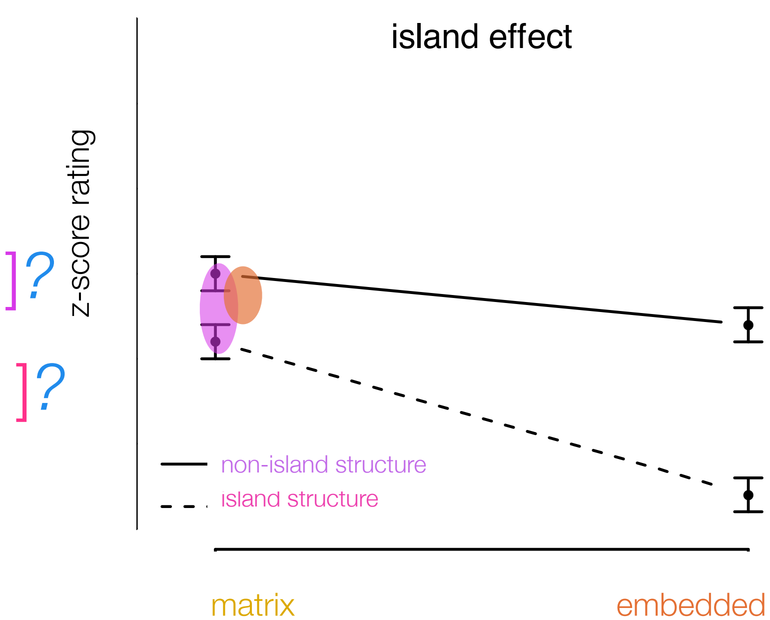
Adult knowledge as measured by acceptability judgment behavior

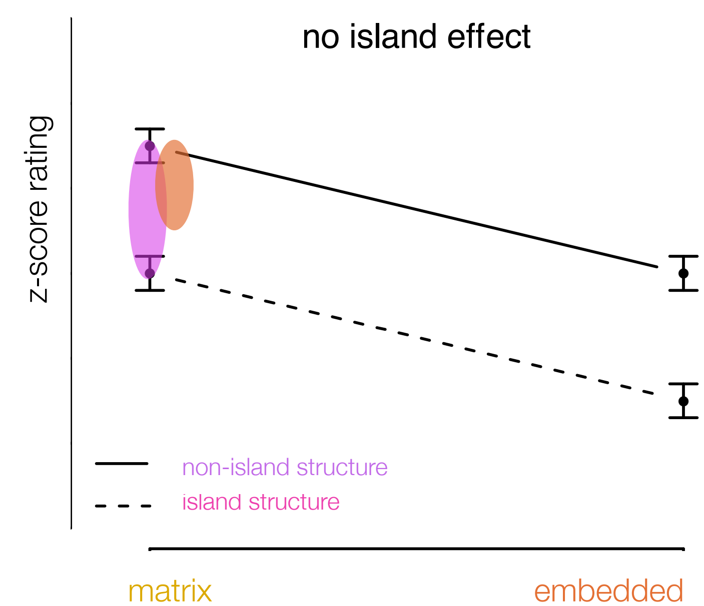
length of dependency (**matrix** vs. **embedded**) X presence of an island structure (**non-island** vs. **island**)

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

Who [non-island]?

Who [island]?

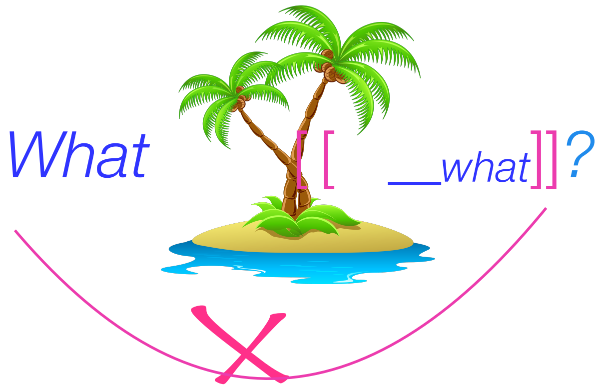




Who __who?

Who [CP… __who]?

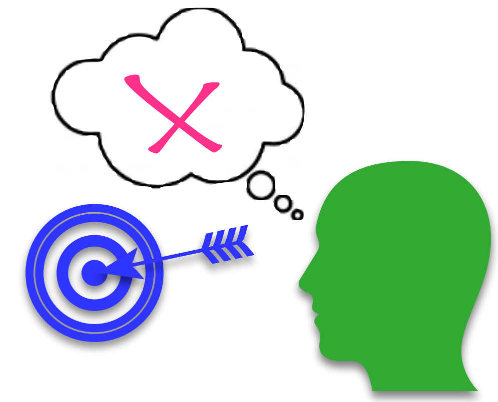
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



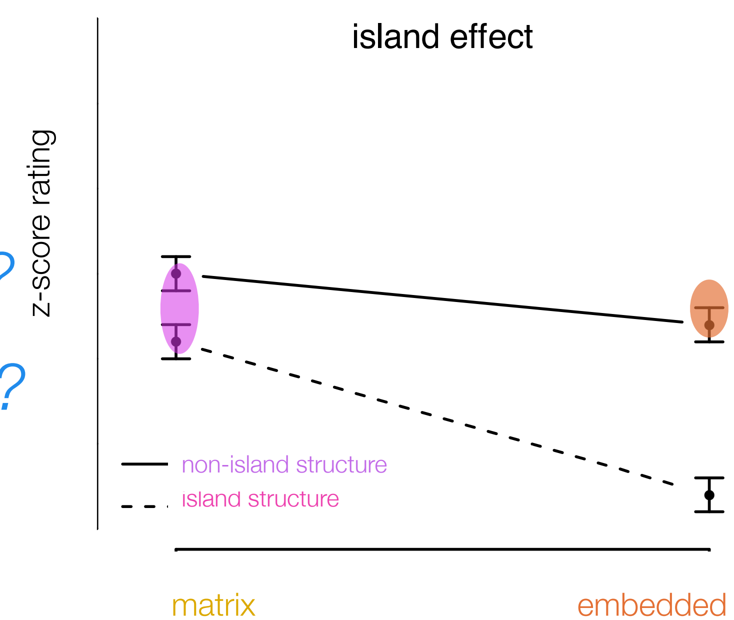
Adult knowledge as measured by acceptability judgment behavior

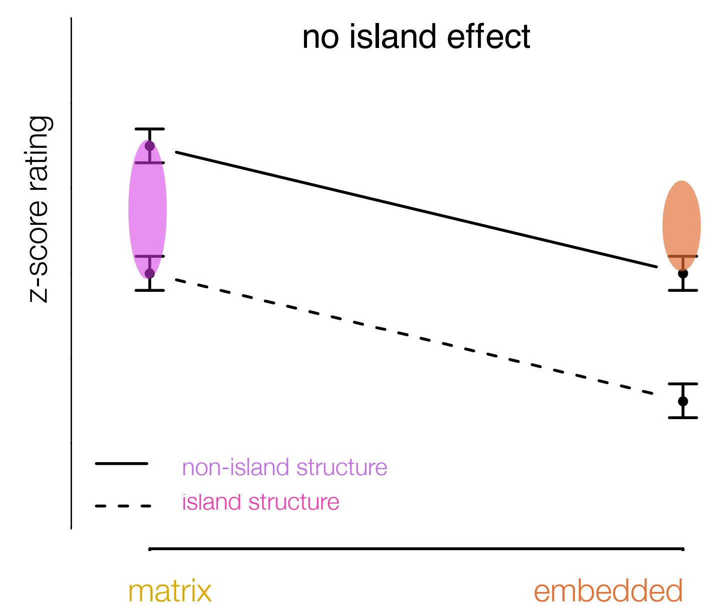
length of dependency (**matrix vs. embedded**) X presence of an island structure (**non-island vs. island**)

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

Who [non-island]?

Who [island]?

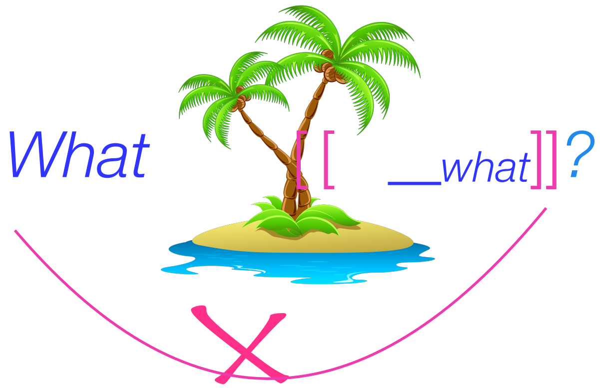




Who __who?

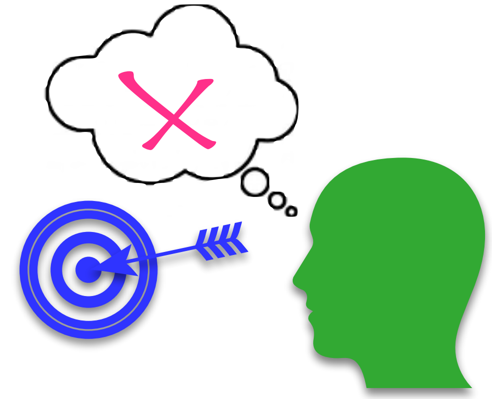
Who [CP… __who]?

Sprouse et al. 2012



Syntactic islands

Adult judgments
= behavioral target outcome



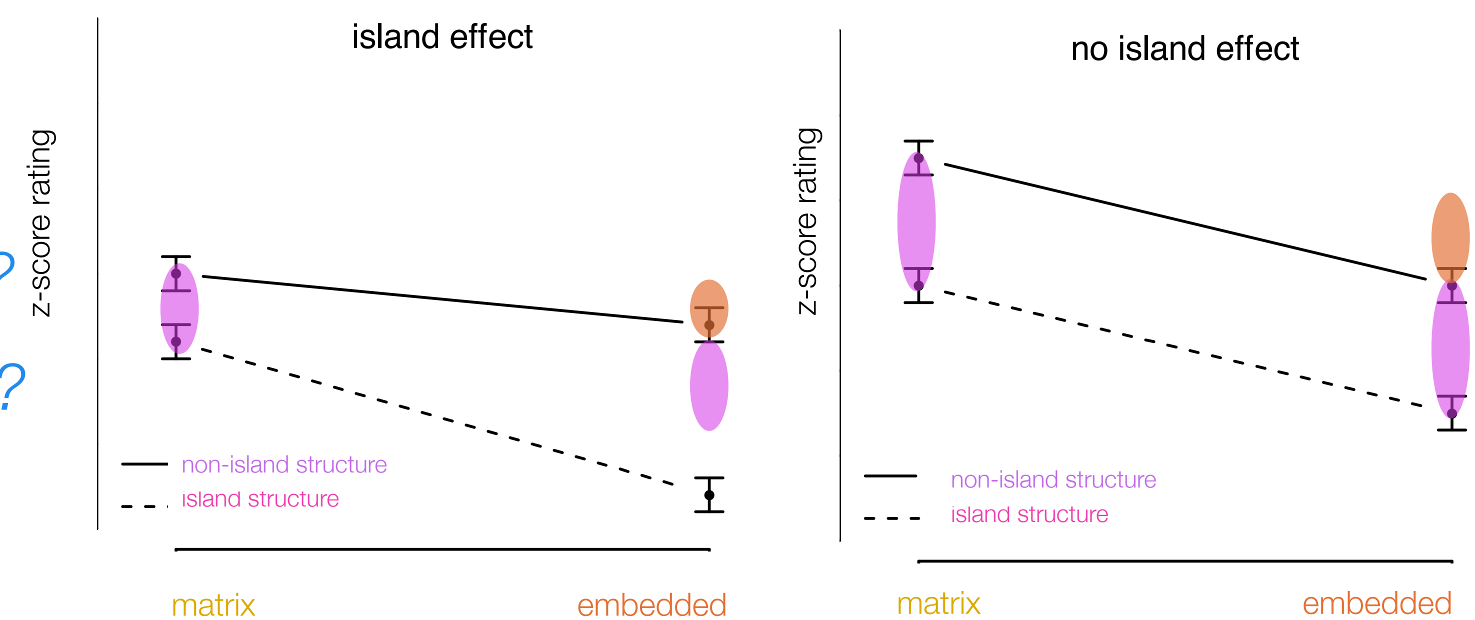
Adult knowledge as measured by acceptability judgment behavior

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

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

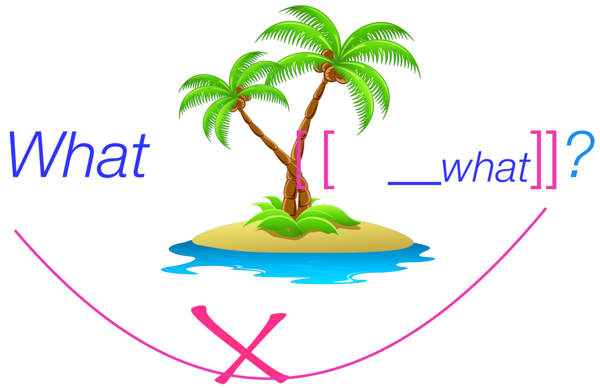
Who [non-island]?

Who [island]?



Who *__who*?   *Who* [CP… *__who*]?

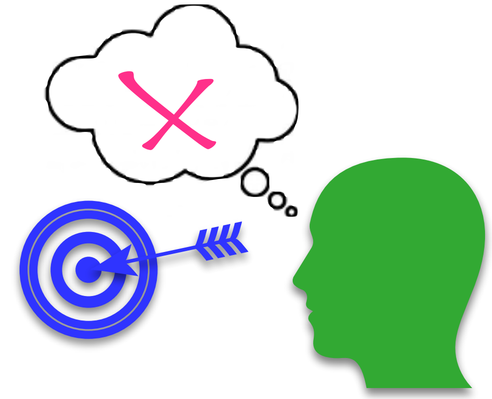
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome



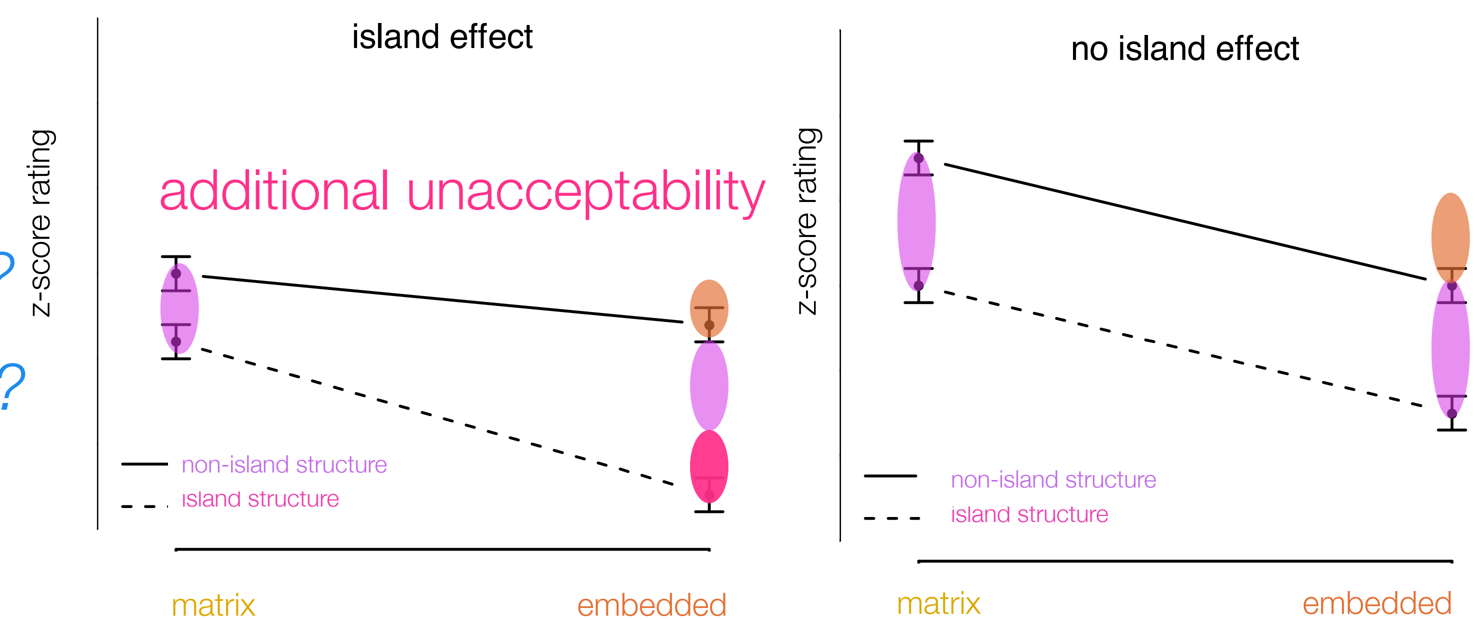
Adult knowledge as measured by acceptability judgment behavior

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

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

Who [non-island]?

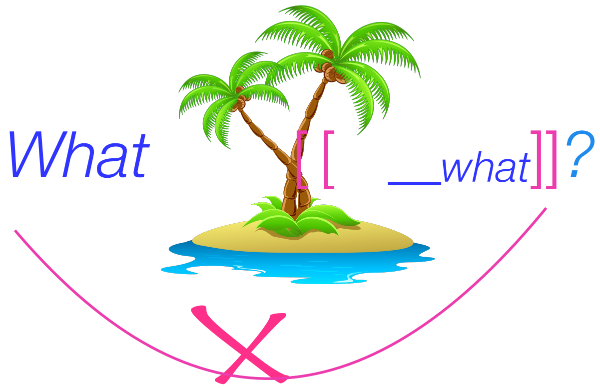
Who [island]?



Who *__who*?

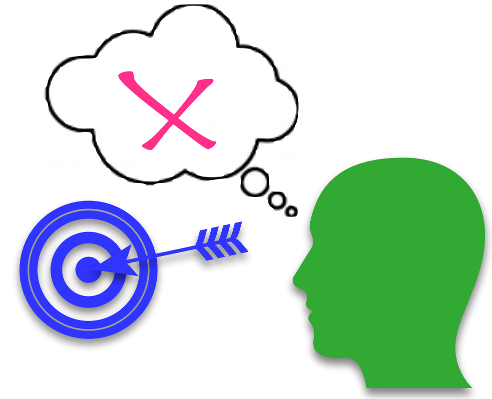
Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands

Adult judgments
= behavioral target outcome

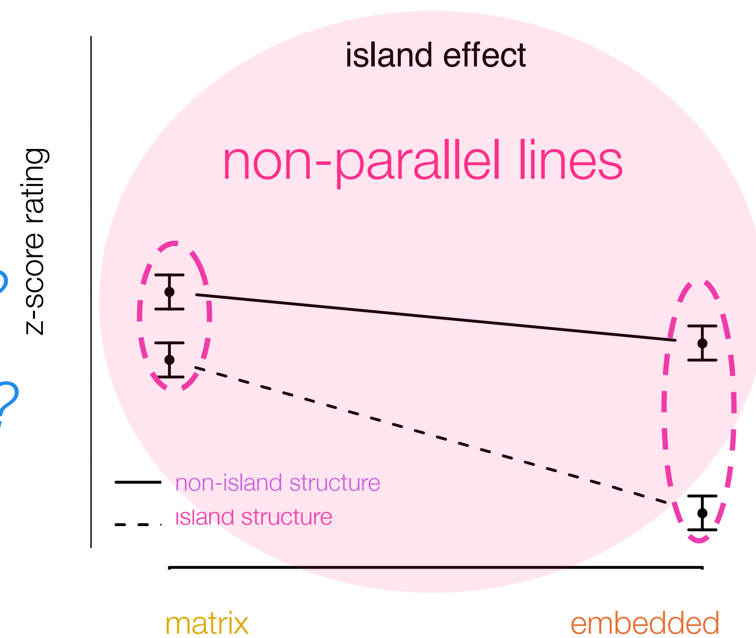


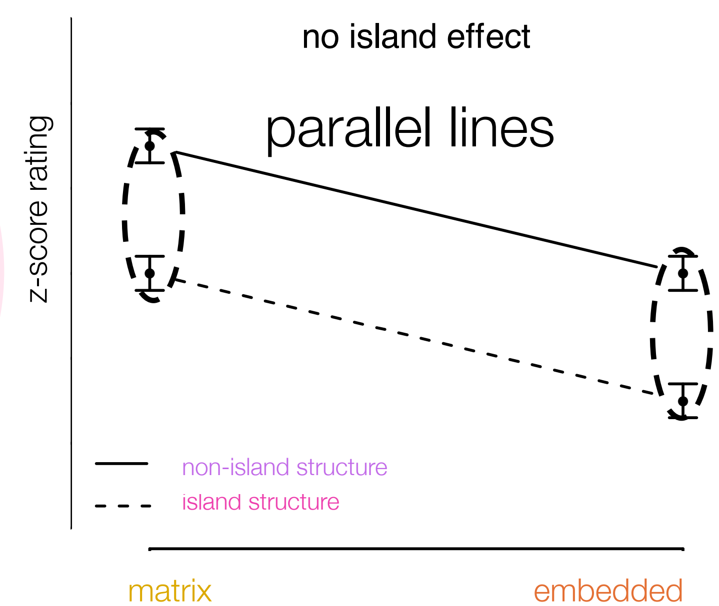
Adult knowledge as measured by acceptability judgment behavior

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

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

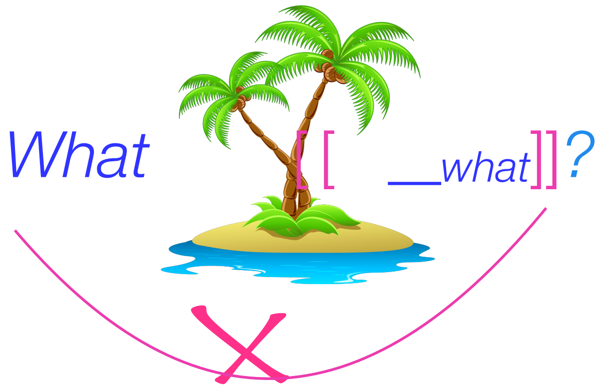
Who [non-island]?
Who [island]?





Who __who?
Who [CP… __who]?

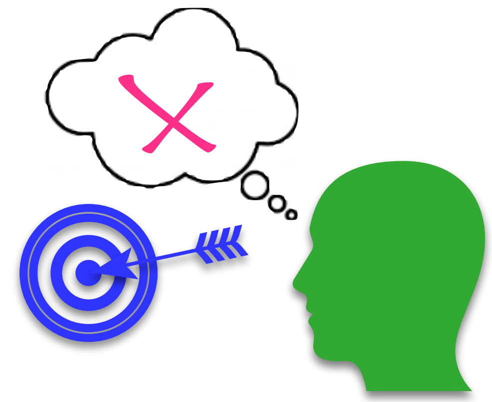
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome

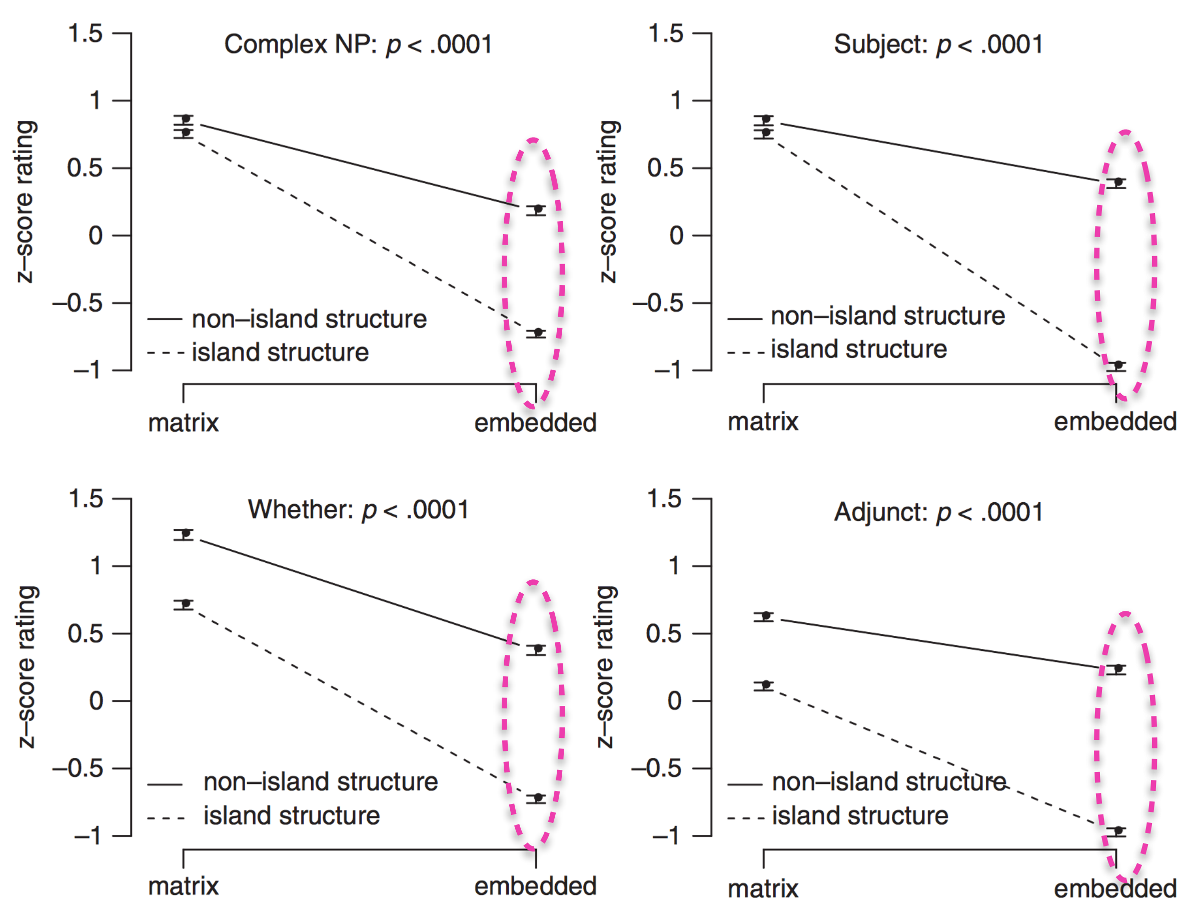


Adult knowledge as measured by acceptability judgment behavior

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

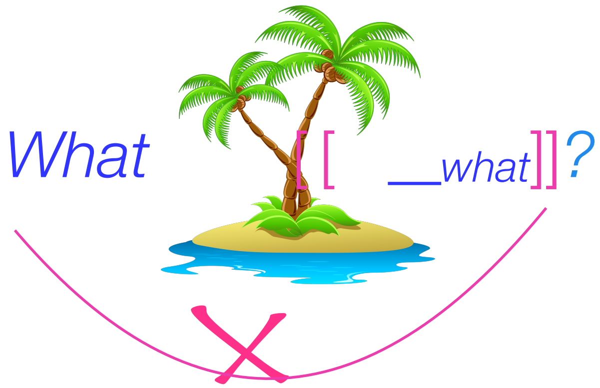
Syntactic island = **superadditive** interaction of the two factors

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



superadditivity for all four island types

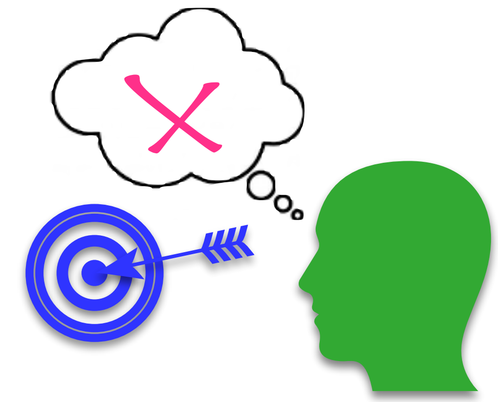
Sprouse et al. 2012



Syntactic islands

Adult judgments

= behavioral target outcome

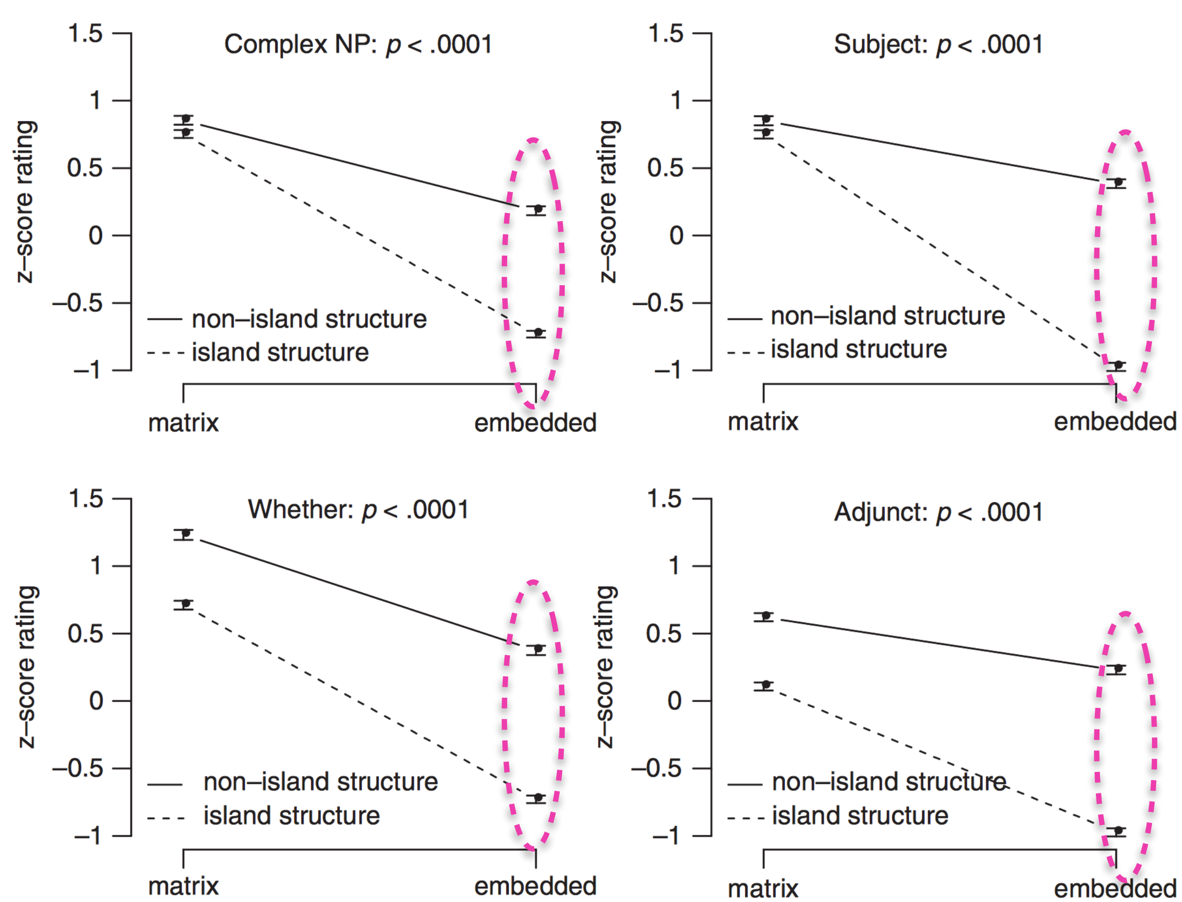


Adult knowledge as measured by acceptability judgment behavior

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

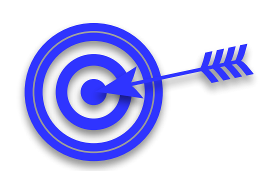
Syntactic island = **superadditive** interaction of the two factors

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

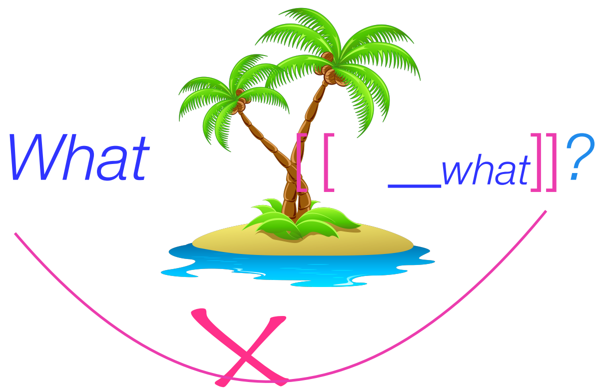


superadditivity for all four island types

= knowledge that dependencies crossing these island structures are dispreferred.



Sprouse et al. 2012



Syntactic islands

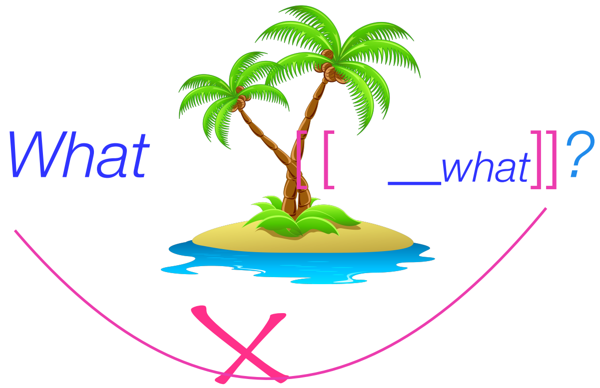
Child judgments

= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

De Villiers et al. 2008:

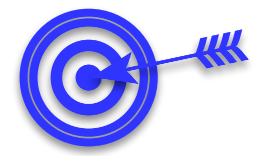
How do children prefer to interpret potentially ambiguous *wh*-questions?



Syntactic islands

Child judgments

= behavioral target outcome

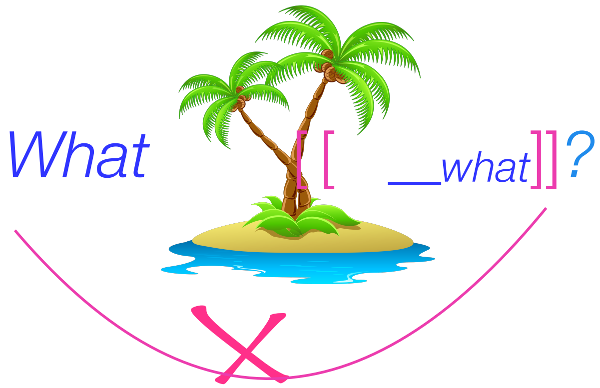


Child knowledge as measured by preferred interpretation behavior

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

$\boxed{\text{context}}$

De Villiers et al. 2008



Syntactic islands

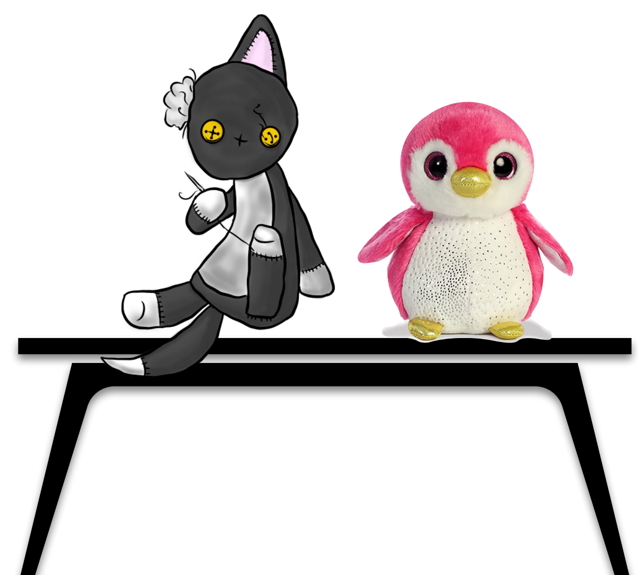
Child judgments

= behavioral target outcome

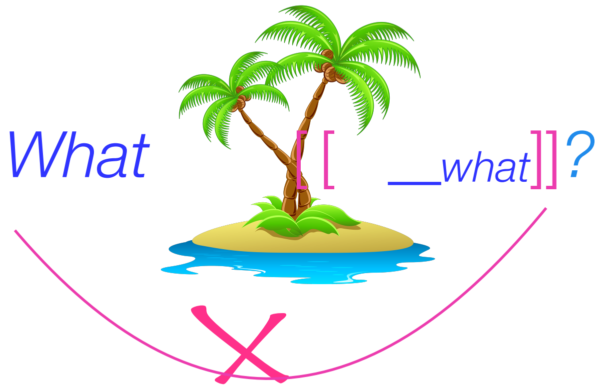
Child knowledge as measured by preferred interpretation behavior

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

context



De Villiers et al. 2008



Syntactic islands

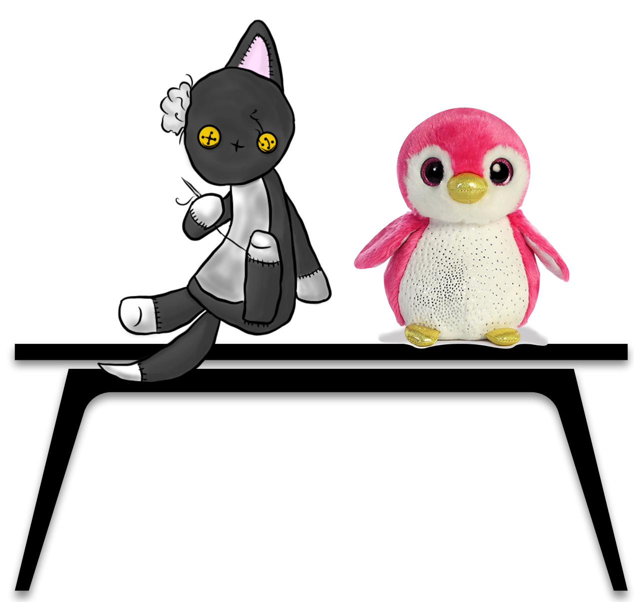
Child judgments

= behavioral target outcome

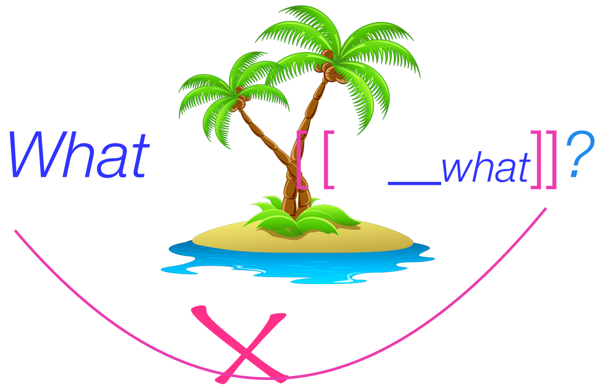
Child knowledge as measured by preferred interpretation behavior

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*?



De Villiers et al. 2008



Syntactic islands

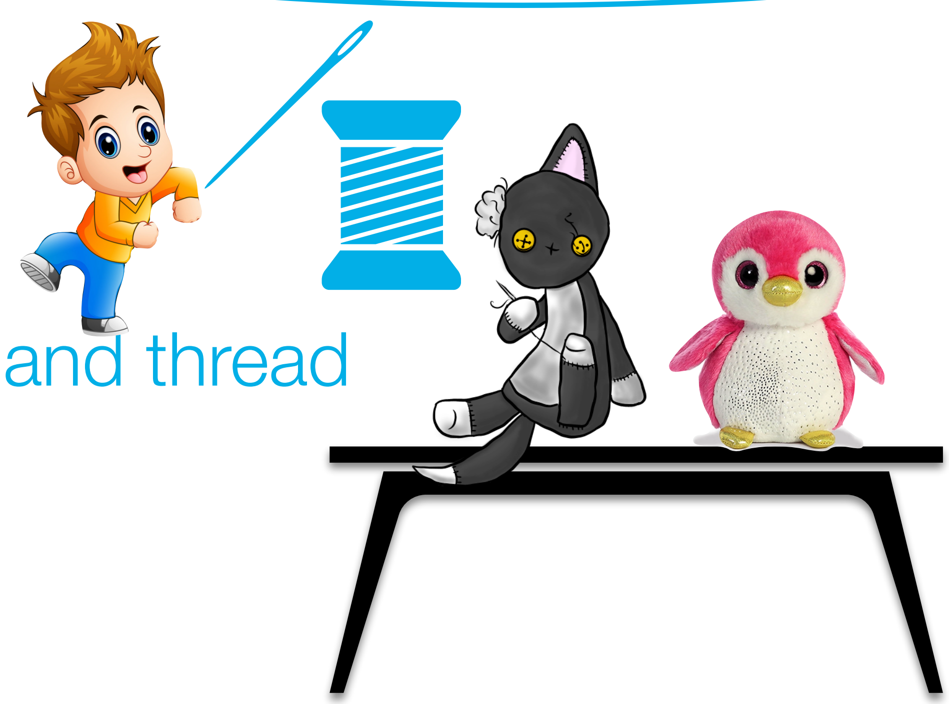
Child judgments

= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

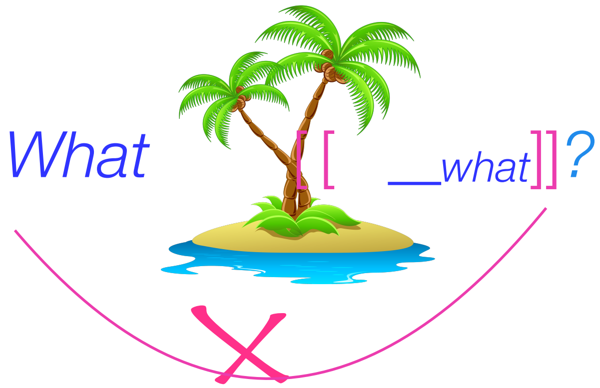
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

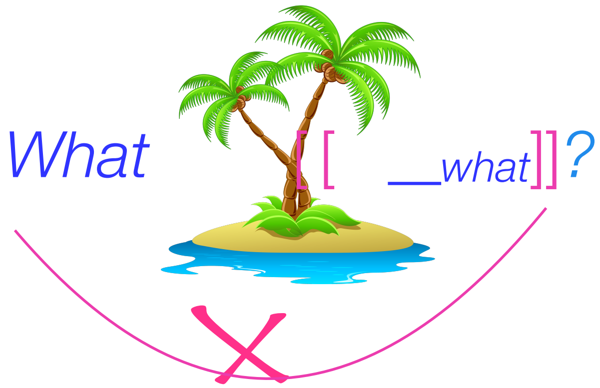
Child knowledge as measured by preferred interpretation behavior

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

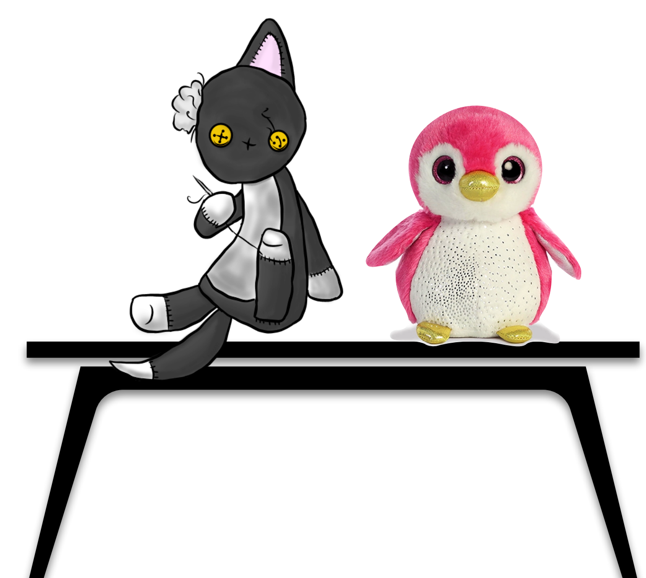
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

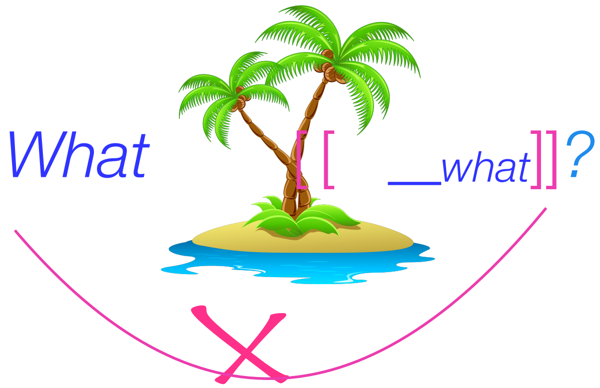
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

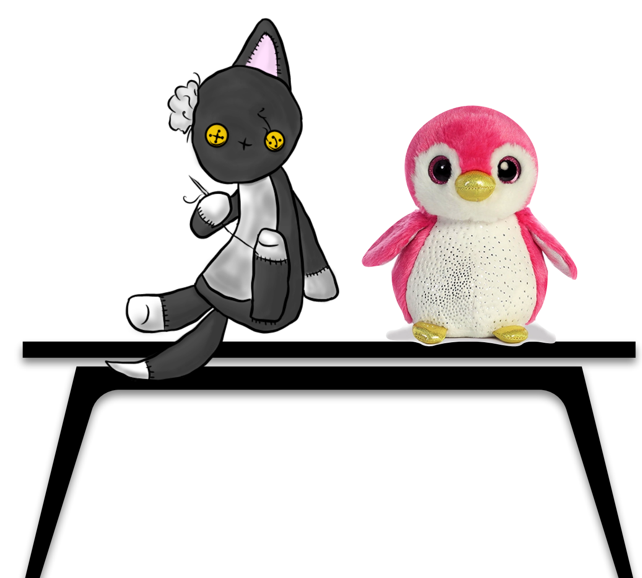
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

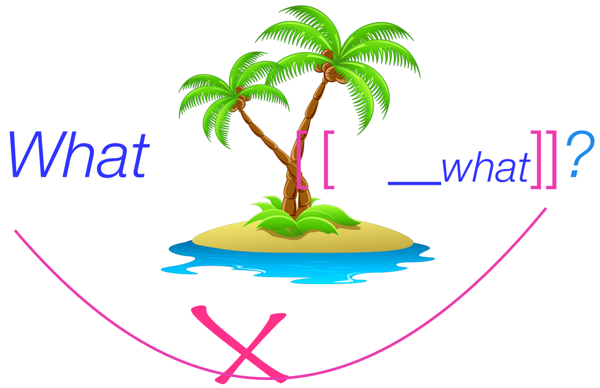
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

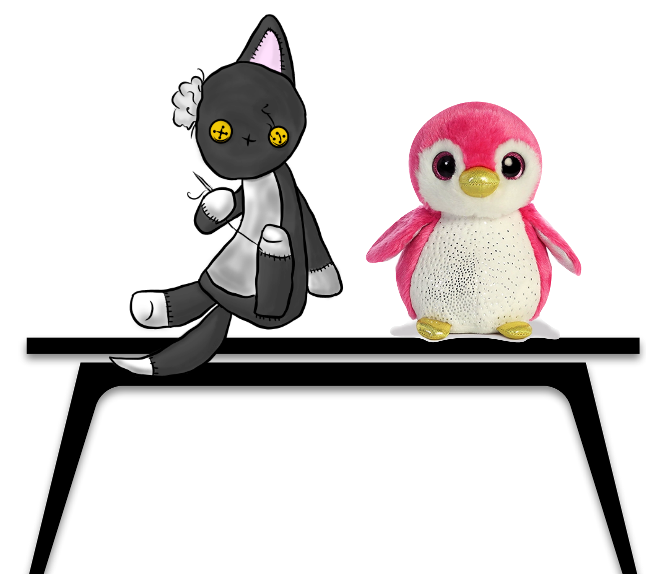
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

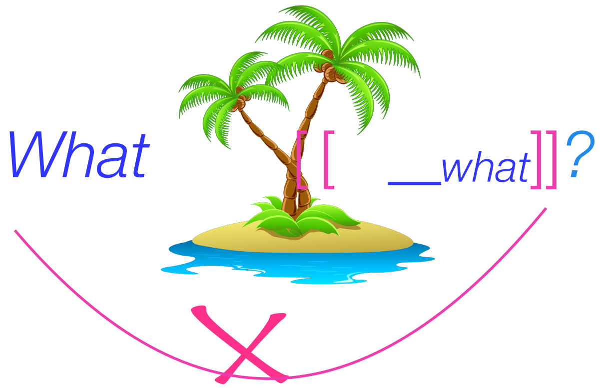
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

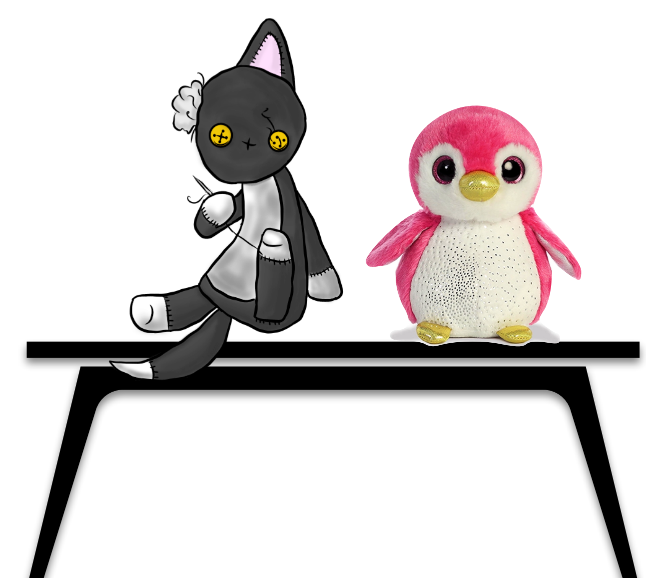
= behavioral target outcome

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.

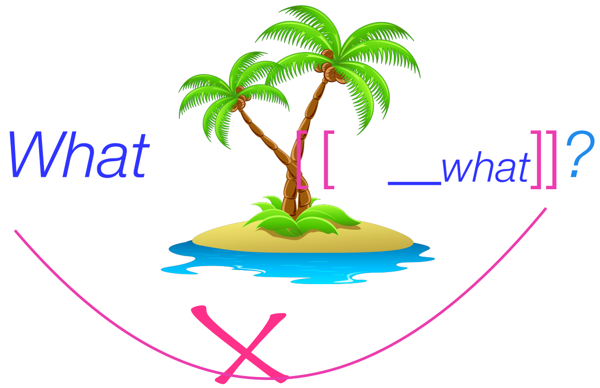


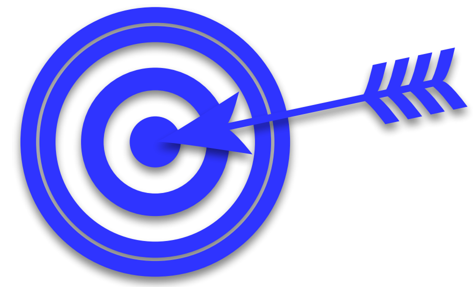
De Villiers et al. 2008

Syntactic islands

Adult & child judgments

= behavioral target outcome



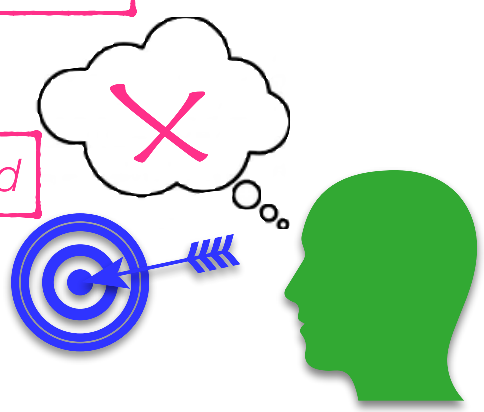


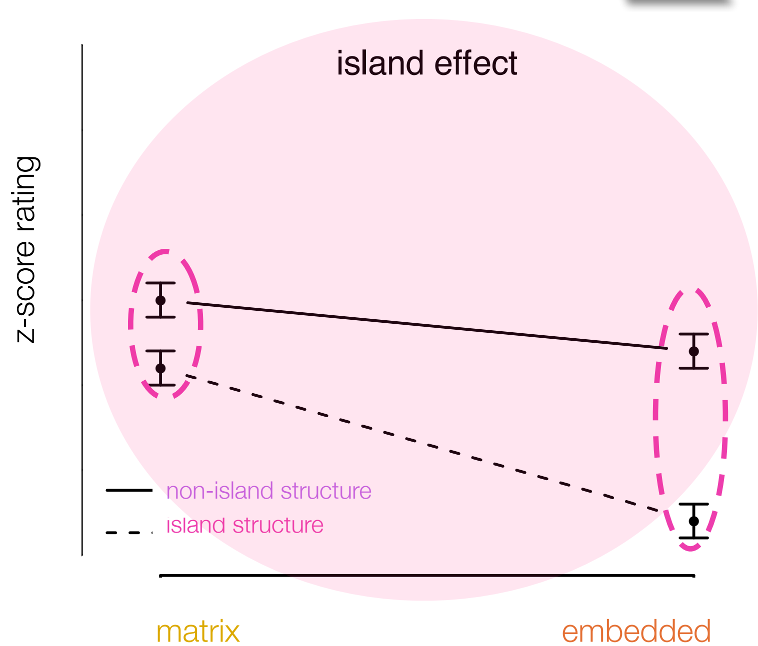
Subject island

Complex NP island

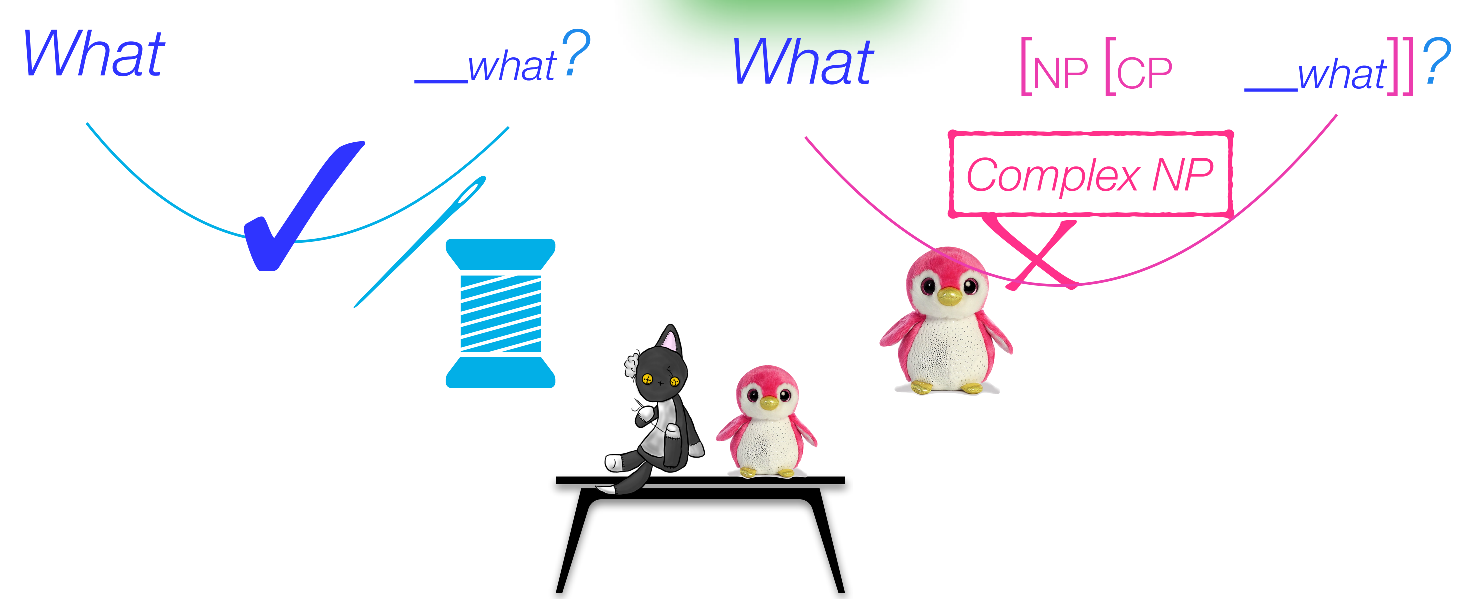
Whether island

Adjunct island



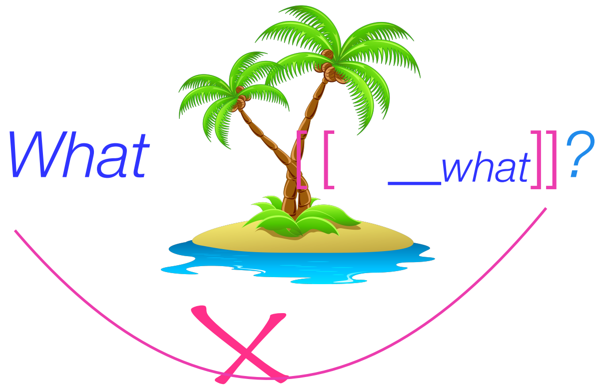


Sprouse et al. 2012



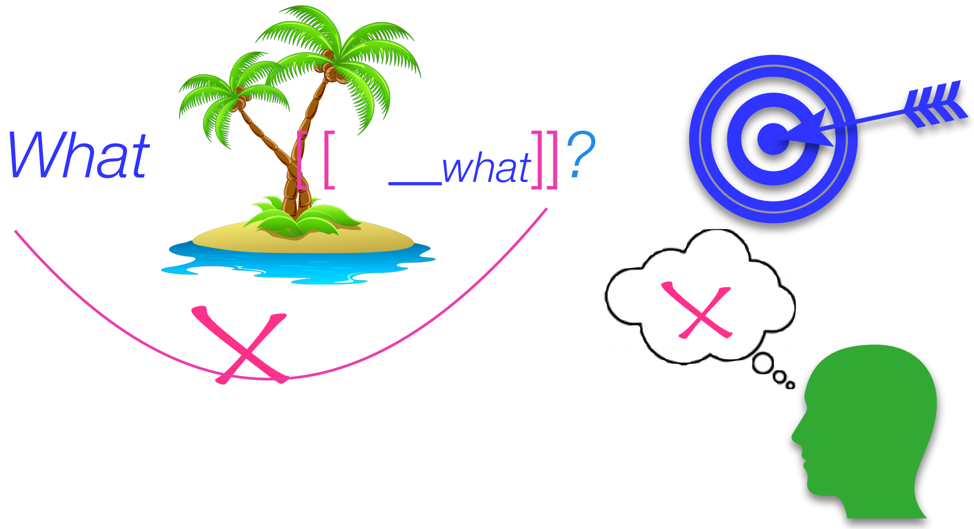
De Villiers et al. 2008

Syntactic islands



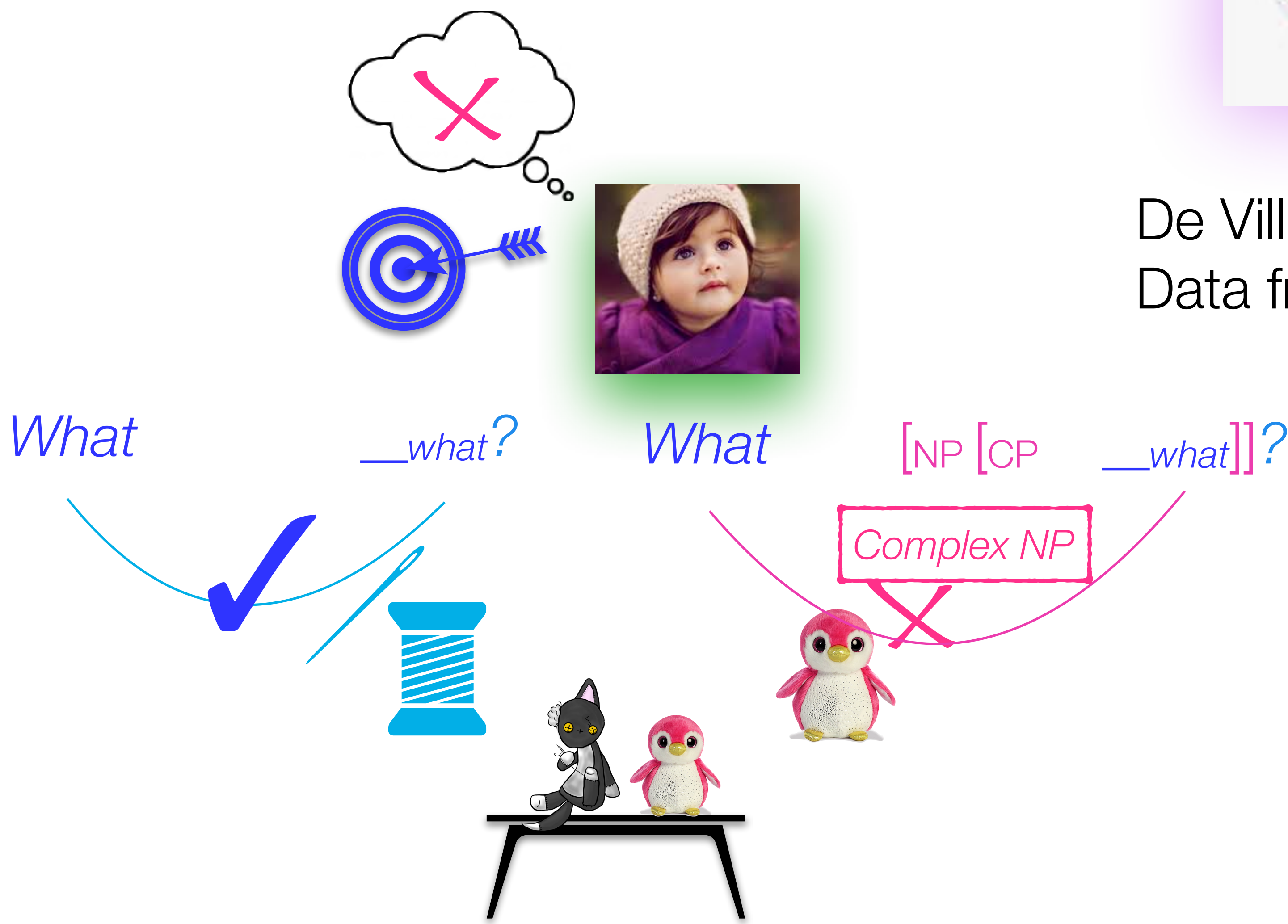
How long do children have to learn?

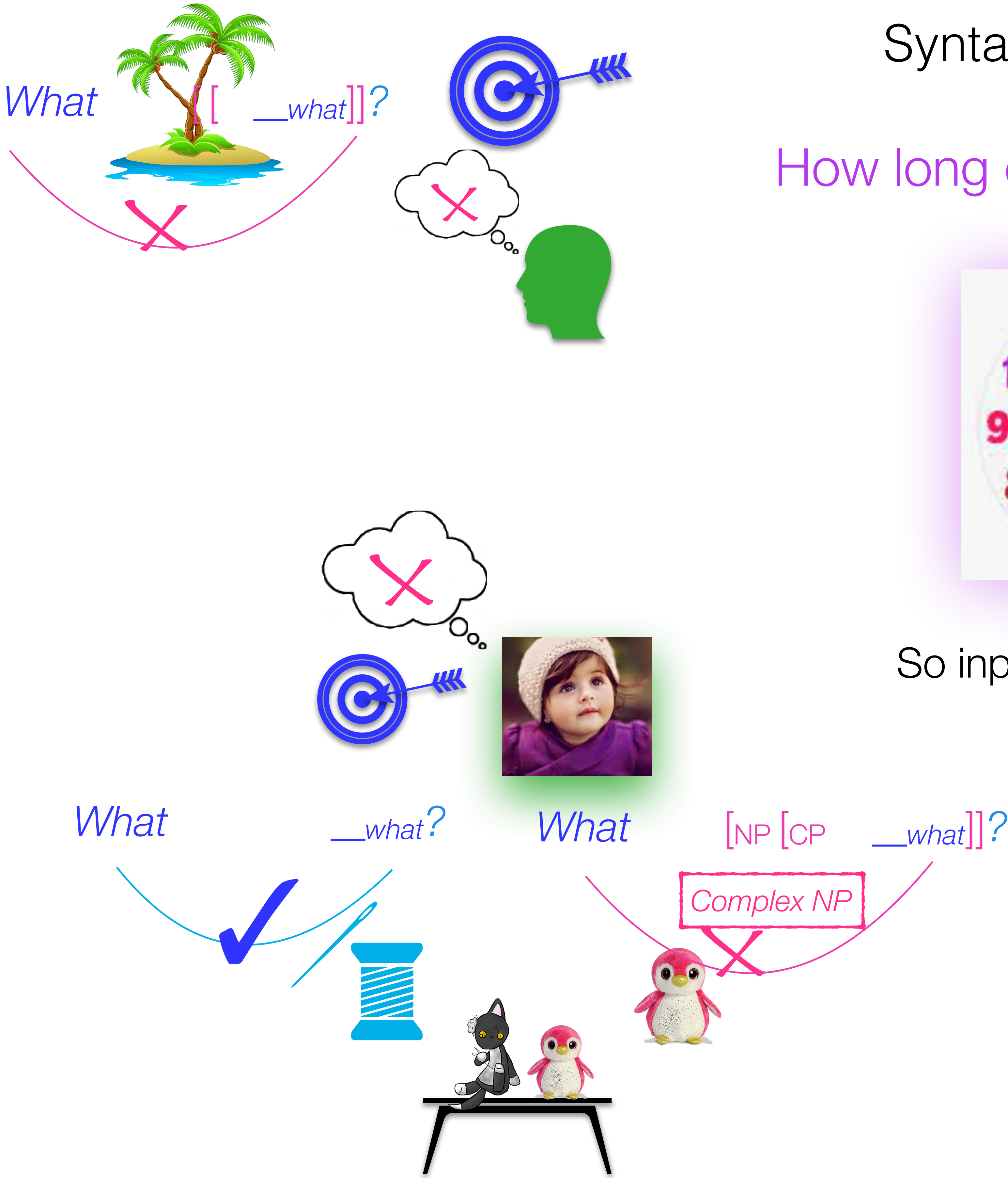
Syntactic islands



How long do children have to learn?

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



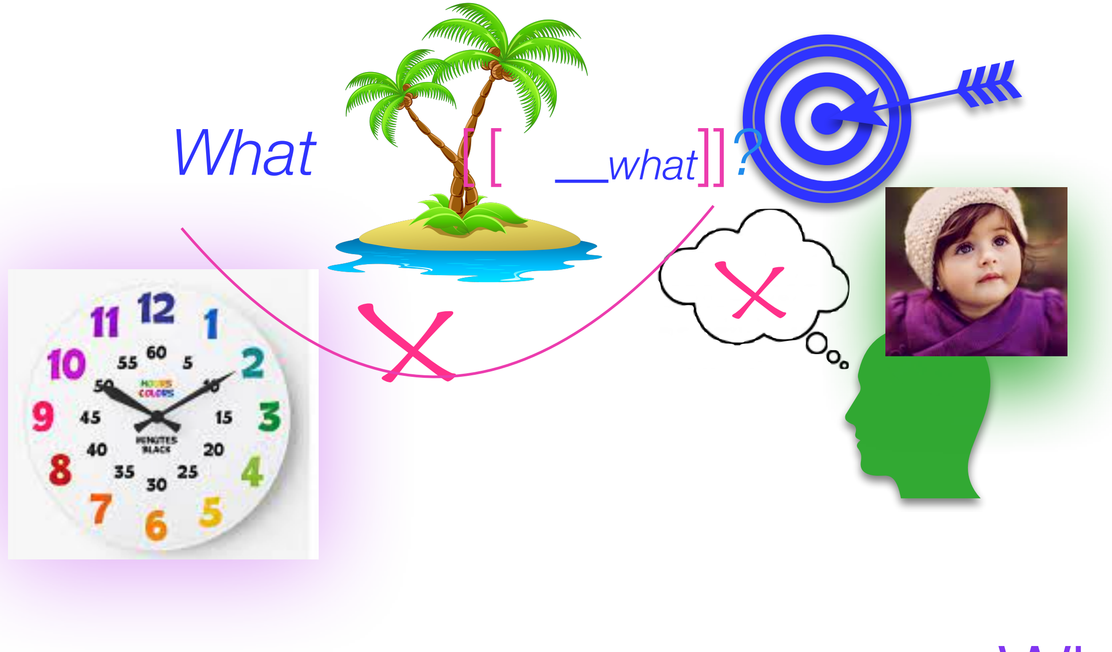


Syntactic islands

How long do children have to learn?

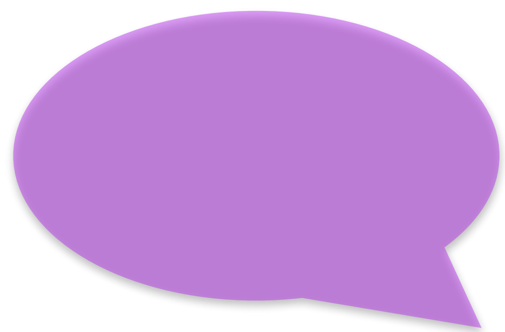
So input through age four.

(<60 months)

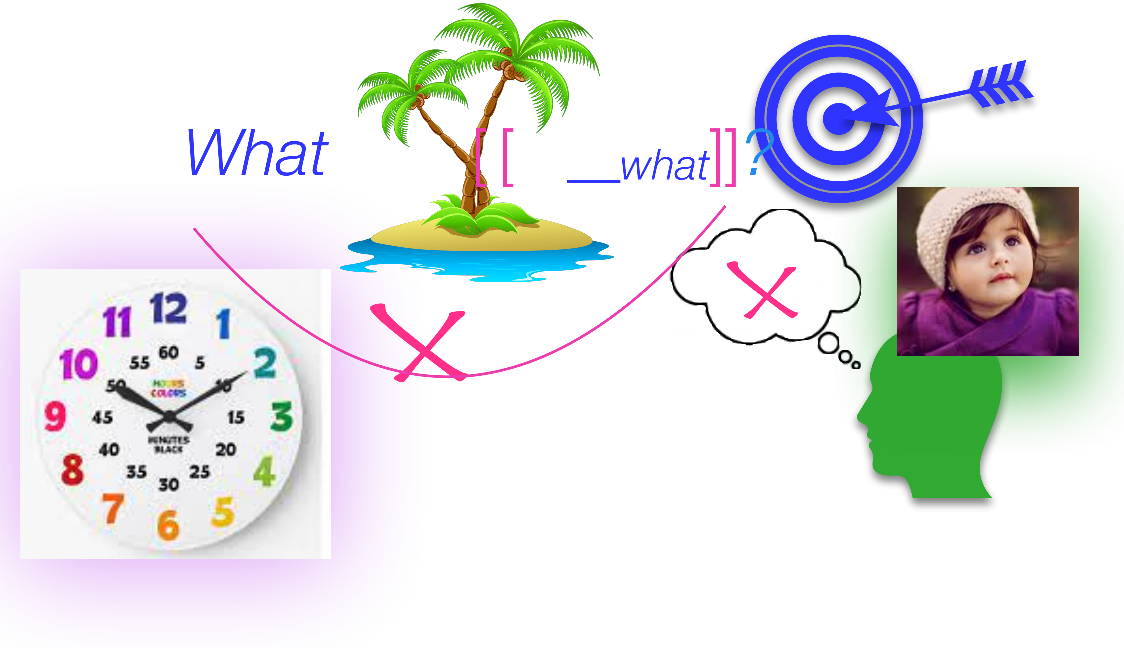


Syntactic islands

What input do children get?



Syntactic islands

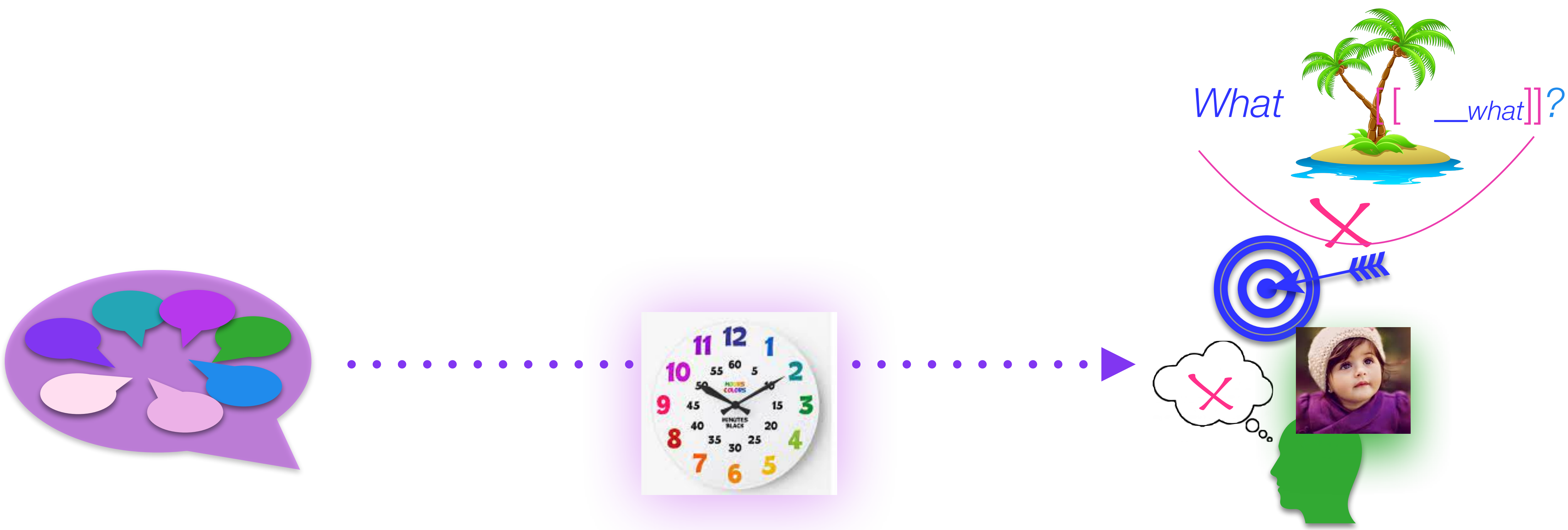


What input do children get?

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

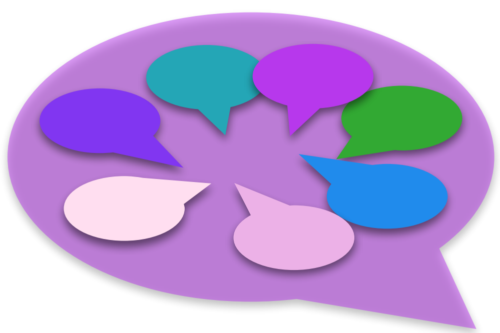
Syntactic islands

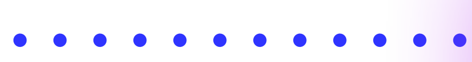
This is the acquisition problem



Syntactic islands

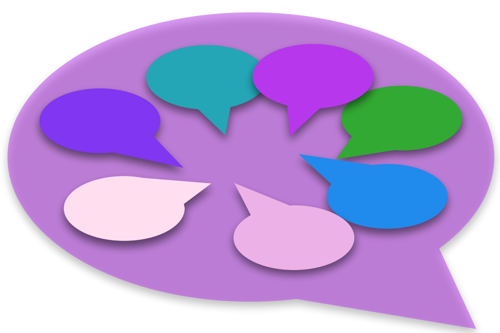
…which is where a theory of acquisition comes in.





Syntactic islands

…which is where a theory of acquisition comes in.

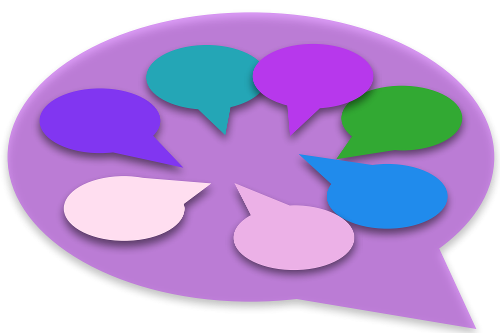


1: probabilities of pre-specified pieces

2: what the pieces are and their probabilities

Syntactic islands

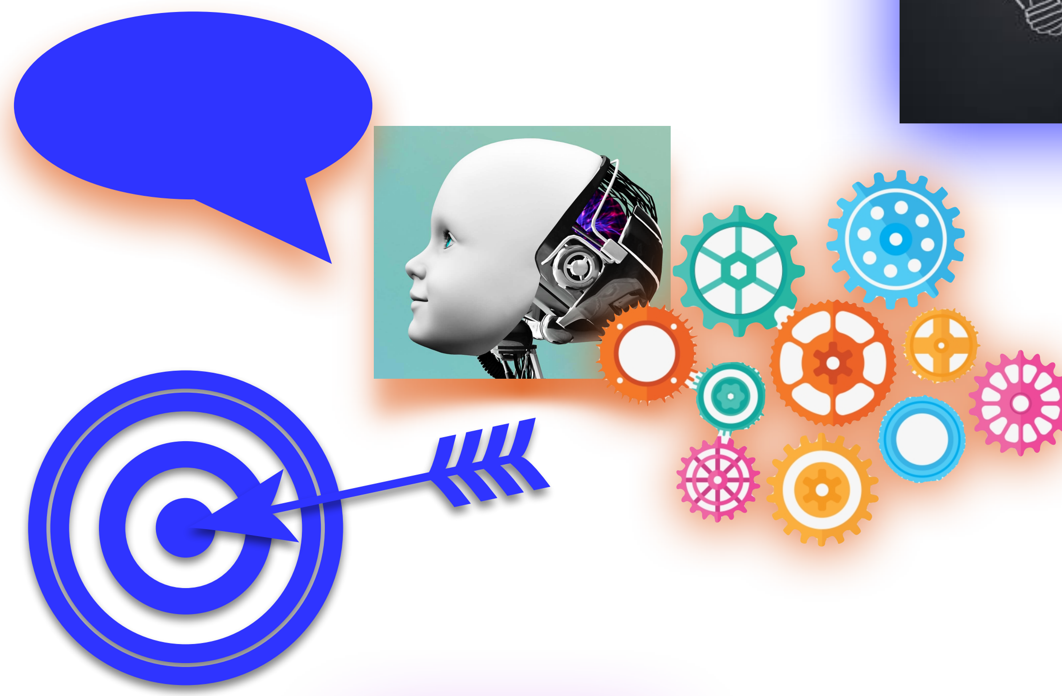
We can evaluate a theory by implementing it concretely in a computational cognitive model.



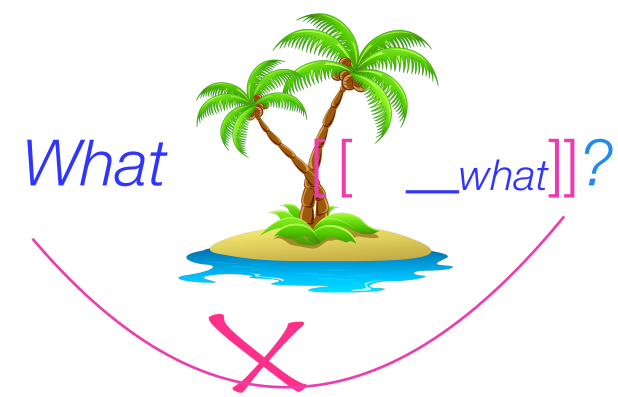
1: probabilities of pre-specified pieces

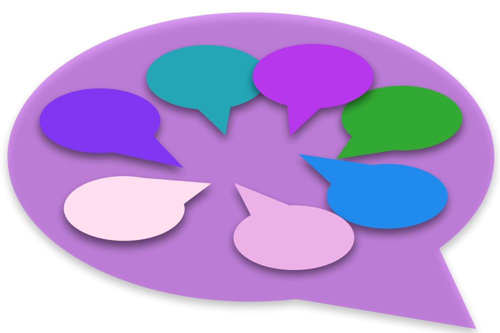
2: what the pieces are and their probabilities

Syntactic islands



The model generates predictions that can be compared with available empirical data.



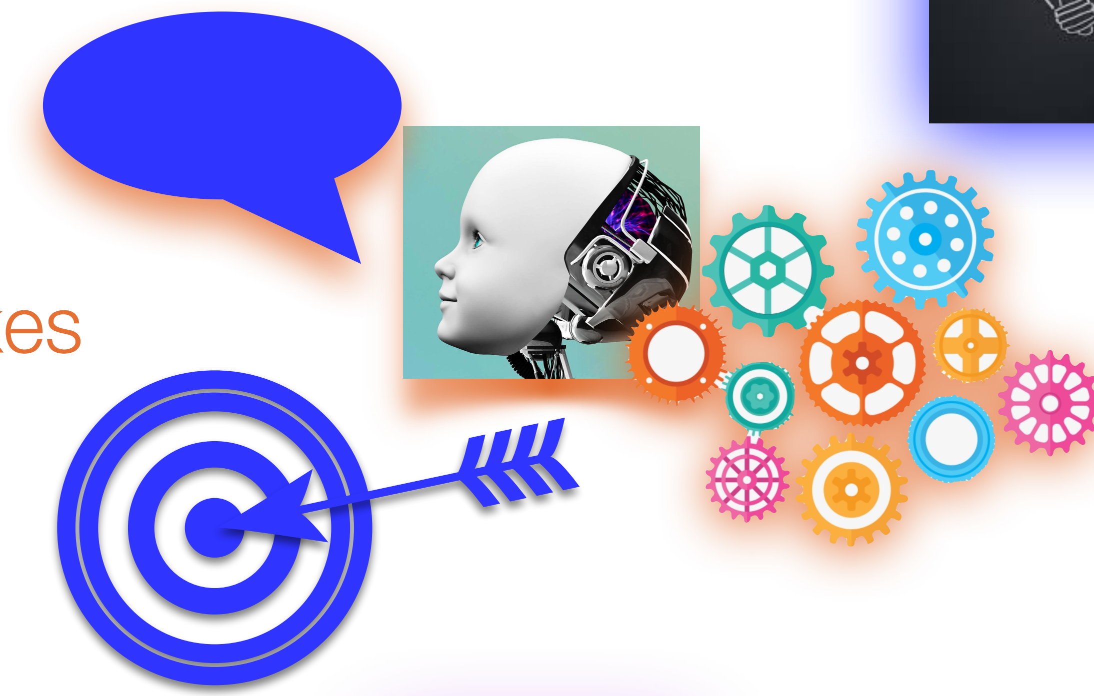


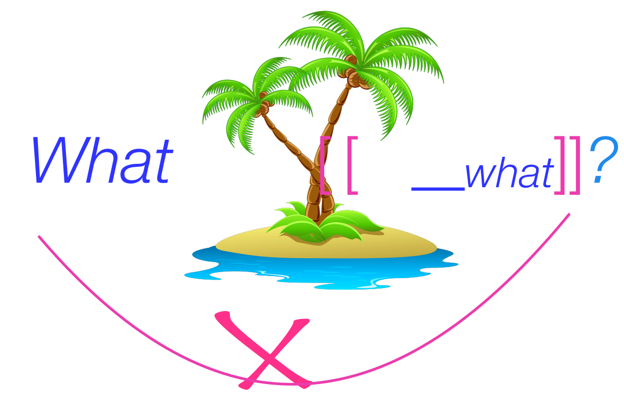
1: probabilities of pre-specified pieces

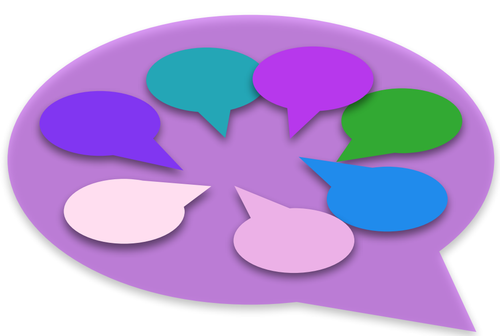
2: what the pieces are and their probabilities

Syntactic islands

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



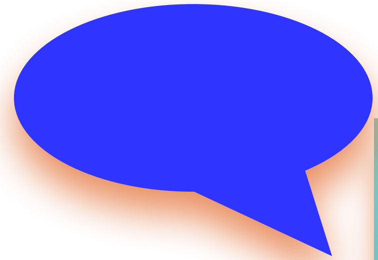


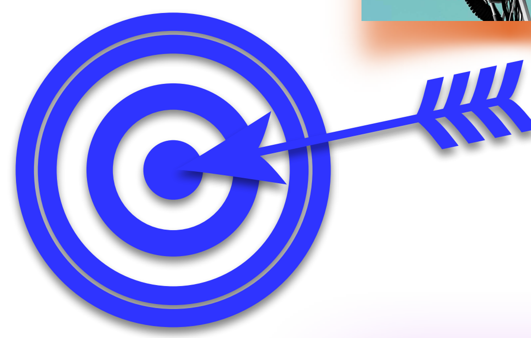


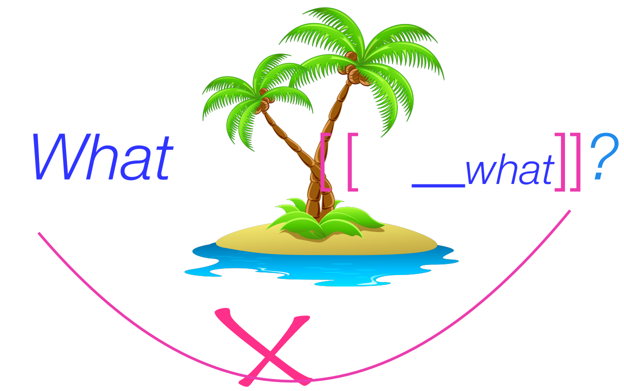
1: probabilities of pre-specified pieces

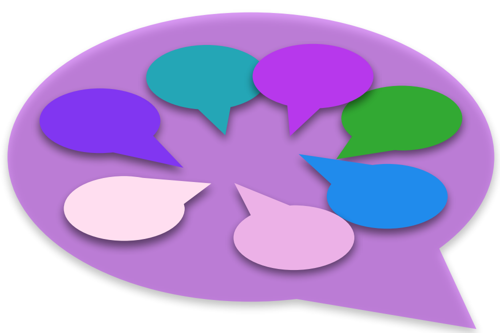
2: what the pieces are and their probabilities

Syntactic islands







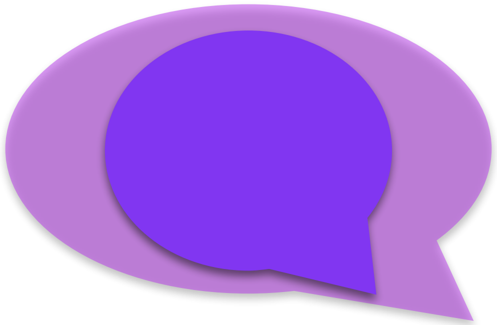


So let's do this for
our two theories.

1: probabilities of
pre-specified pieces

2: what the pieces are
and their probabilities

Both theories

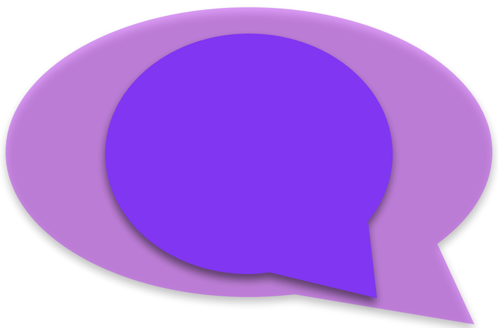


Intuition:

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

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, in prep.

Both theories



Intuition:

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

- Apply it to generate behavior for *wh*-dependencies you haven't seen before, like those crossing syntactic islands (or other longer *wh*-dependencies).

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, in prep.

Both theories

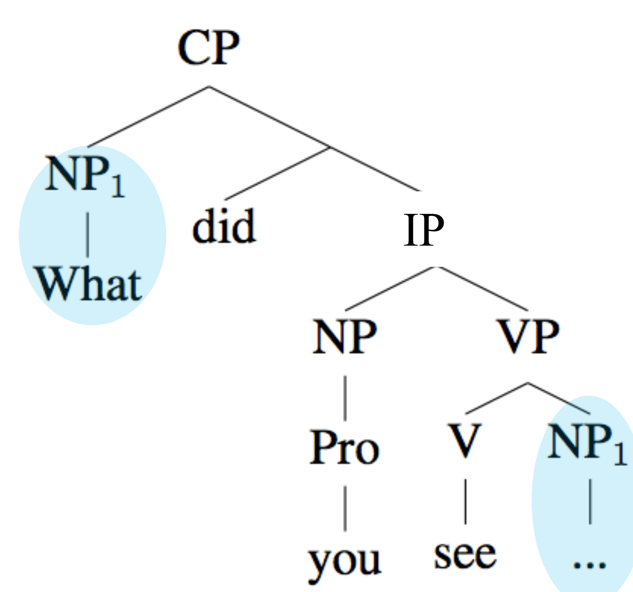
View *wh*-dependencies in terms of their building blocks and track those building blocks in the input.

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, in prep.

Both theories

Dependencies represented as a sequence of container nodes

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

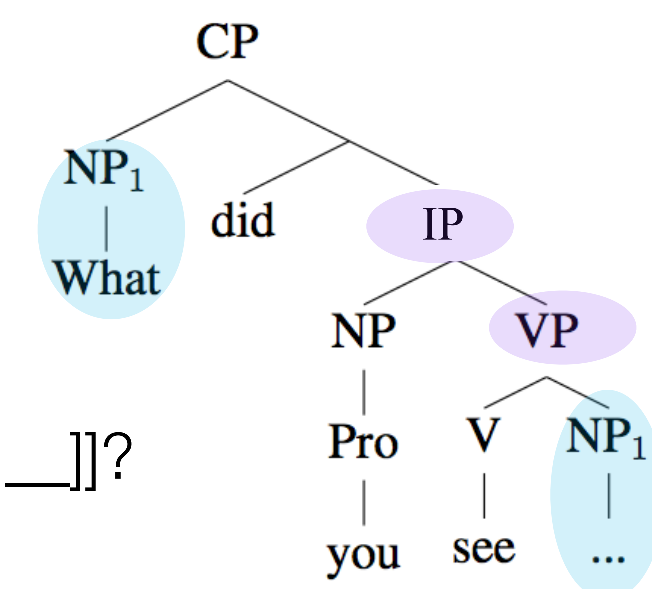


Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, in prep.

Both theories

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
Dickson, Pearl, & Futrell 2022, in prep.

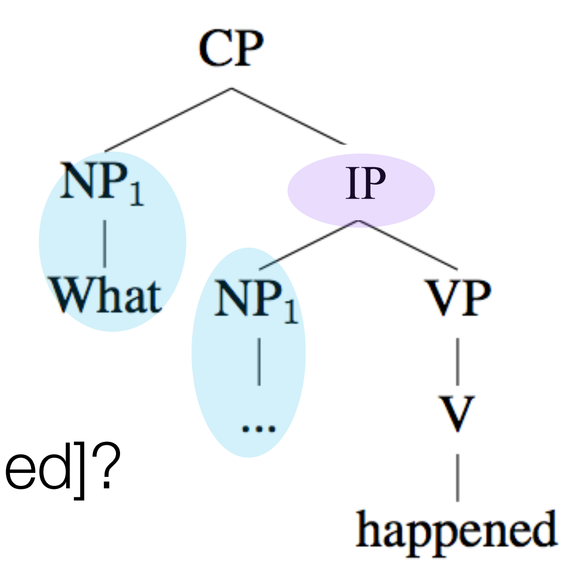
Both theories

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*

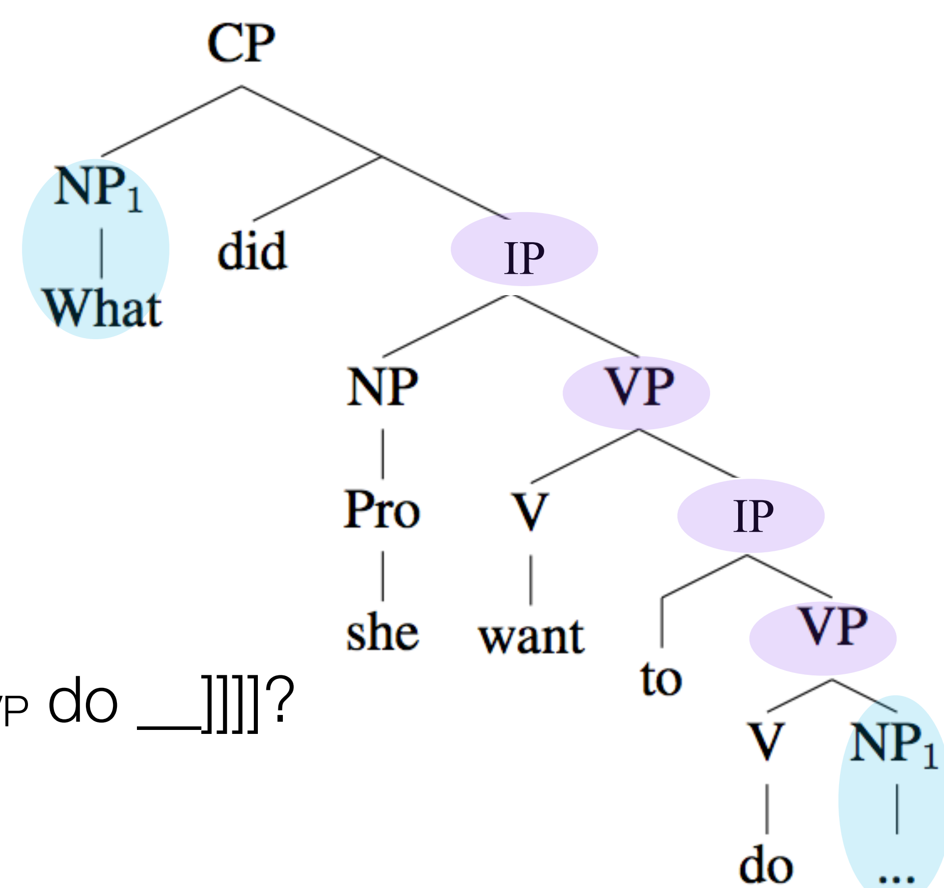


Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, in prep.

Both theories

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
Dickson, Pearl, & Futrell 2022, in prep.

Both theories

What did you see __?
= What did [$_{\text{IP}}$ you [$_{\text{VP}}$ see __]]?
= *start*-IP-VP-*end*

What __ happened?
= What [$_{\text{IP}}$ __ happened]?
= *start*-IP-*end*

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

(Much) less acceptable dependencies have low probability segments

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

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

Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, in prep.

Both theories

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* [$_{CP\text{-}that}$ [$_{IP}$ [$_{NP}$ *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
Dickson, Pearl, & Futrell 2022, in prep.

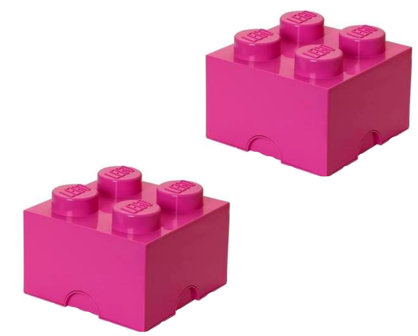
1: probabilities of pre-specified pieces

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

1: probabilities of pre-specified pieces

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

1: probabilities of pre-specified pieces

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

1: probabilities of pre-specified pieces

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

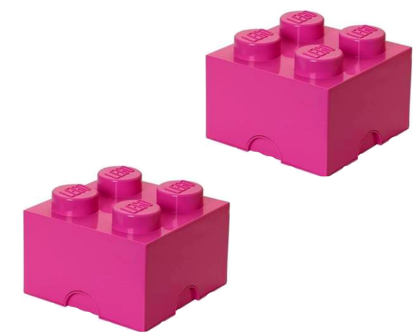
1: probabilities of pre-specified pieces

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

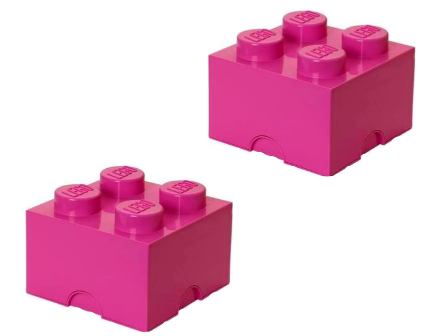
1: probabilities of pre-specified pieces

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

1: probabilities of pre-specified pieces

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

NP-PP-*end*

VP-CP$_{that}$-IP

Pearl & Sprouse 2013

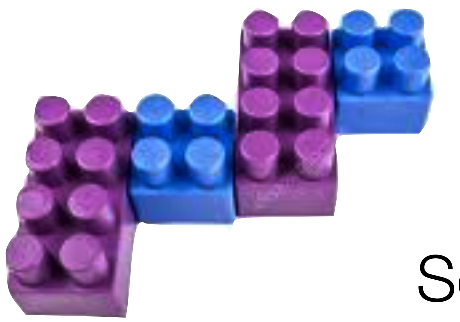
1: probabilities of pre-specified pieces

start-IP-VP-*end*

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

IP-VP-IP

IP-VP-PP

VP-PP-*end*

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

IP-NP-PP

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

VP-IP-VP

NP-PP-*end*

Pearl & Sprouse 2013

1: probabilities of pre-specified pieces

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

start-IP-VP

start-IP-*end*

IP-VP-*end*

IP-VP-CP$_{that}$

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-CP$_{that}$-IP

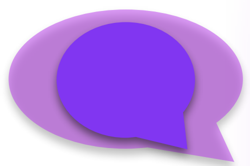
VP-IP-VP

NP-PP-*end*

CP$_{that}$-IP-NP

IP-NP-PP

Pearl & Sprouse 2013



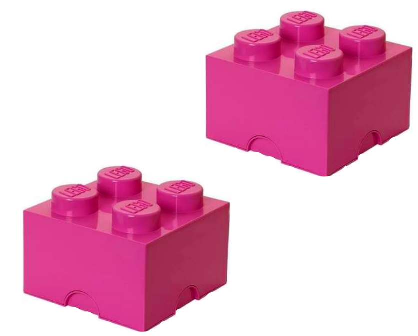
1: probabilities of pre-specified pieces

start-IP-VP-*end*

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

IP-VP-IP

IP-VP-PP

VP-PP-*end*

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

IP-NP-PP

VP-IP-VP

NP-PP-*end*

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

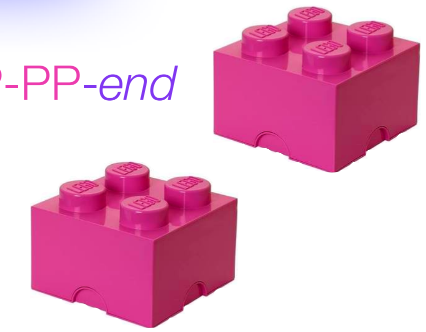
Pearl & Sprouse 2013

1: probabilities of pre-specified pieces

start-IP-*end*

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

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



start-IP-VP-*end*

start-IP-VP

IP-VP-*end*

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

start-IP-*end*

IP-VP-CP$_{that}$

IP-VP-IP

VP-CP$_{that}$-IP

IP-VP-PP

VP-IP-VP

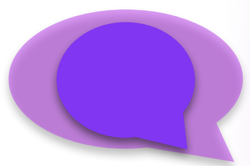
VP-PP-*end*

NP-PP-*end*

CP$_{that}$-IP-NP

IP-NP-PP

Pearl & Sprouse 2013

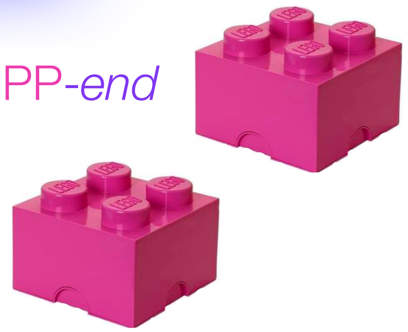


1: probabilities of pre-specified pieces

start-IP-*end*

start-IP-VP-*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

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

IP-VP-*end*

start-IP-*end*

IP-VP-CP_{that}

IP-VP-PP

VP-PP-*end*

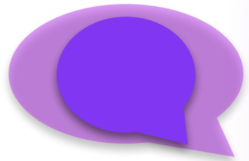
CP_{that}-IP-NP

NP-PP-*end*

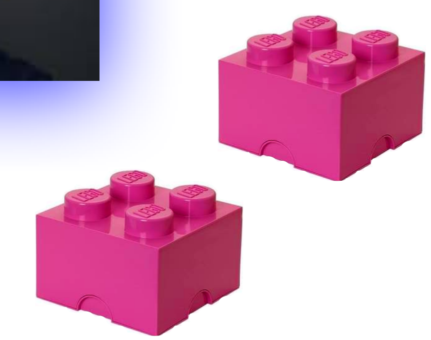
IP-NP-PP

VP-CP_{that}-IP

Pearl & Sprouse 2013



1: probabilities of pre-specified pieces



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*

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

start-IP-*end*

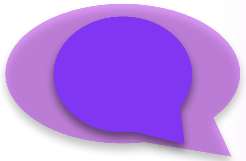
IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

Pearl & Sprouse 2013



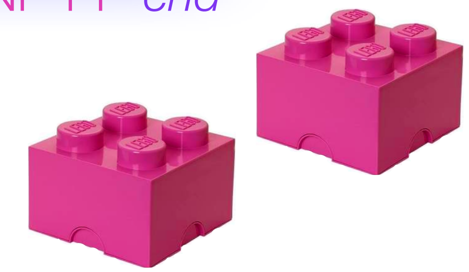
1: probabilities of pre-specified pieces

start-IP-*end*

start-IP-VP-*end*

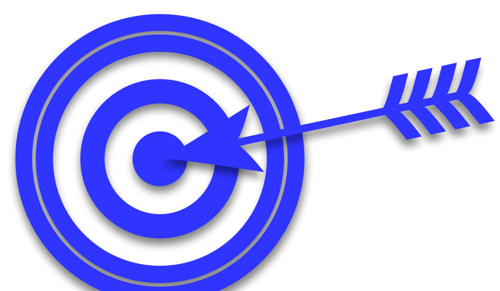
start-IP-VP-IP-VP-*end*

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

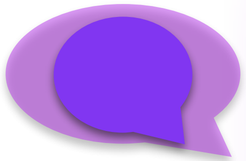


$$\prod_{t \in trigrams} p(t)$$

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



Pearl & Sprouse 2013



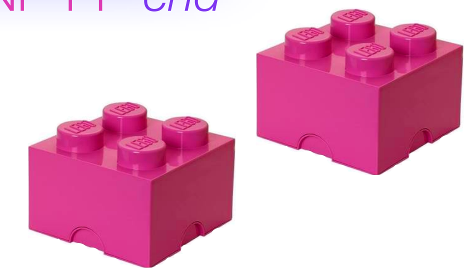

1: probabilities of
pre-specified pieces


start-IP-*end*

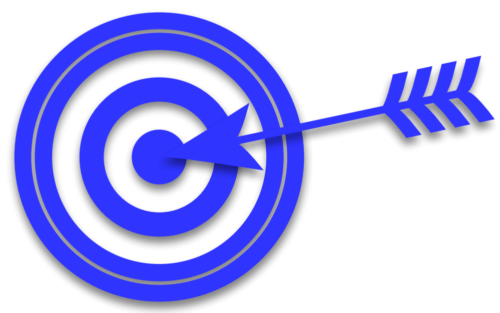
start-IP-VP-*end*

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

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



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

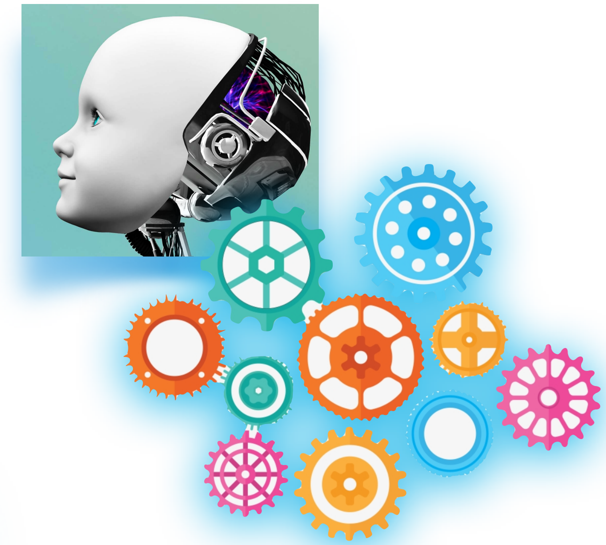


Lower probability dependencies are predicted to be less acceptable (dispreferred), compared to higher probability dependencies.

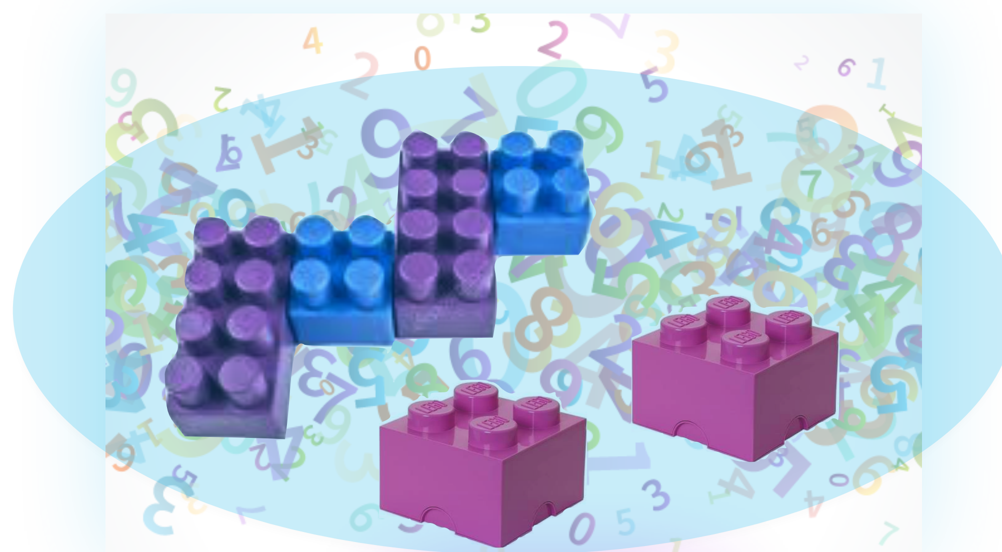
Pearl & Sprouse 2013

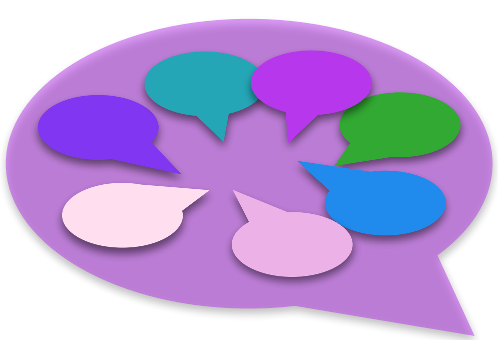
Both theories

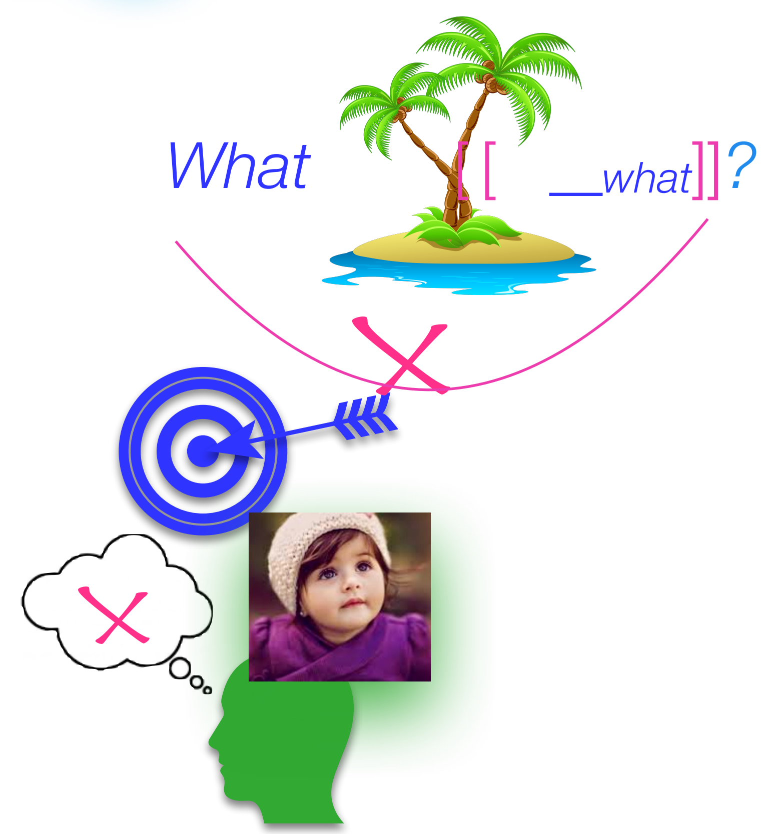
Evaluating the theory



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



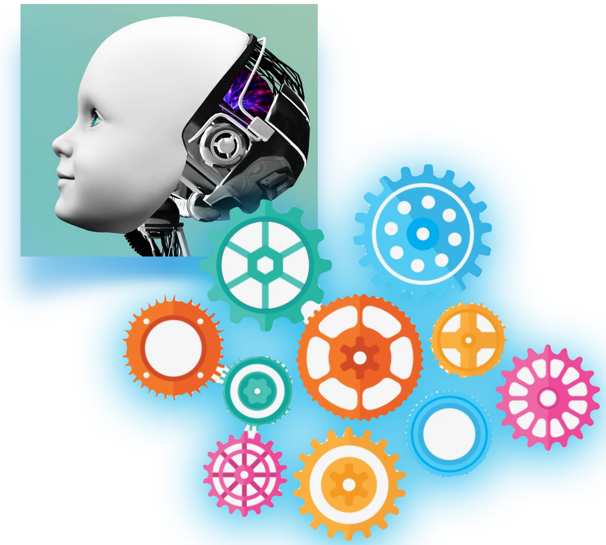


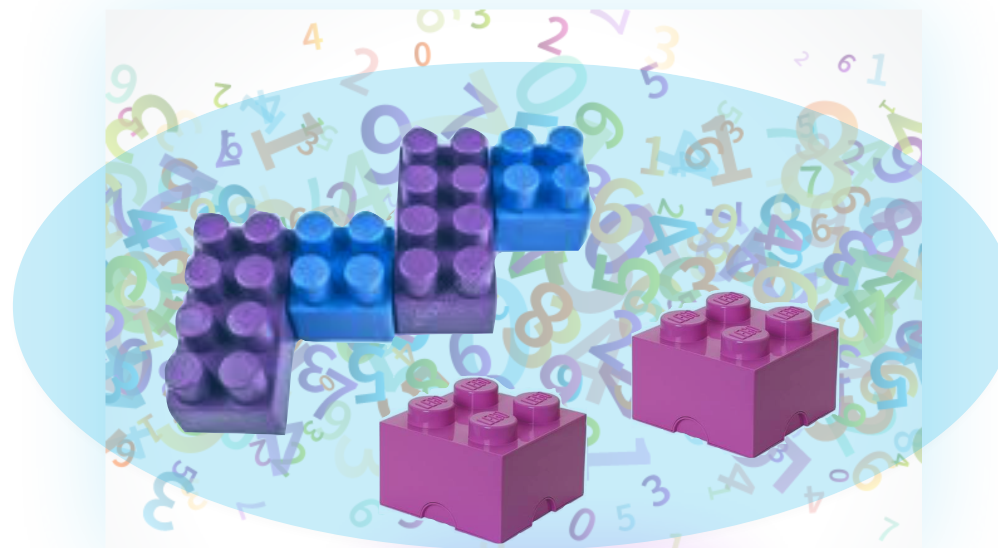


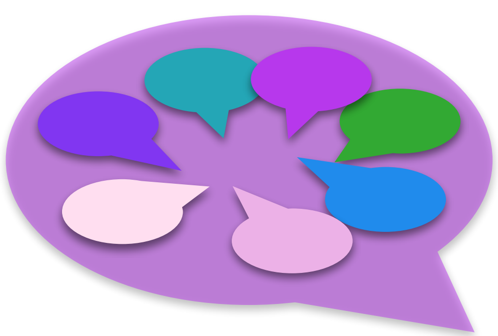
Pearl & Sprouse 2013

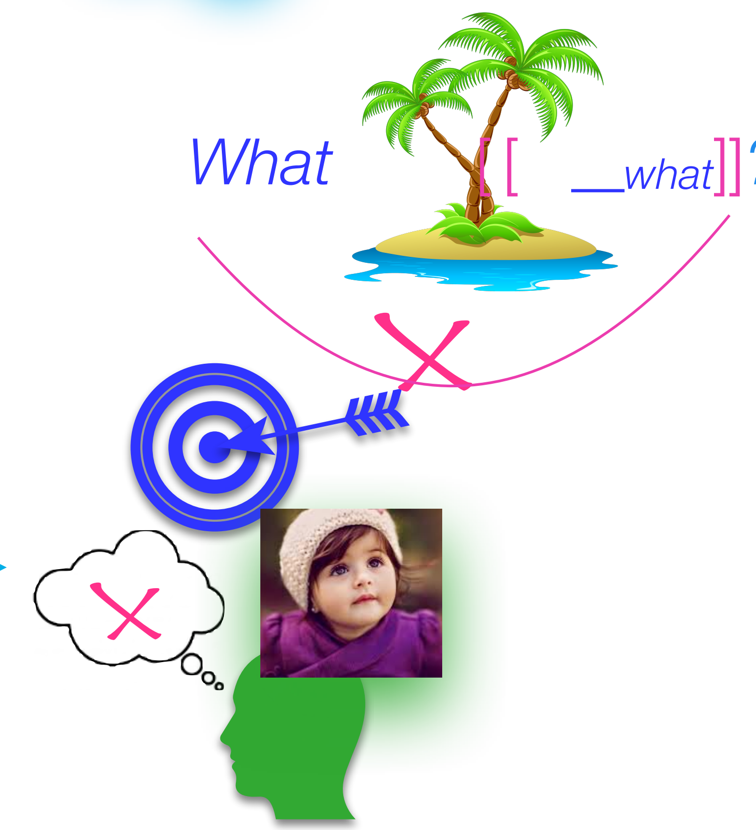
Both theories

Evaluating the theory







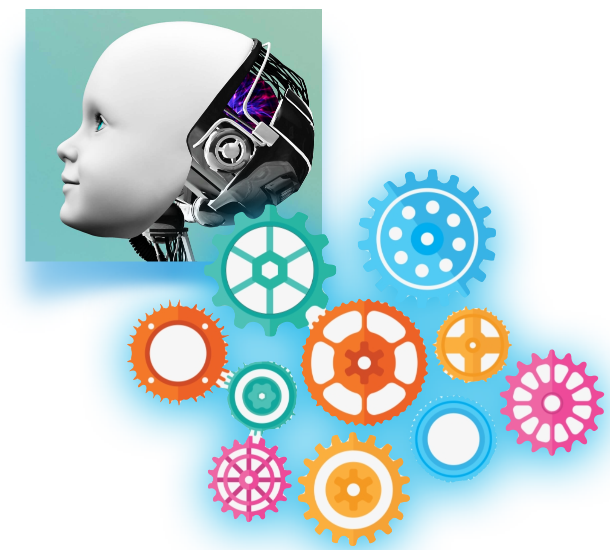


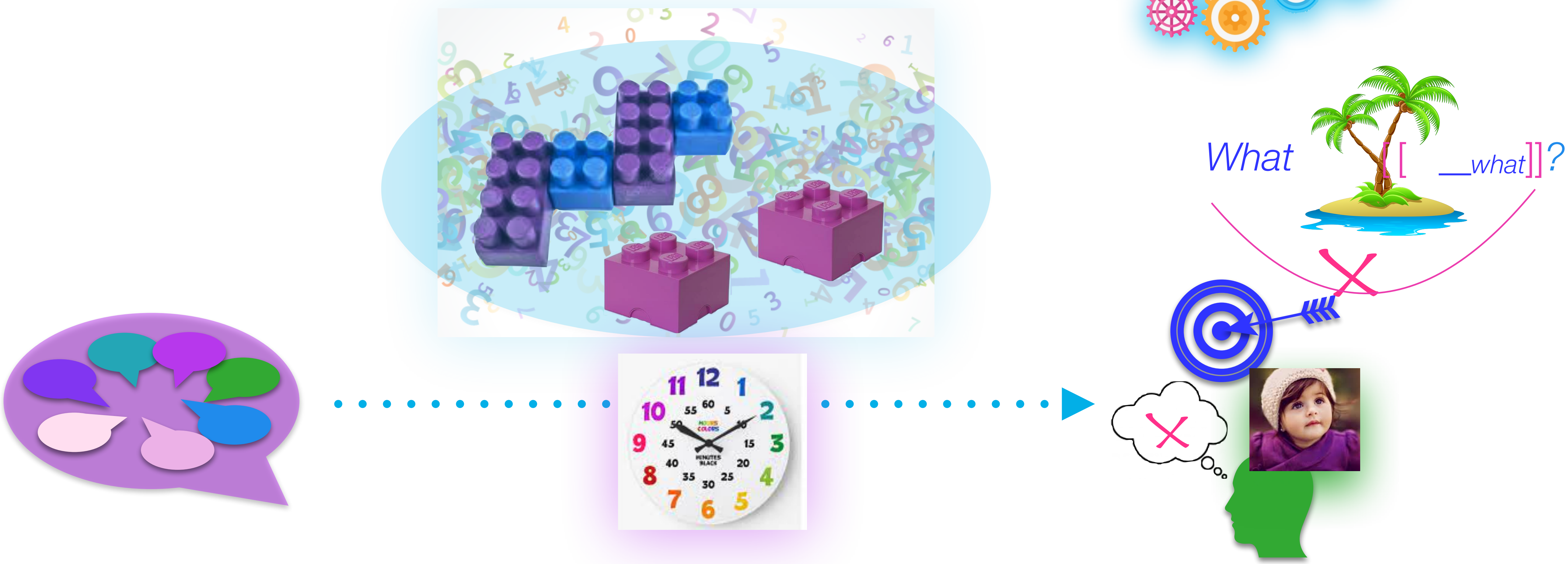
What's the input look like?

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

Both theories

Evaluating the theory



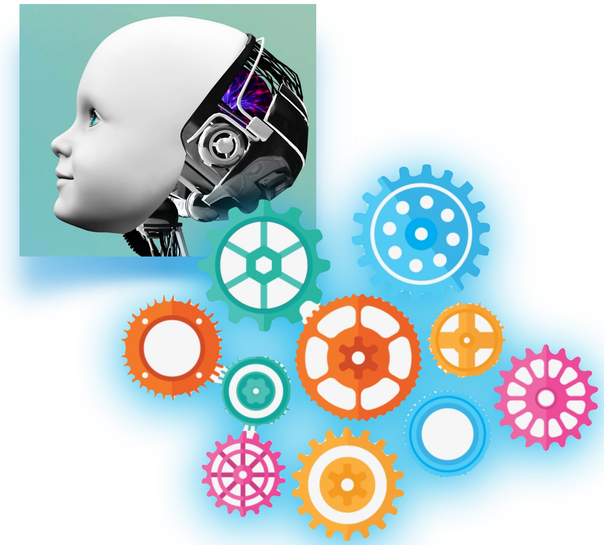


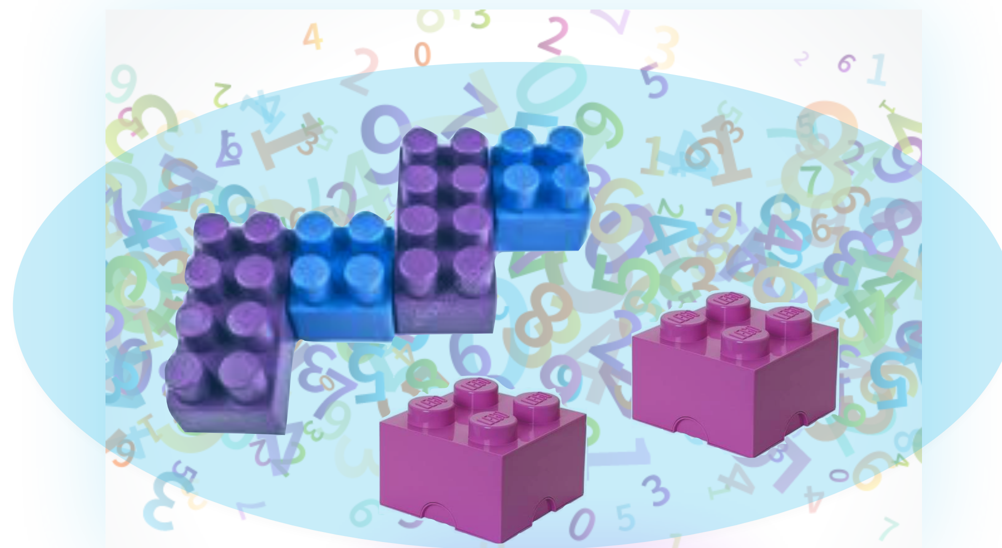
102K utterances (≈21K *wh*-dependencies) from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 children between the ages of 1 and 5 years old.

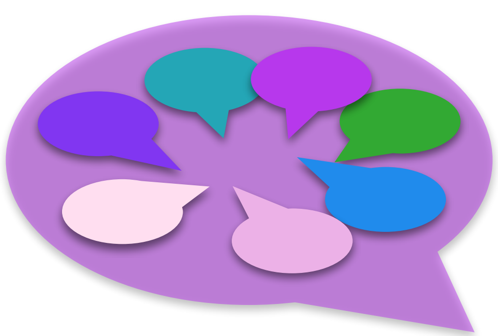
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

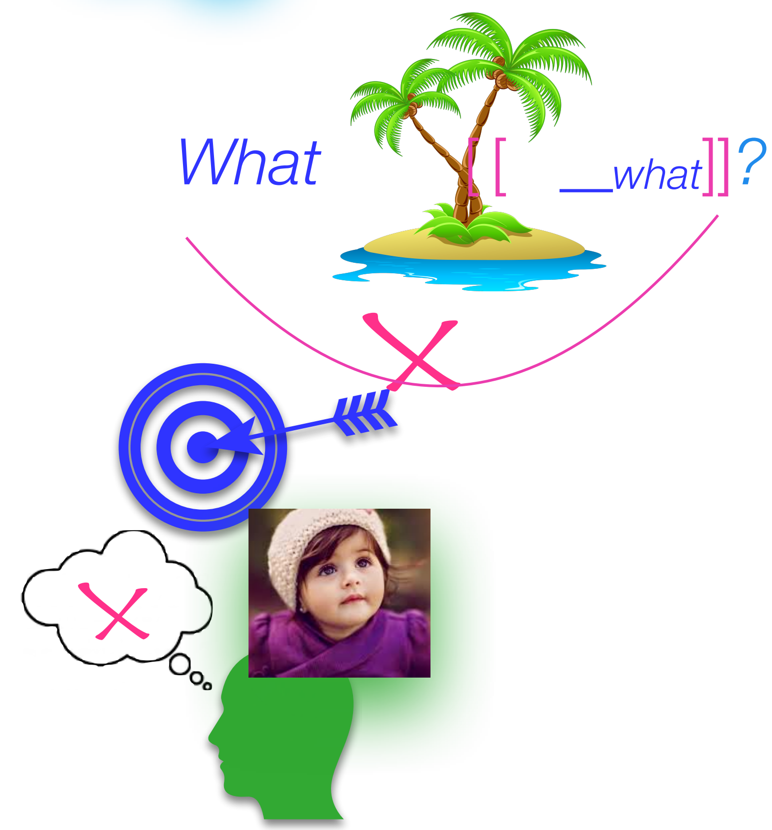
Both theories

Evaluating the theory

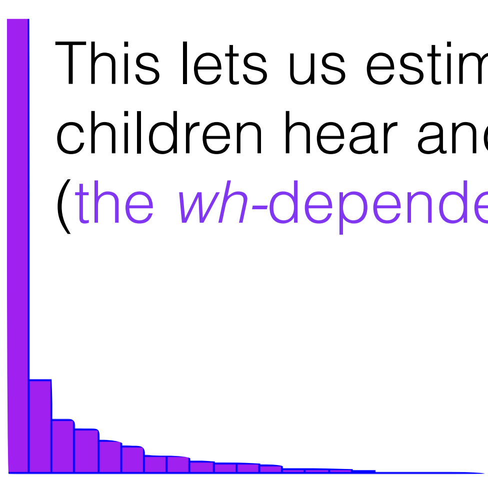








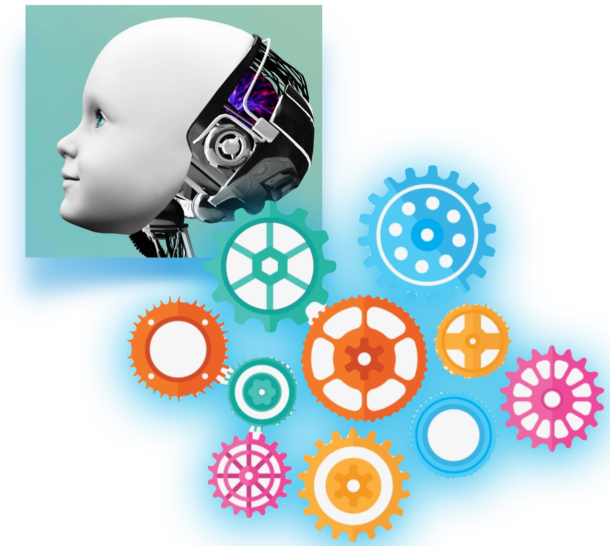
This lets us estimate which *wh*-dependencies children hear and how often they hear them (the *wh*-dependency distribution).

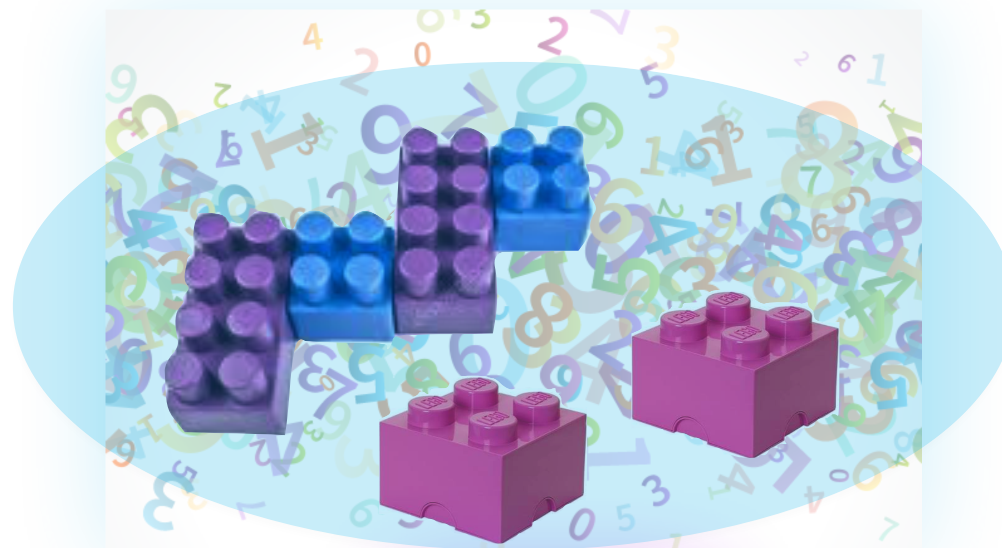


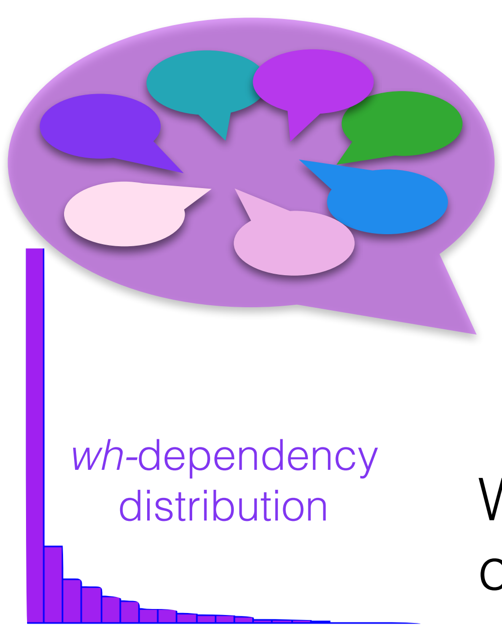
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

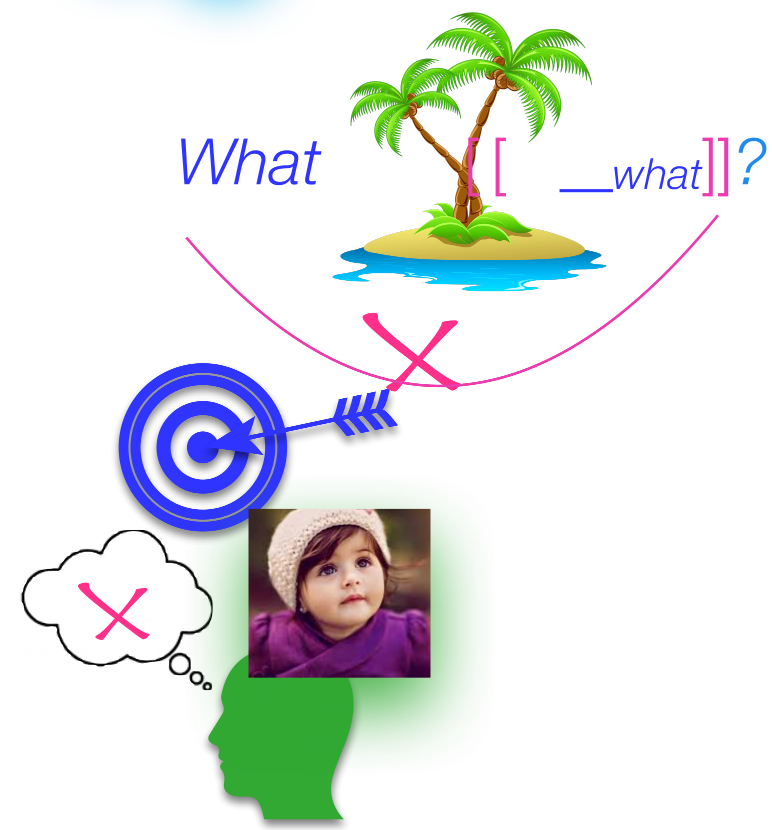
Both theories

Evaluating the theory









#

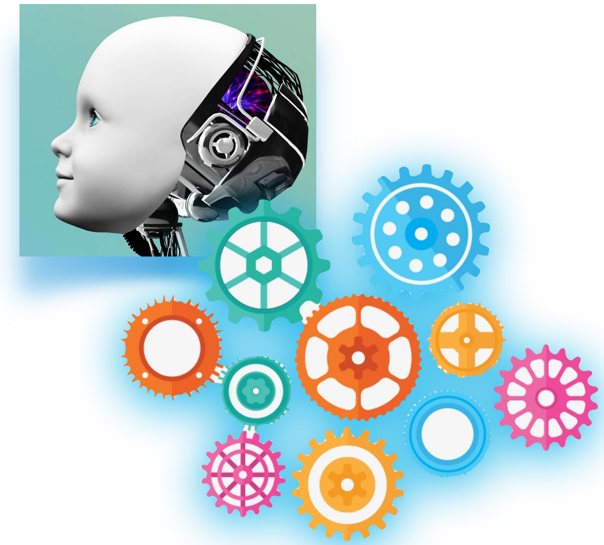
We can then estimate how many *wh*-dependencies children hear during the learning period.

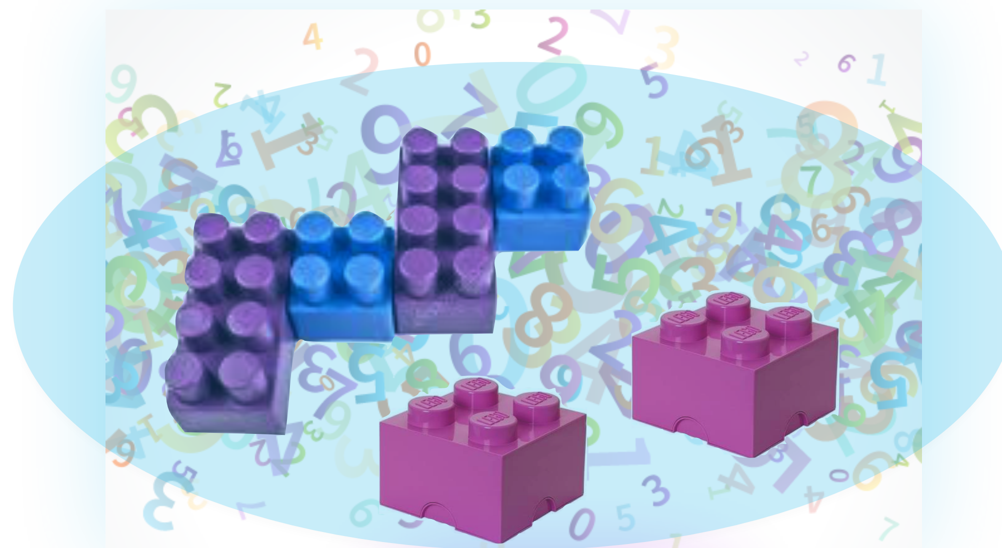
(<60 months)

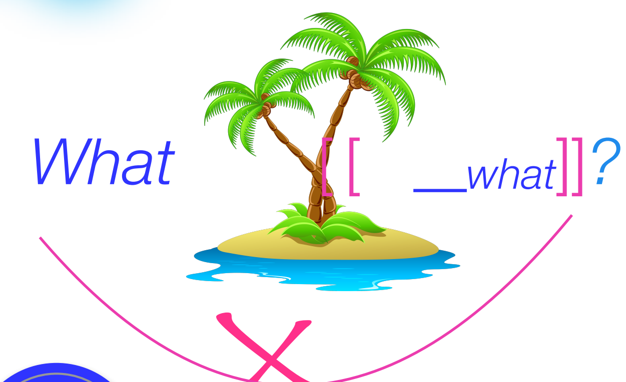
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

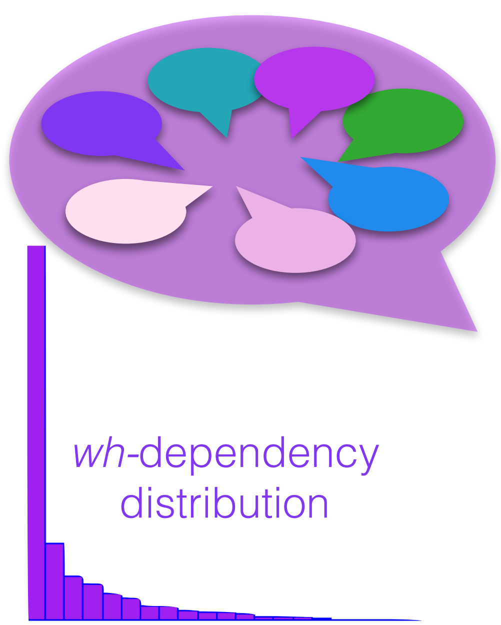
Both theories

Evaluating the theory









#

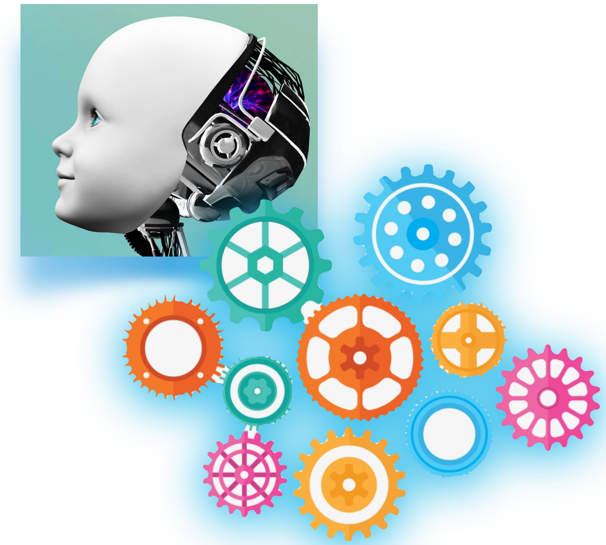
(<60 months)

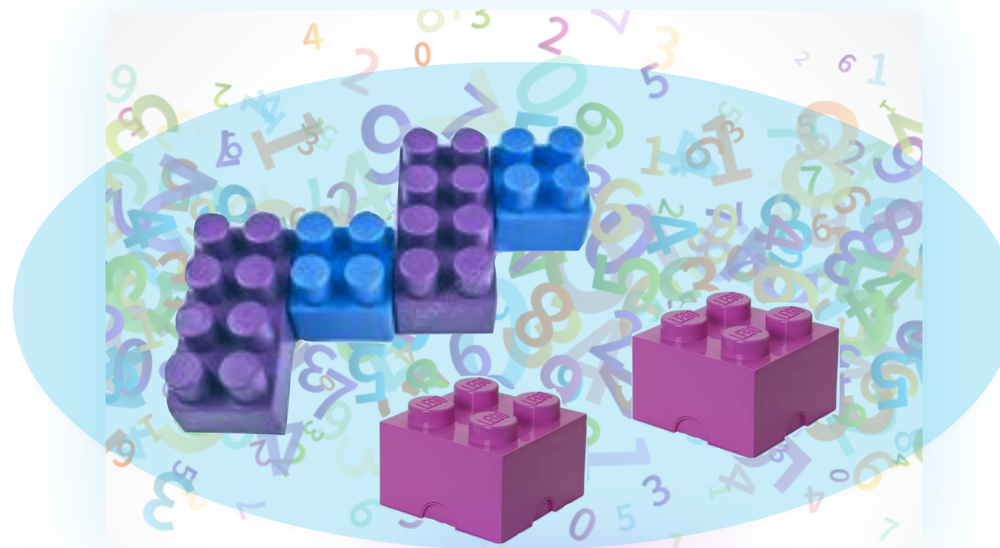
Children begin to represent the full structure of *wh*-dependencies (e.g., *wh*-questions and relative clauses) around 20 months: Seidl et al. 2003, Gagliardi et al. 2016, Perkins & Lidz 2020.

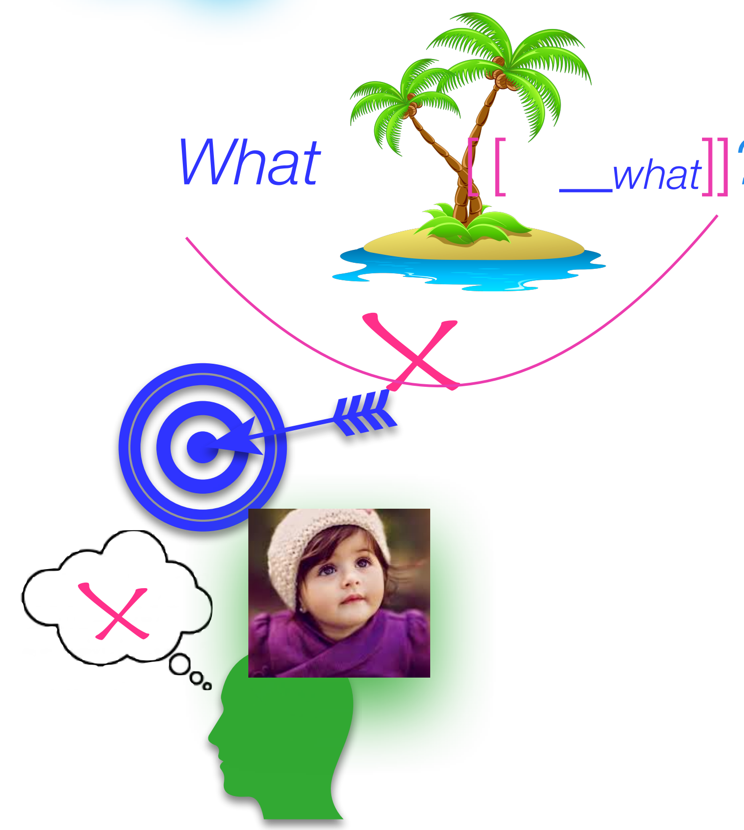
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

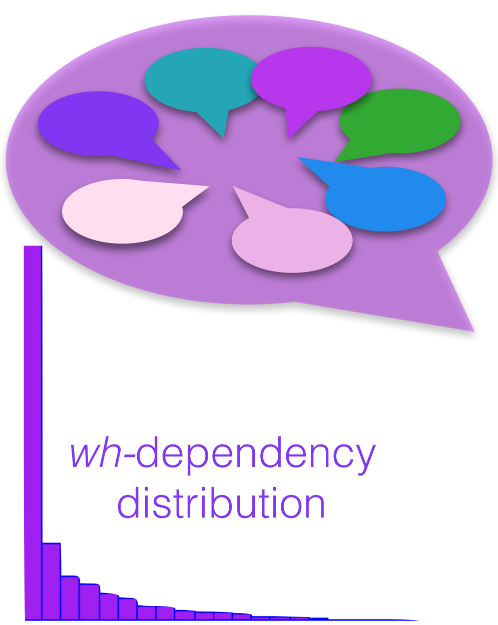
Both theories

Evaluating the theory



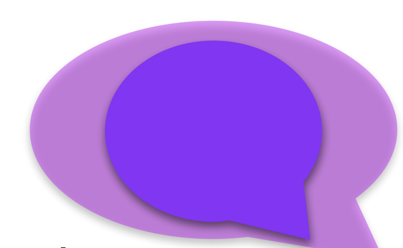






\#

(20 months ≤ age < 60 months)

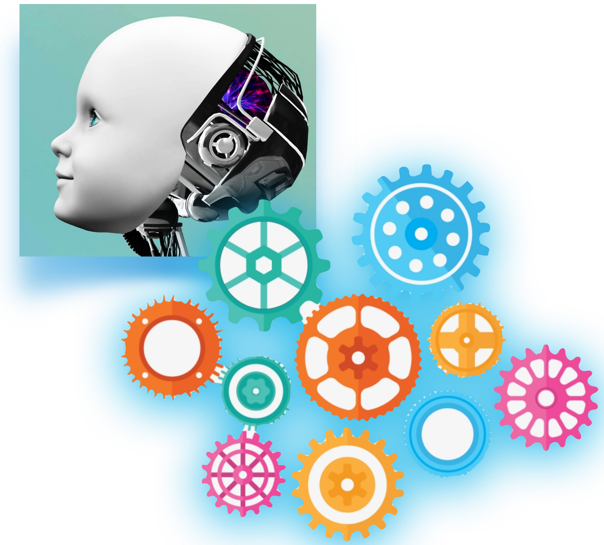


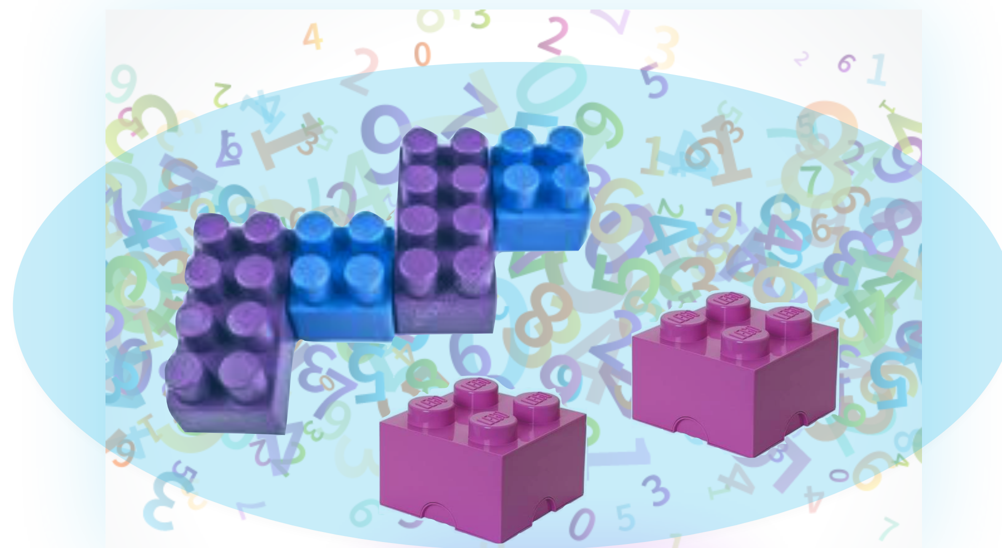
Educated guess: This is when children can start processing *wh*-dependencies reliably from their input.

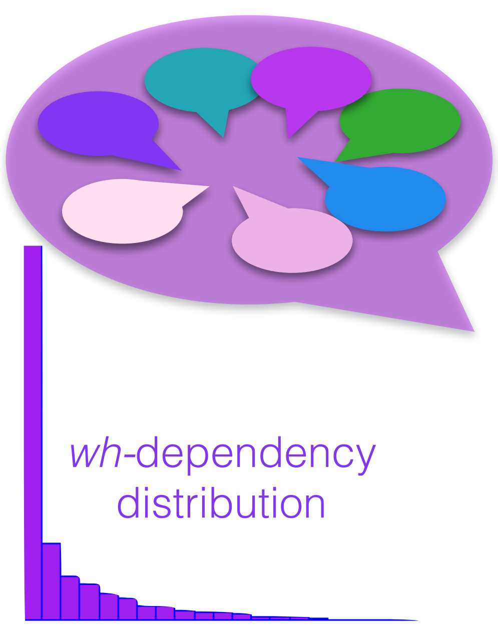
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

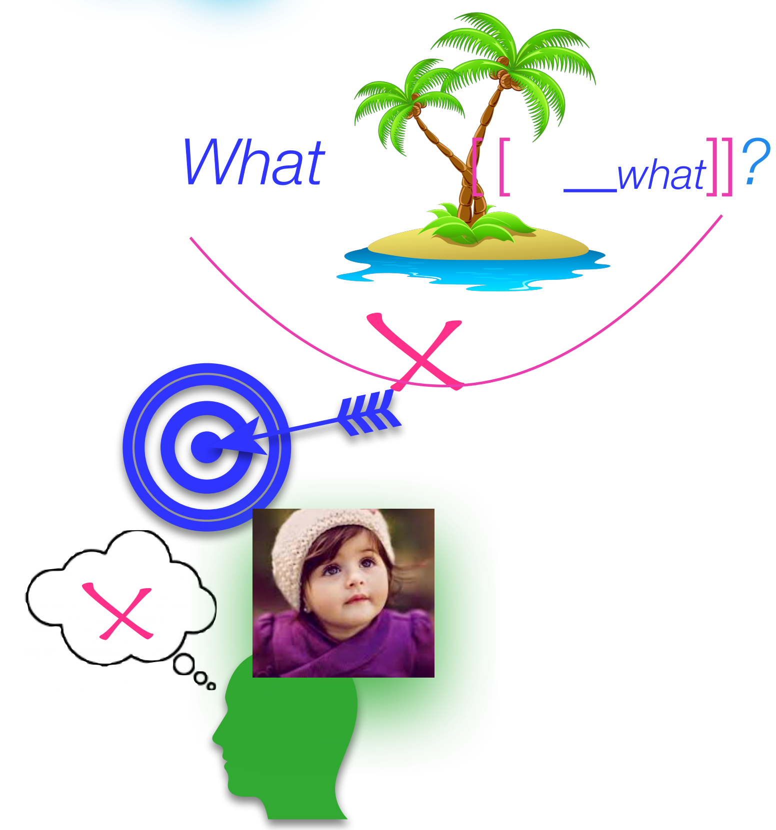
Both theories

Evaluating the theory









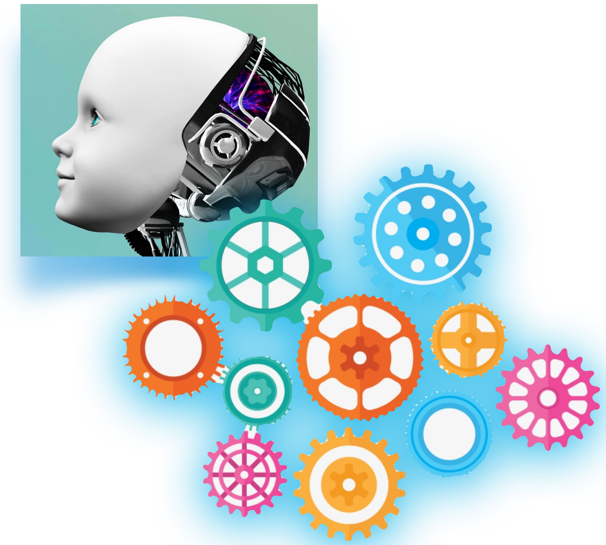
(20 months ≤ age < 60 months)

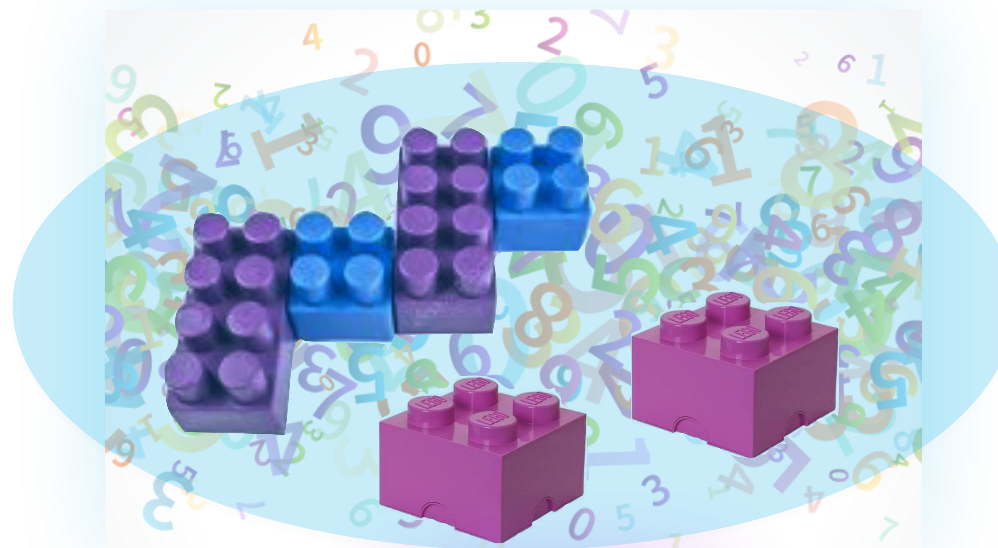
How many minutes is this? In particular, children are awake for only a certain portion of the day at different ages (Davis et al. 2004).

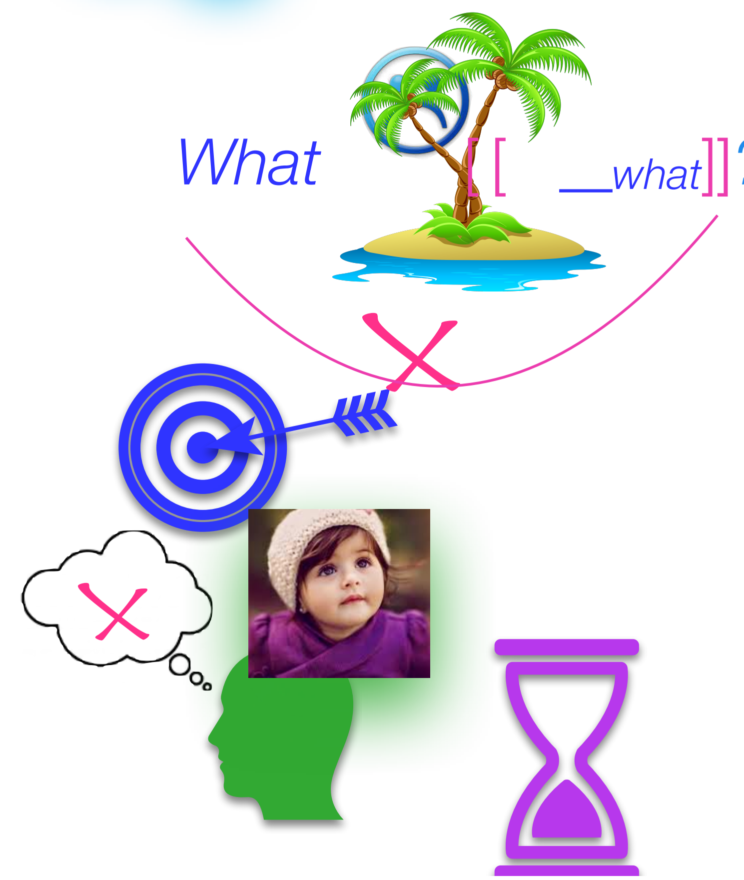
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

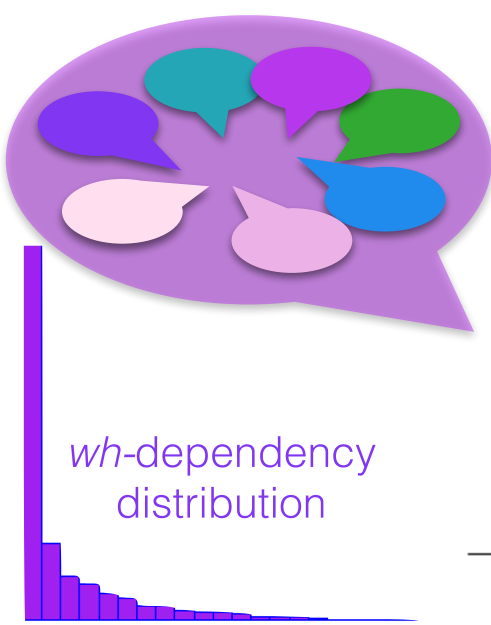
Both theories

Evaluating the theory









#

(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

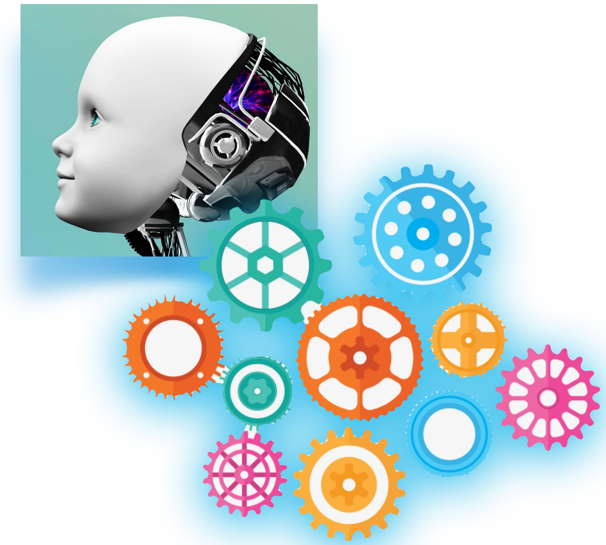
14174.17 * 60 min/hour

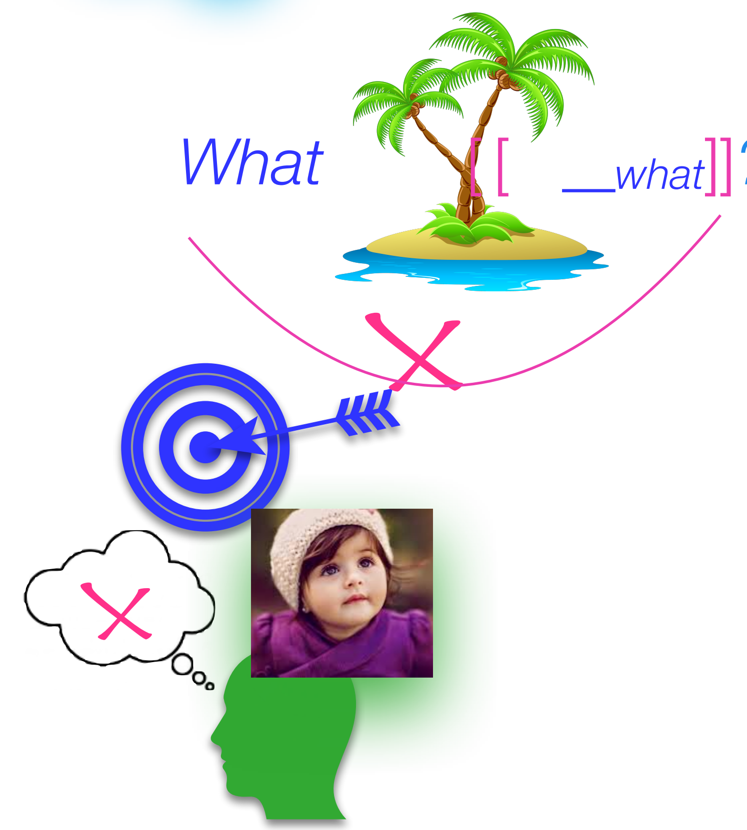
How many minutes 850450.2

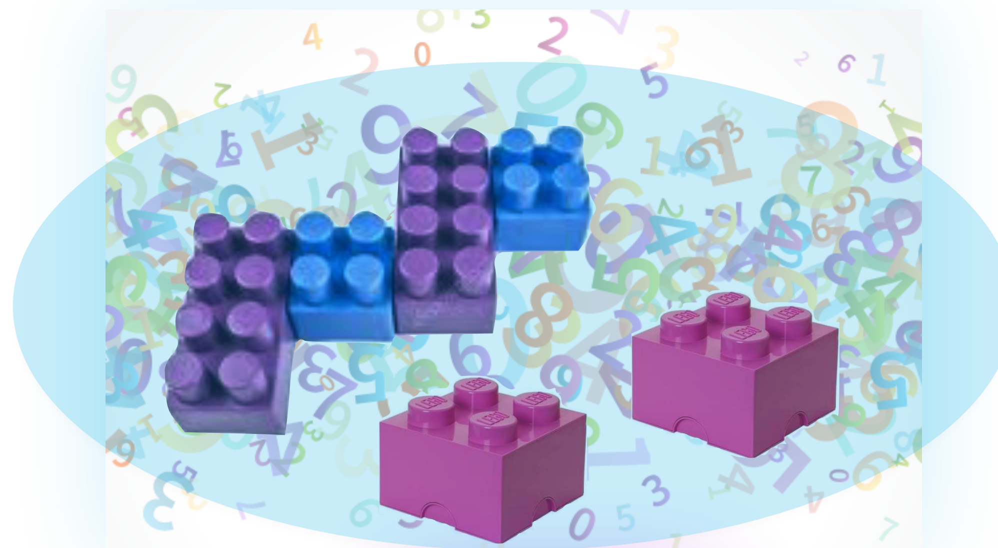
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

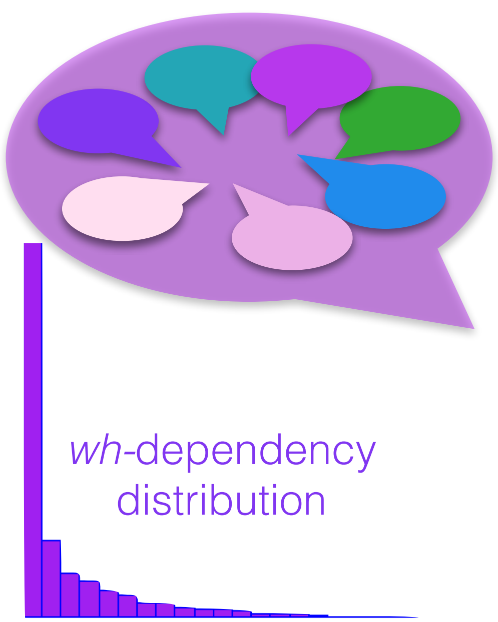
Both theories

Evaluating the theory

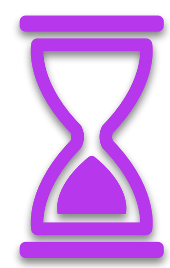








(≈850450 minutes)

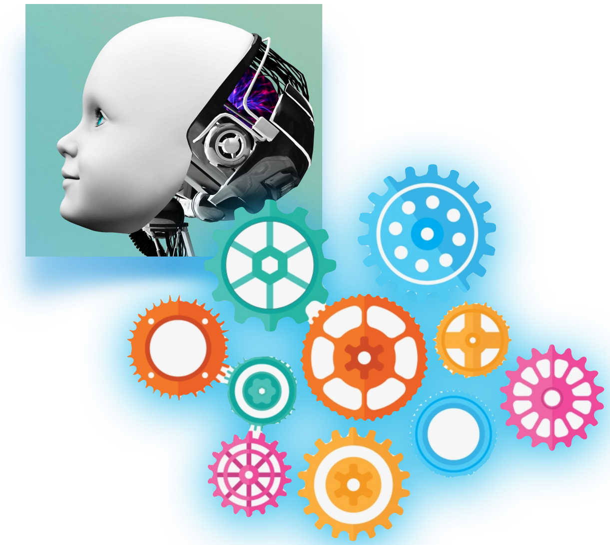


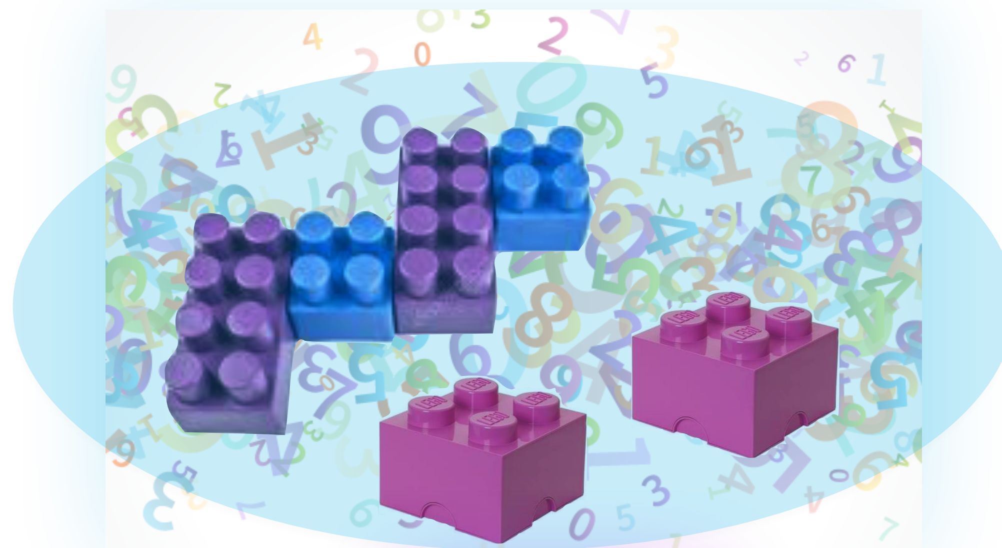
How many *wh*-dependencies is this? #

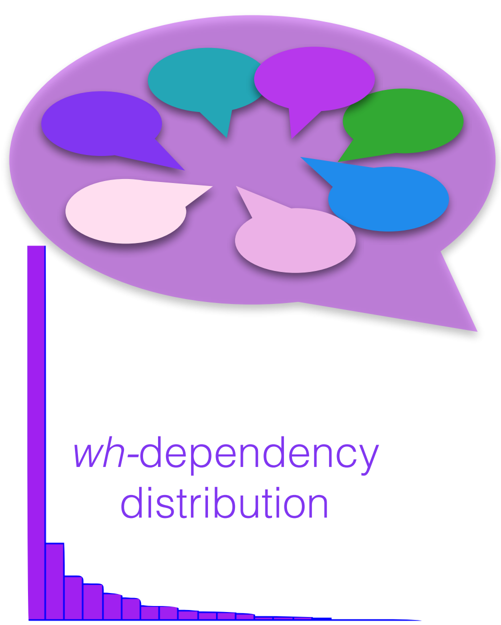
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Both theories

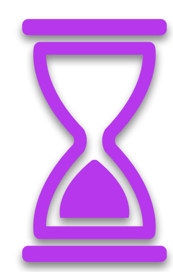
Evaluating the theory







(≈850450 minutes)



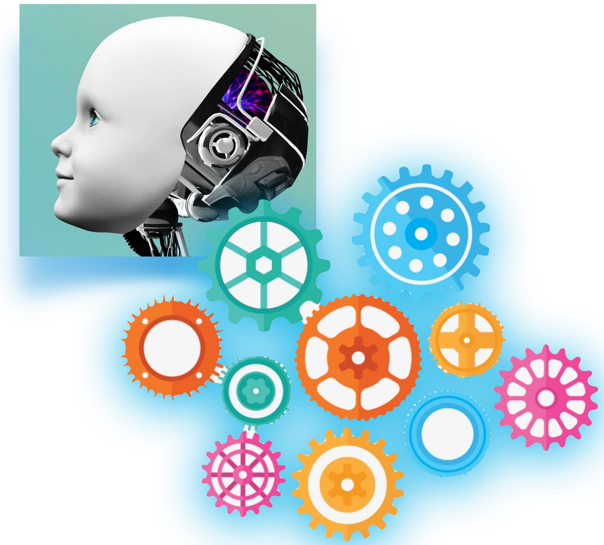
#

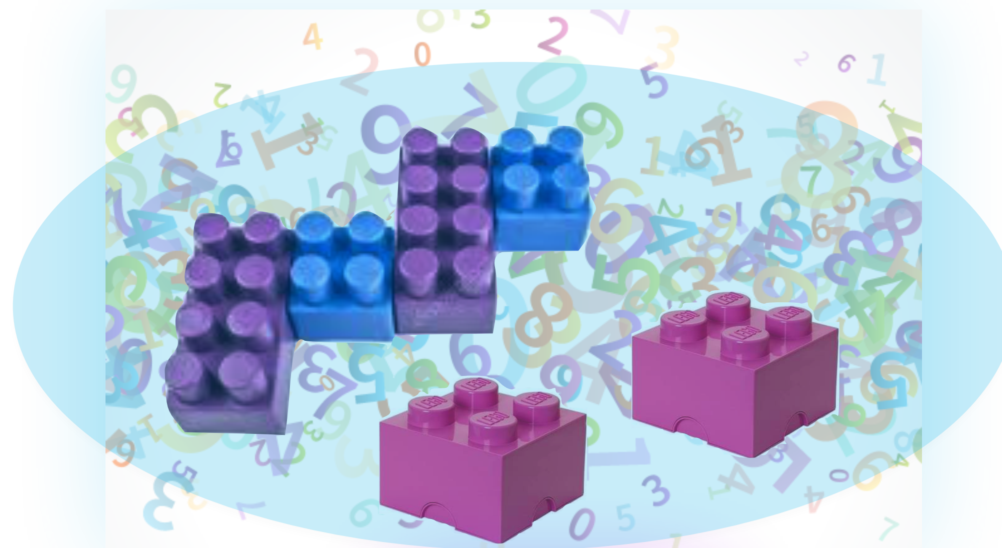
Hoff-Ginsberg (1998) and Rowe (2012):
Estimates of utterances per minute in speech directed at children from different backgrounds.

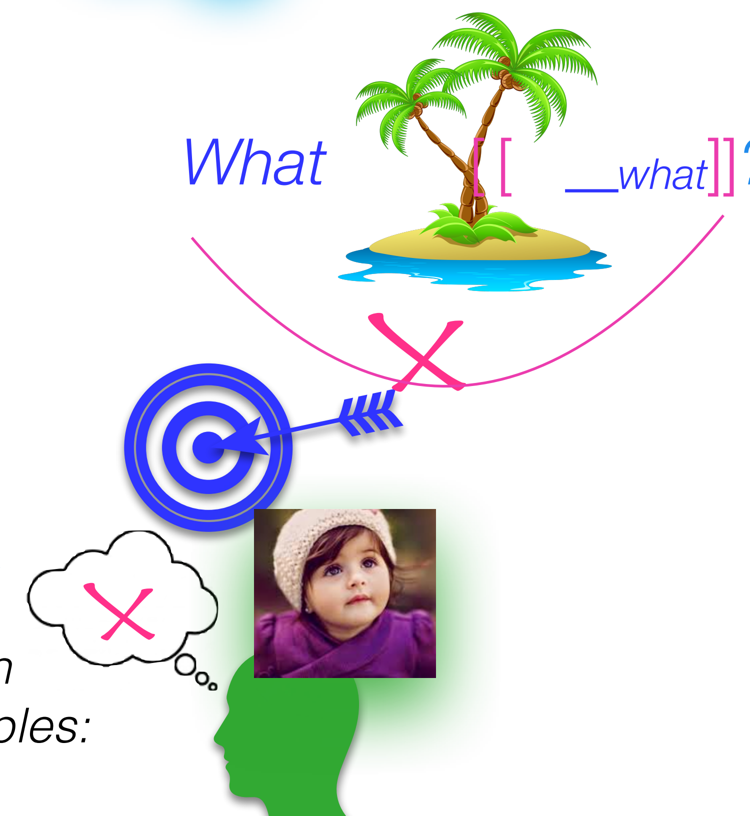
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

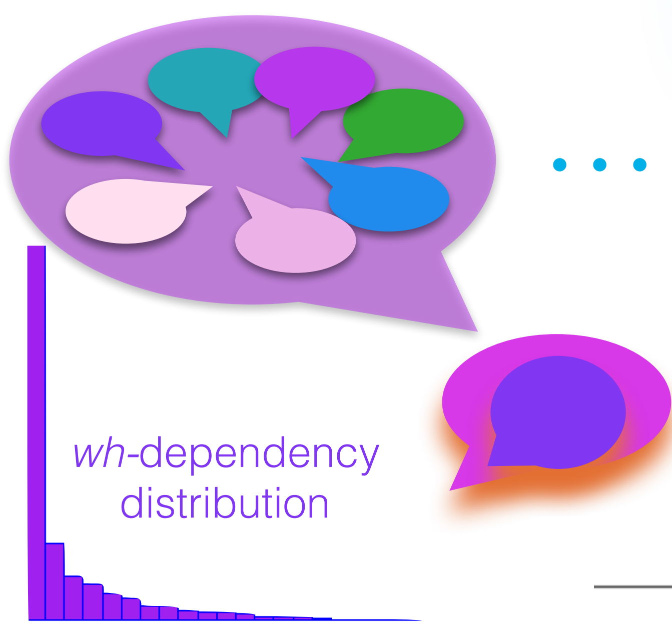
Both theories

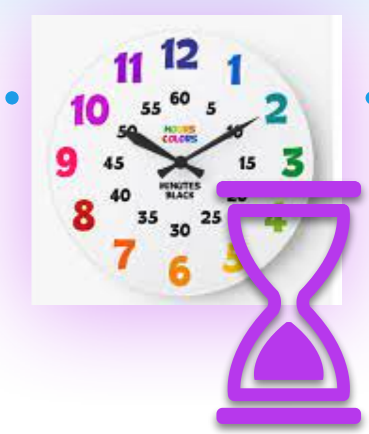
Evaluating the theory











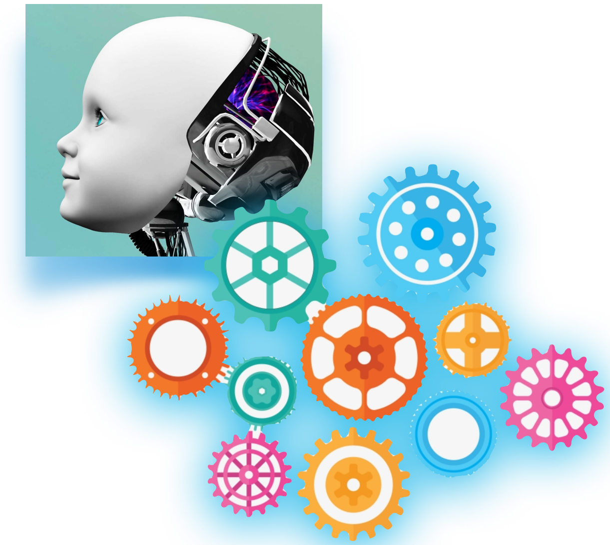
from our own corpus samples: rate of wh-dependencies/ utterance

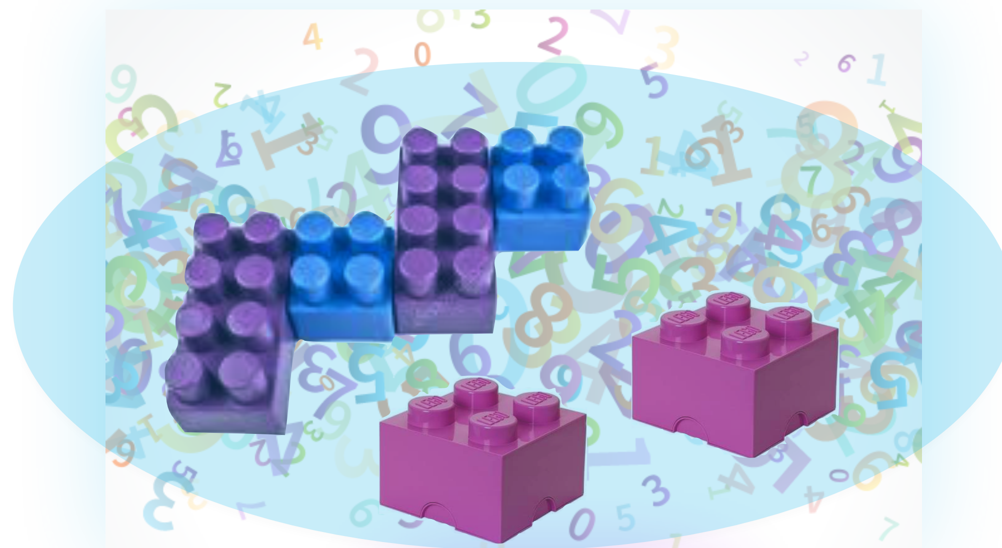
	utt/min	*	min	*	***wh*-dep/utt**	=	**total *wh*-dep**
		*	850,450.2	*	20,932/101,838		
- 2 s.d.	7.4					=	1,293,545
- 1 s.d.	11.6					=	2,027,719
average	15.8		#			=	2,761,893
+ 1 s.d.	20.0					=	3,496,067
+ 2 s.d.	24.2					=	4,230,241

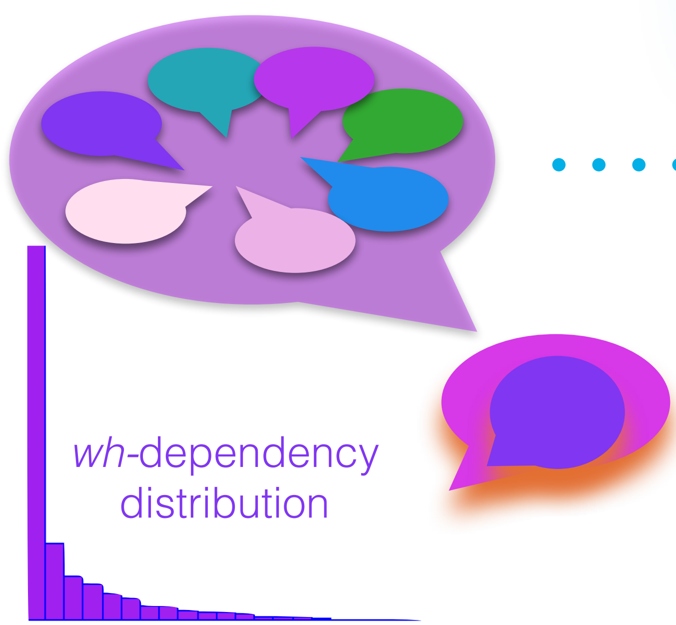
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

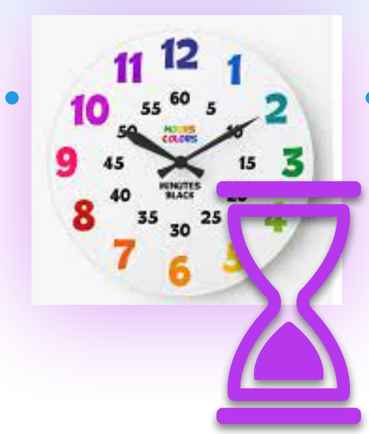
Both theories

Evaluating the theory









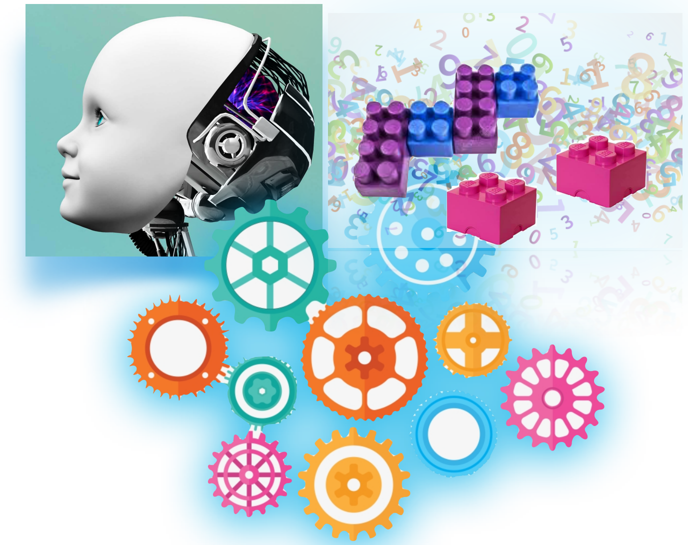
#
≈1.3 million - 4.2 million
wh-dependencies

Can the modeled learner produce the appropriate observable behavior?

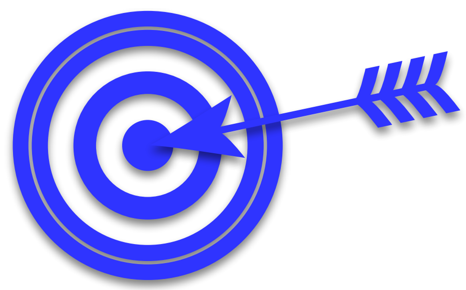
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Both theories

Evaluating the theory



Reminder: Target behavior

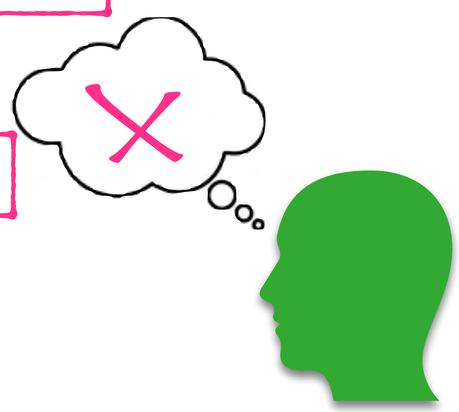


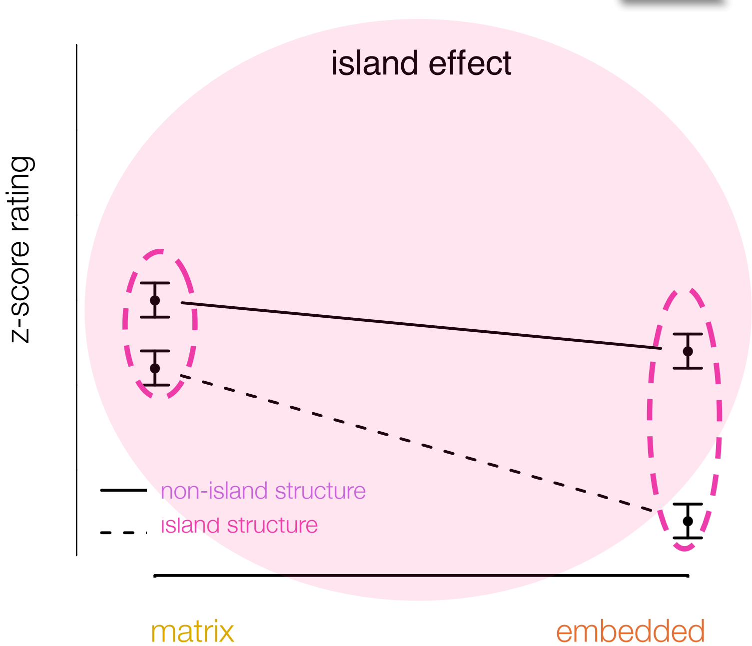
Subject island

Complex NP island

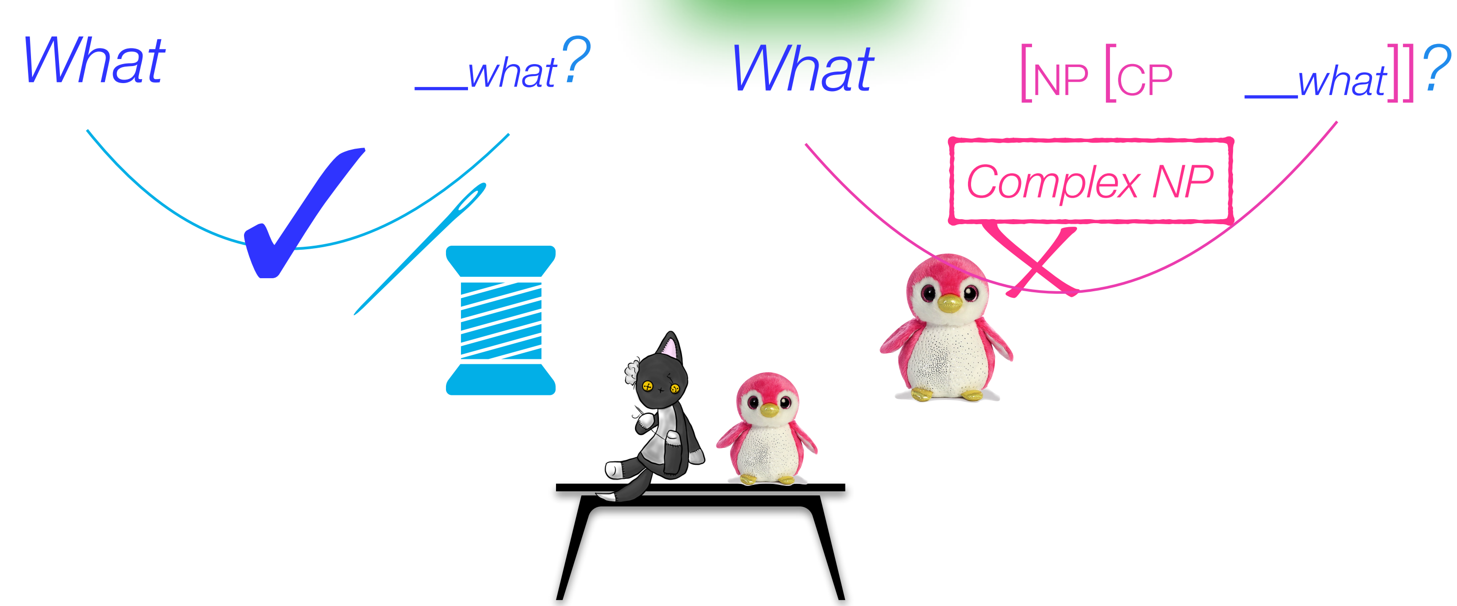
Whether island

Adjunct island





Sprouse et al. 2012

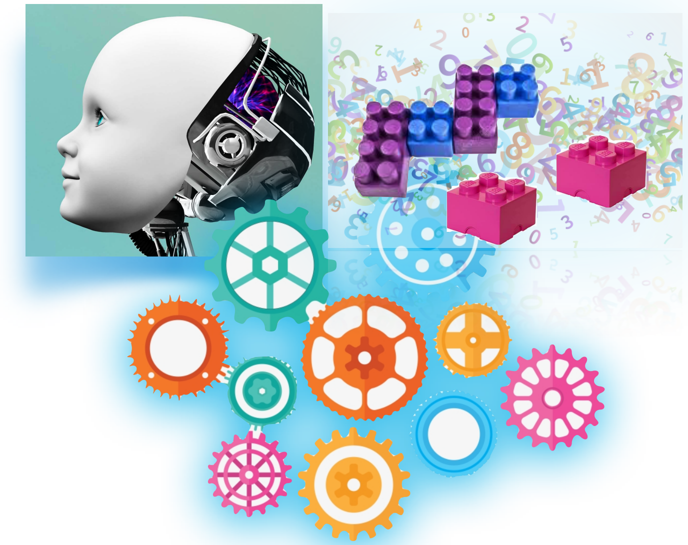


De Villiers et al. 2008

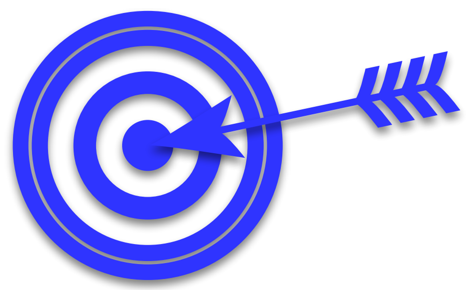
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Both theories

Evaluating the theory



Reminder: Target behavior

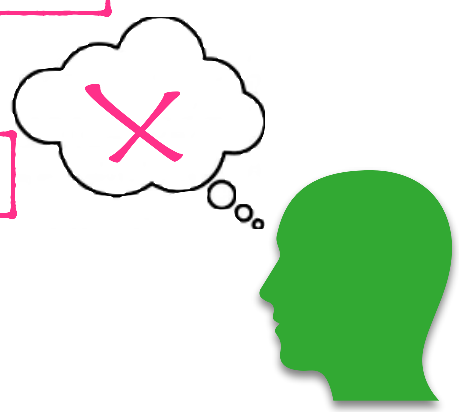


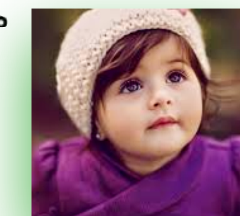
Subject island

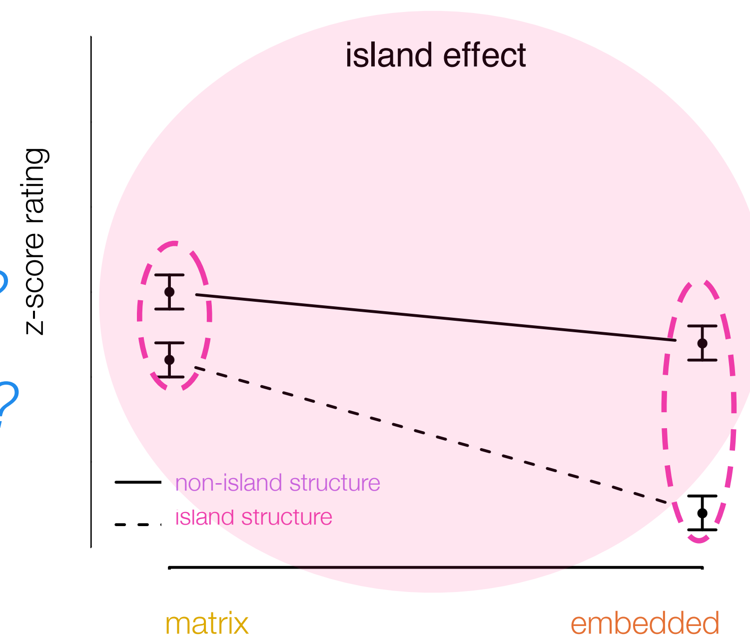
Complex NP island

Whether island

Adjunct island







Who [non-island]?

Who [island]?

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

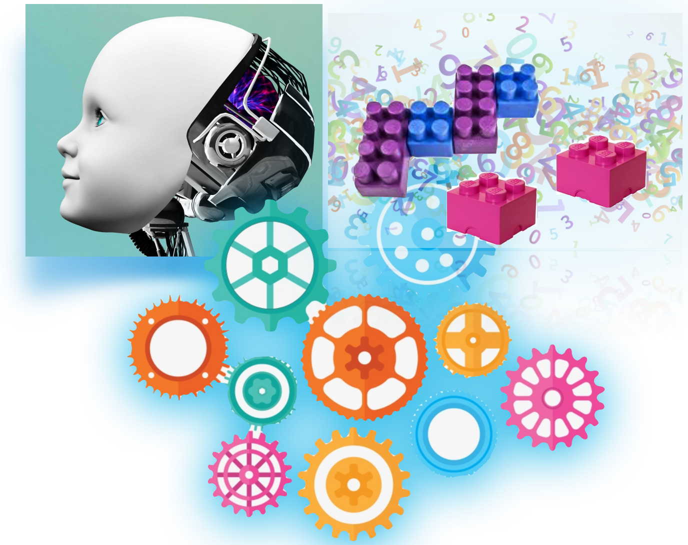
Who __*who*?

Who [CP… __*who*]?

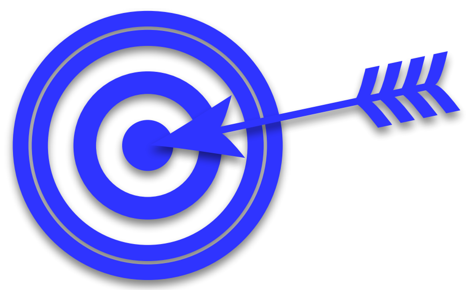
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

Both theories

Evaluating the theory



Reminder: Target behavior

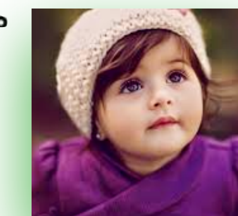


Subject island

Complex NP island

Whether island

Adjunct island



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

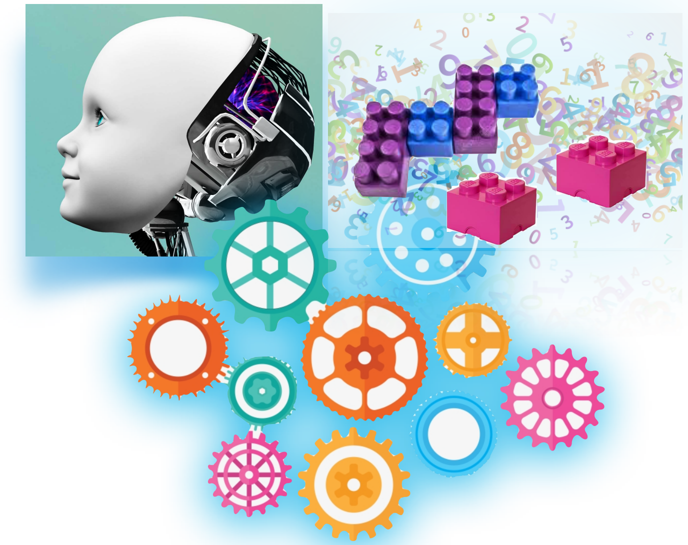
Complex NP island stimuli

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

Both theories

Evaluating the theory



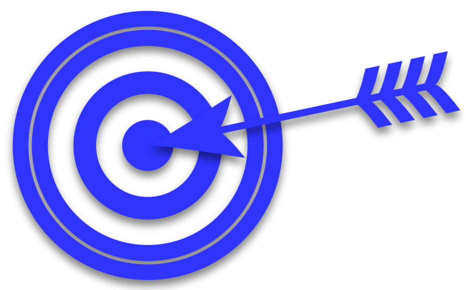
Reminder: Target behavior

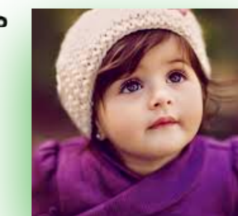
Subject island

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

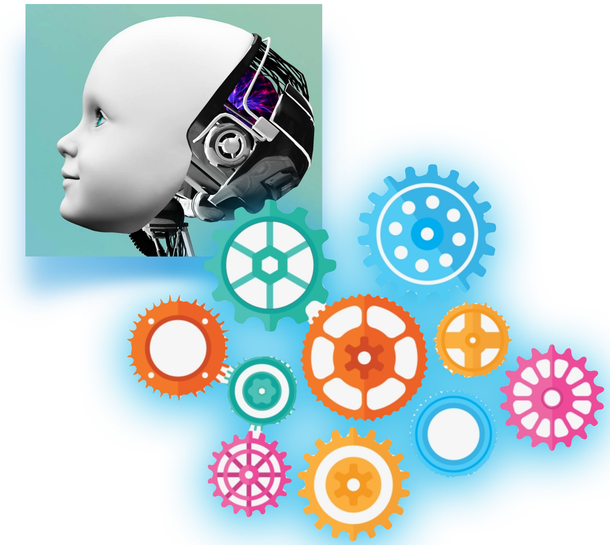
Complex NP island stimuli

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

Both theories

Evaluating the theory



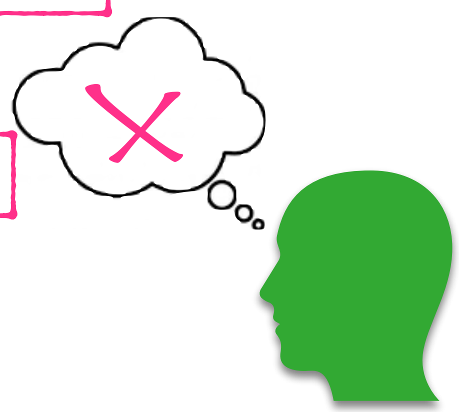
Reminder: Target behavior

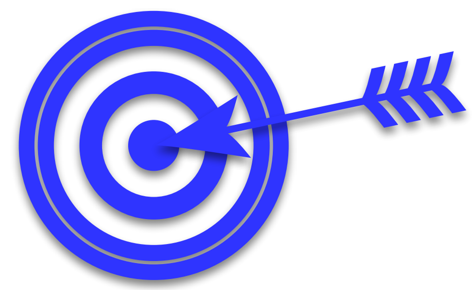
Subject island

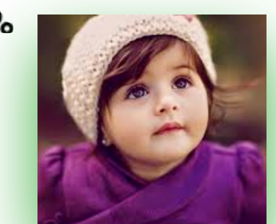
Complex NP island

Whether island

Adjunct island

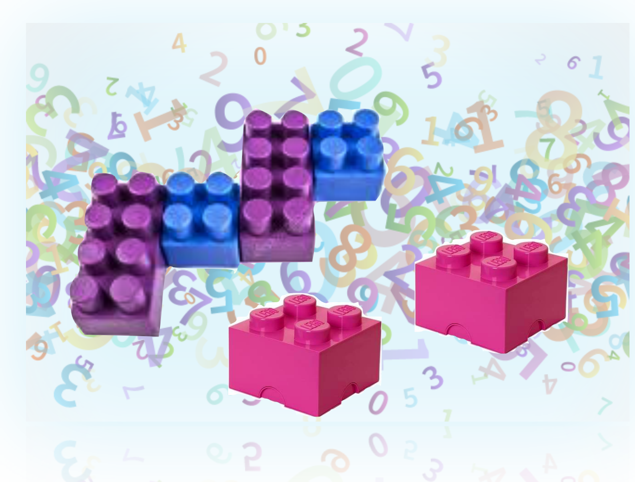






Each *wh*-dependency from the island stimuli of Sprouse et al. 2012
- can be transformed into container node sequences
- can be broken into syntactic trigram building blocks and have its probability calculated

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



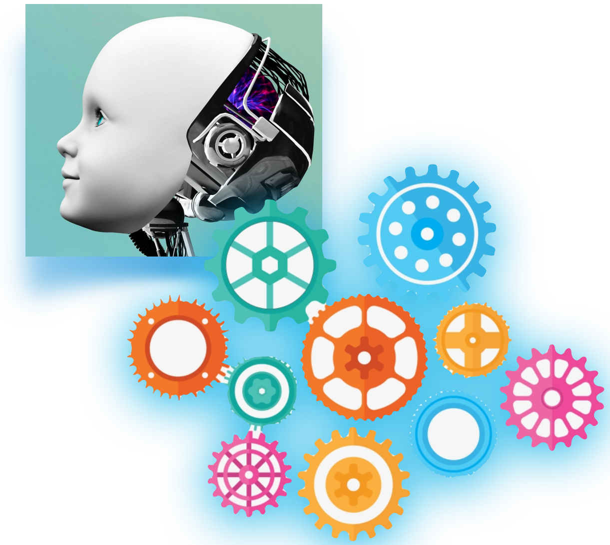
Complex NP island stimuli

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

Both theories

Evaluating the theory



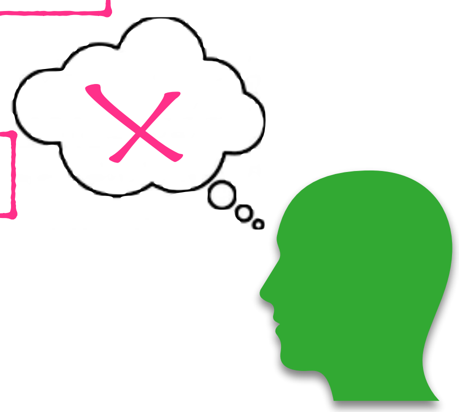
Reminder: Target behavior

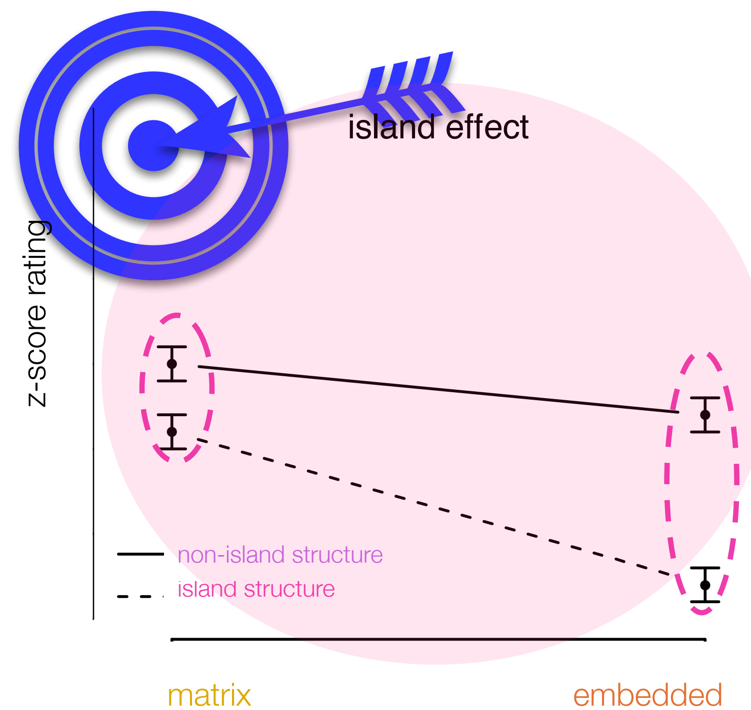
Subject island

Complex NP island

Whether island

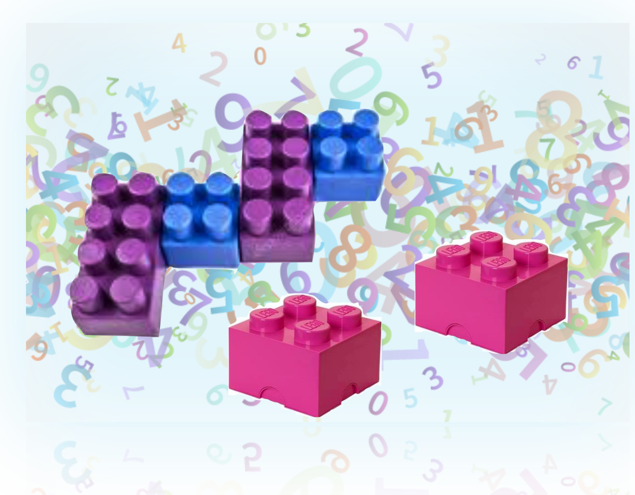
Adjunct island





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

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



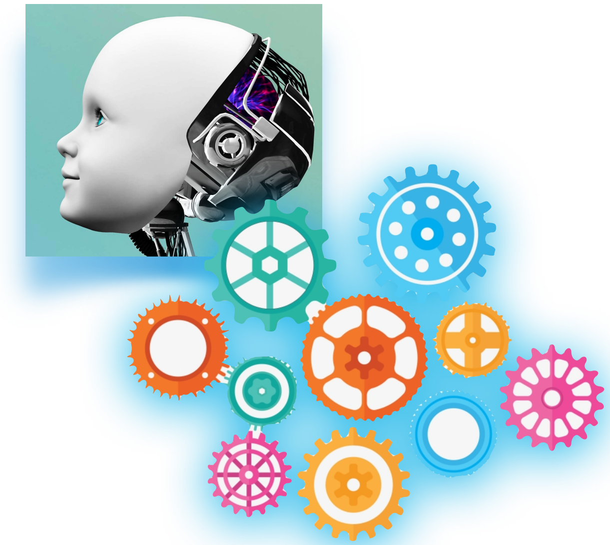
Complex NP island stimuli

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

Both theories

Evaluating the theory



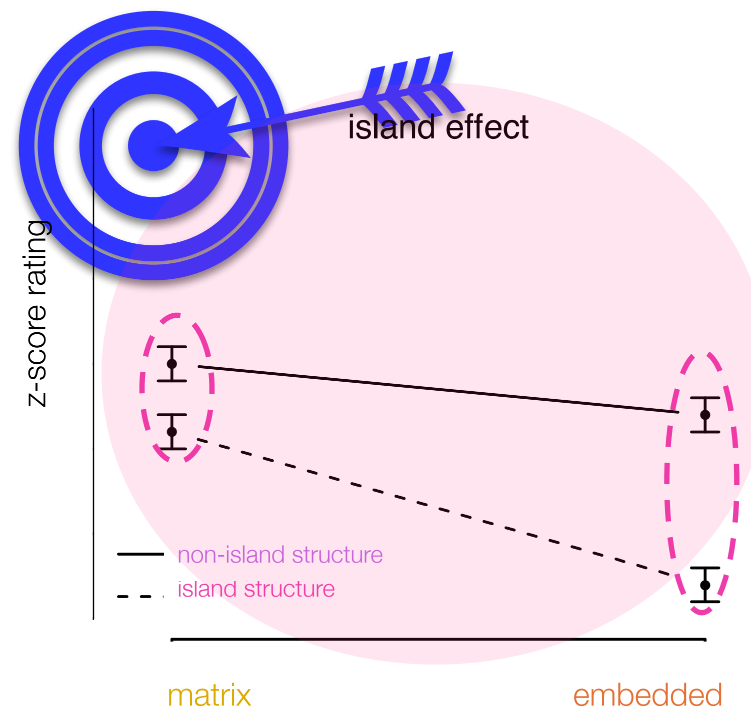
Reminder: Target behavior

Subject island

Complex NP island

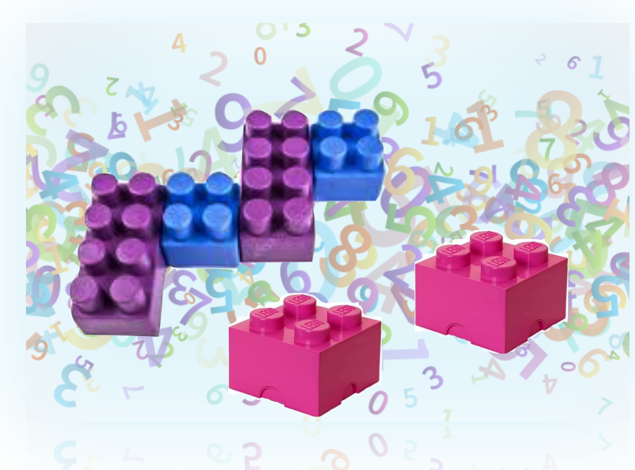
Whether island

Adjunct island

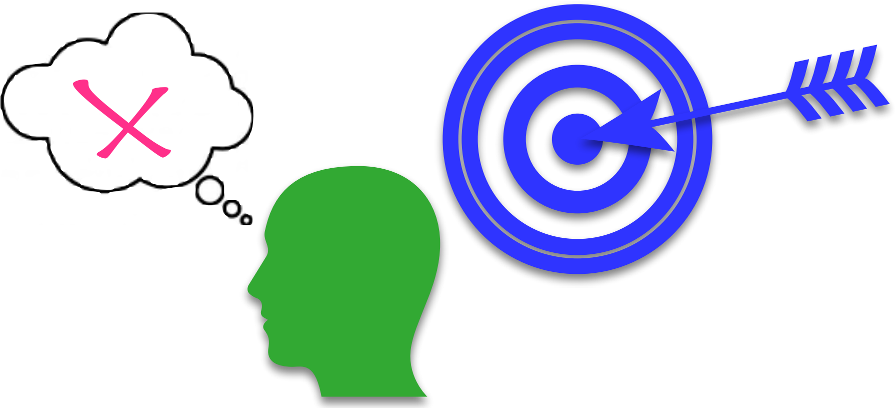


If so, then we predict the modeled child has syntactic island knowledge that allows the same judgment pattern as adults, learned from the building blocks in children's input.

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

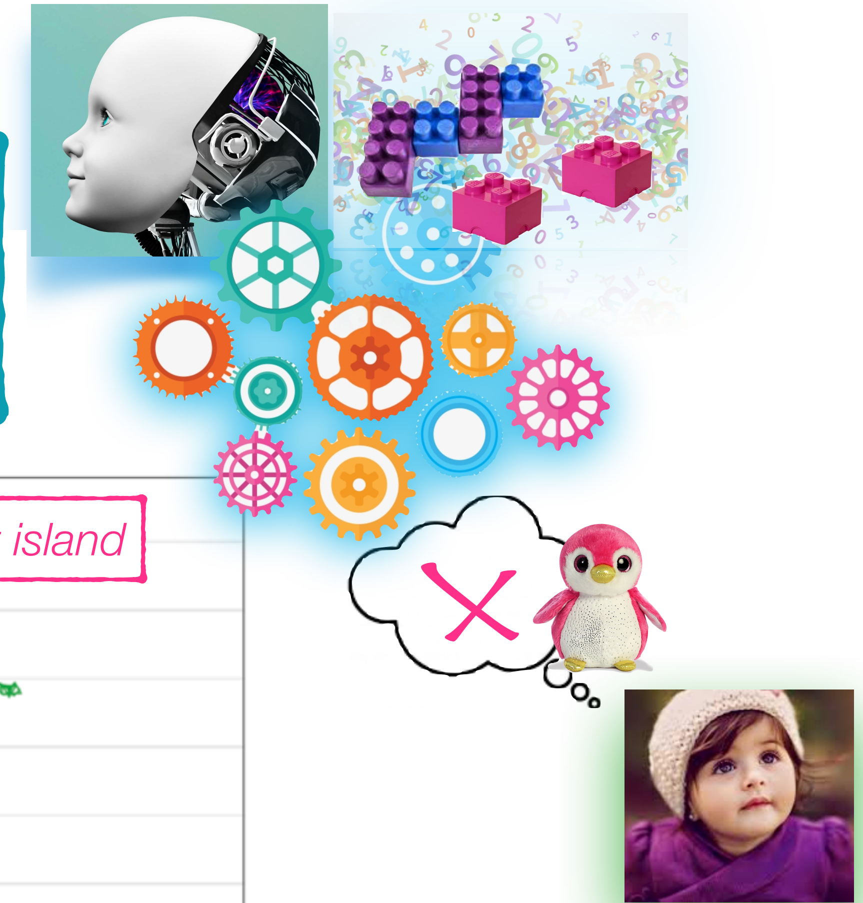


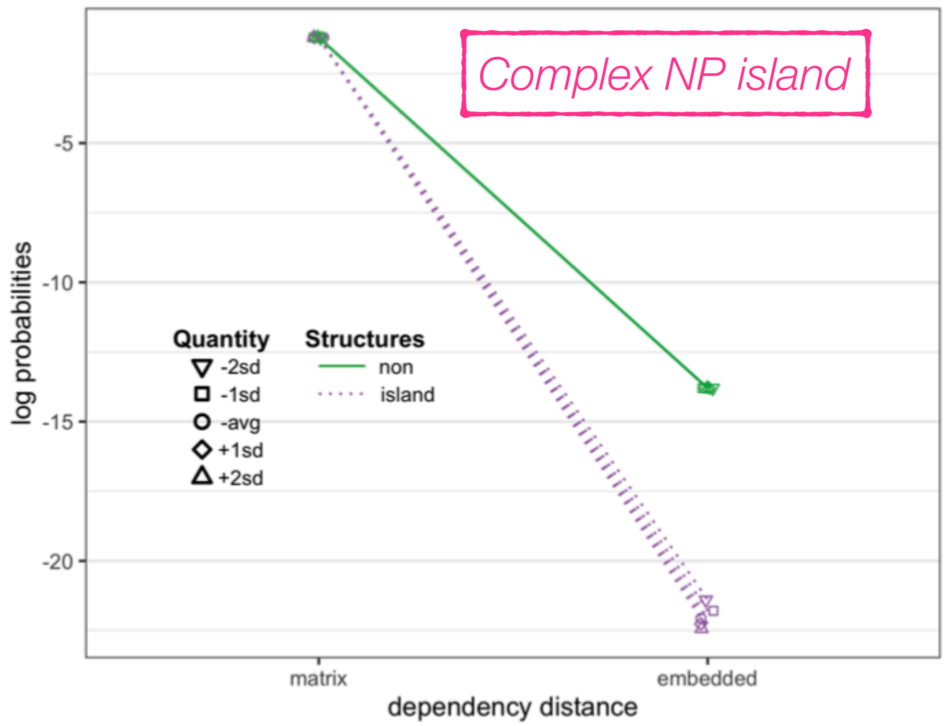
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

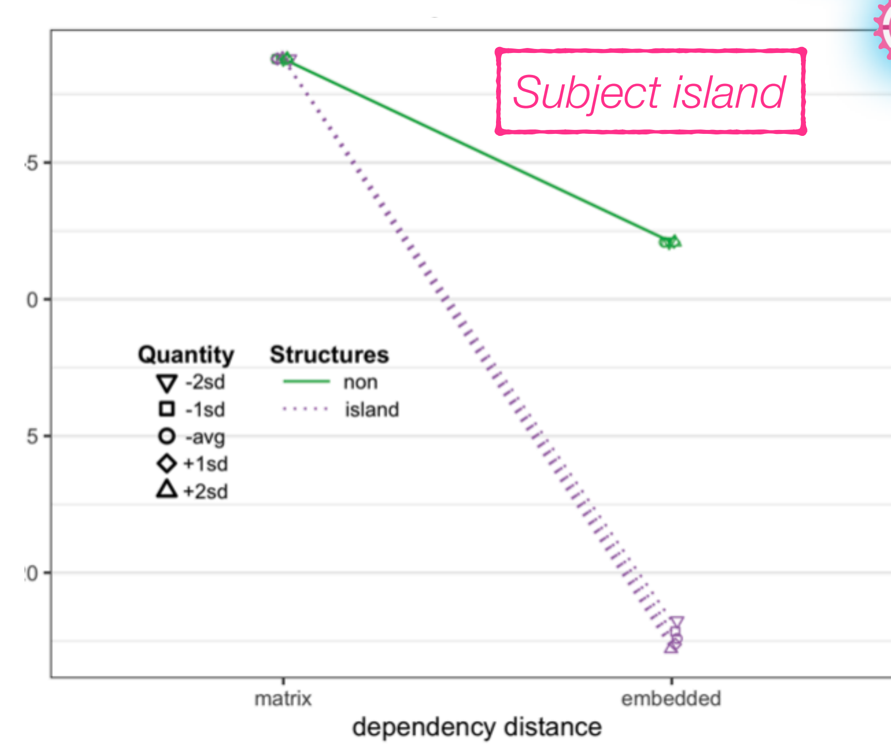


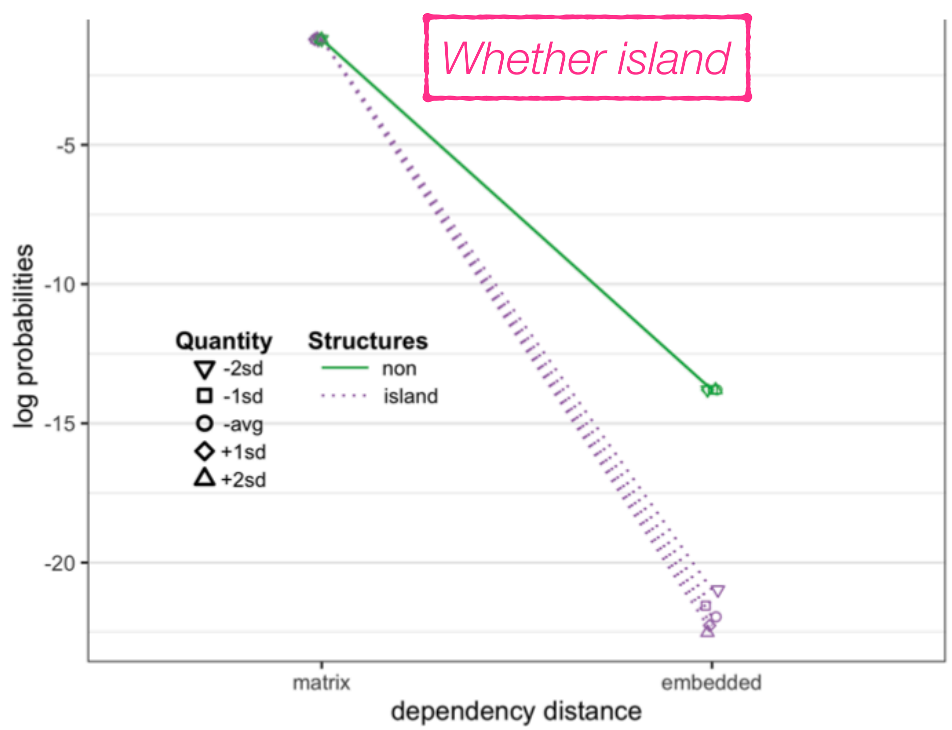
Evaluating the theory

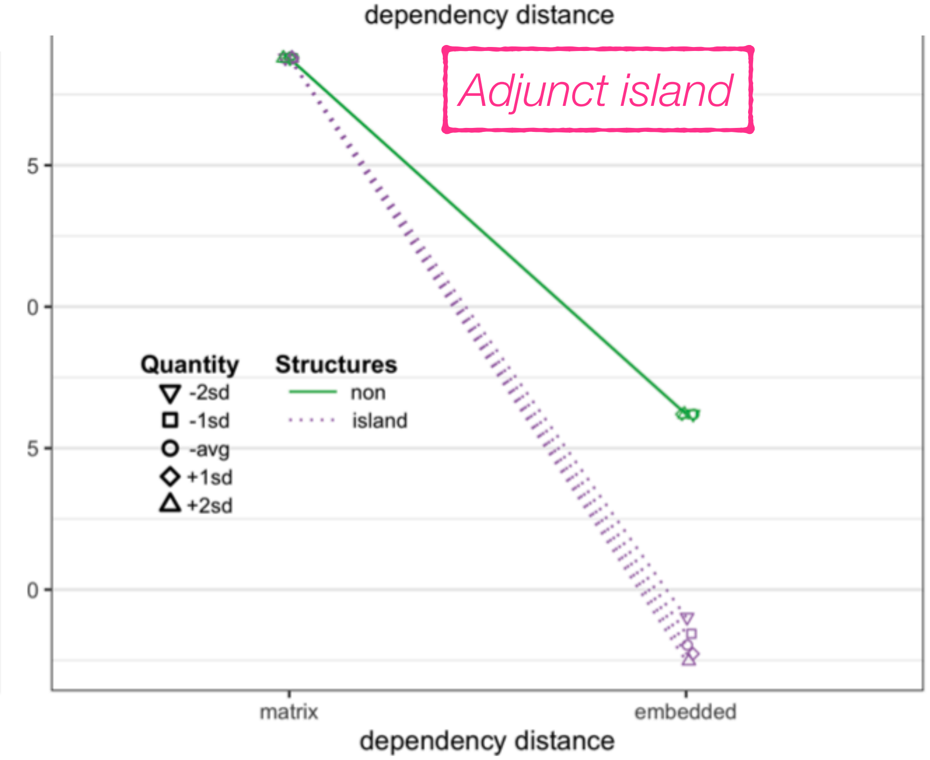
1: probabilities of pre-specified pieces



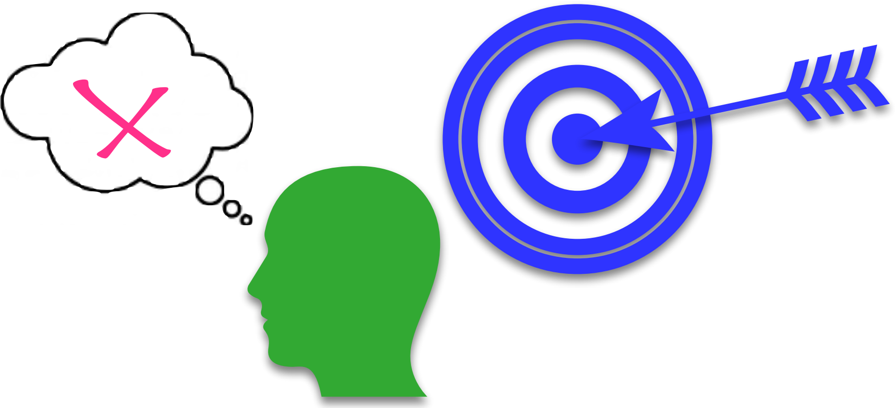






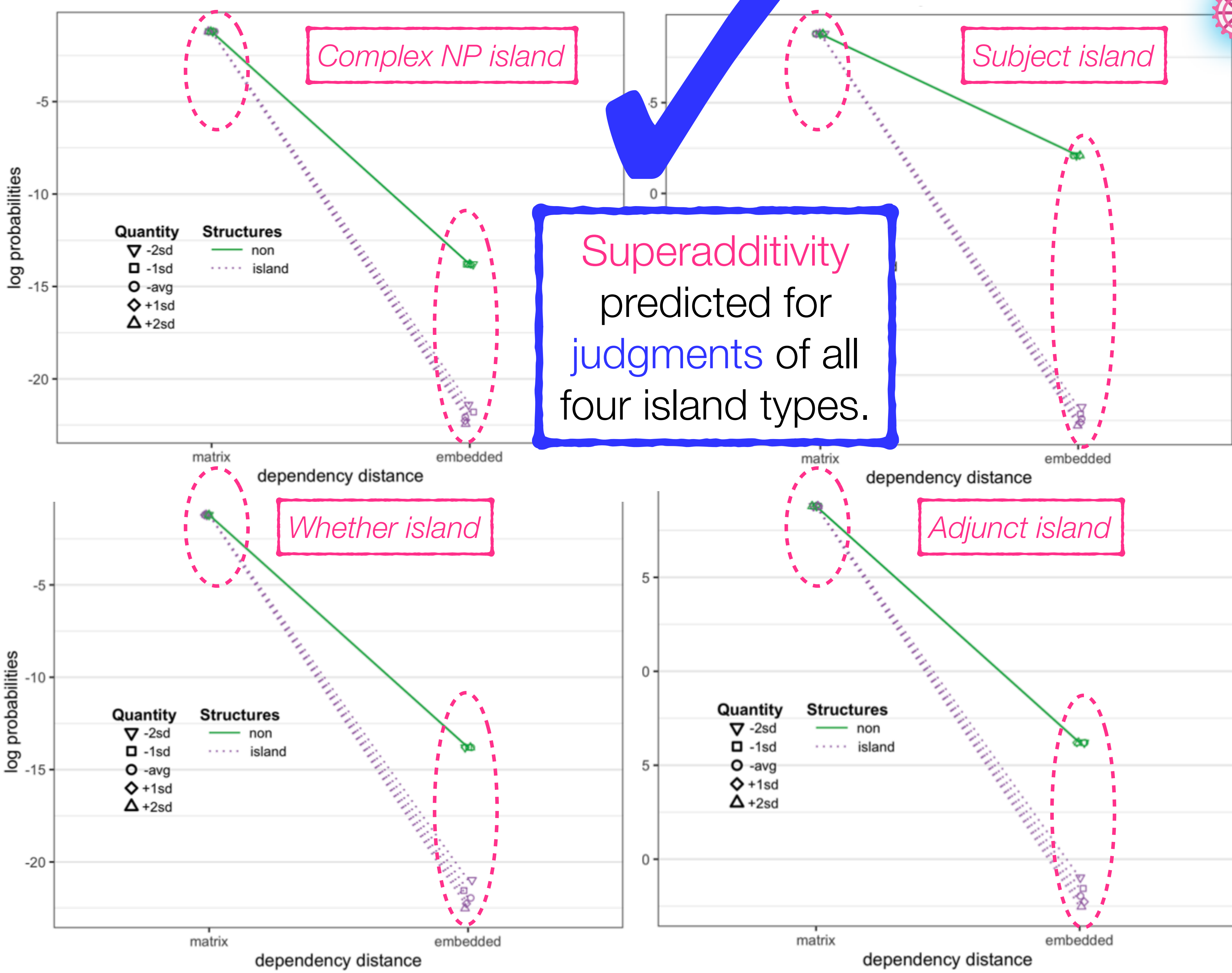


Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press



Evaluating the theory

1: probabilities of pre-specified pieces

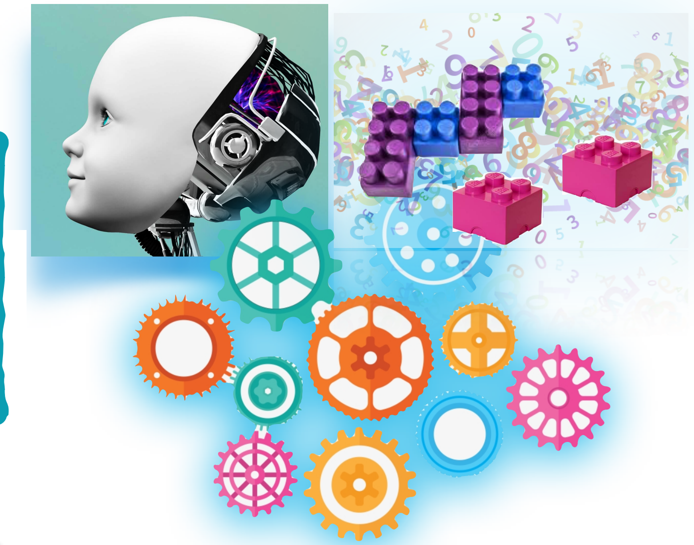


Superadditivity predicted for judgments of all four island types.

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

Evaluating the theory

1: probabilities of
pre-specified pieces

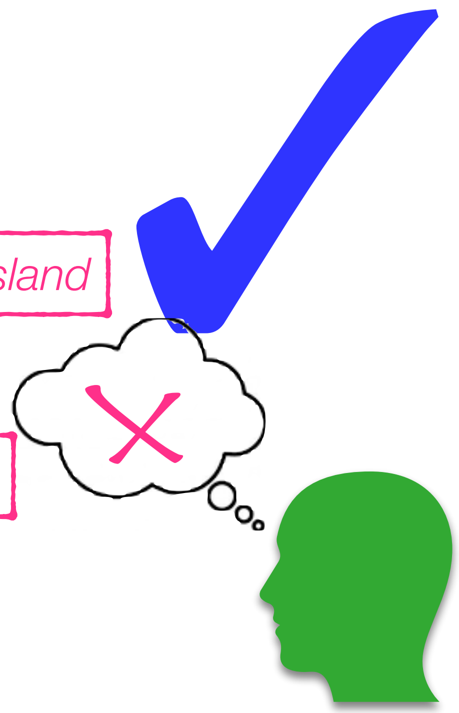


Subject island

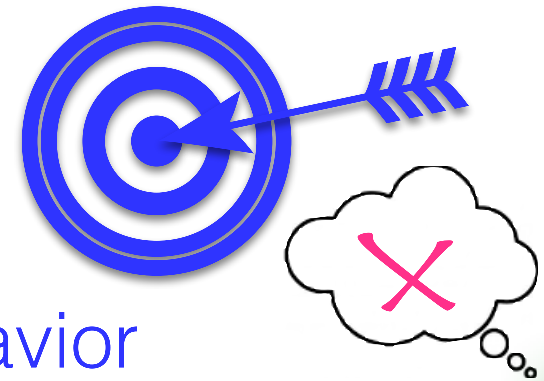
Complex NP island

Whether island

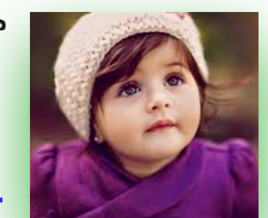
Adjunct island

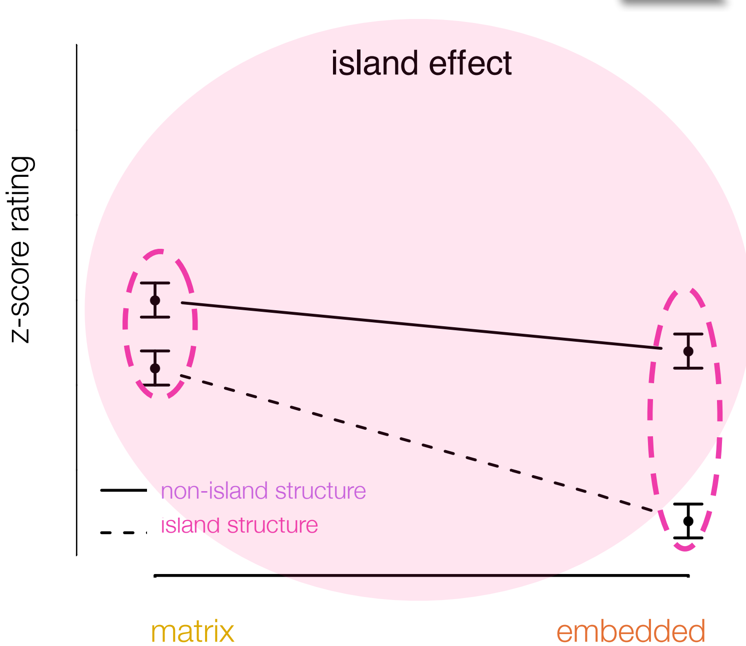


Reminder:
Target behavior

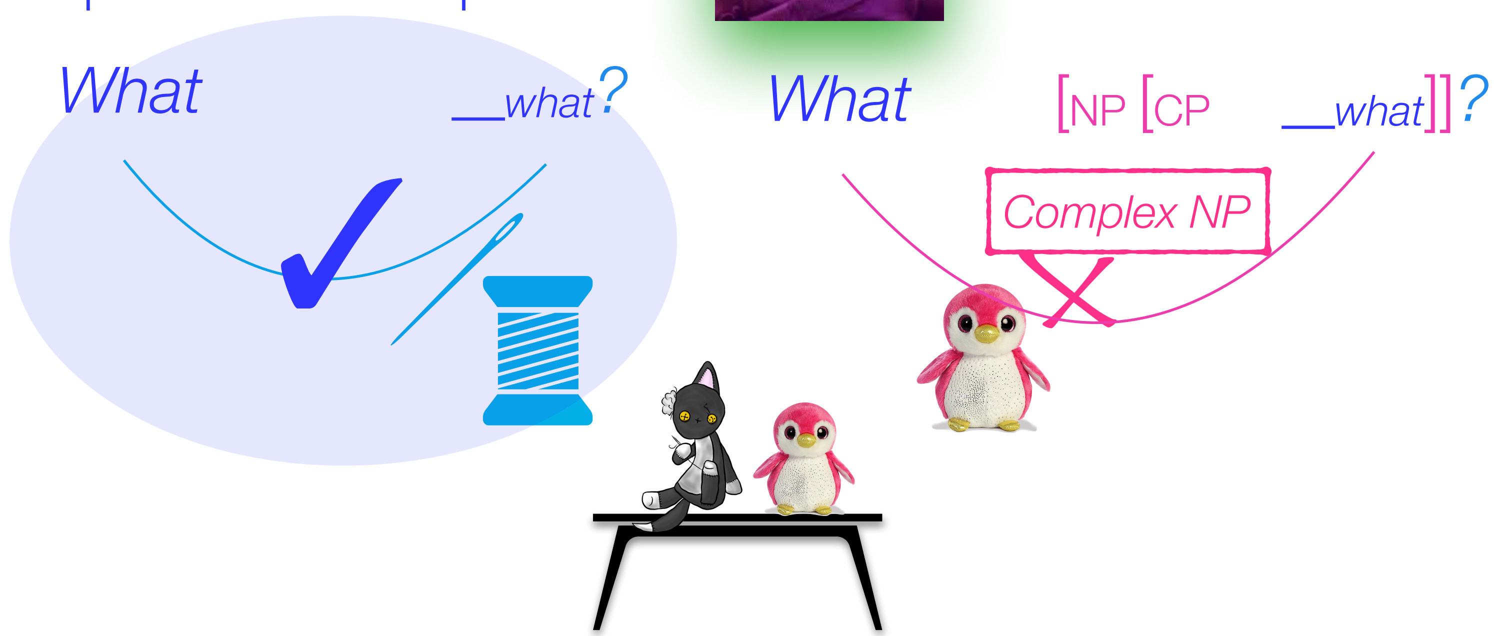


Children prefer this interpretation.





Sprouse et al. 2012

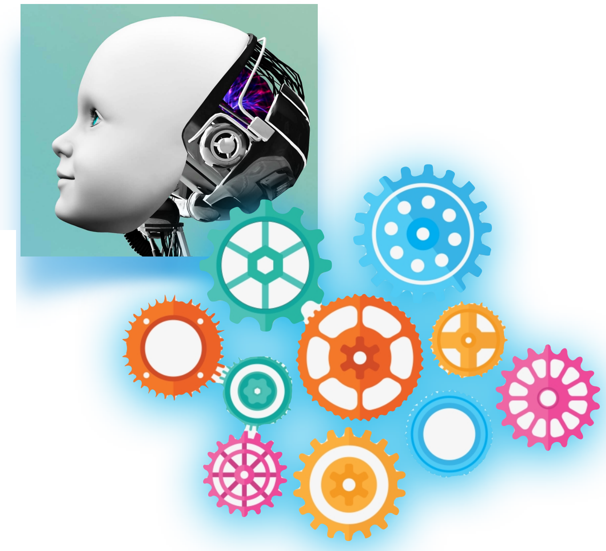


De Villiers et al. 2008

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

Evaluating the theory

1: probabilities of pre-specified pieces

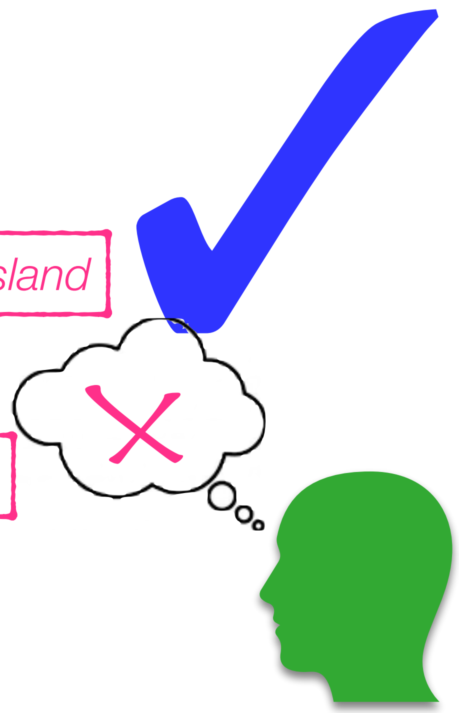


Subject island

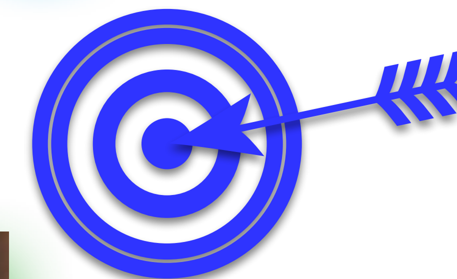
Complex NP island

Whether island

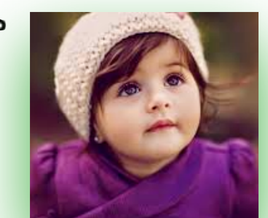
Adjunct island

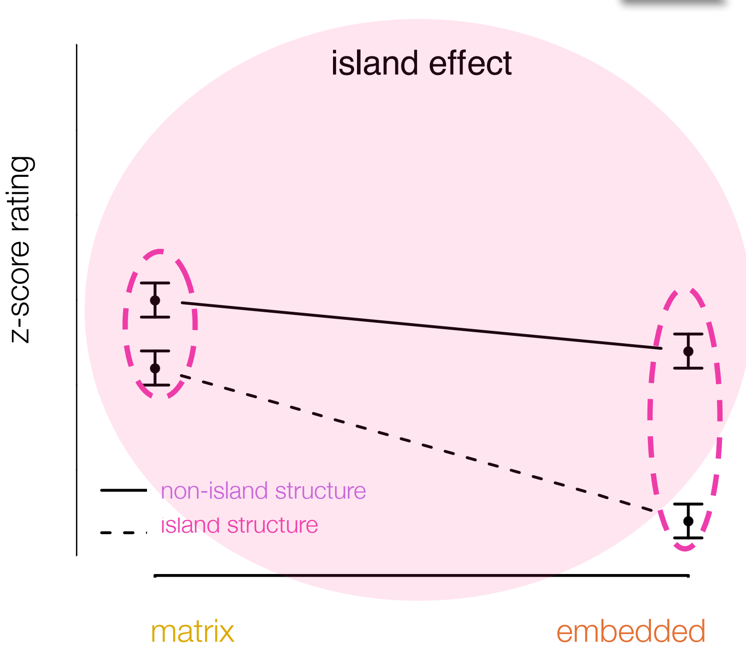


$\prod_{t \in trigrams} p(t)$

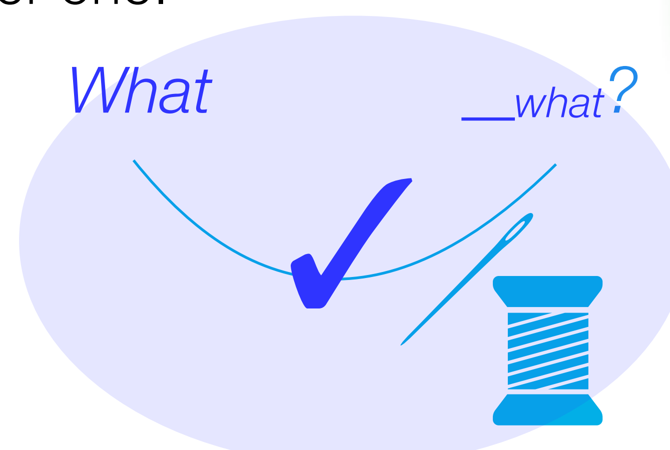


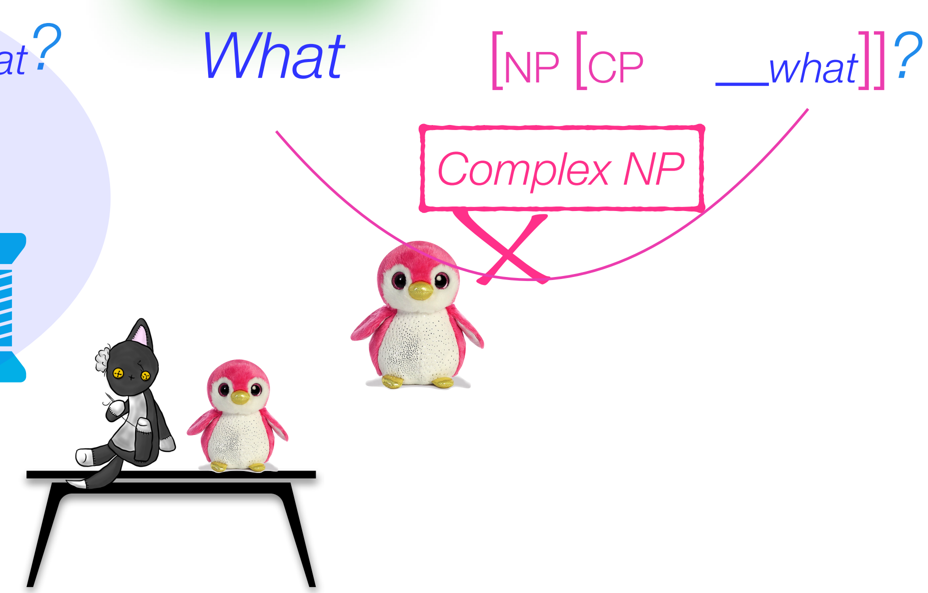
The *wh*-dependency this interpretation relies on is 10^{18} times more probable than the other one.





Sprouse et al. 2012



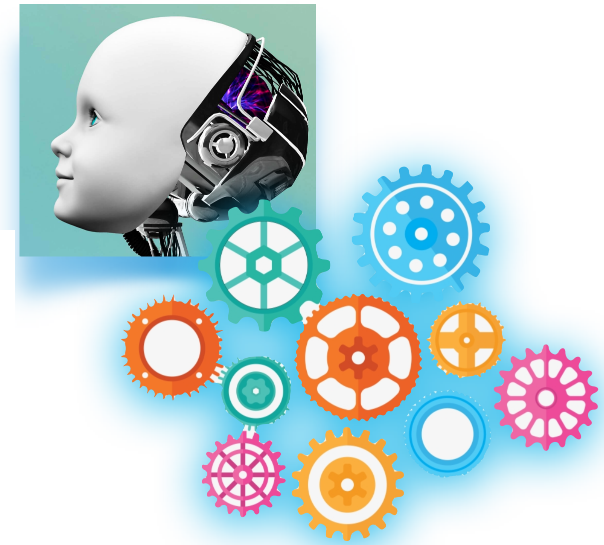


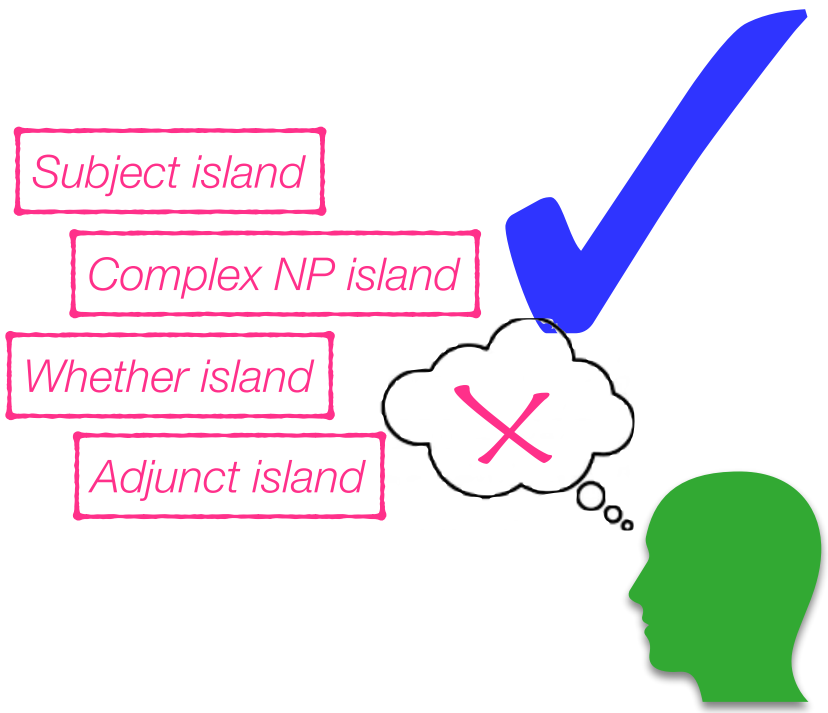
De Villiers et al. 2008

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

Evaluating the theory

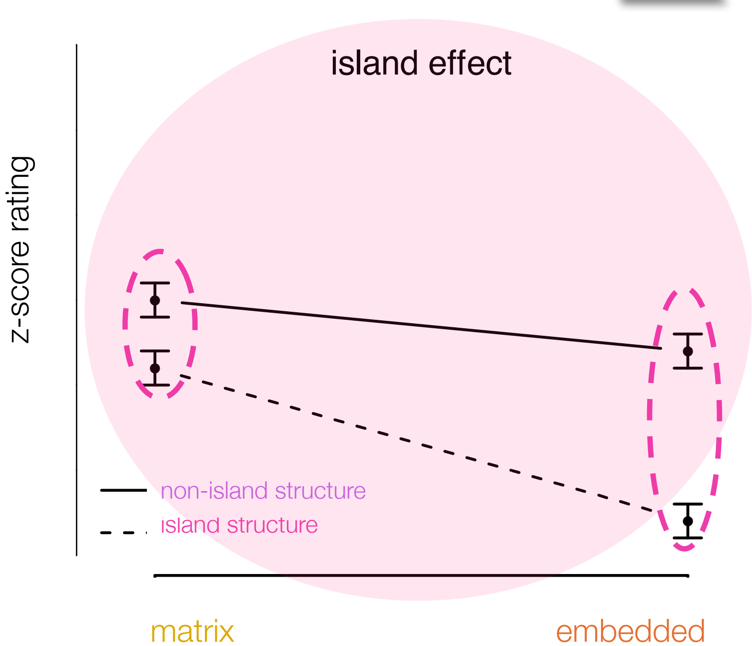
1: probabilities of
pre-specified pieces



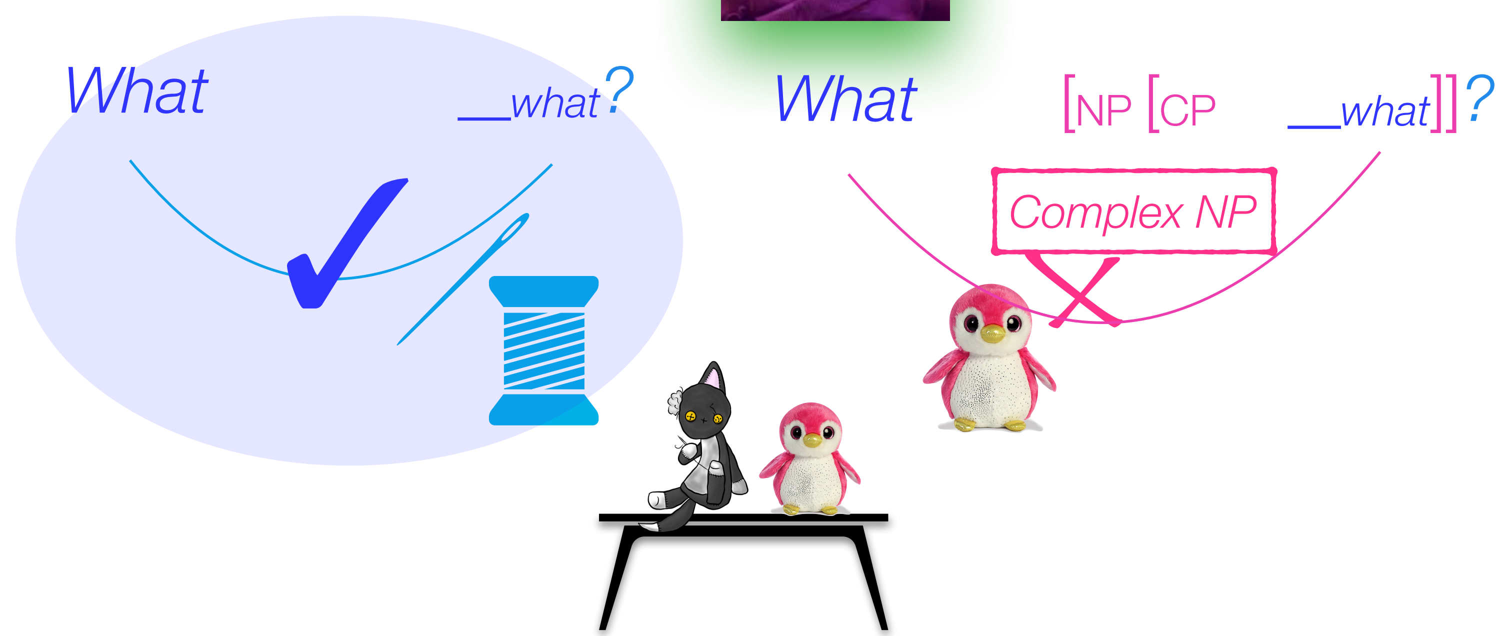


$$\prod_{t \in trigrams} p(t)$$

So, the modeled child prefers it.



Sprouse et al. 2012

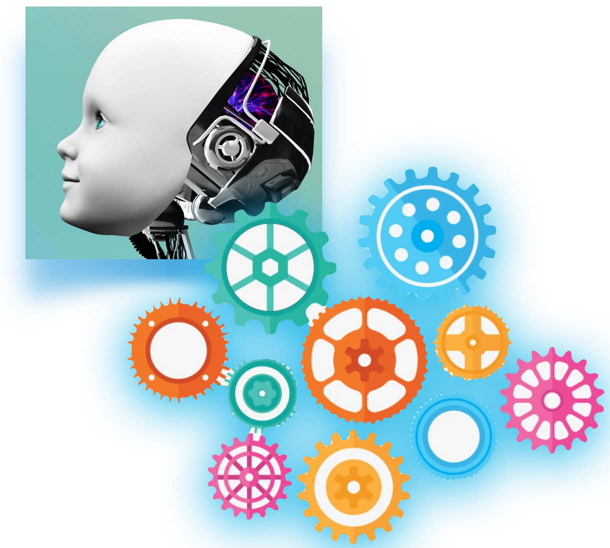


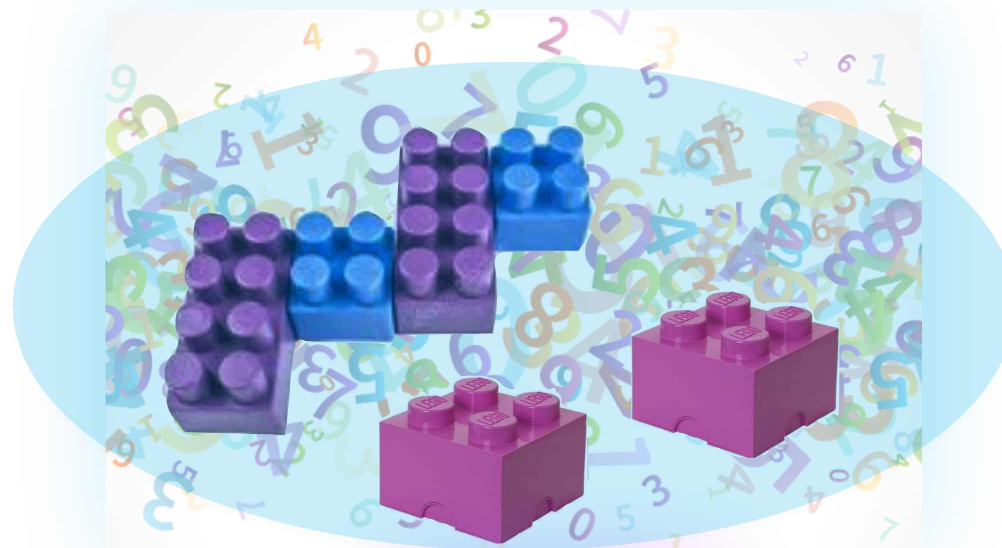
De Villiers et al. 2008

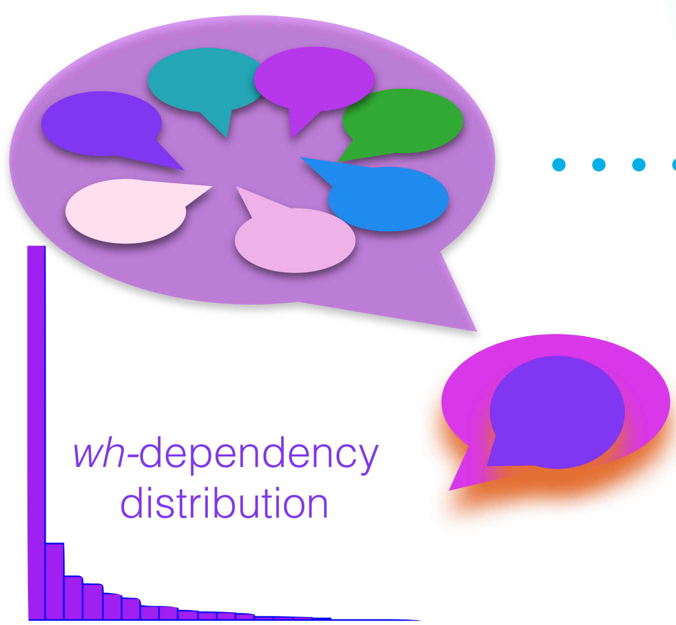
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press

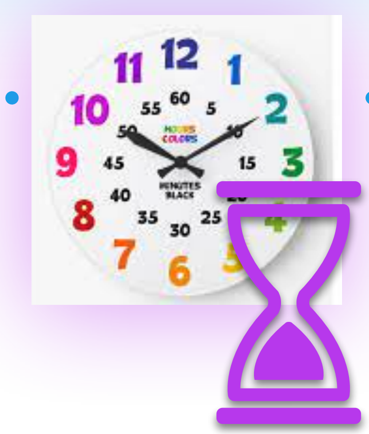
1: probabilities of pre-specified pieces

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







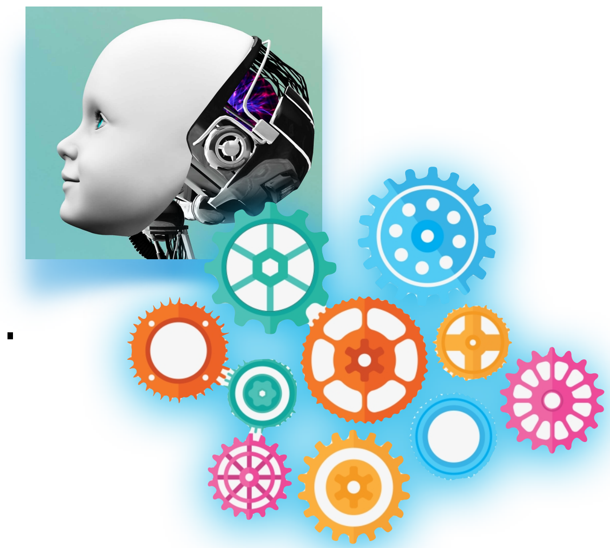


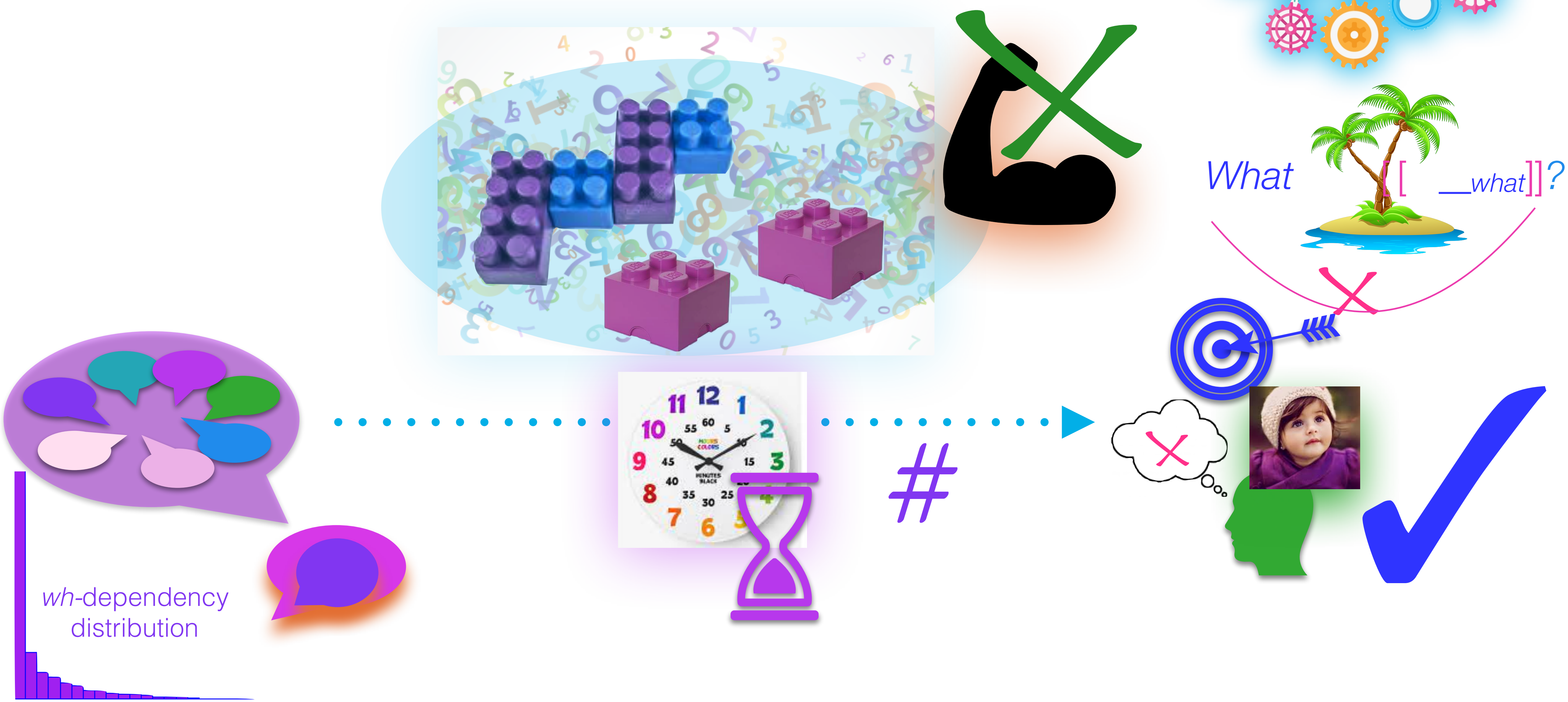
#

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

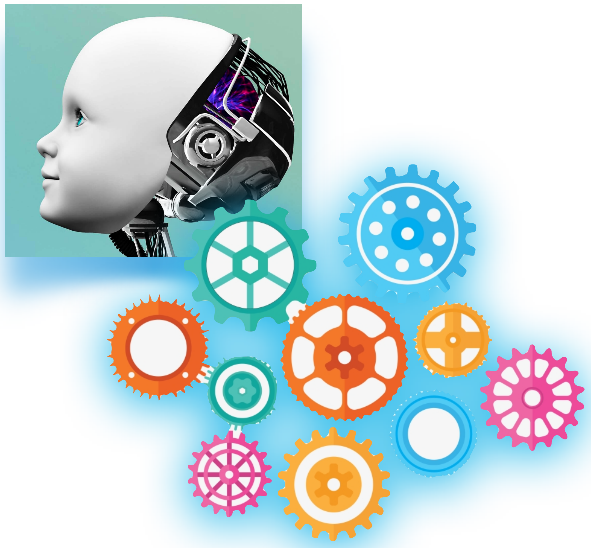
1: probabilities of pre-specified pieces

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





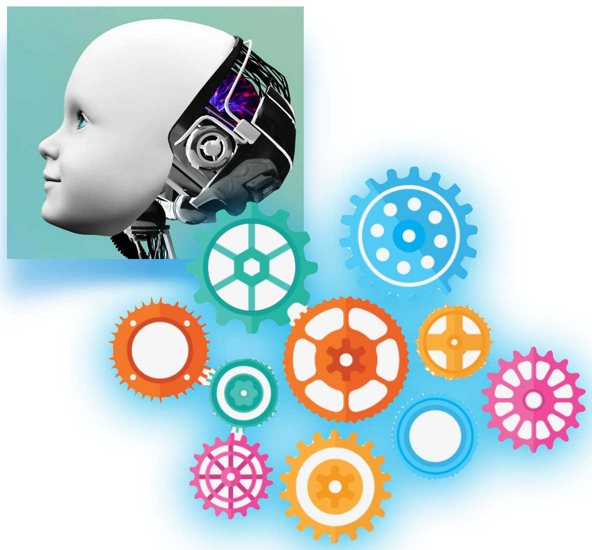
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates in press



What about the other theory that learns what the building blocks of *wh*-dependencies are at the same time as it learns their probabilities?

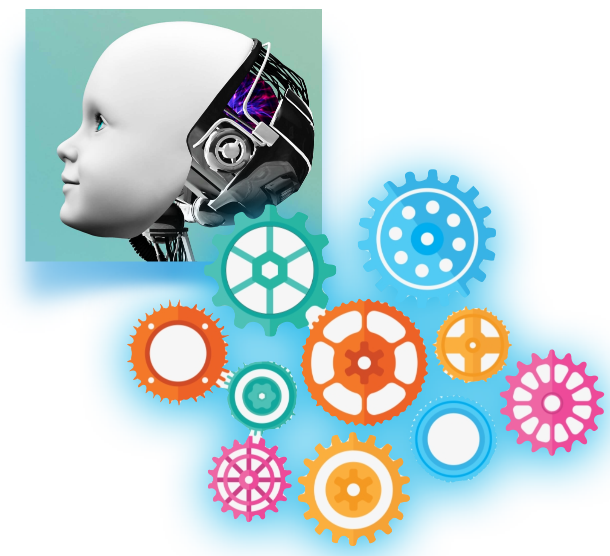
Dickson, Pearl, & Futrell 2022, in prep.

2: what the pieces are and their probabilities

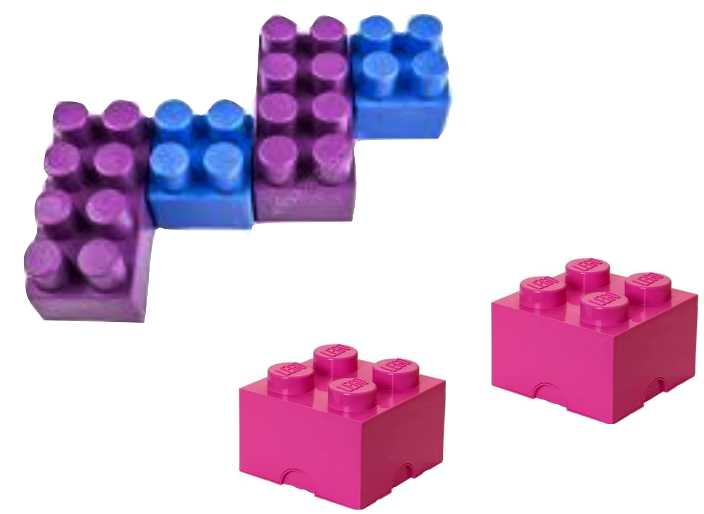


Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



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

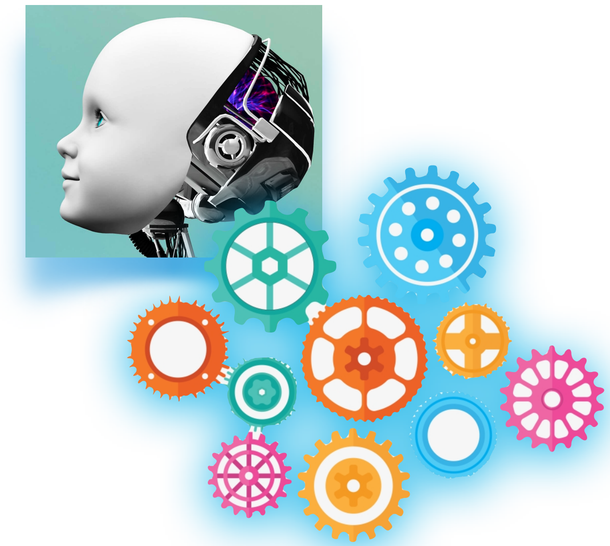
NP-PP-*end*

CP_{that}-IP-NP

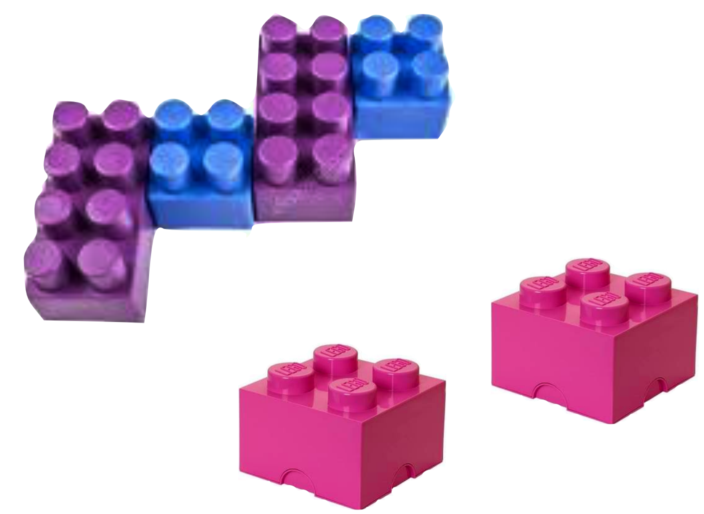
IP-NP-PP

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



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}

VP-PP-*end*

IP-VP-PP

VP-CP_{that}-IP

NP-PP-*end*

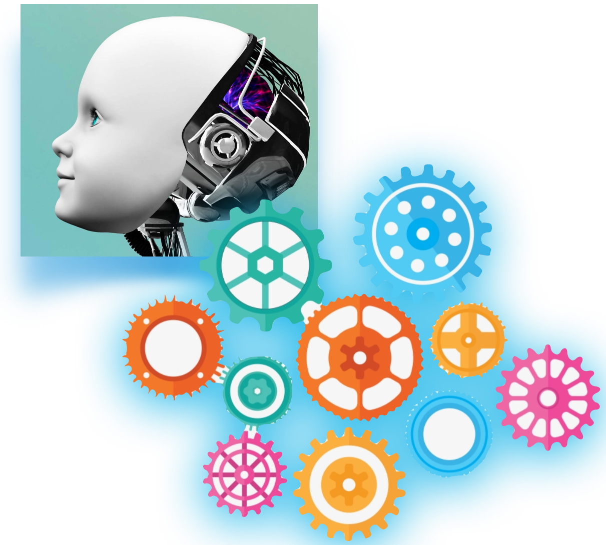
CP_{that}-IP-NP

IP-NP-PP

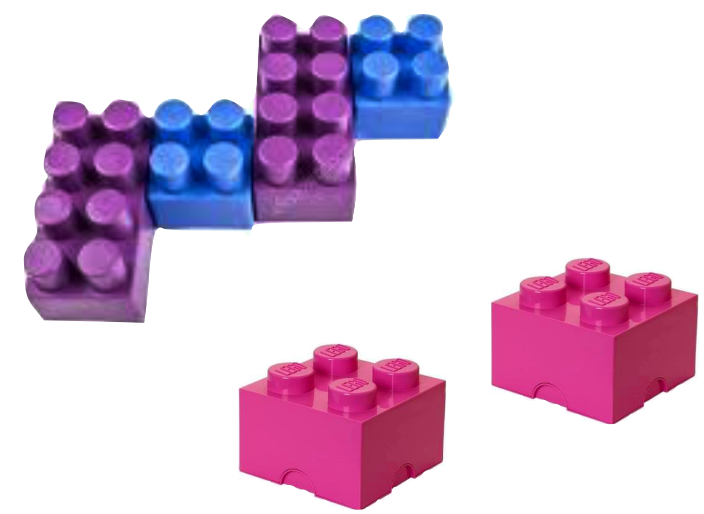
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

…CP_{that}…



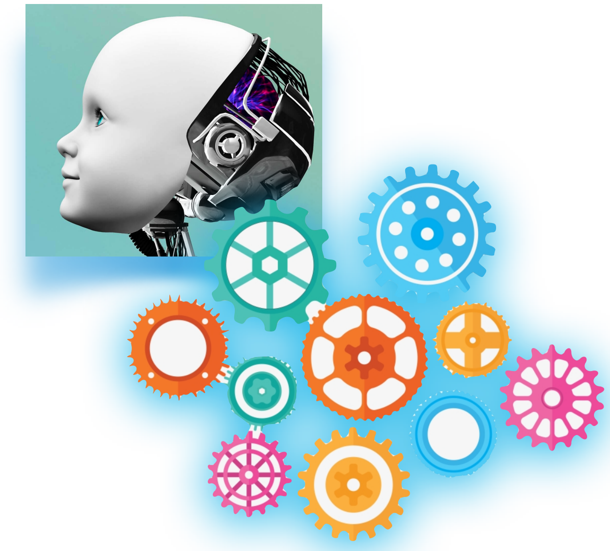
Empirical motivation for the CP lexical item: Two of the islands (Whether and Adjunct) only differ from more acceptable *wh*-dependencies by the complementizer used.



	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

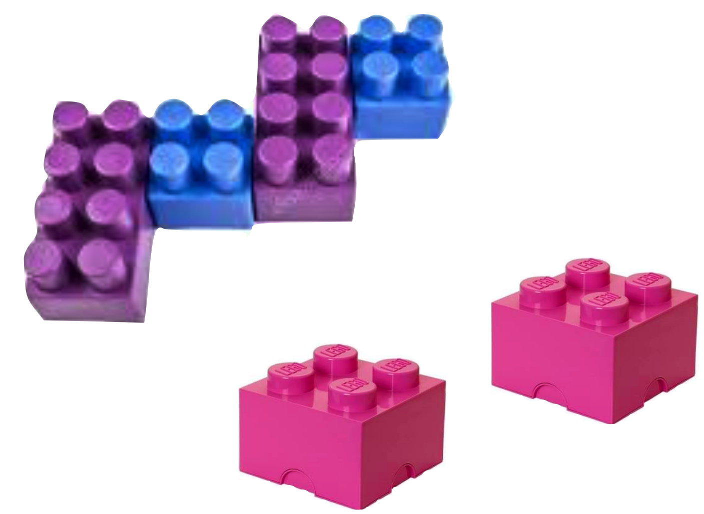
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



…CP_{that}…

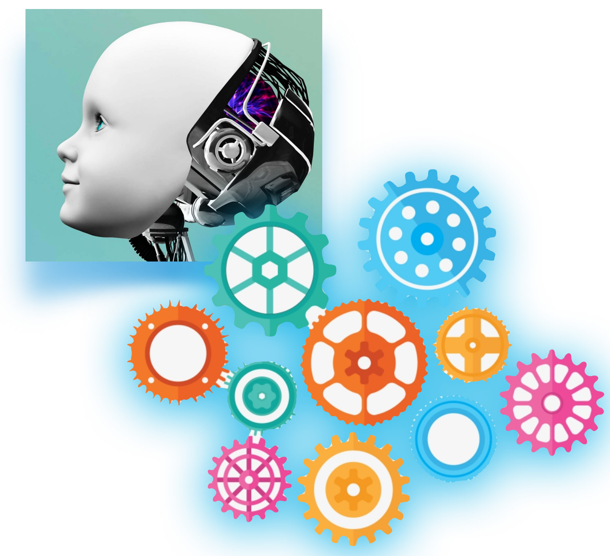
Empirical motivation for the CP lexical item: Two of the islands (Whether and Adjunct) only differ from more acceptable *wh*-dependencies by the complementizer used.



	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

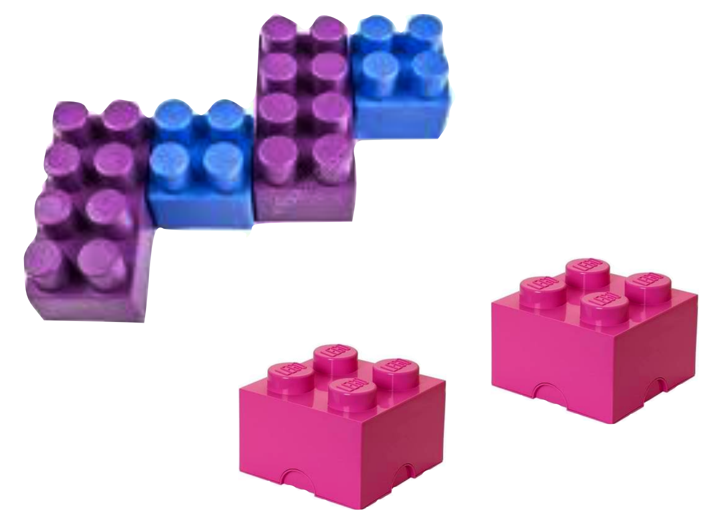
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



…CP_{that}…

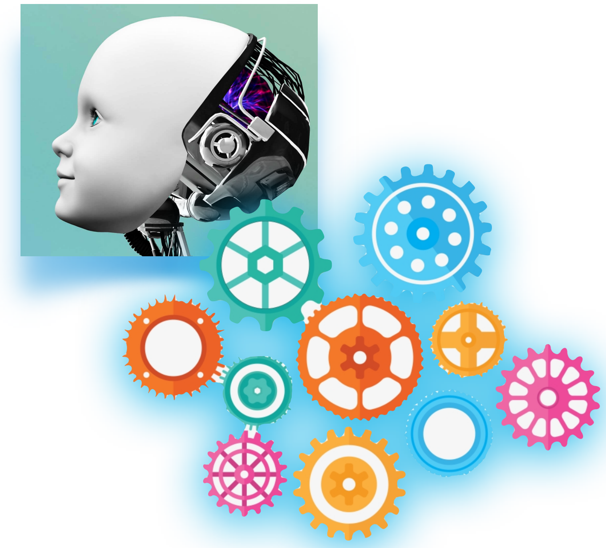
So the building blocks need to include this lexical item type.



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

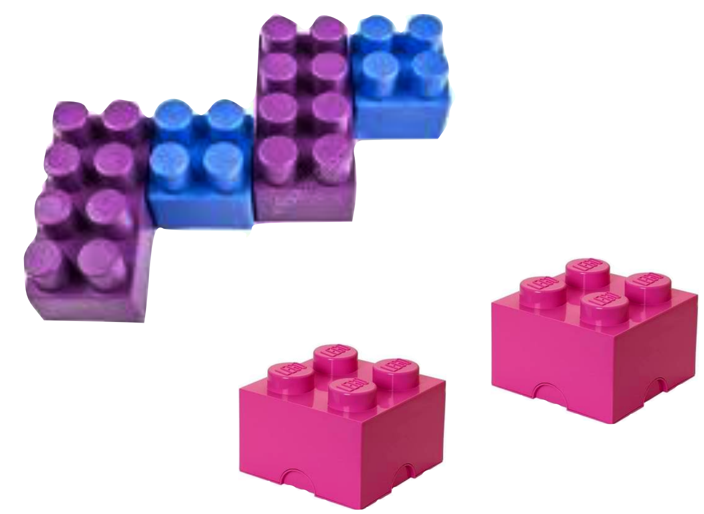
Dickson et al. 2022, in prep.

2: what the pieces are
and their probabilities



…CP_{that}…

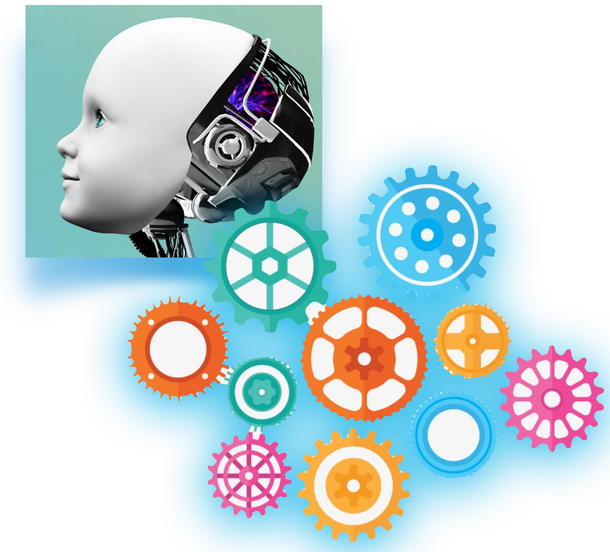
Is this the only one needed?



	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

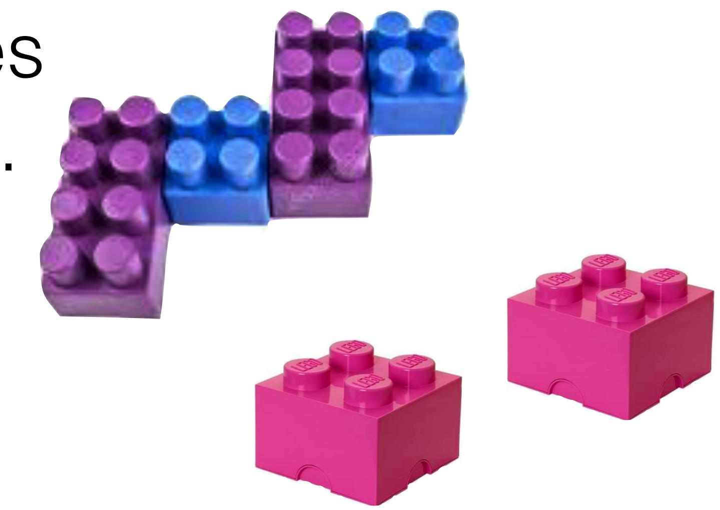
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



…CP_{that}…

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

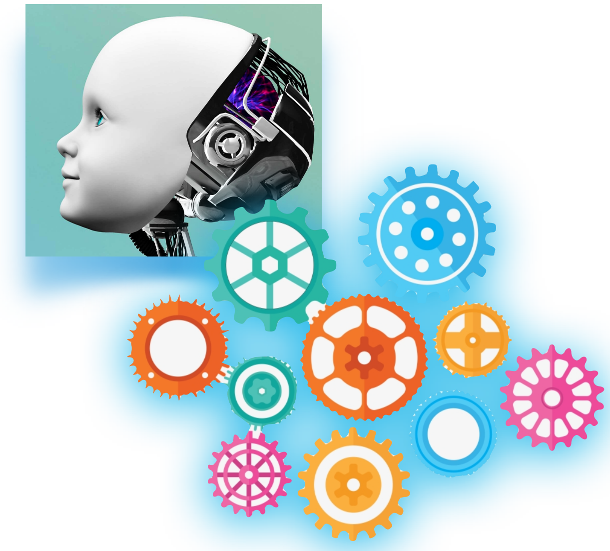


What did she think [that he saw __]?
What did she say [that he saw __]?

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

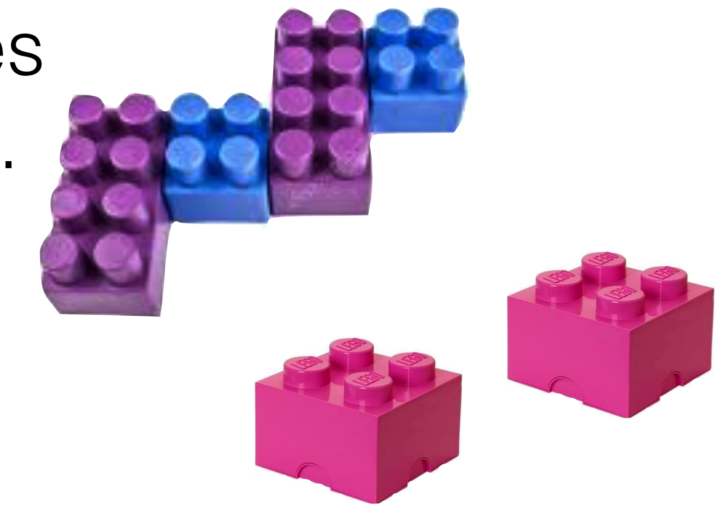
Dickson et al. 2022, in prep.

2: what the pieces are
and their probabilities



…CP$_{\text{that}}$…

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

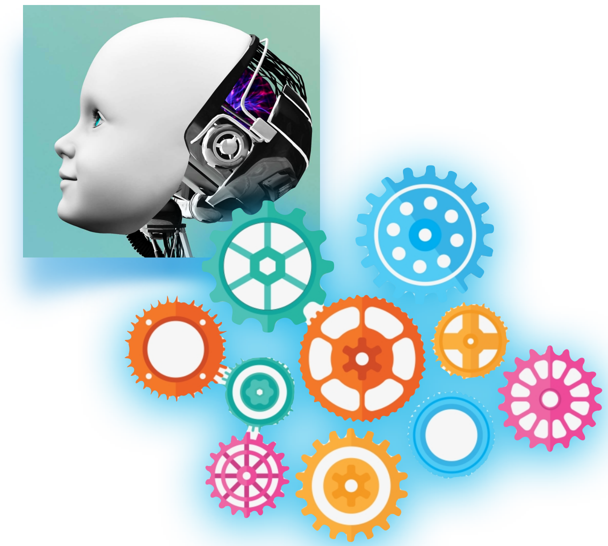


start-IP-VP$_{\text{think}}$ -CP$_{\text{that}}$-IP-VP-*end*
start-IP-VP$_{\text{say}}$ -CP$_{\text{that}}$-IP-VP-*end*

start-IP-VP$_{\text{whine}}$ -CP$_{\text{that}}$-IP-VP-*end*
start-IP-VP$_{\text{mumble}}$ -CP$_{\text{that}}$-IP-VP-*end*

Dickson et al. 2022, in prep.

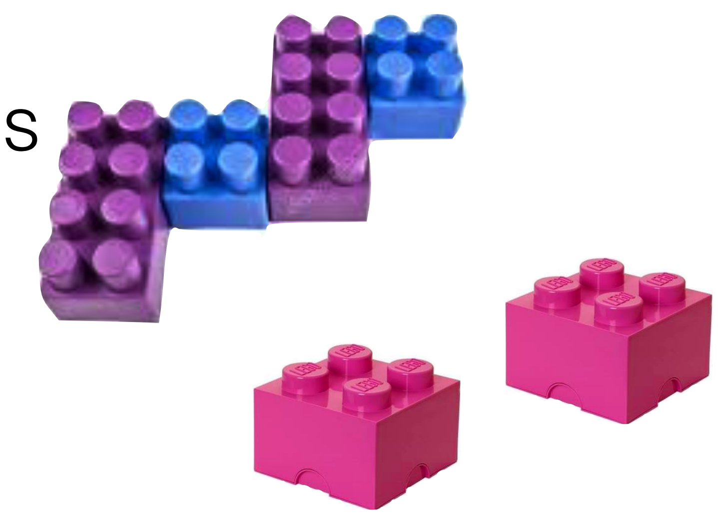
2: what the pieces are and their probabilities



…CP_{that}…

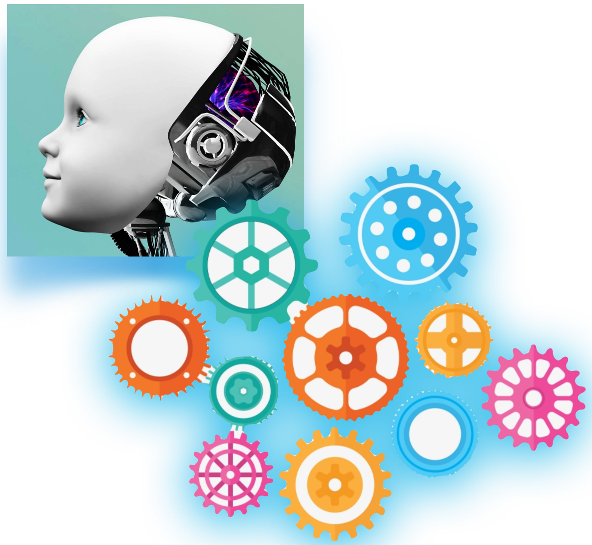
…VP_{think}…

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

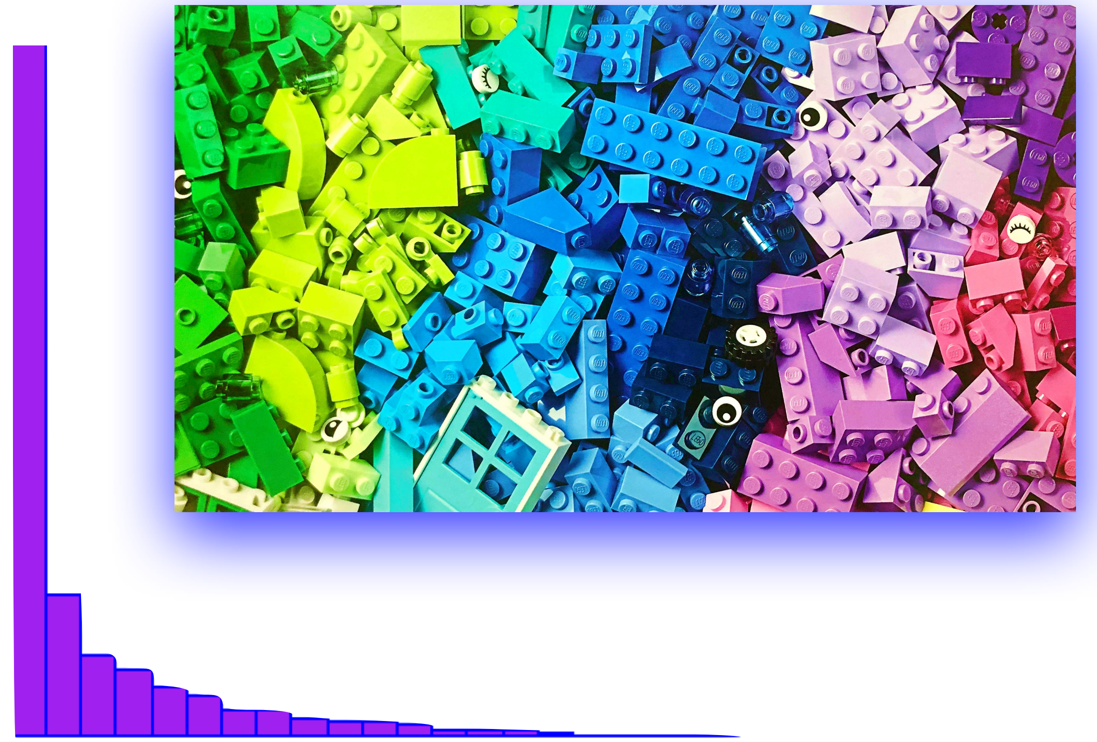


Dickson et al. 2022, in prep.

2: what the pieces are
and their probabilities

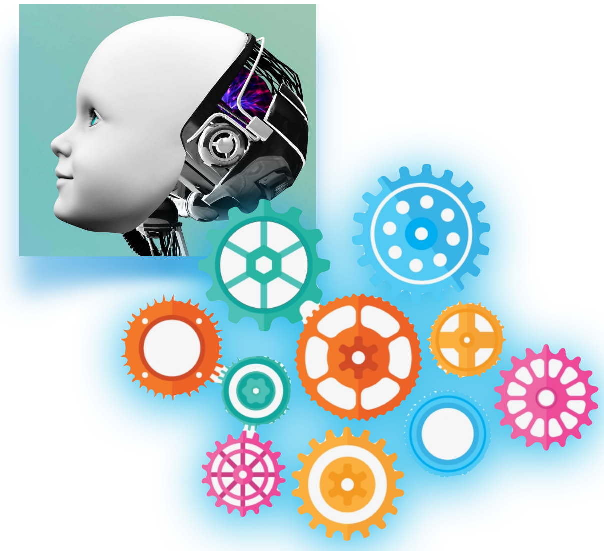


Theory: The child tries to learn
what the “best” building blocks are
at the same time she learns about their
distributions in the input.



Dickson et al. 2022, in prep.

2: what the pieces are
and their probabilities

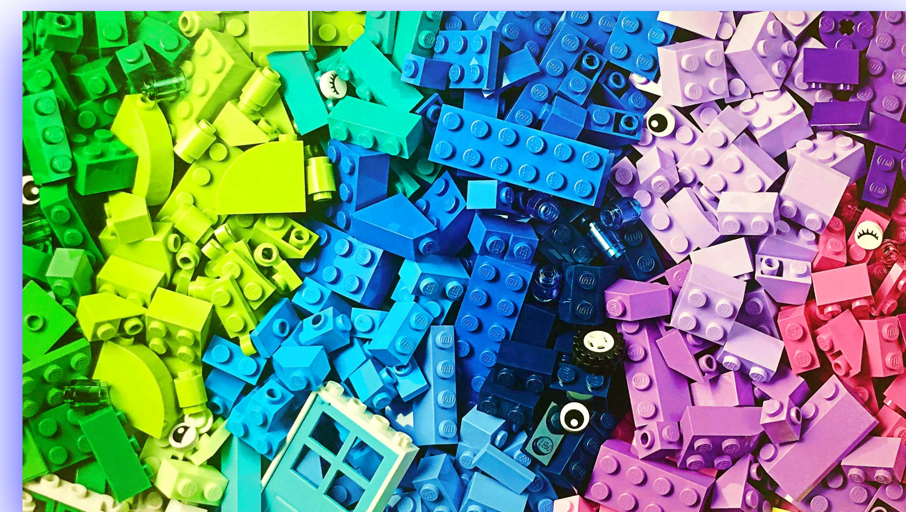


the best building blocks

Before:

(1) Look for groups of three units

(2) If the unit is a CP, include the lexical item



start-IP-VP

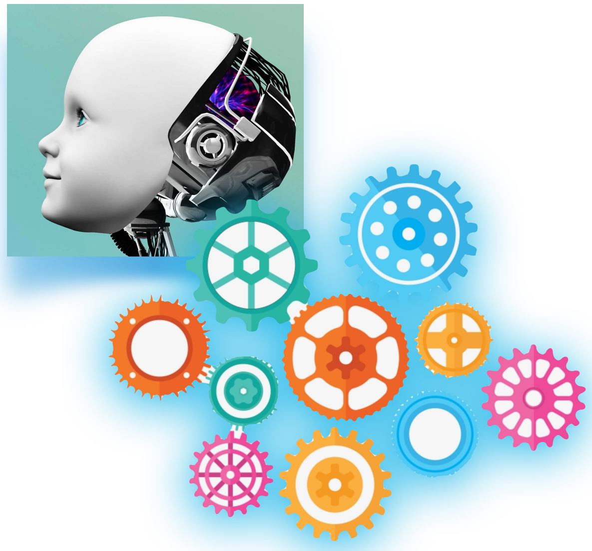
IP-VP-*end*

IP-VP-CP_{that}

VP-CP_{that}-IP

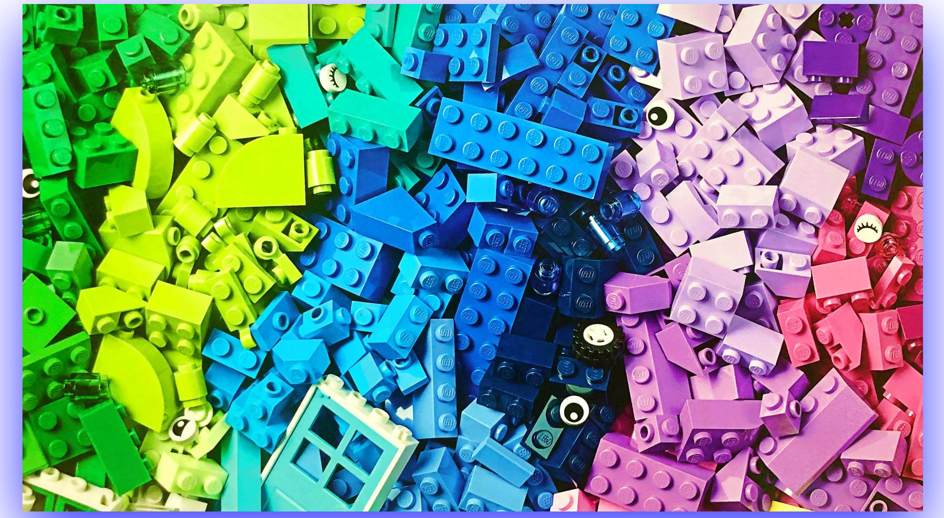
Dickson et al. 2022, in prep.

2: what the pieces are
and their probabilities



the best building blocks

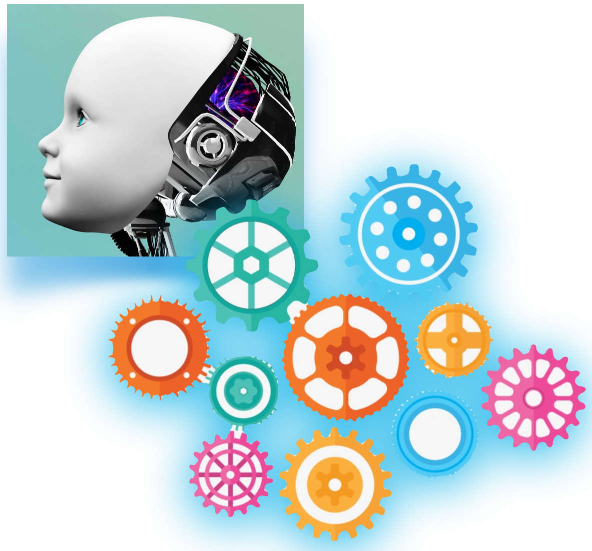
(1) Look for groups of ~~three units~~
(2) If the unit is a CP, include the lexical item



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

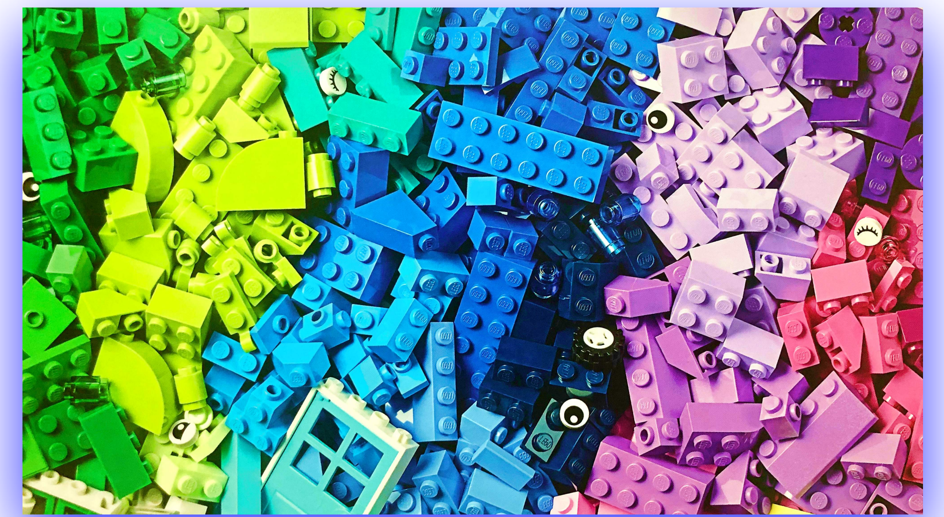
Dickson et al. 2022, in prep.

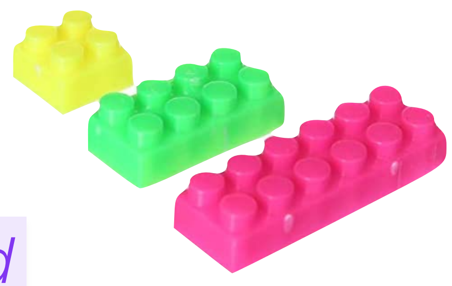
2: what the pieces are and their probabilities



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

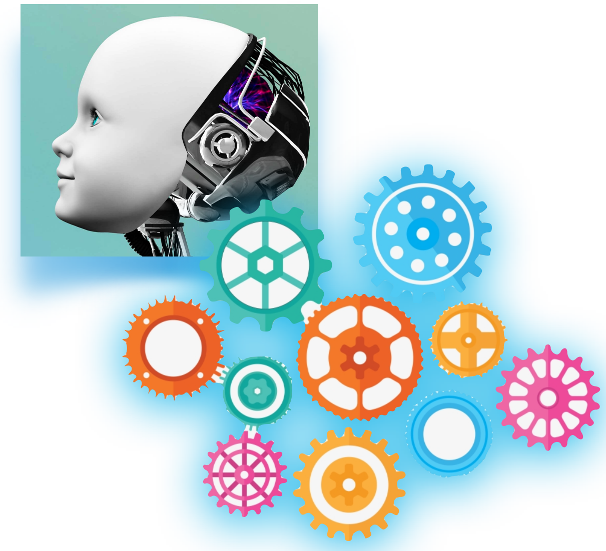
IP-VP-*end*

IP-VP-CP_{that}

VP-CP_{that}-IP

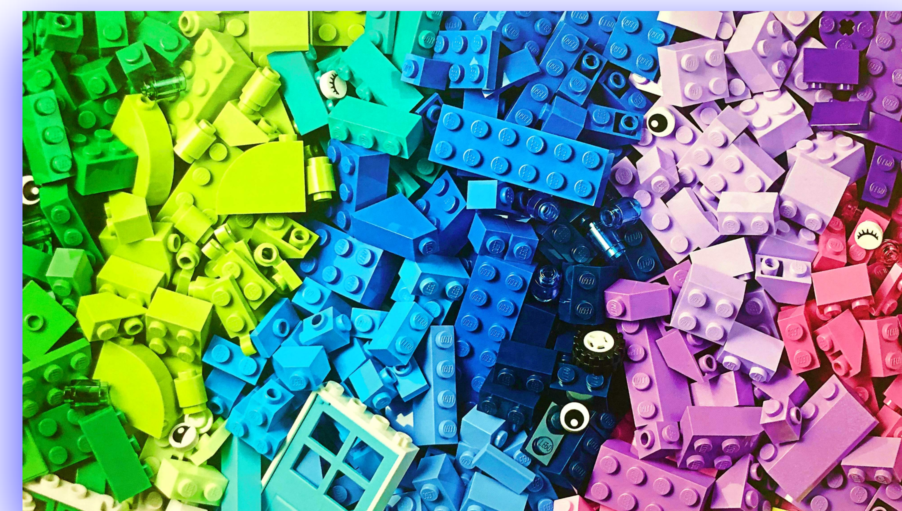
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



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}

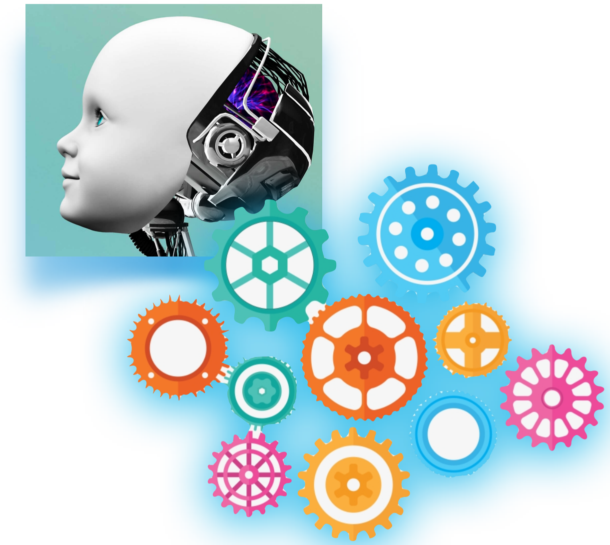
start-IP

IP-VP-*end*

CP_{that}-IP

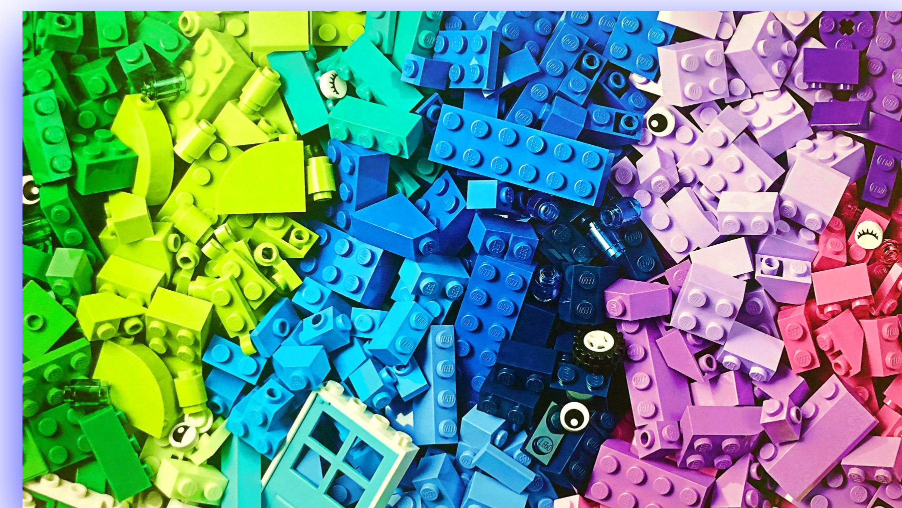
Dickson et al. 2022, in prep.

2: what the pieces are
and their probabilities



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$_{\text{think}}$-CP$_{\text{that}}$

start-IP

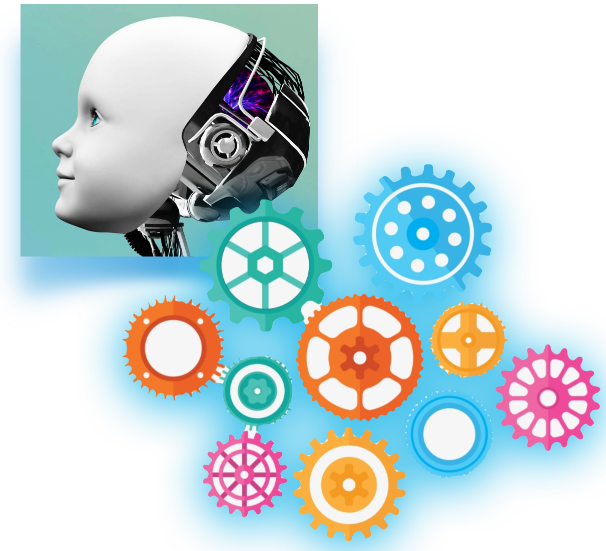
IP-VP-*end*

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

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

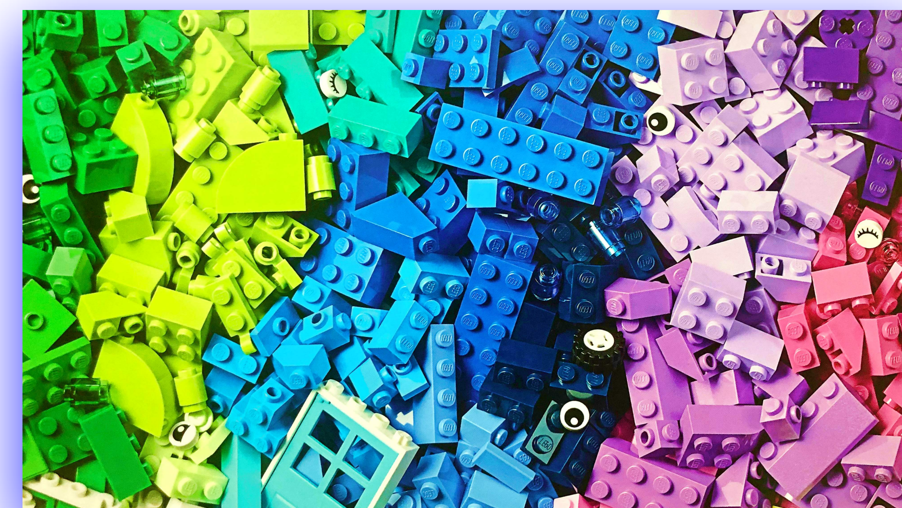
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



the best building blocks

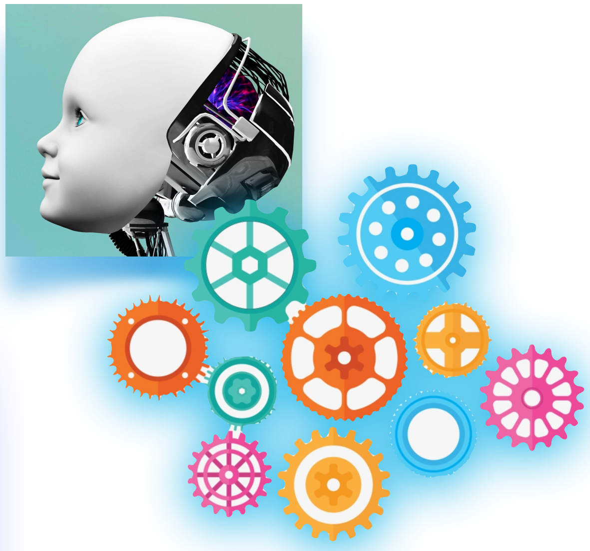
(1) Look for the best-sized units
(2) Sometimes include the lexical item



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

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

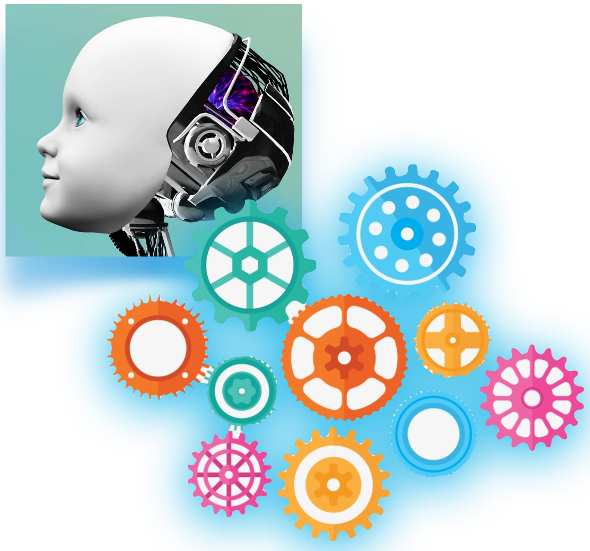


How can the child learn what the best building blocks are?

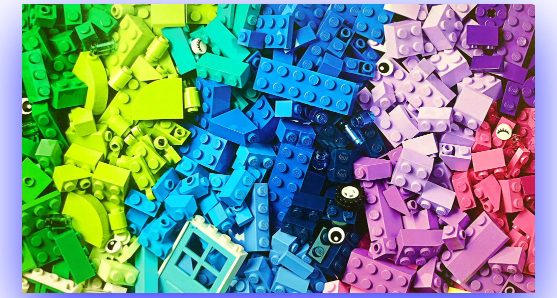
(1) Look for the best-sized units
(2) Sometimes include the lexical item

Dickson et al. 2022, *in prep.*

2: what the pieces are
and their probabilities



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

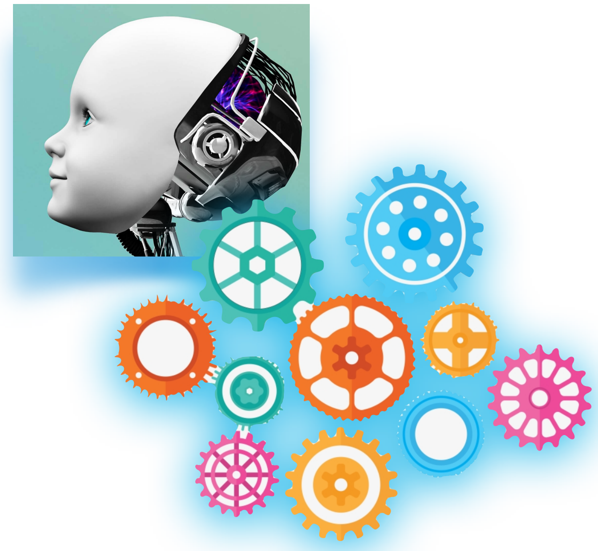


(1) Look for the best-sized units
(2) Sometimes include the lexical item

Theory: Look for an "efficient" set of building blocks.

Dickson et al. 2022, *in prep.*

2: what the pieces are and their probabilities



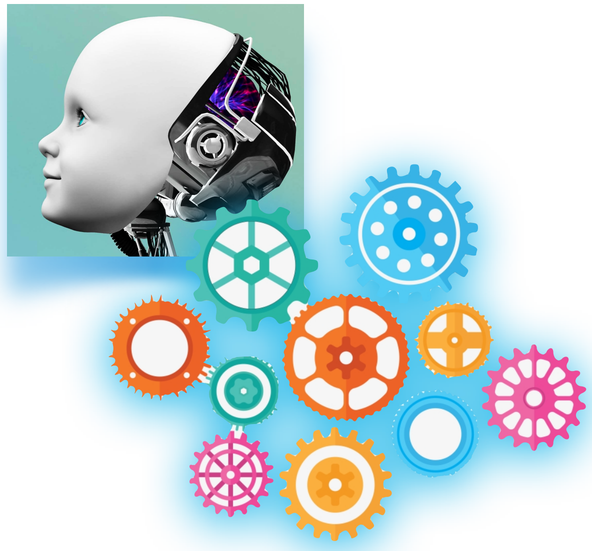
How can the child learn what the best building blocks are?

(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.

2: what the pieces are and their probabilities



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

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

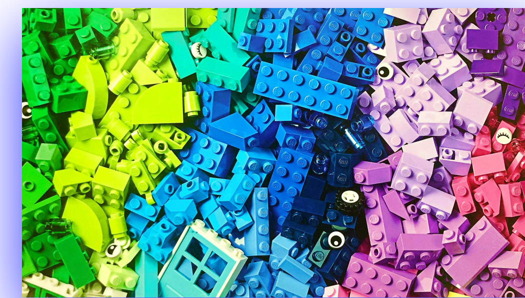
Efficient building blocks allow the representation of current and future *wh*-dependencies to be more probable.

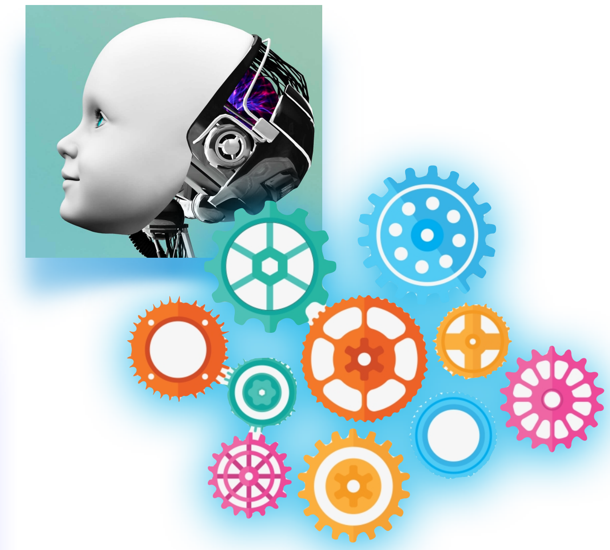
Why? One idea: Higher probability *wh*-dependencies are faster to process (comprehending or producing).

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

learning efficient building blocks

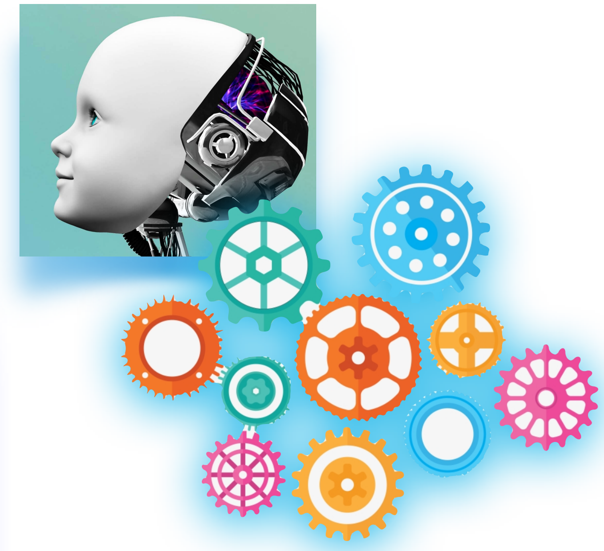




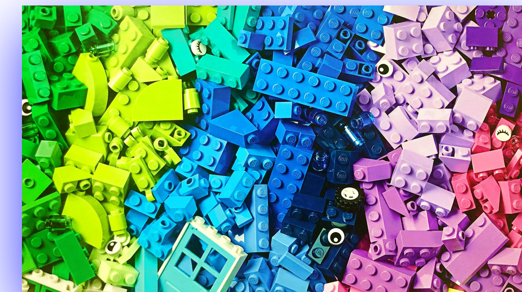
How? Look for building blocks that are a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

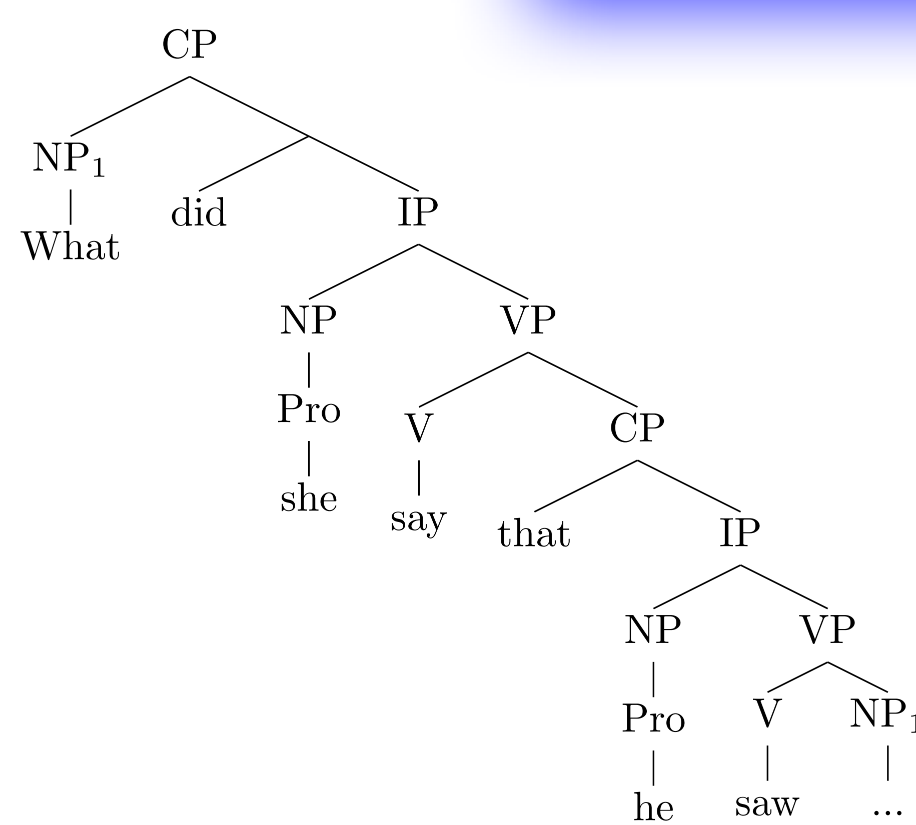


learning efficient building blocks



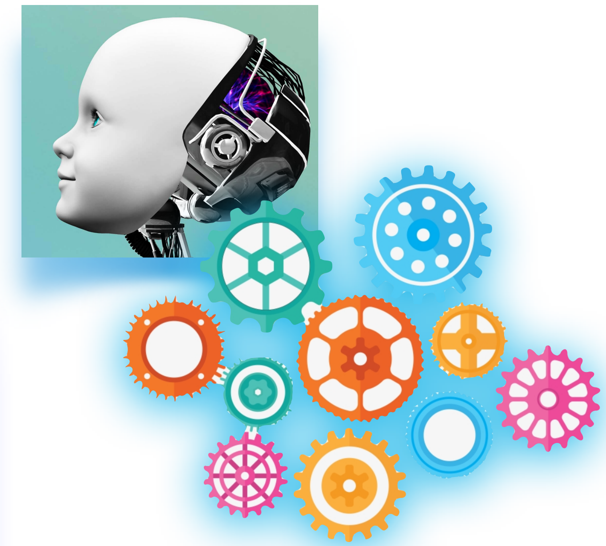
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

What did she say that he saw __ ?



Dickson et al. 2022, in prep.

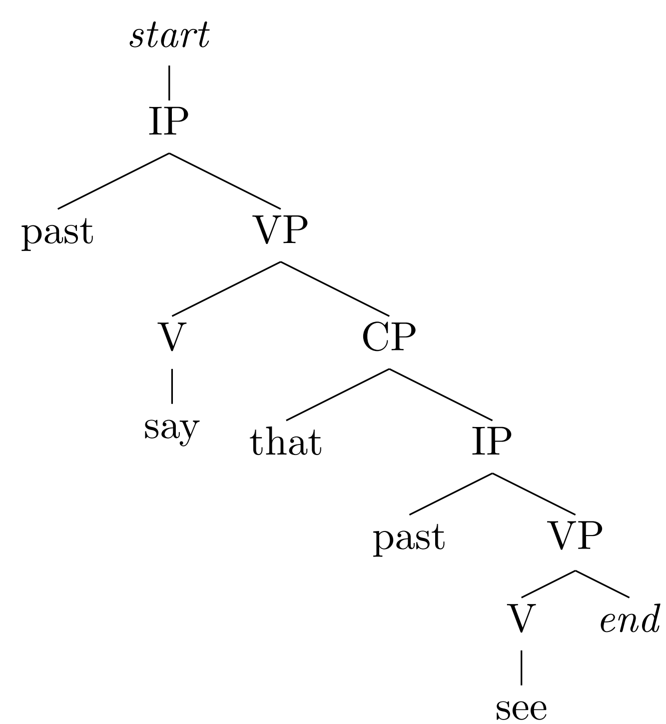
2: what the pieces are and their probabilities



learning efficient building blocks

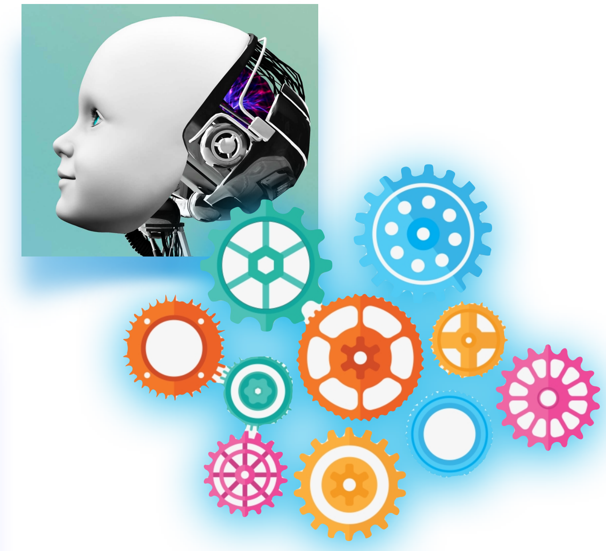
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

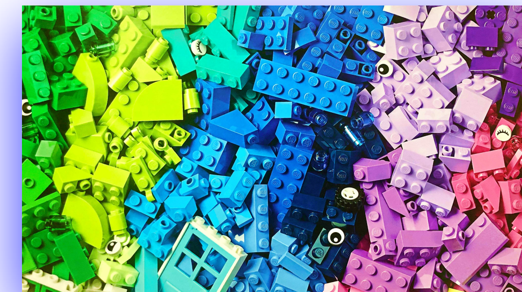


Dickson et al. 2022, *in prep.*

2: what the pieces are and their probabilities



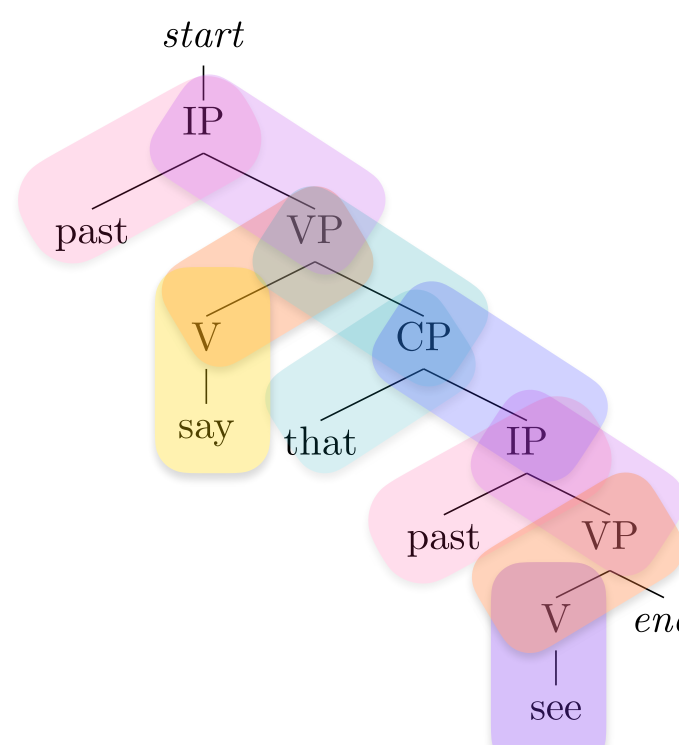
learning efficient building blocks



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

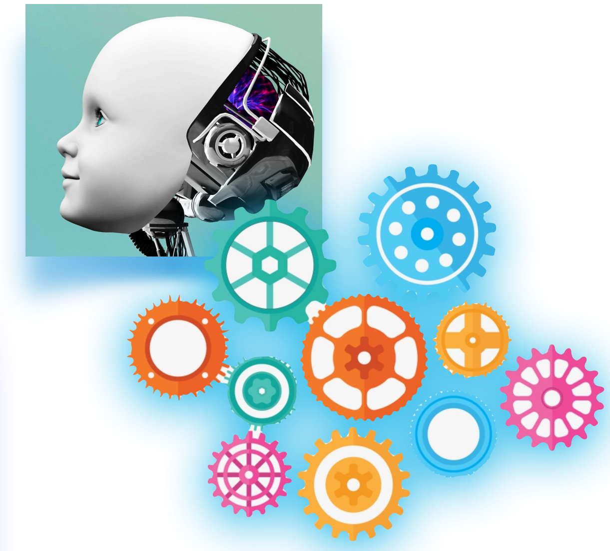
$\textit{start}\text{-IP}_{\text{past}}\text{-VP}_{\text{say}}\text{-CP}_{\text{that}}\text{-IP}_{\text{past}}\text{-VP}_{\text{see}}\text{-}\textit{end}$

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

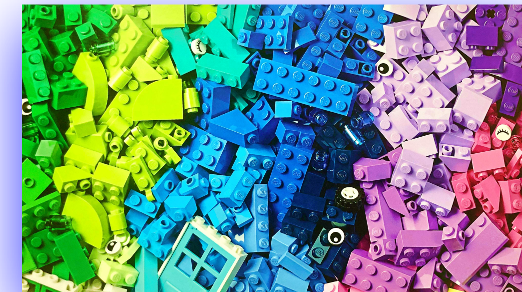


Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



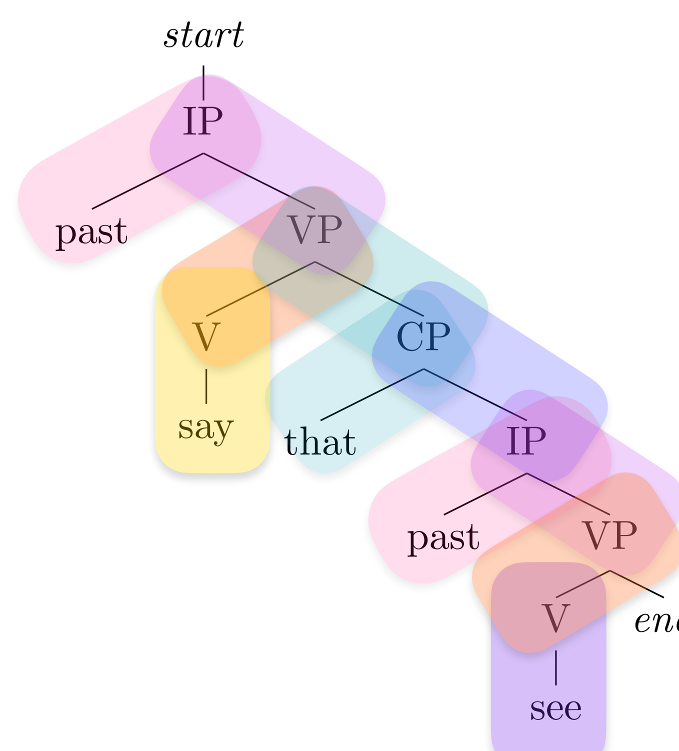
learning efficient building blocks



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

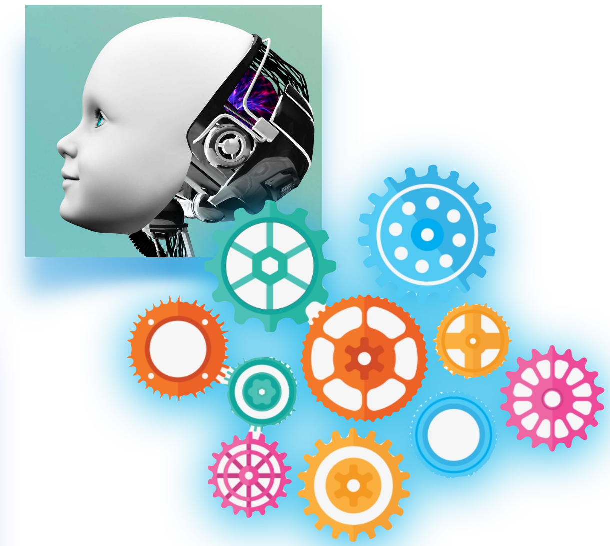
start-IP$_{\text{past}}$-VP$_{\text{say}}$-CP$_{\text{that}}$-IP$_{\text{past}}$-VP$_{\text{see}}$-*end*

It may be slower to put together many small pieces.

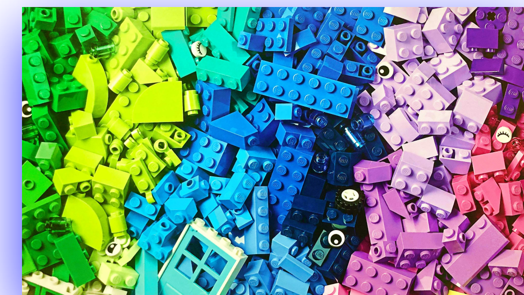


Dickson et al. 2022, *in prep.*

2: what the pieces are and their probabilities



learning efficient building blocks

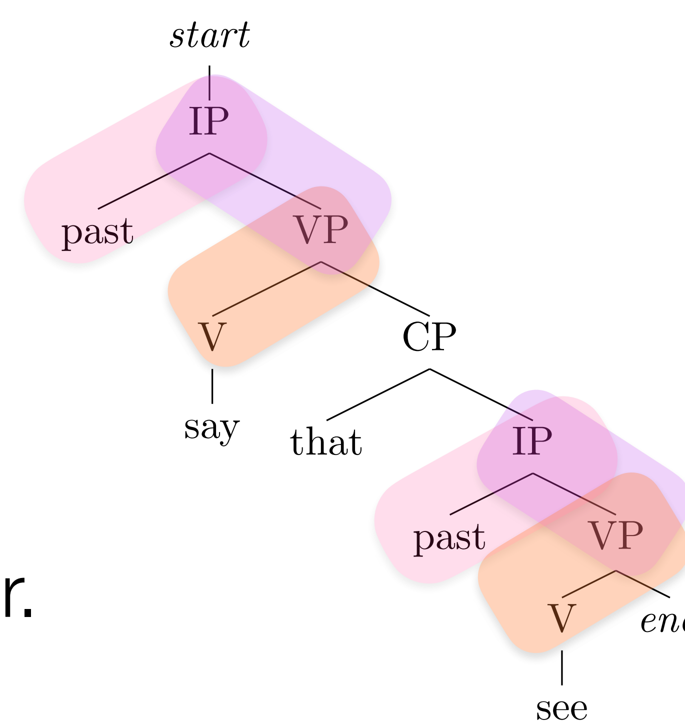


a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

many smaller

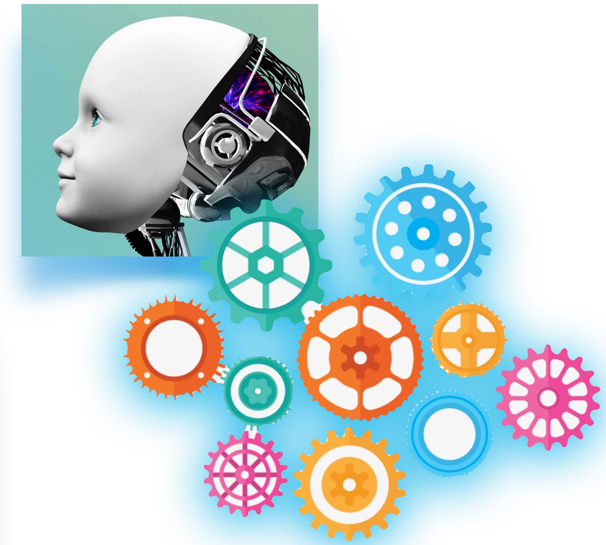
slower because many



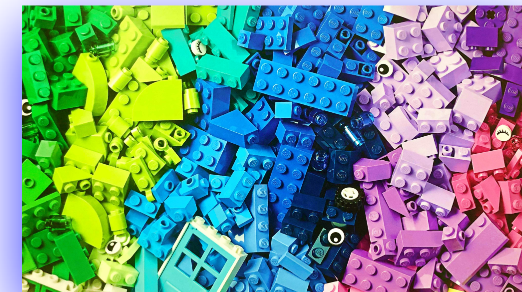
But these pieces may get reused, so that makes them faster to put together.

Dickson et al. 2022, *in prep.*

2: what the pieces are and their probabilities



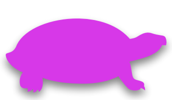
learning efficient building blocks



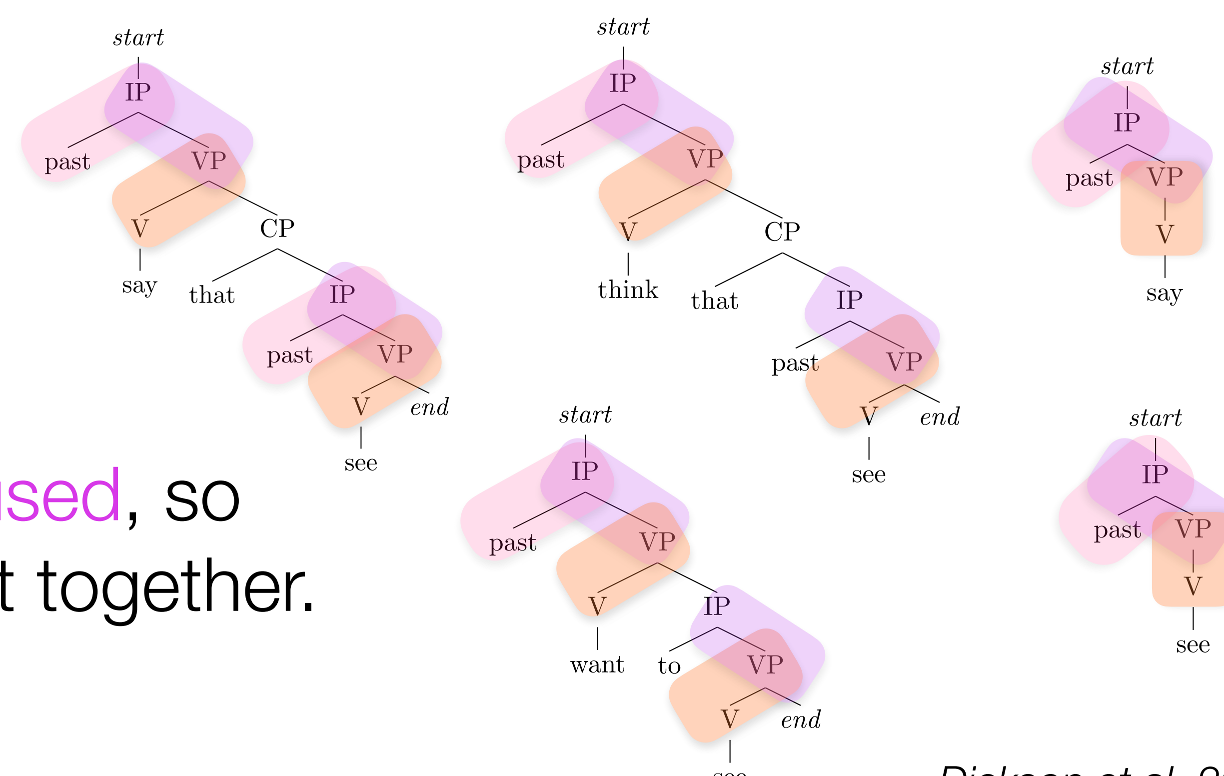
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

many smaller



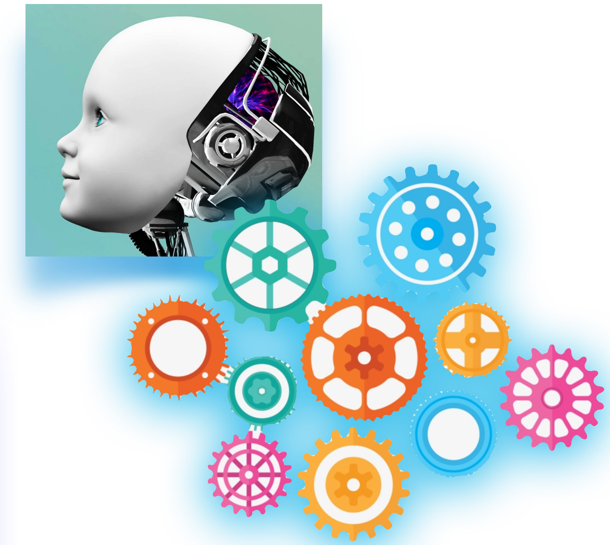
slower because many



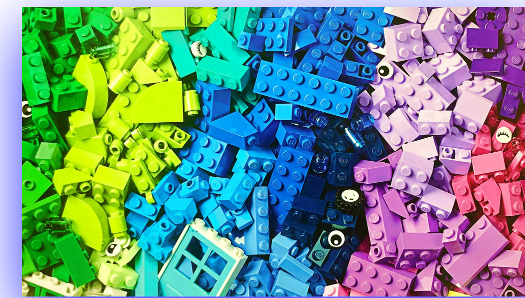
But these pieces may get reused, so that makes them faster to put together.

Dickson et al. 2022, *in prep.*

2: what the pieces are and their probabilities



learning efficient building blocks



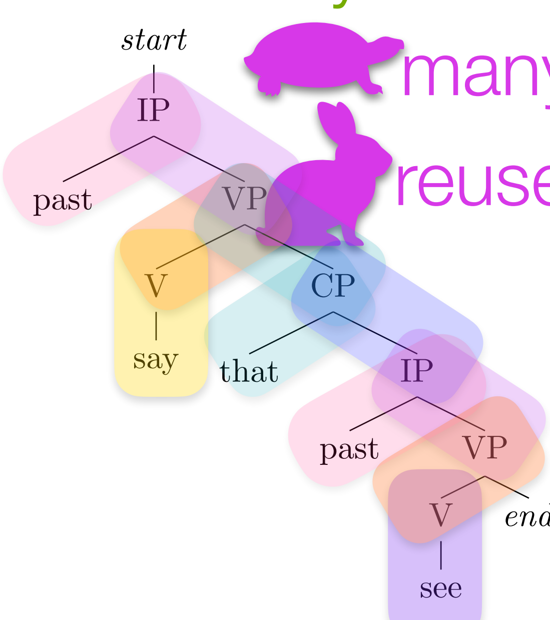
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

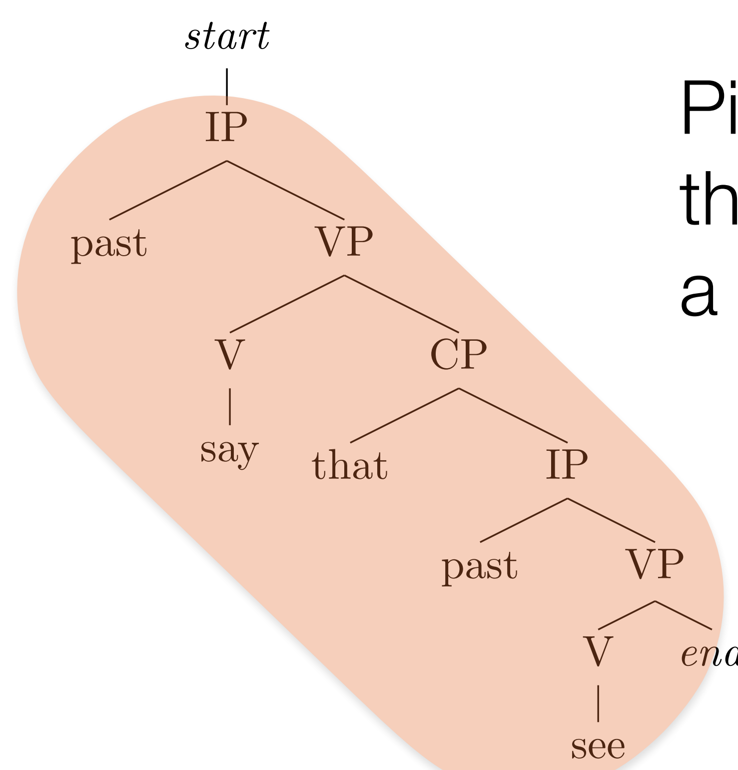
start-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller

many

reused

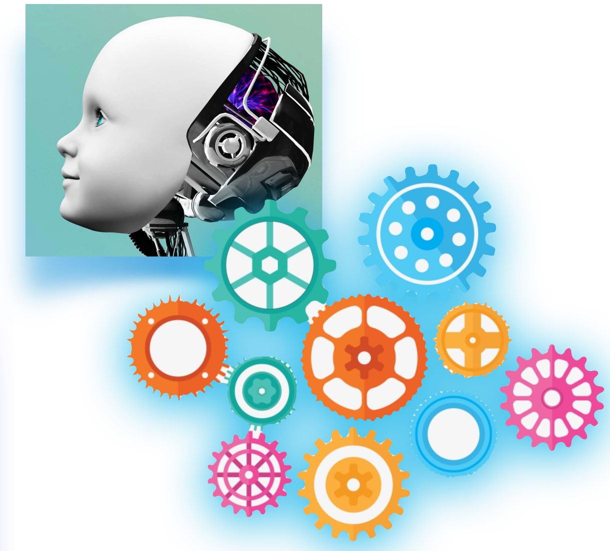




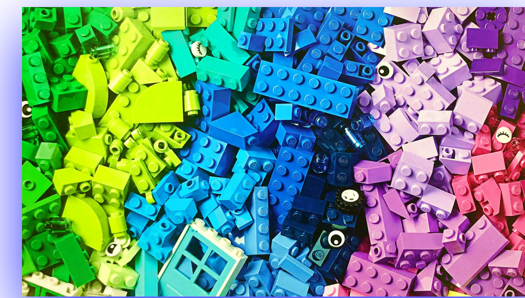
Pieces can be big, so that only one makes up a *wh*-dependency

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



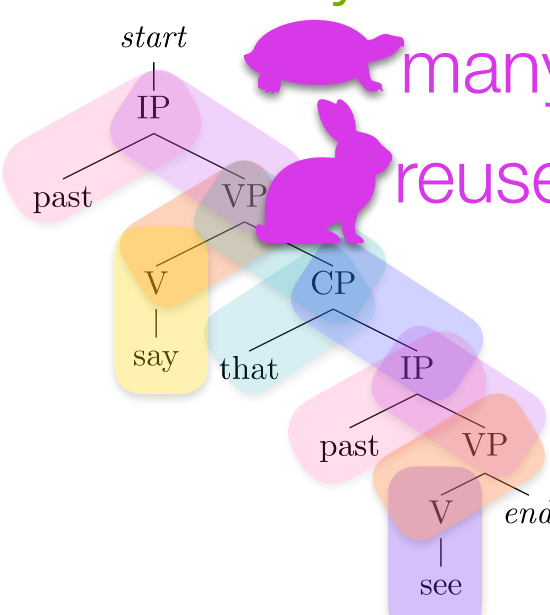
learning efficient building blocks

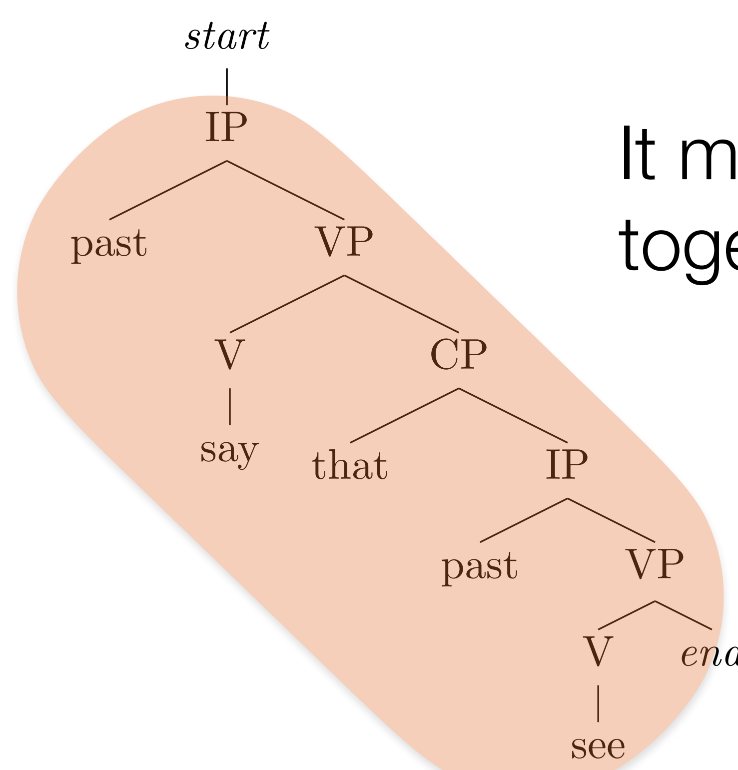


a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

many smaller
many
reused

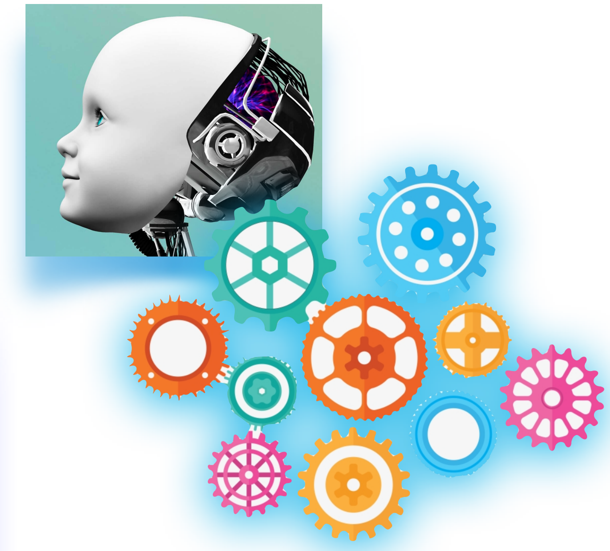




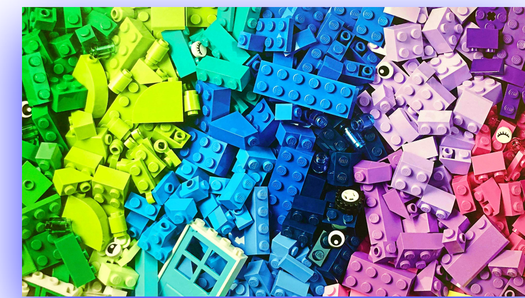
It may be faster to put together one big piece.

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



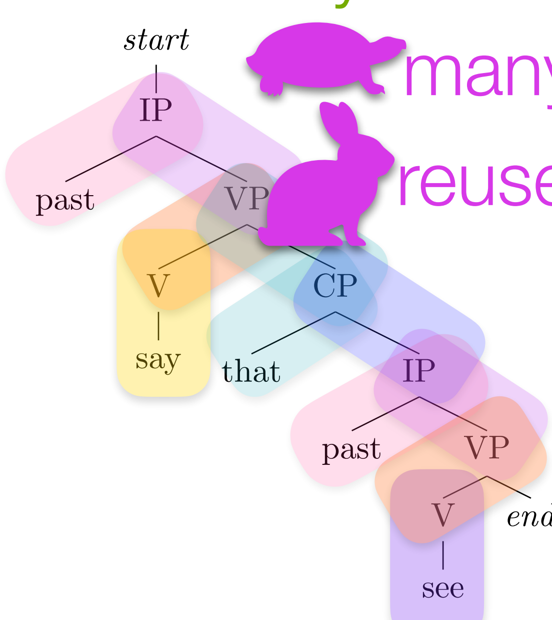
learning efficient building blocks



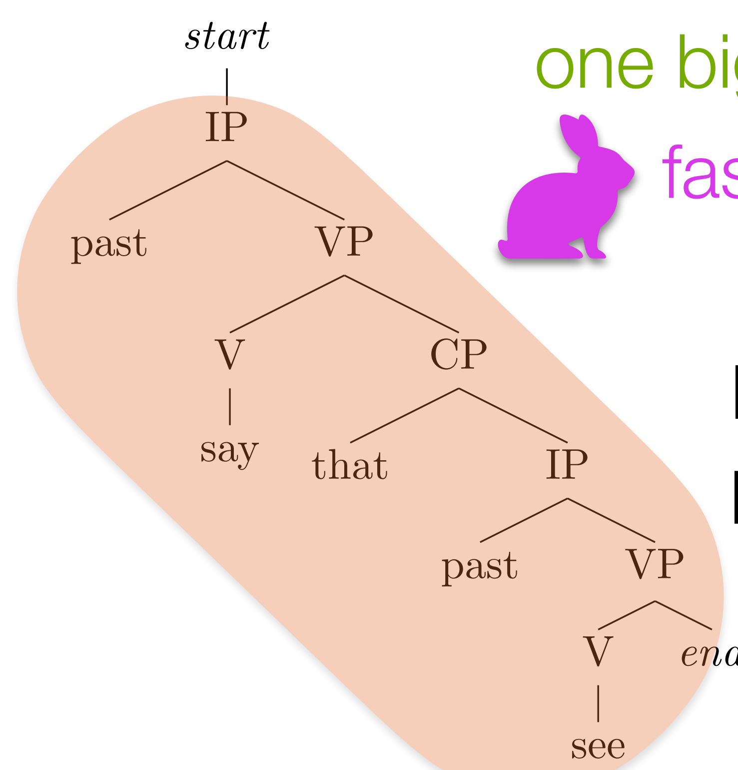
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

many smaller



many
reused

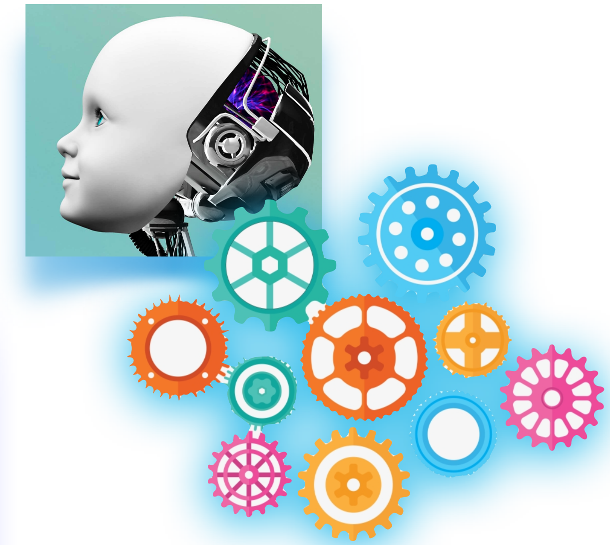


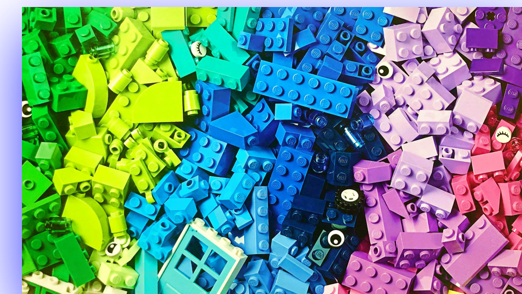
one big
faster because one

It may be slower if the piece is used rarely.

Dickson et al. 2022, *in prep.*

2: what the pieces are and their probabilities





learning efficient building blocks

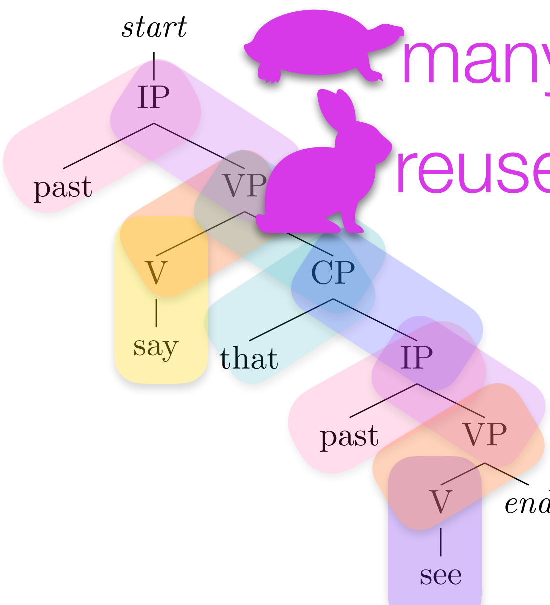
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

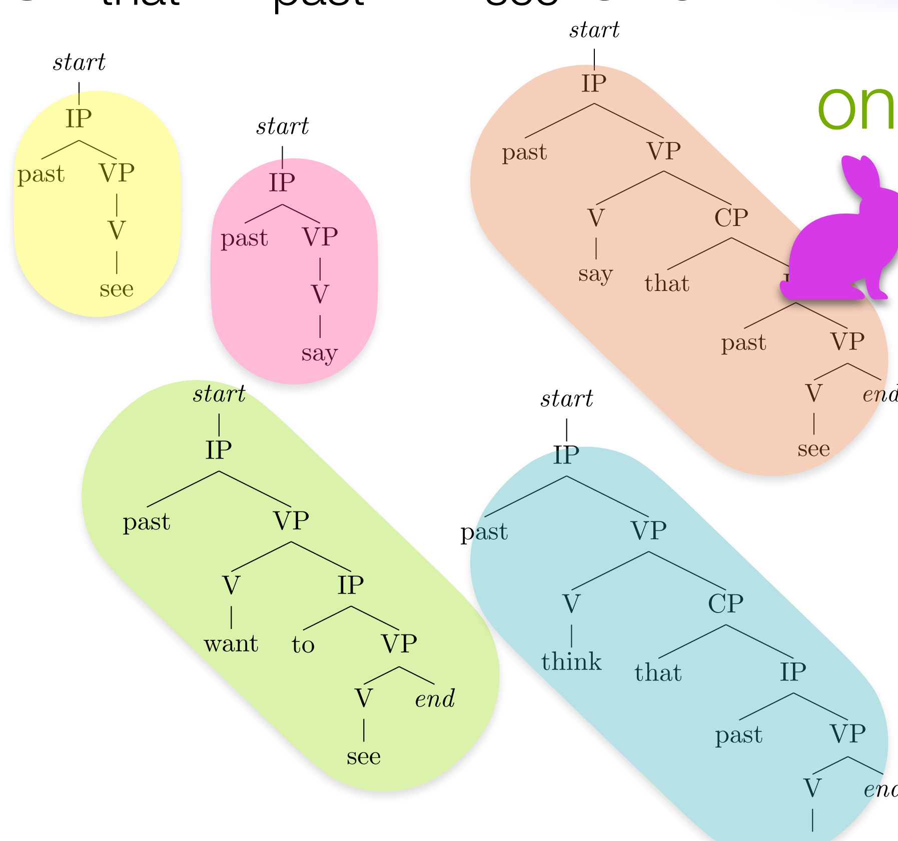
start-IP$_{\text{past}}$-VP$_{\text{say}}$-CP$_{\text{that}}$-IP$_{\text{past}}$-VP$_{\text{see}}$-*end*

many smaller

many

reused





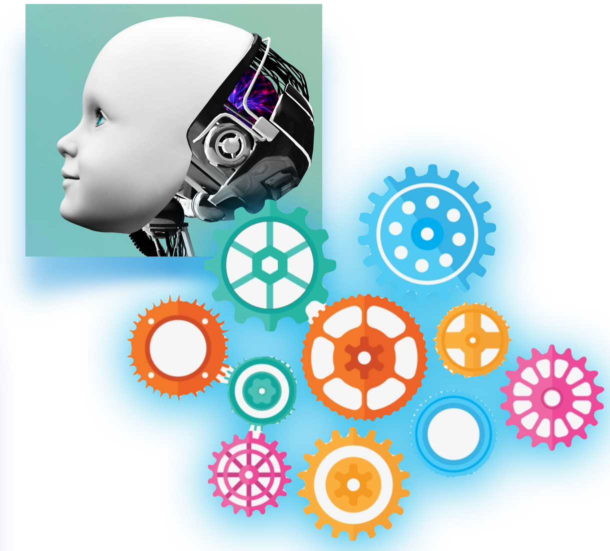
one big

faster because one

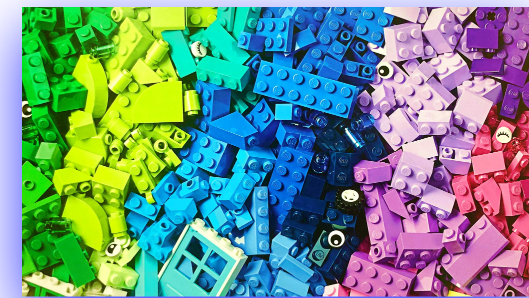
It may be slower if the piece is used rarely.

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



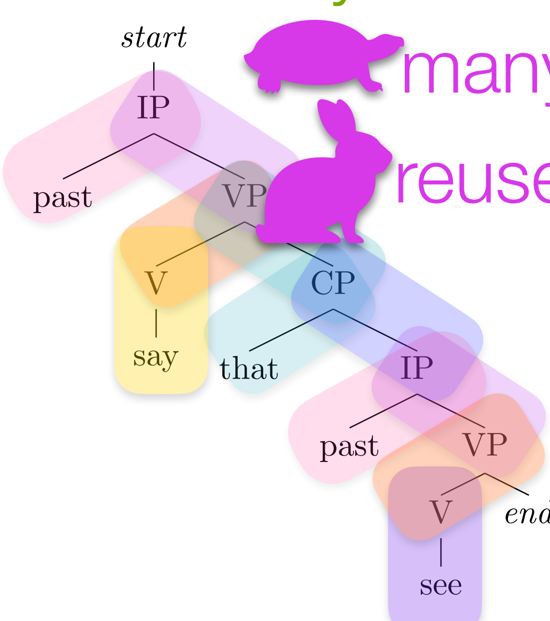
learning efficient building blocks



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

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

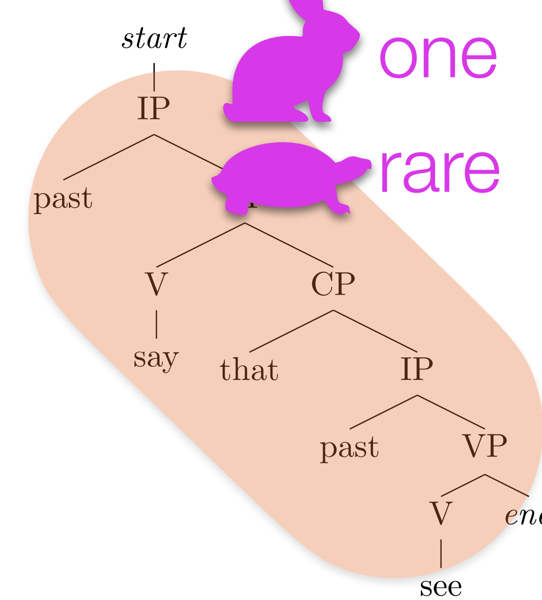
many smaller



many
reused

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

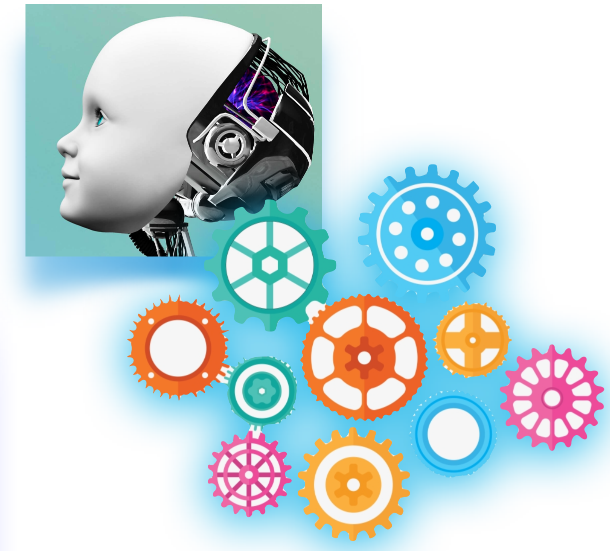
one big



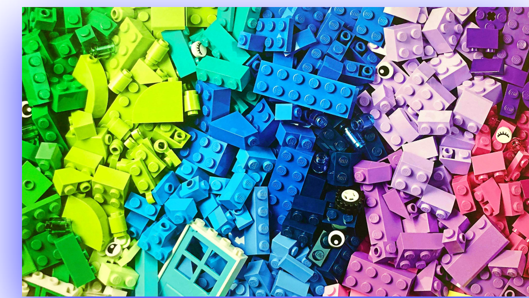
one
rare

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



learning efficient building blocks



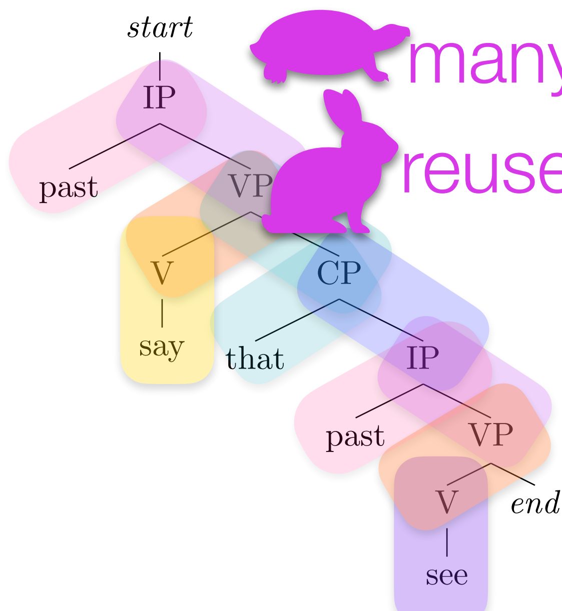
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

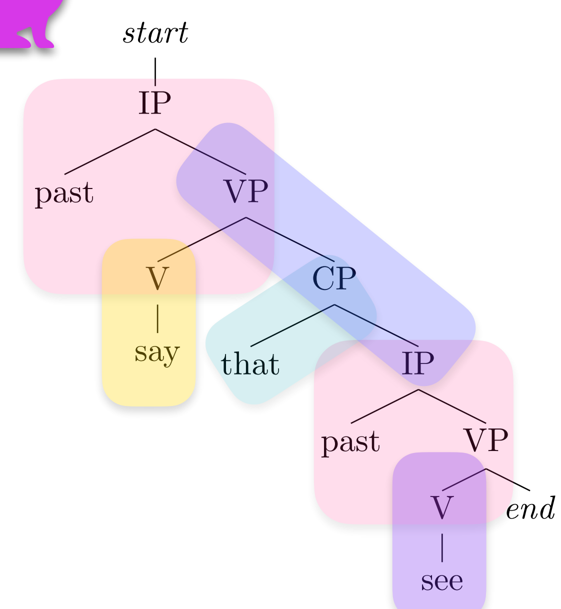
start-IP$_{\text{past}}$-VP$_{\text{say}}$-CP$_{\text{that}}$-IP$_{\text{past}}$-VP$_{\text{see}}$-*end*

many smaller

many

reused

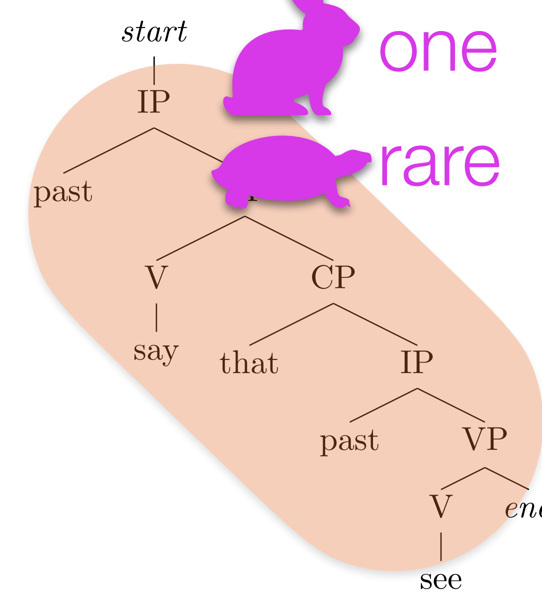




one big

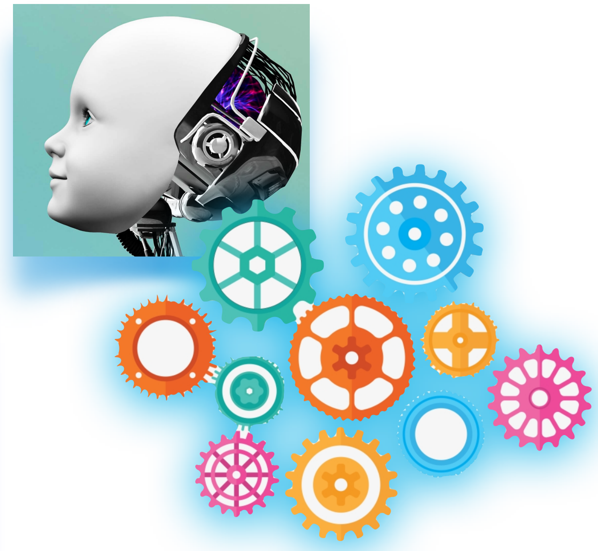
one

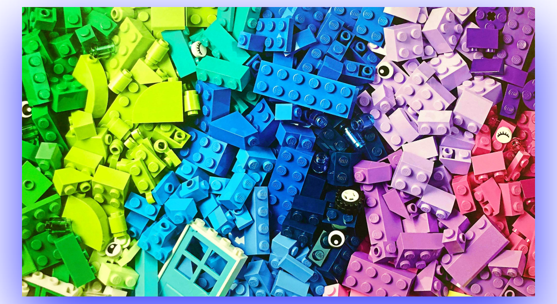
rare



Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

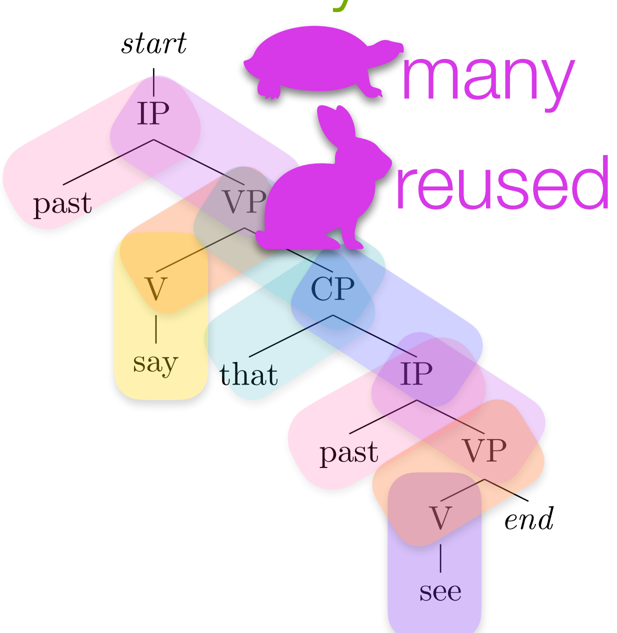


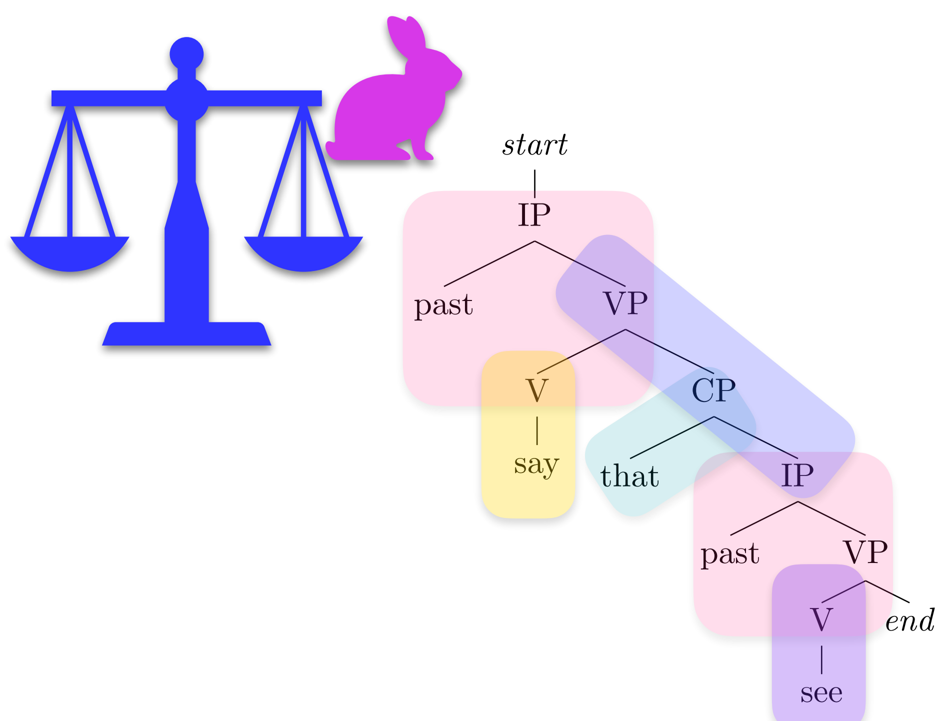


learning efficient building blocks

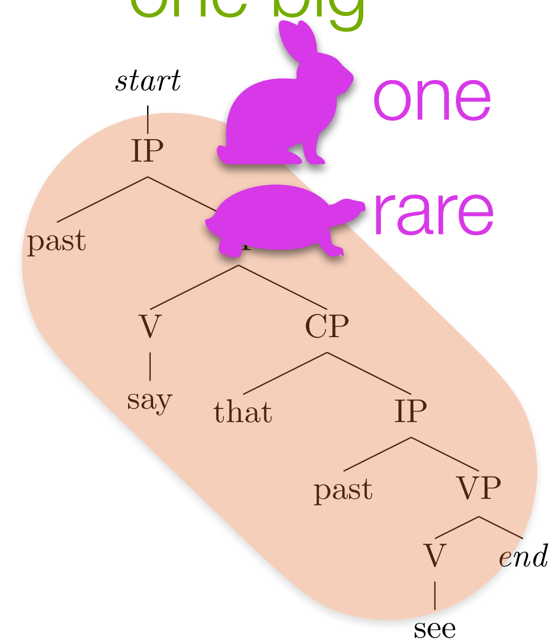
How can children find the best balance?

many smaller



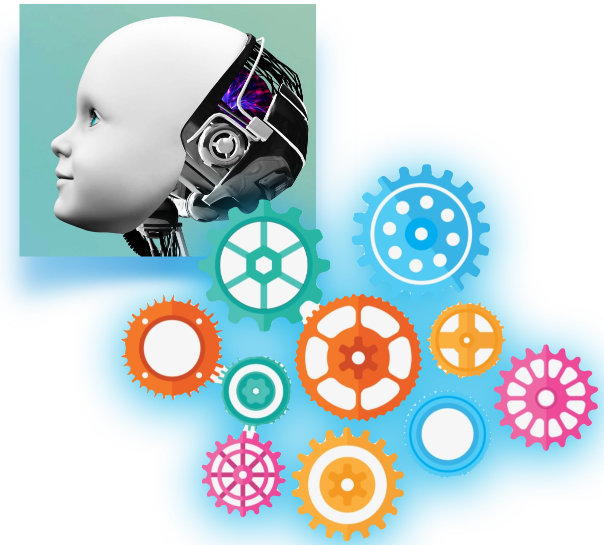


one big



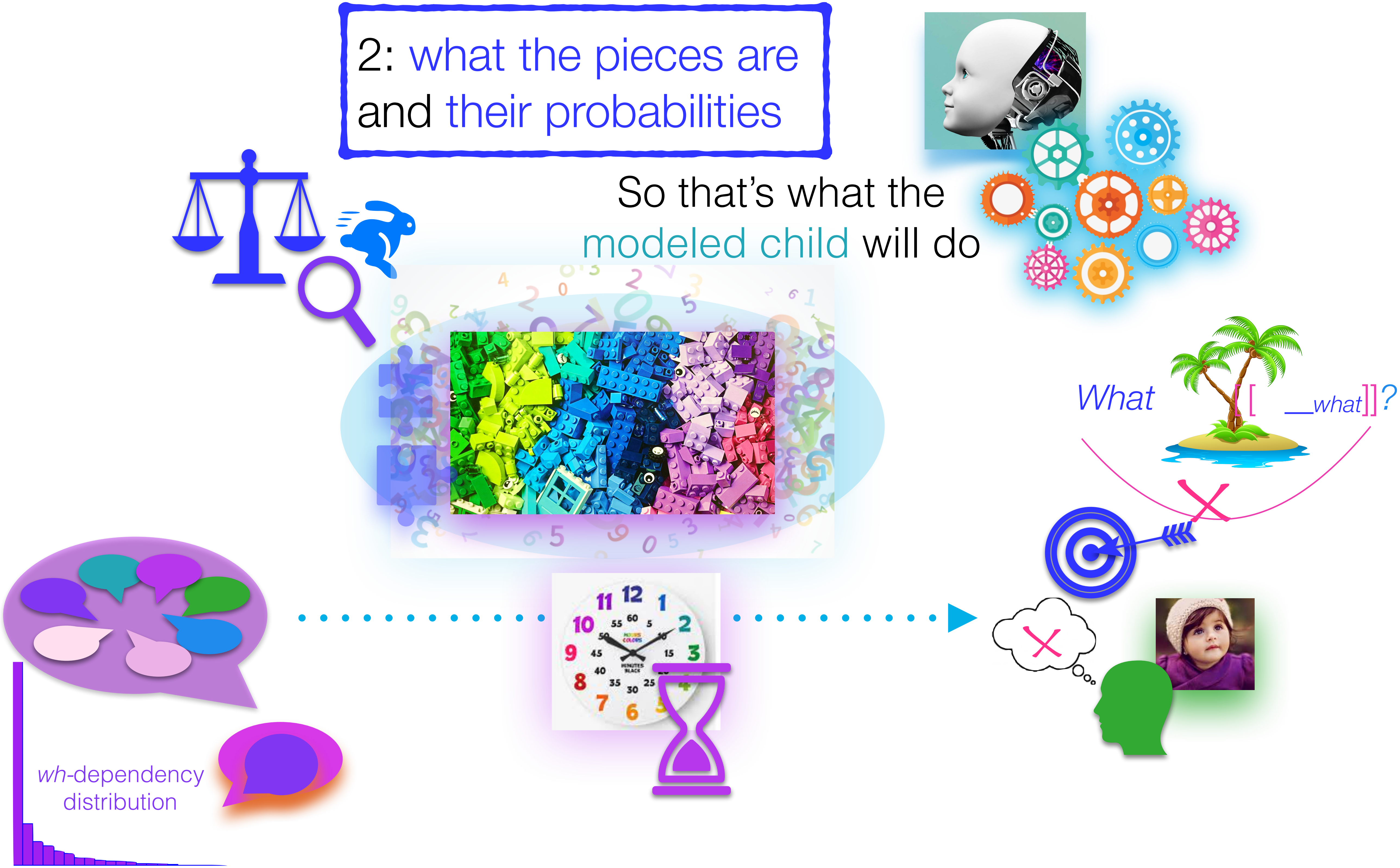
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



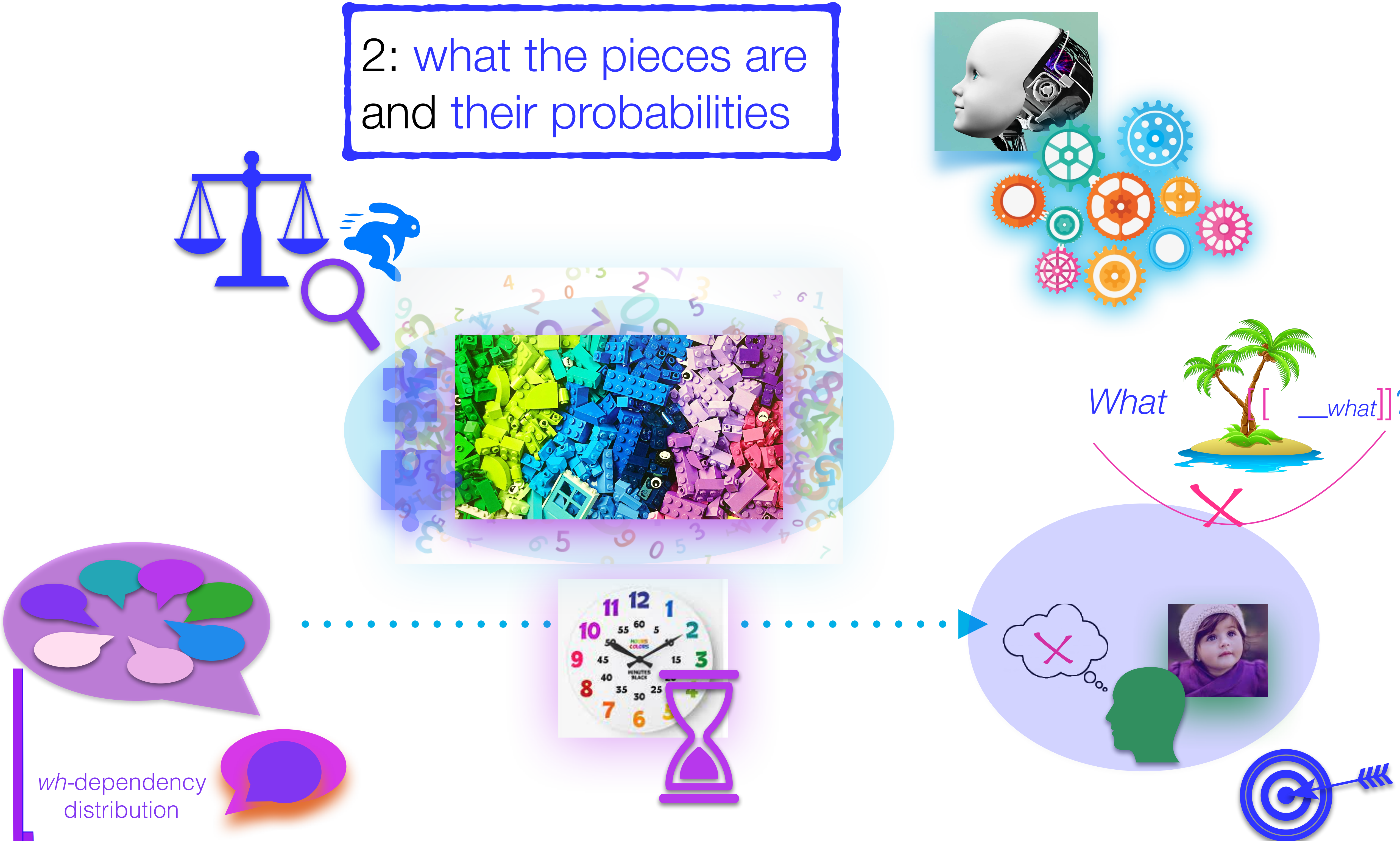
Use Bayesian inference to search through the hypothesis space of all possible building blocks (O'Donnell 2015) and find an efficient set for children's input.

Dickson et al. 2022, in prep.



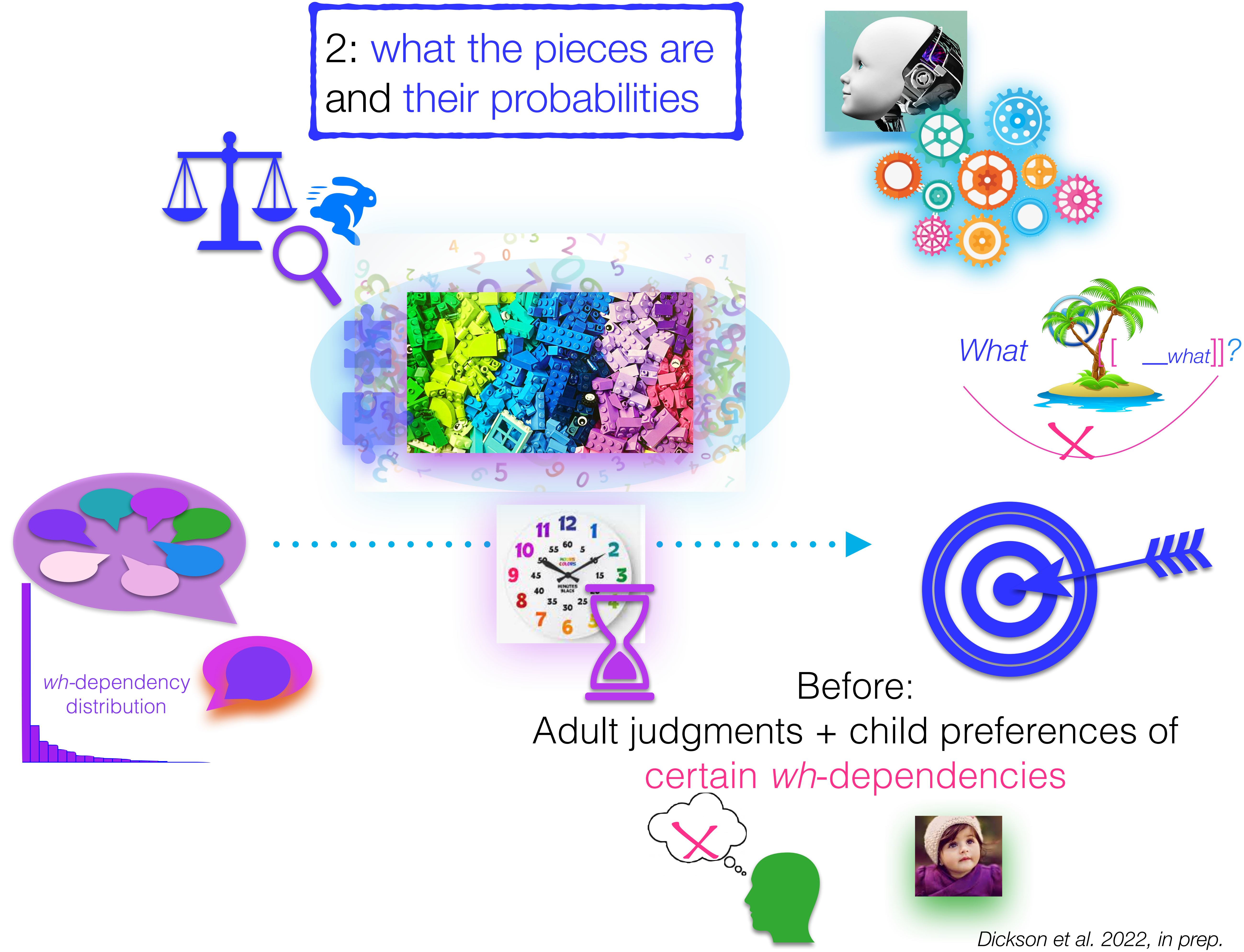
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



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

Dickson et al. 2022, in prep.


2: what the pieces are and their probabilities
What [__what]]?
wh-dependency distribution
Before:
Adult judgments + child preferences of certain wh-dependencies
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

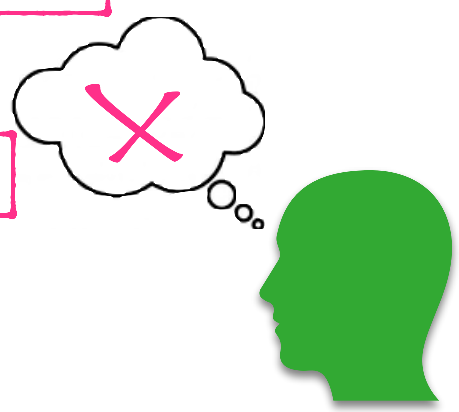
certain *wh*-dependencies

Subject island

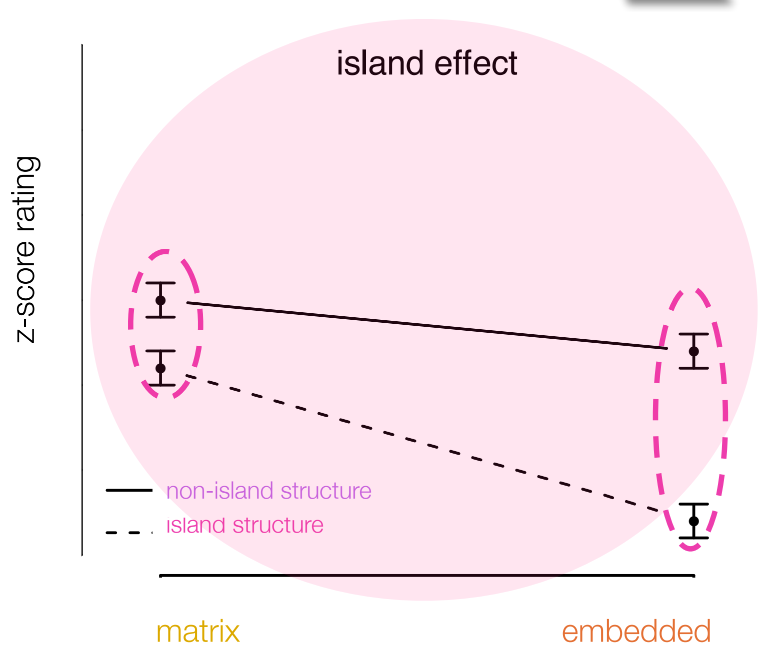
Complex NP island

Whether island

Adjunct island



Before: Target behavior

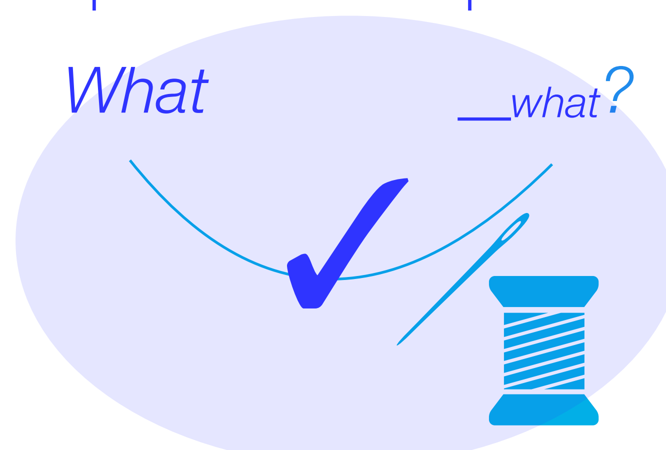


Sprouse et al. 2012

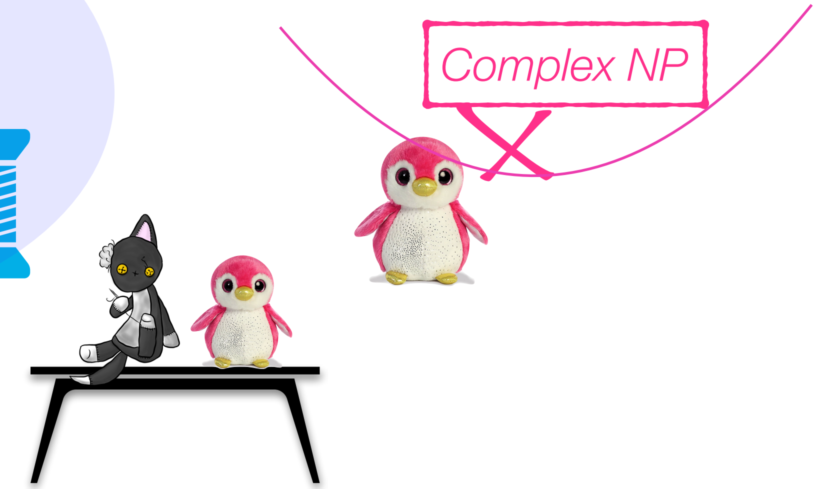
Bates et al. in prep.

certain *wh*-dependencies

Children prefer this interpretation.



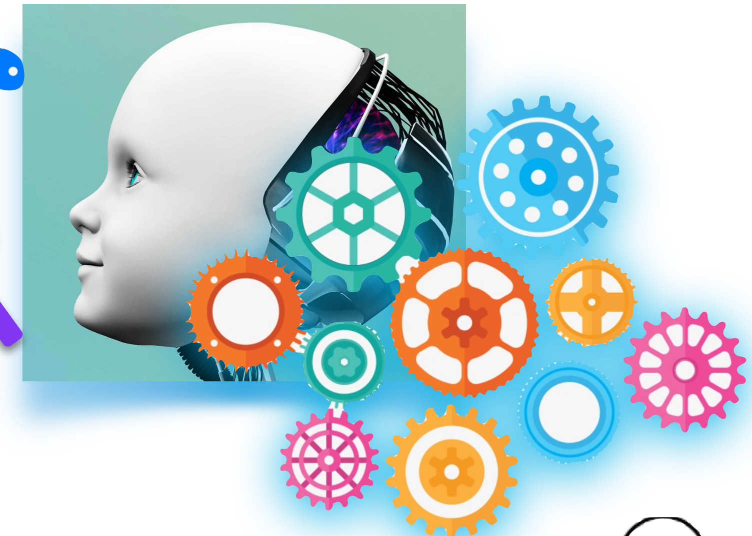
What [NP [CP __*what*]]?



De Villiers et al. 2008

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



certain *wh*-dependencies

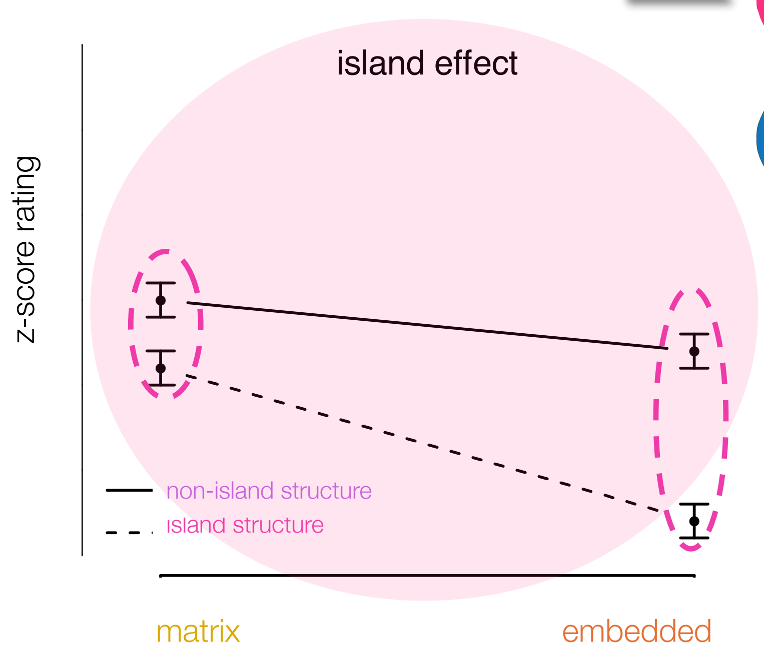
Subject island

Complex NP island

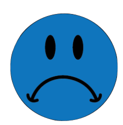
Whether island

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 __]?

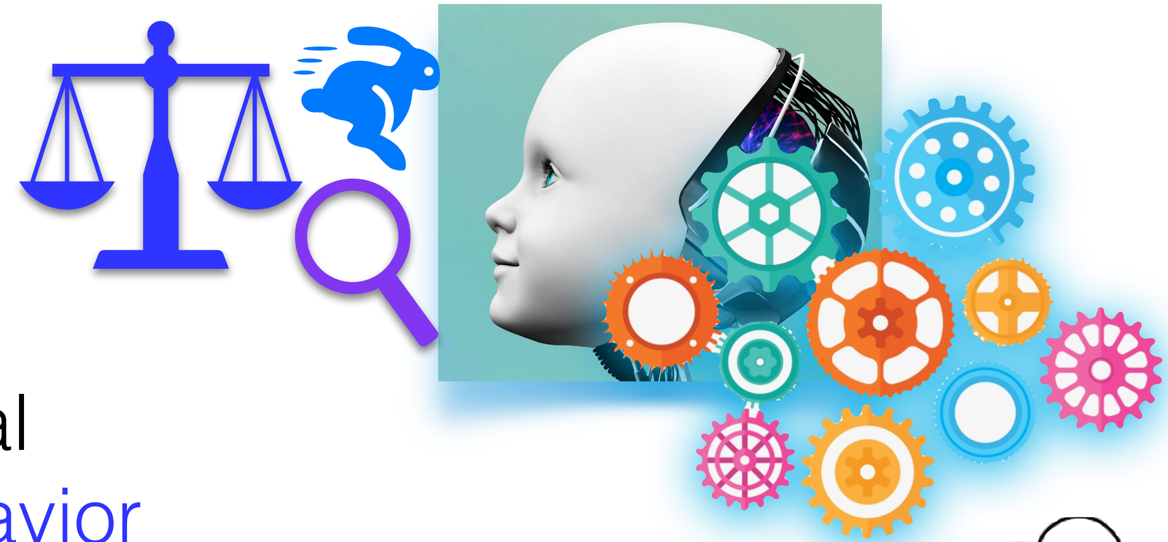


What did she whine [that he saw __]?
What did she mumble [that he saw __]?

Liu et al. 2019

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



certain *wh*-dependencies

Subject island

Complex NP island

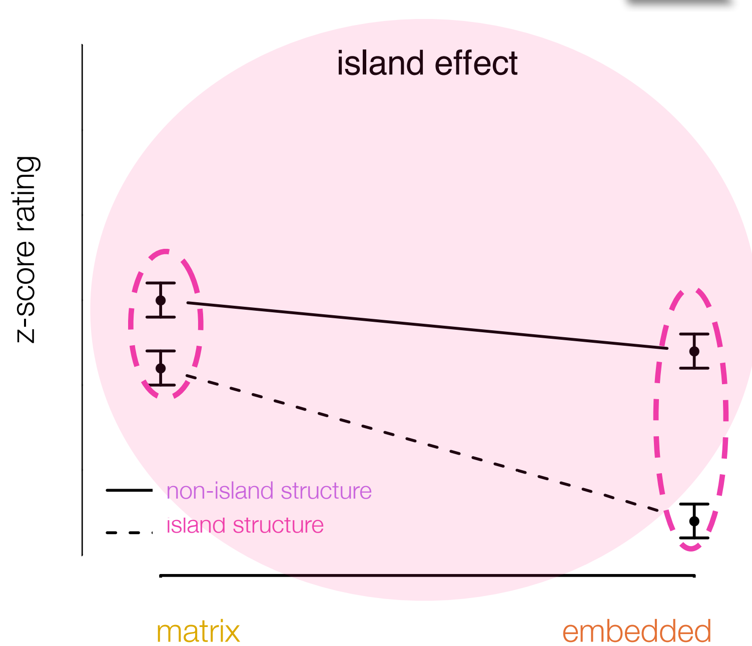
Whether island

Adjunct island

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

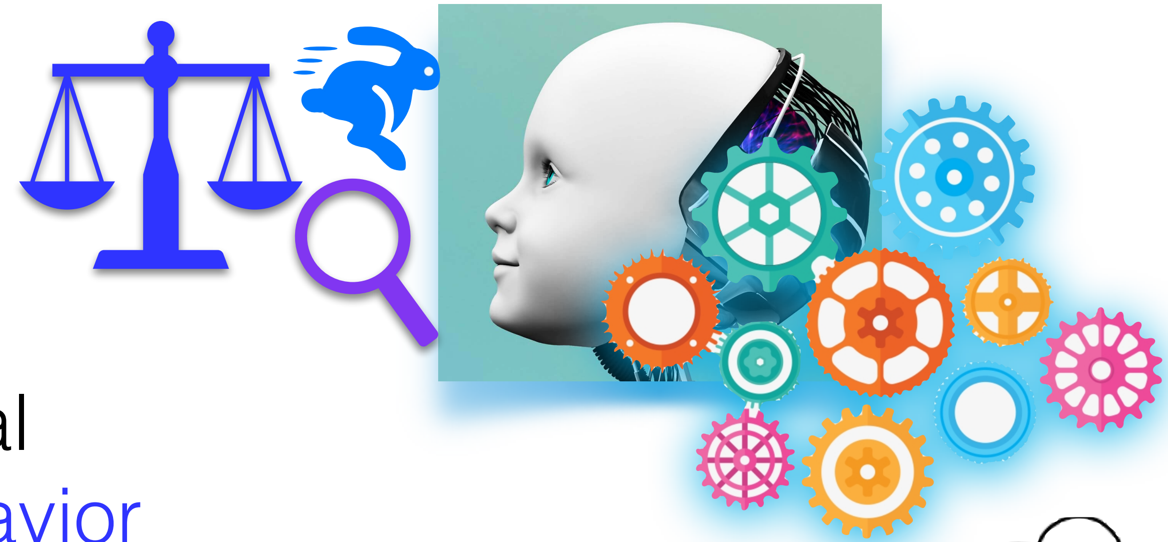
What did she VERB [that he saw __]?

Liu et al. 2019



Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



certain *wh*-dependencies

Subject island

Complex NP island

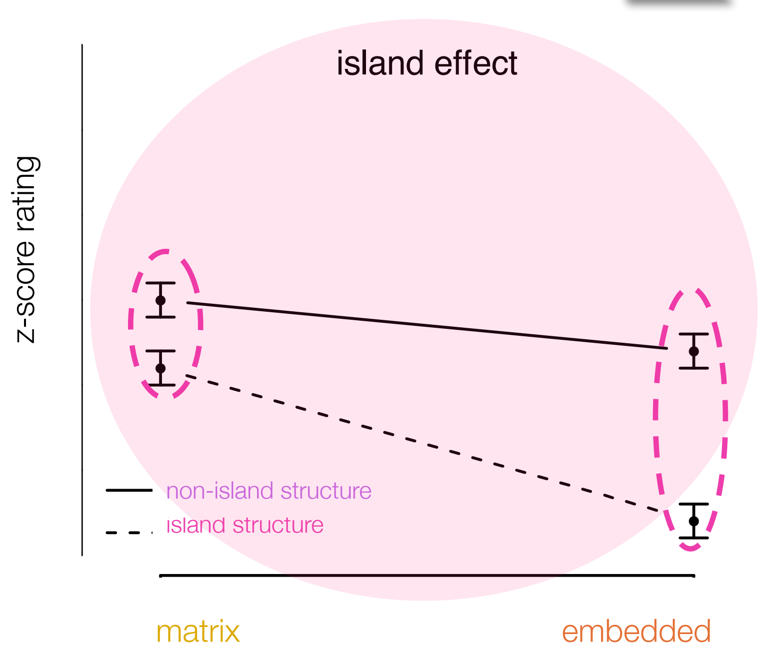
Whether island

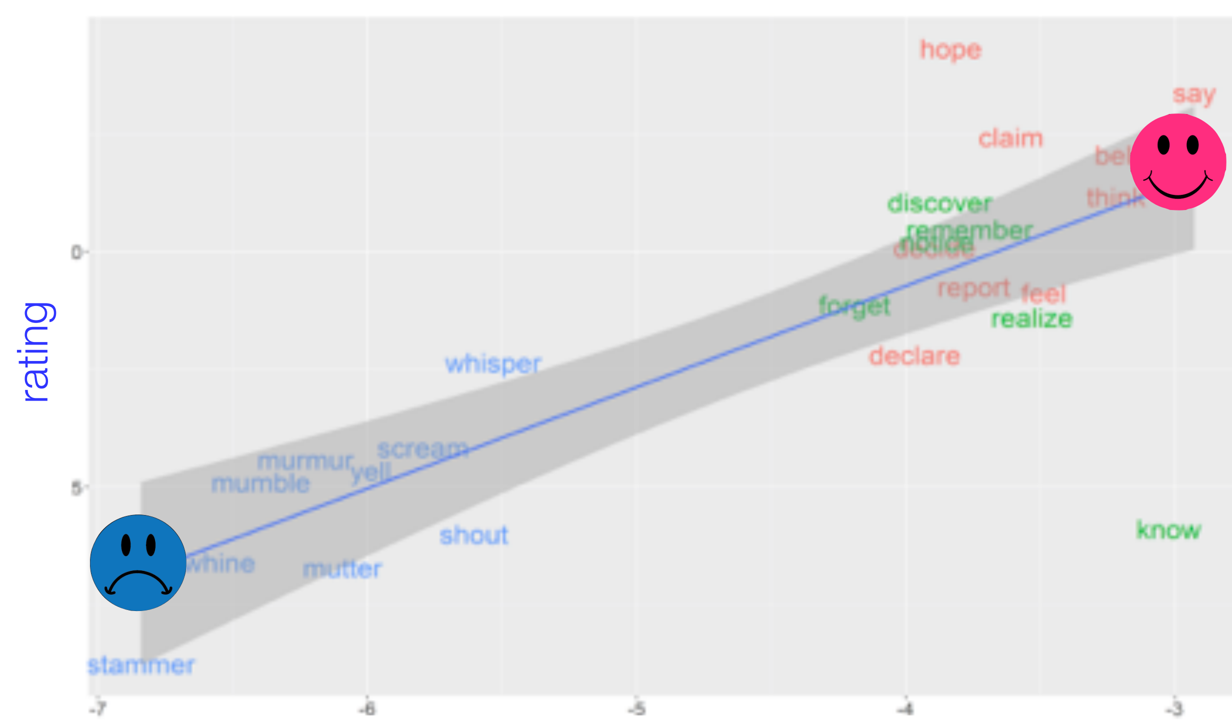
Adjunct island

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

What did she VERB [that he saw __]?

Liu et al. 2019

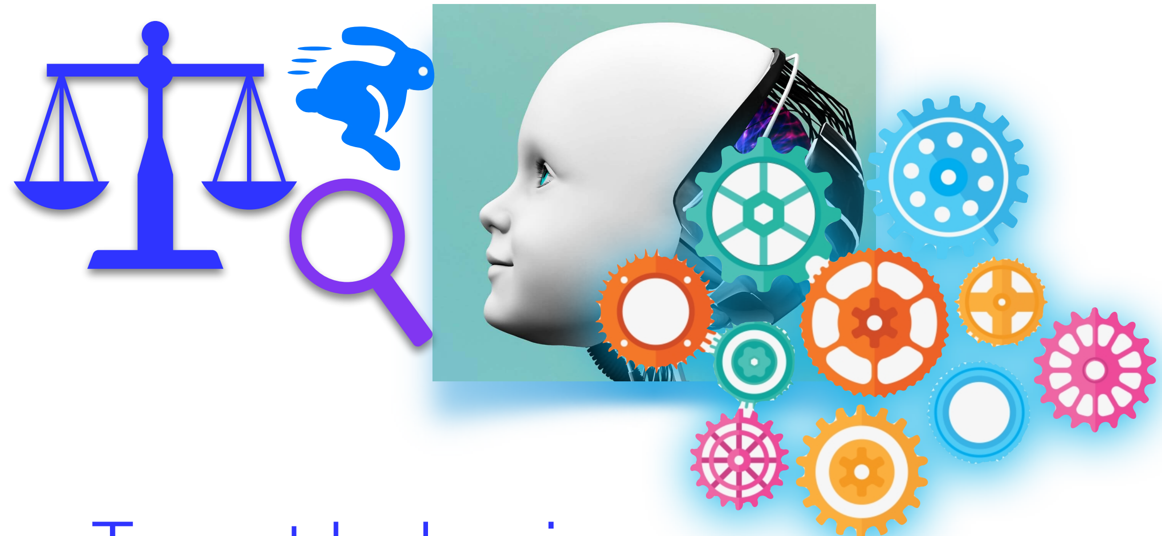




Important pattern:
Positive correlation between main verb frequency and judged acceptability.

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



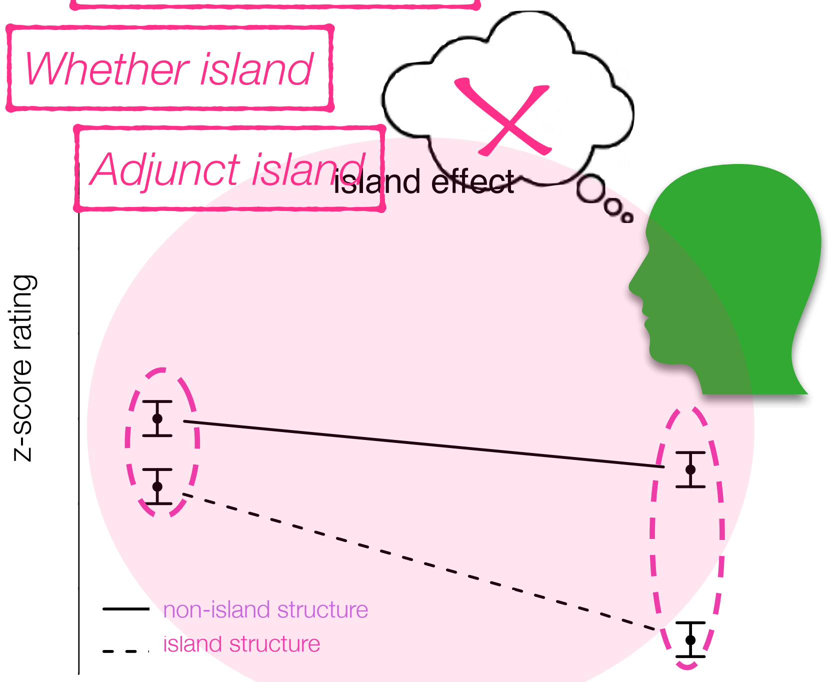
Before: Target behavior

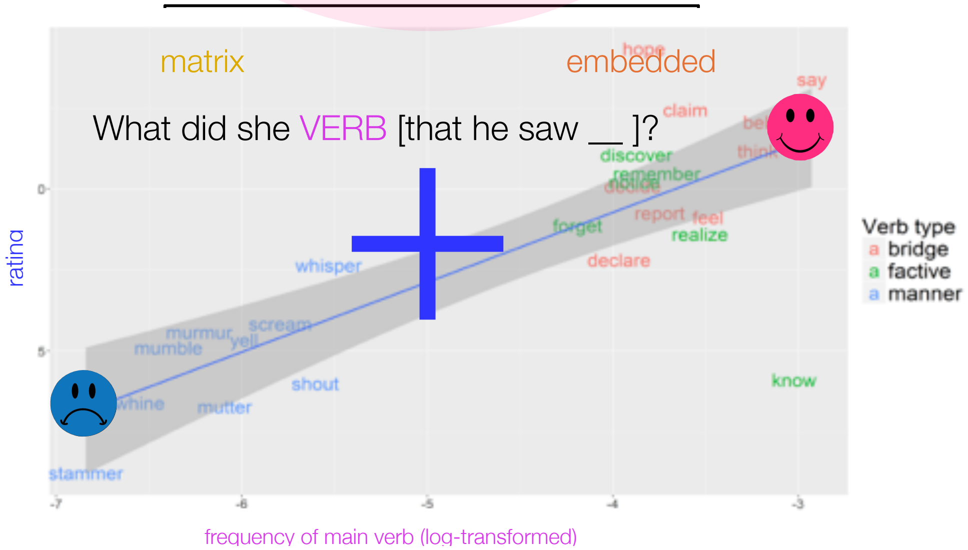
Subject island

Complex NP island

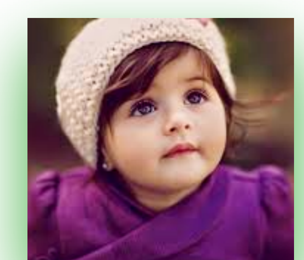
Whether island

Adjunct island



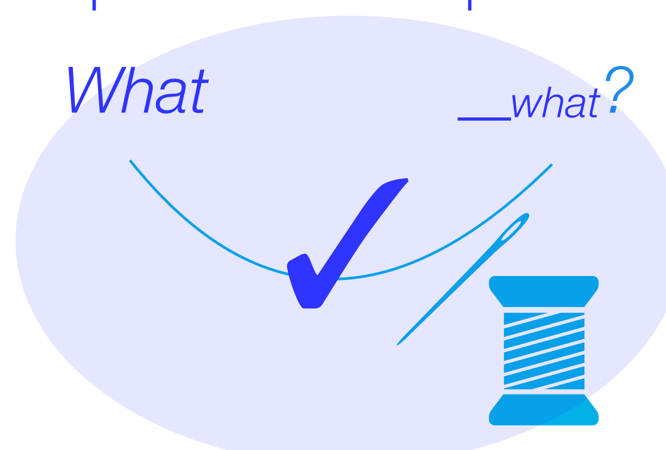


certain *wh*-dependencies



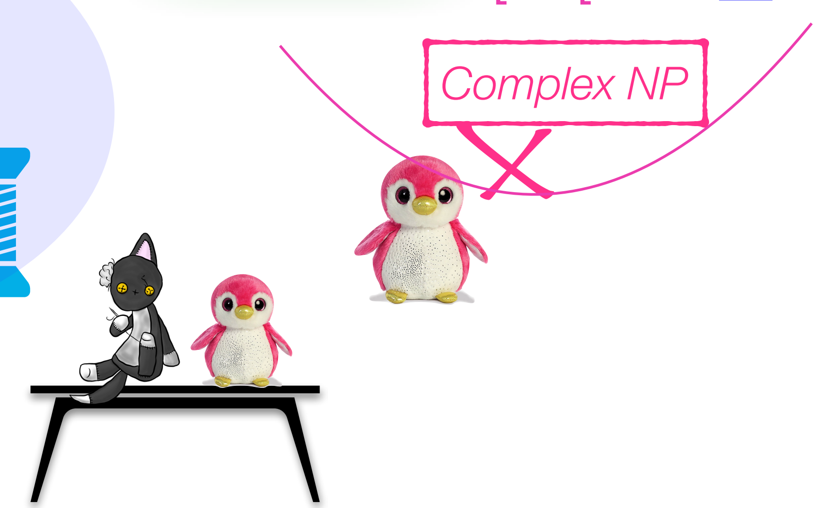
Children prefer this interpretation.

What *__what*?



What [NP [CP *__what*]]?

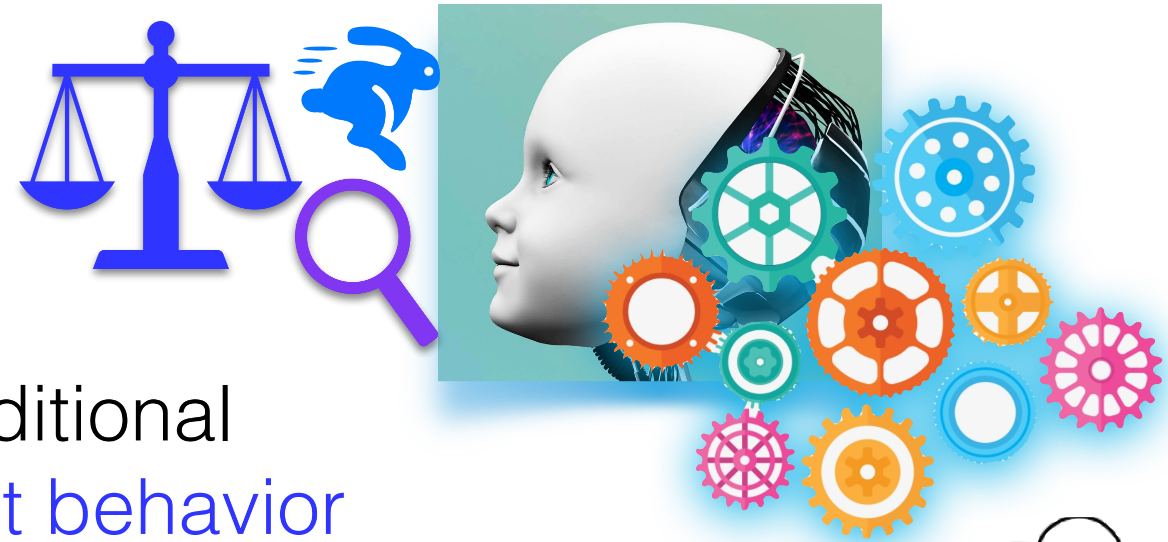
Complex NP



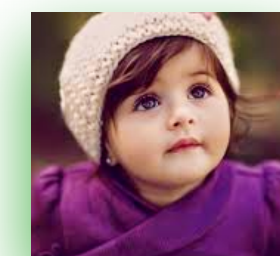
De Villiers et al. 2008

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



+ 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?

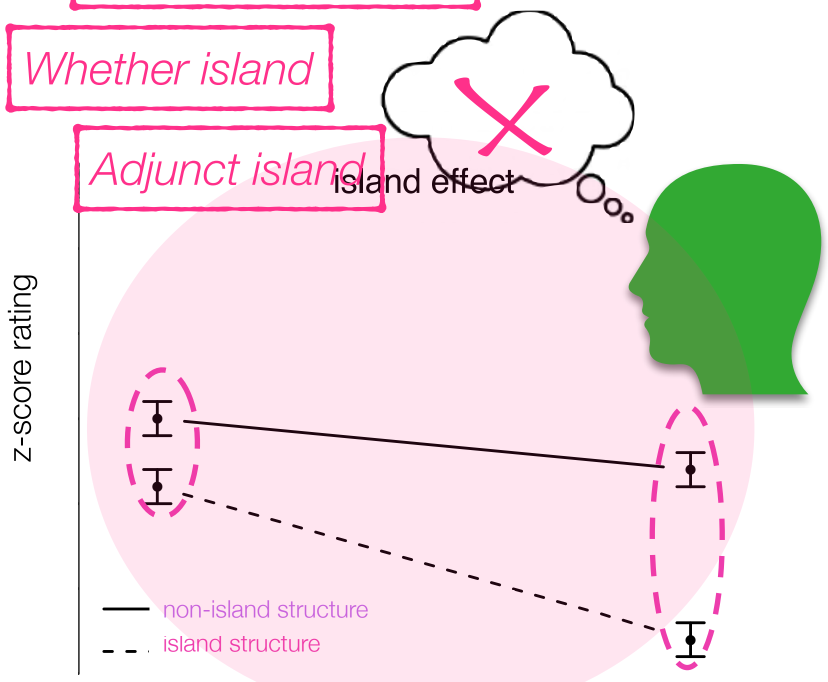
How did the boy who sneezed drink the milk?

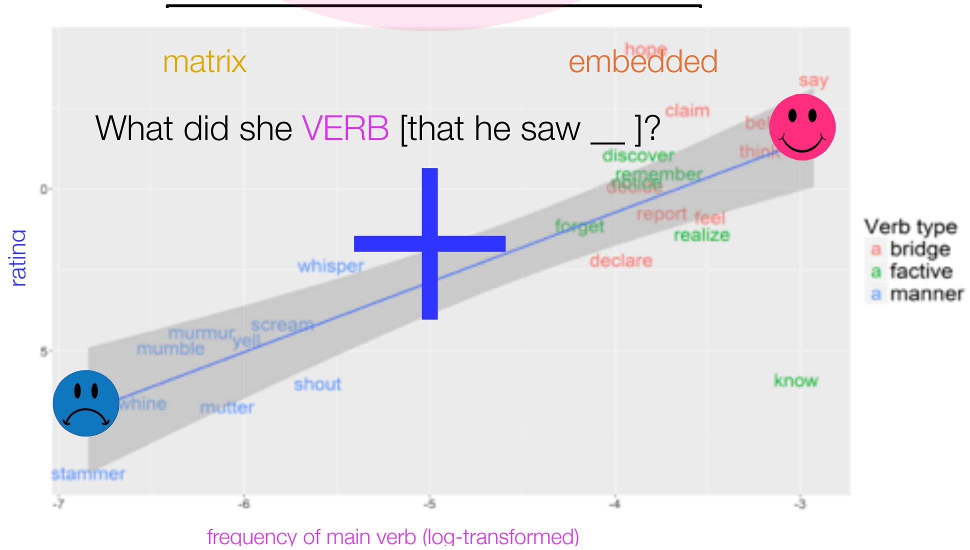
Subject island

Complex NP island

Whether island

Adjunct island





Who ... *__who*?
How ... *__how*?

✓

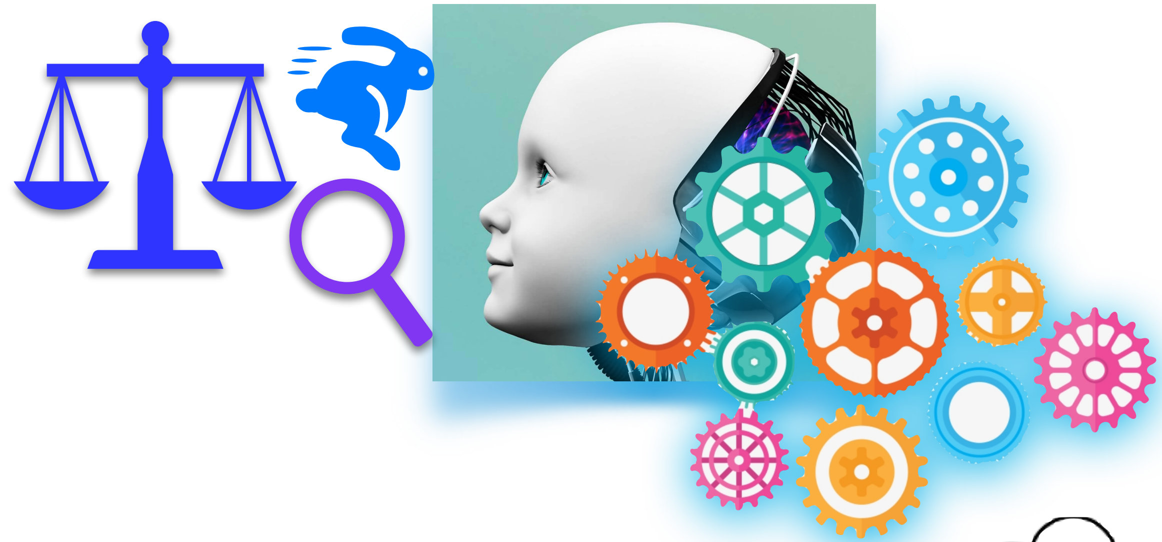
Who [CP-how *__who*]?
Who [CP-what *__who*]?
How [CP-what *__how*]?
How [CP-where *__how*]?
How [NP [CP-who *__how*]]?

✗

De Villiers et al. 2008

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

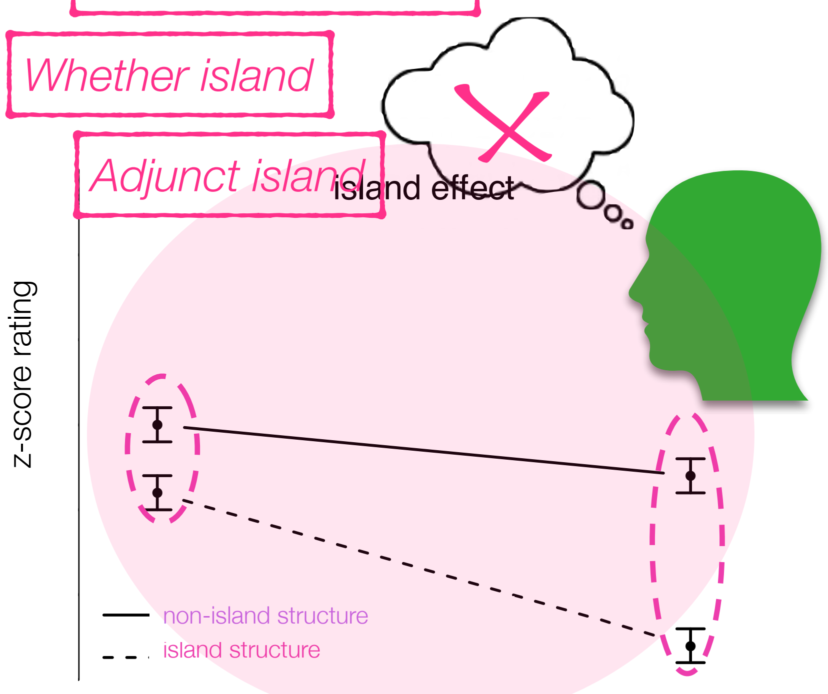


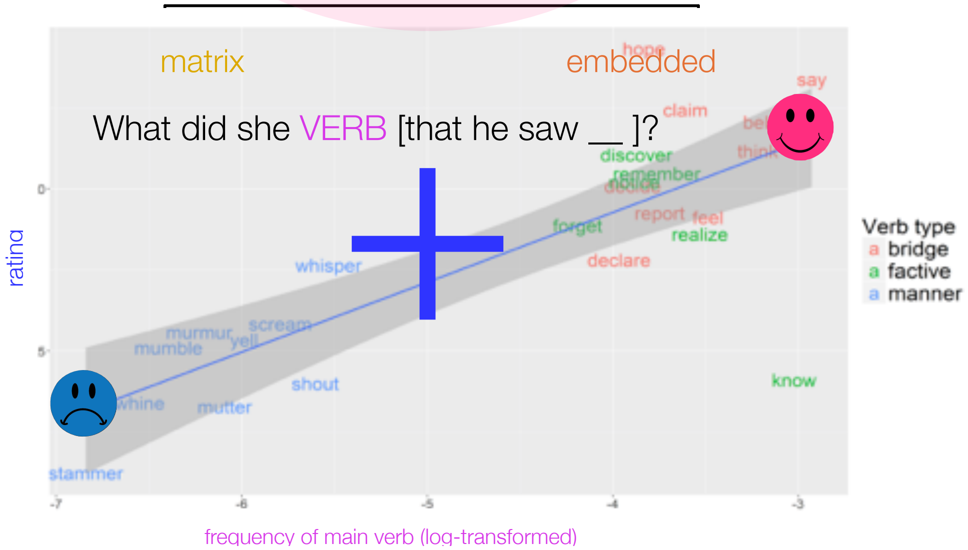
Subject island

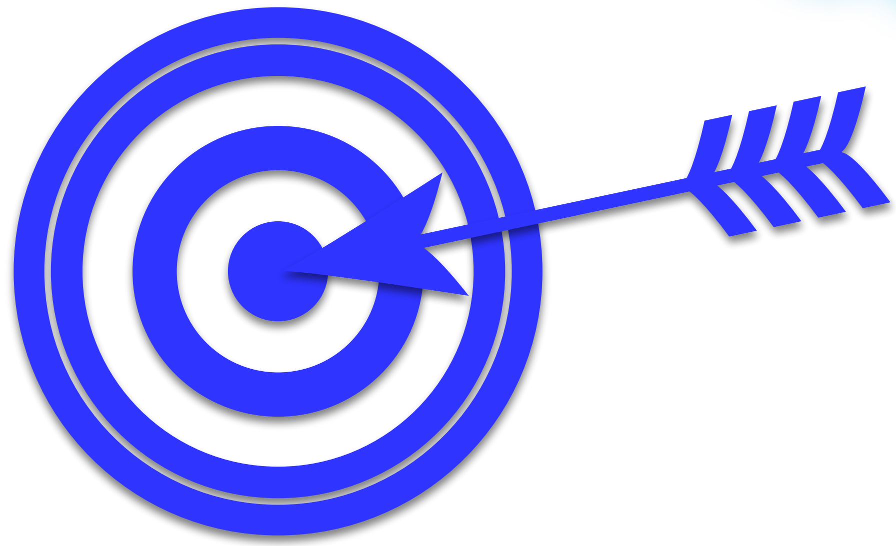
Complex NP island

Whether island

Adjunct island







Who *__who*?

How *__how*?

Who [CP-how *__who*]?

Who [CP-what *__who*]?

How [CP-what *__how*]?

How [CP-where *__how*]?

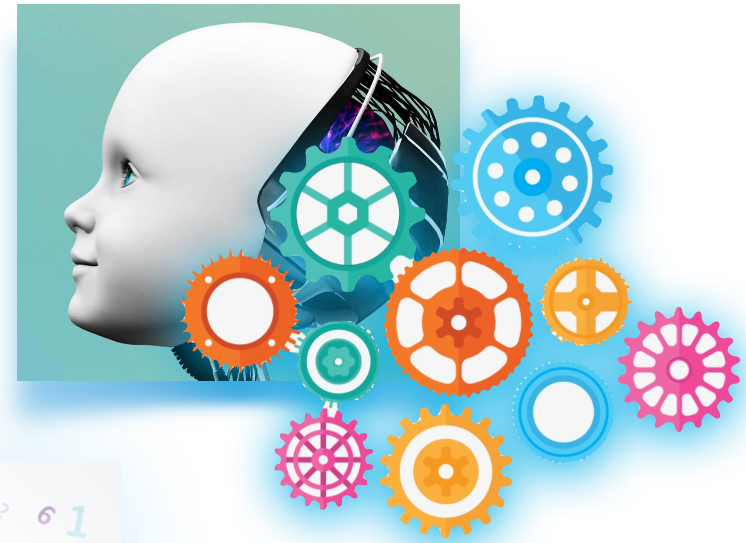
How [NP [CP-who *__how*]]?

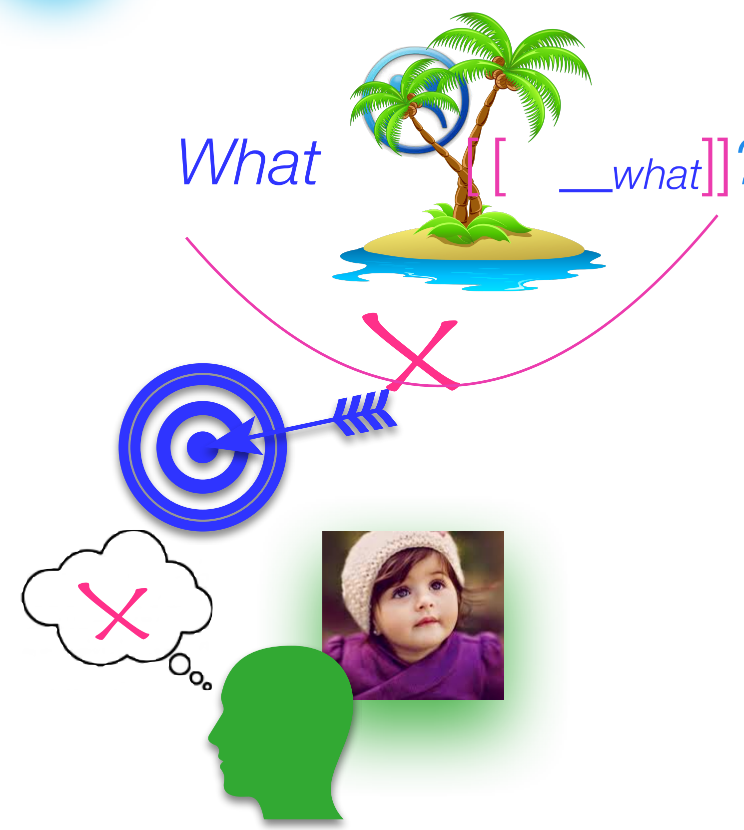
De Villiers et al. 2008

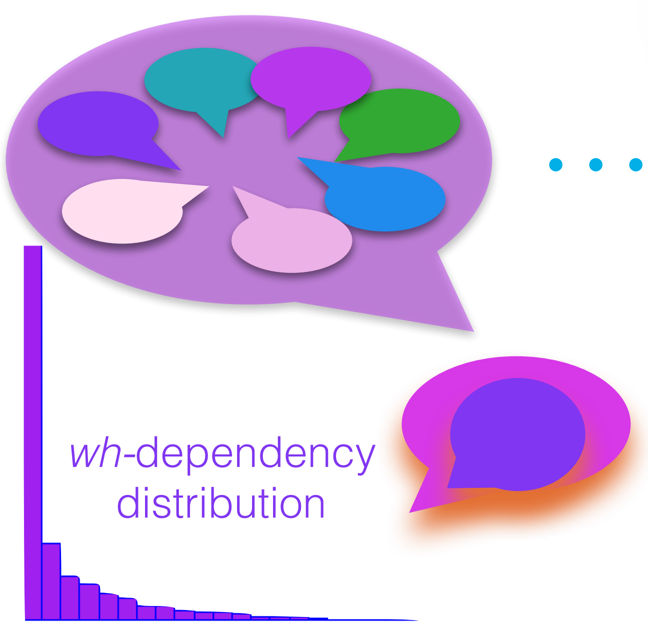
Dickson et al. 2022, in prep.

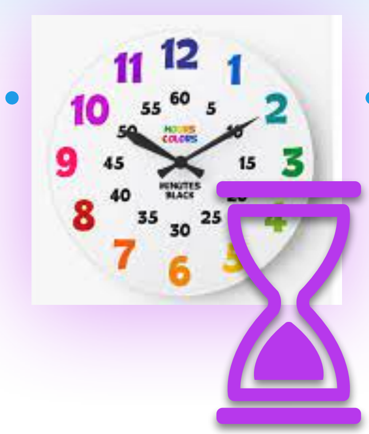
2: what the pieces are and their probabilities

So what does the modeled child do?

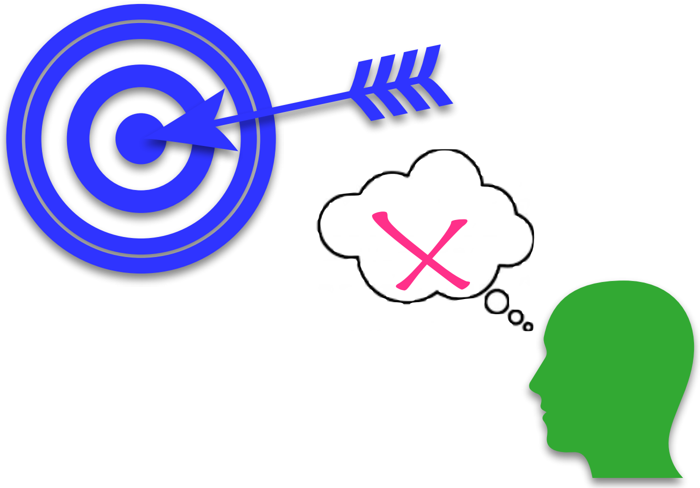






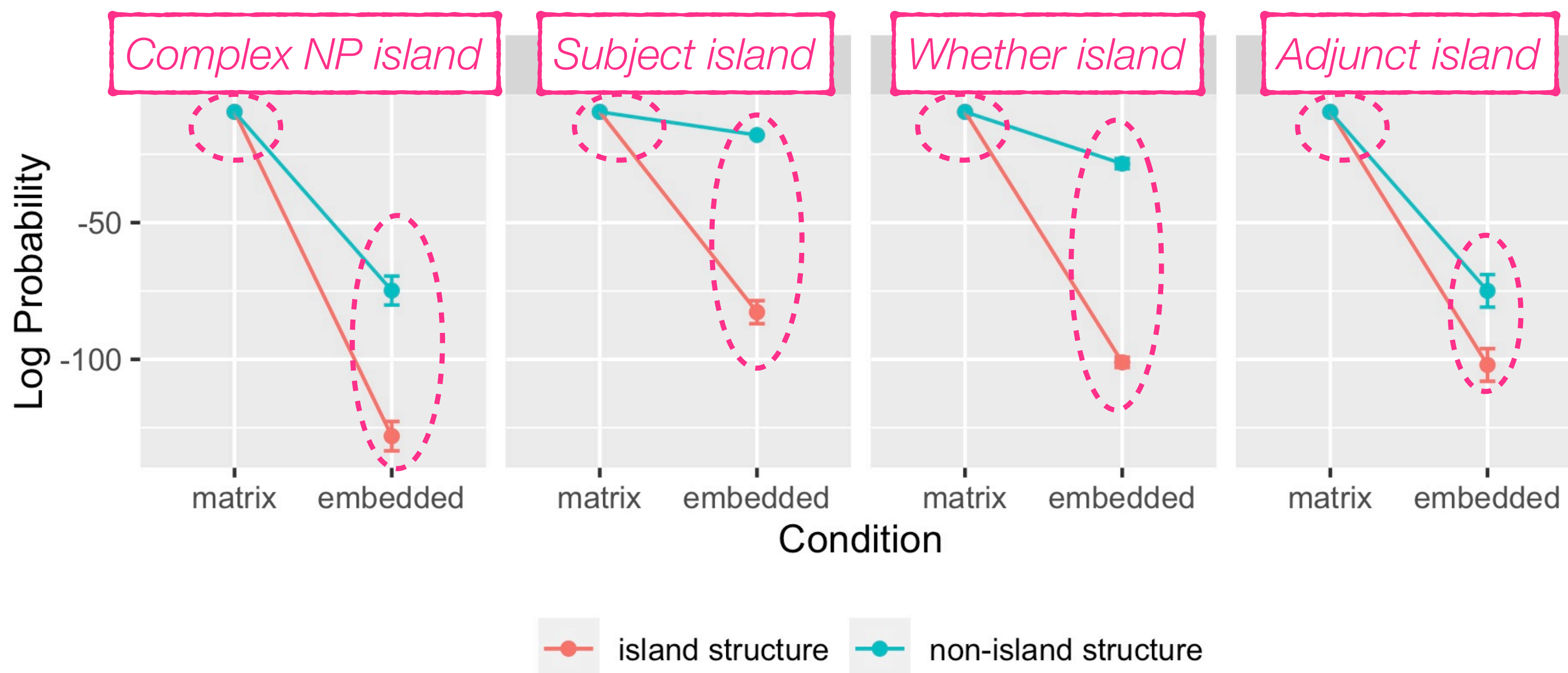


Dickson et al. 2022, in prep.

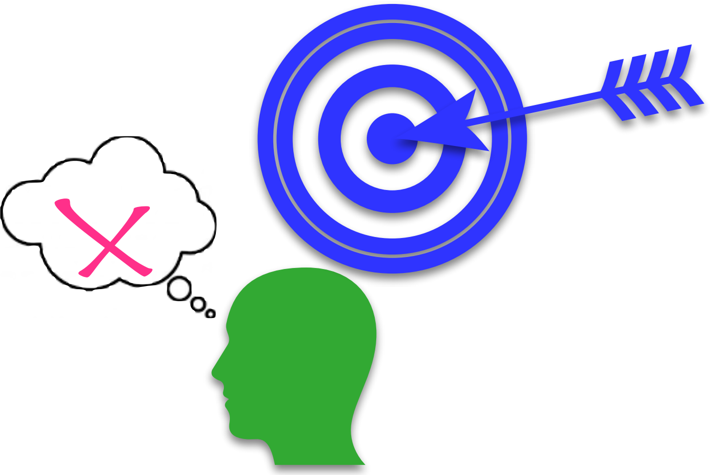


2: what the pieces are and their probabilities

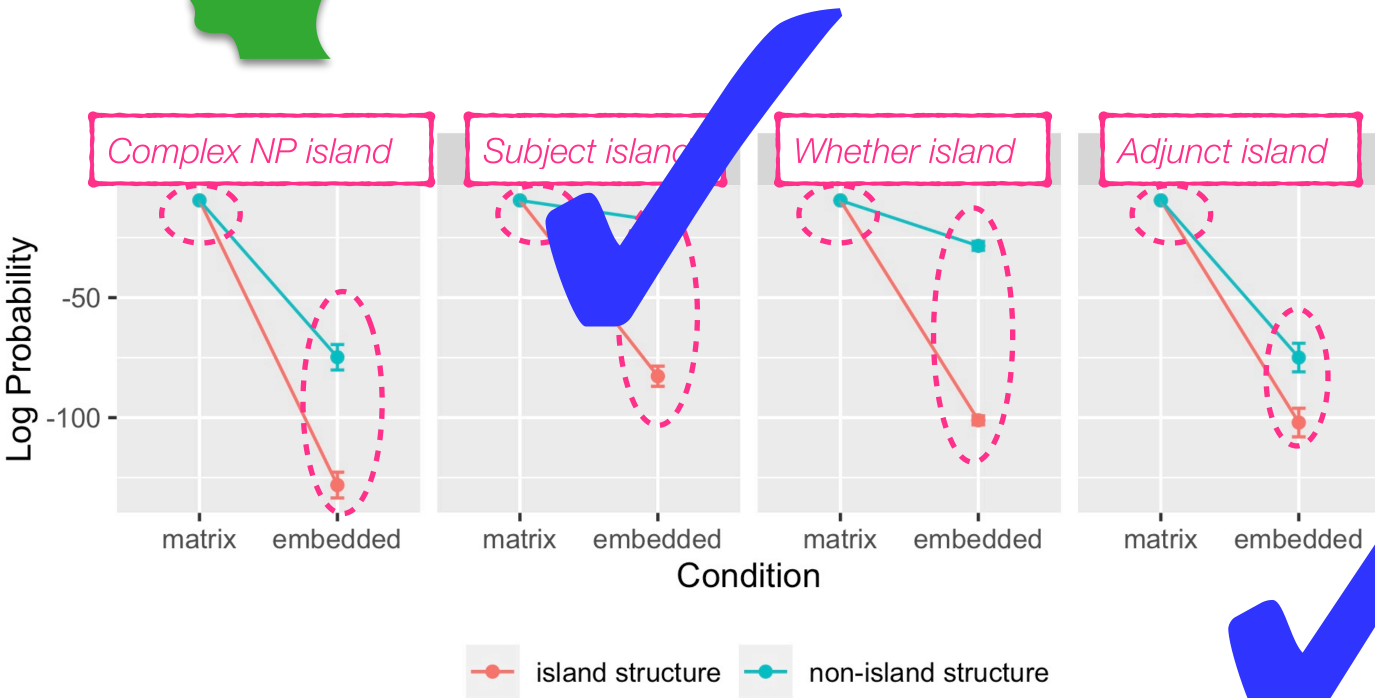
Superadditivity predicted for judgments of all four island types.



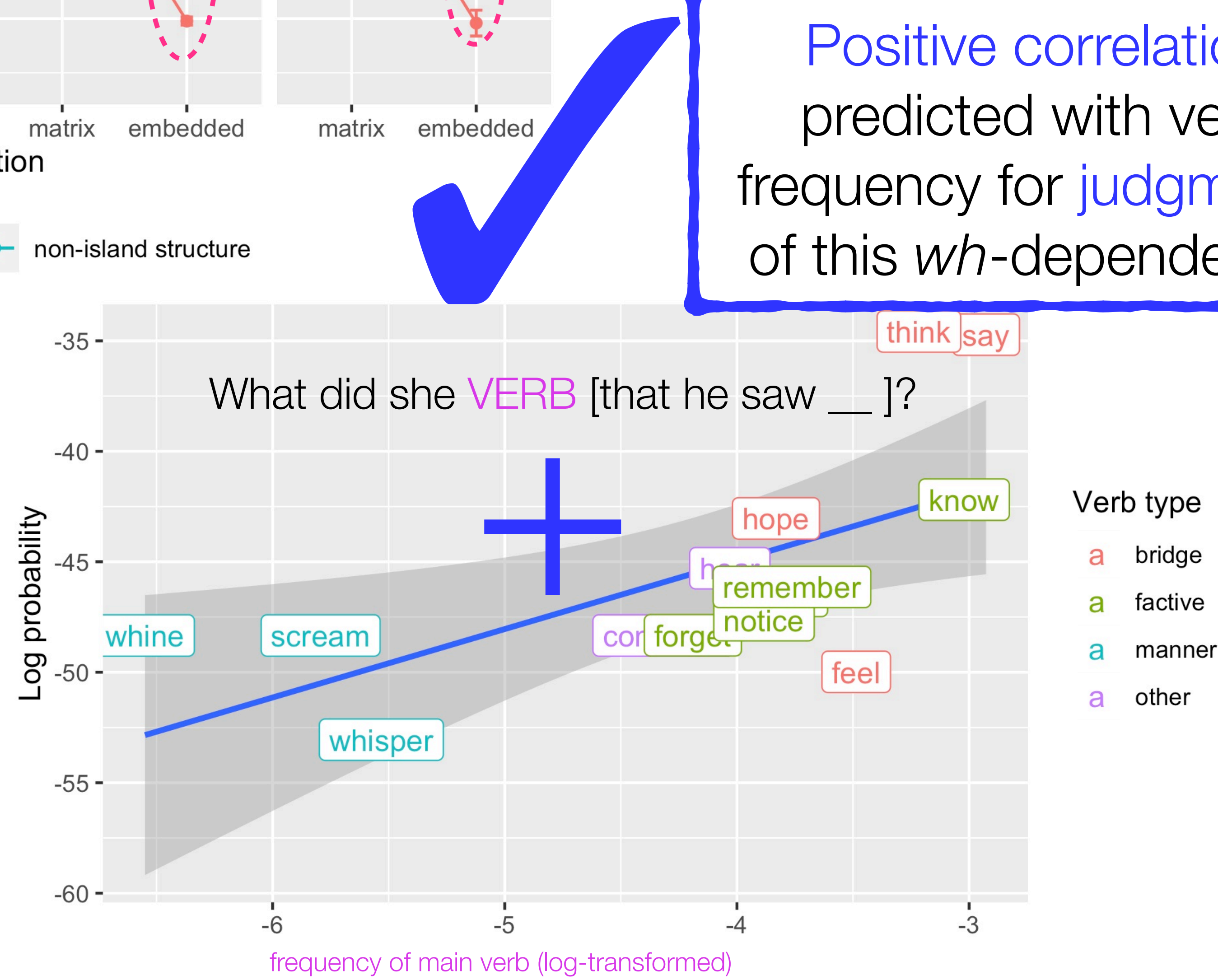
Dickson et al. 2022, in prep.



2: what the pieces are and their probabilities

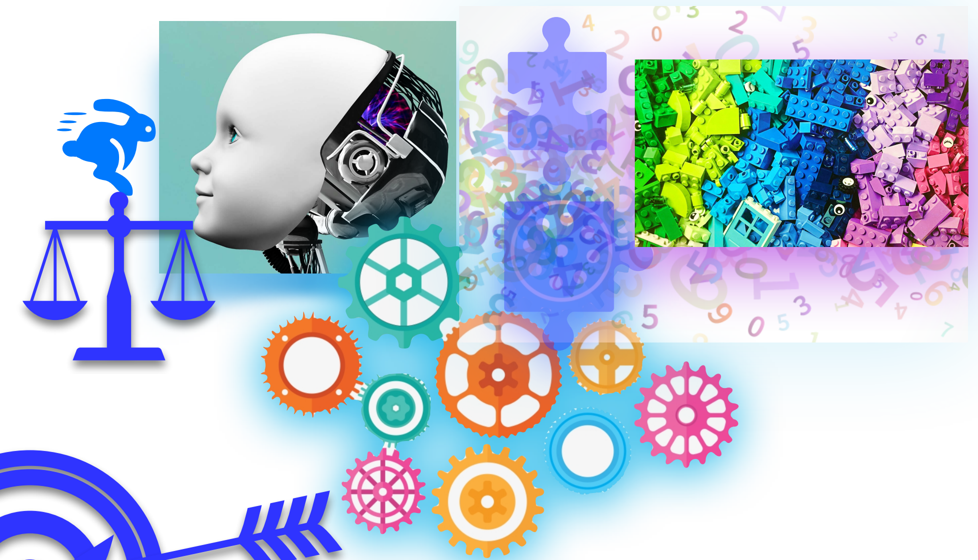


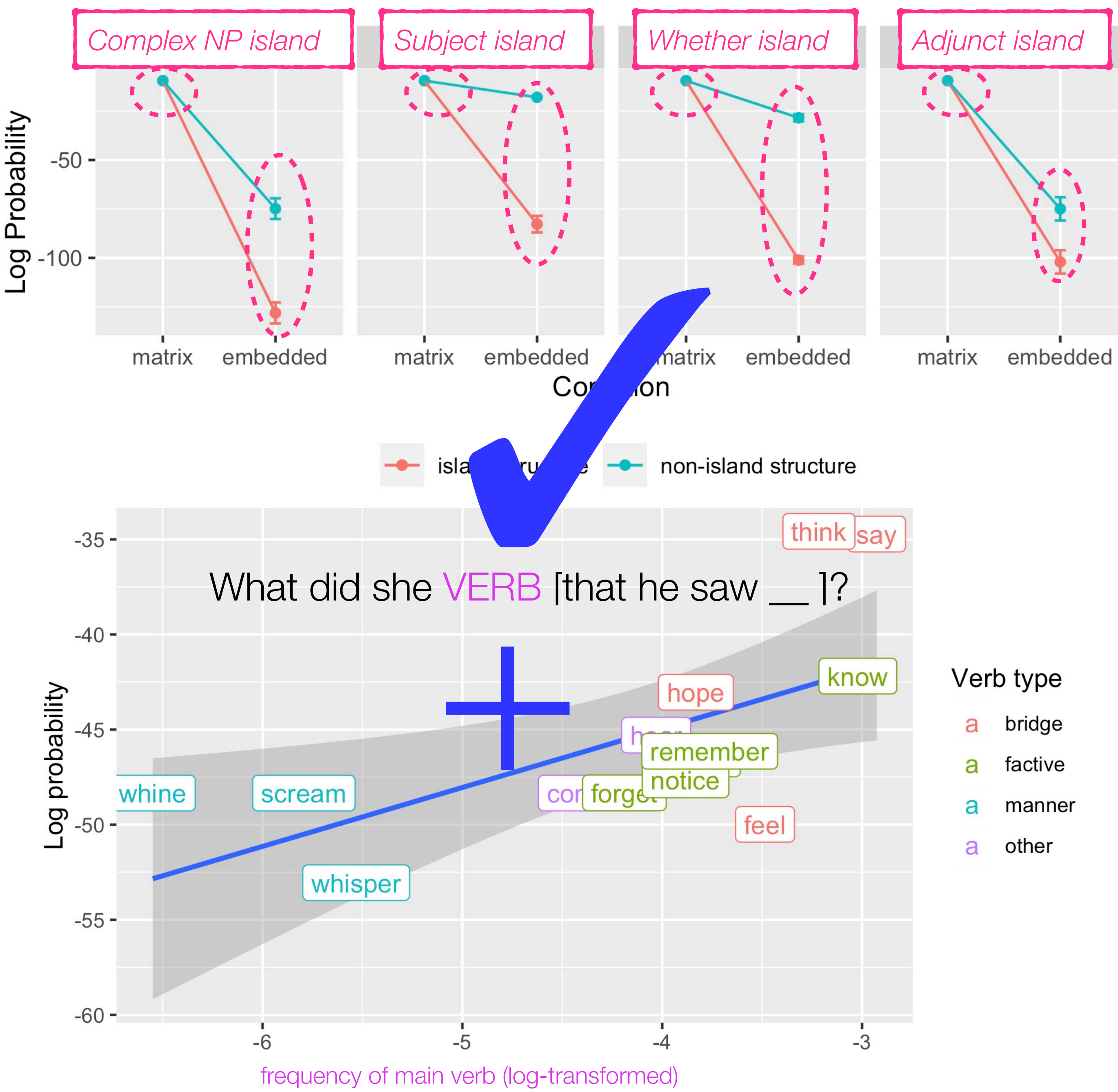
Positive correlation predicted with verb frequency for judgments of this *wh*-dependency.

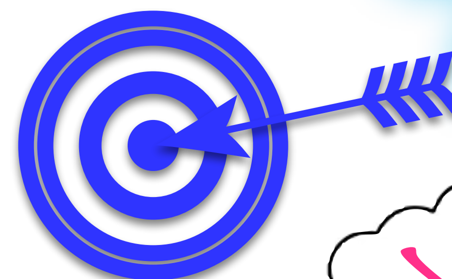


Dickson et al. 2022, in prep.

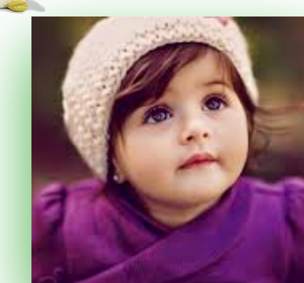
2: what the pieces are and their probabilities

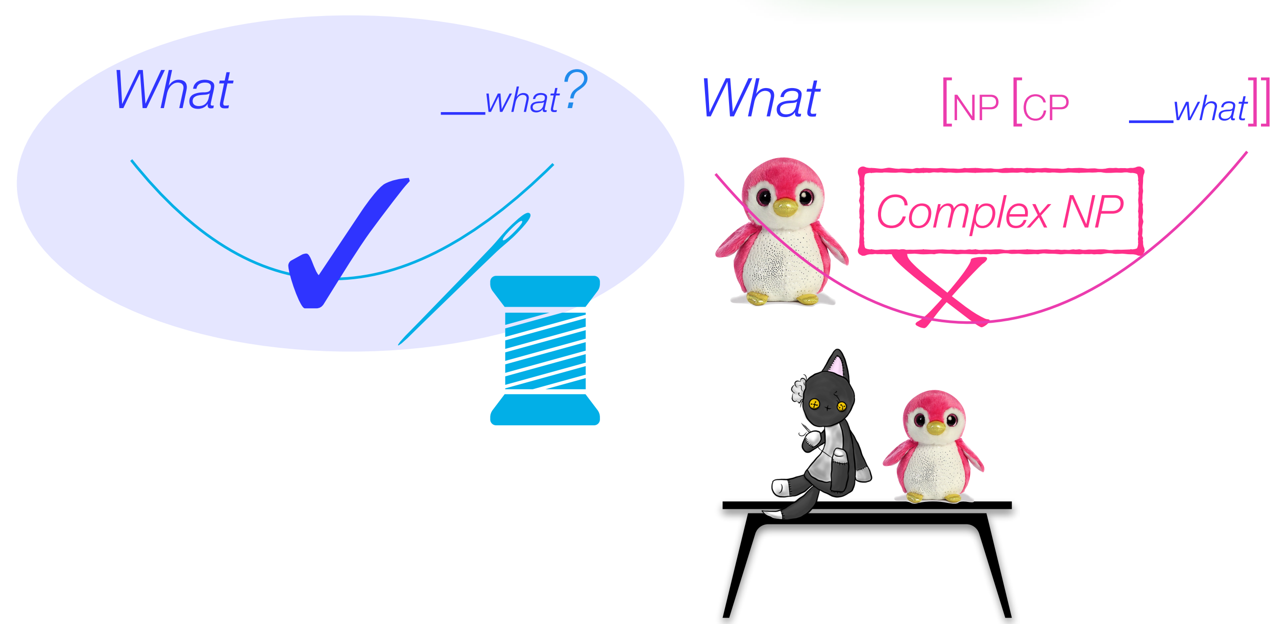






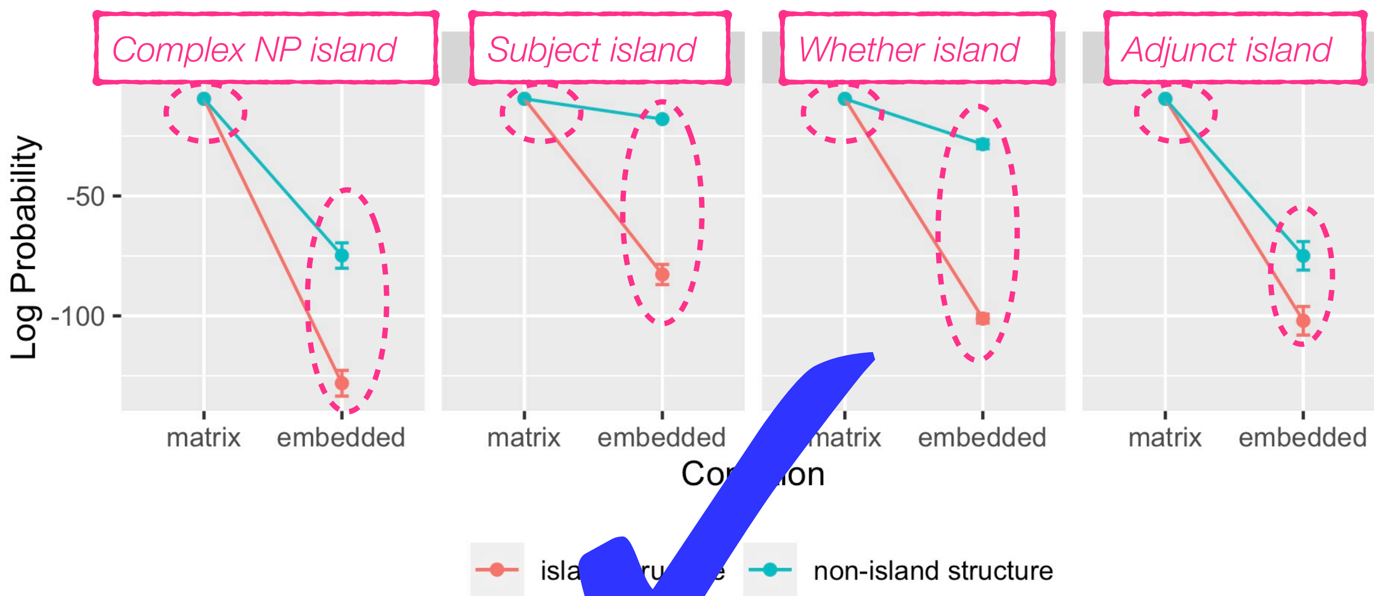
Children prefer this interpretation.



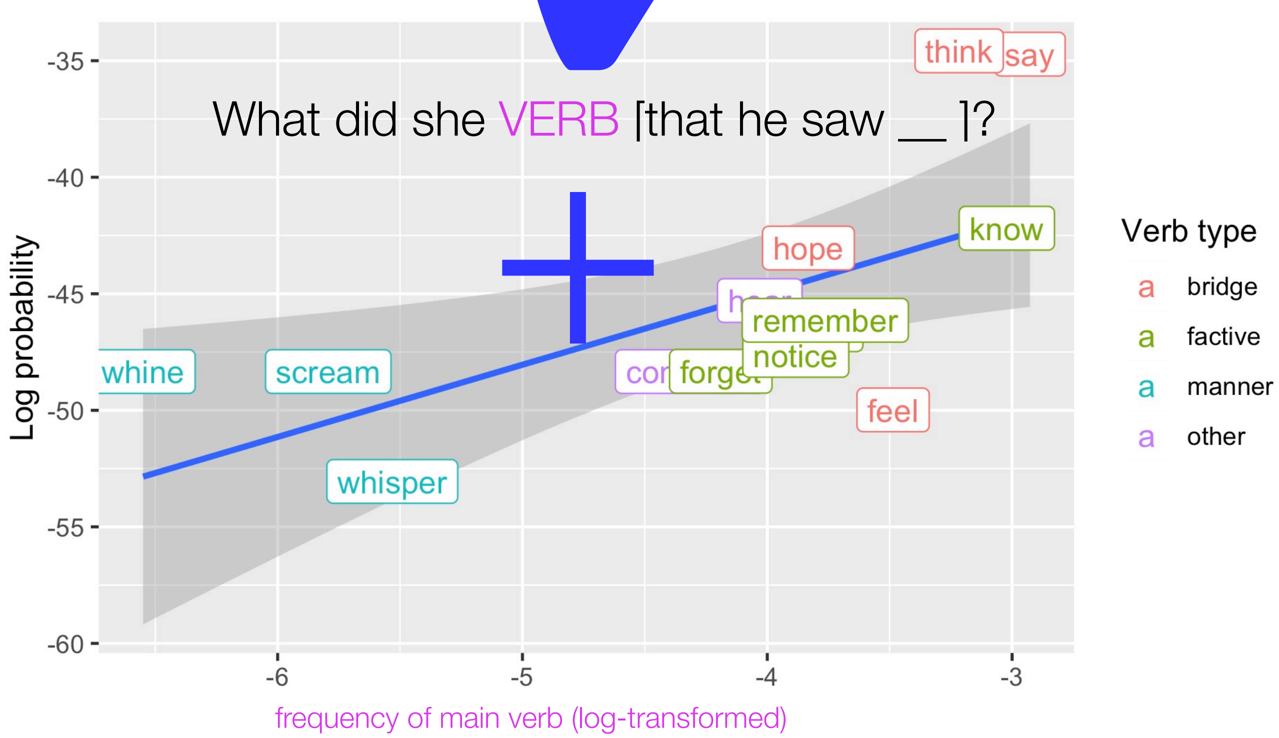


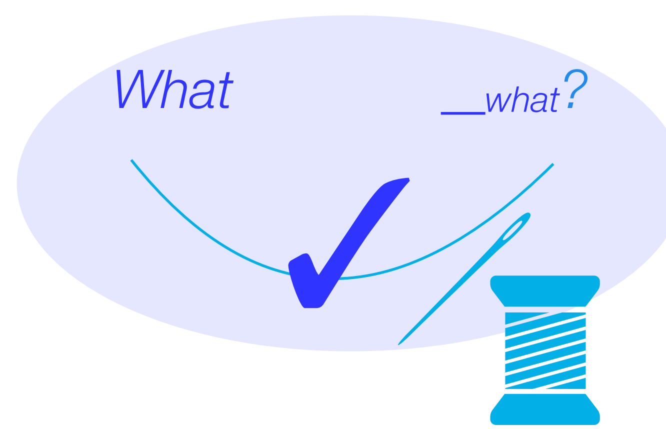
Dickson et al. 2022, in prep.

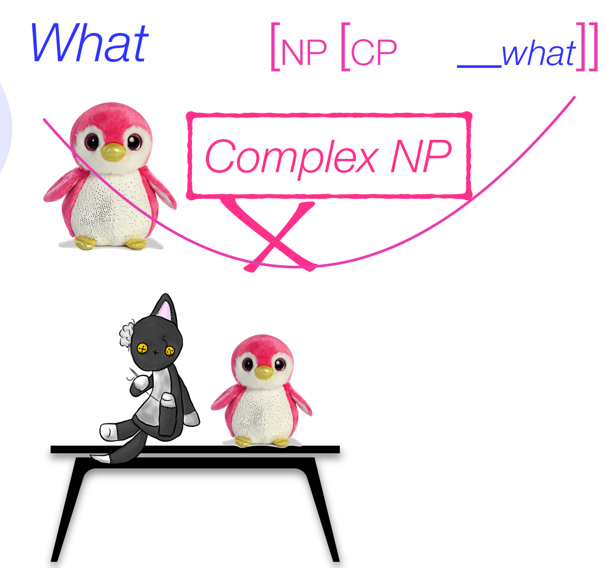
2: what the pieces are and their probabilities



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

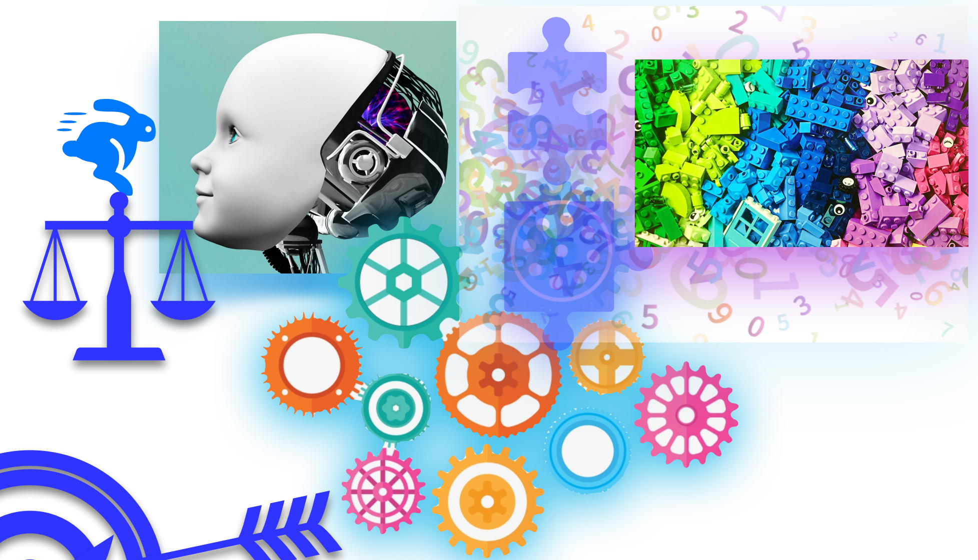


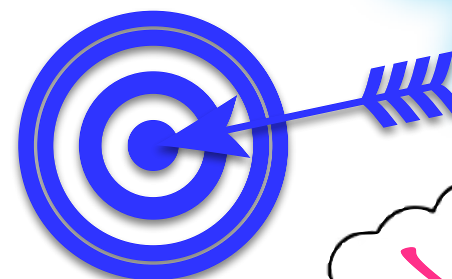


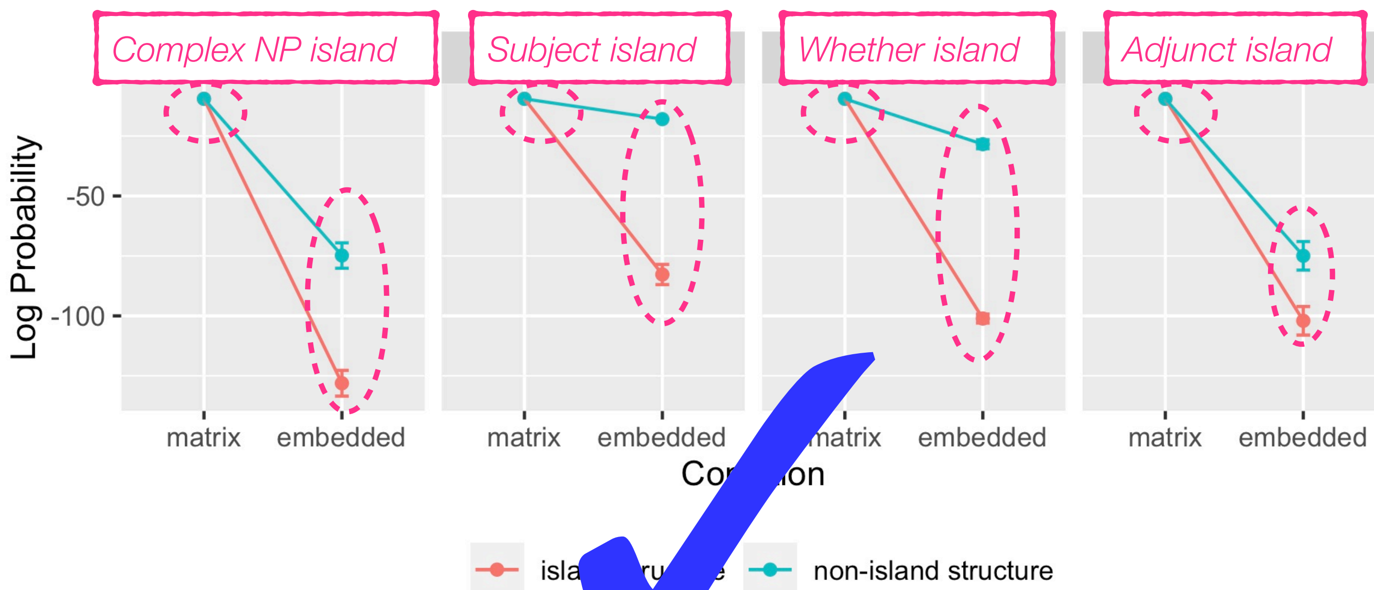


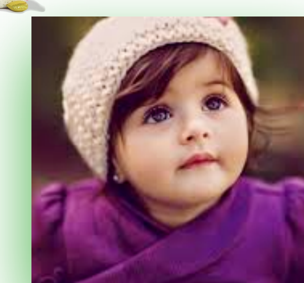
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

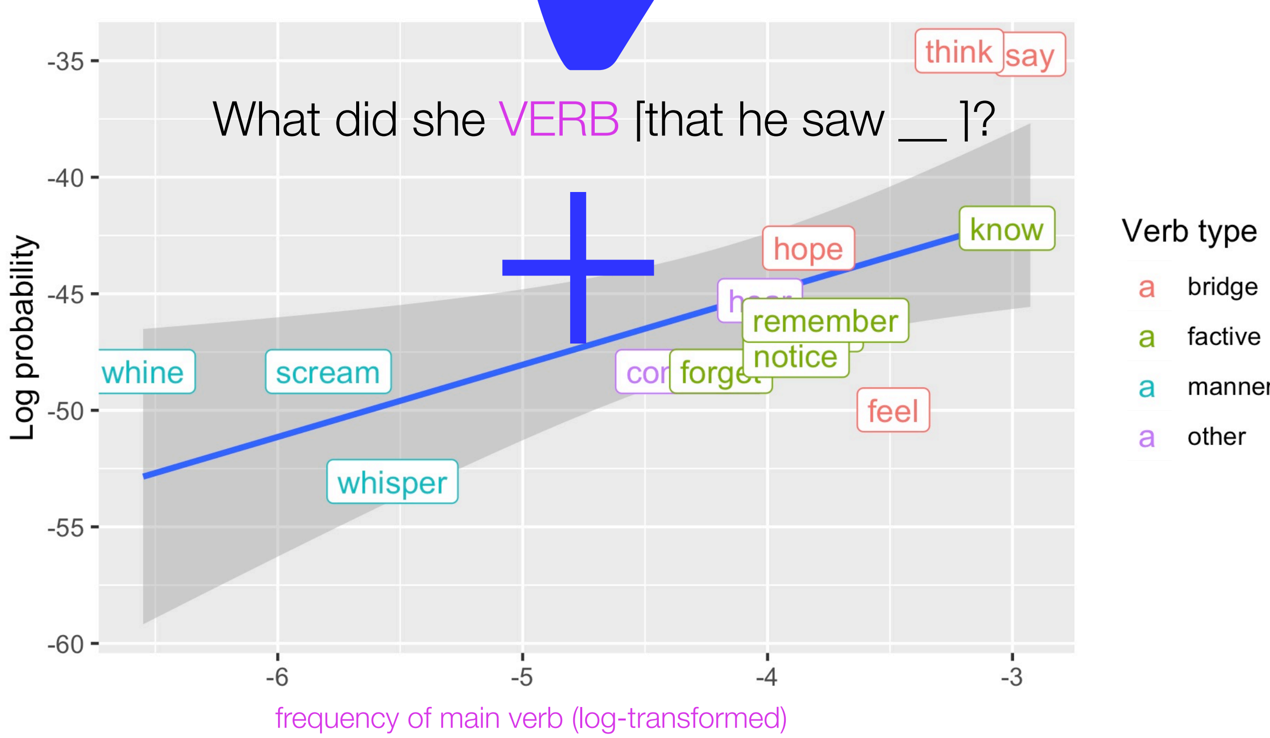


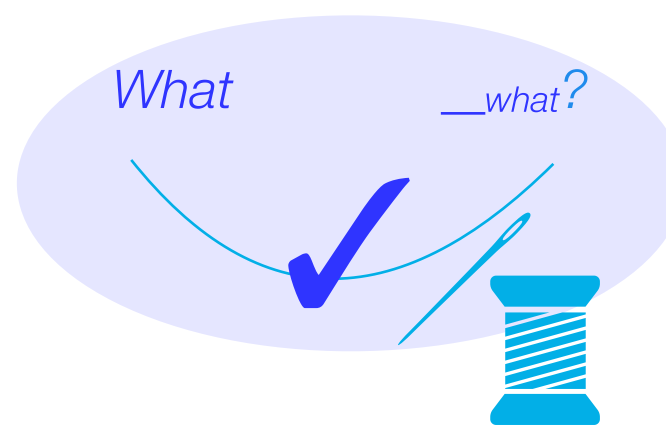


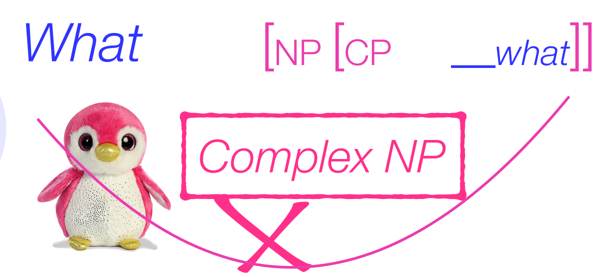


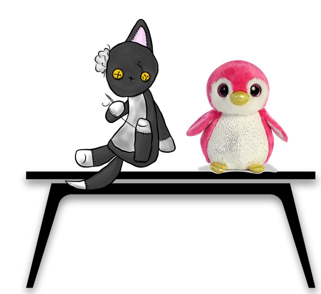


So, the modeled child prefers it.



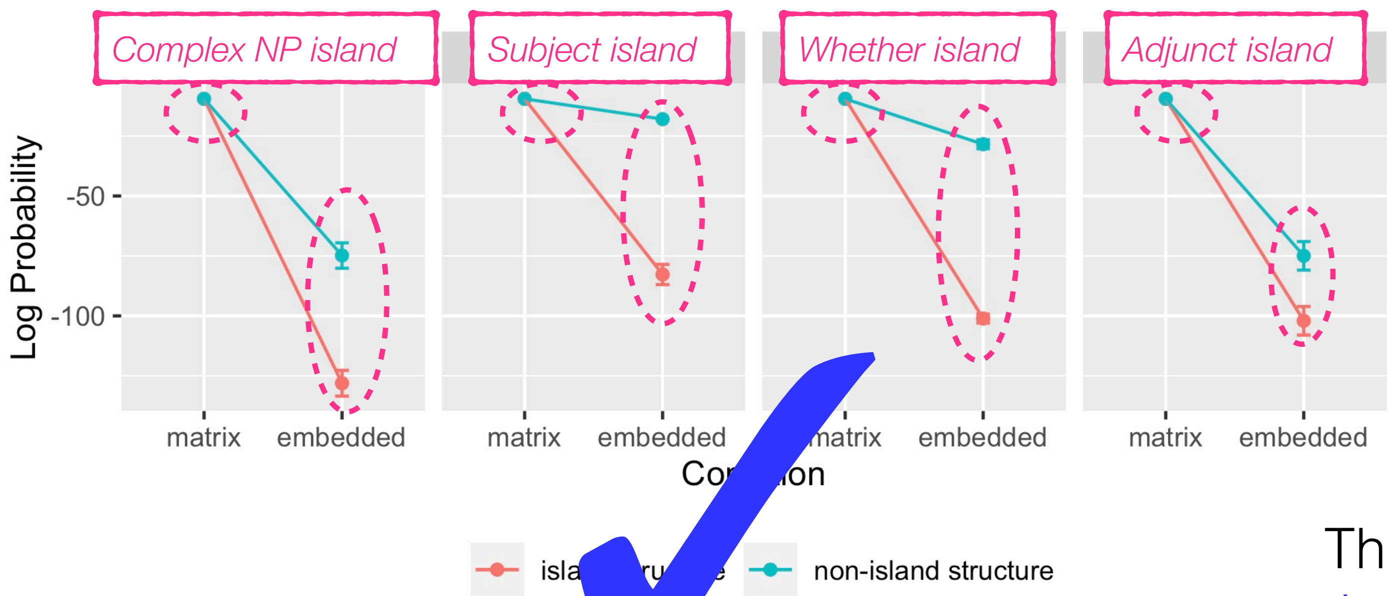


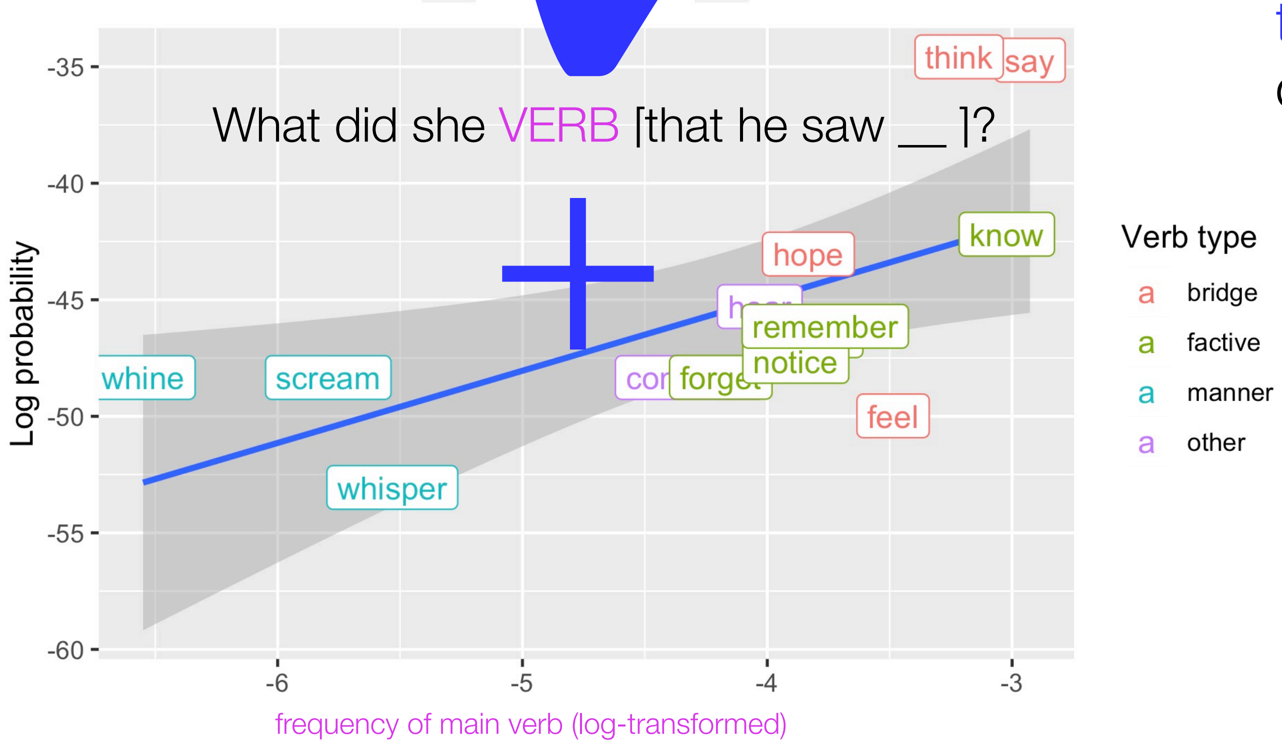


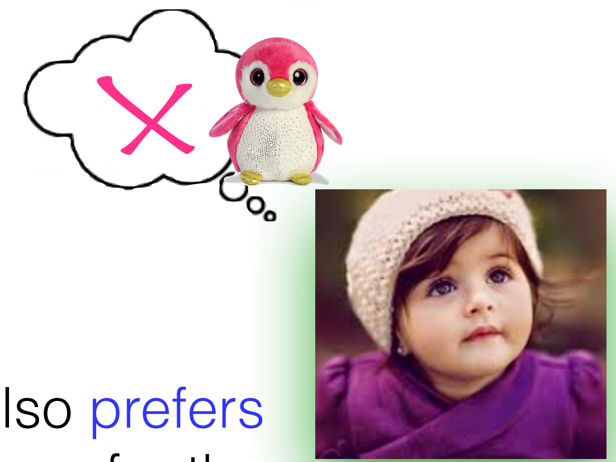


Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities







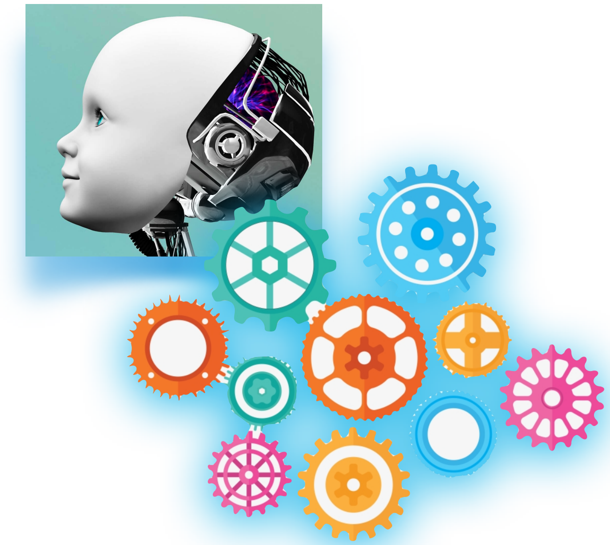
The modeled child also prefers the child-preferred ones for the other *wh*-dependencies.

Who __*who*?
How __*how*?

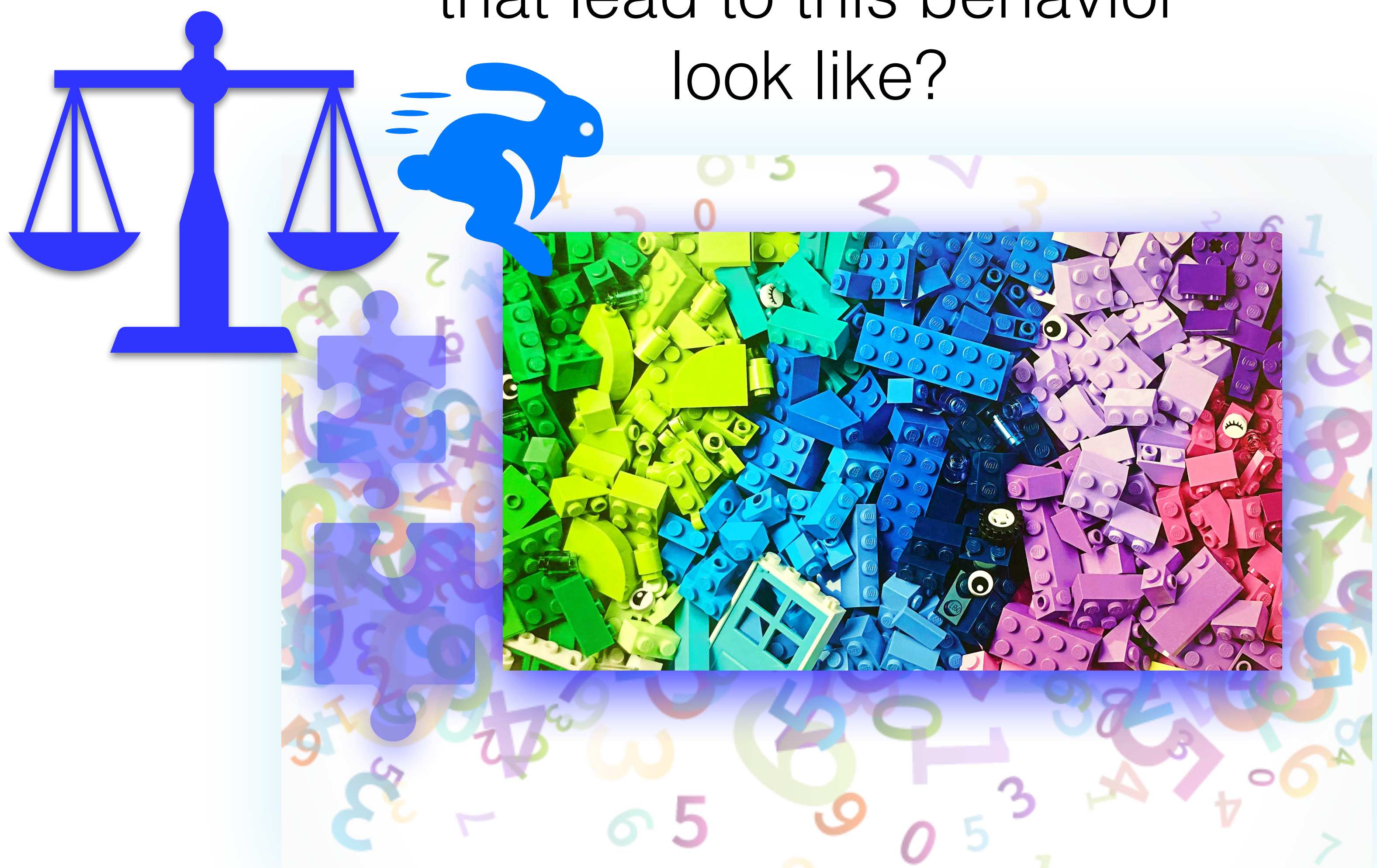
Who [CP-how __*who*]?
Who [CP-what __*who*]?
How [CP-what __*how*]?
How [CP-where __*how*]?
How [NP [CP-who __*how*]]?

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

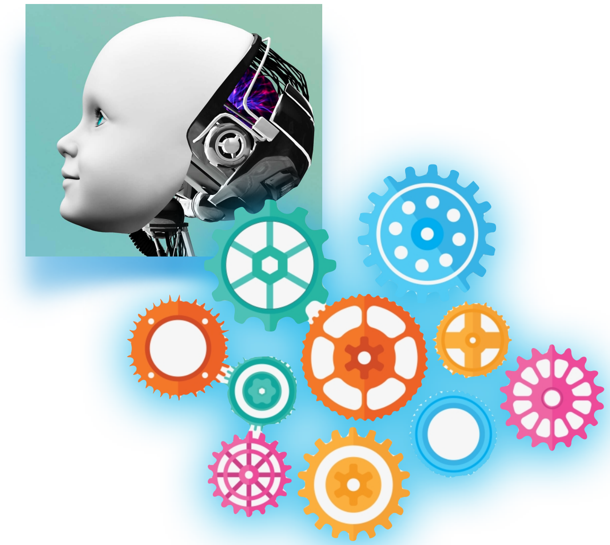


What do the learned building blocks that lead to this behavior look like?



Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities



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

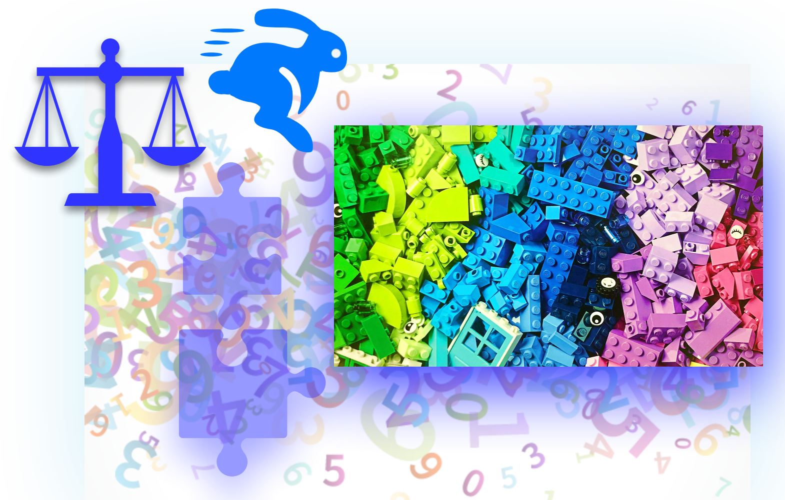
$IP_{I:present}$-VP_{V}-$IP_{I:-finite}$-$VP_{V:talk}$-$PP_{P:to}$

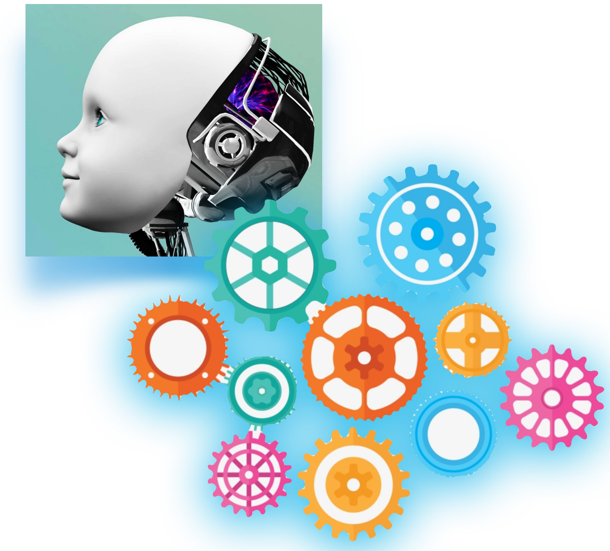
$VP_{V:think}$-$CP_{C:null}$-IP

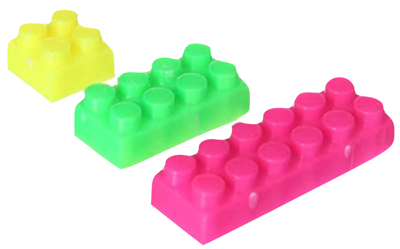
$IP_{I:past}$-$VP_{V:say}$-CP

Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities







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$_{\text{say}}$

V$_{\text{go}}$

VP$_{\text{V:know}}$

VP$_{\text{V}}$

IP$_{\text{I:present}}$-VP$_{\text{V:be}}$

IP$_{\text{I:past}}$-VP

CP$_{\text{C:null}}$-IP

NP$_{\text{N}}$-PP$_{\text{P:of}}$

IP$_{\text{I:-finite}}$-VP$_{\text{V}}$

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

IP$_{\text{I:present}}$-VP$_{\text{V}}$-IP$_{\text{I:-finite}}$-VP$_{\text{V:talk}}$-PP$_{\text{P:to}}$

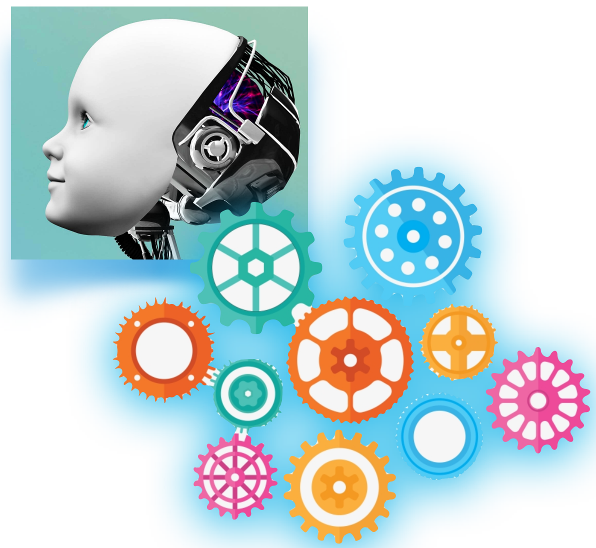
VP$_{\text{V:think}}$-CP$_{\text{C:null}}$-IP

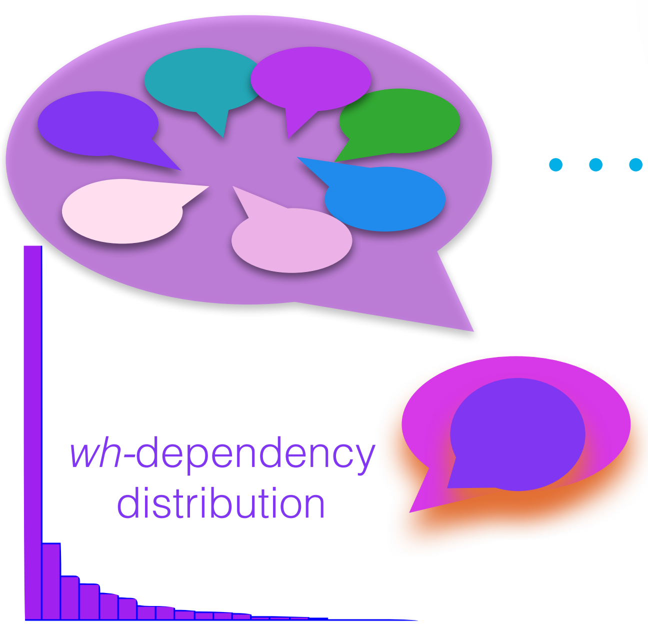
IP$_{\text{I:past}}$-VP$_{\text{V:say}}$-CP

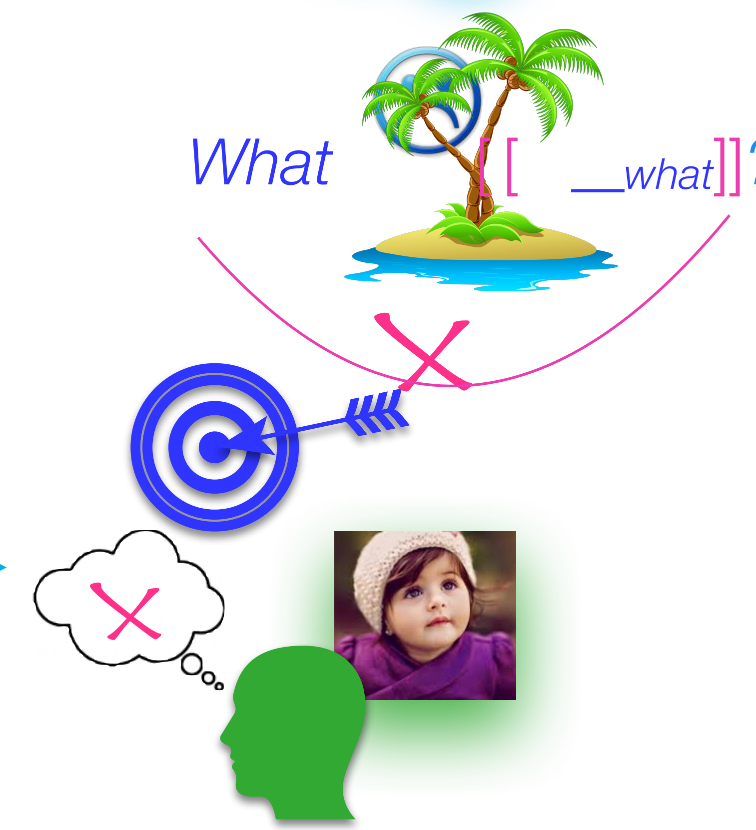
Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

Takeaway:
This theory can work (even better) for learning knowledge about syntactic islands.



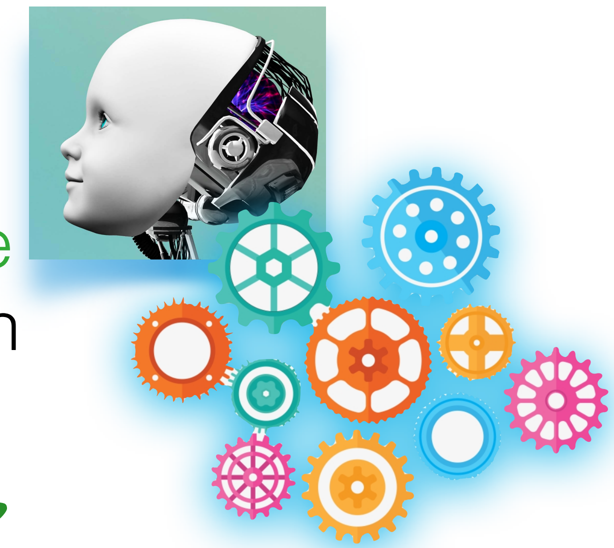


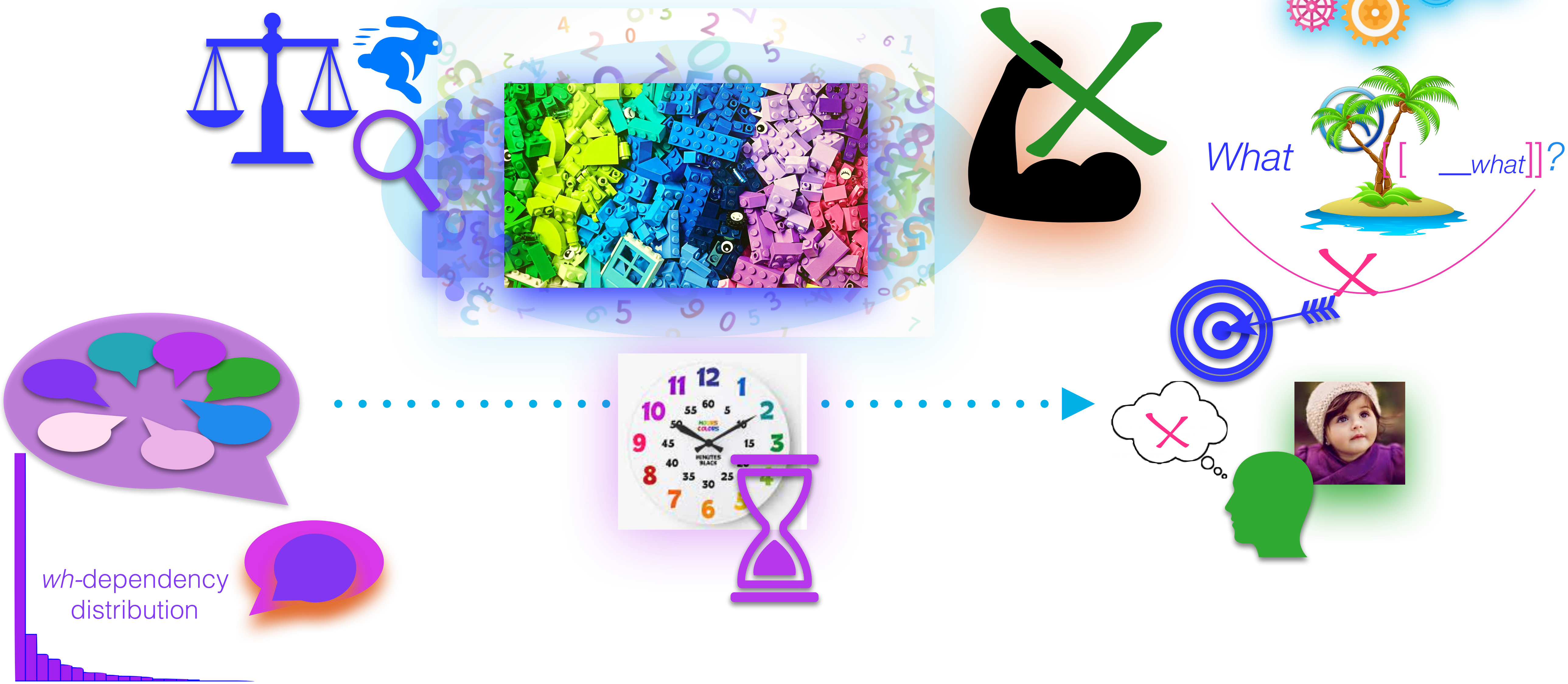


Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

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.

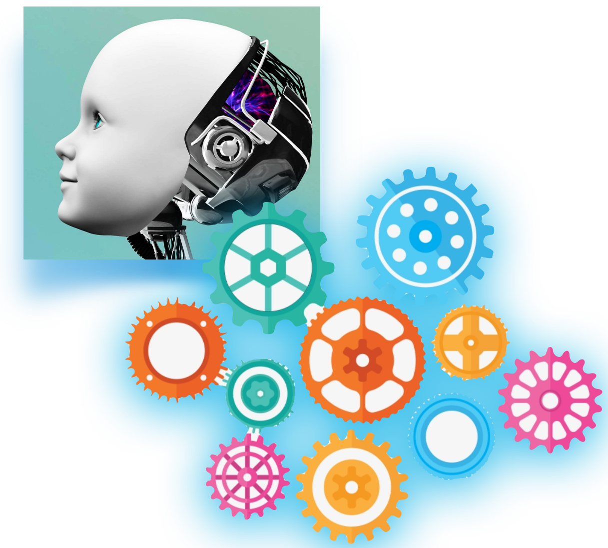


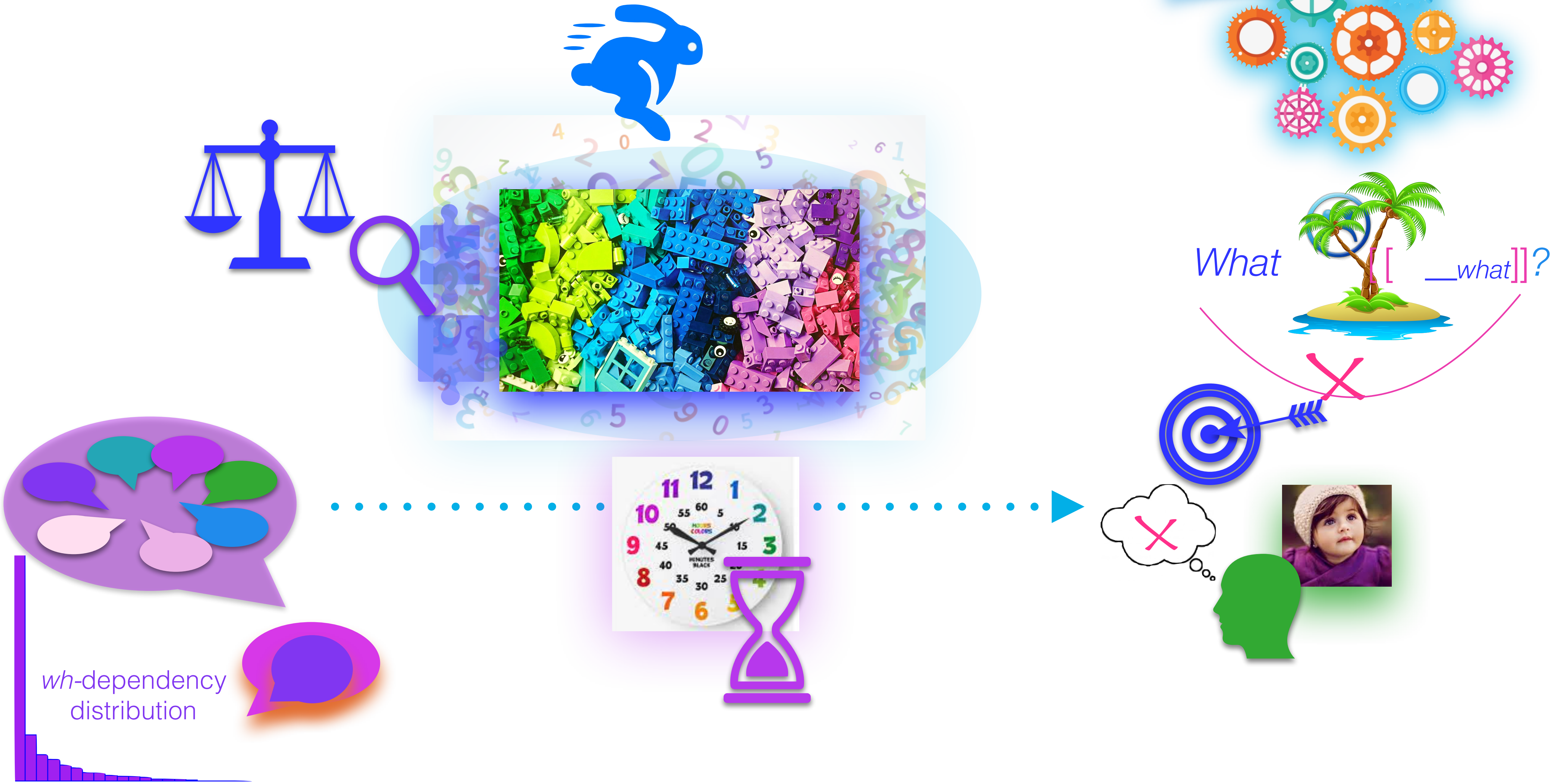


Dickson et al. 2022, in prep.

2: what the pieces are and their probabilities

Key idea: This strategy works when the child's goal is finding efficient building blocks.





Dickson et al. 2022, in prep.

The big picture

One way to succeed at learning about constraints on *wh*-dependencies (syntactic islands) is to learn them indirectly.

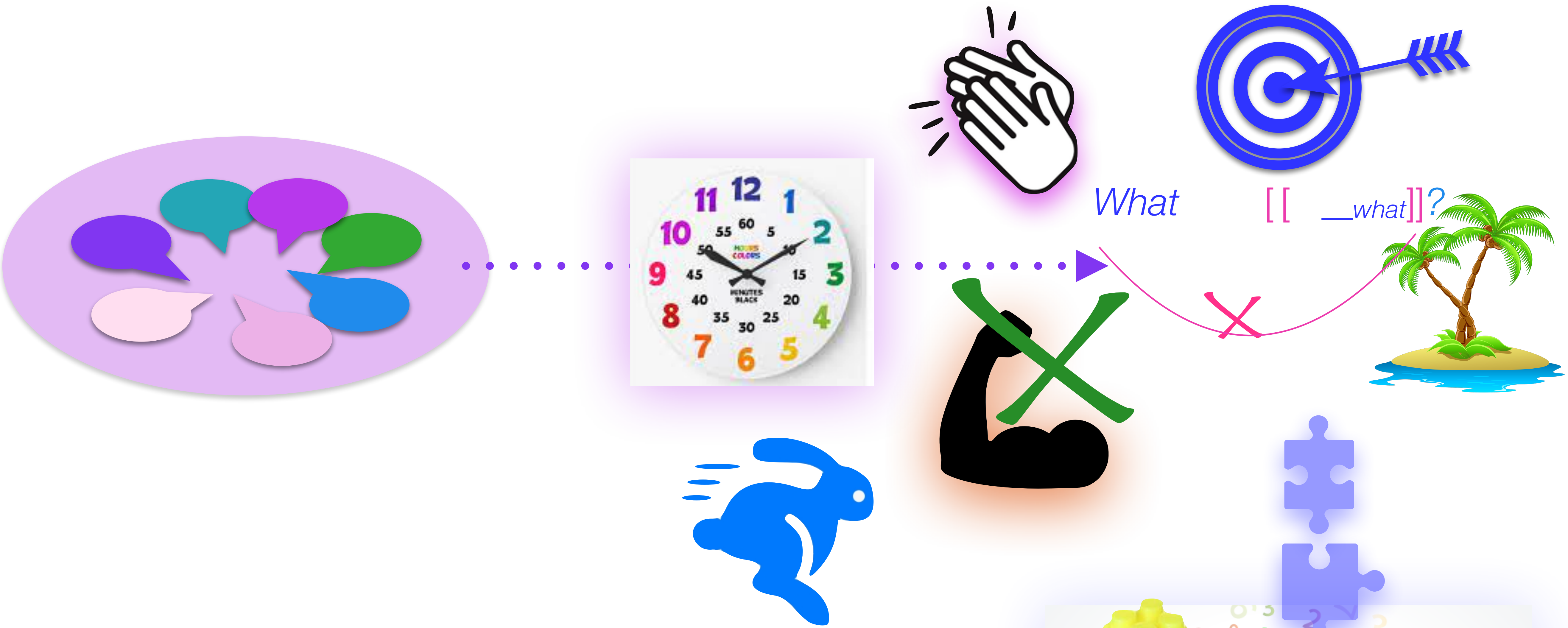
Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.

The big picture

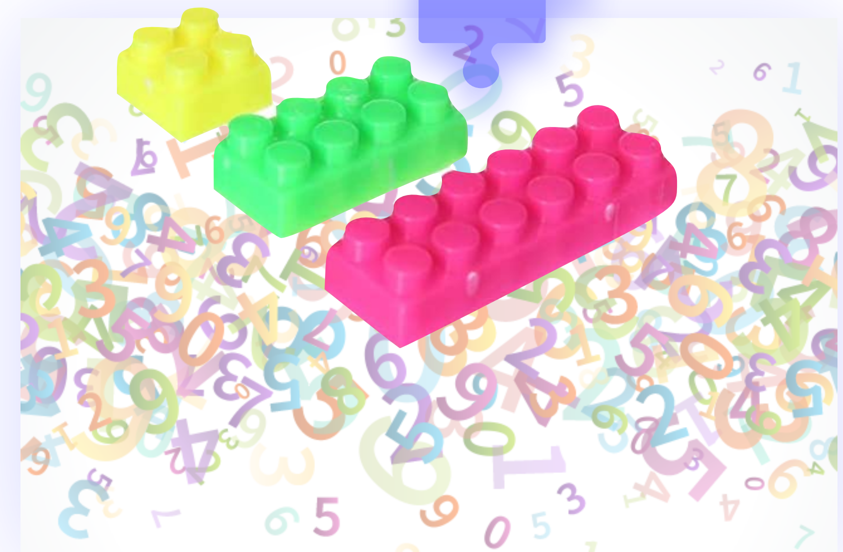
In particular, learn how to represent *wh*-dependencies efficiently using pieces that can be composed locally to represent any *wh*-dependency.

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

The big picture

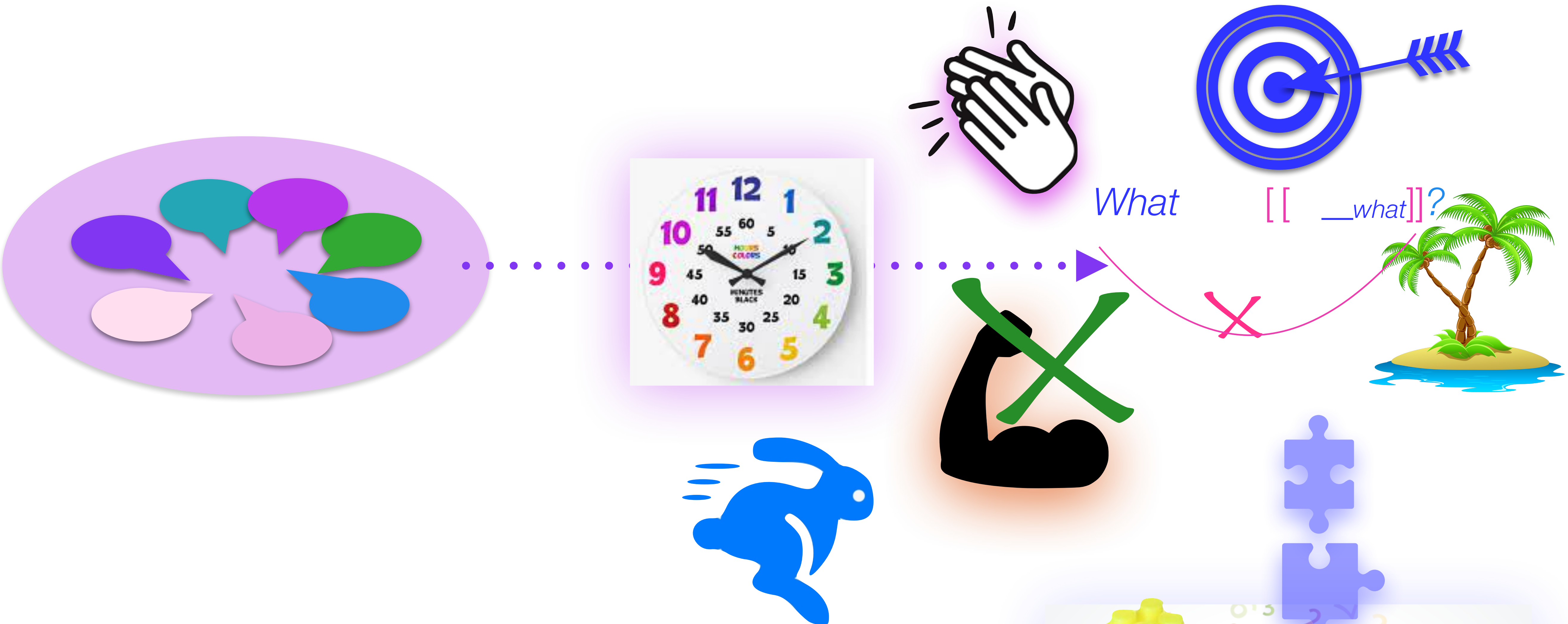


The efficient pieces that are learned, with their associated probabilities, allow the constraints to emerge automatically.

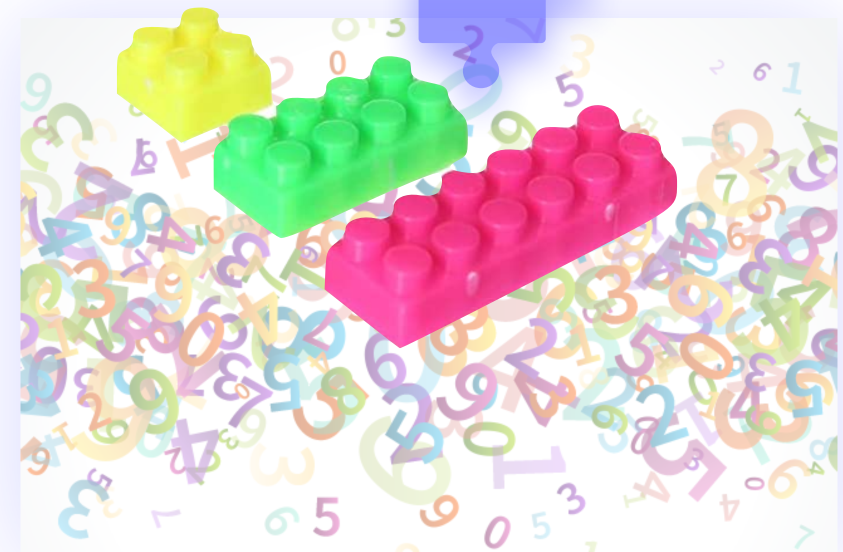


Pearl & Sprouse 2013, Bates & Pearl 2019, Dickson et al. 2022, Pearl & Bates in press, Dickson et al. in prep.

The big picture



So this is one way locality could play a key role in the acquisition of complex syntactic knowledge.



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

UArizona Linguistics 2022 UChicago LEAP 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