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


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

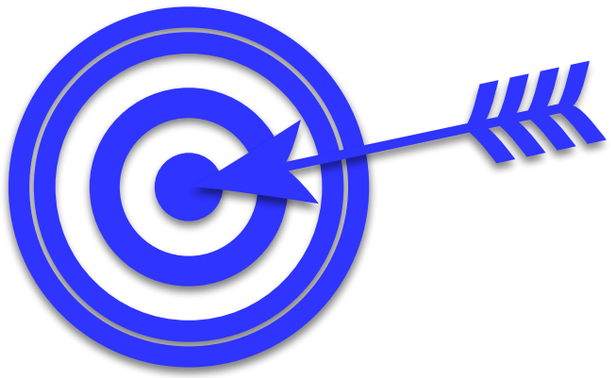
lpearl@uci.edu


February 23, 2024

Pomona Acquisition Workshop (PAW)

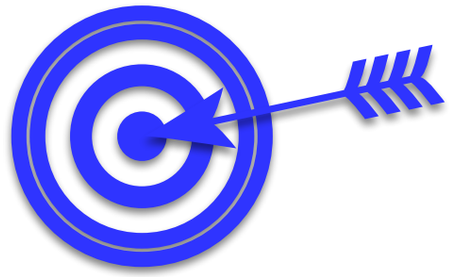
Pomona College

What does it mean to succeed
at syntactic island acquisition?

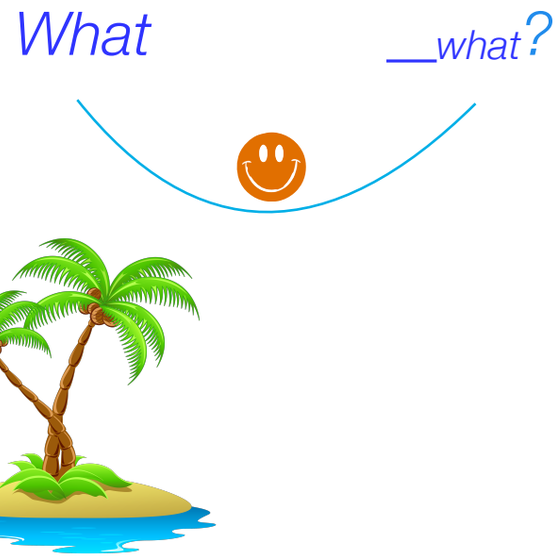


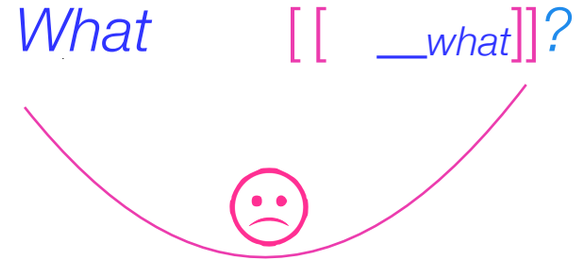
One answer: To develop the target behavior we observe about syntactic islands…

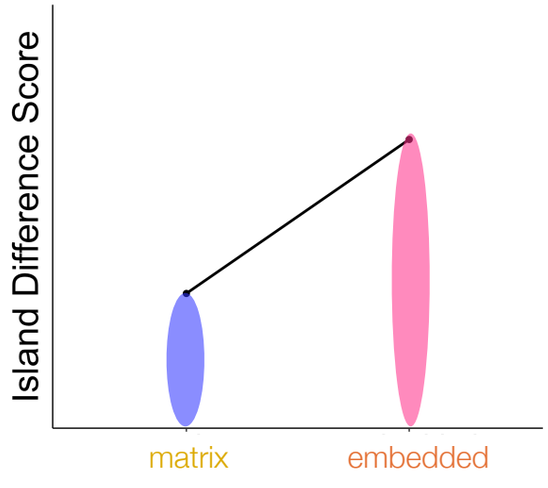
One answer: To develop the target behavior we observe about syntactic islands…

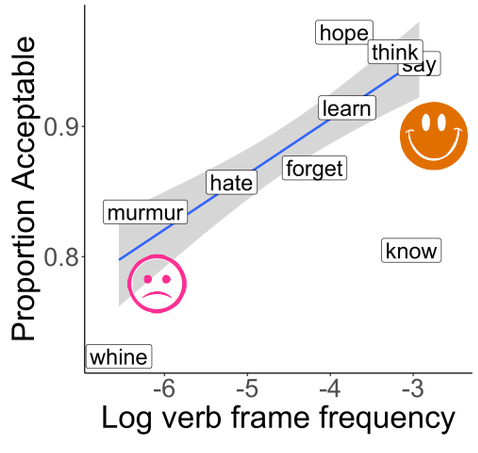


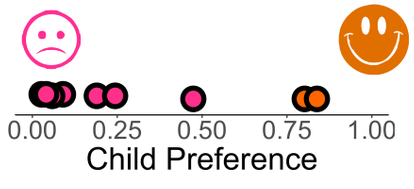
Some example behavior:
judgment patterns and (dis)preferences for certain utterances related to syntactic islands



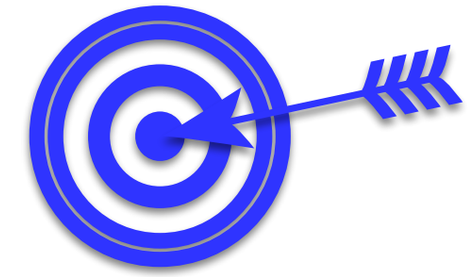



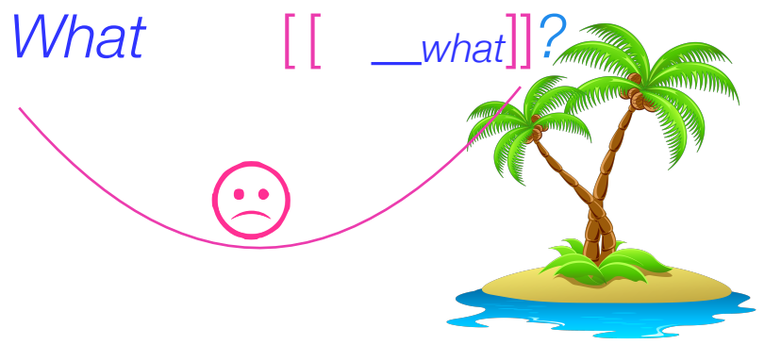



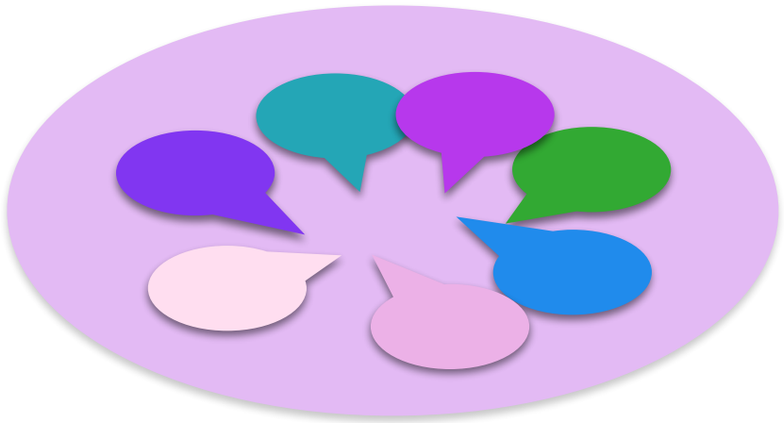



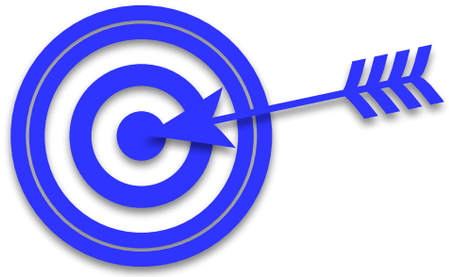
One answer: To develop the target behavior we observe about syntactic islands, given the input children get and the time they have to learn.



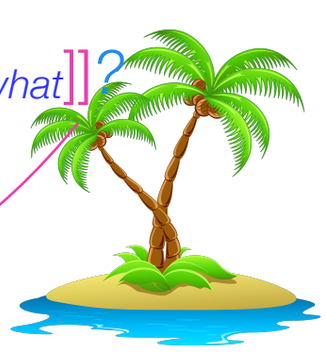


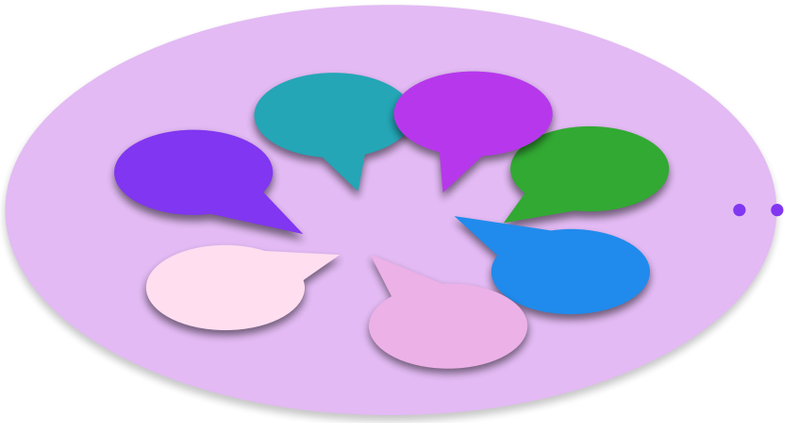



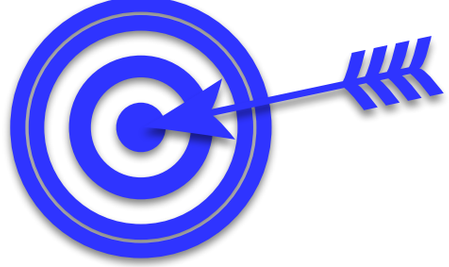
# Acquisition success for syntactic islands

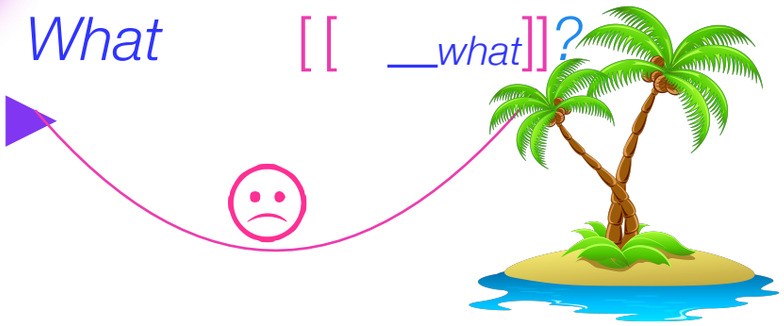


*What* [ [ __*what*]]?



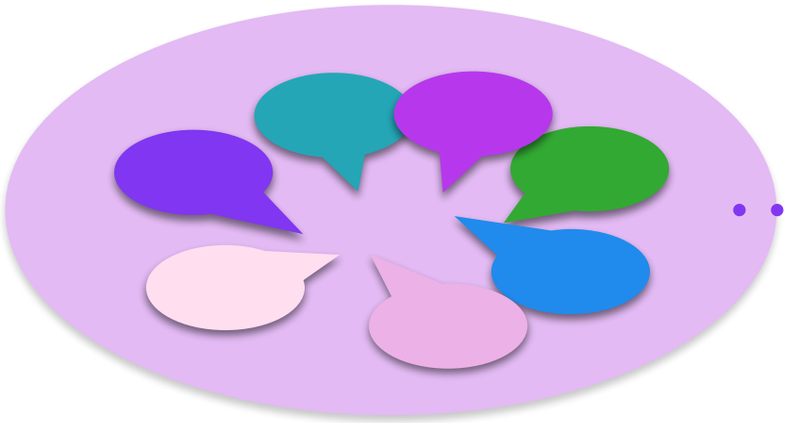


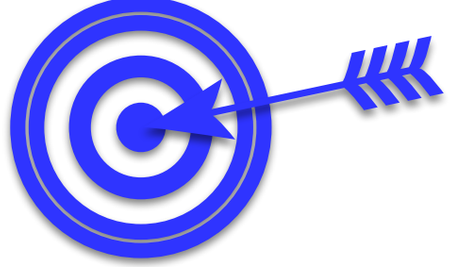





"…without really trying…"

What does it mean to try?

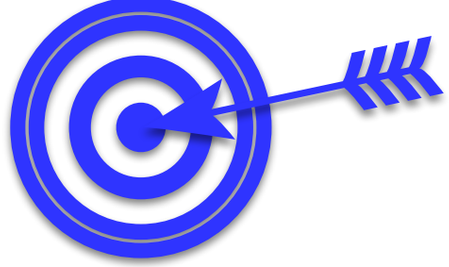




“trying”

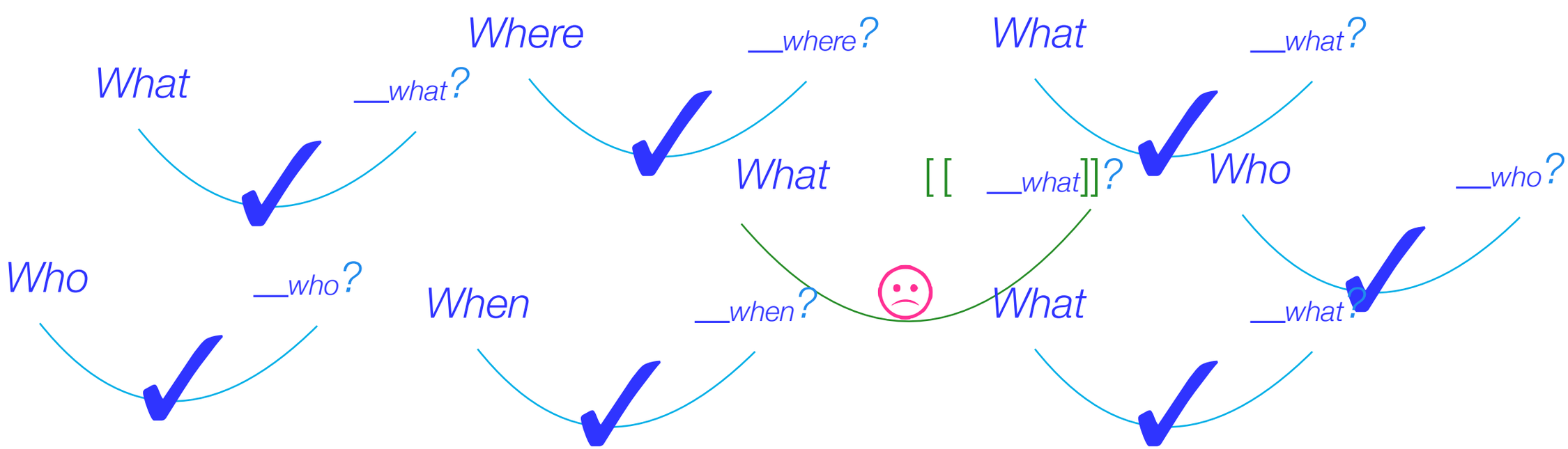
One answer: Learn about syntactic islands directly. For instance, look for language-specific “bounding nodes” (Subjacency: Chomsky 1973, Huang 1982, Lasnik & Saito 1984) that signal syntactic island structure in *wh*-dependencies.

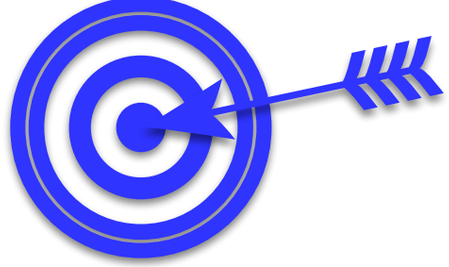
*What* [BN1 [BN2 __*what*]]*?*

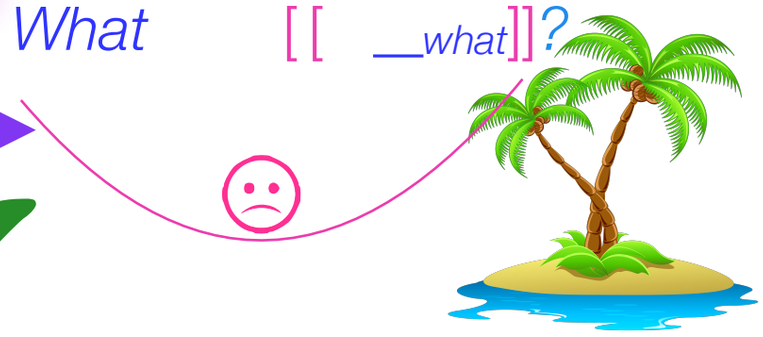


“without really trying”

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

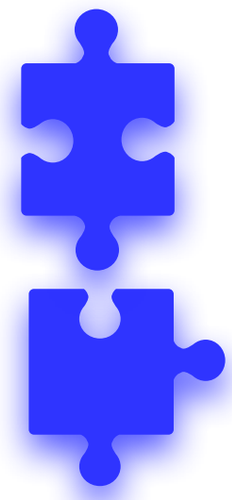


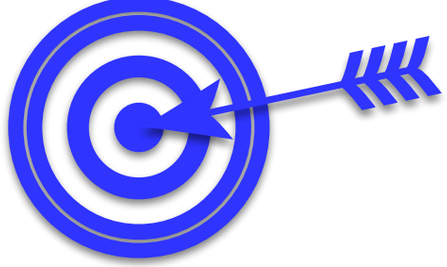


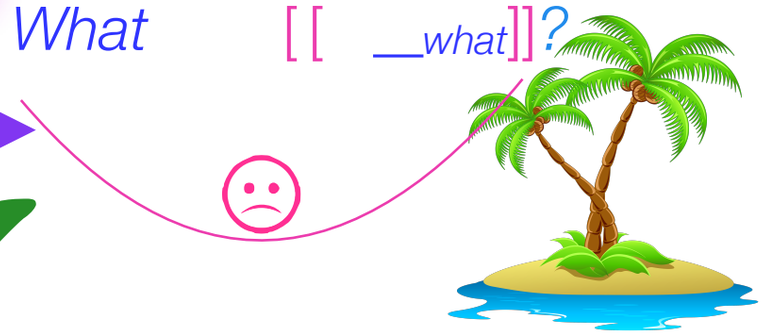



“Learning the right building blocks”

Proposal: The child is trying to learn about the building blocks that combine into *wh*-dependencies.

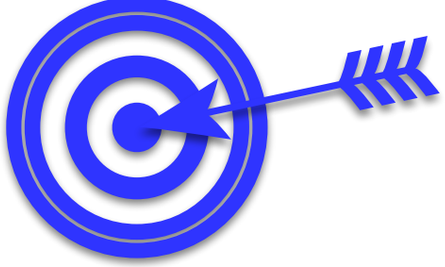




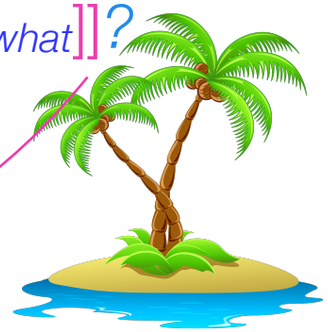



“Learning the right building blocks”

Proposal: Learn about syntactic islands indirectly by learning about the probabilities of the building blocks for *wh*-dependencies.



What [[ __what]]?



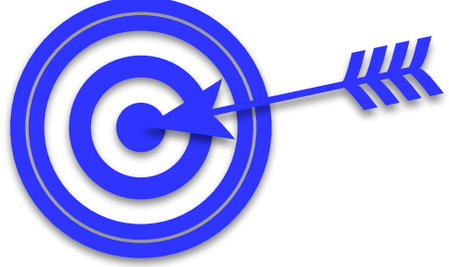
Proposal: Learn simultaneously from the input

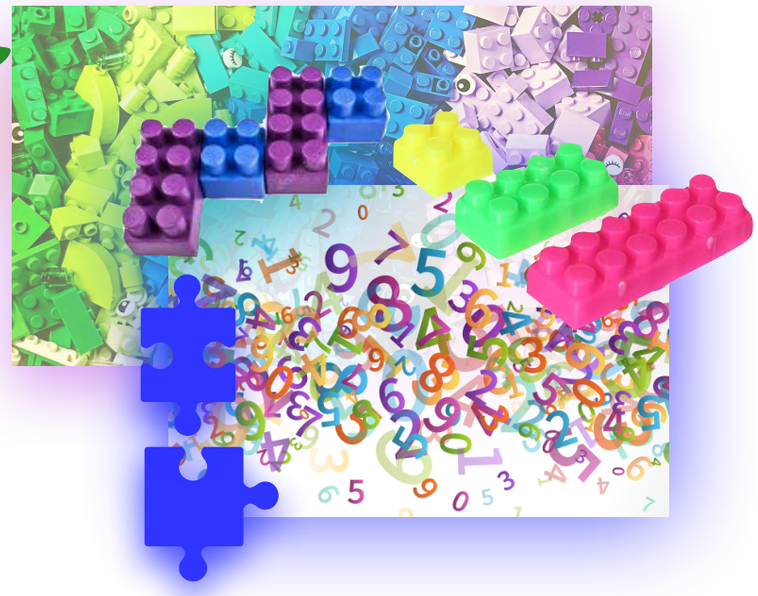
(i) what the building blocks are, and

(ii) their probabilities

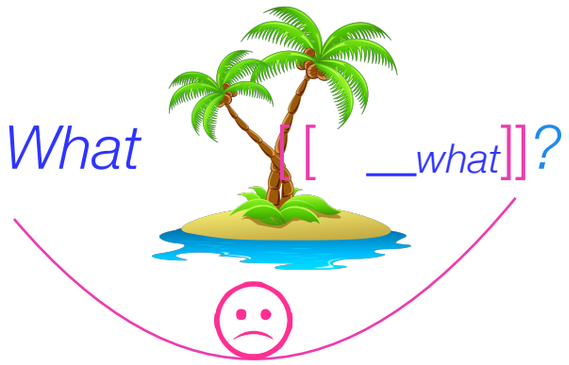
*This turns out to work really well.*

*Dickson, Pearl & Futrell 2022, Dickson, Futrell, & Pearl 2024, Dickson, Futrell, & Pearl in prep*





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




# Syntactic islands

involve *wh*-dependencies.

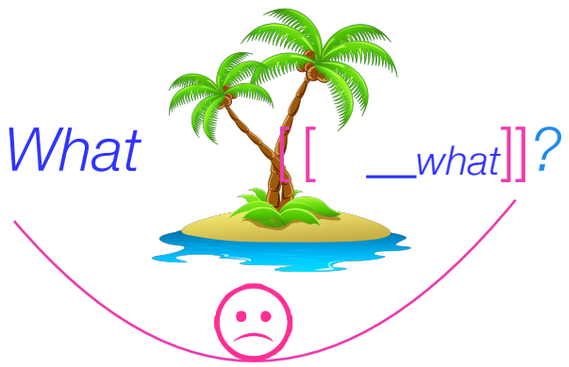
*This kitty was bought as a present for someone.*

*Lily thinks this kitty is pretty.*

What's going on here?

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

*What does Lily think is pretty, and who does she think it's for?*



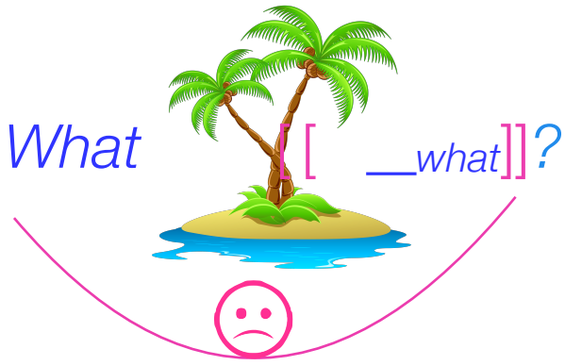

# Syntactic islands

involve *wh*-dependencies.

What's going on here?

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

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




# Syntactic islands

involve *wh*-dependencies.

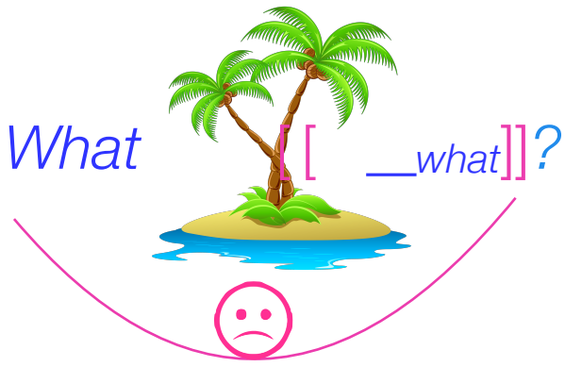
What's going on here?

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

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

This dependency is strongly dispreferred in English.

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

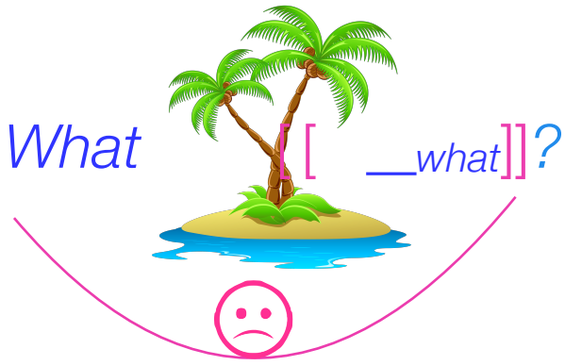



# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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




# Syntactic islands

involve *wh*-dependencies.

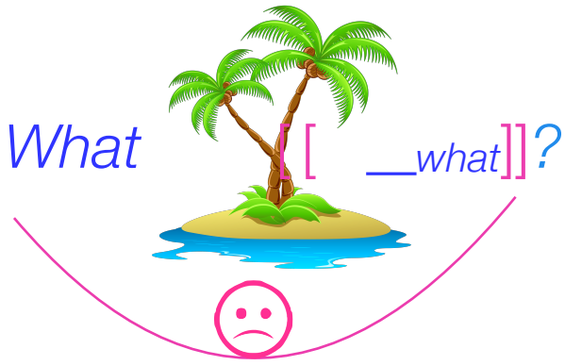
syntactic island (Ross 1967)

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

Jack is somewhat tricksy.

He claimed he bought something.

*What did Jack make the claim that he bought __what?*




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

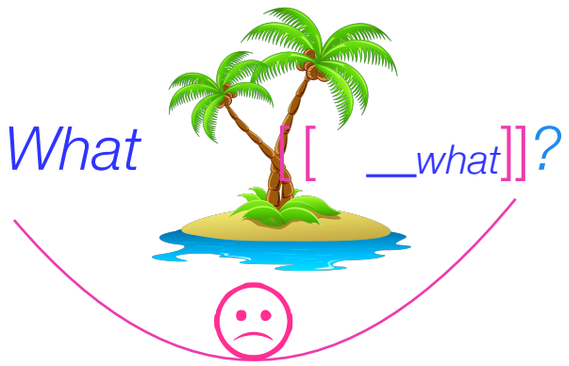
*What did Jack make the claim that he bought __what?* Complex NP island

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth wondered if he actually did and what it was.

*What did Elizabeth wonder whether Jack bought __what?*




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

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

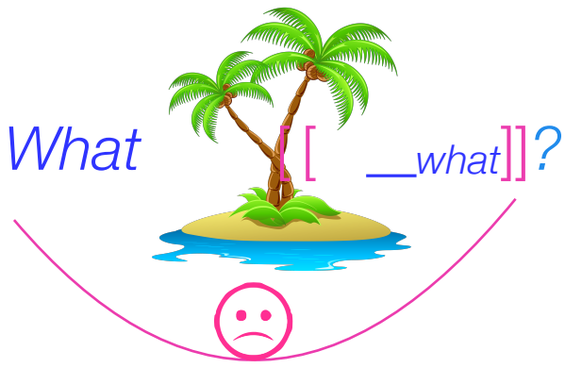
*What did Elizabeth wonder whether Jack bought __what?* Whether island

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth worried it was something dangerous.

*What did Elizabeth worry if Jack bought __what?*




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

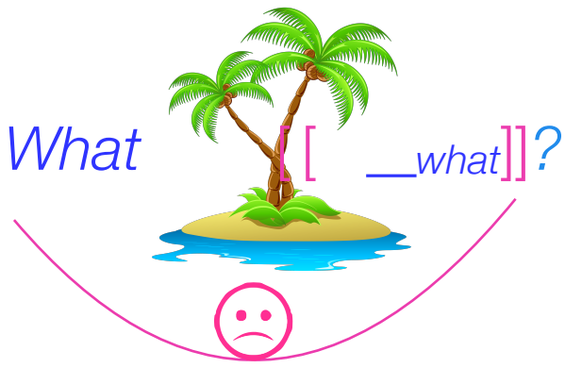
*What did Jack make the claim that he bought __what?* Complex NP island

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

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

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

(Chomsky 1965, Ross 1967)




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

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

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

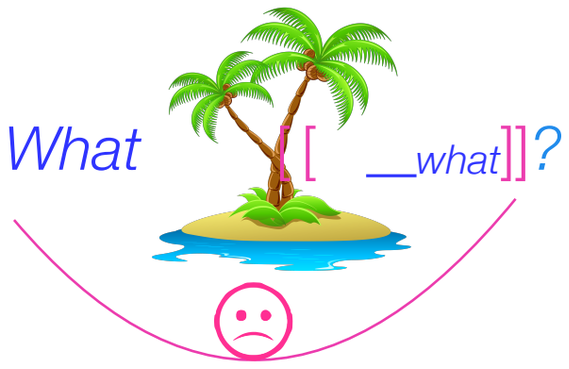
*What did Elizabeth worry if Jack bought __what?* Adjunct island

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

*What did Elizabeth think __what?*

✔

Elizabeth




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

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

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

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

Elizabeth

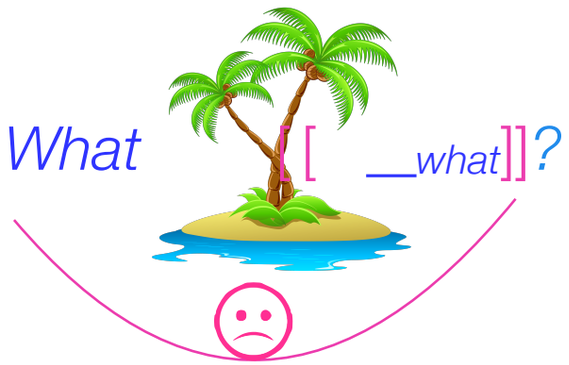


Jack

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

*What did Elizabeth think Jack said __what?*

✔




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

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

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

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

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

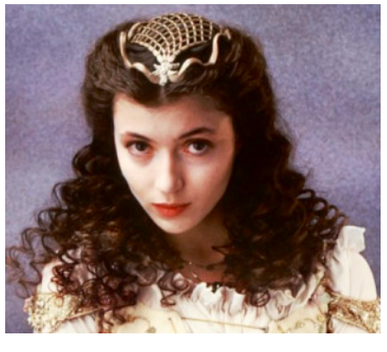
Elizabeth

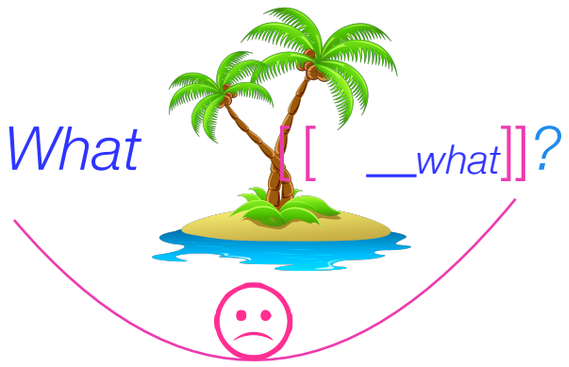
Jack

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

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

Lily






# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

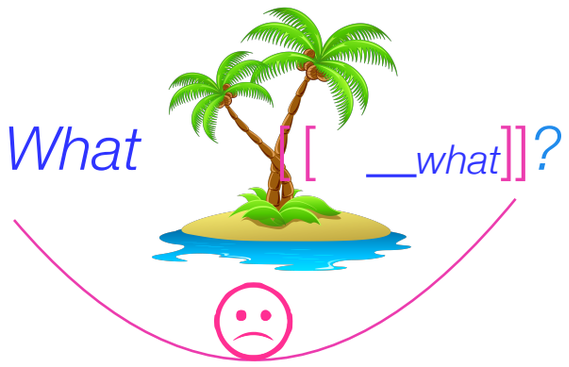
*Who does Lily think the kitty for __who is pretty?* Subject island

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

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

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

English adults judge these island-crossing dependencies to be far less acceptable than many others, including others that are very similar except that they don't cross syntactic islands (Sprouse et al. 2012).




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

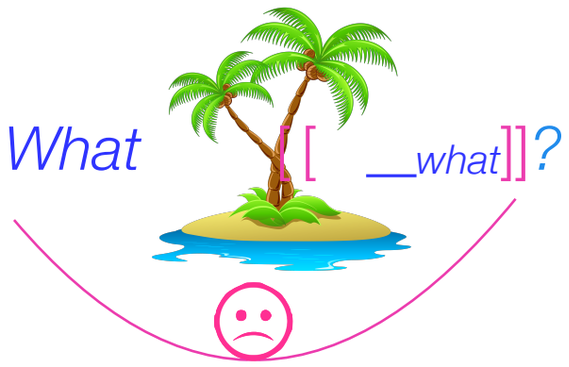
*Who does Lily think the kitty for __who is pretty?* Subject island

*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

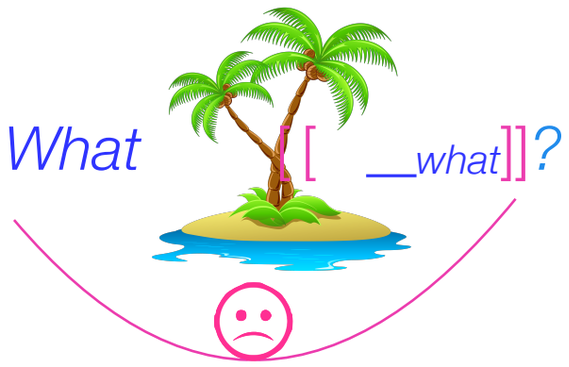
*What did Elizabeth wonder whether Jack bought __what?* Whether island

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

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

*What did Elizabeth say that Jack saw __what?*

*What did Elizabeth whine that Jack saw __what?*




# Syntactic islands

involve *wh*-dependencies.

syntactic island (Ross 1967)

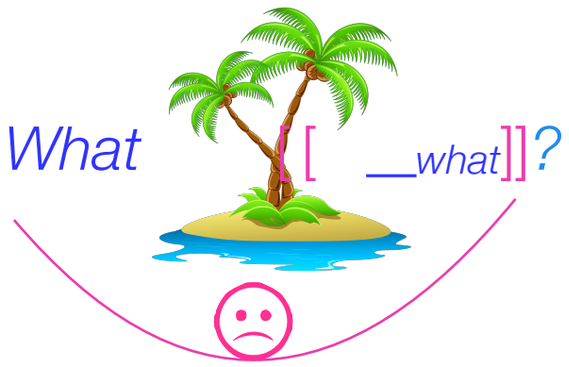
*Who does Lily think the kitty for __who is pretty?* *Subject island*

*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.

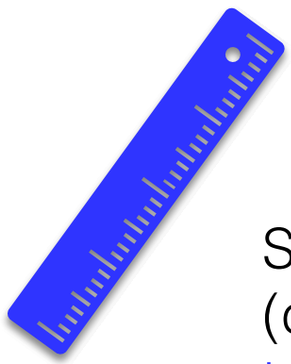
syntactic island (Ross 1967)

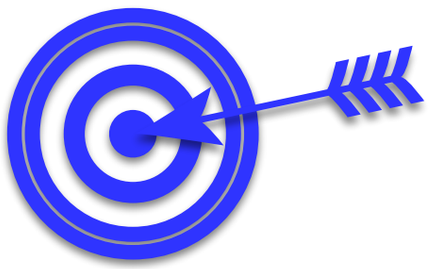
*Who does Lily think the kitty for __who is pretty?* Subject island

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

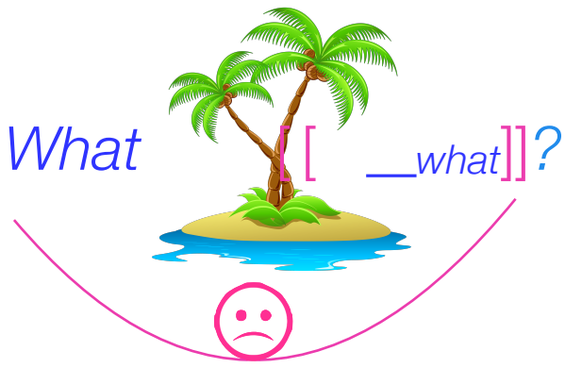
*What did Elizabeth wonder whether Jack bought __what?* Whether island

*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)**

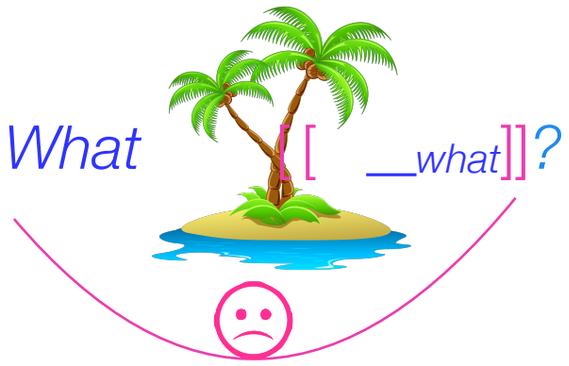
*Who* __*who*?

*Who* [CP… __*who*]?

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

*Who* [non-island ]?

*Who* [island ]?




# Syntactic islands

Adult judgments

= behavioral target outcome

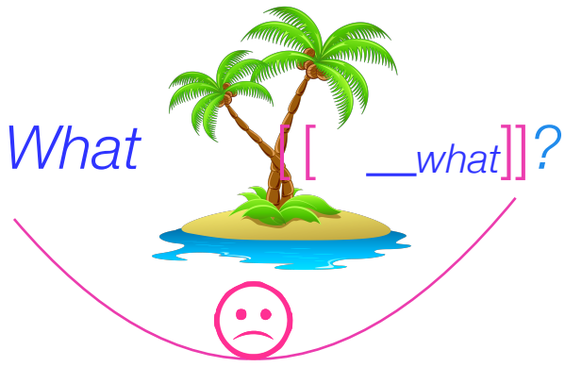
Adult knowledge as measured by acceptability judgment behavior

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

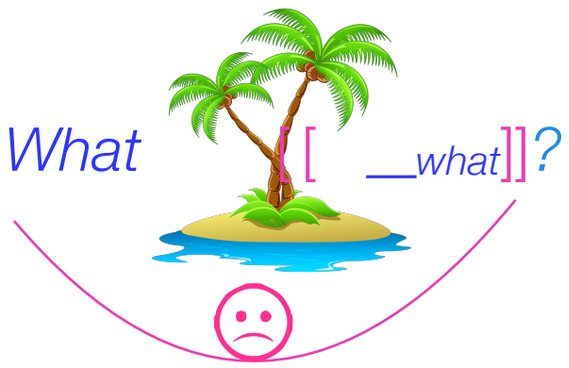
Adult knowledge as measured by acceptability judgment behavior

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

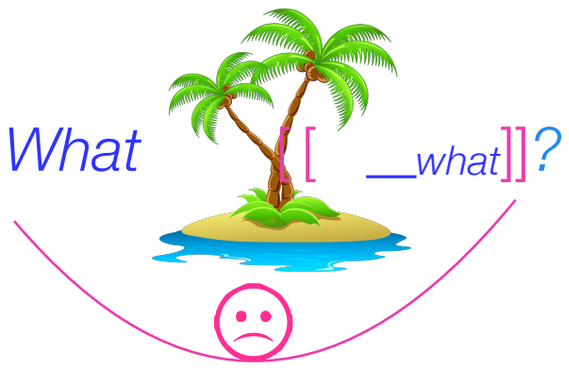
Adult knowledge as measured by acceptability judgment behavior

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

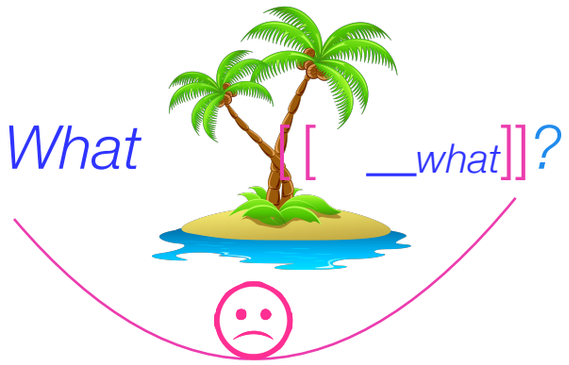
Adult knowledge as measured by acceptability judgment behavior

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

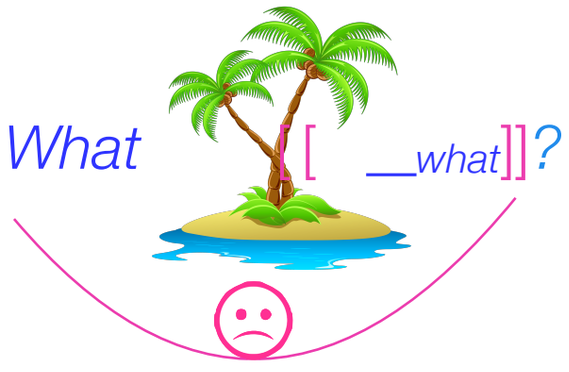
= behavioral target outcome

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

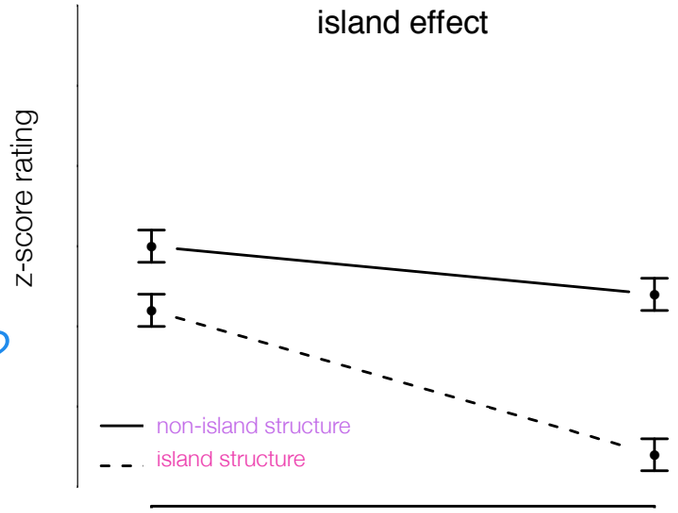
Adult knowledge as measured by acceptability judgment behavior

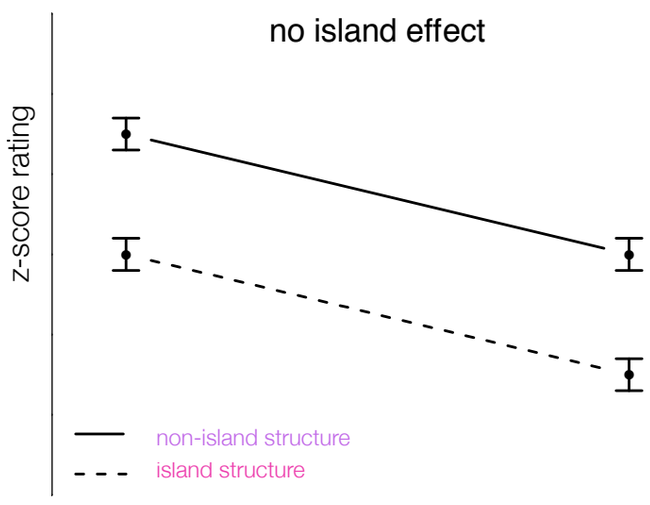
length of dependency (**matrix** **vs.** **embedded**) X presence of an island structure (**non-island** **vs.** **island**)

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

*Who* [non-island ]?

*Who* [island ]?

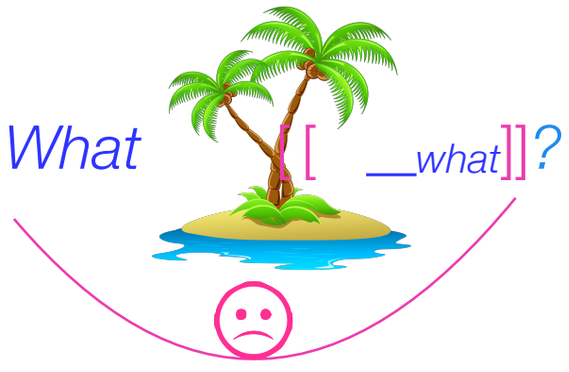



*Who* __who?

*Who* [CP… __who]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

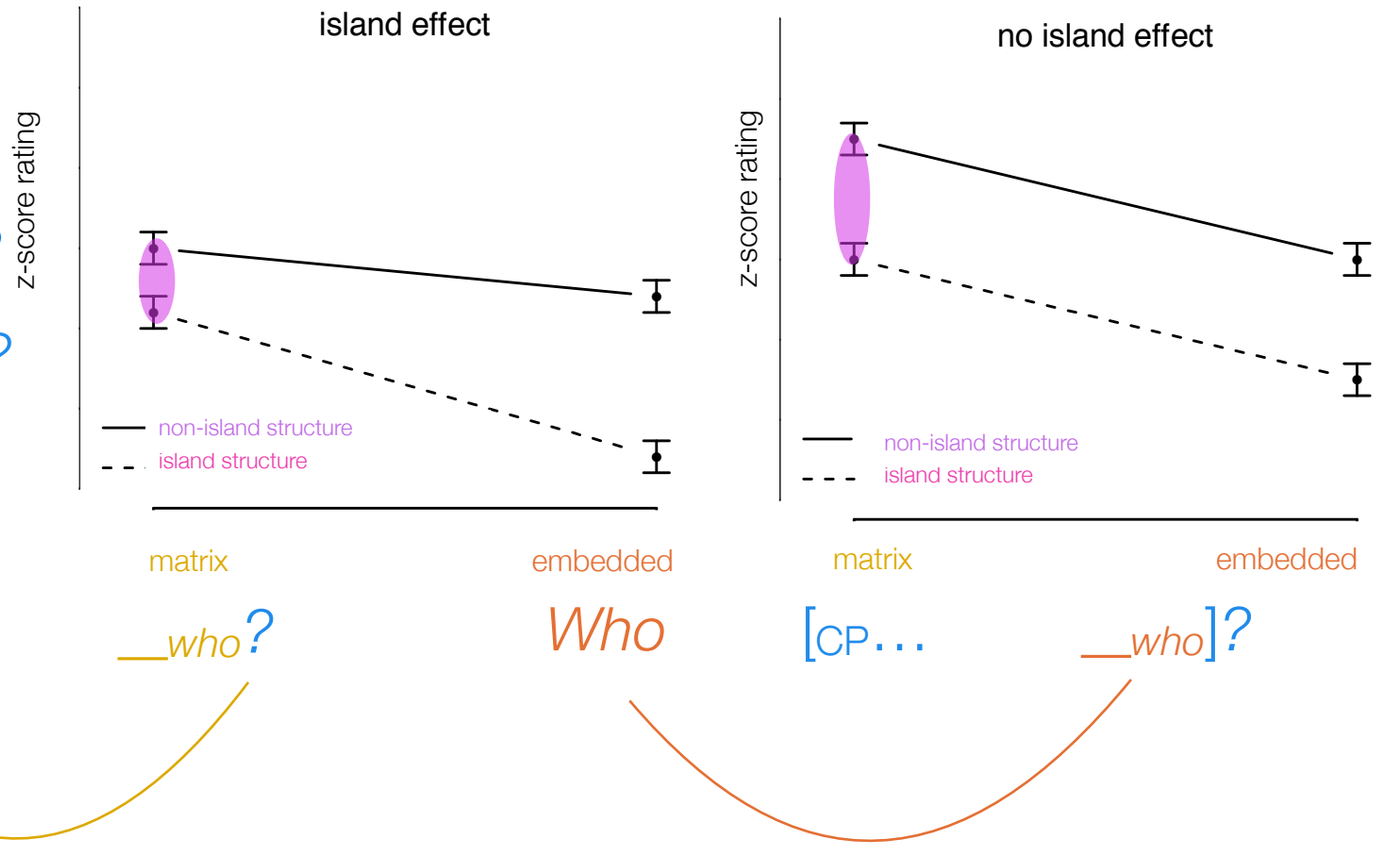
Adult knowledge as measured by acceptability judgment behavior

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 ]?

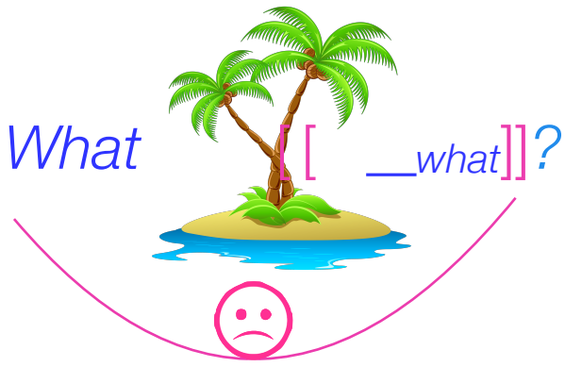
*Who* [island ]?




*Who* __who?

*Who* [CP… __who]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

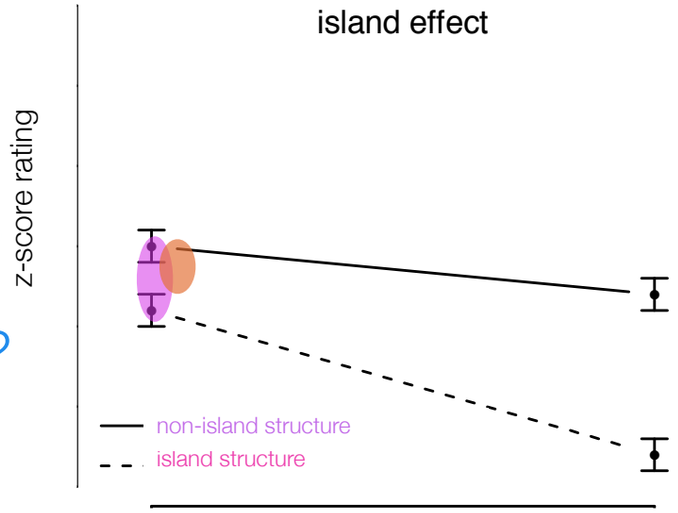
Adult knowledge as measured by acceptability judgment behavior

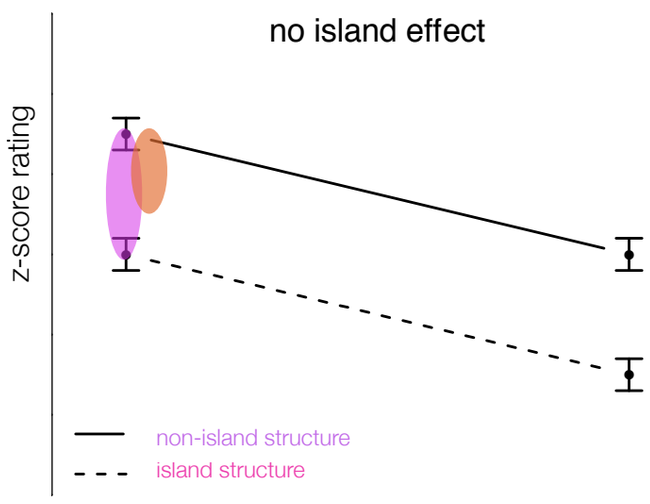
length of dependency (**matrix** **vs.** **embedded**) X presence of an island structure (**non-island** **vs.** **island**)

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

*Who* [non-island ]?

*Who* [island ]?

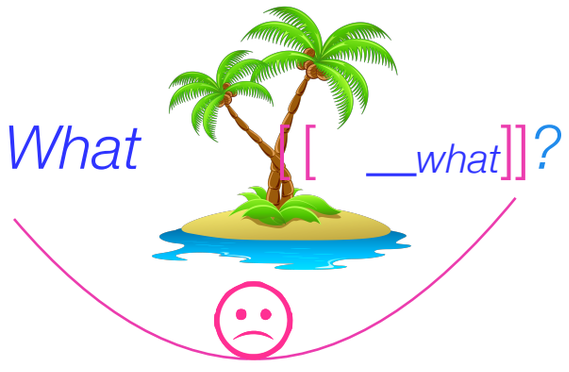



*Who* __*who*?

*Who* [CP… __*who*]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

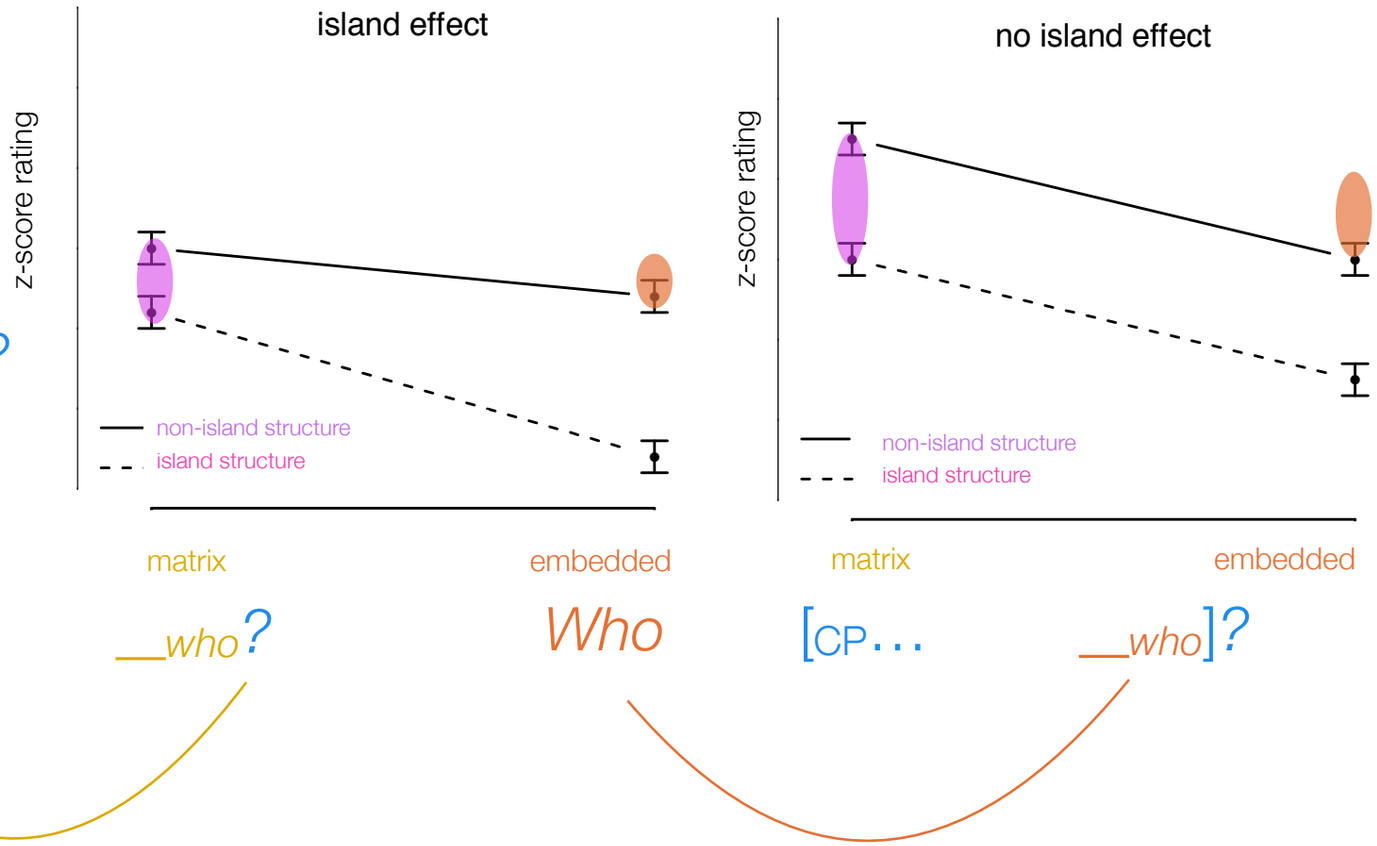
Adult knowledge as measured by acceptability judgment behavior

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 ]?

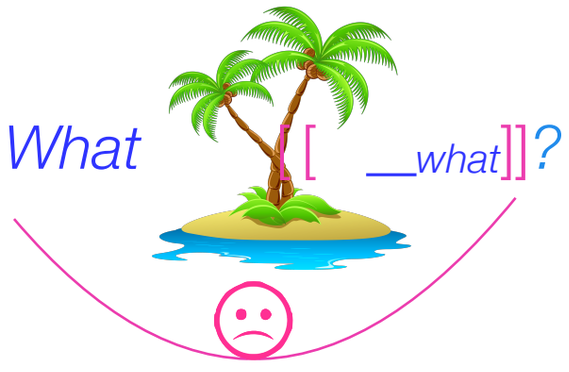
*Who* [island ]?




*Who* __*who*?

*Who* [CP… __*who*]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

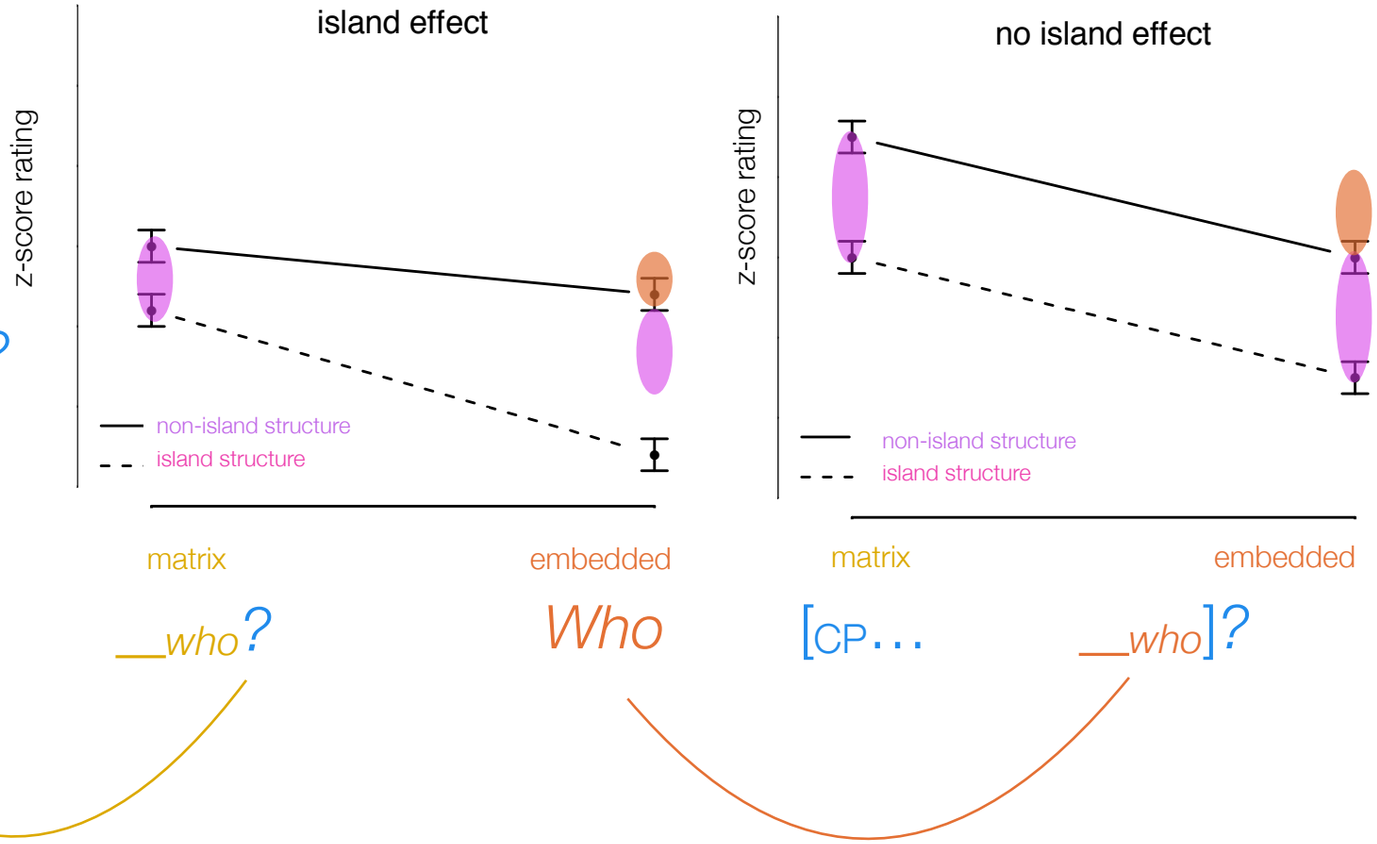
Adult knowledge as measured by acceptability judgment behavior

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 ]?

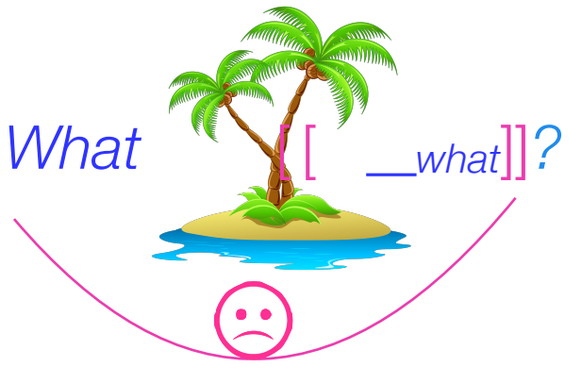
*Who* [island ]?




*Who* __*who*?

*Who* [CP… __*who*]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

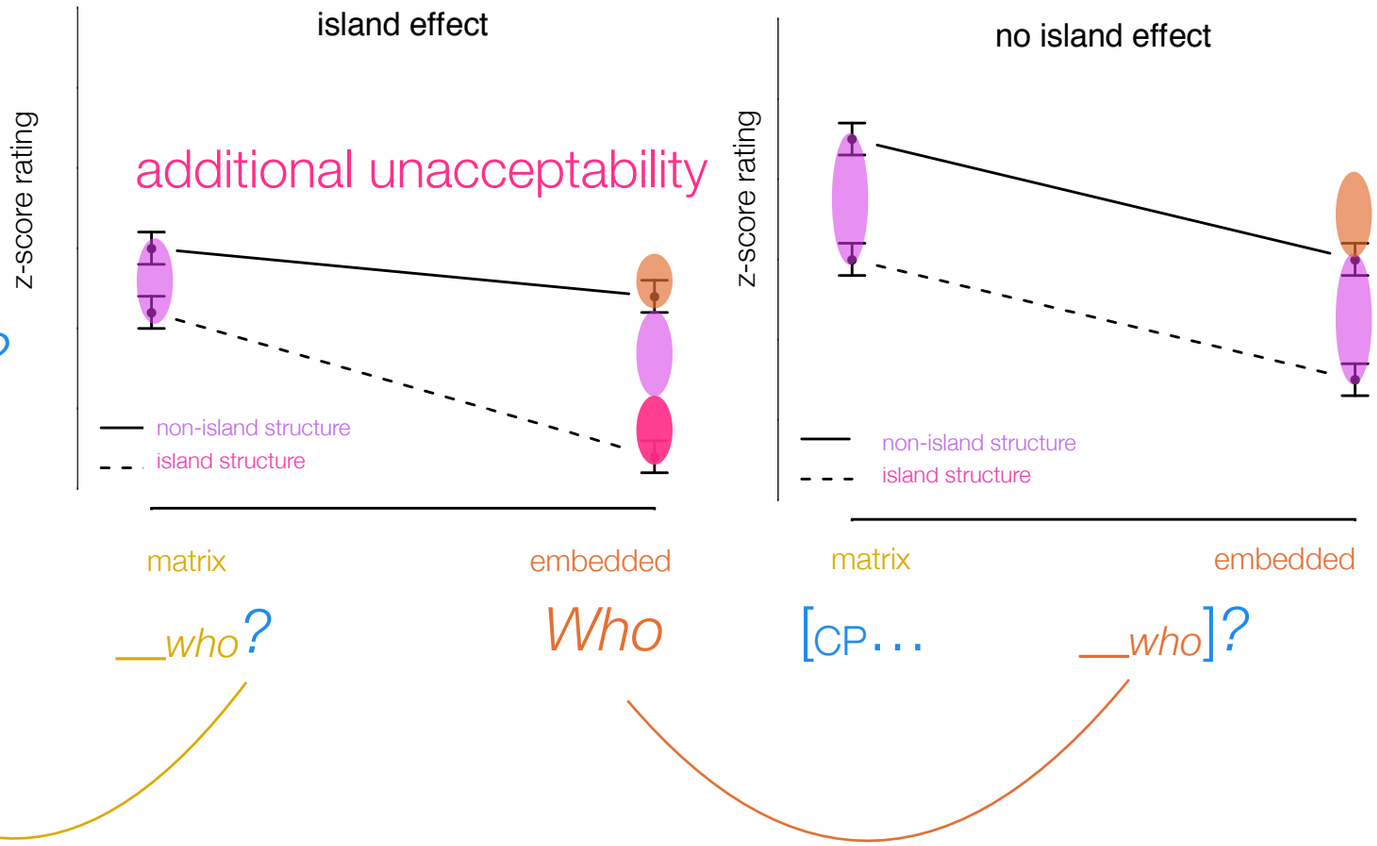
Adult knowledge as measured by acceptability judgment behavior

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

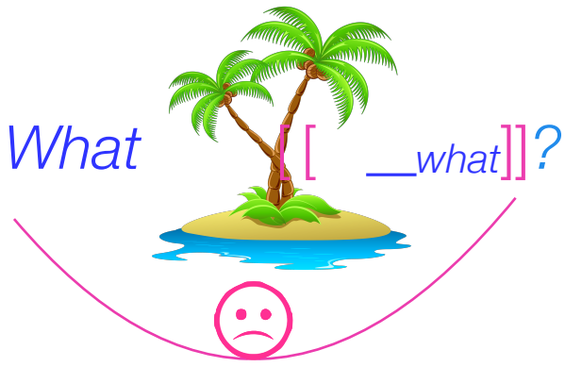
Syntactic island = **superadditive** interaction of the two factors

*Who* [non-island ]?

*Who* [island ]?




*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

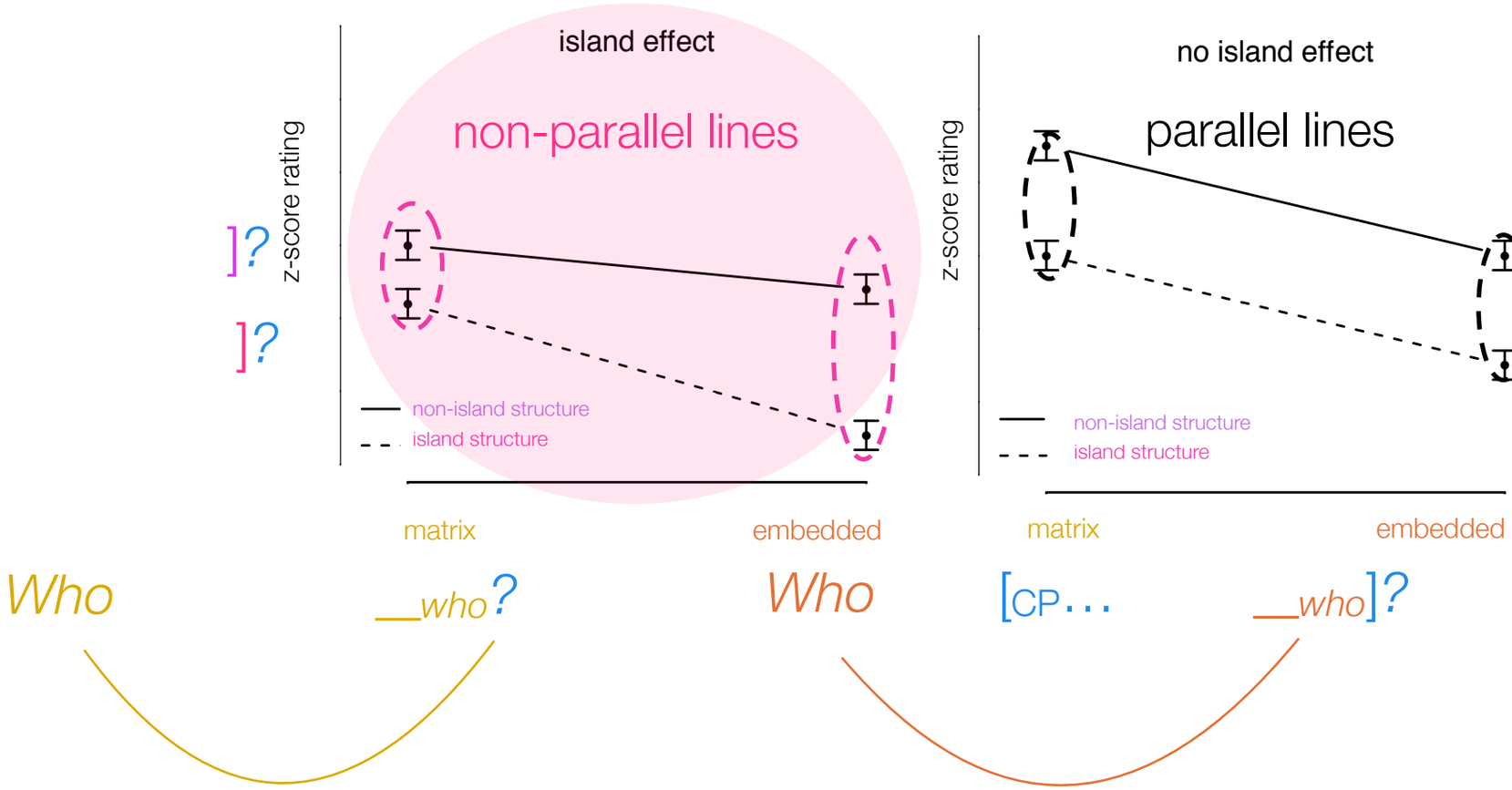
Adult knowledge as measured by acceptability judgment behavior

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

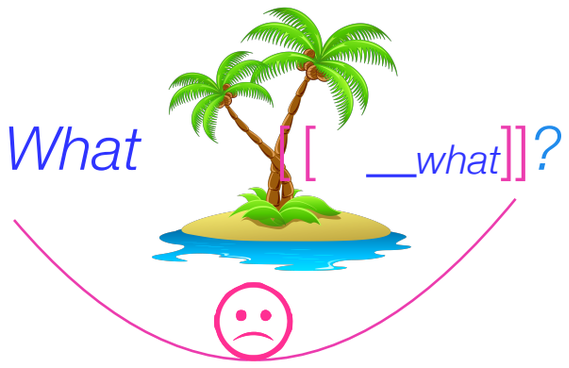
Syntactic island = **superadditive** interaction of the two factors

*Who* [non-island ]?

*Who* [island ]?




*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

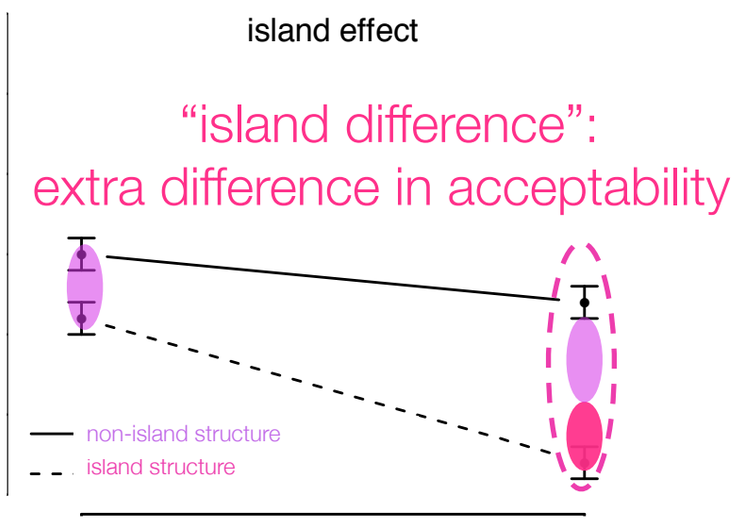
Adult knowledge as measured by acceptability judgment behavior

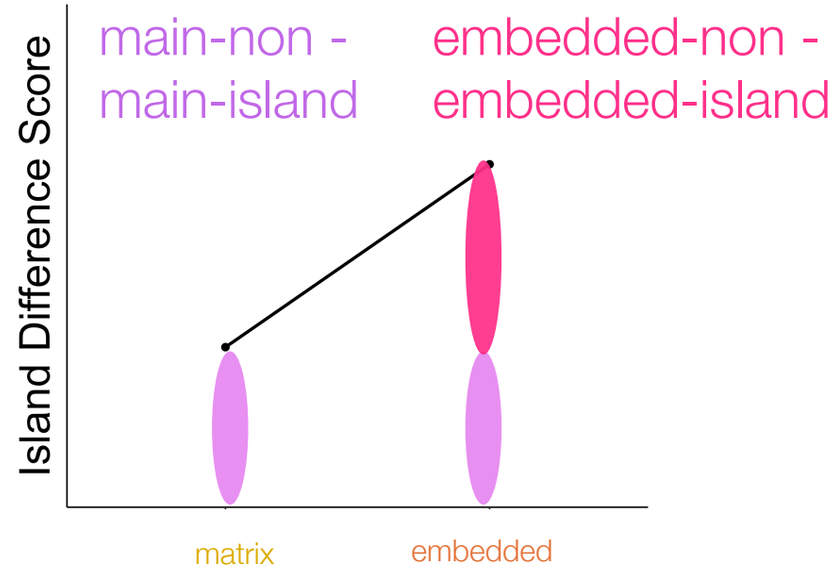
length of dependency (**matrix** **vs.** **embedded**) X presence of an island structure (**non-island** **vs.** **island**)

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

*Who* [non-island ]?

*Who* [island ]?

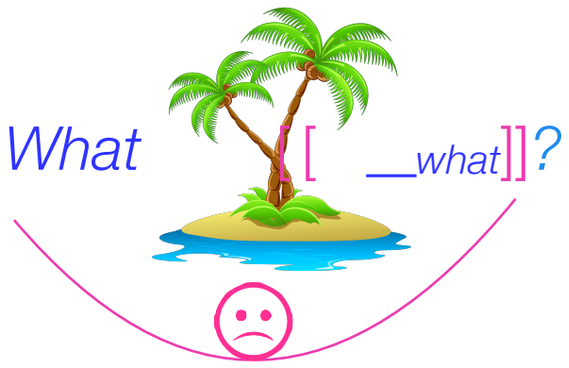



*Who* __*who*?

*Who* [CP… __*who*]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

= behavioral target outcome

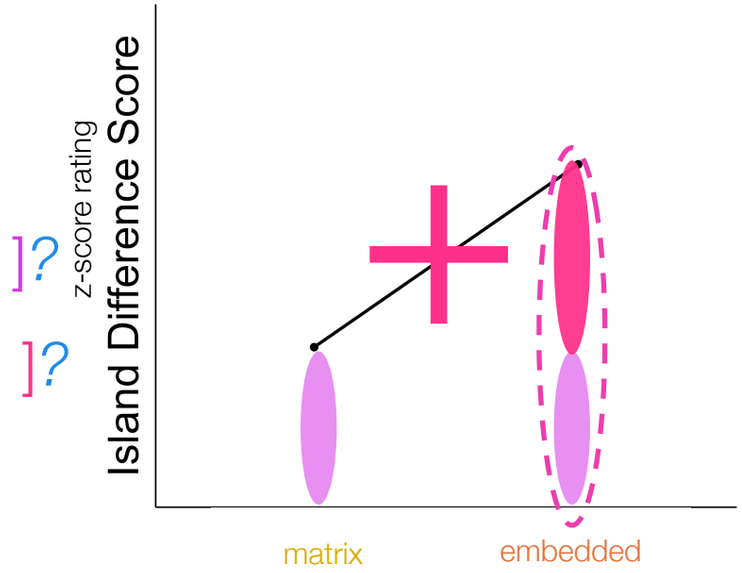
Adult knowledge as measured by acceptability judgment behavior

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 ]?

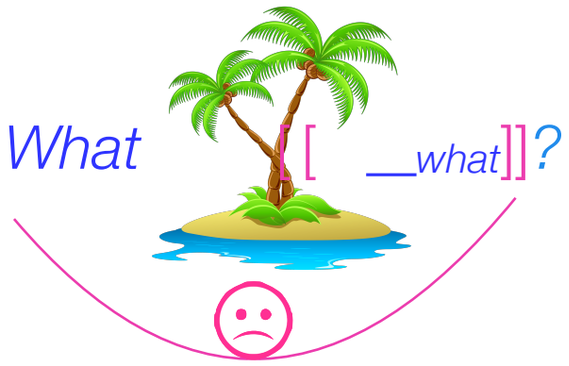



"island difference":
superadditivity = positive difference
(positive slope)

*Who* __*who*?

*Who* [CP… __*who*]?

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

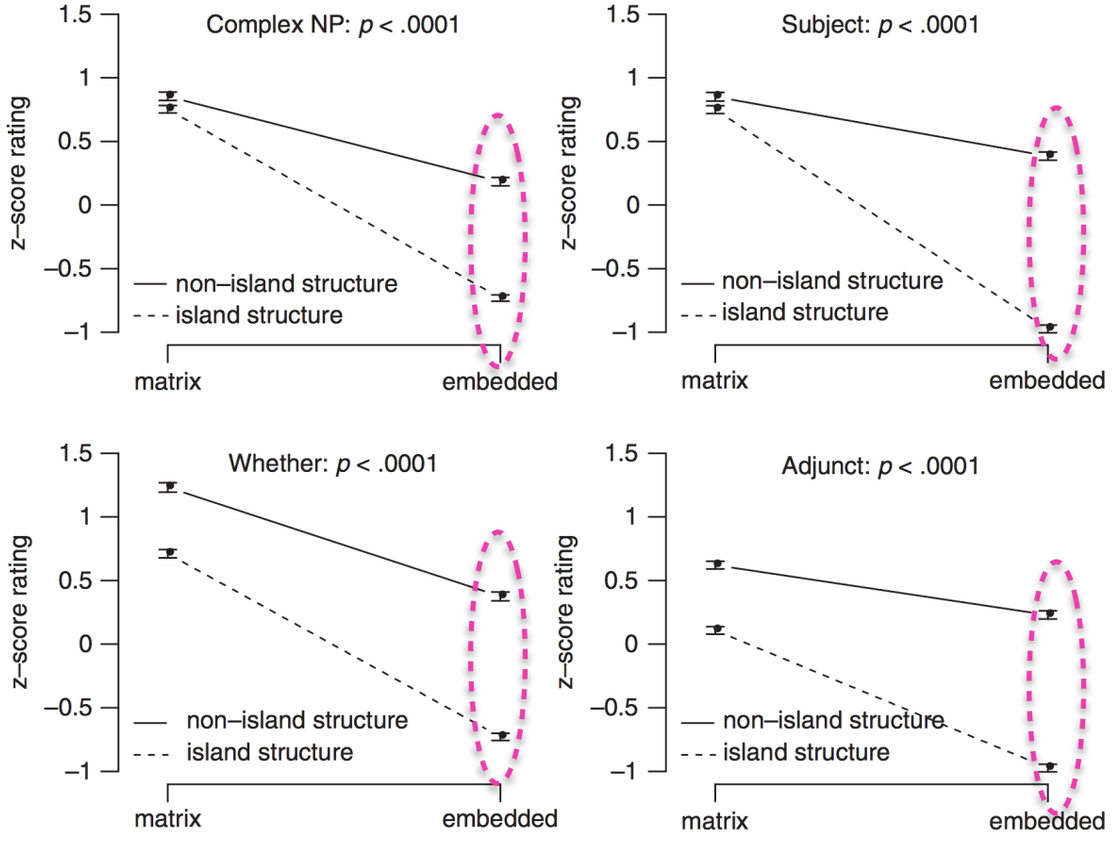
= behavioral target outcome

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

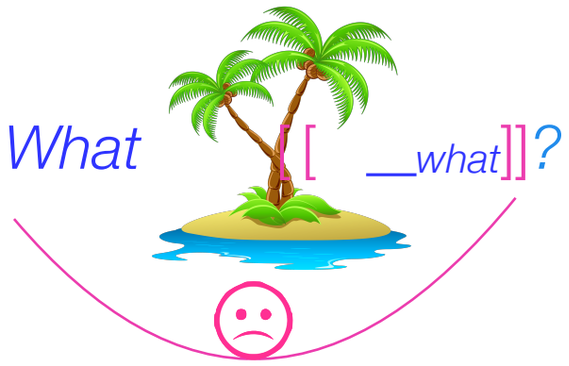



(non-parallel lines)

✓

superadditivity for all four island types

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

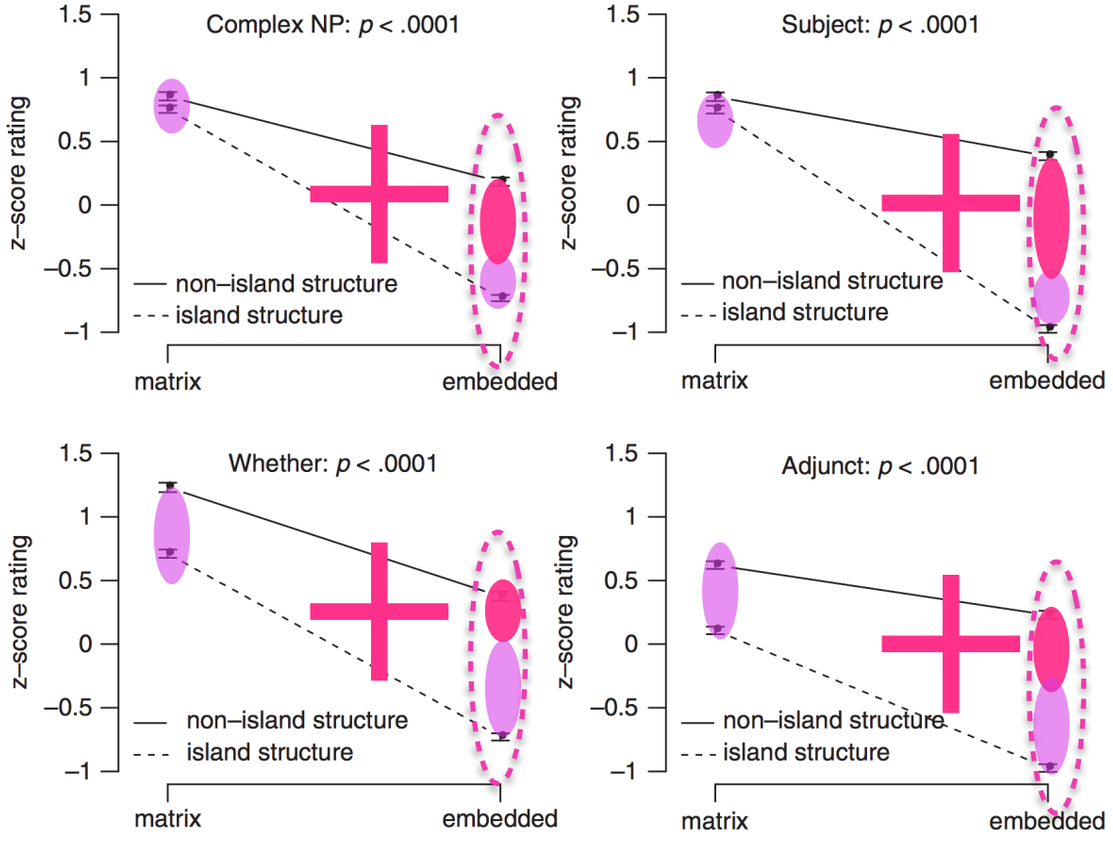
= behavioral target outcome

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

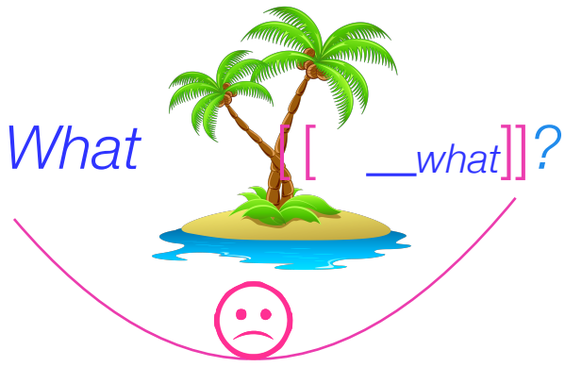



(positive difference)

✓

superadditivity for all four island types

*Sprouse et al. 2012*




# Syntactic islands

Adult judgments

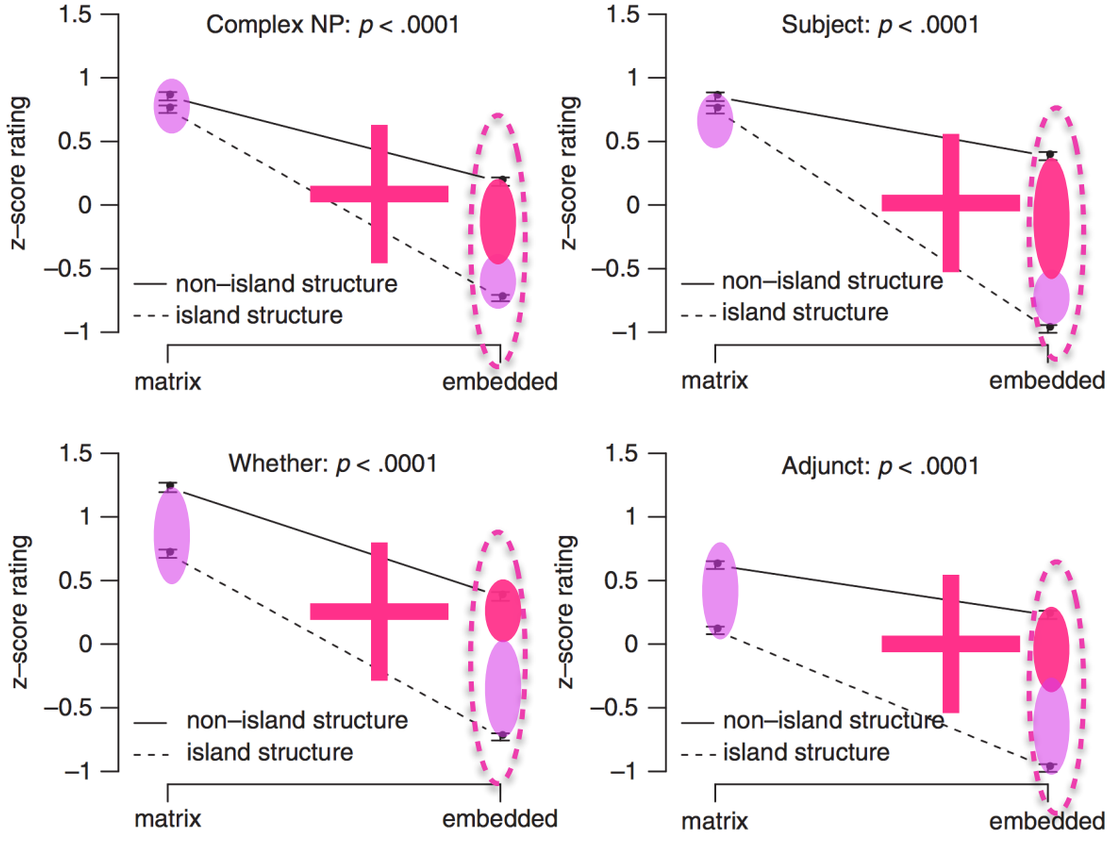
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

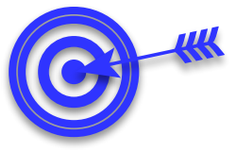
Syntactic island = **superadditive** interaction of the two factors

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

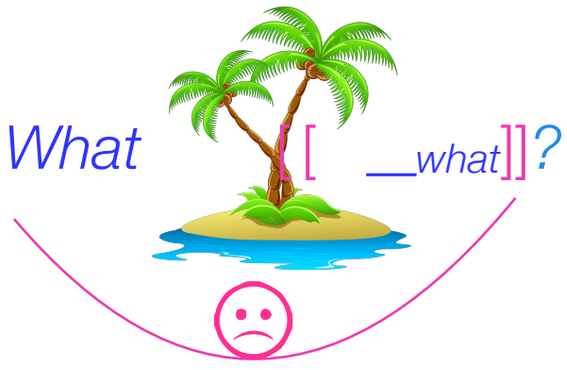



✓ superadditivity for all four island types

= knowledge that dependencies crossing these island structures are dispreferred.



*Sprouse et al. 2012*




## Syntactic islands

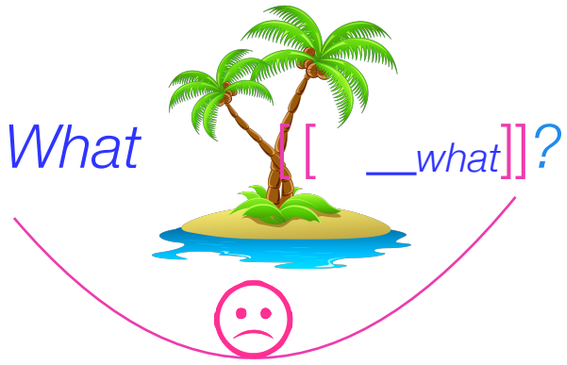
Child judgments

= 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

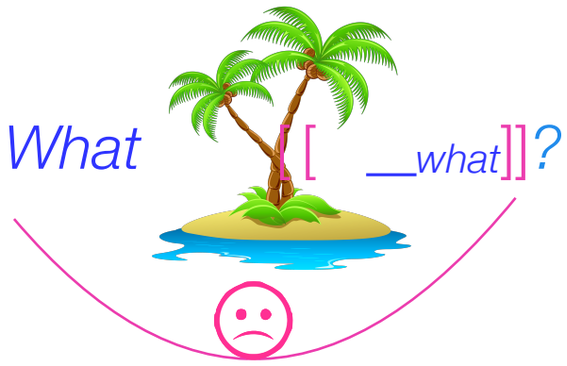
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

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

context

*De Villiers et al. 2008*




## Syntactic islands

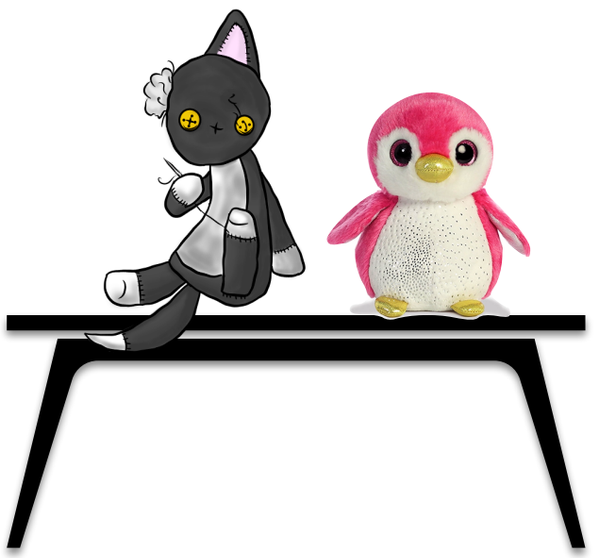
Child judgments

= behavioral target outcome

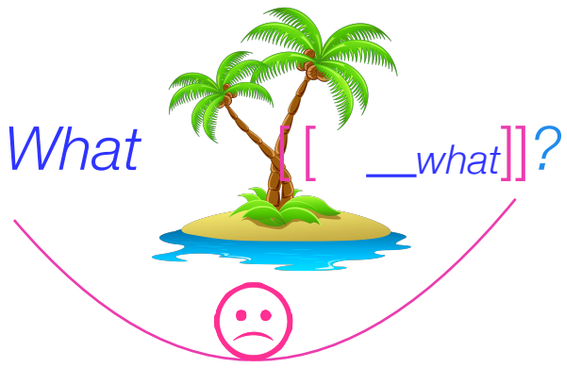
Child knowledge as measured by preferred interpretation behavior

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

context



*De Villiers et al. 2008*




## Syntactic islands

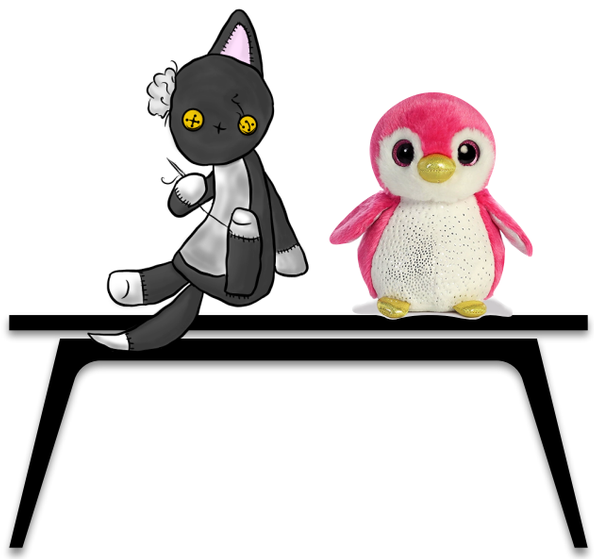
Child judgments

= behavioral target outcome

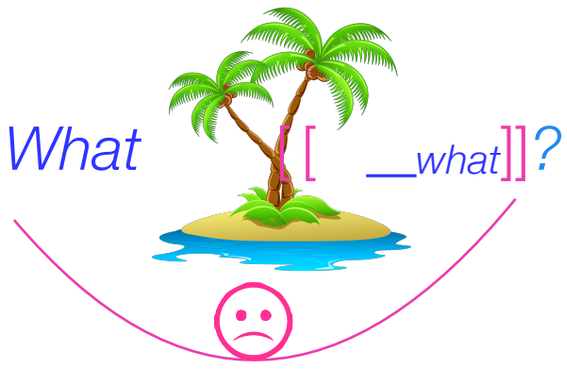
Child knowledge as measured by preferred interpretation behavior

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

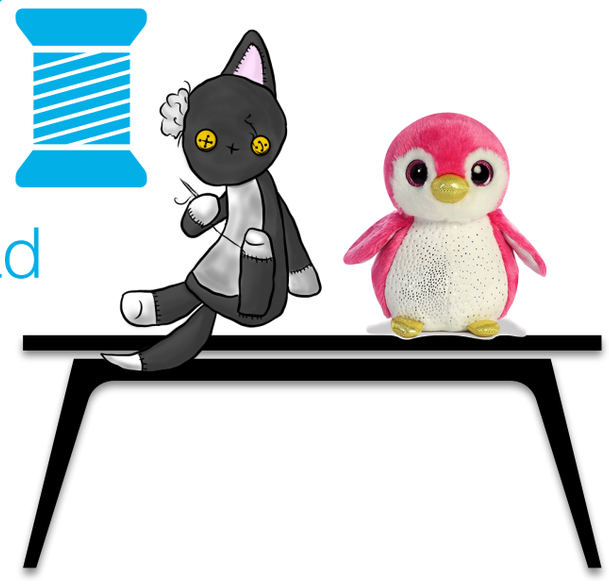
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

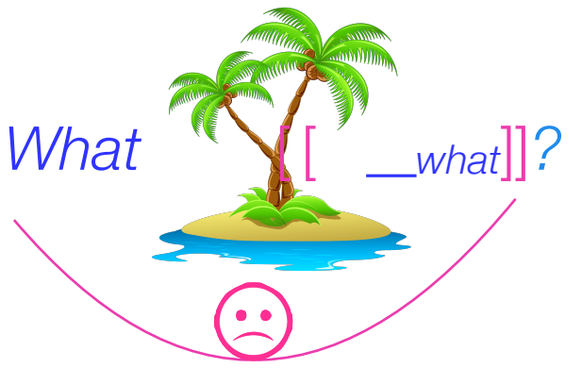
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

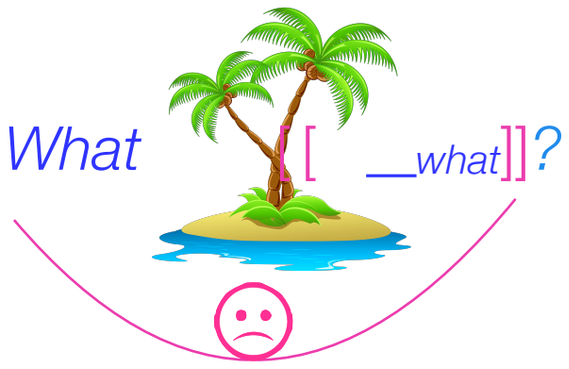
Child knowledge as measured by preferred interpretation behavior

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

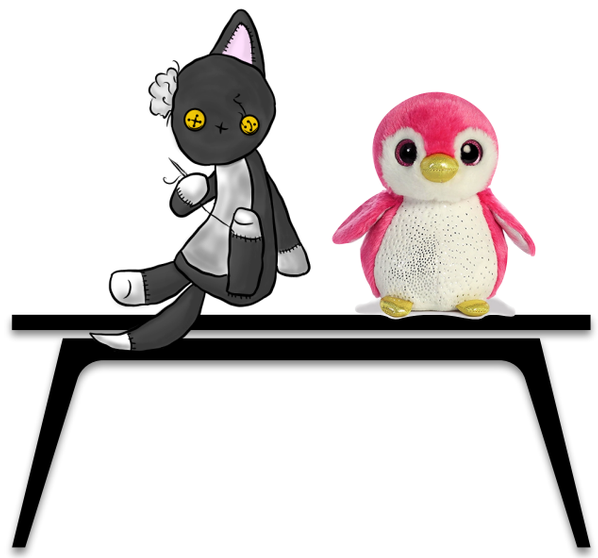
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

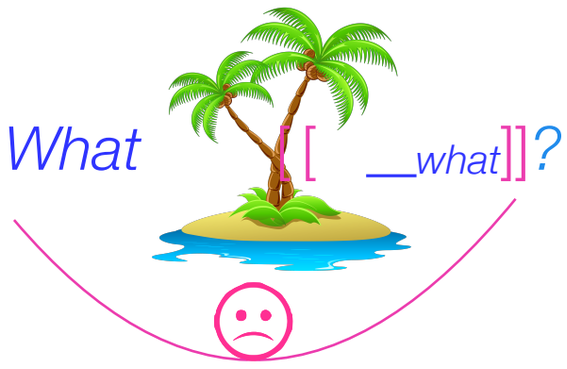
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

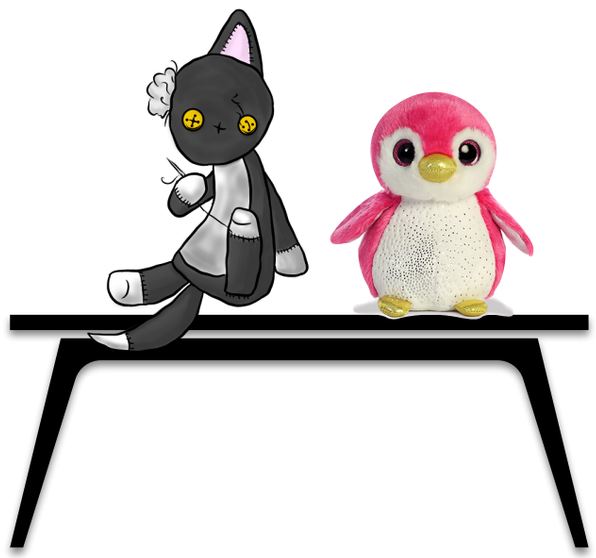
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

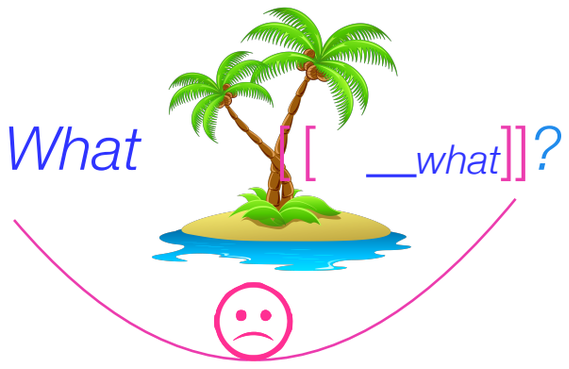
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

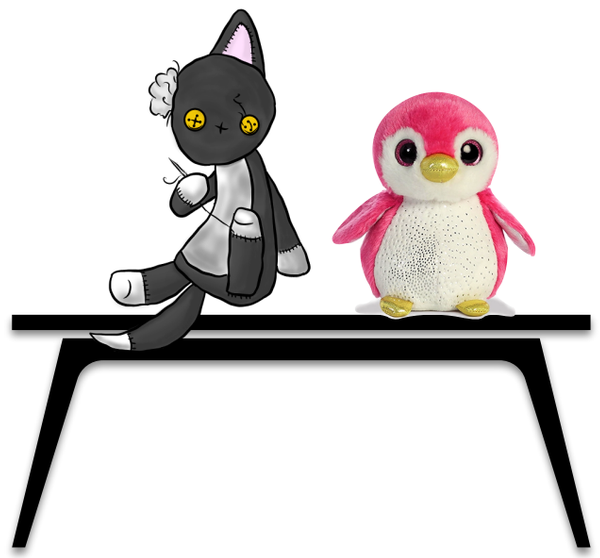
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

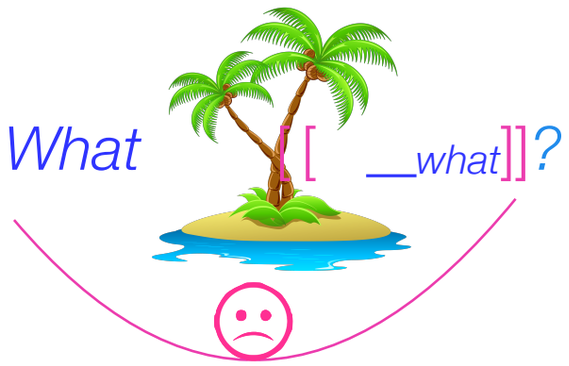
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

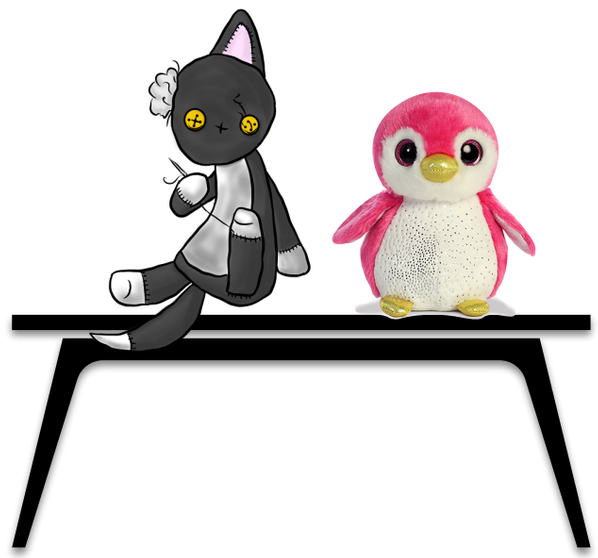
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

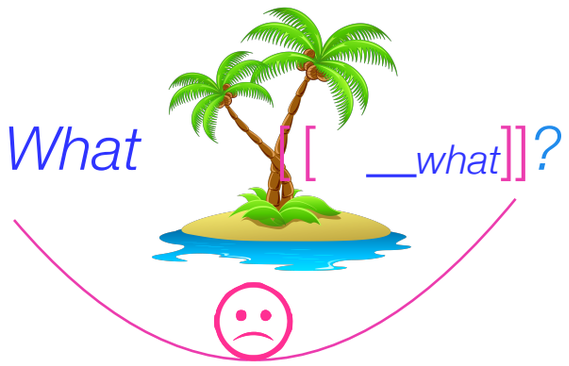
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

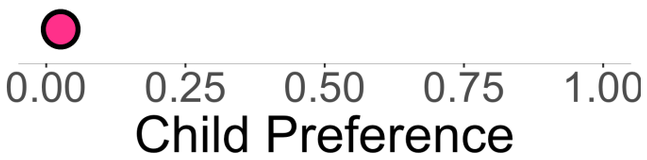
Child judgments

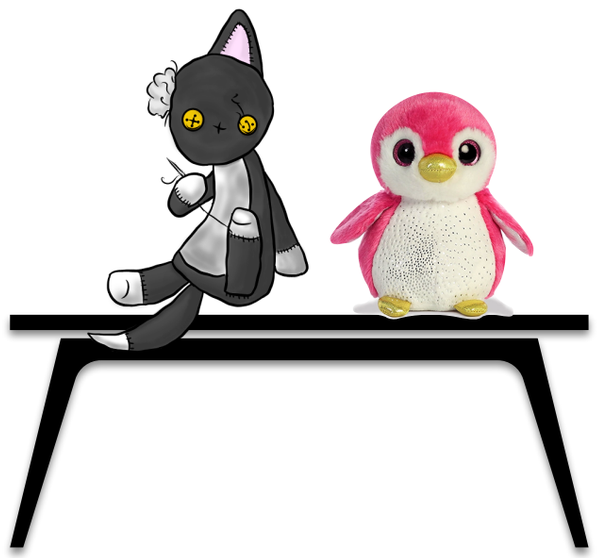
= behavioral target outcome

Child knowledge as measured by preferred interpretation behavior

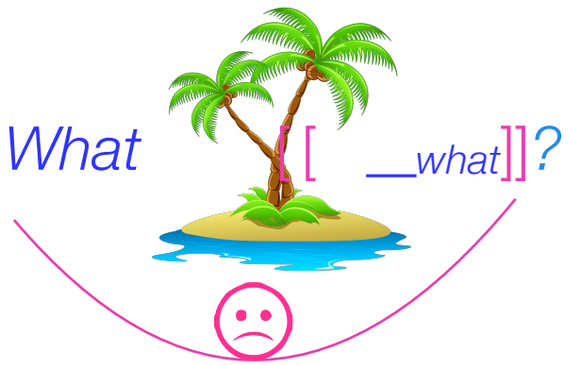
How do children prefer to interpret potentially ambiguous *wh*-questions?

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



*De Villiers et al. 2008*




# Syntactic islands

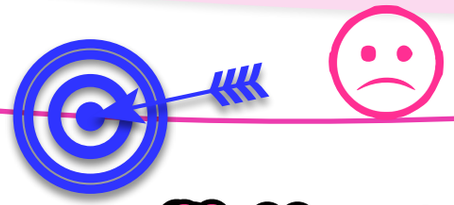
Child judgments

= behavioral target outcome

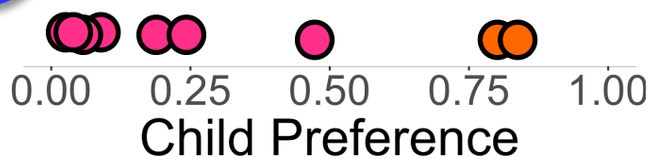
Child knowledge as measured by preferred interpretation behavior

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]]]]]]]]?*



+ other *wh*-dependencies




*Who did the little sister ask how to see __?*

*Who did the boy ask what to bring __?*

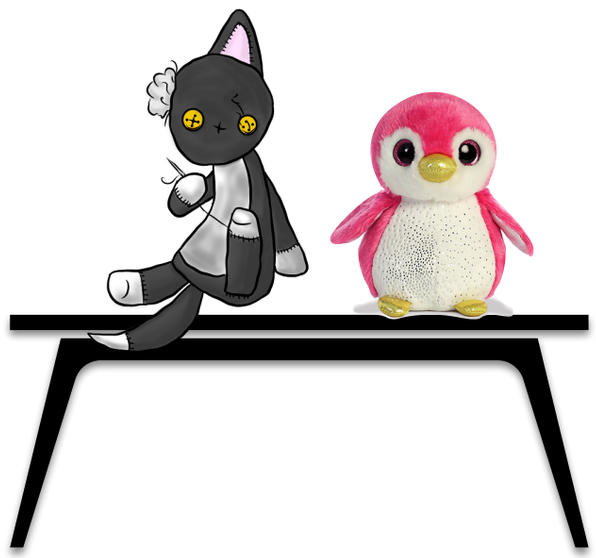
*How did the mom learn what to bake __?*

*How did the girl ask where to ride __?*

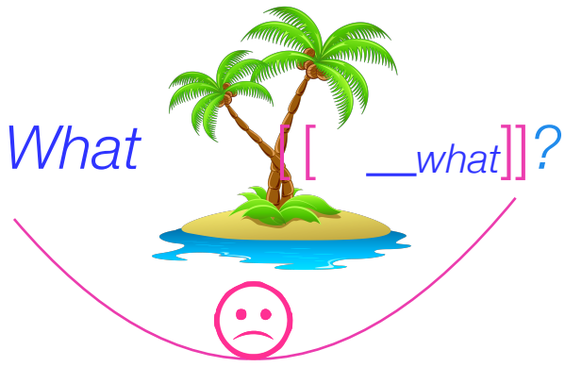
*How did the boy who sneezed drink the milk ___?*

*What did the mother say she bought __?*

*Who did the policewoman help to call __?*



*De Villiers et al. 2008*




## Syntactic islands

More adult judgments

= behavioral target outcome

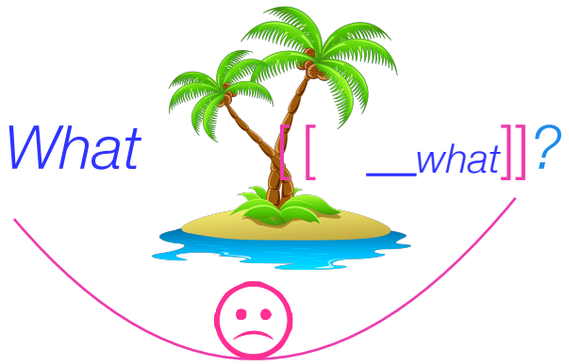
Adult knowledge as measured by acceptability judgment behavior

Liu et al 2019, 2022: 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 __ ]?

*Liu et al. 2019, 2022*




## Syntactic islands

More adult judgments

= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

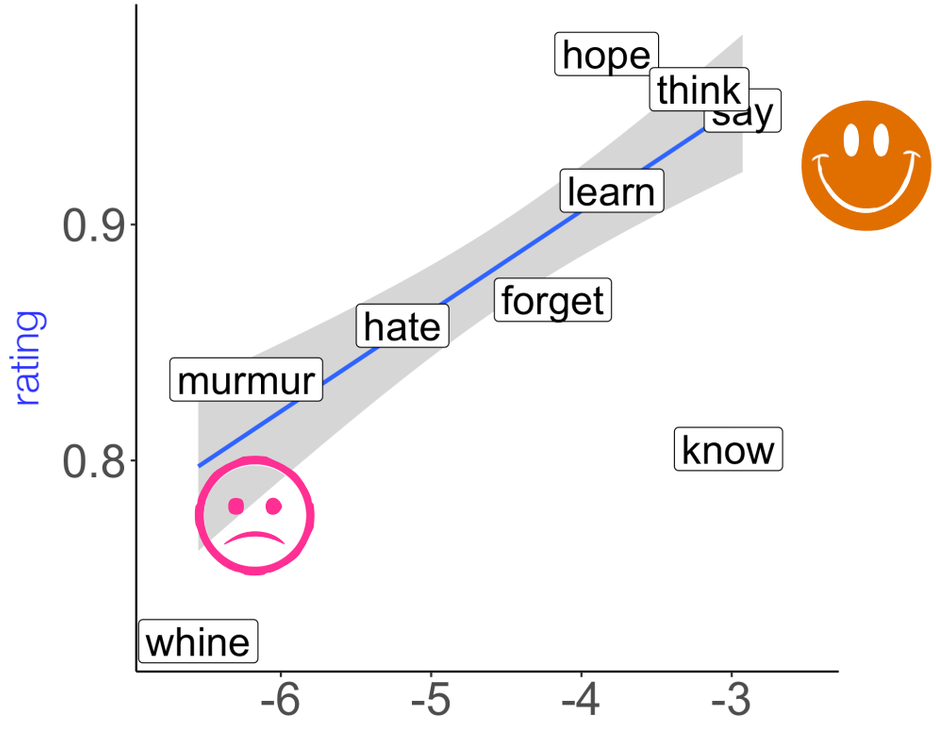
Liu et al 2019, 2022: Acceptability of *wh*-dependencies can depend on the lexical item in the main verb.

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

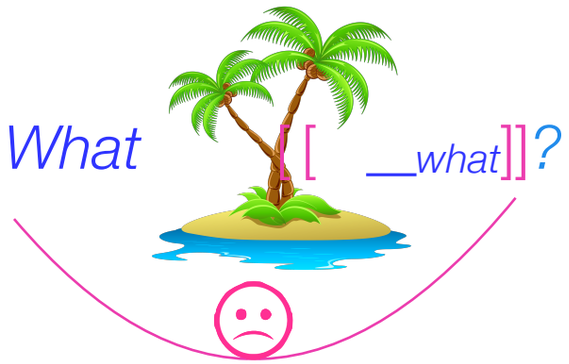
What did she think [that he saw __ ]?
What did she say [that he saw __ ]?

CP

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




*Liu et al. 2019, 2022*




## Syntactic islands

More adult judgments

= behavioral target outcome

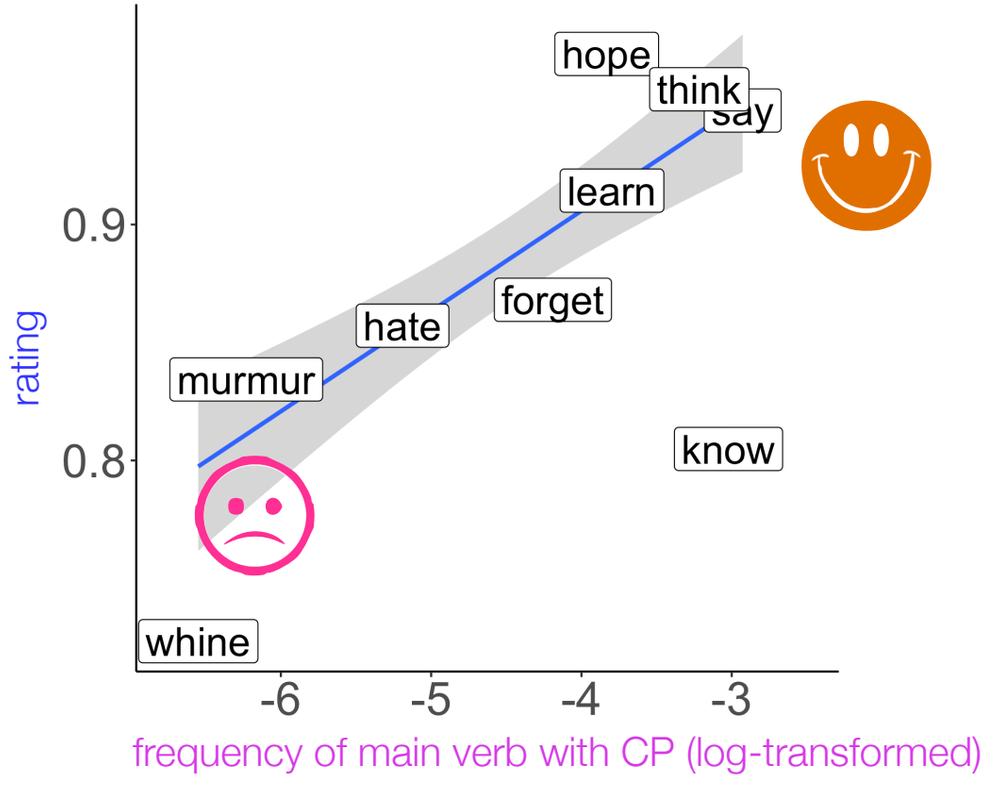
Adult knowledge as measured by acceptability judgment behavior

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

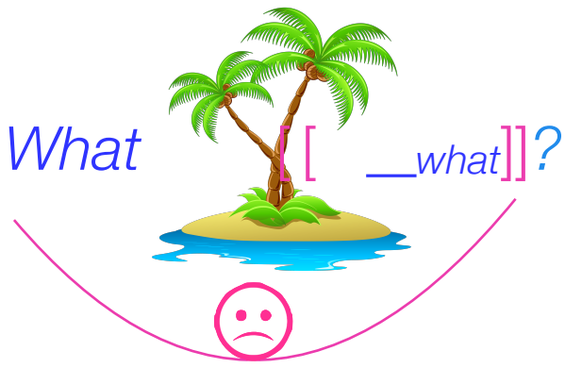
**What did she VERB [that he saw __ ]?**

Important pattern:

Positive correlation between main verb with CP frequency and judged acceptability.



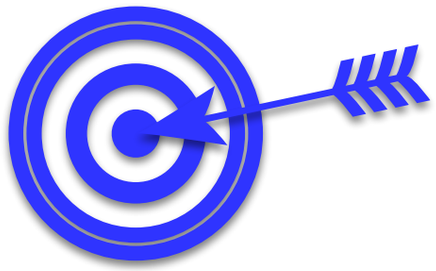

*Liu et al. 2019, 2022*




# Syntactic islands

Adult & child judgments

= behavioral target outcome

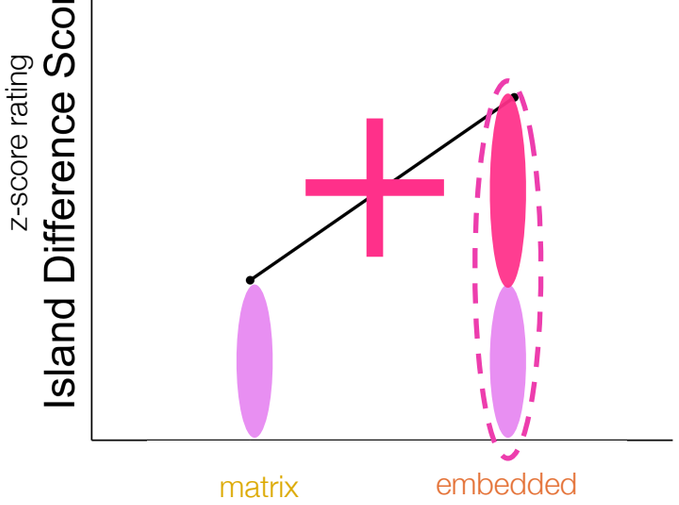


*Subject island*

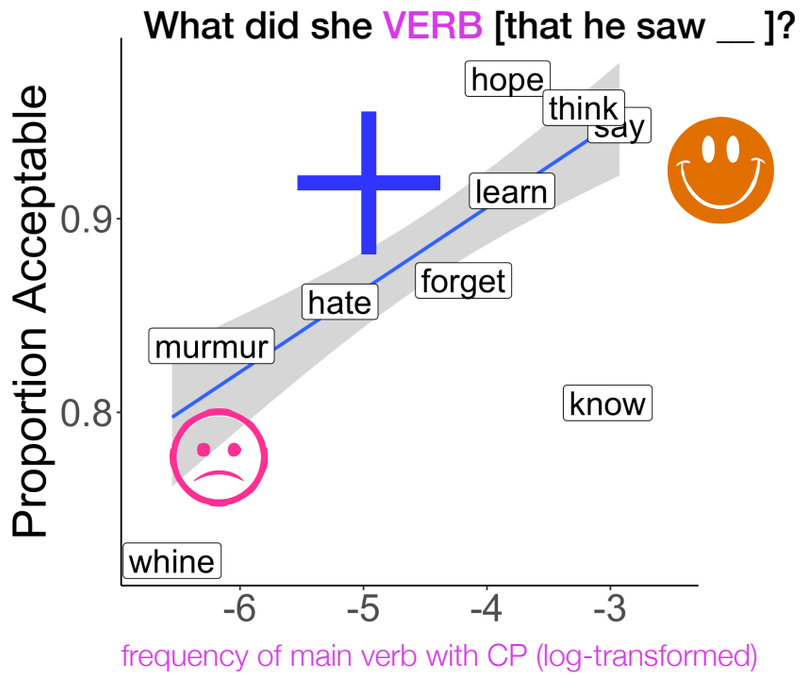
*Complex NP island*

*Whether island*

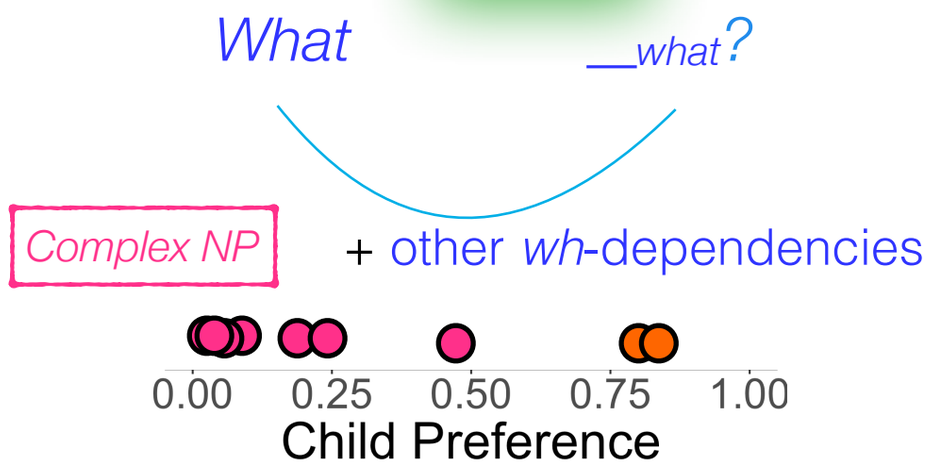
*Adjunct island*



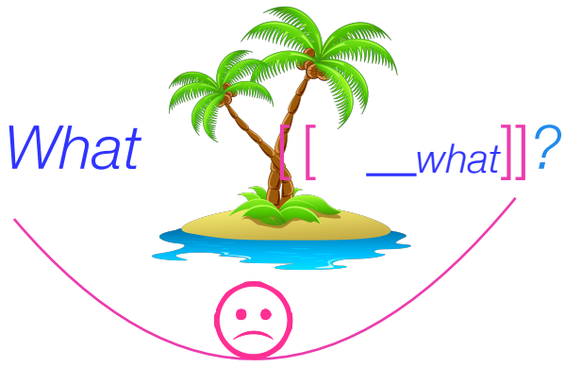

*Sprouse et al. 2012*




*Liu et al. 2019, 2022*



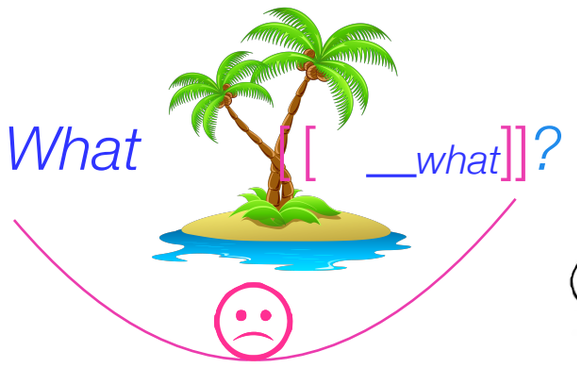

*De Villiers et al. 2008*




# Syntactic islands

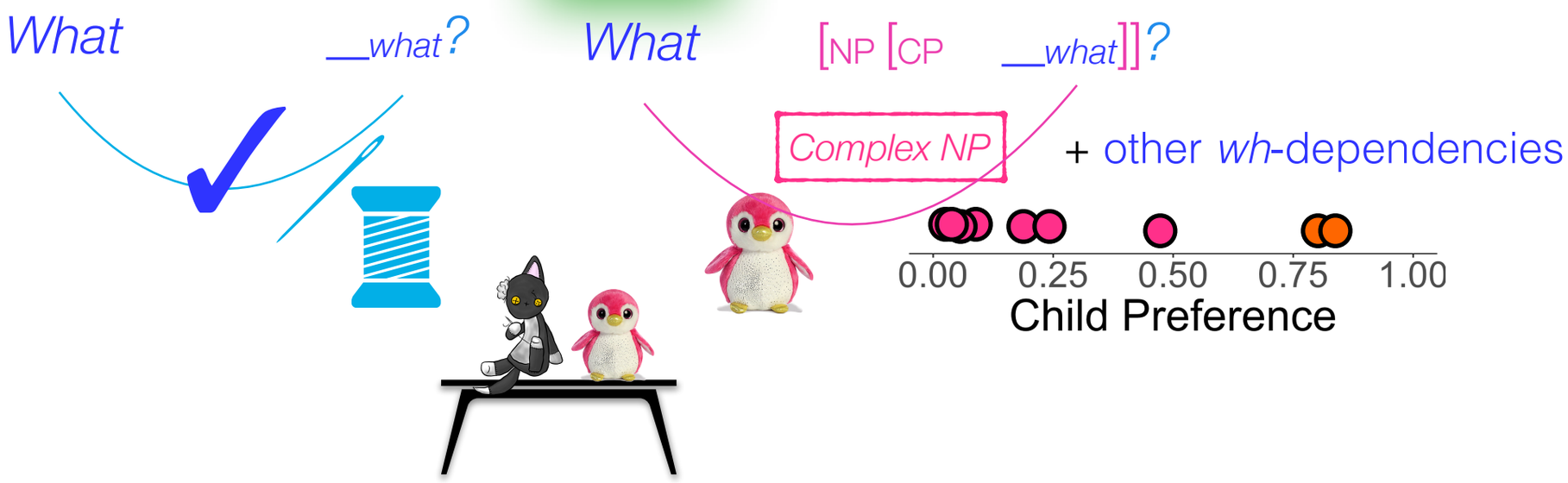
## How long do children have to learn?

# Syntactic islands

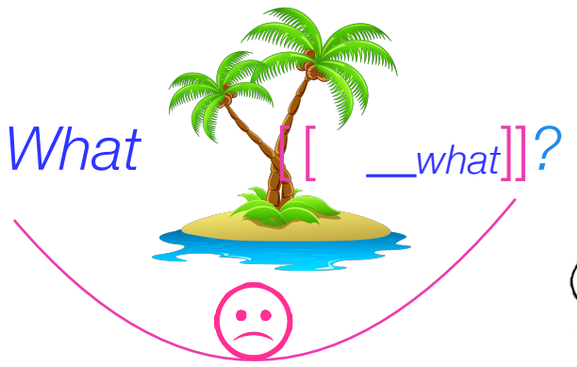



## How long do children have to learn?

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



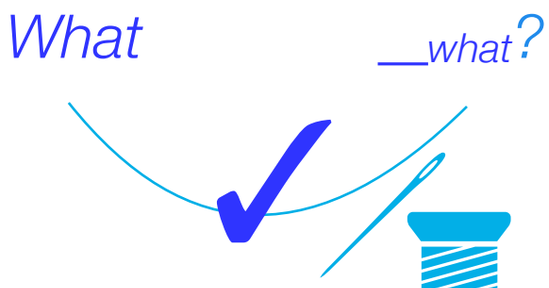
# Syntactic islands

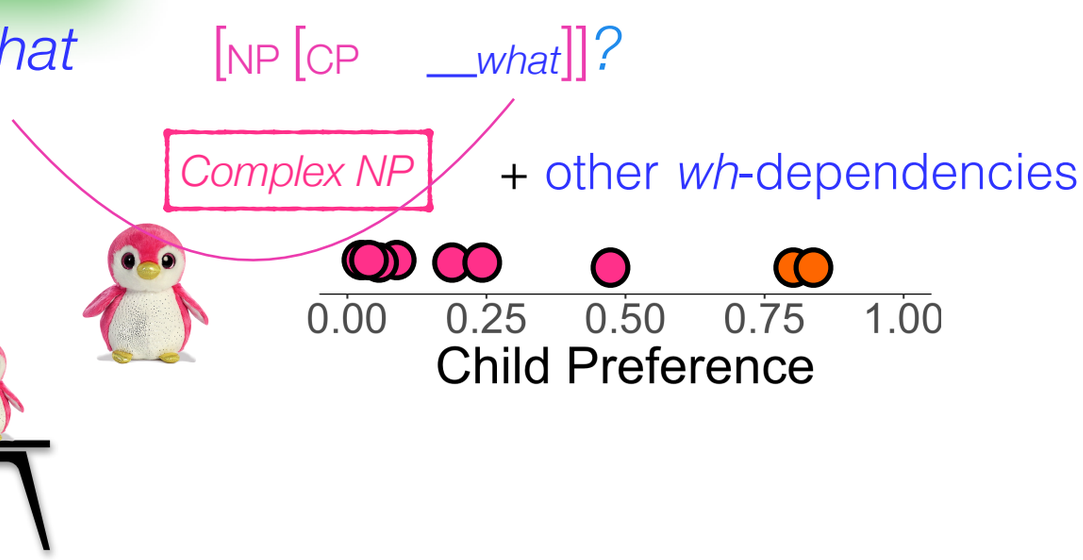



## How long do children have to learn?

So input through age four.

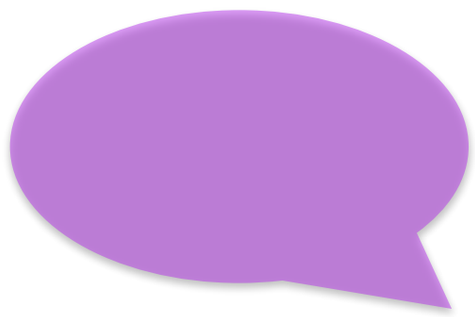
(<60 months)

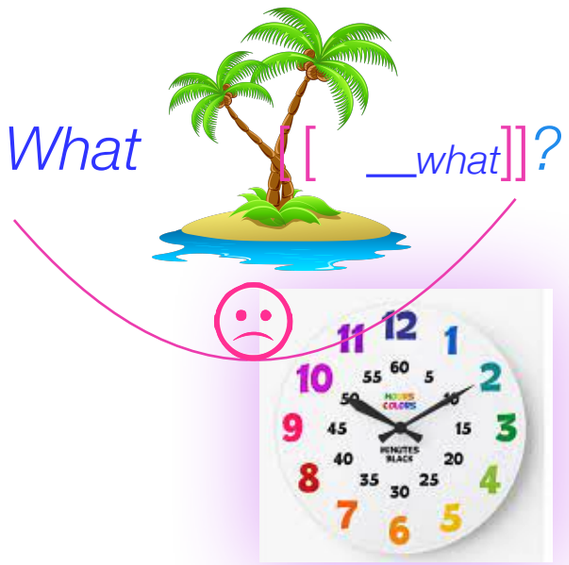


# Syntactic islands

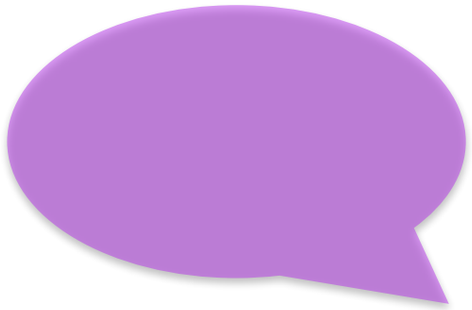

## What input do children get?






# Syntactic islands

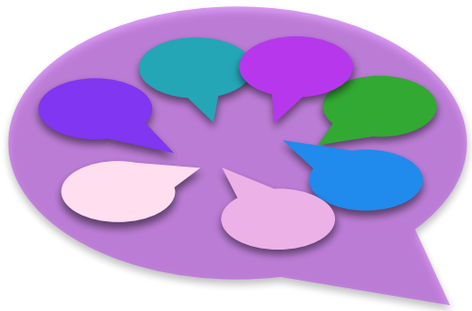
## What input do children get?

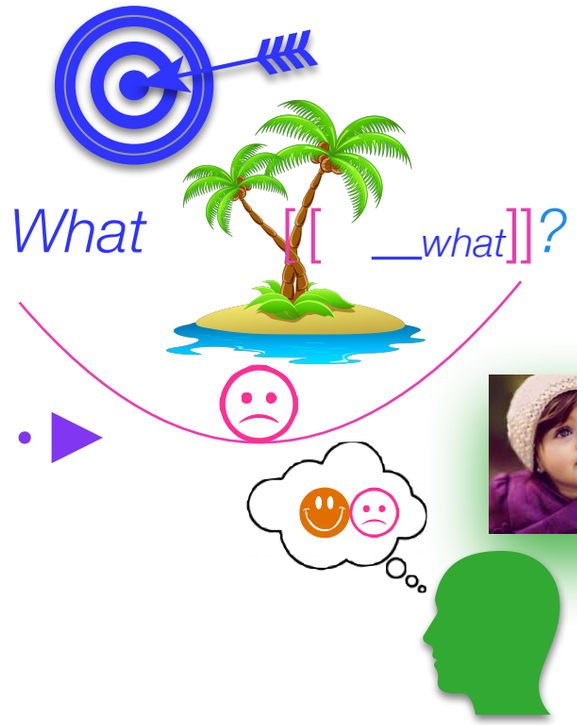


We can estimate this from samples of child-directed speech.

# Syntactic islands

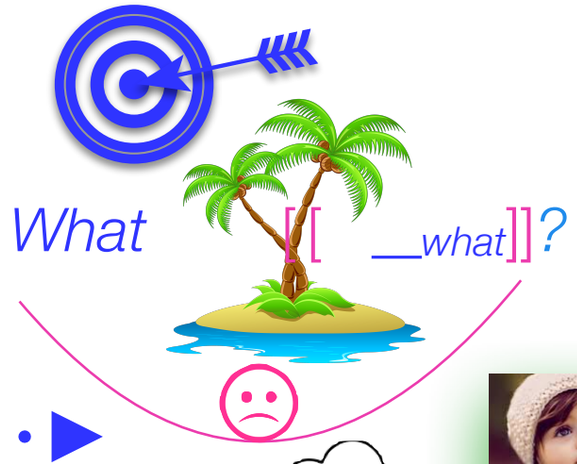
This is the acquisition problem

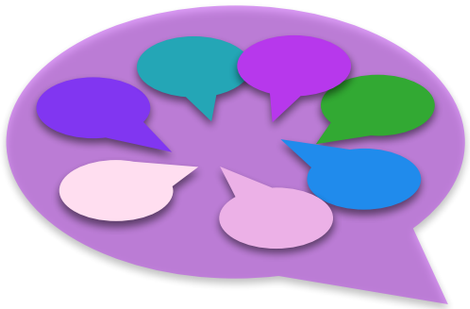


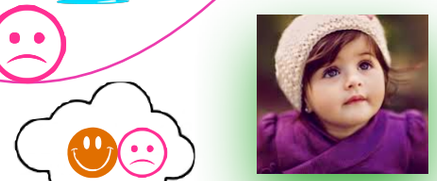


# Syntactic islands

…which is where a theory of acquisition comes in.

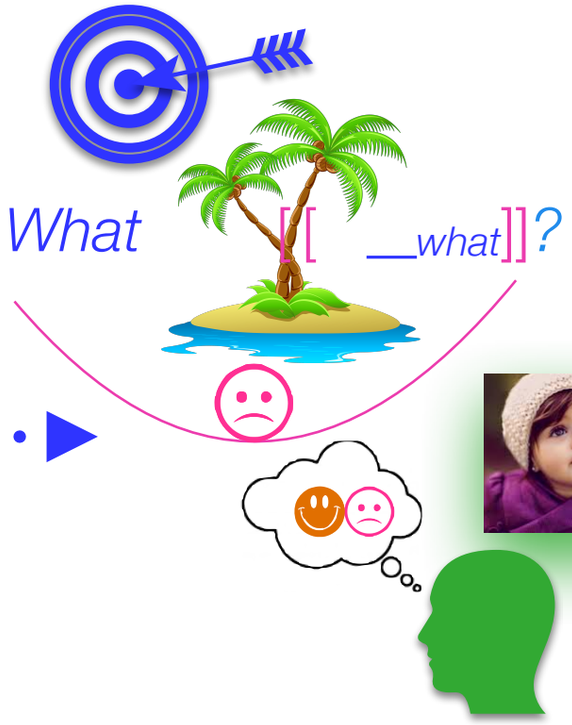


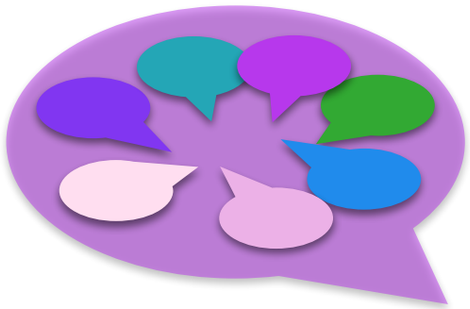





# Syntactic islands

…which is where a theory of acquisition comes in.

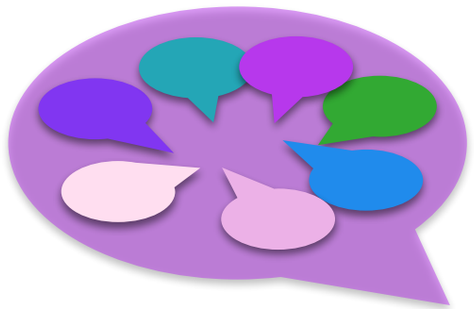


Learn the right building blocks

# Syntactic islands

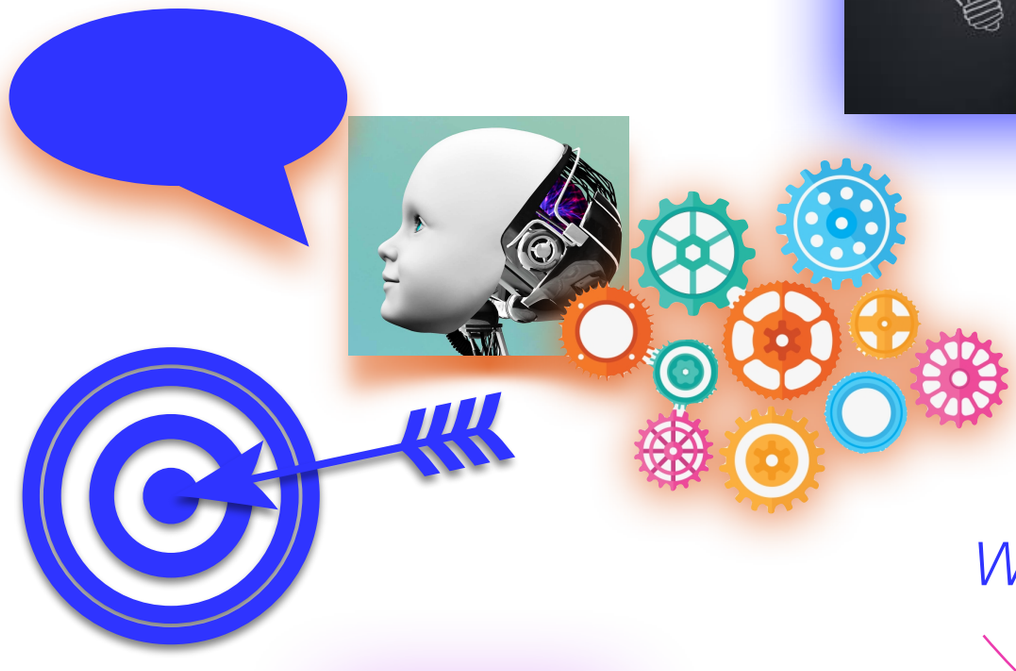
We can evaluate a theory by implementing it concretely in a computational cognitive model.

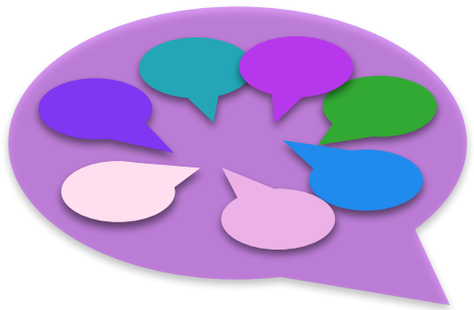


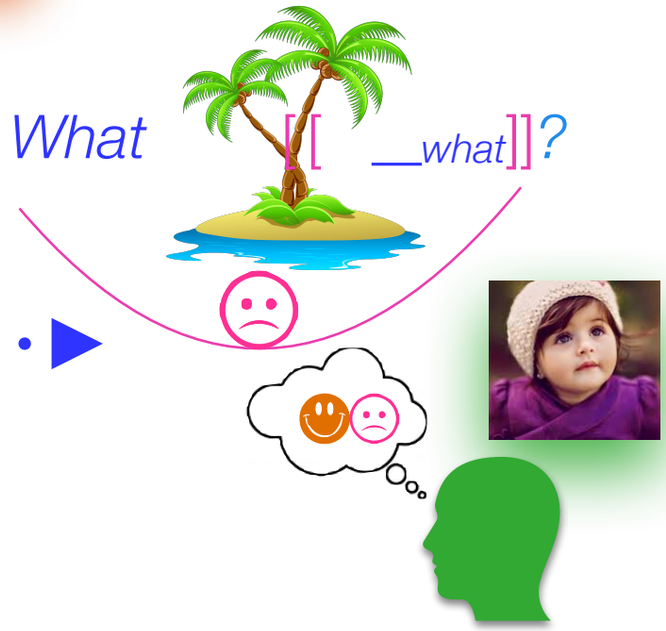

Learn the right building blocks

# Syntactic islands

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



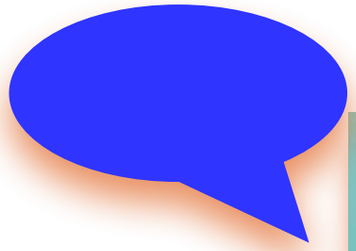


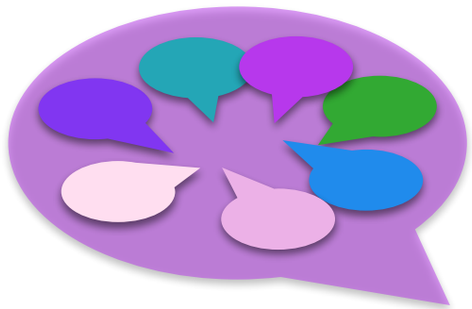



Learn the right building blocks

Syntactic islands

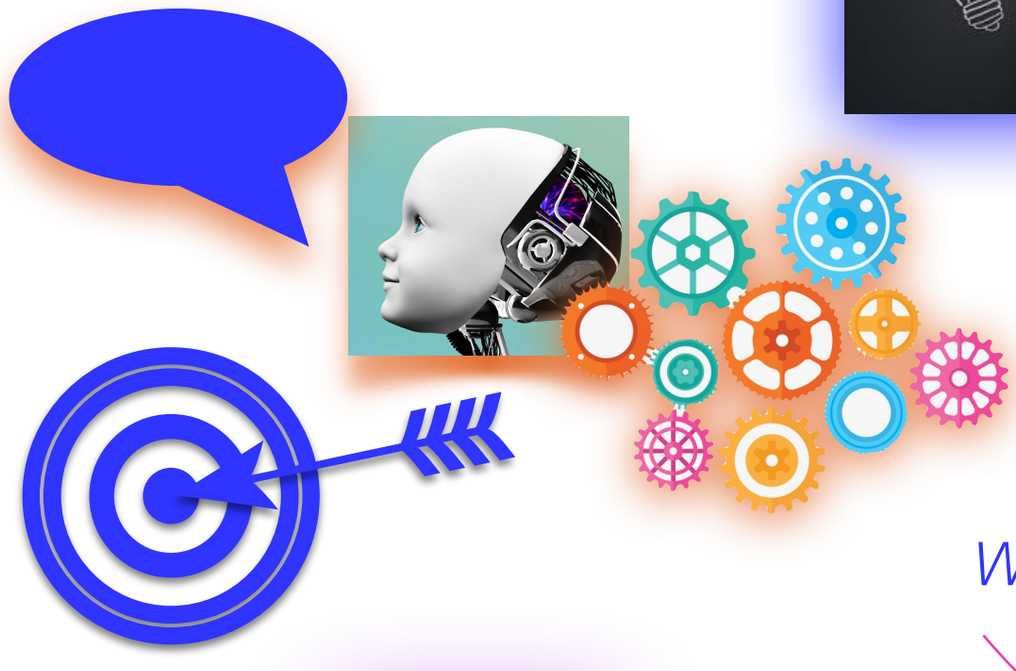
And then we can look inside it to see what makes it work (or not work).

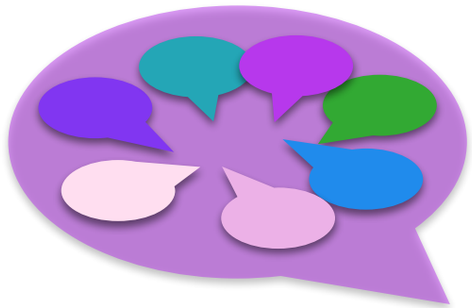





Learn the right building blocks

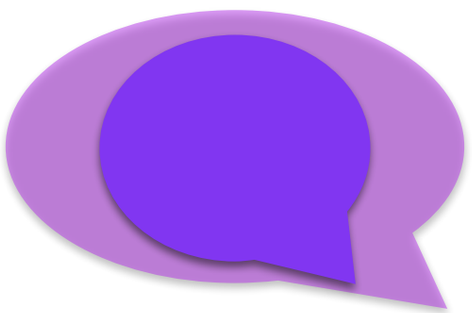
# Syntactic islands





So let's do this for our theory.

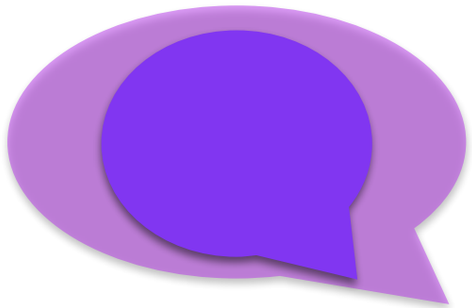

Learn the right building blocks



Intuition:

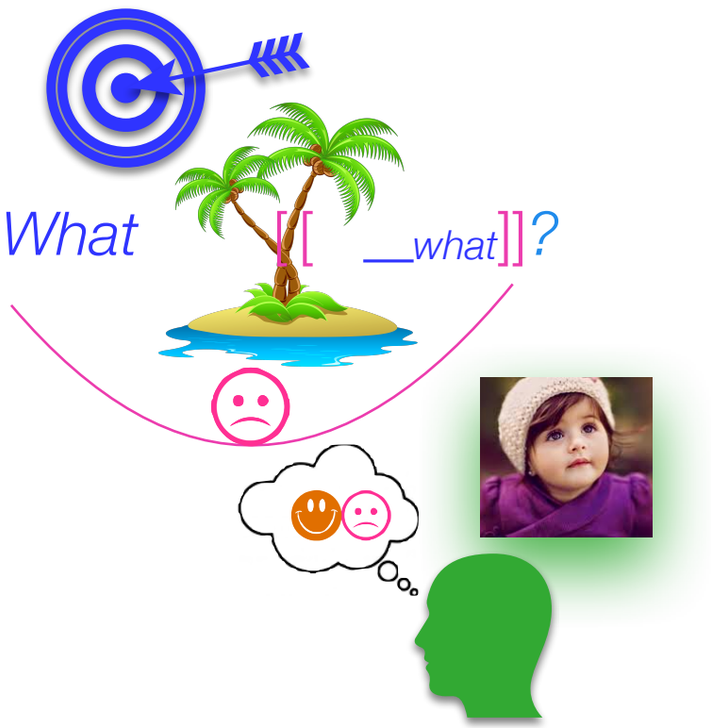
- Learn what you can from the *wh*-dependencies you observe in the input over time

*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*



Intuition:

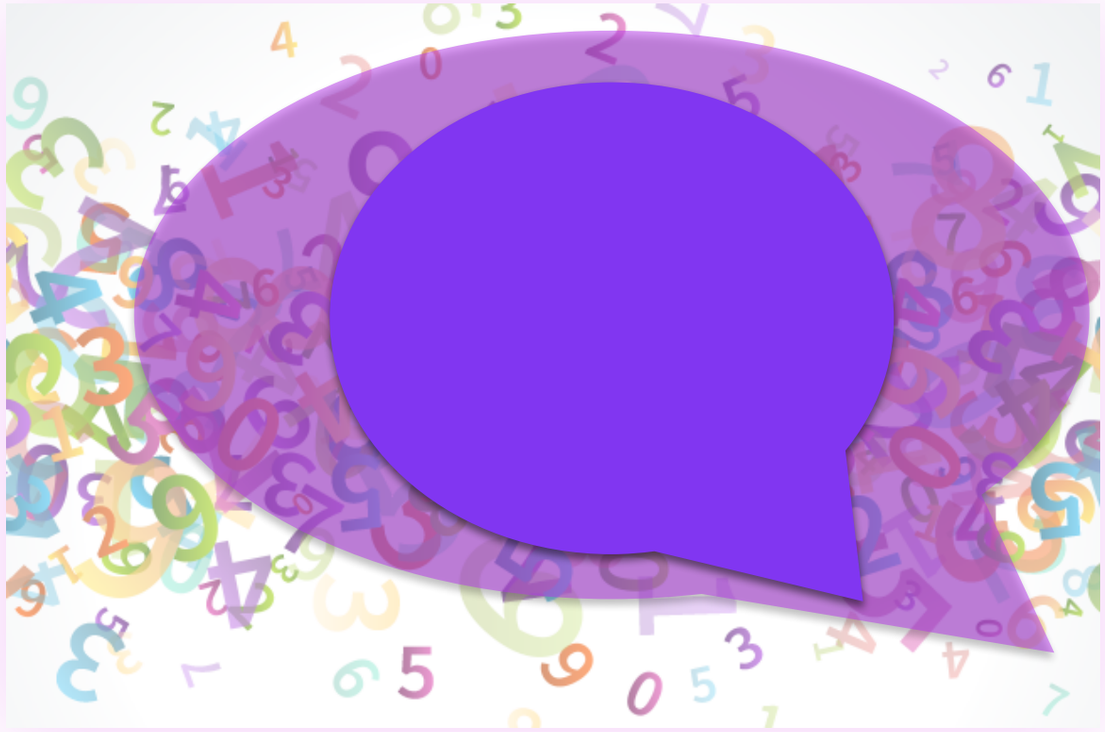
- Learn what you can from the *wh*-dependencies you observe in the input over time
- Apply it to generate behavior for *wh*-dependencies you haven't seen before, like those crossing syntactic islands (or other longer *wh*-dependencies).




*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*



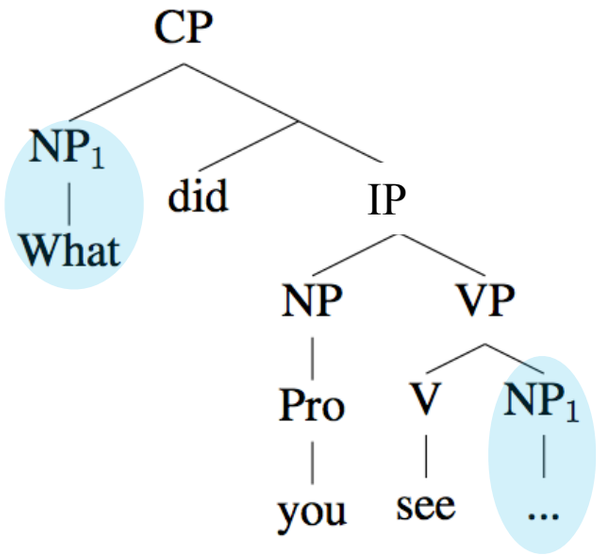
View *wh*-dependencies in terms of their building blocks and track those building blocks in the input.



*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*

Dependencies represented as a sequence of container nodes

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

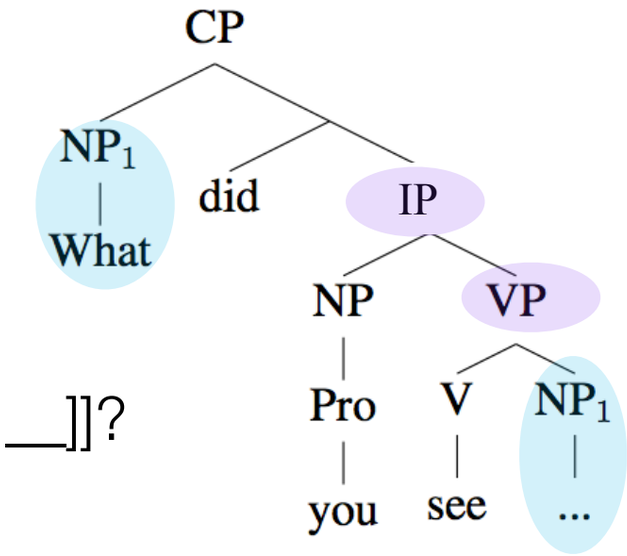



*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*

Dependencies represented as a sequence of container nodes

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

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



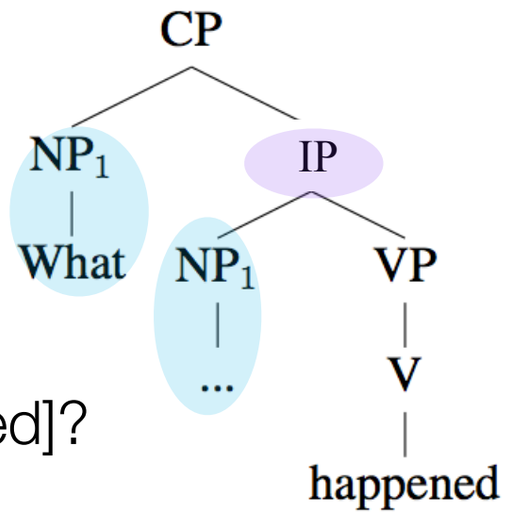

*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*

Dependencies represented as a sequence of container nodes

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

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

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




*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*

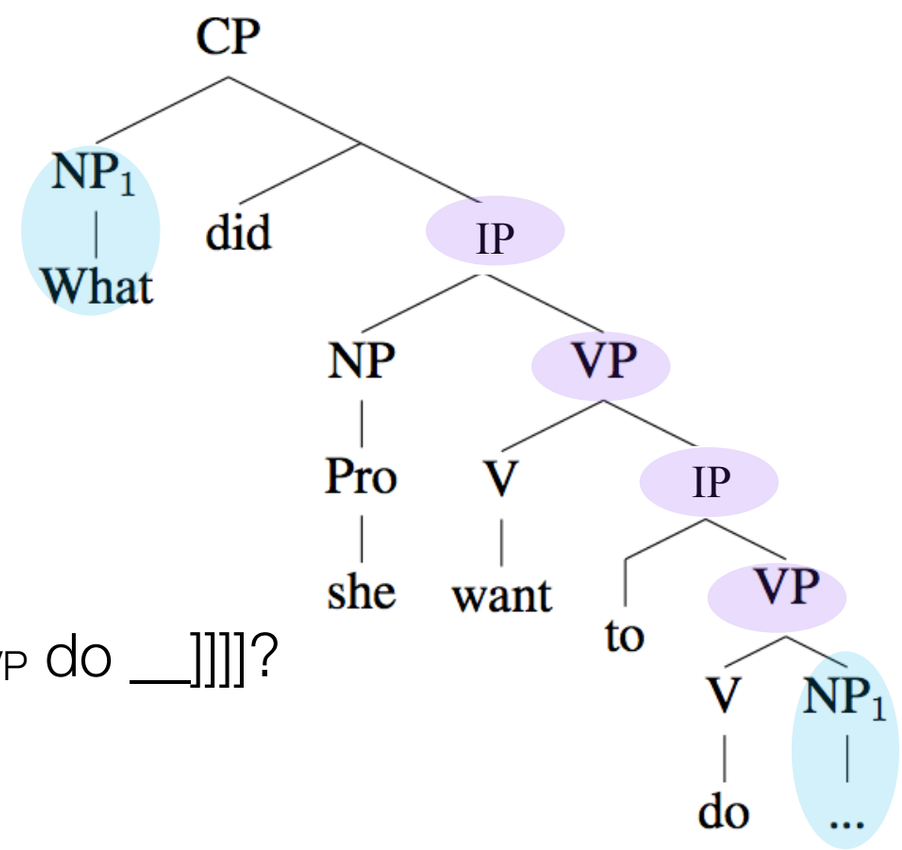
Dependencies represented as a sequence of container nodes

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

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

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, 2024, in prep.*

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

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

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

(Much) less acceptable dependencies have low probability segments

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

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

*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*

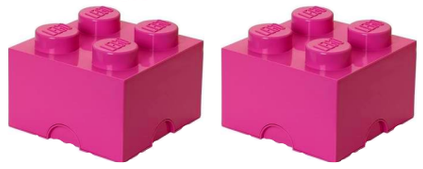
What did you see __?
= What did [IP you [VP see __]]?
= *start*-IP-VP-*end*

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

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

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

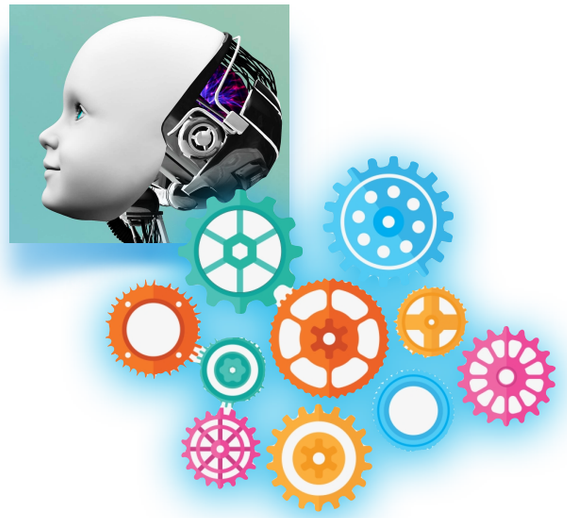
*start*-IP-VP-CP$_{that}$-IP-NP-PP-*end*



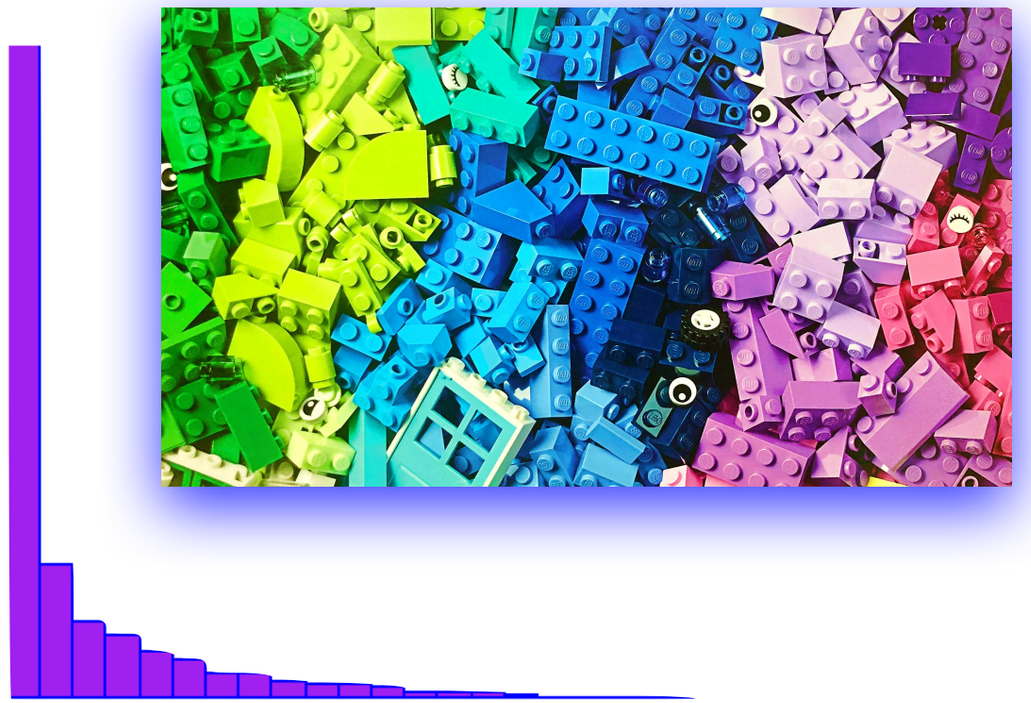
So if children break these dependencies into smaller building blocks, they can identify if a dependency has bad segments (made up of one or more low probability building blocks).

*Pearl & Sprouse 2013*
*Dickson, Pearl, & Futrell 2022, 2024, in prep.*

Learn the right building blocks

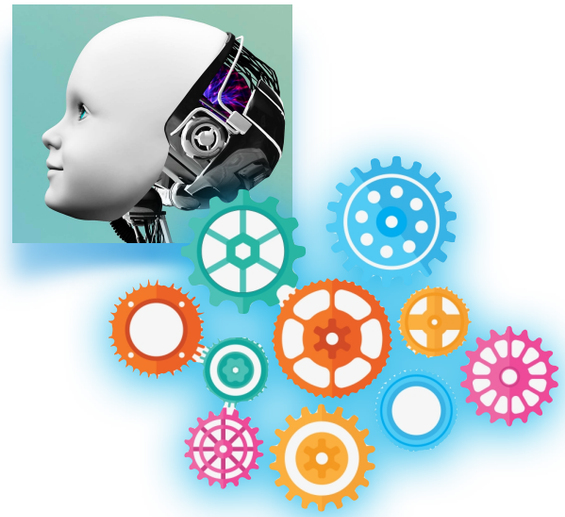


Theory: The child tries to learn
what the “best” building blocks are
at the same time she learns about their distributions in the input.



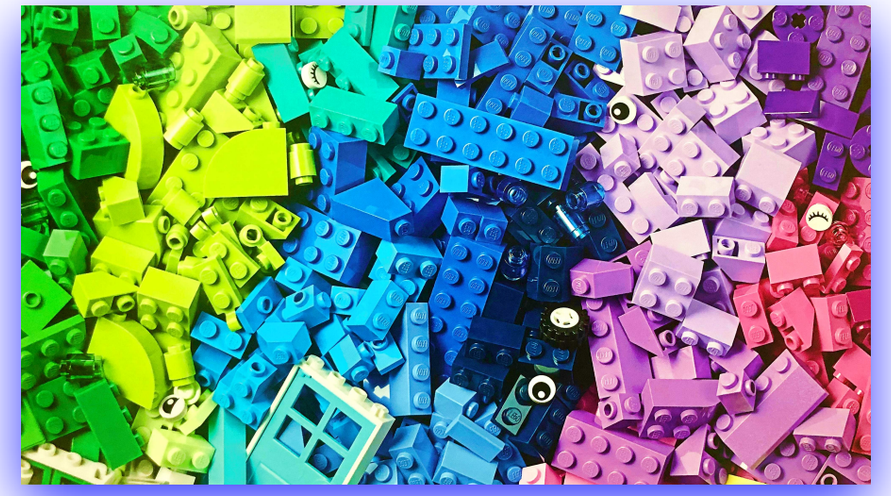
*Dickson et al. 2022, 2024, in prep.*

Learn the right building blocks



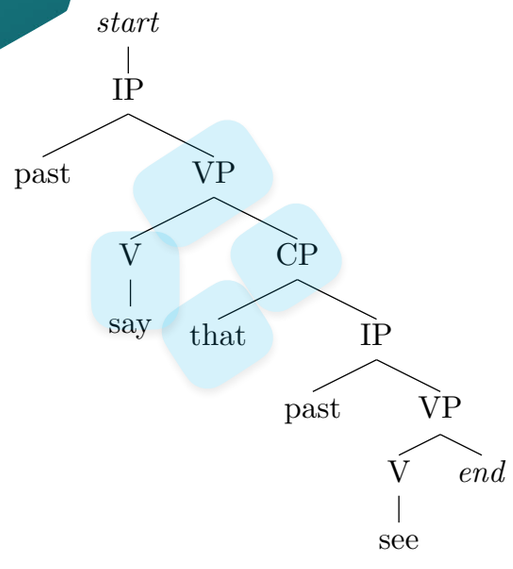
the best building blocks

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



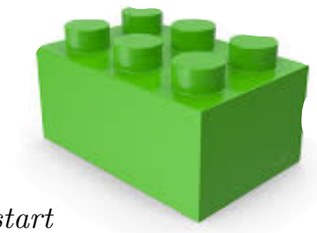
VP$_{V}$

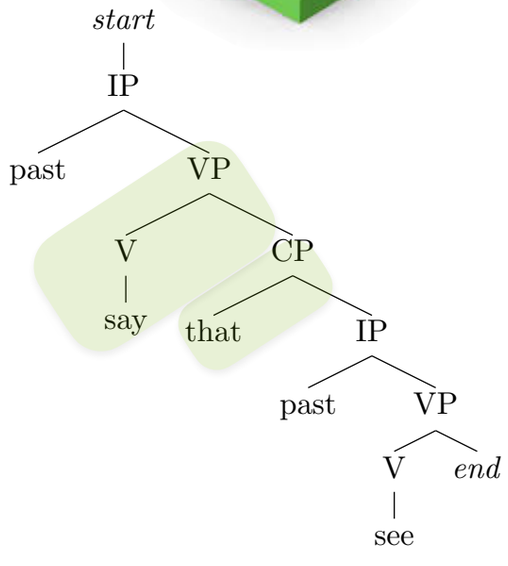
CP$_{C}$




CP$_{C=that}$

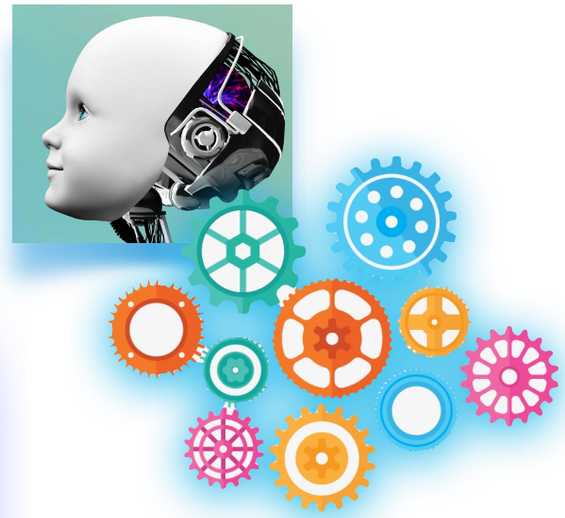
VP$_{V=say}$






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

# Learn the right building blocks

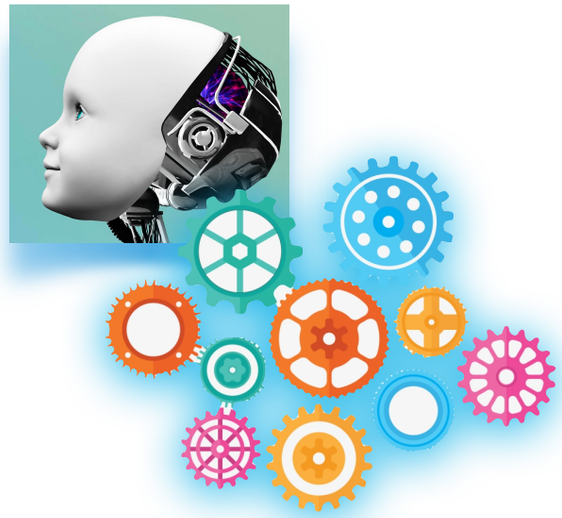


How can the child learn what the best building blocks are?

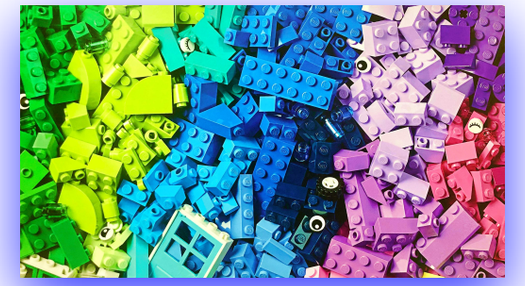
(1) Look for the best-sized units
(2) Sometimes include the lexical item

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

# Learn the right building blocks



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

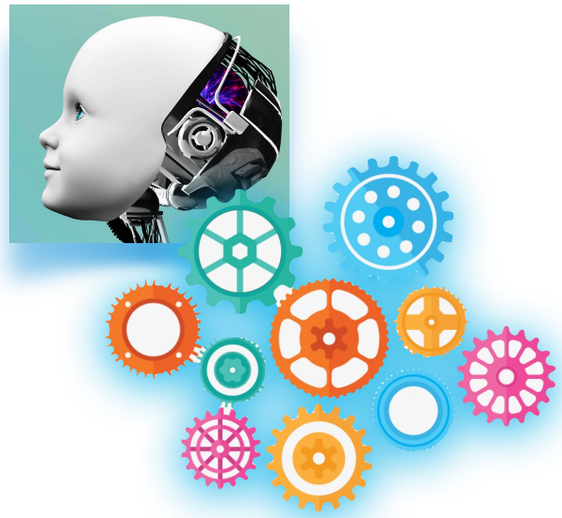


(1) Look for the best-sized units
(2) Sometimes include the lexical item

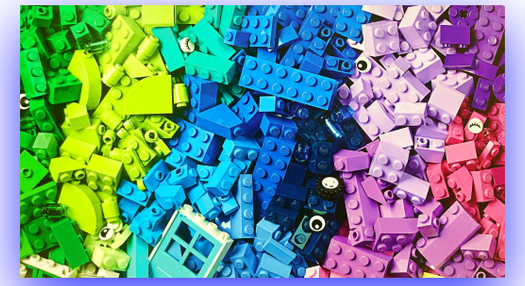
Theory: Look for an “efficient” set of building blocks.

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

# Learn the right building blocks



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

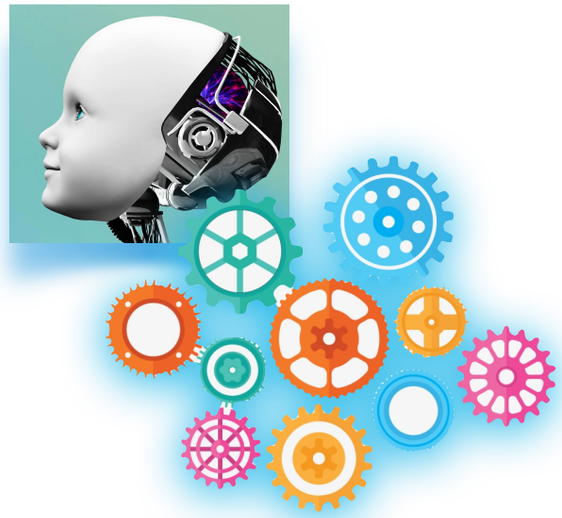


(1) Look for the best-sized units
(2) Sometimes include the lexical item

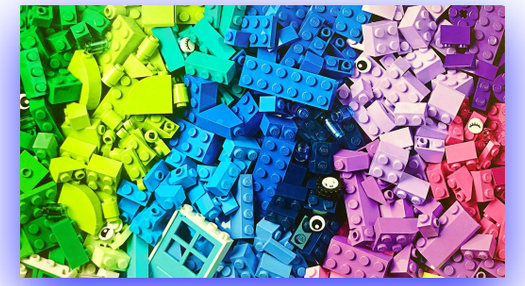
Efficient building blocks allow the representation of current and future *wh*-dependencies to be more probable.

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

# Learn the right building blocks



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



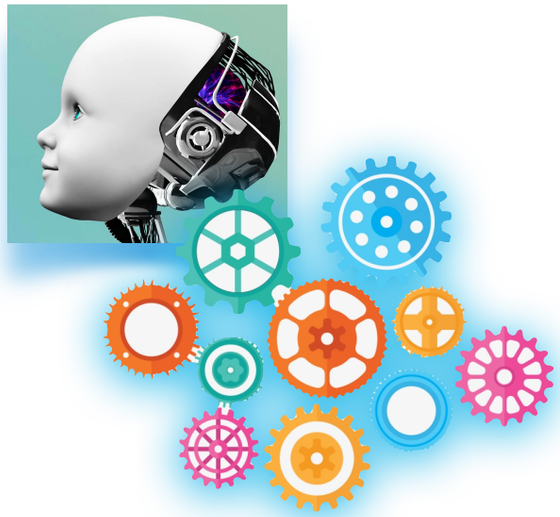
(1) Look for the best-sized units
(2) Sometimes include the lexical item

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

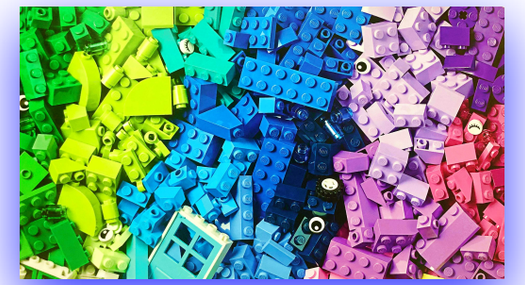
Why? One idea: Higher probability *wh*-dependencies are faster to process (comprehending or producing).

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

Learn the right building blocks



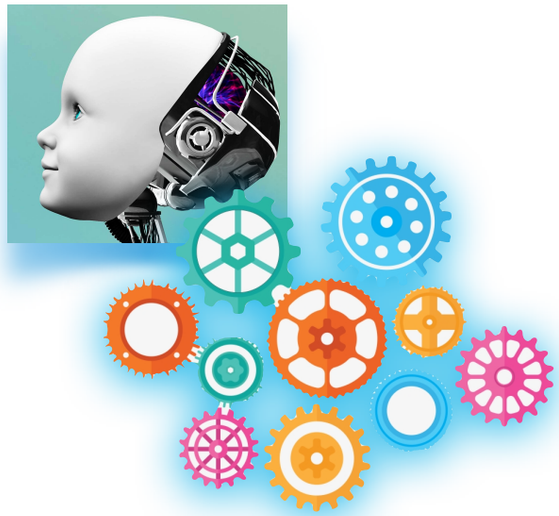
learning efficient building blocks

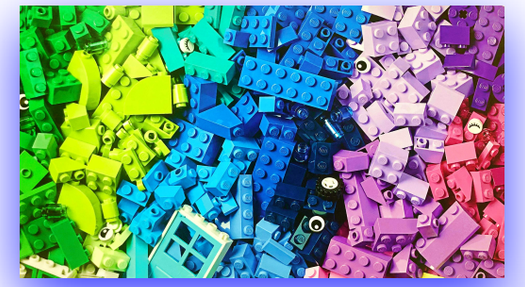


How? Look for building blocks that are a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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

Learn the right building blocks

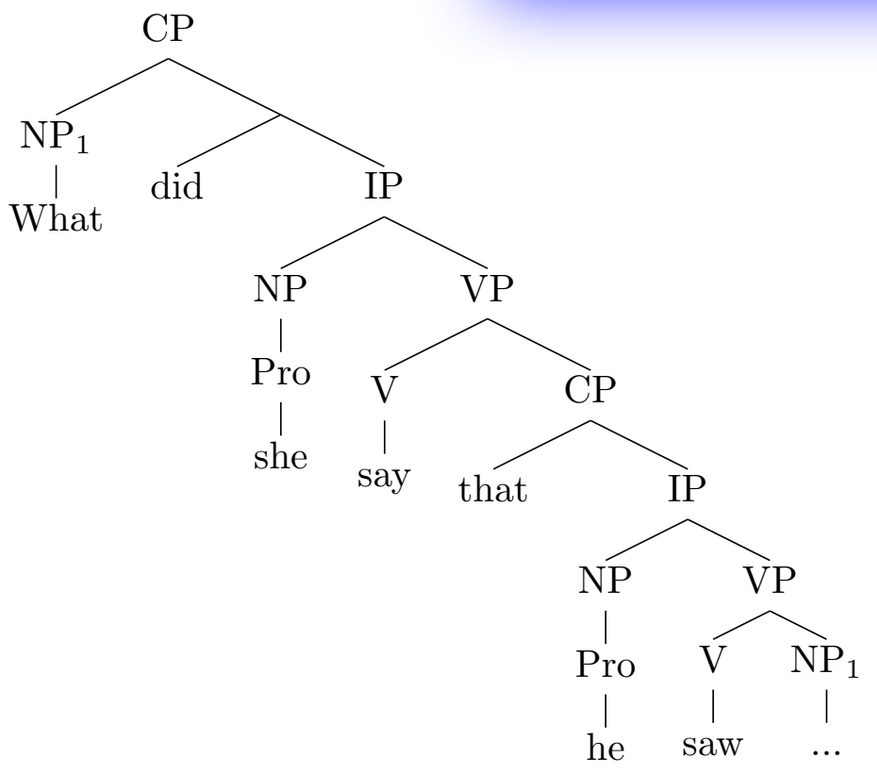




learning efficient building blocks

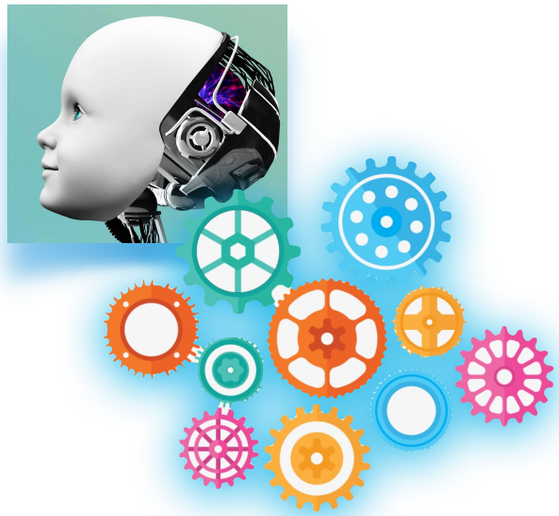
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

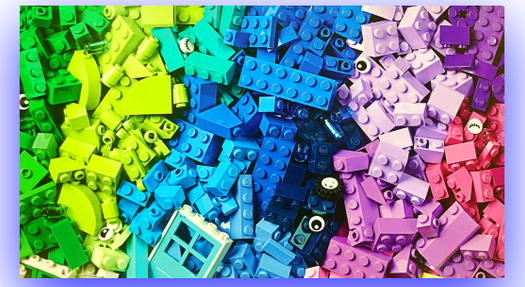
What did she say that he saw __ ?




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

Learn the right building blocks

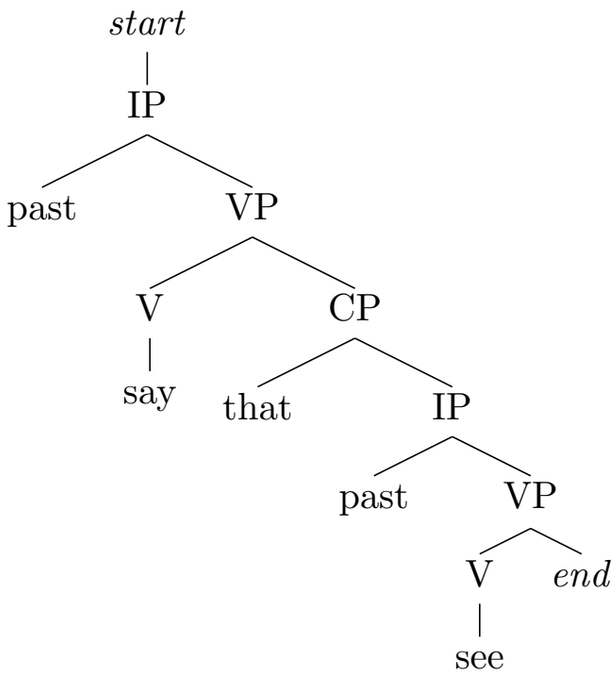




learning efficient building blocks

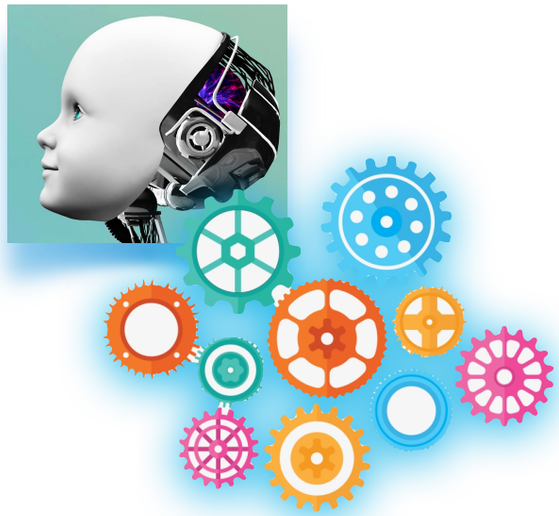
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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




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

# Learn the right building blocks

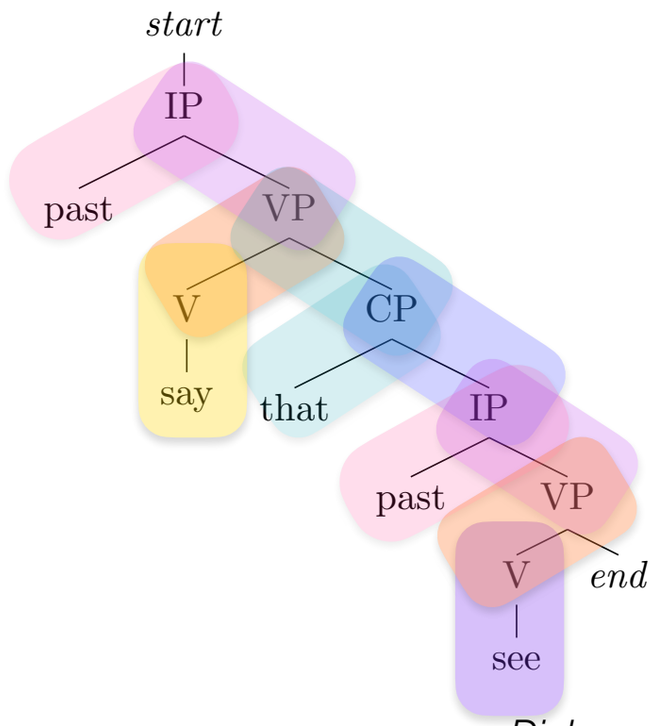


learning efficient building blocks

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

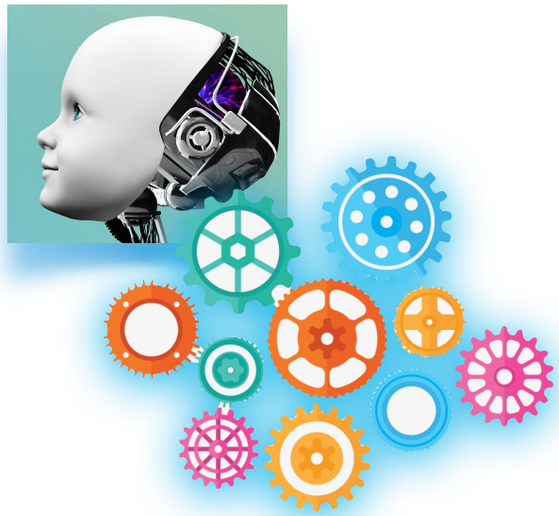
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

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




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

# Learn the right building blocks

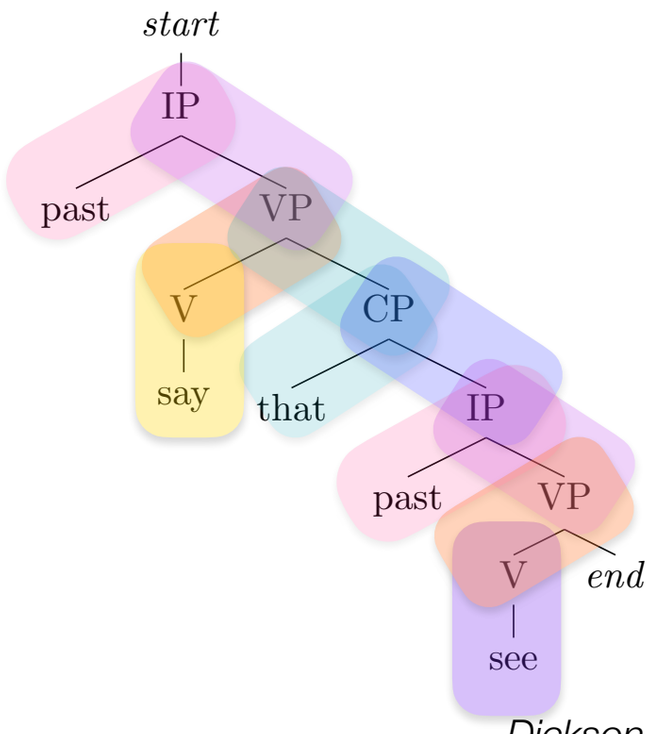


learning efficient building blocks

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

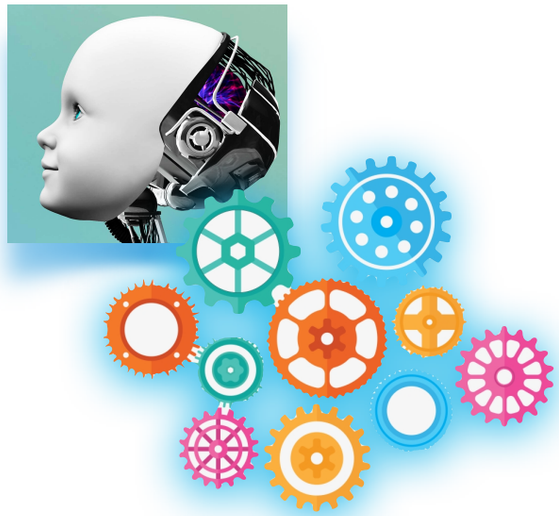
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

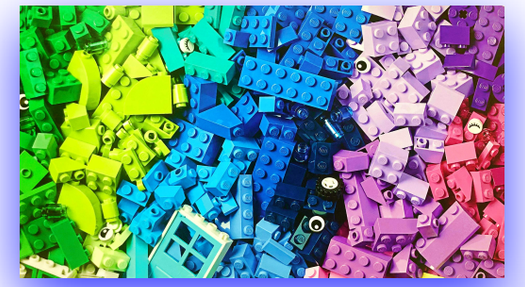
It may be slower to put together many small pieces.




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

# Learn the right building blocks





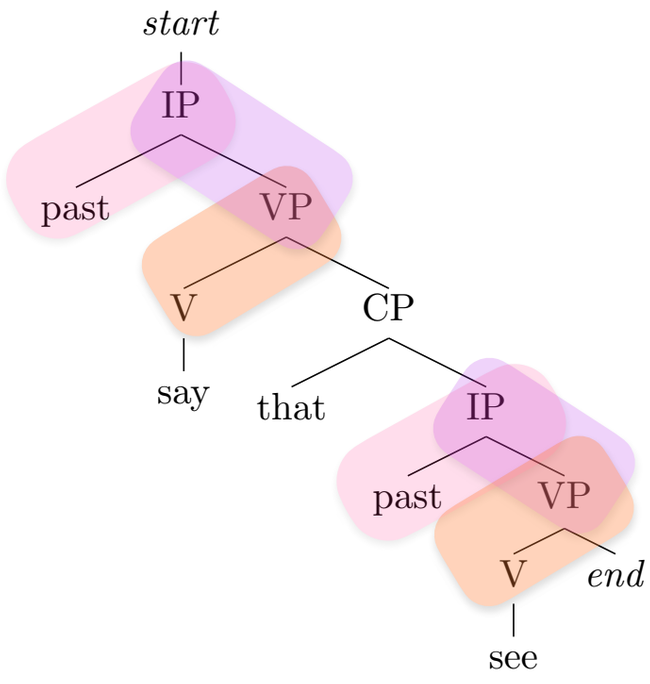
learning efficient building blocks

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

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

many smaller

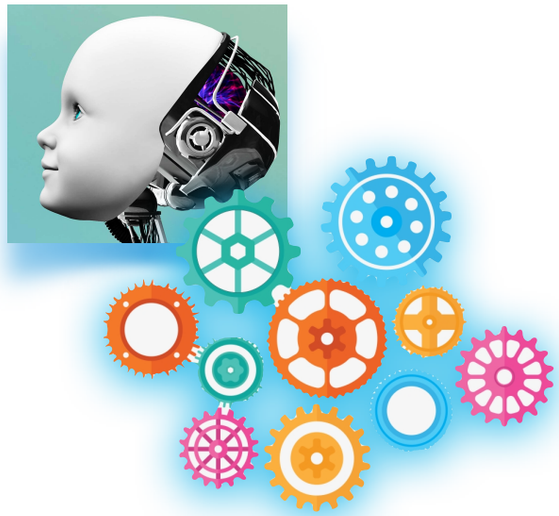
slower because many




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

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

# Learn the right building blocks



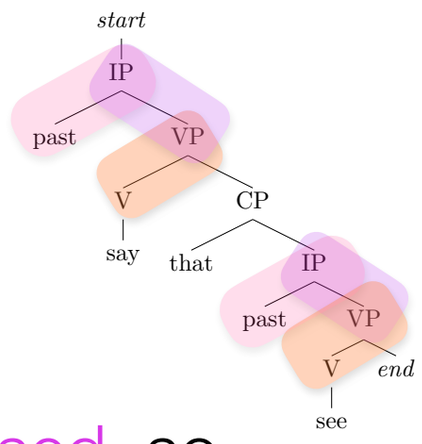
learning efficient building blocks

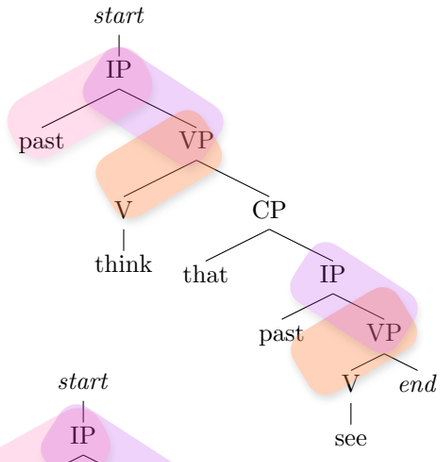
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

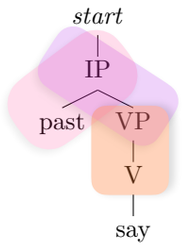
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

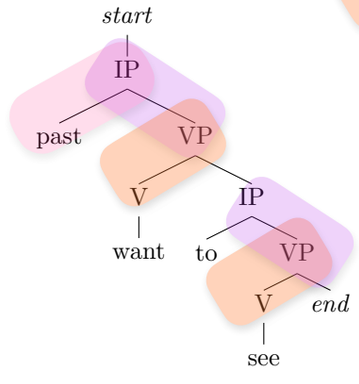
many smaller

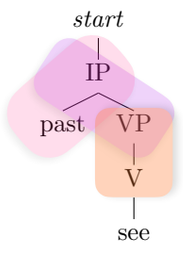
slower because many



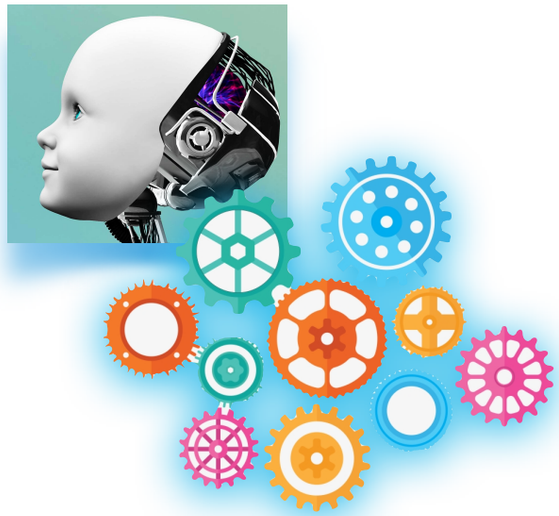



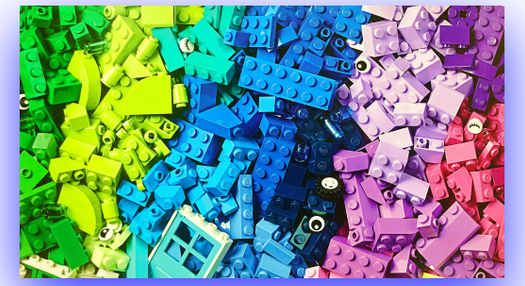

But these pieces may get reused, so that makes them faster to access.

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

# Learn the right building blocks



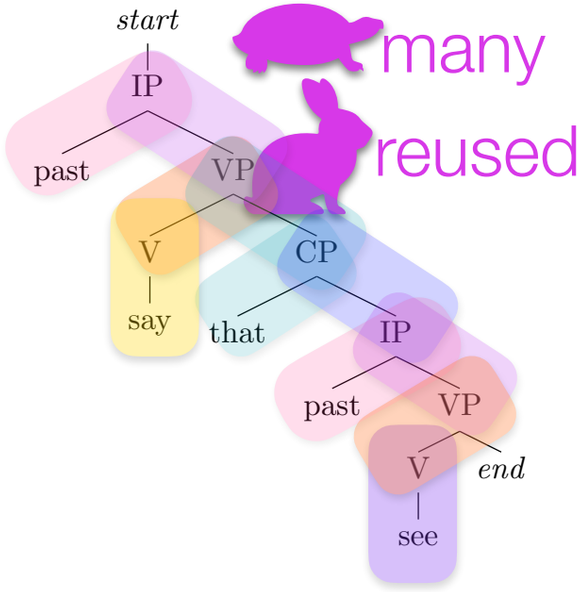
learning efficient building blocks



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

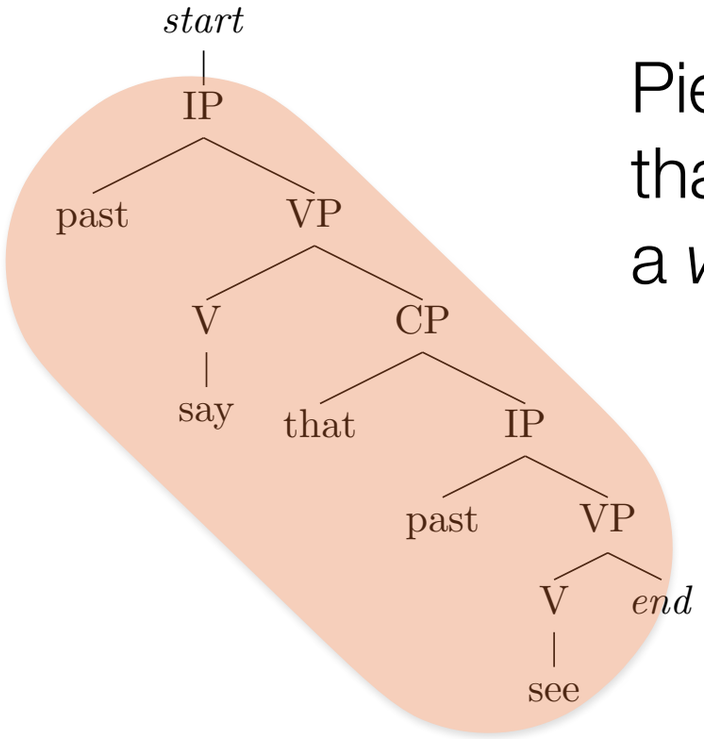
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller




many

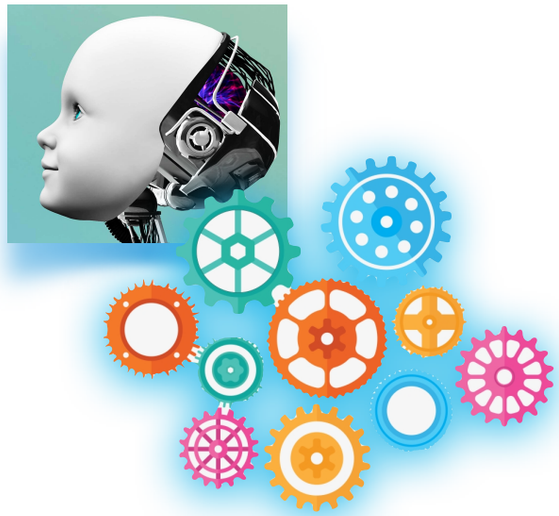
reused



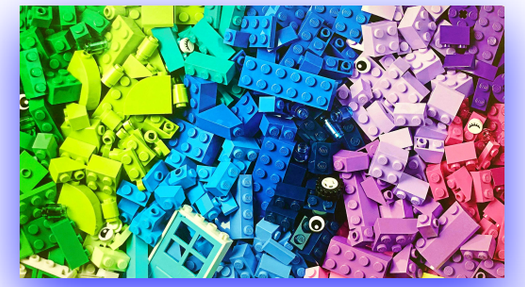

Pieces can be big, so that only one makes up a *wh*-dependency

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

# Learn the right building blocks



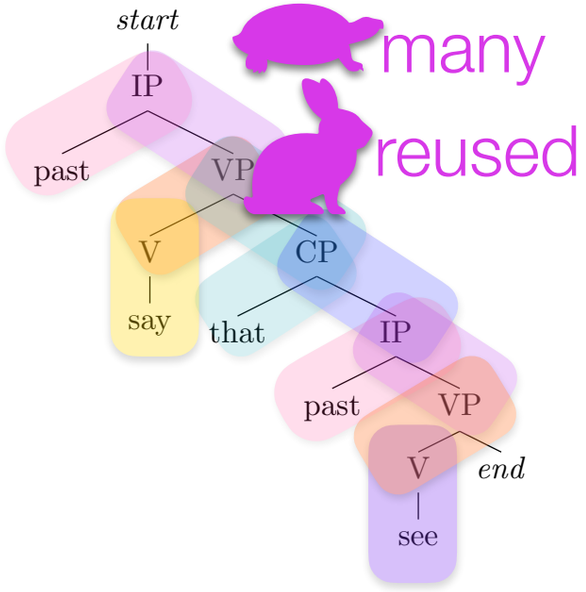
learning efficient building blocks



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

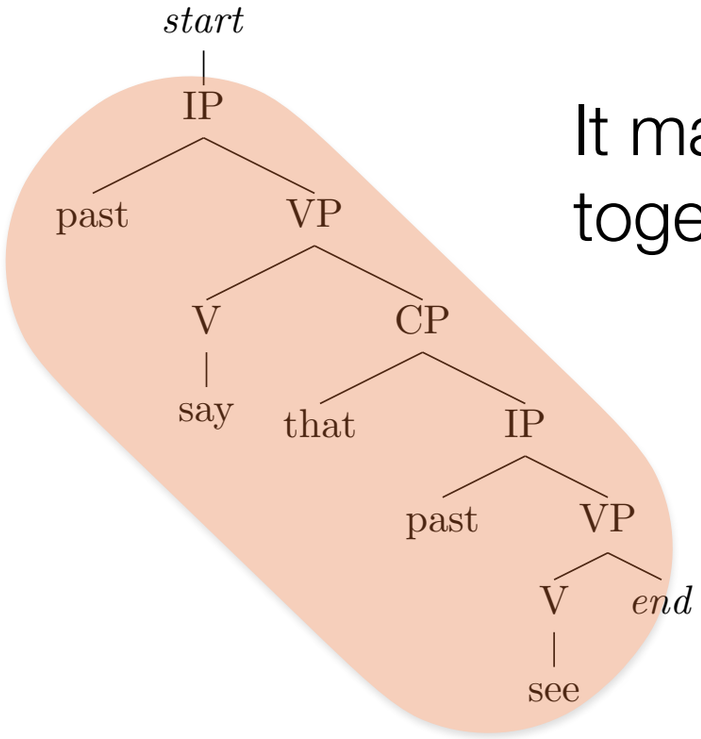
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller




many

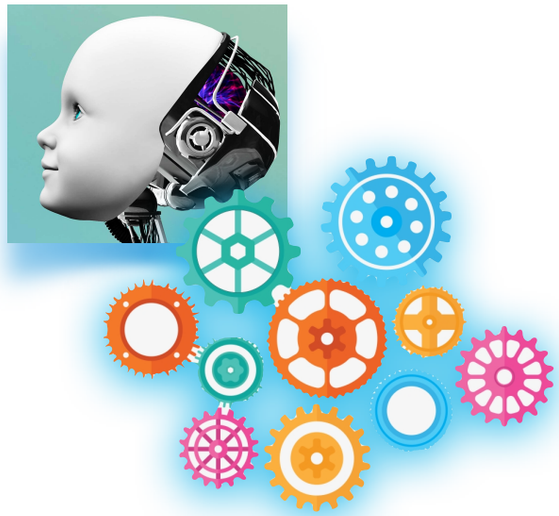
reused



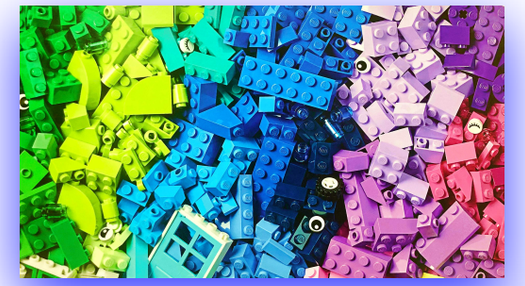

It may be faster to put together one big piece.

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

# Learn the right building blocks



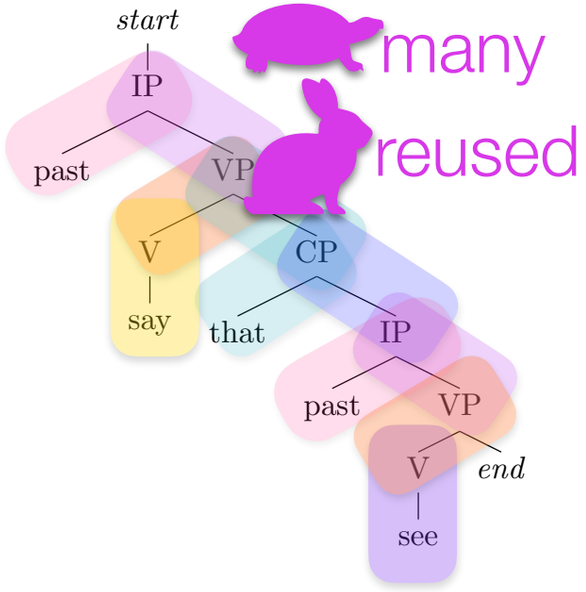
learning efficient building blocks



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

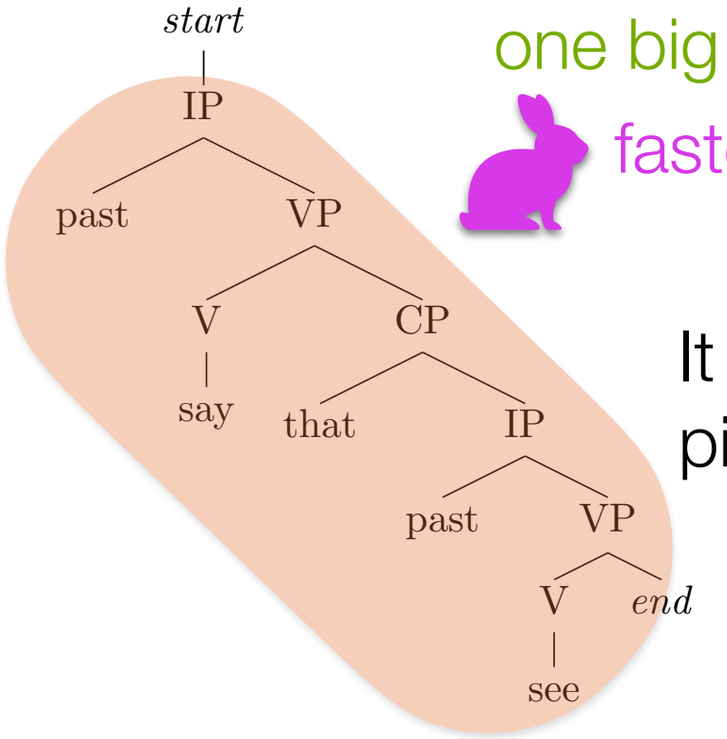
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller




many reused

one big

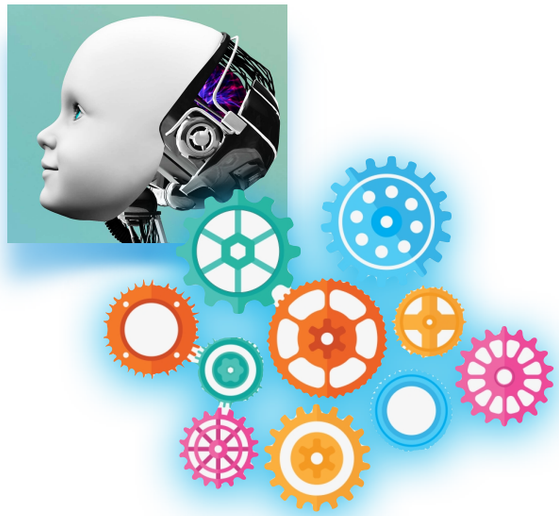



faster because one

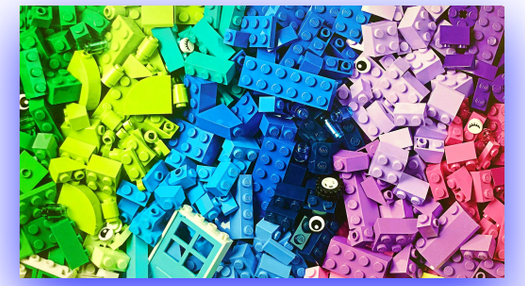
It may be slower if the piece is used rarely.

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

# Learn the right building blocks



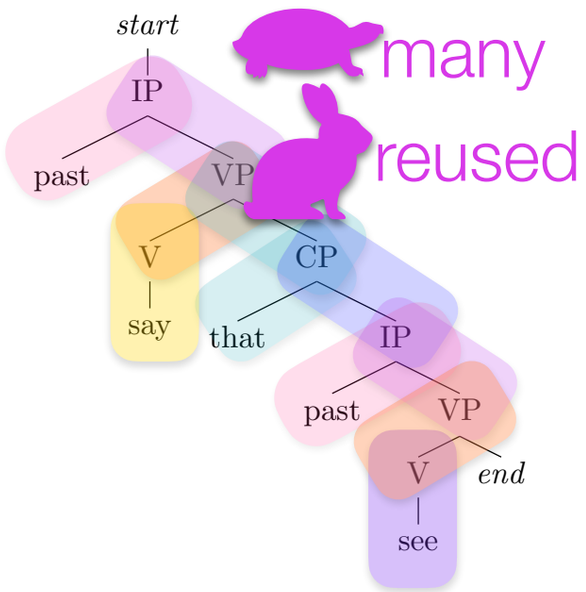
learning efficient building blocks



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

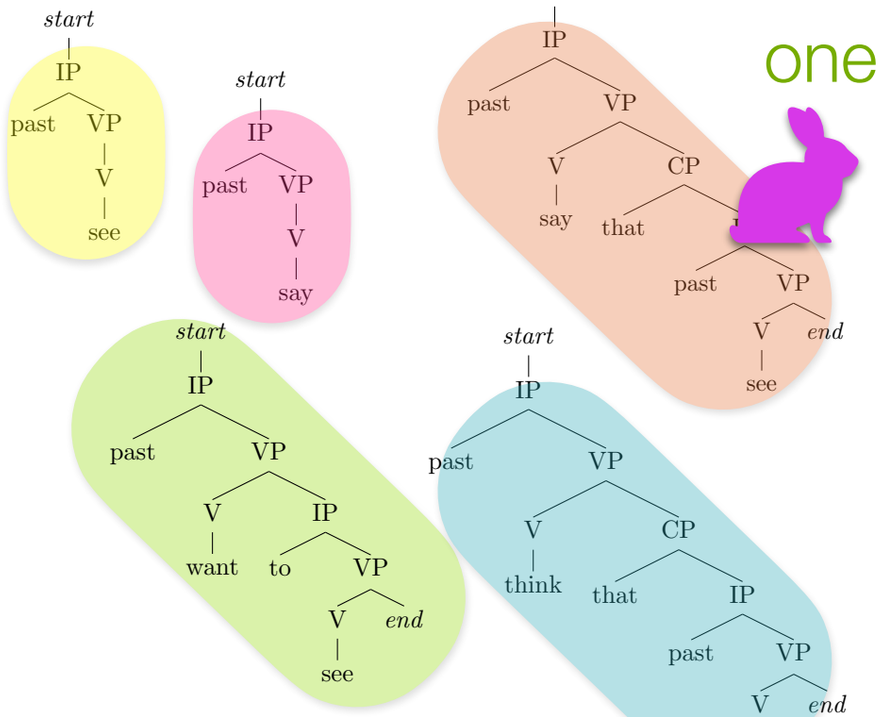
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

many smaller




many

reused



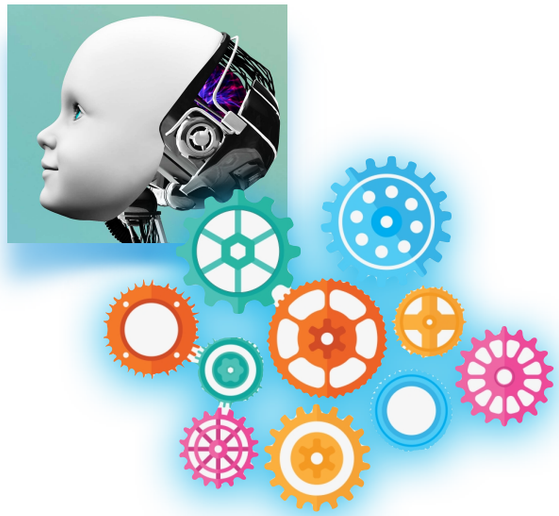

one big

faster because one

It may be slower if the piece is used rarely.

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

# Learn the right building blocks

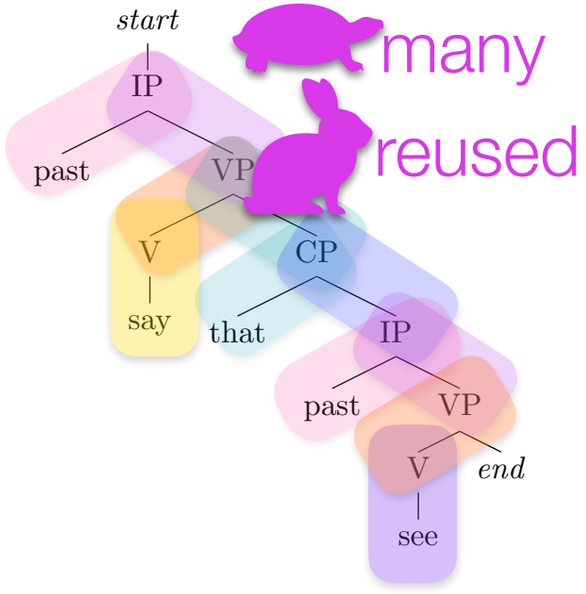


learning efficient building blocks

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

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

many smaller

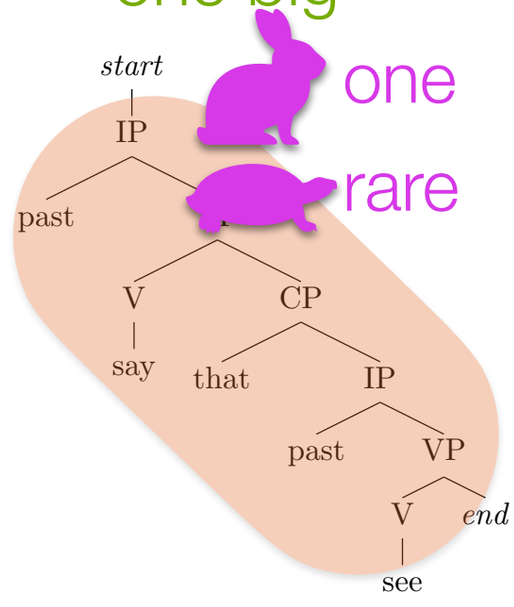



many

reused

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

one big

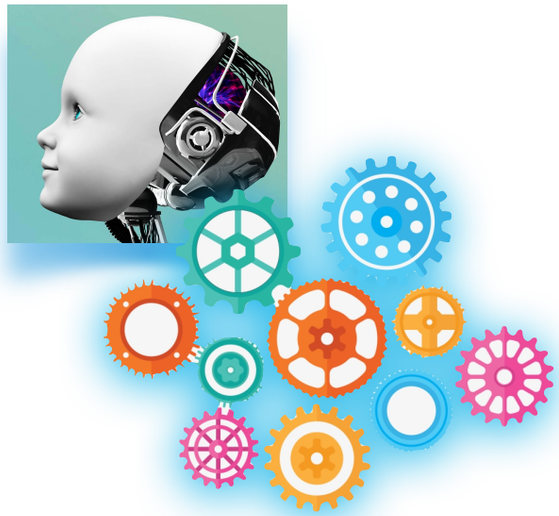


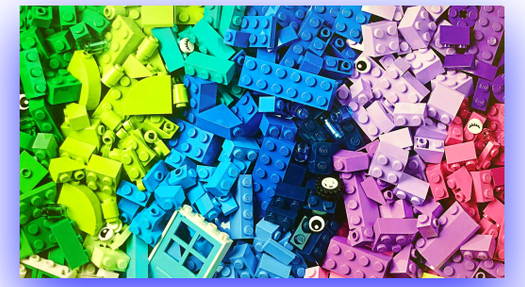

one

rare

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

# Learn the right building blocks



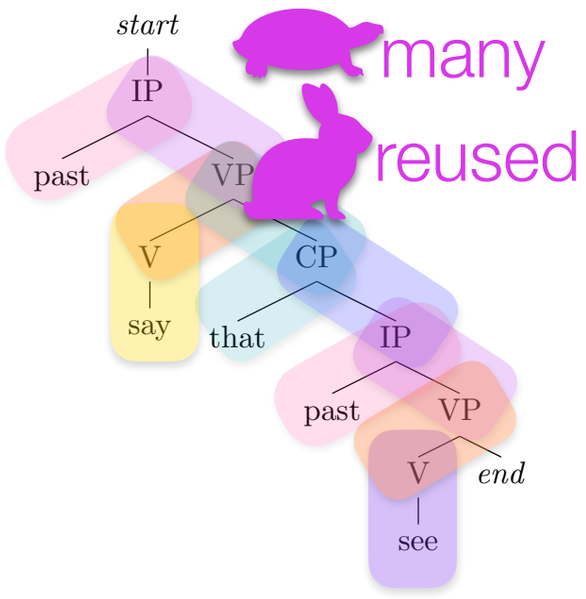


learning efficient building blocks

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

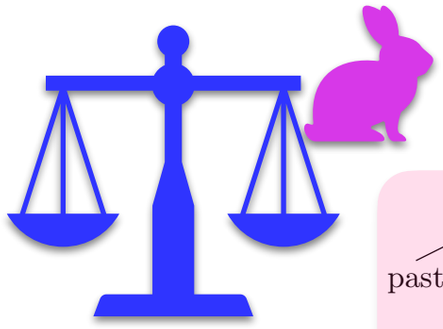
*start*-IP$_{past}$-VP$_{say}$-CP$_{that}$-IP$_{past}$-VP$_{see}$-*end*

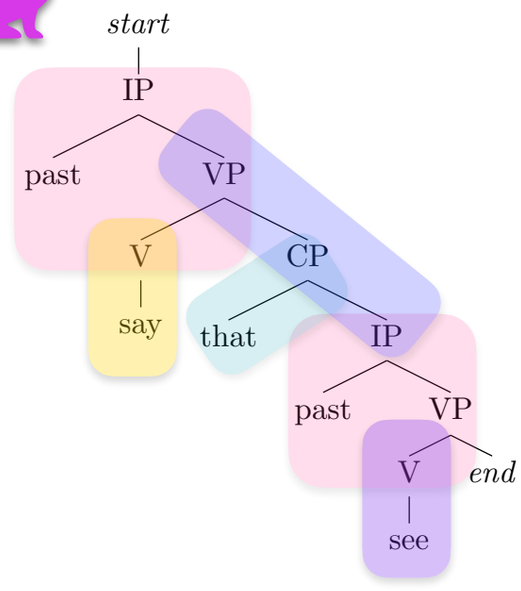
many smaller



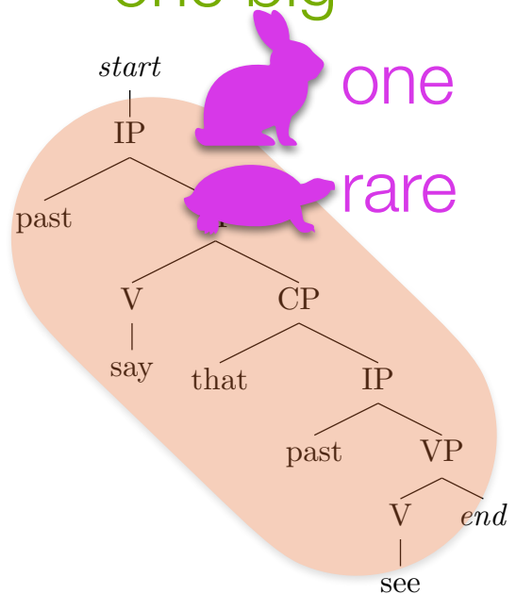

many

reused






one big

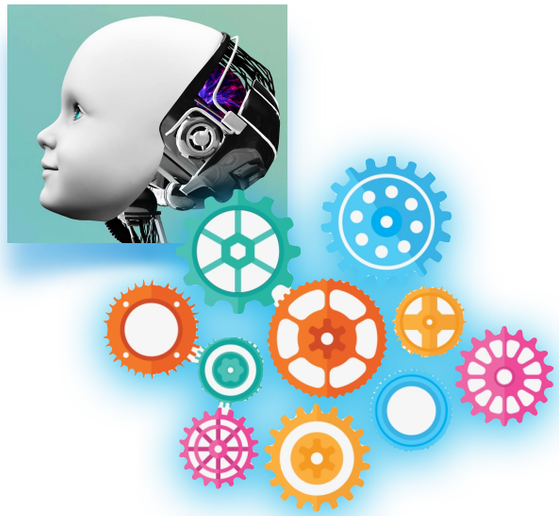



one

rare

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

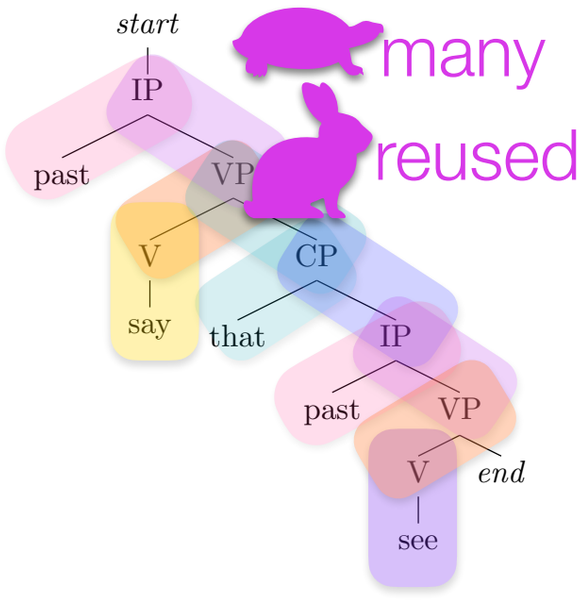
Learn the right building blocks

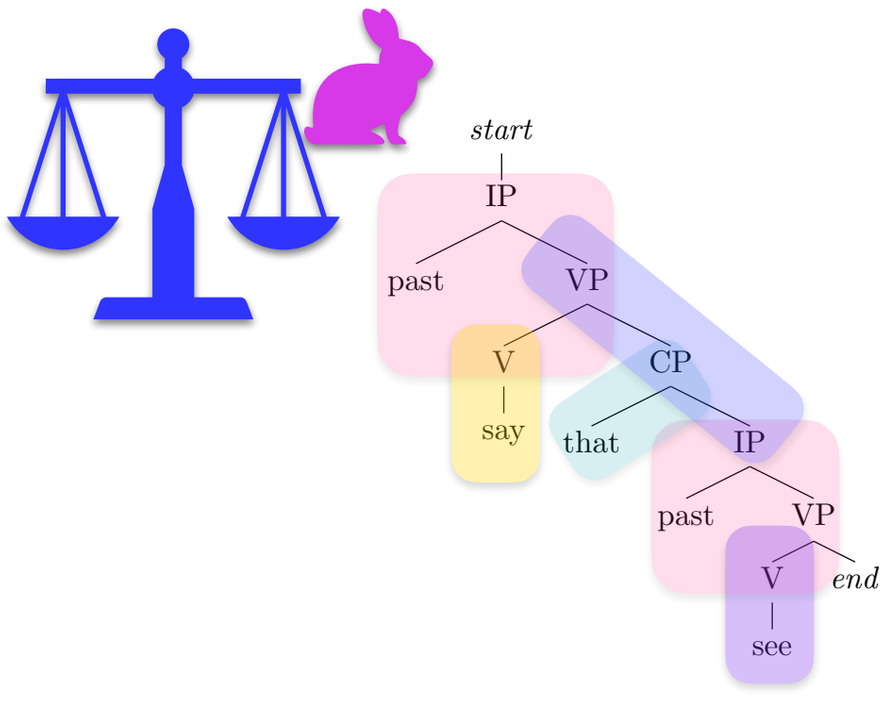


learning efficient building blocks

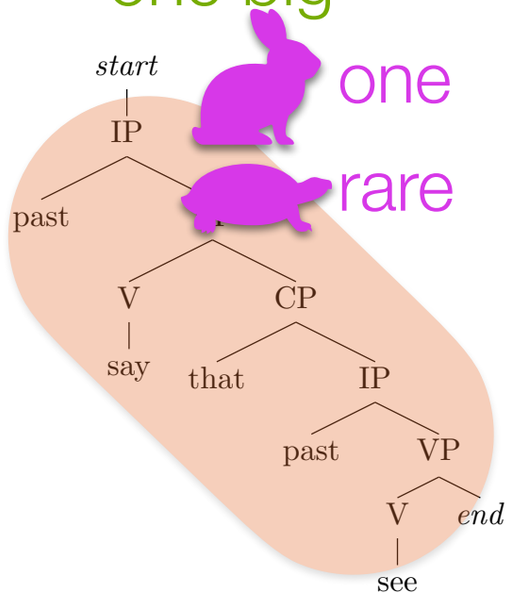
How can children find the best balance?

many smaller many reused



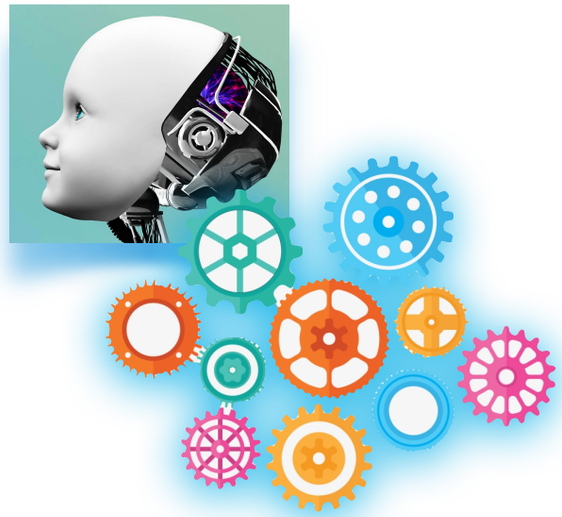




one big one rare




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

# Learn the right building blocks

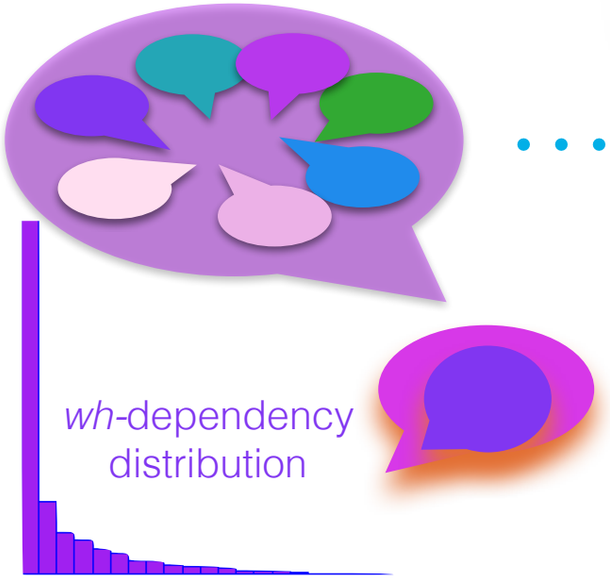


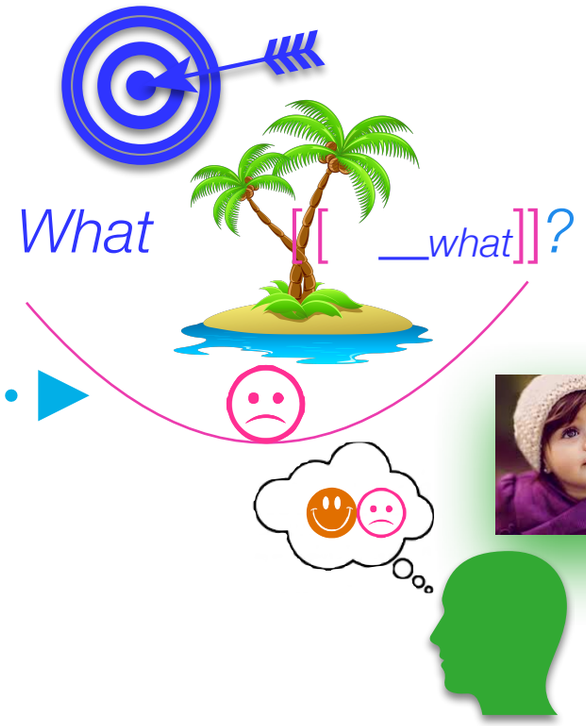
Use Bayesian inference to search through the hypothesis space of all possible building blocks (O'Donnell 2015) and find an efficient set for children's input.

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

# Learn the right building blocks

So that's what the modeled child will do

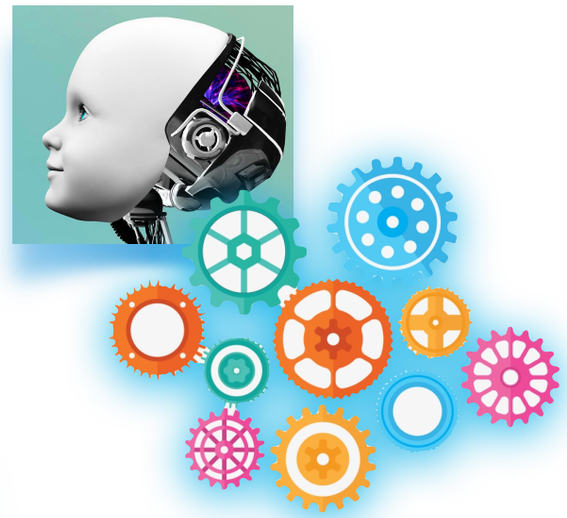


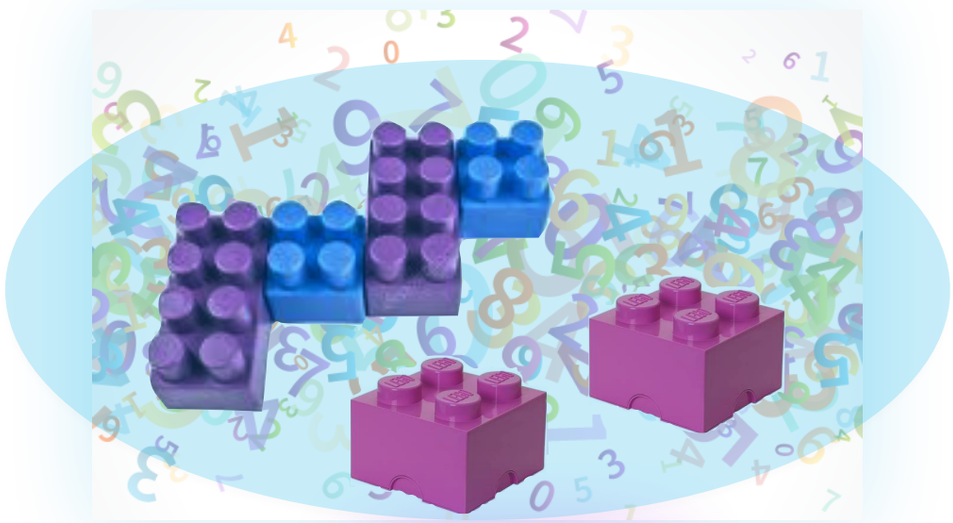



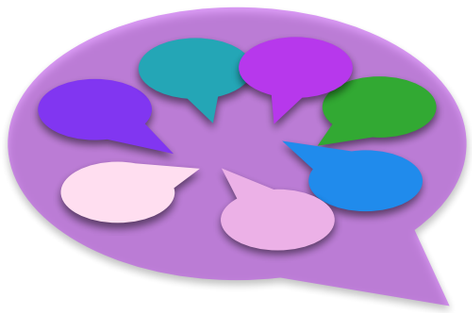

*Dickson et al. 2022, 2024, in prep.*

# 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?

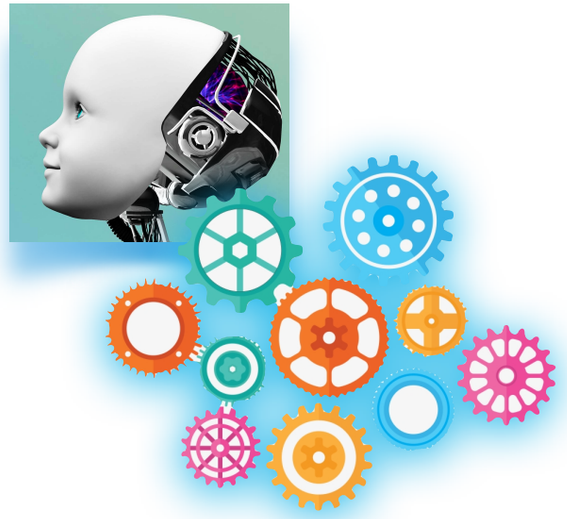


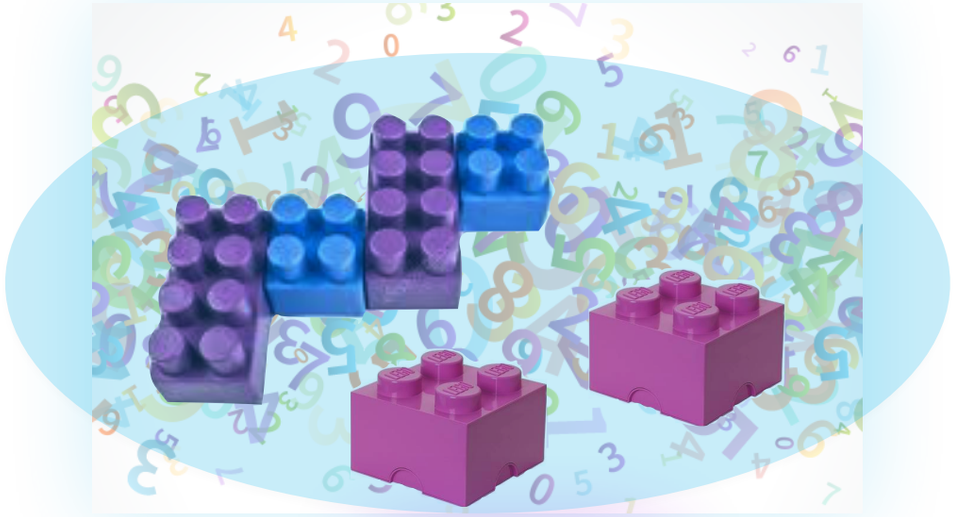


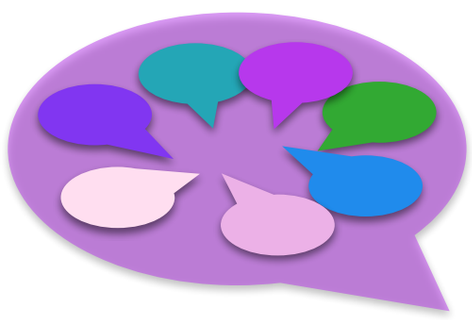



*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory



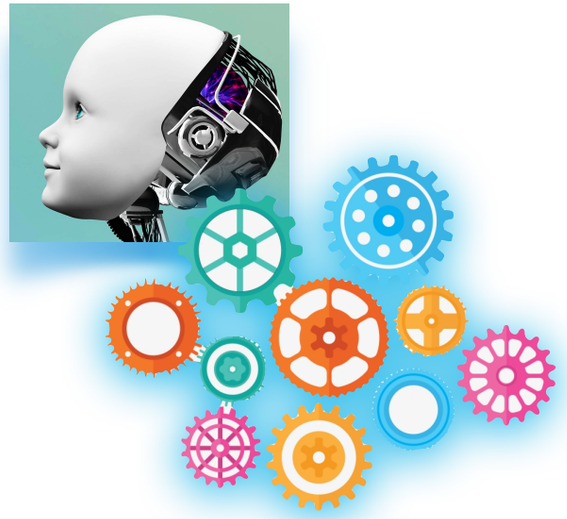


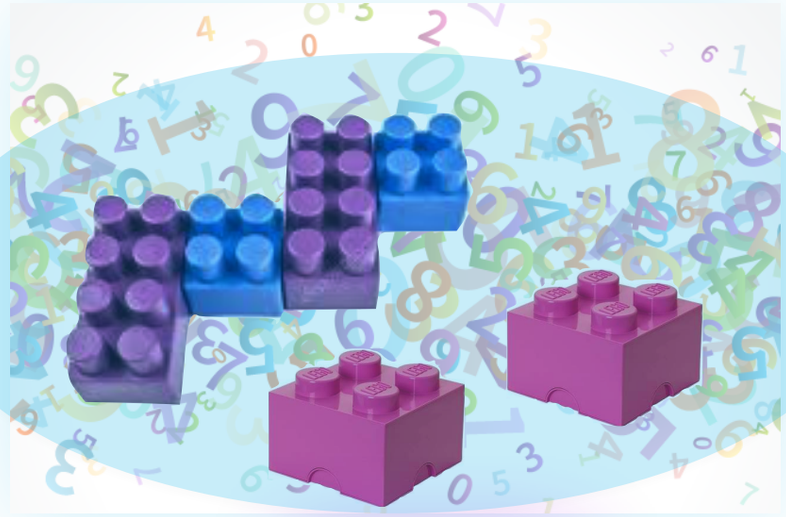


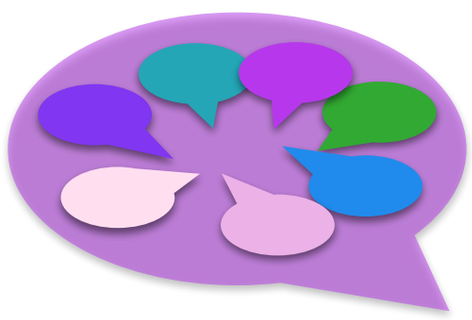

What's the input look like?

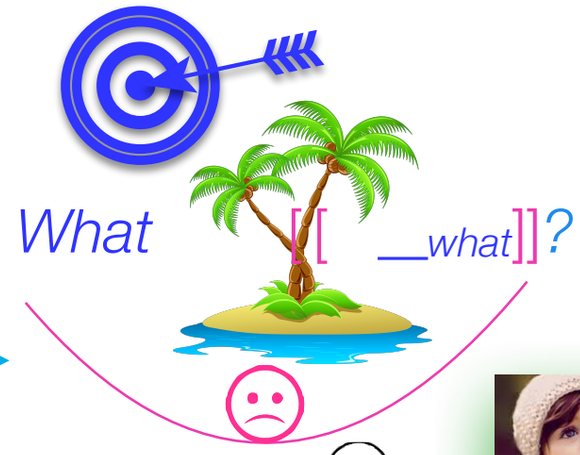
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory





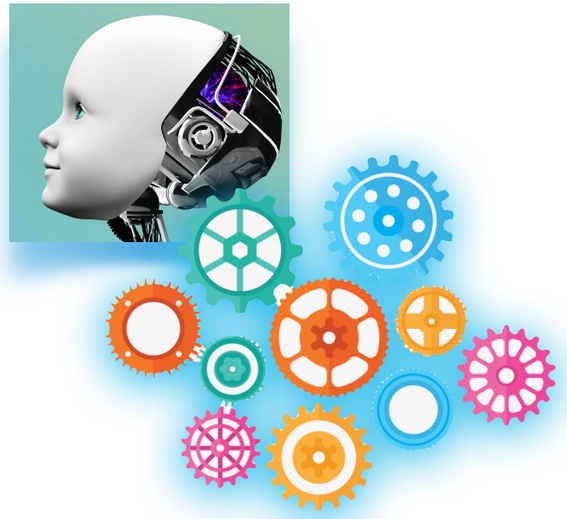


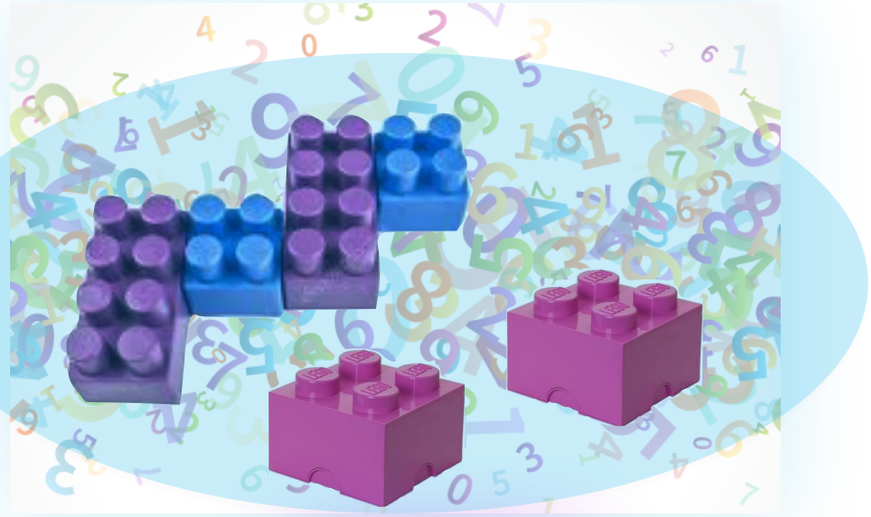


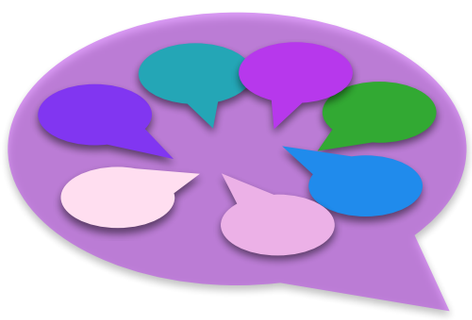

12.7K *wh*-dependencies from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 children between the ages of 1 and 5 years old.

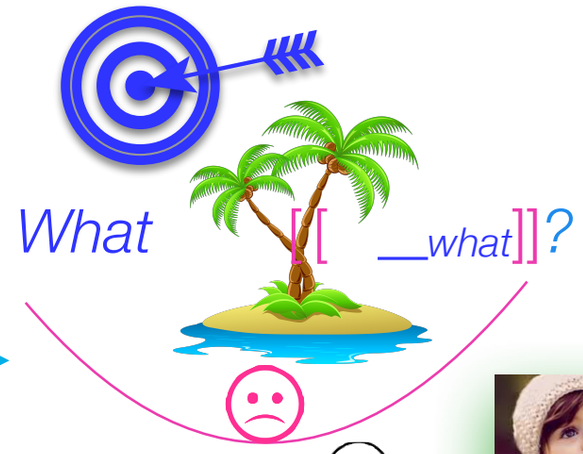
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

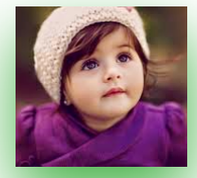
# Evaluating the theory



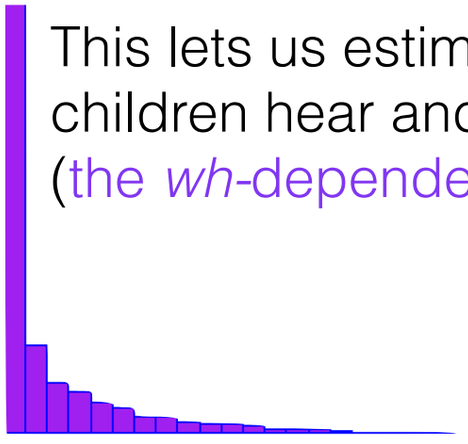






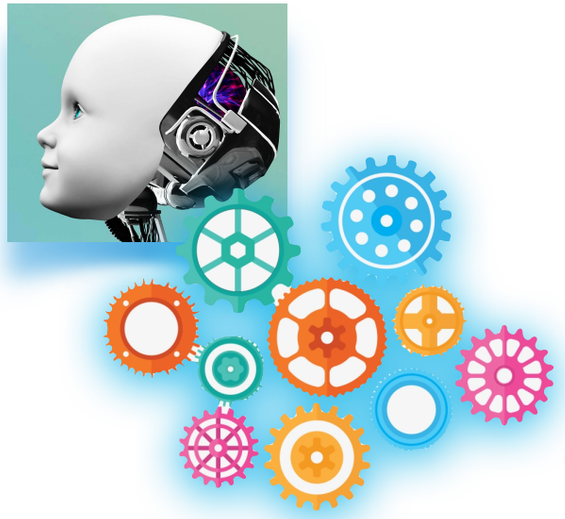



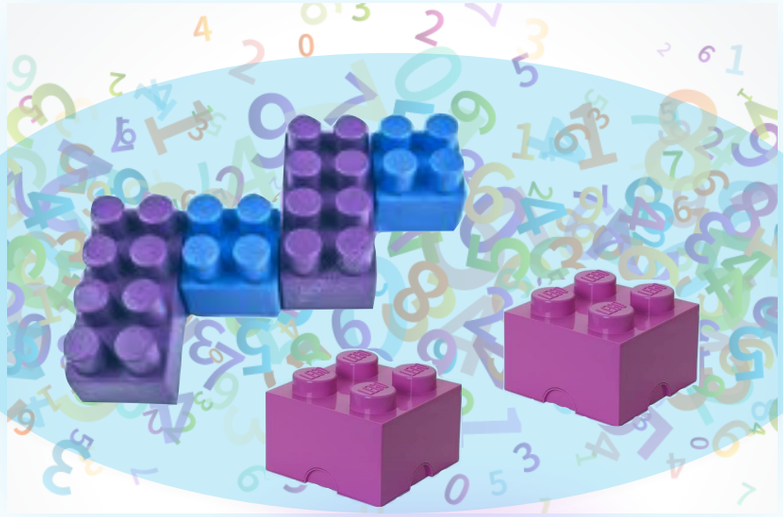
This lets us estimate which *wh*-dependencies children hear and how often they hear them (the *wh*-dependency distribution).

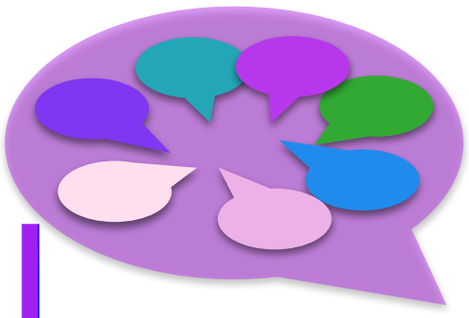


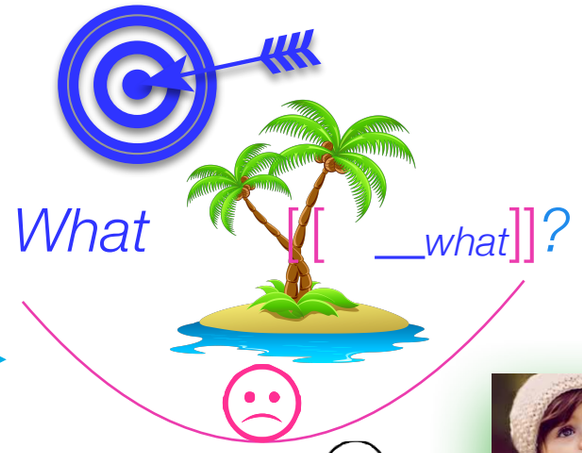
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

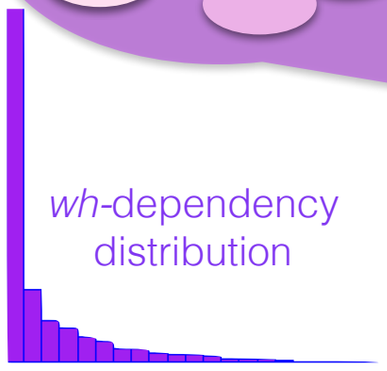
# Evaluating the theory







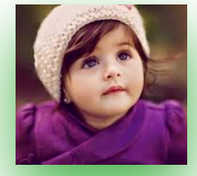



#

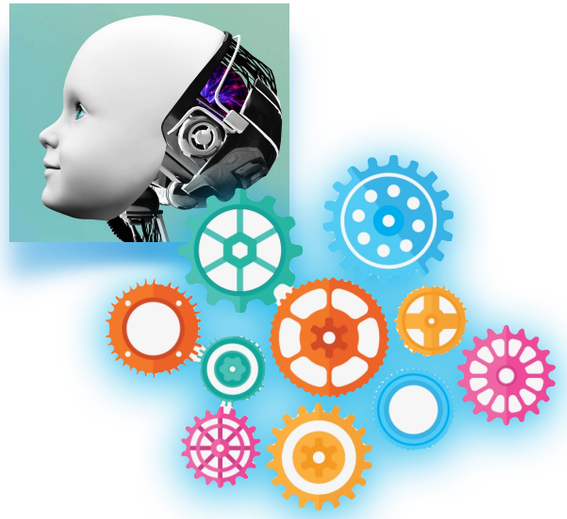
We can then estimate how many *wh*-dependencies children hear during the learning period.

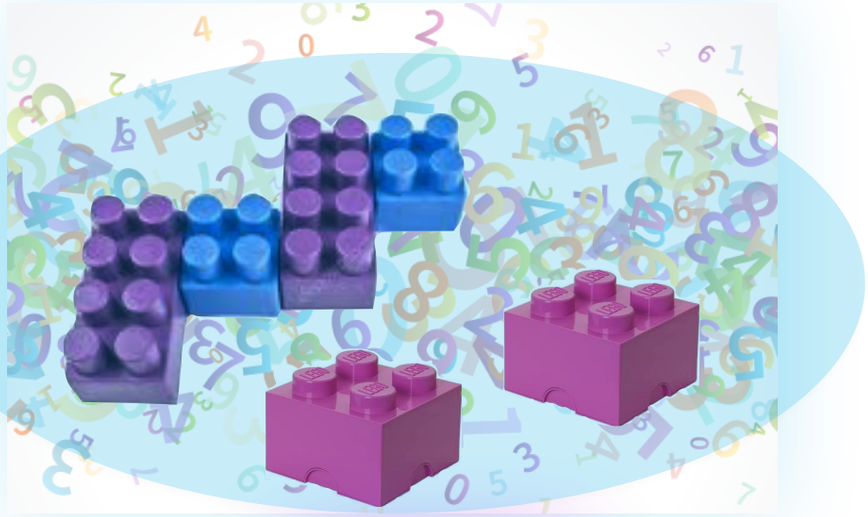
(<60 months)

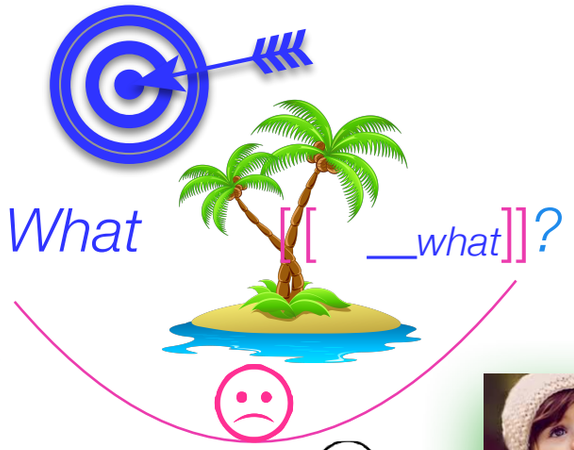


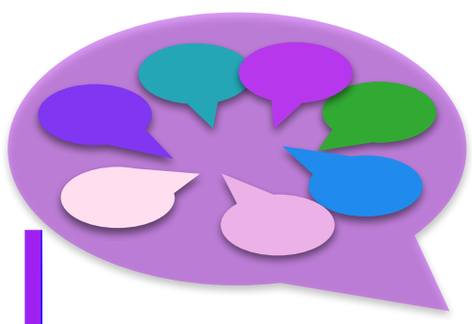
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory







*wh*-dependency distribution

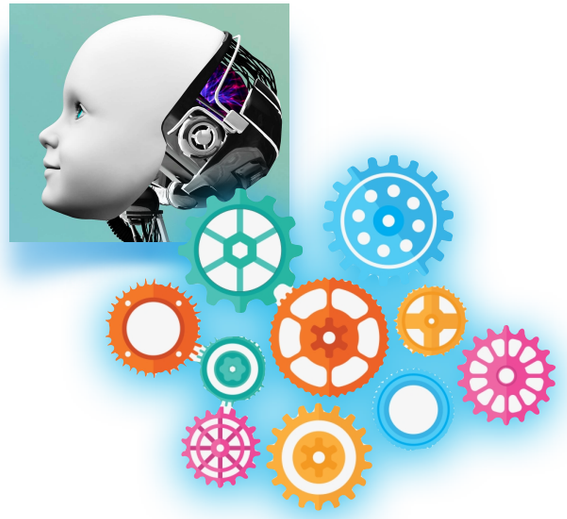
#

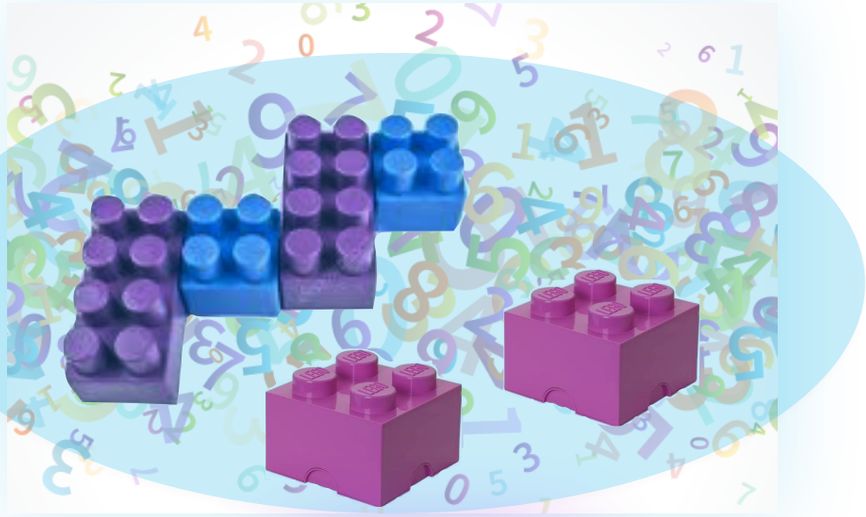
(<60 months)

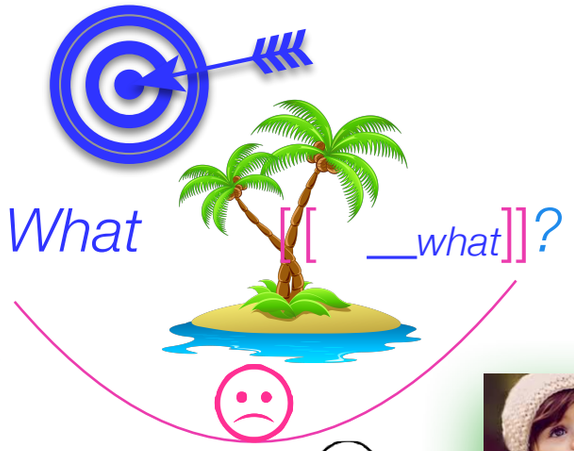
Children begin to represent the full structure of *wh*-dependencies (e.g., *wh*-questions and relative clauses) around 18 months: Perkins & Lidz 2021.

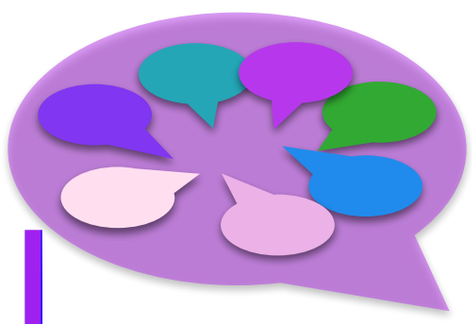
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory







*wh*-dependency distribution

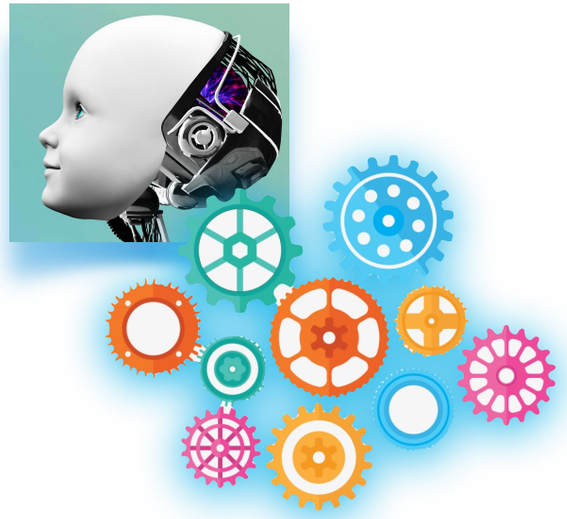
#

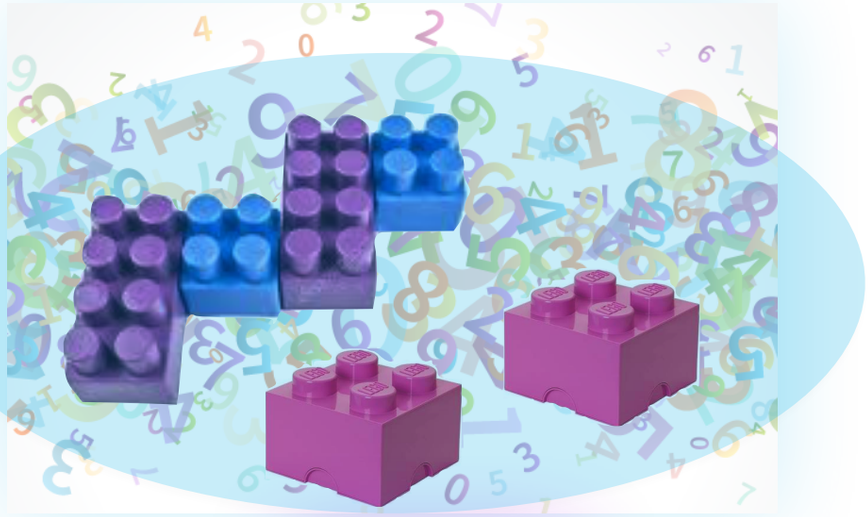
(18 months ≤ age < 60 months)

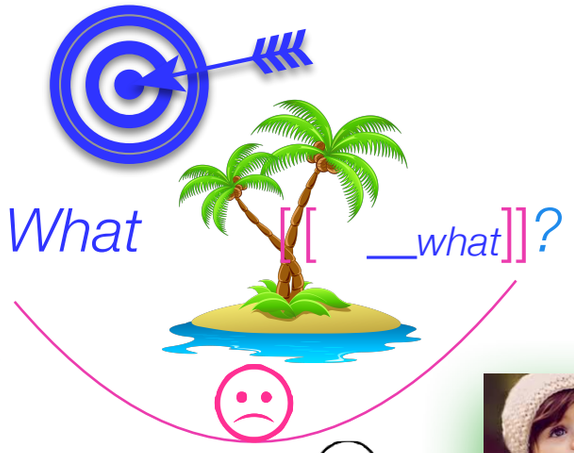
Educated guess: This is when children can start processing *wh*-dependencies reliably from their input.

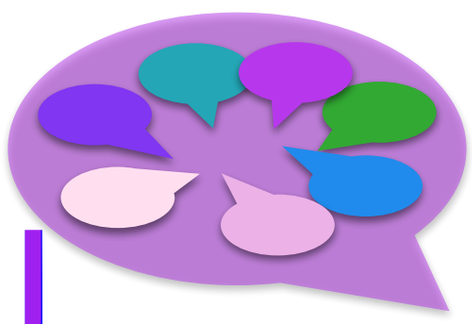
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory





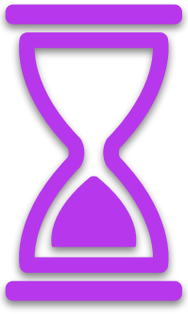


*wh*-dependency distribution

#

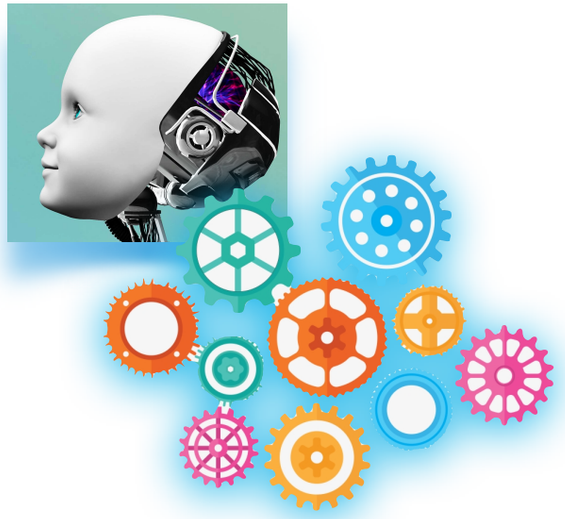
(18 months ≤ age < 60 months)

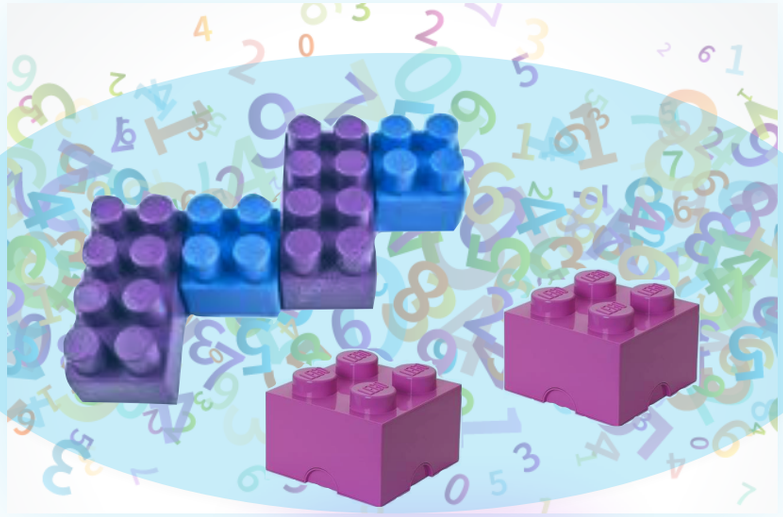


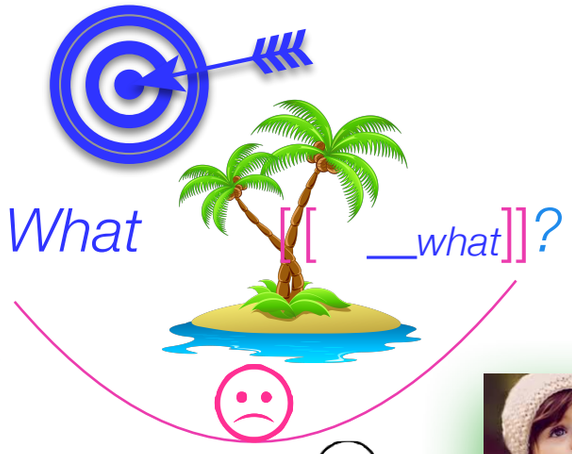
How many minutes is this? In particular, children are awake for only a certain portion of the day at different ages (Davis et al. 2004).

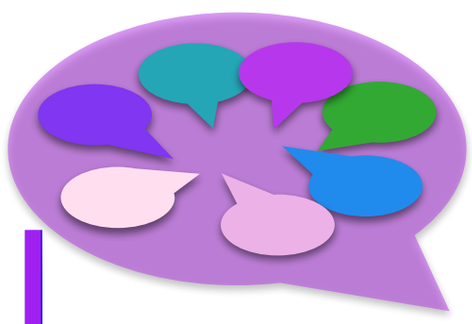
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory







*wh*-dependency distribution

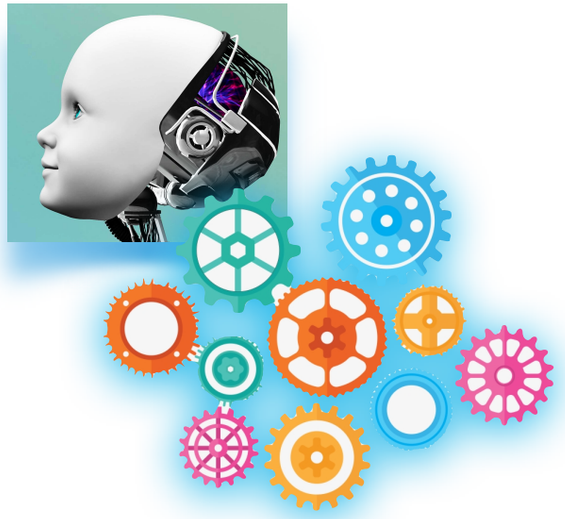
#

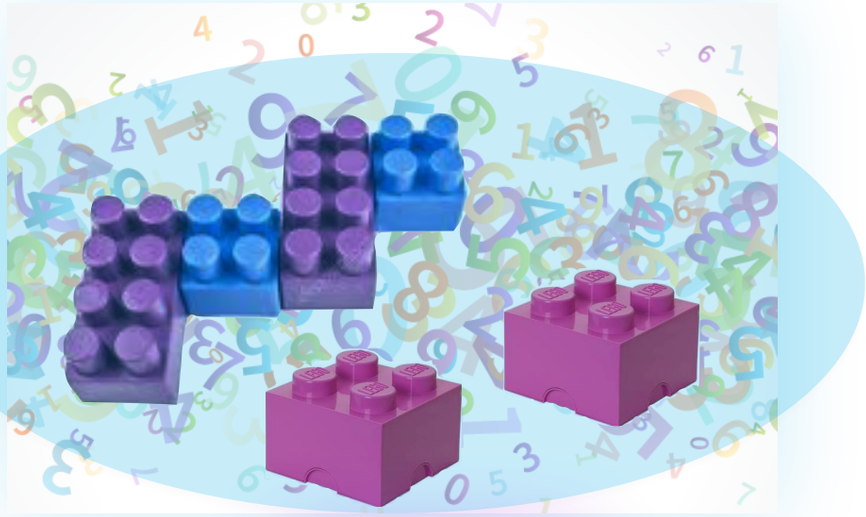
(18 months ≤ age < 60 months)

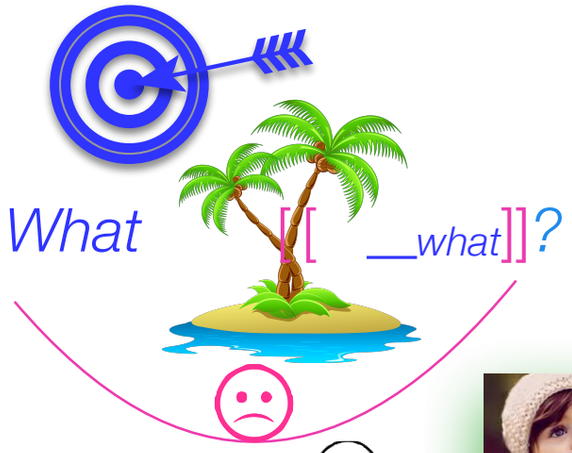
How many minutes:
≈10,442,258

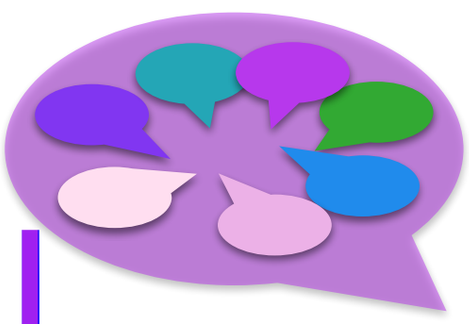
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory







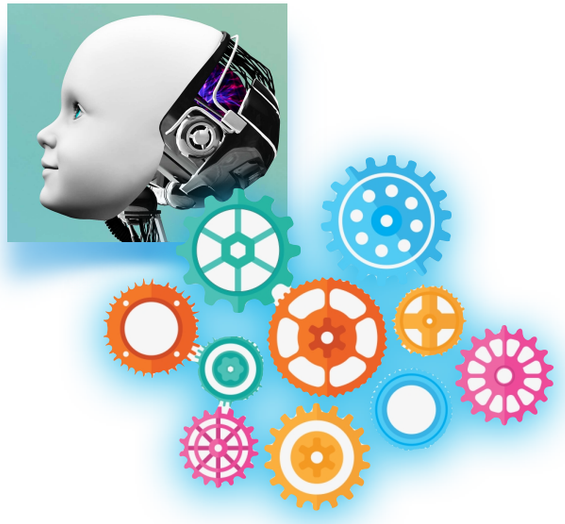
*wh*-dependency distribution

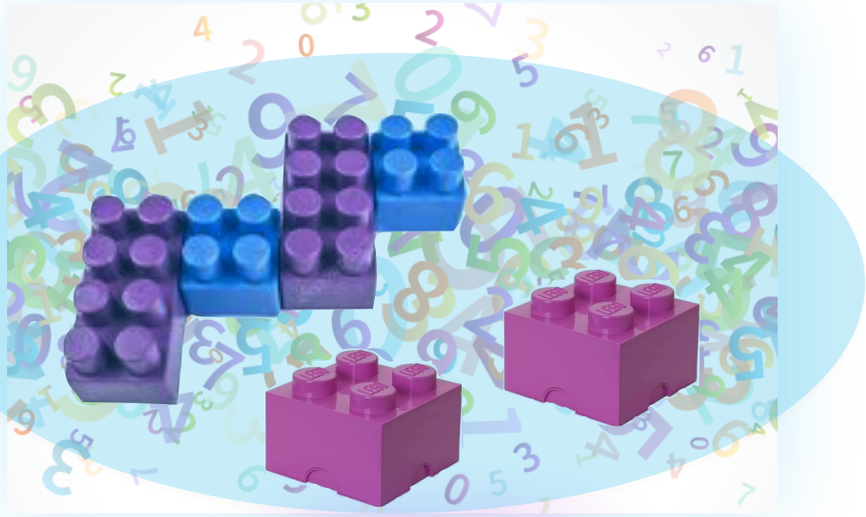
(≈10,442,258 minutes)

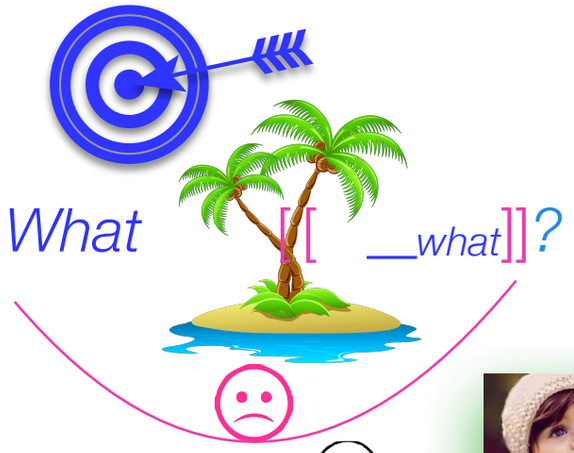
How many *wh*-dependencies is this? #

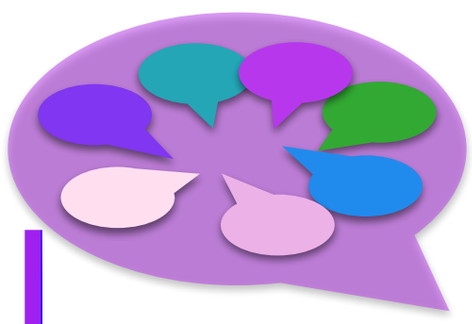
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory







*wh*-dependency distribution

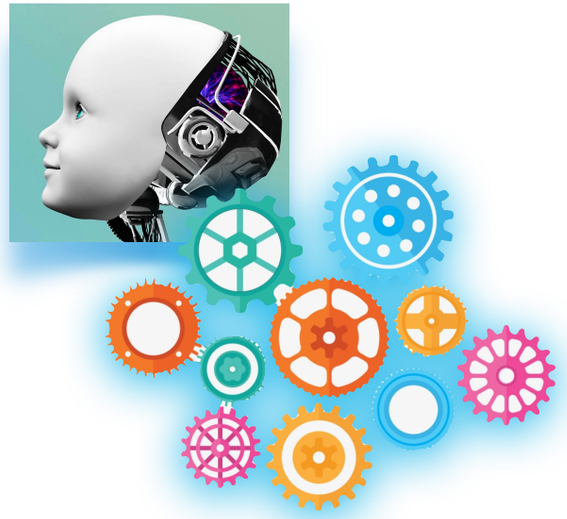
(≈10,442,258 minutes)

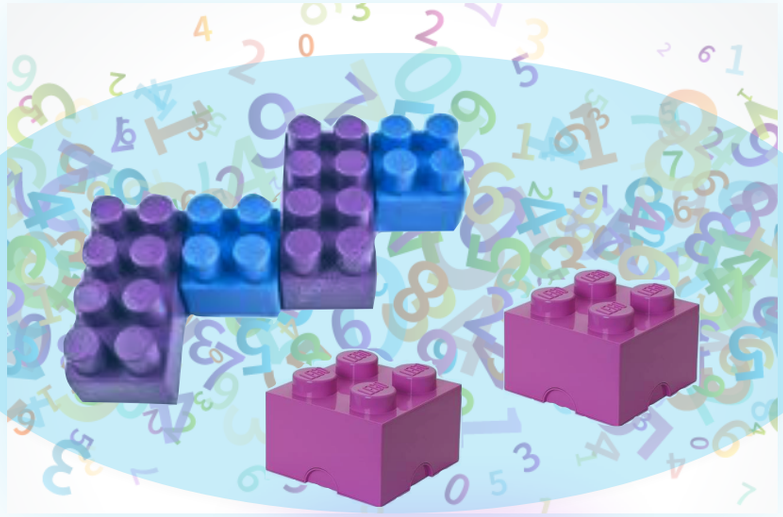
#

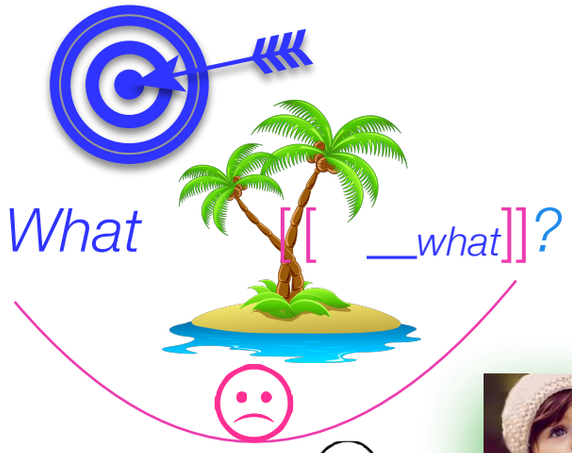
Hoff-Ginsberg (1998) and Rowe (2012): Estimates of utterances per minute in speech directed at children from different backgrounds.

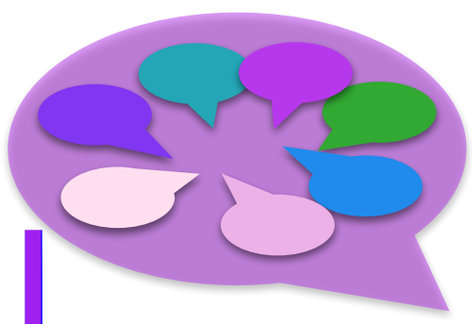
*Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022*
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory



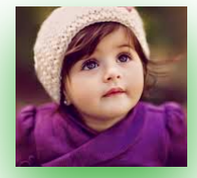




*wh*-dependency distribution

(≈10,442,258 minutes)

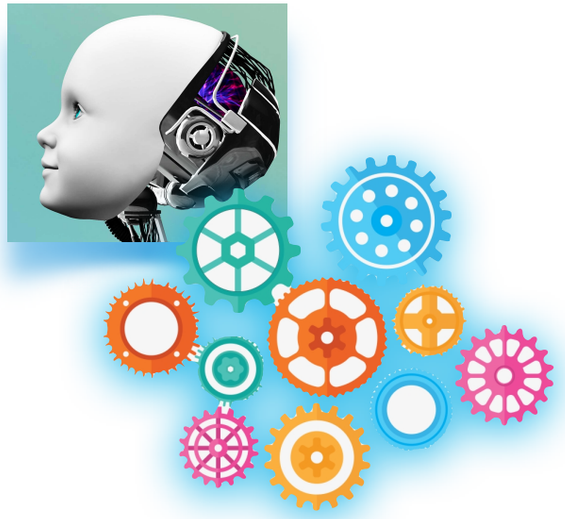


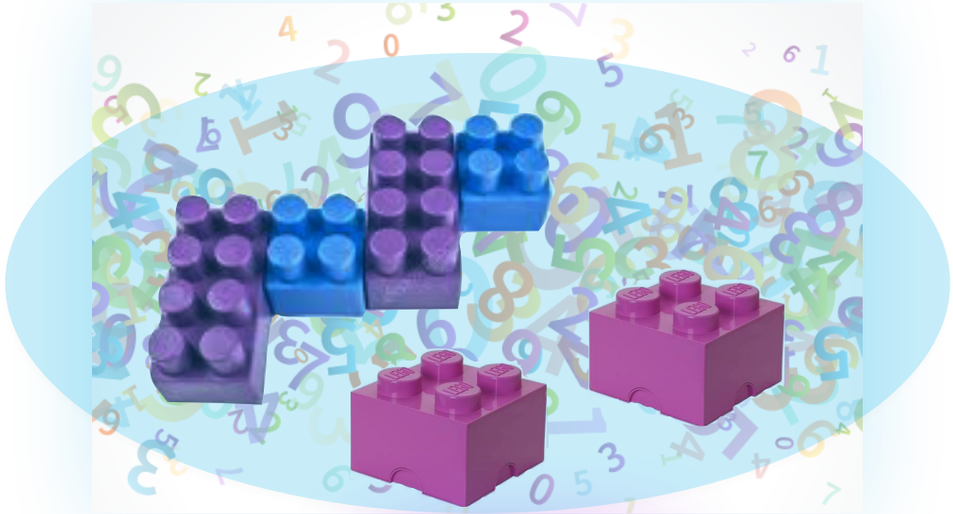
#

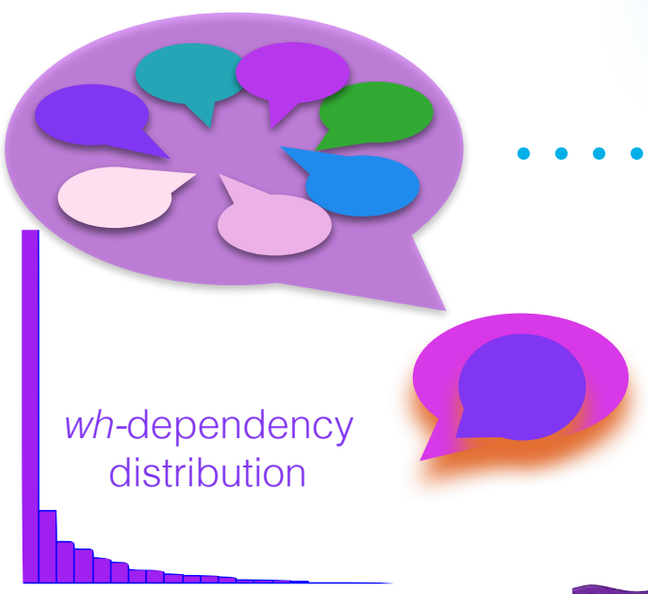
≈ 2.15 million
*wh*-dependencies

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

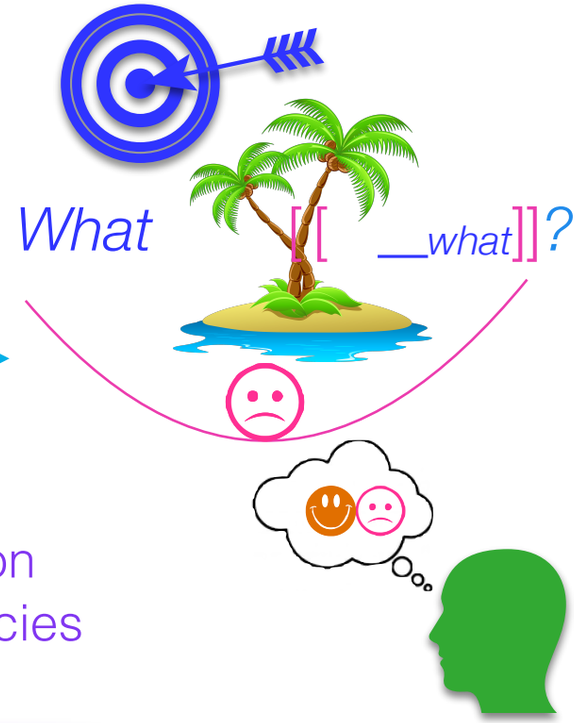
# Evaluating the theory







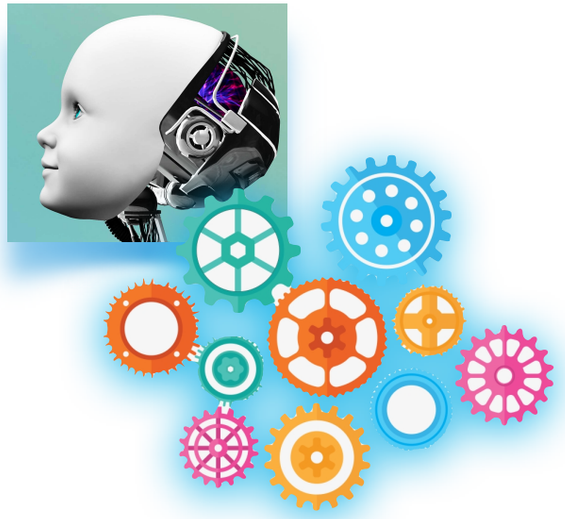

≈ 2.15 million
*wh*-dependencies

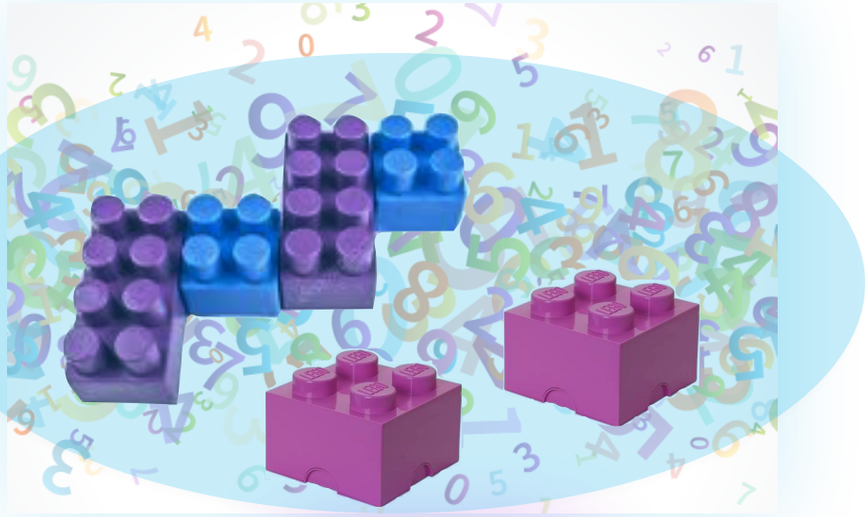


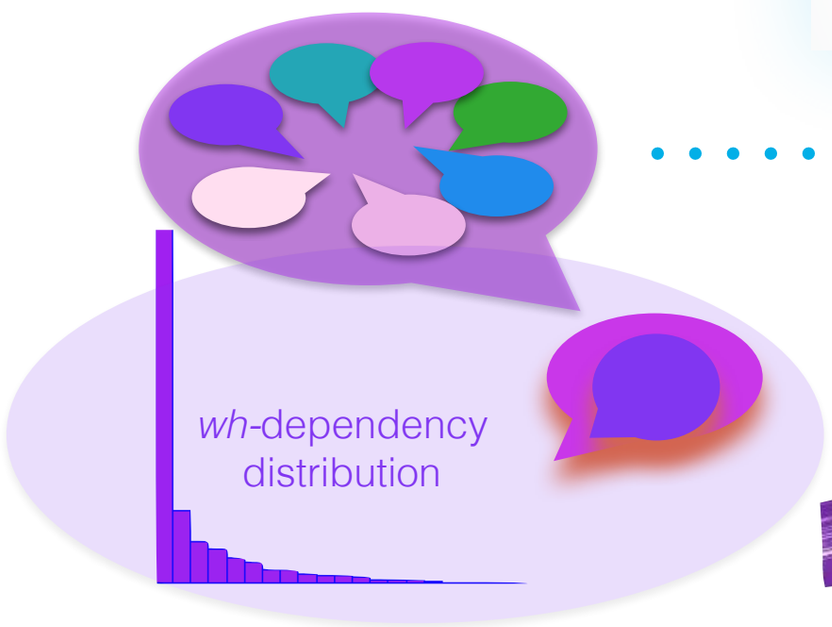

An additional wrinkle:
Children's memory isn't perfect

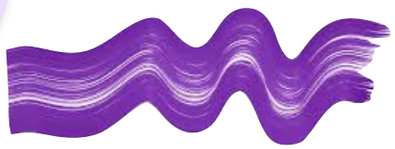
*Dickson et al. 2024*

# Evaluating the theory

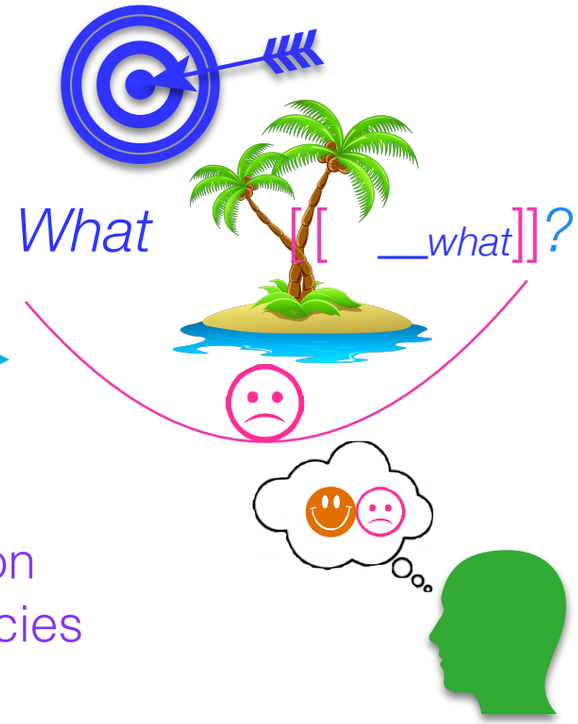






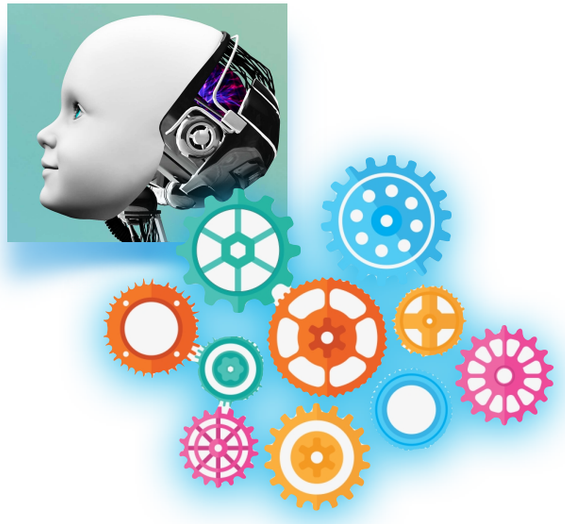
# ≈ 2.15 million *wh*-dependencies

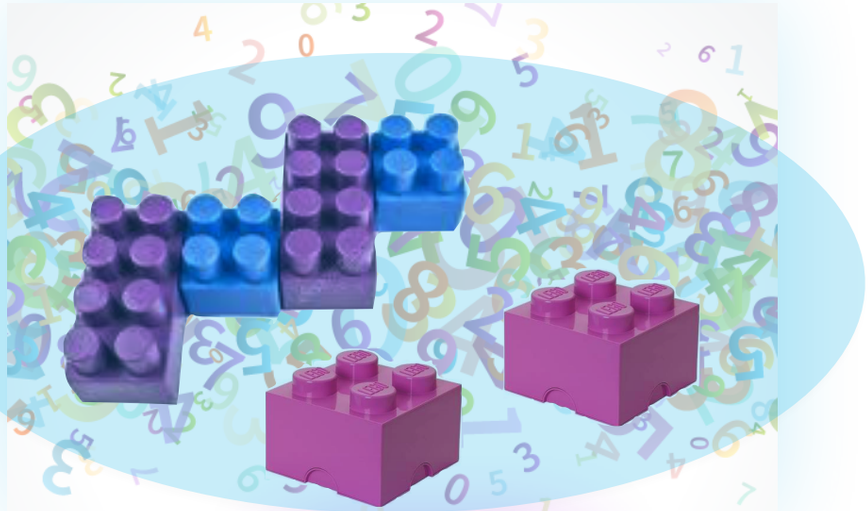


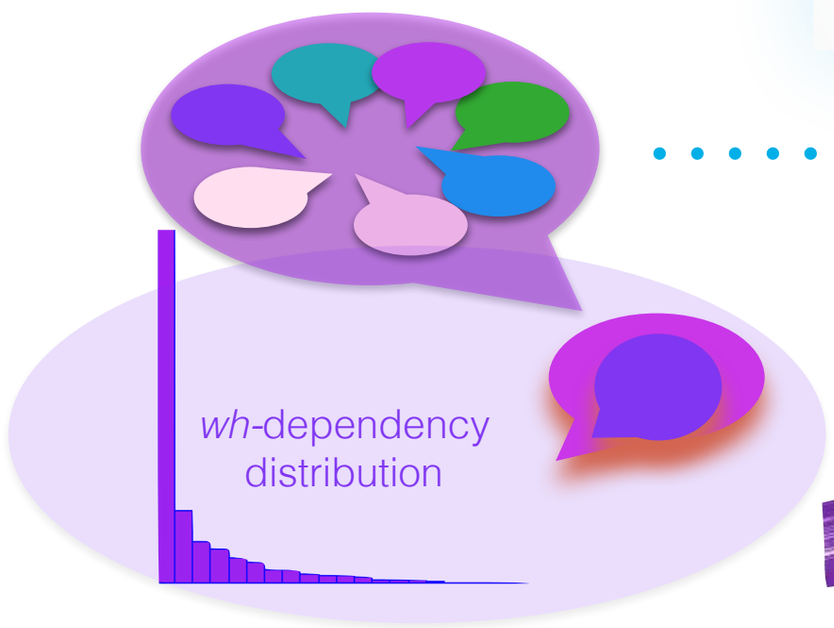

Memory is an important part of processing dependencies (McElree et al. 2003).

*Dickson et al. 2024*

# Evaluating the theory







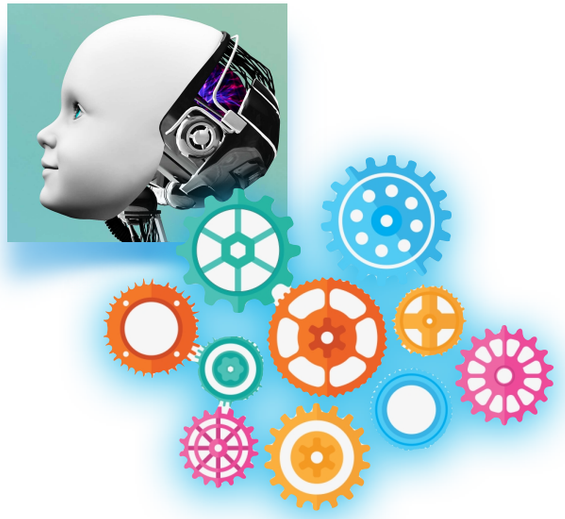

# ≈ 2.15 million *wh*-dependencies

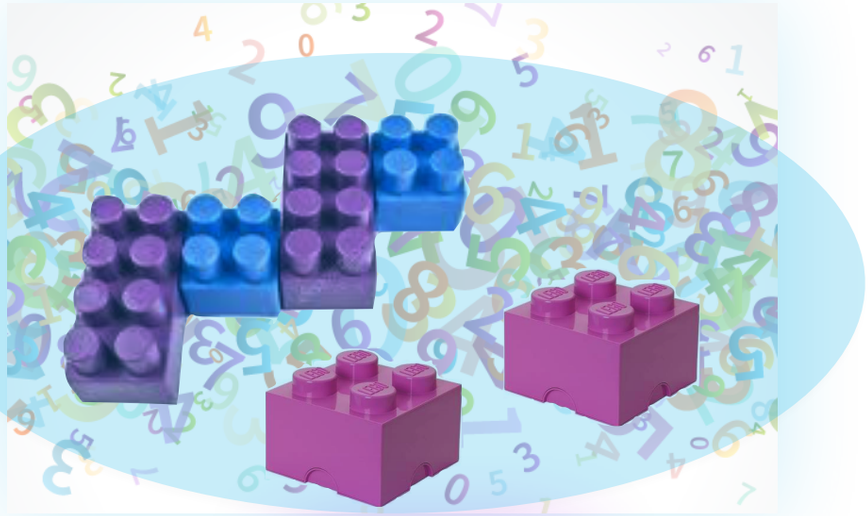


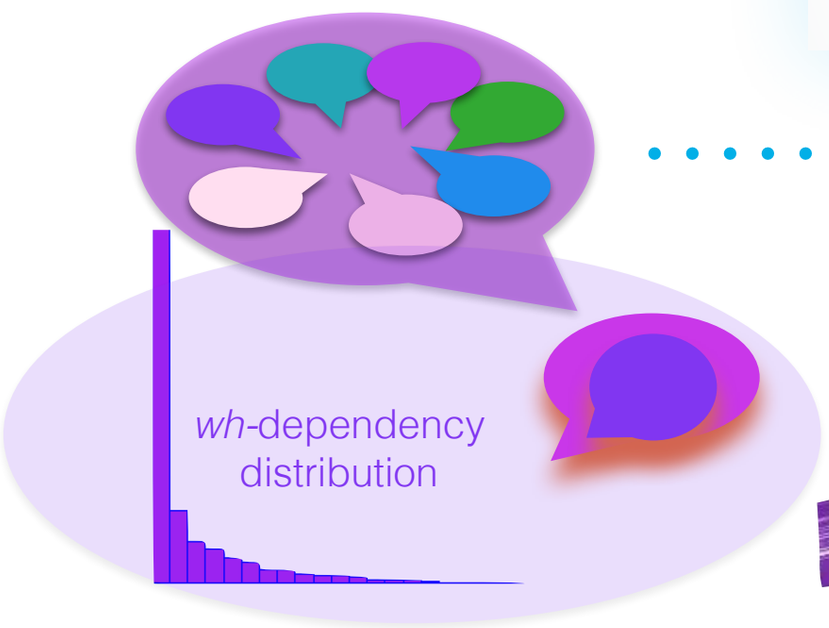

Children's short term memory, along with related abilities like encoding information with context and maintaining attention, develops over time (Paris 1978, Gathercole et al. 2004, Fandakova et al. 2014).

*Dickson et al. 2024*

# Evaluating the theory







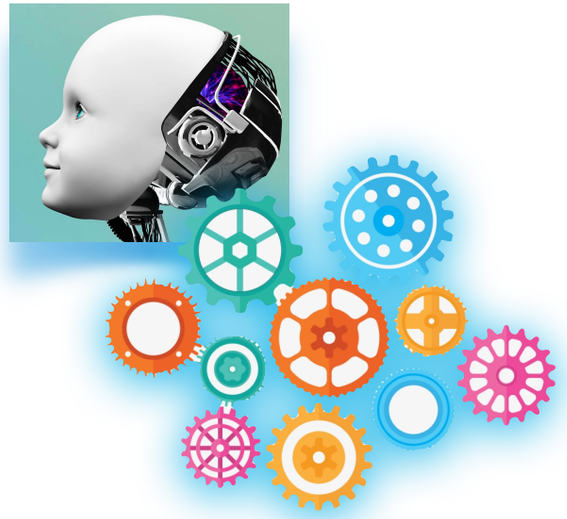

≈ 2.15 million
*wh*-dependencies

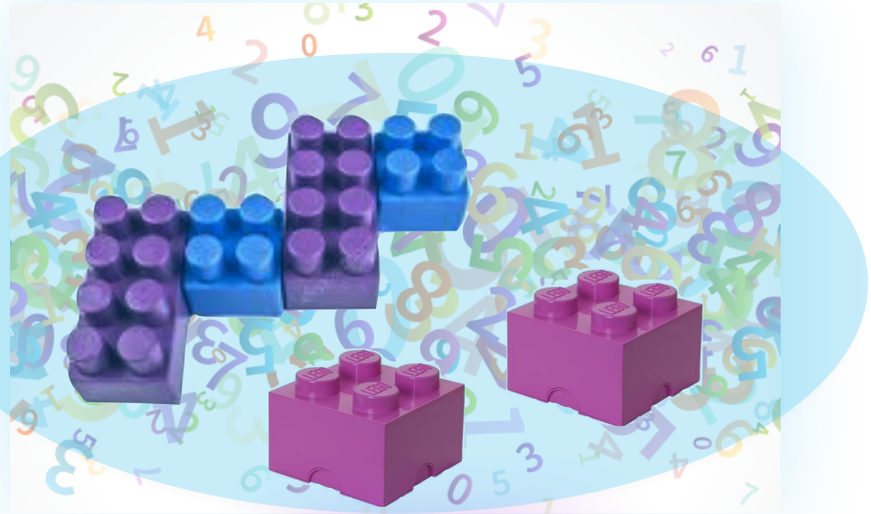


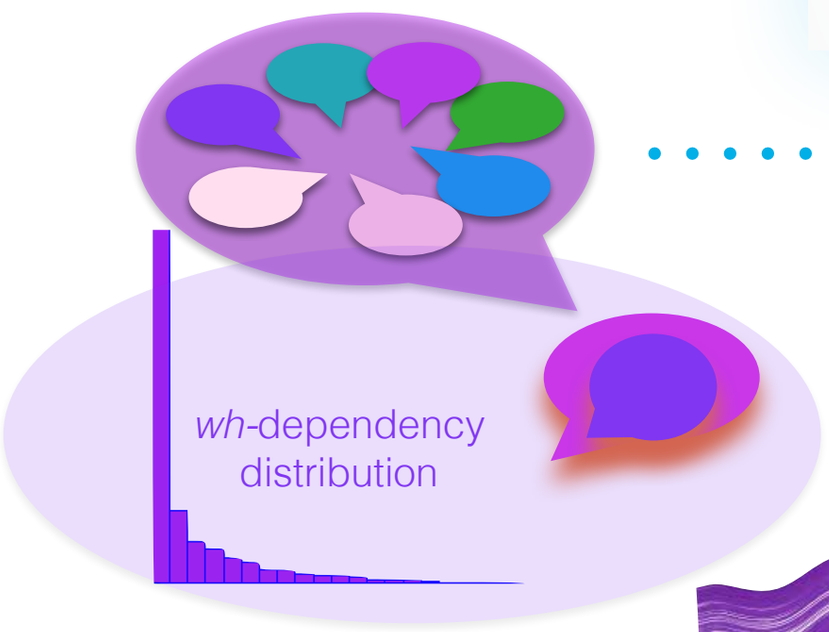

Upshot: Children might **not take in** all the *wh*-dependency distribution information in their input.

*Dickson et al. 2024*

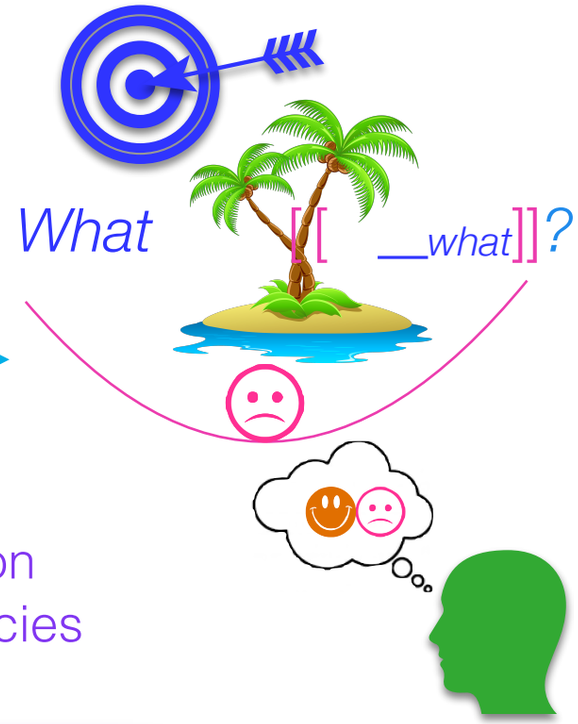
# Evaluating the theory







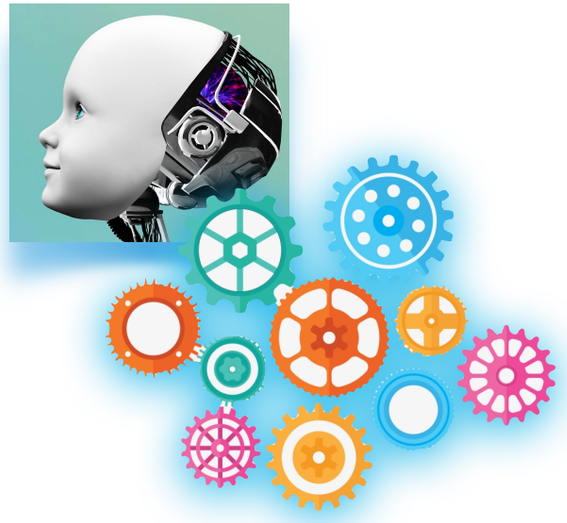

≈ 2.15 million
*wh*-dependencies



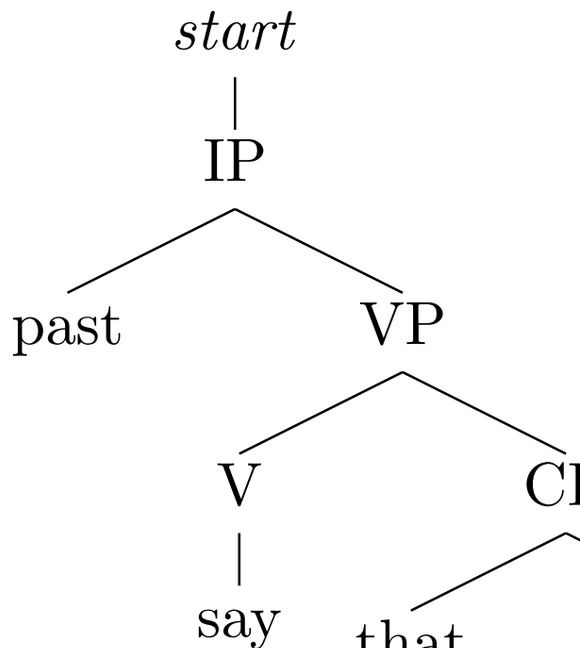

Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

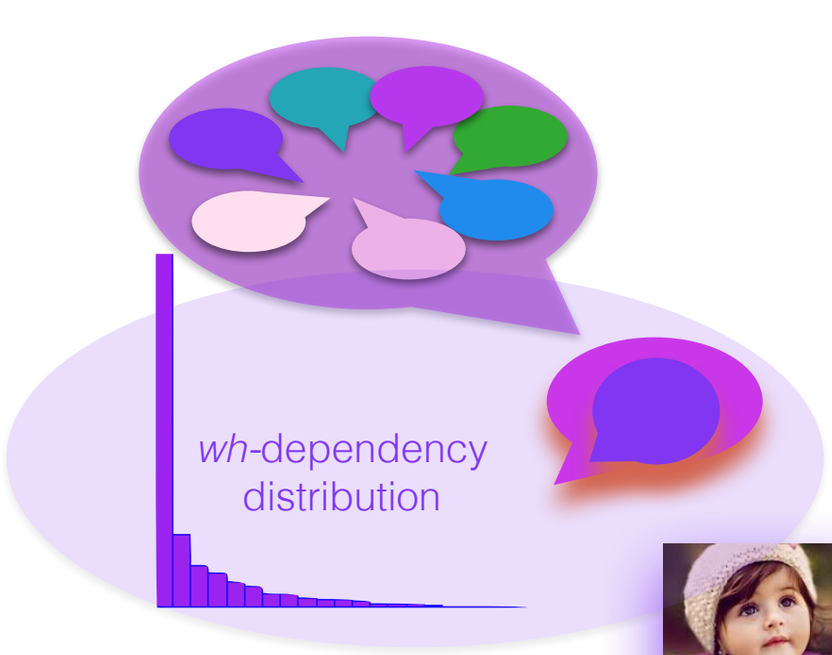
*Dickson et al. 2024*

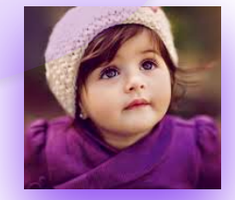
# Evaluating the theory

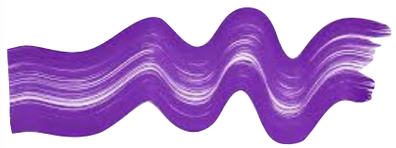


input



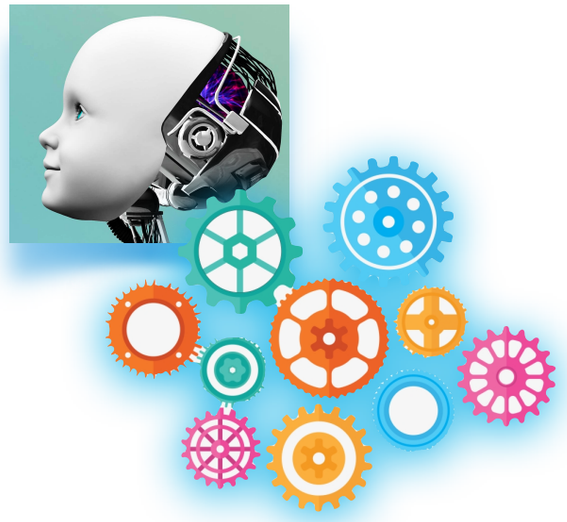





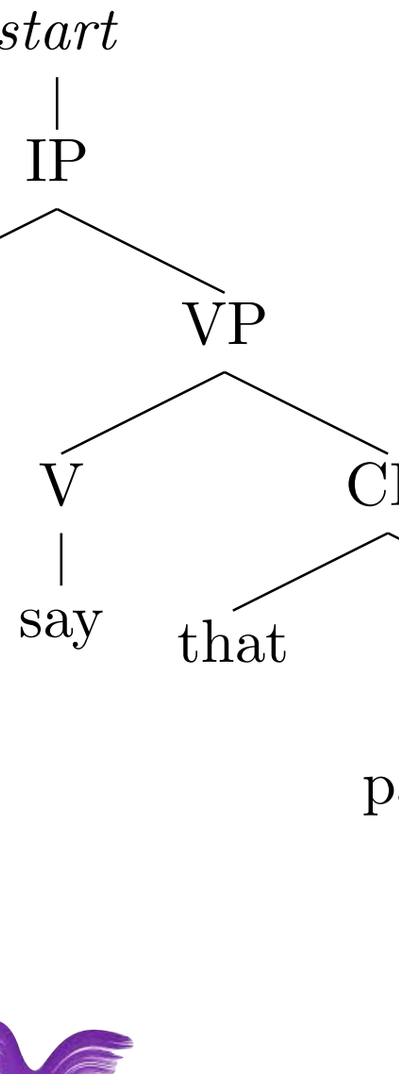
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

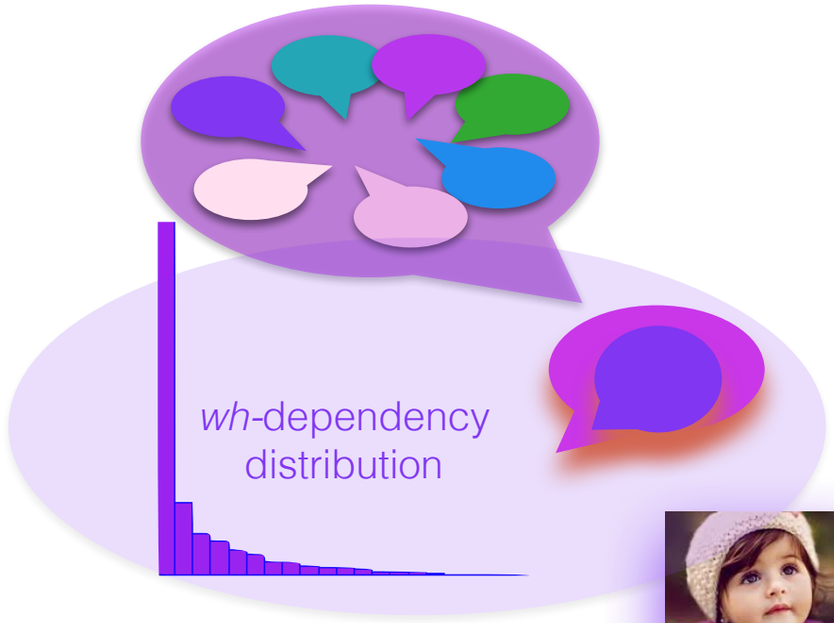
*Dickson et al. 2024*

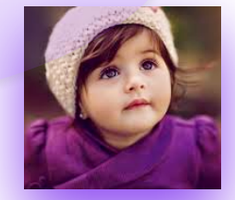
# Evaluating the theory

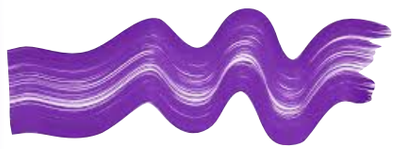


perfect intake



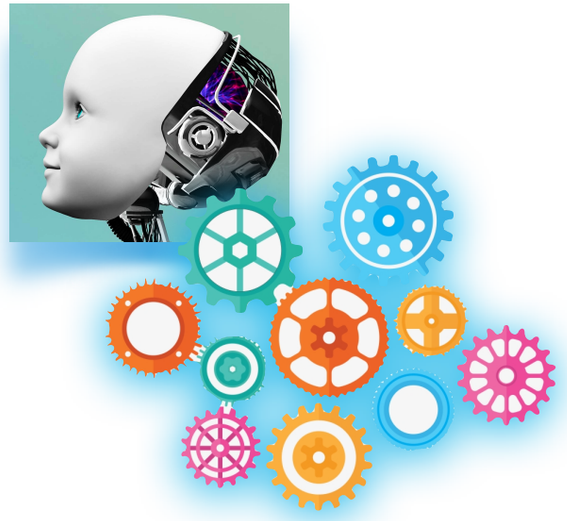





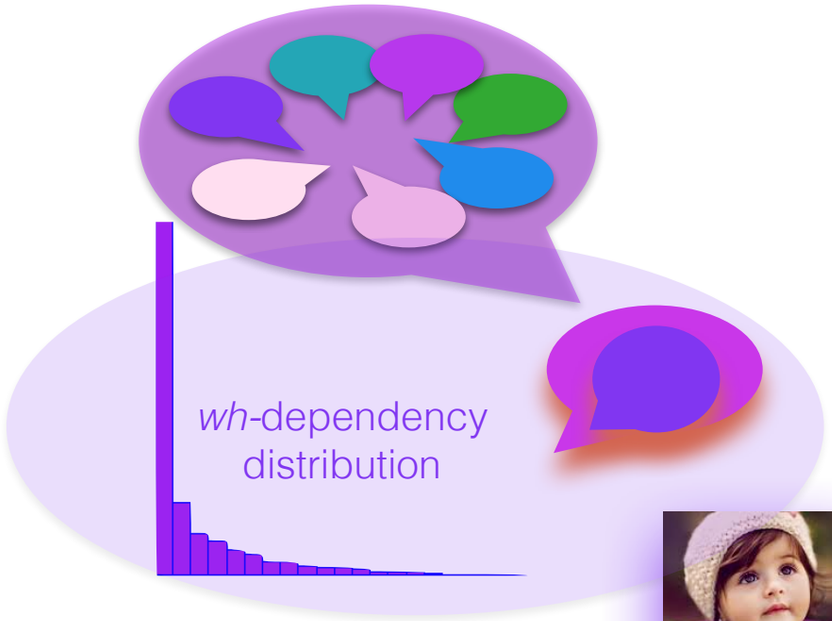
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

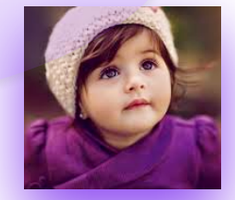
*Dickson et al. 2024*

# Evaluating the theory

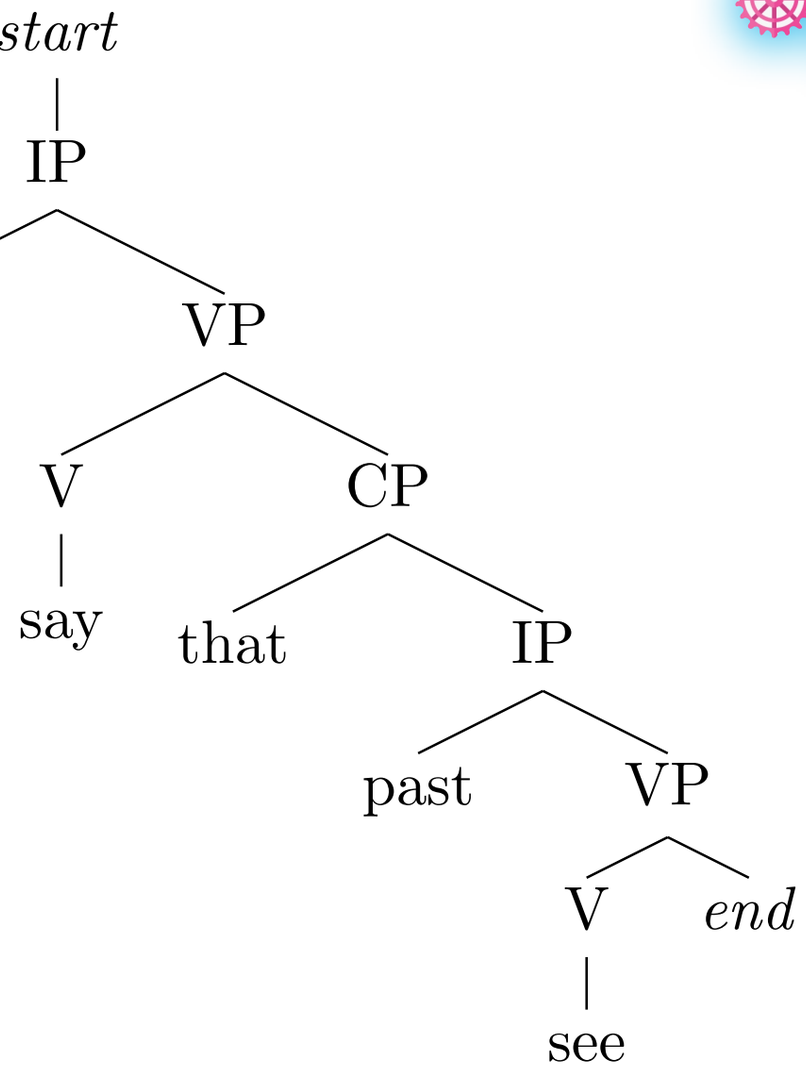


imperfect intake
(a little forgetting)



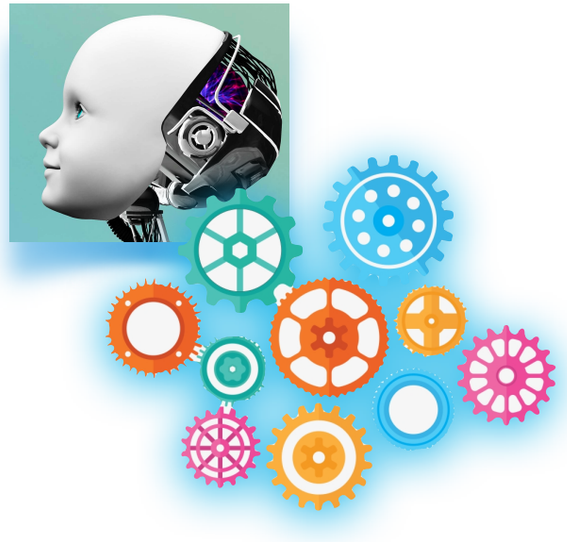



Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

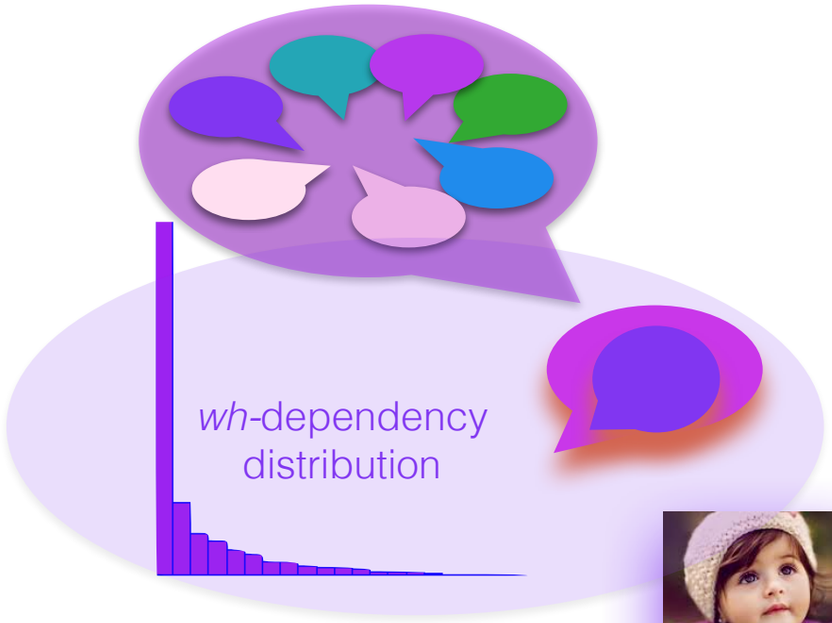


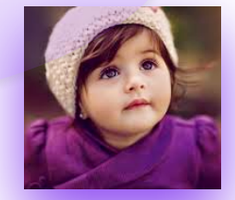

*Dickson et al. 2024*

# Evaluating the theory

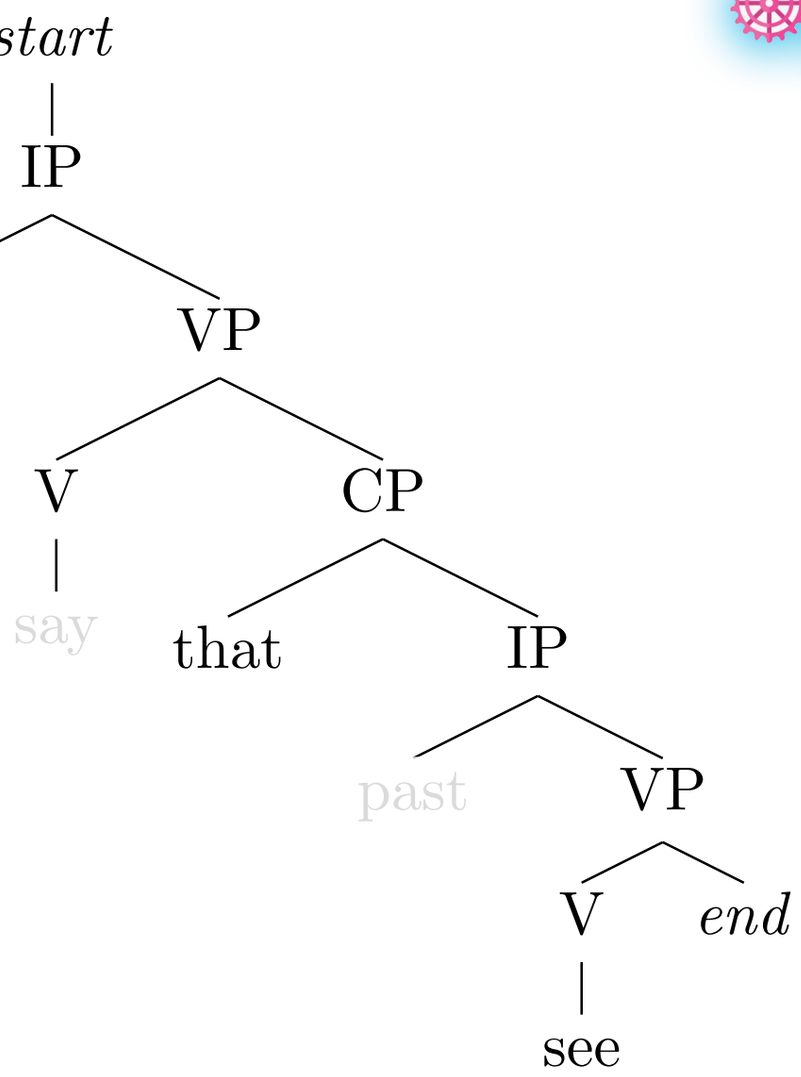


imperfect intake
(a little more forgetting)



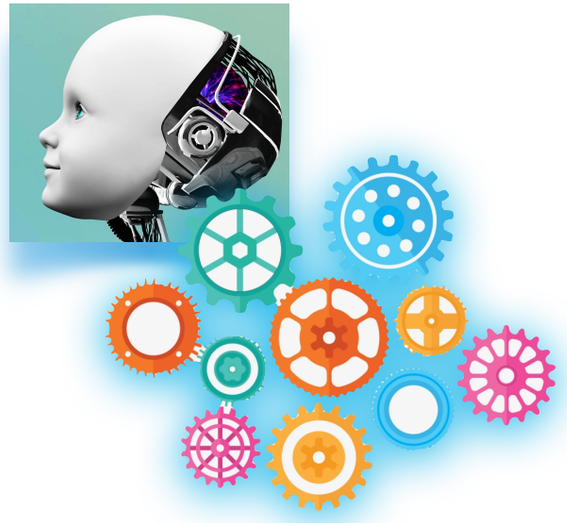



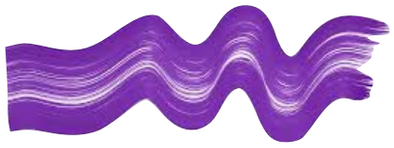
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.



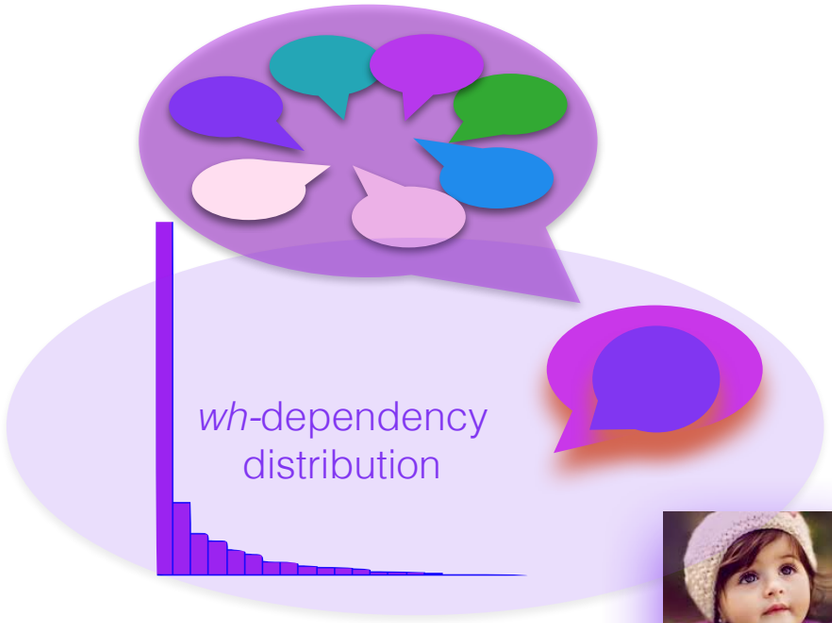

*Dickson et al. 2024*

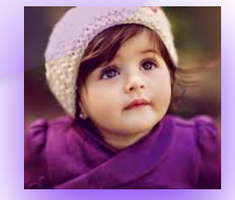
# Evaluating the theory

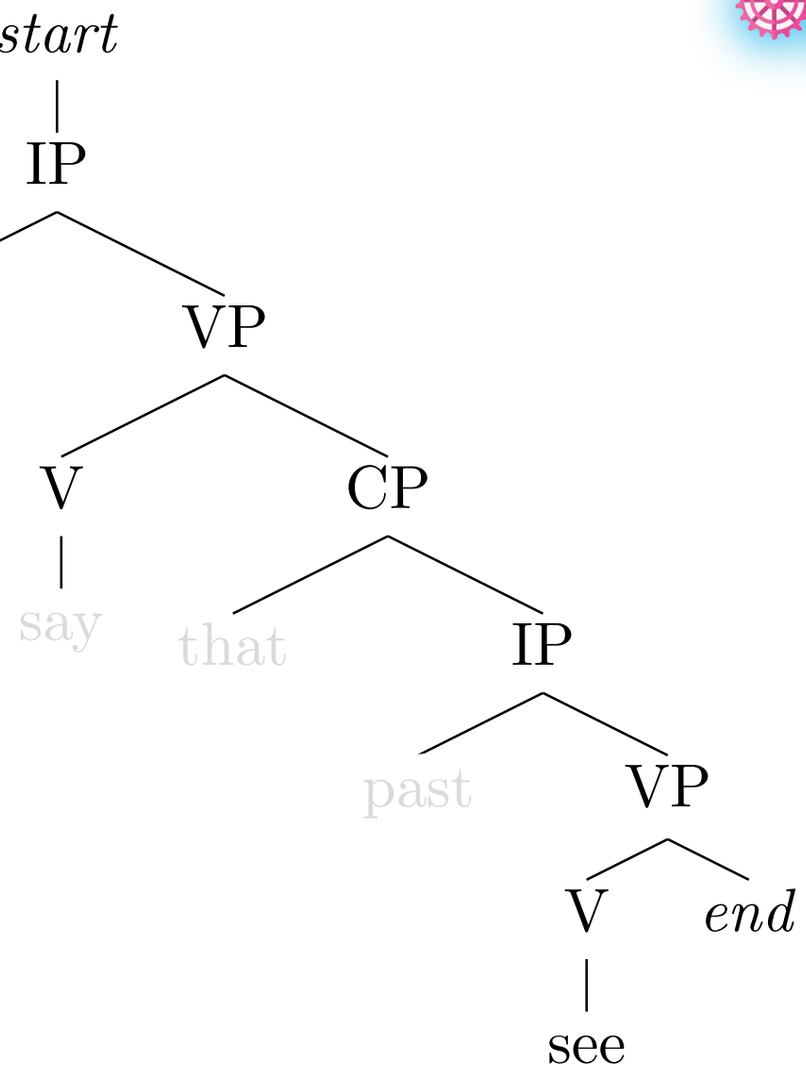




imperfect intake
(a lot of forgetting)



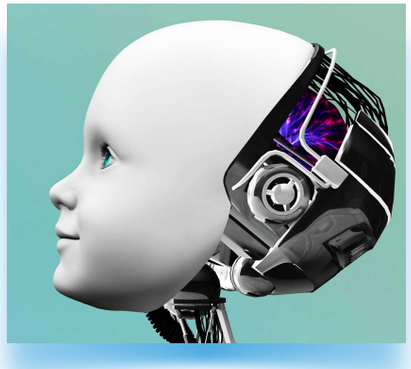



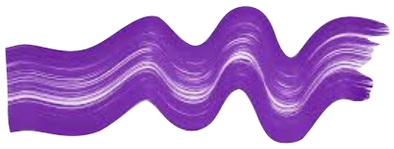



Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

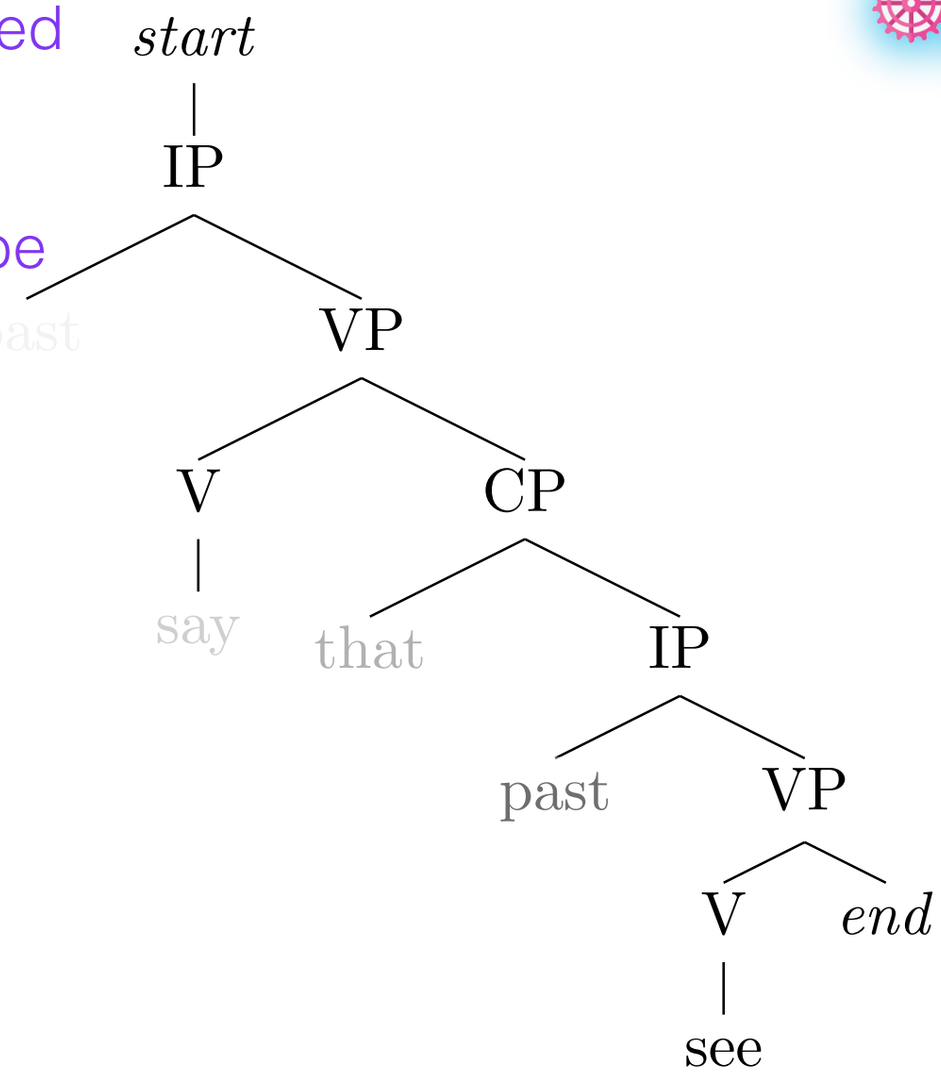
*Dickson et al. 2024*

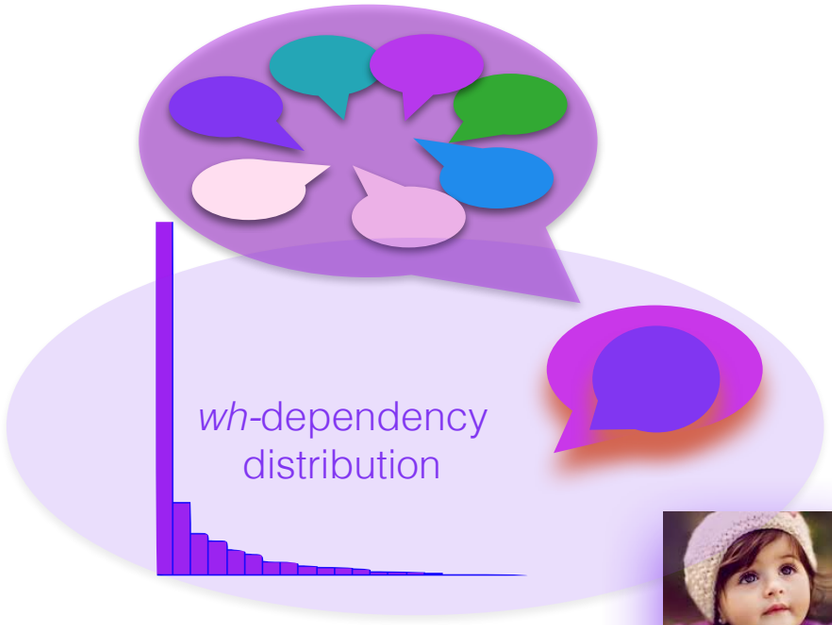
# Evaluating the theory

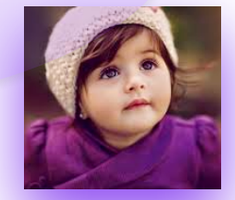




We can investigate memory-impacted learners with a **recency effect** (Anderson & Milson 1989), where more recent information is more likely to be remembered.



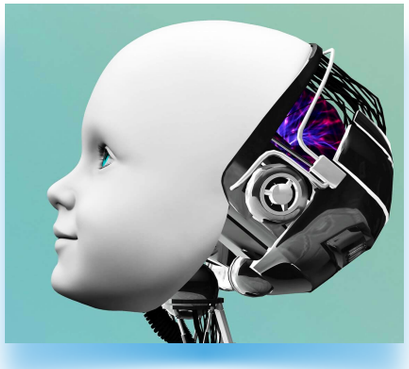



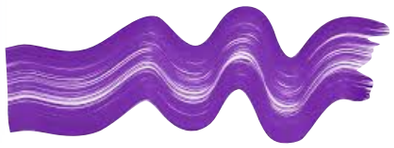



Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

*Dickson et al. 2024*

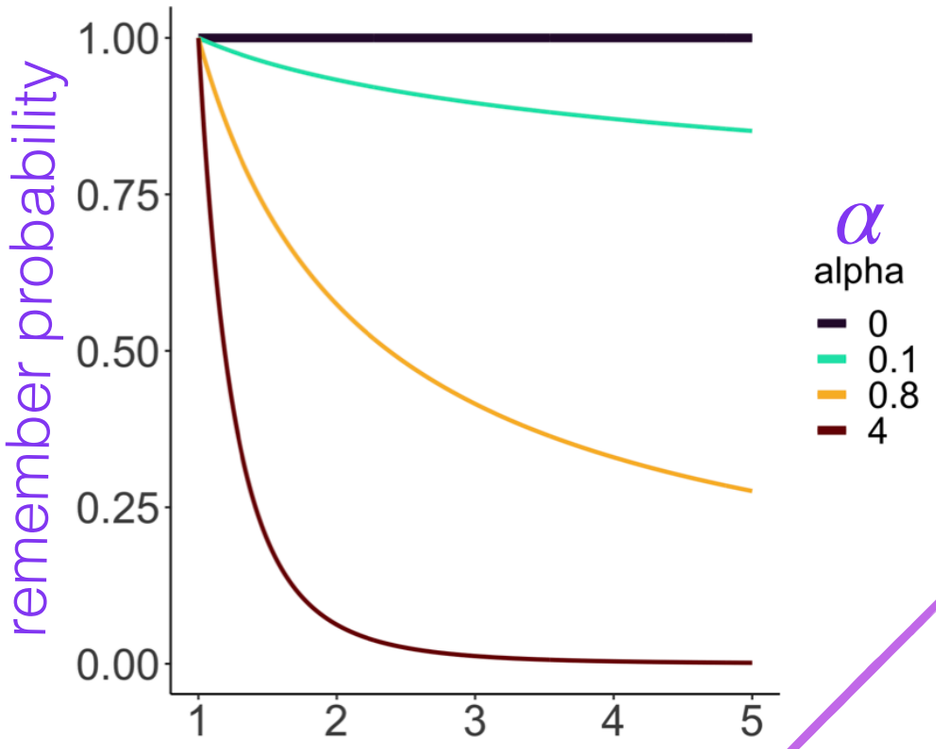
# Evaluating the theory

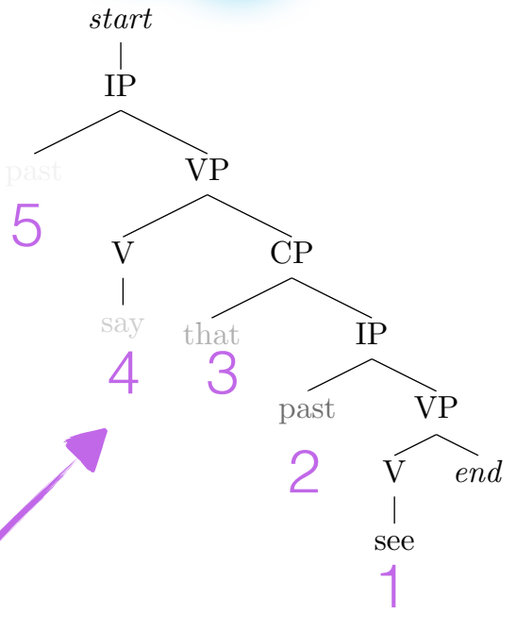


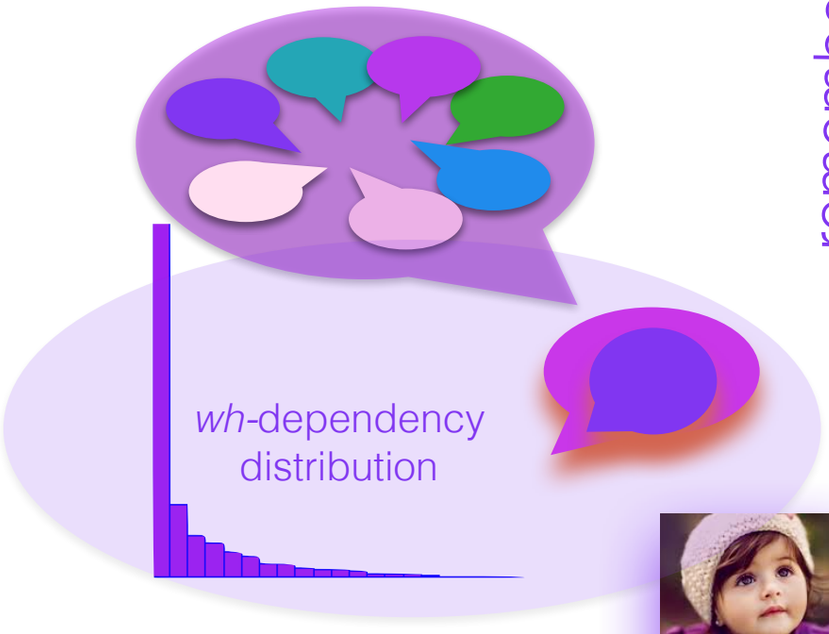


Anderson & Milson 1989

$$\frac{1}{(position+1)^{\alpha}}$$



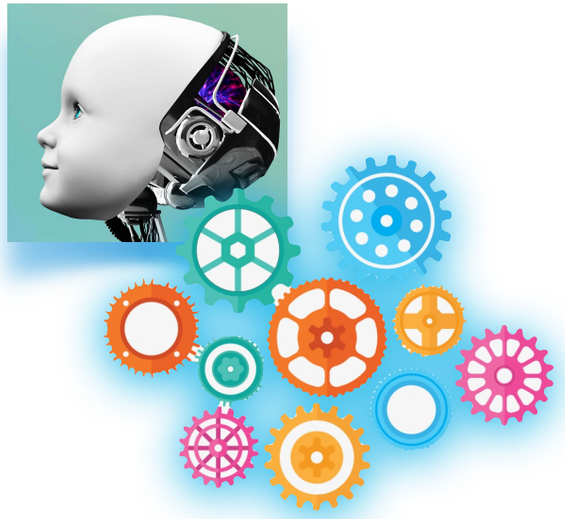



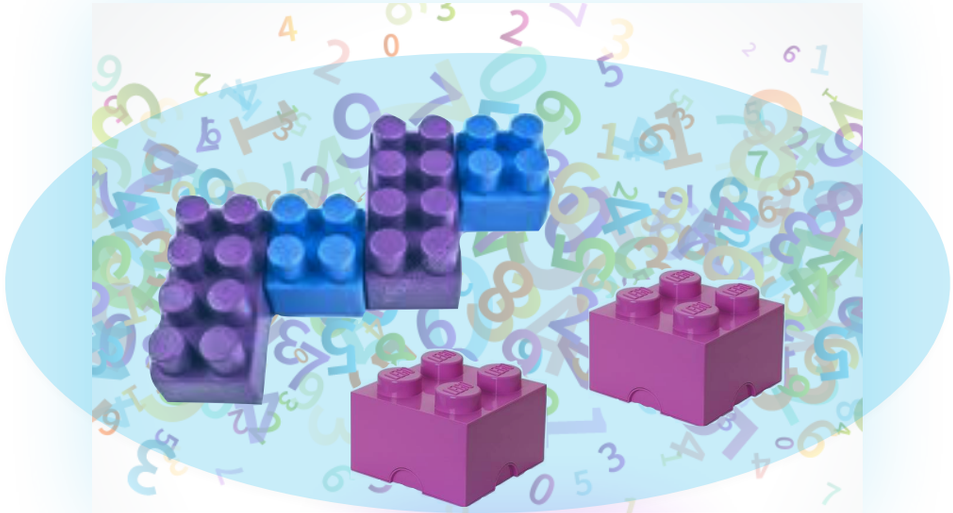



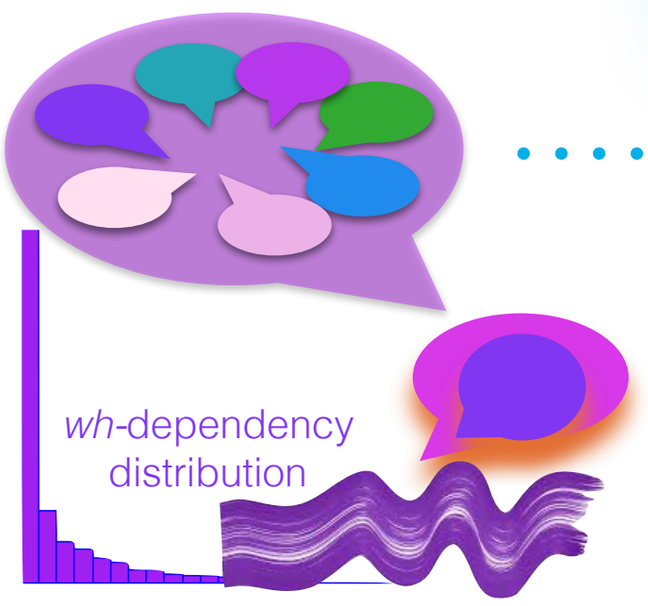

Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

*Dickson et al. 2024*

# Evaluating the theory





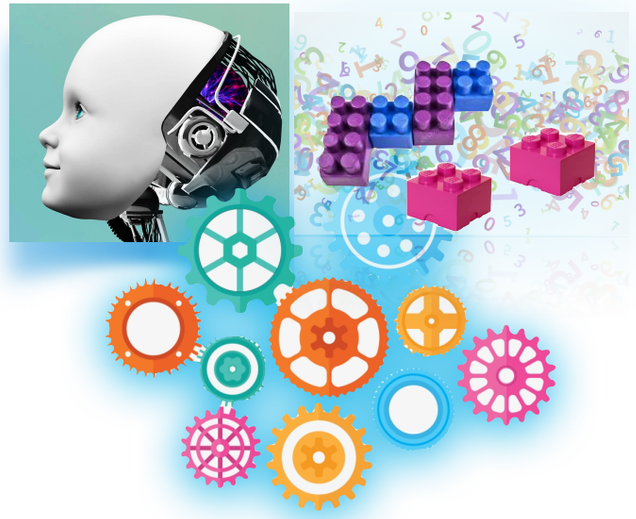



≈ 2.15 million
*wh*-dependencies


Can the modeled learner produce the appropriate observable behavior?

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

# Evaluating the theory



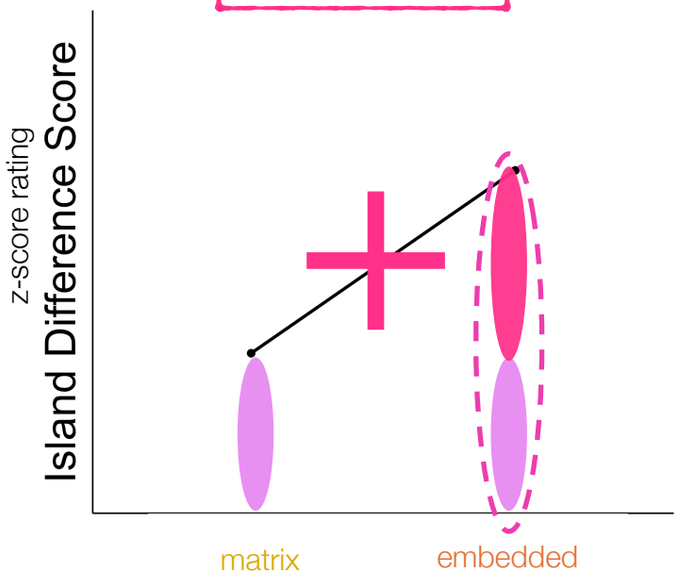
## Reminder: Target behavior

*Subject island*

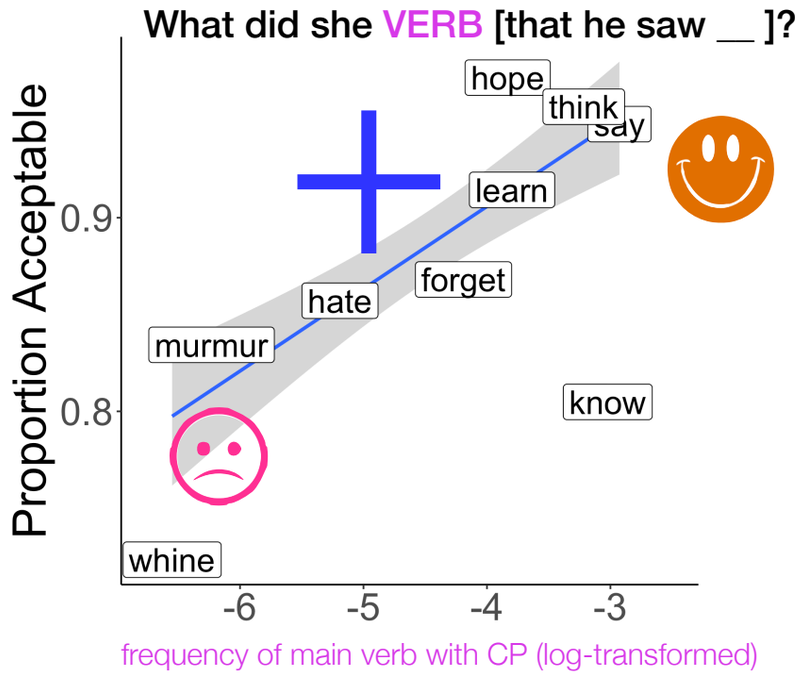
*Complex NP island*

*Whether island*

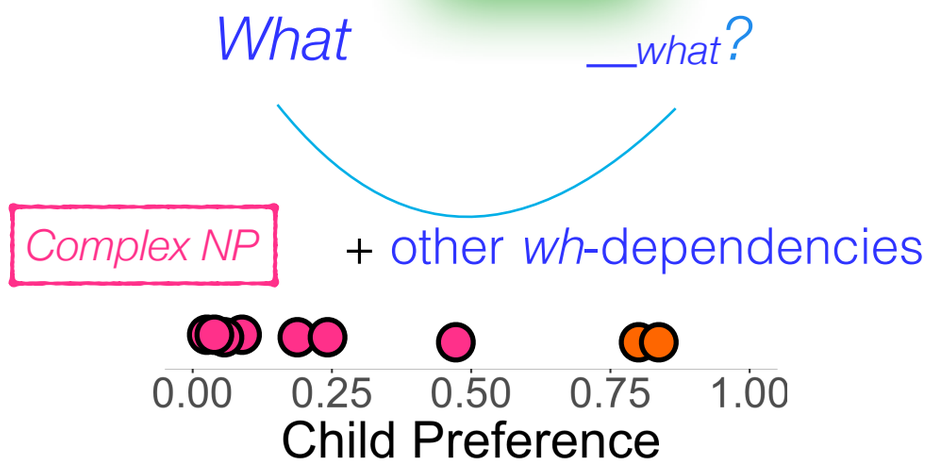
*Adjunct island*




*Sprouse et al. 2012*




*Liu et al. 2019, 2022*




*De Villiers et al. 2008*

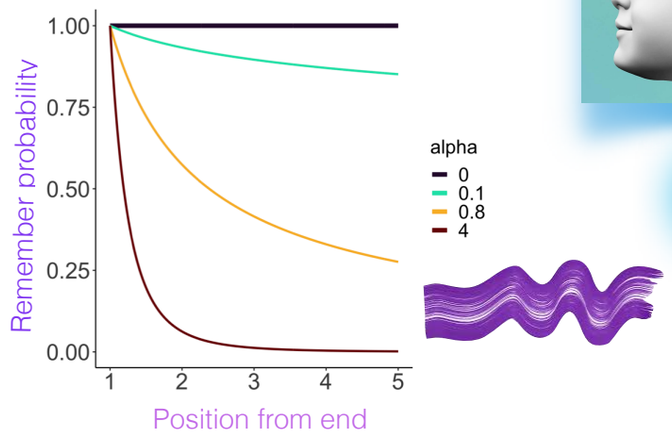
# Evaluating the theory

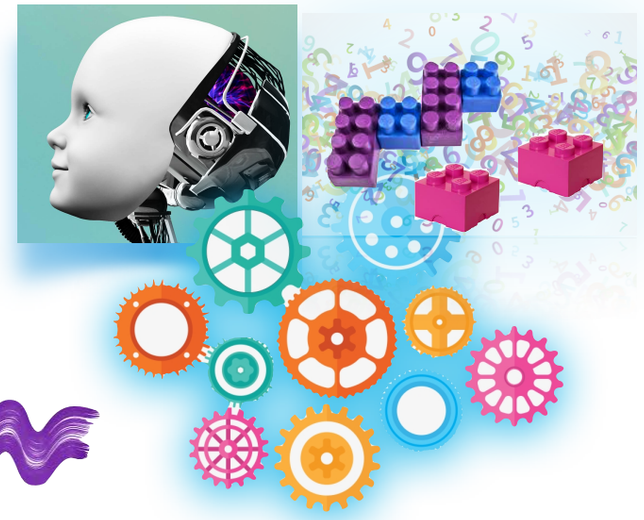
*Subject island*

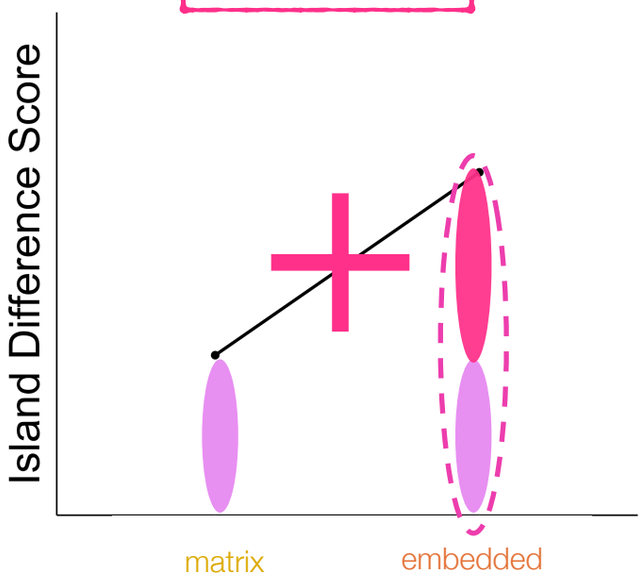
*Complex NP island*

*Whether island*

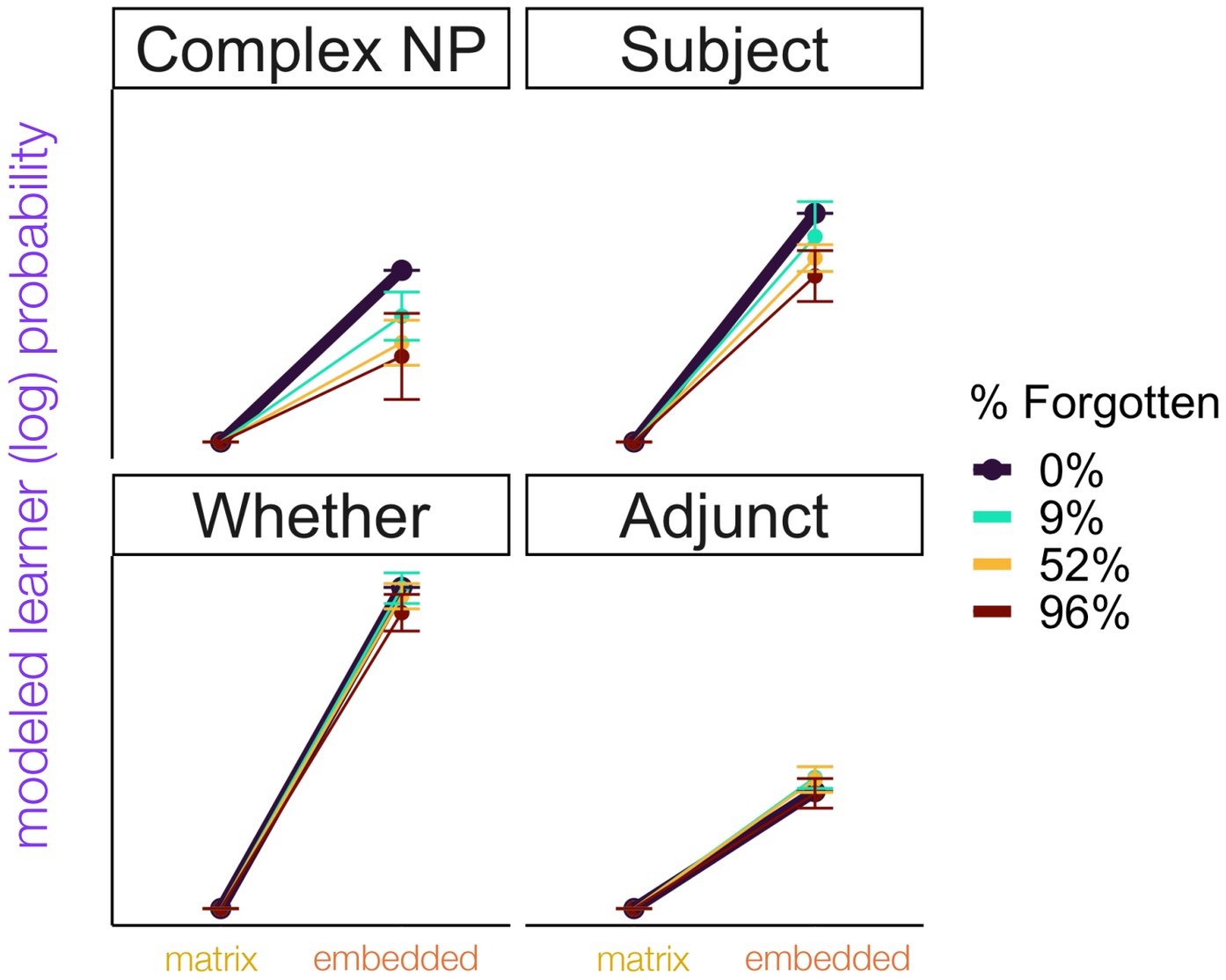
*Adjunct island*






*Sprouse et al. 2012*




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

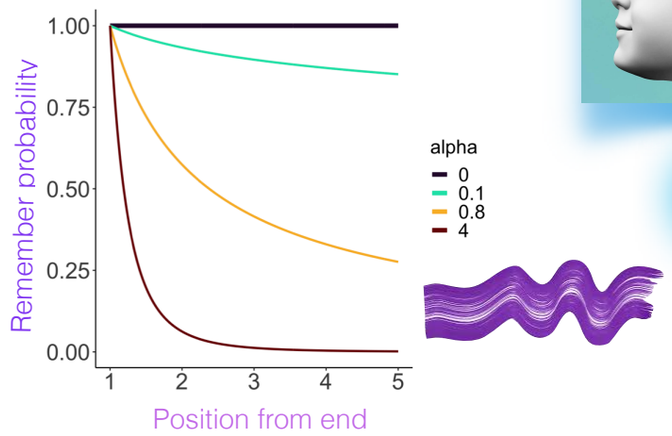
# Evaluating the theory

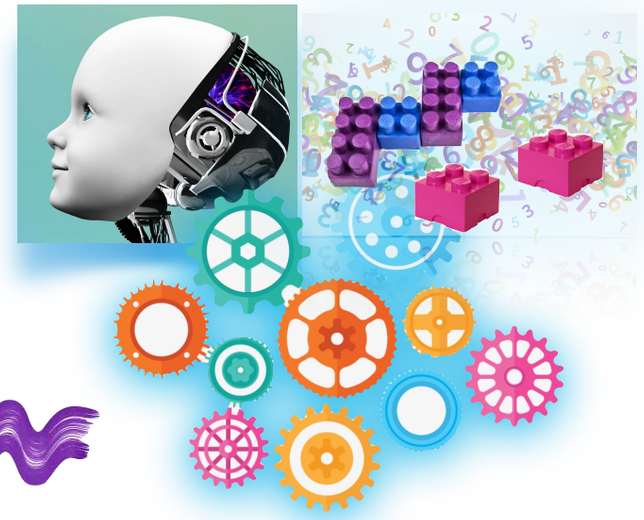
*Subject island*

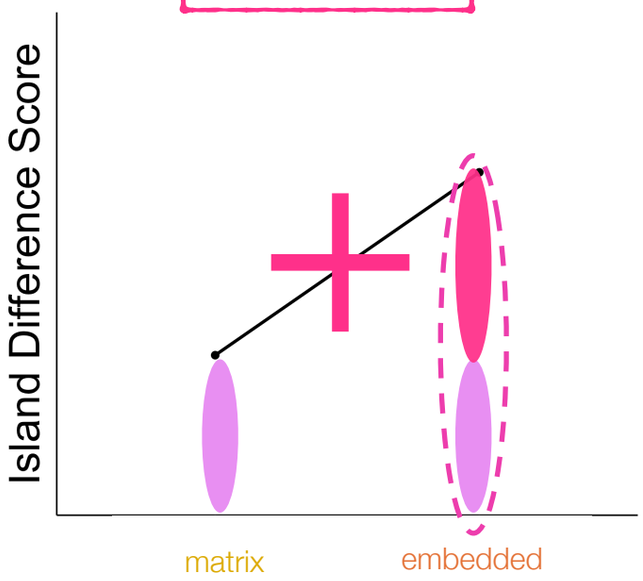
*Complex NP island*

*Whether island*

*Adjunct island*

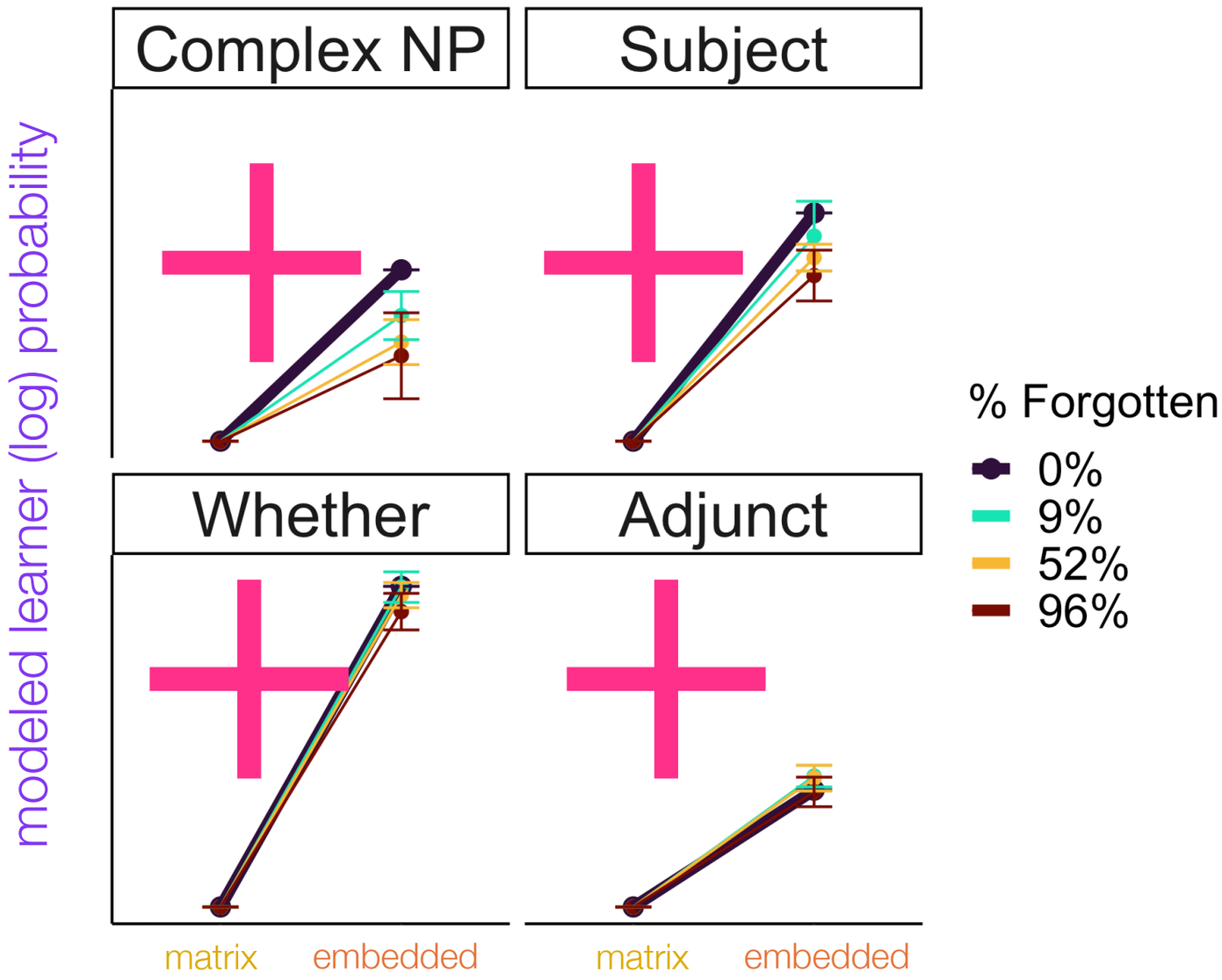





*Sprouse et al. 2012*

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




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

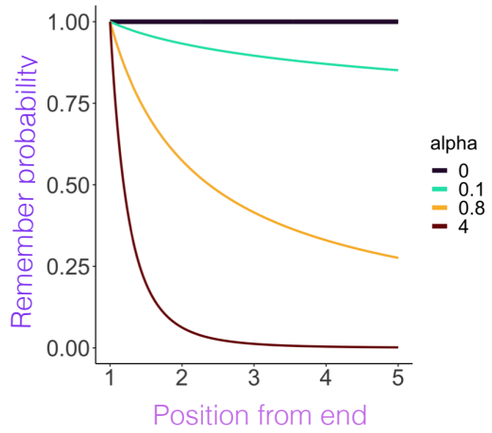
# Evaluating the theory

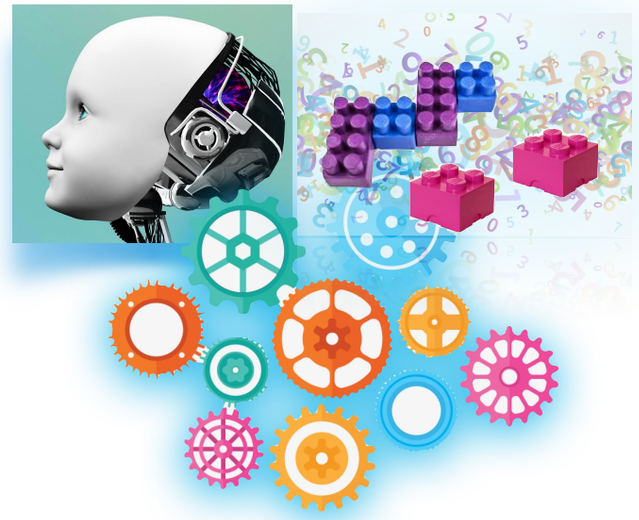
*Subject island*

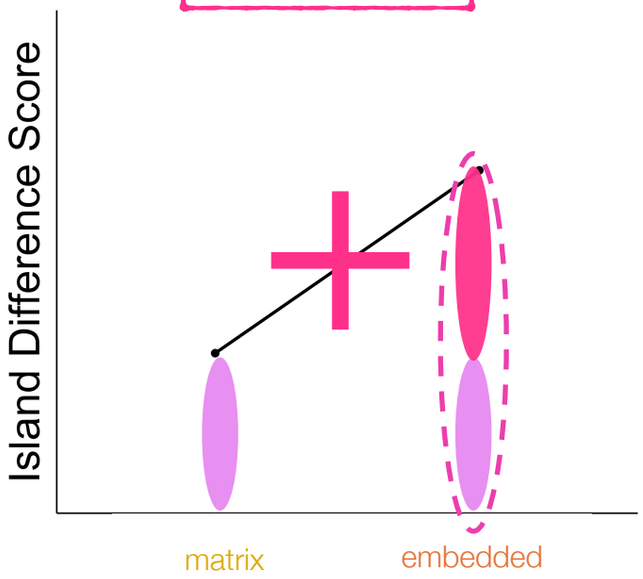
*Complex NP island*

*Whether island*

*Adjunct island*

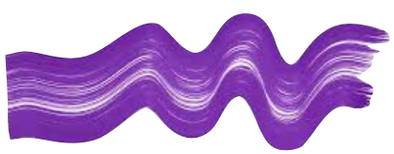


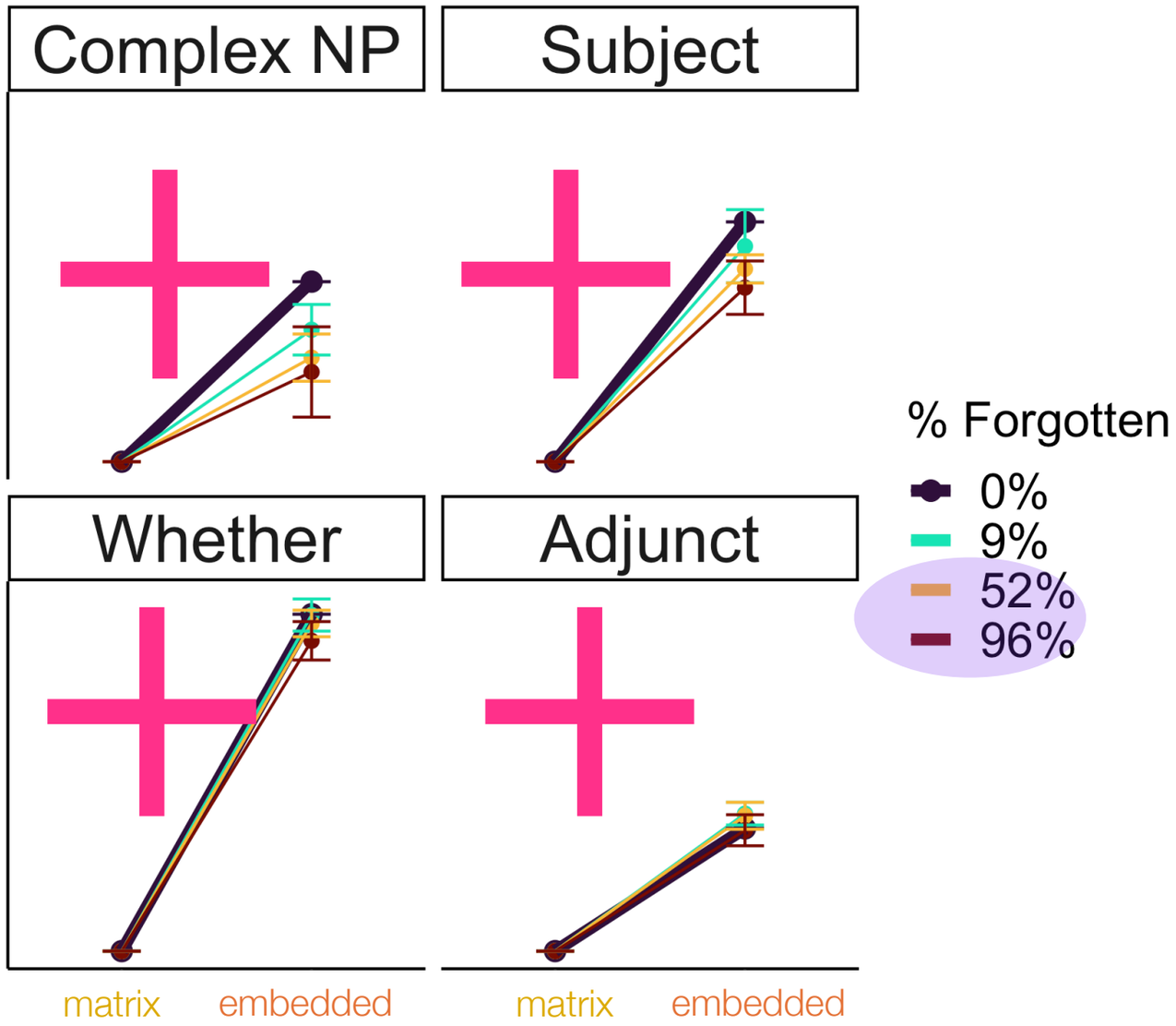






*Sprouse et al. 2012*

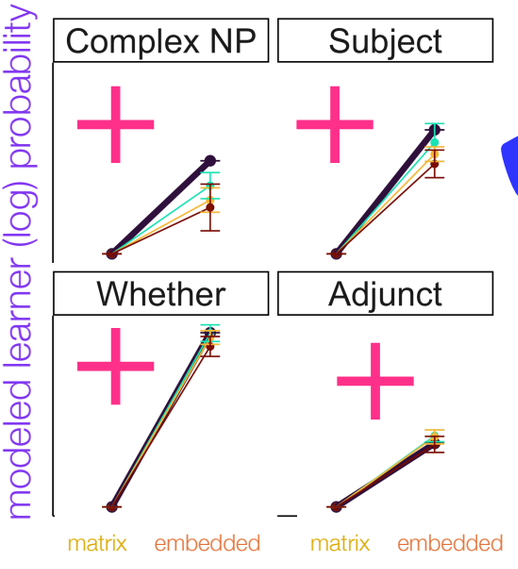
This happens even if they forget a whole lot.

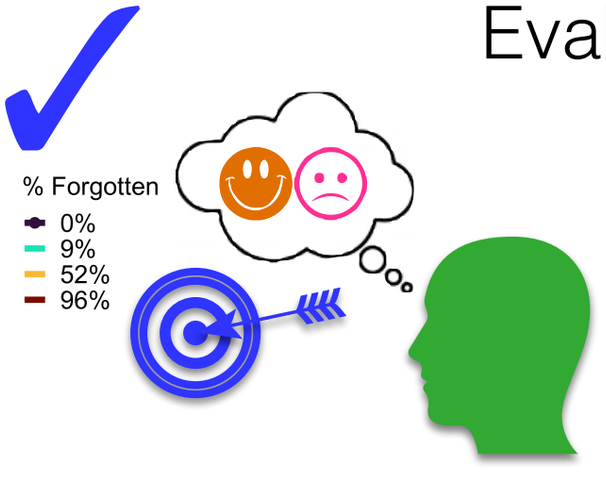


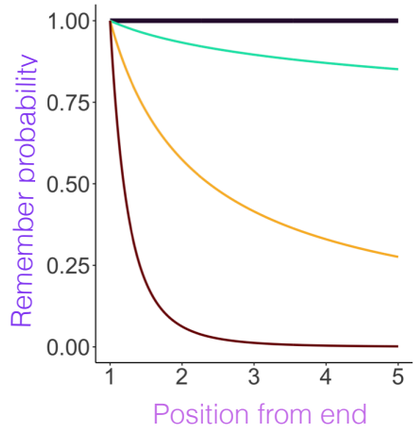


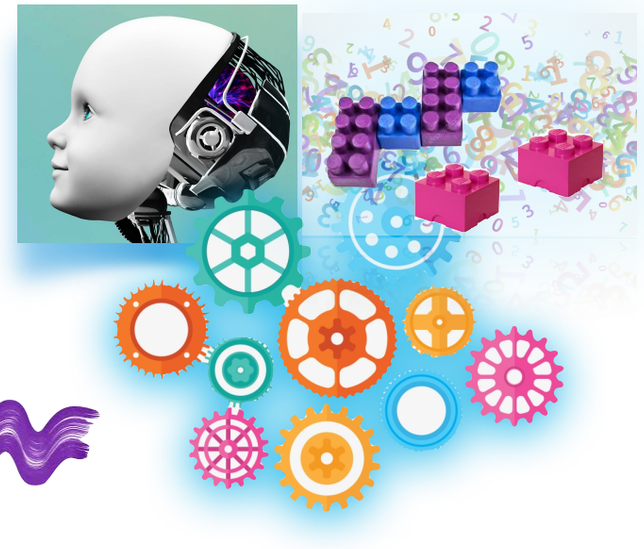

*Dickson et al. 2022, 2024, in prep.*

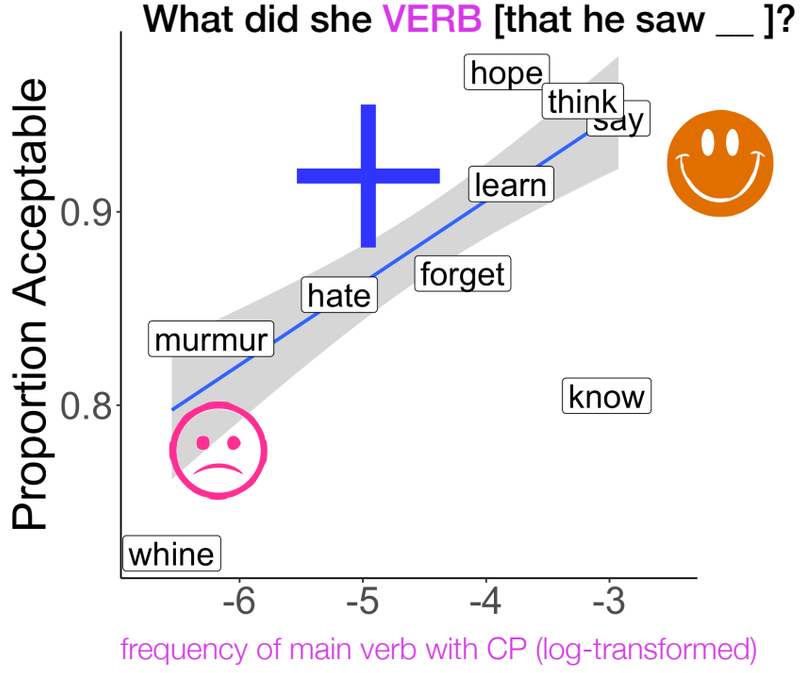
# Evaluating the theory



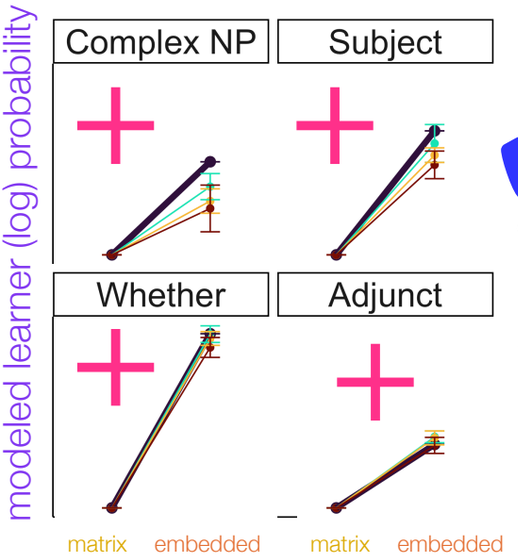



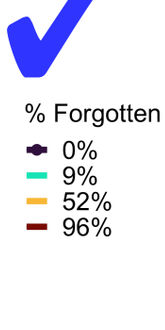


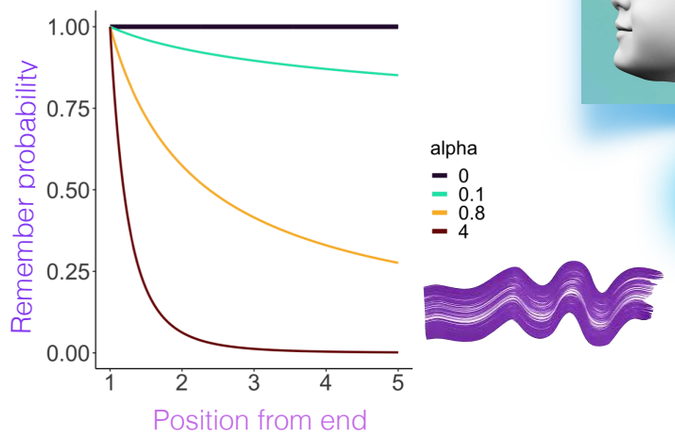

*Liu et al. 2019, 2022*

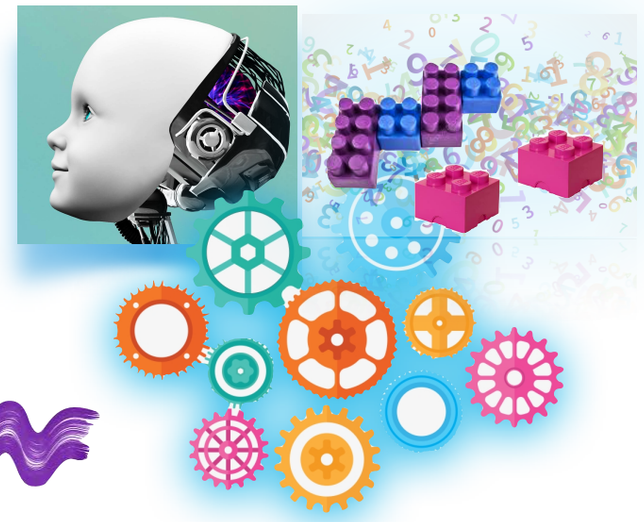
*Dickson et al. 2022, 2024, in prep.*

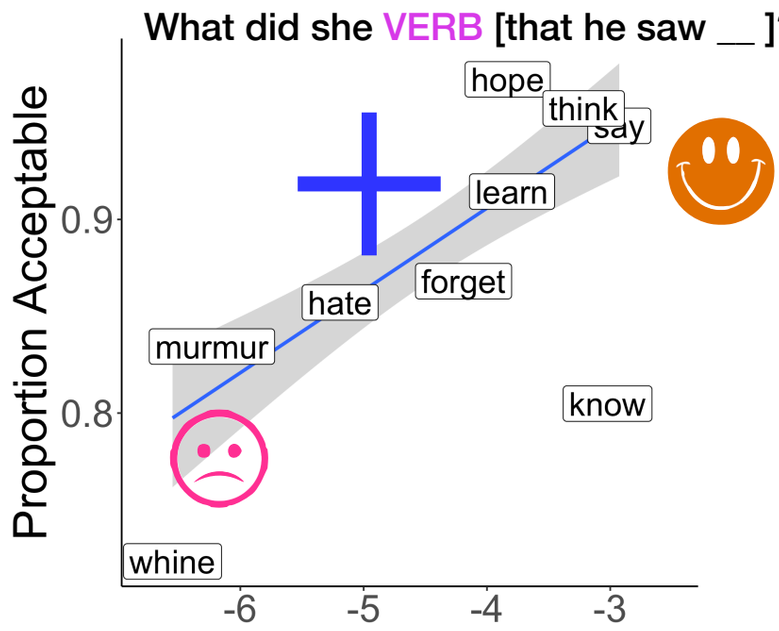
# Evaluating the theory



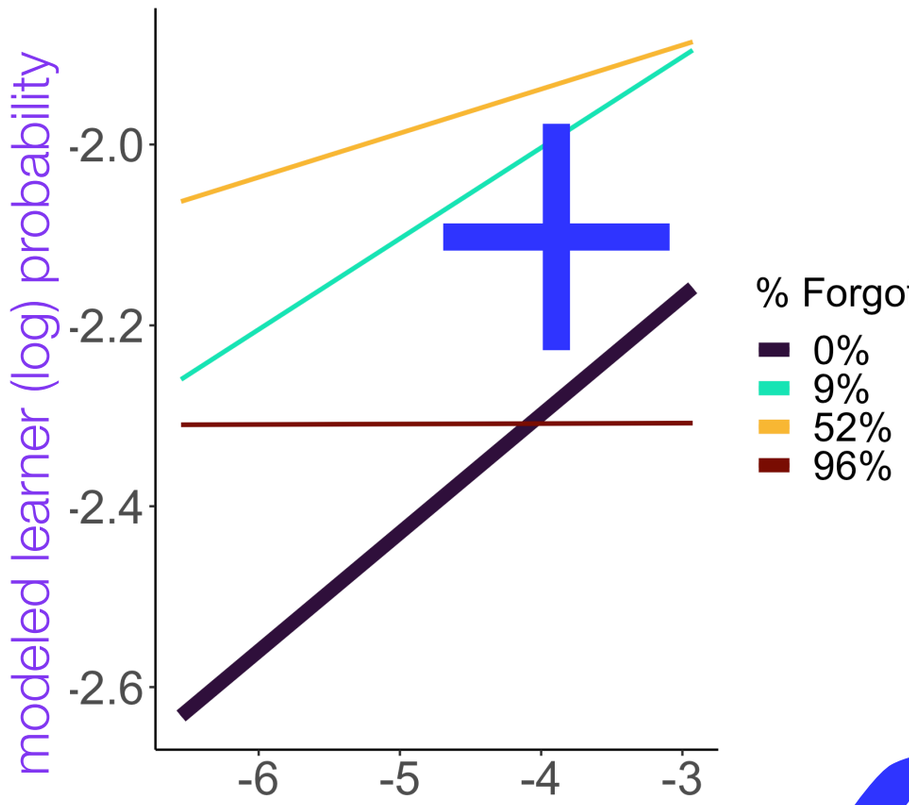





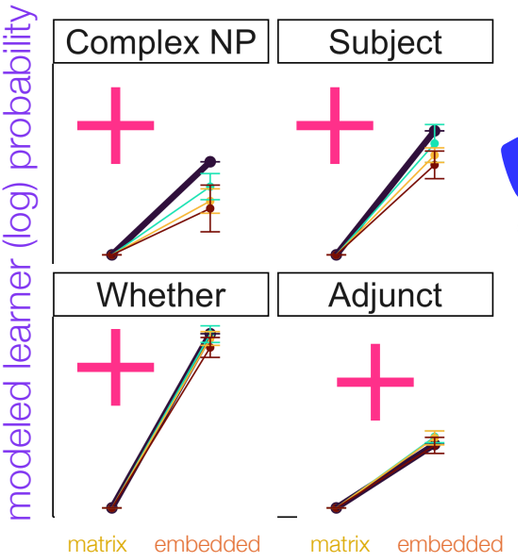

*Liu et al. 2019, 2022*

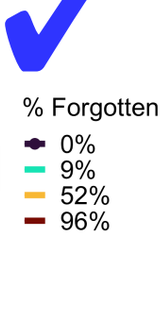


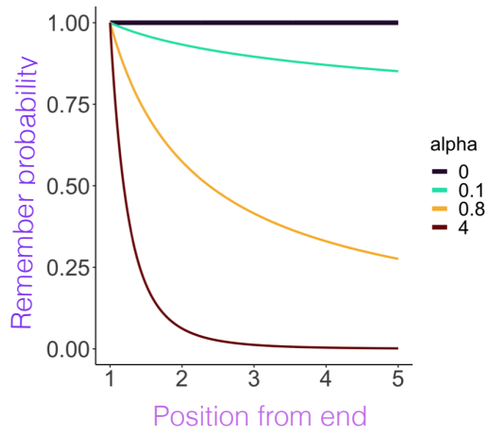

Most modeled learners can predict the observed pattern (positive correlation with frequency).

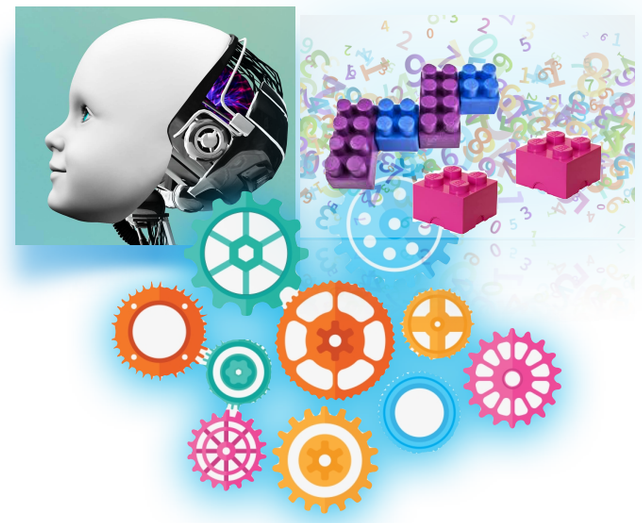
*Dickson et al. 2022, 2024, in prep.*

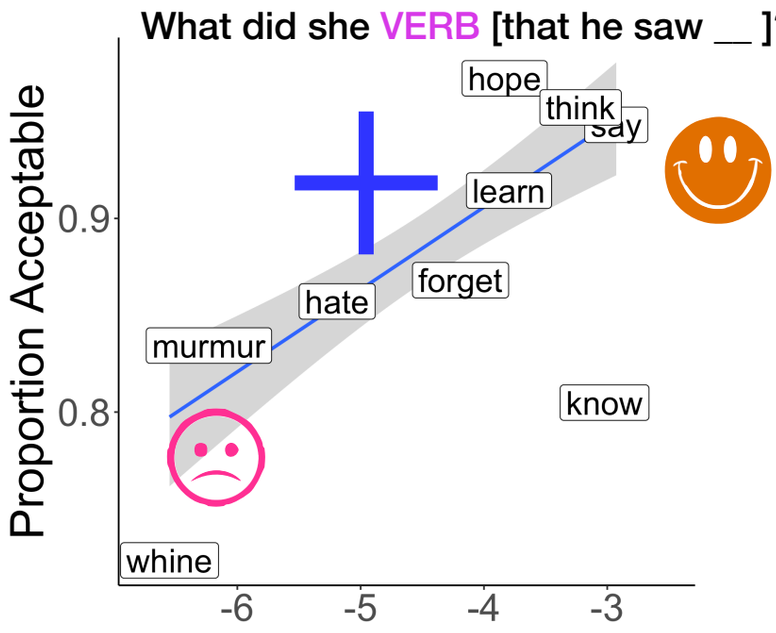
# Evaluating the theory

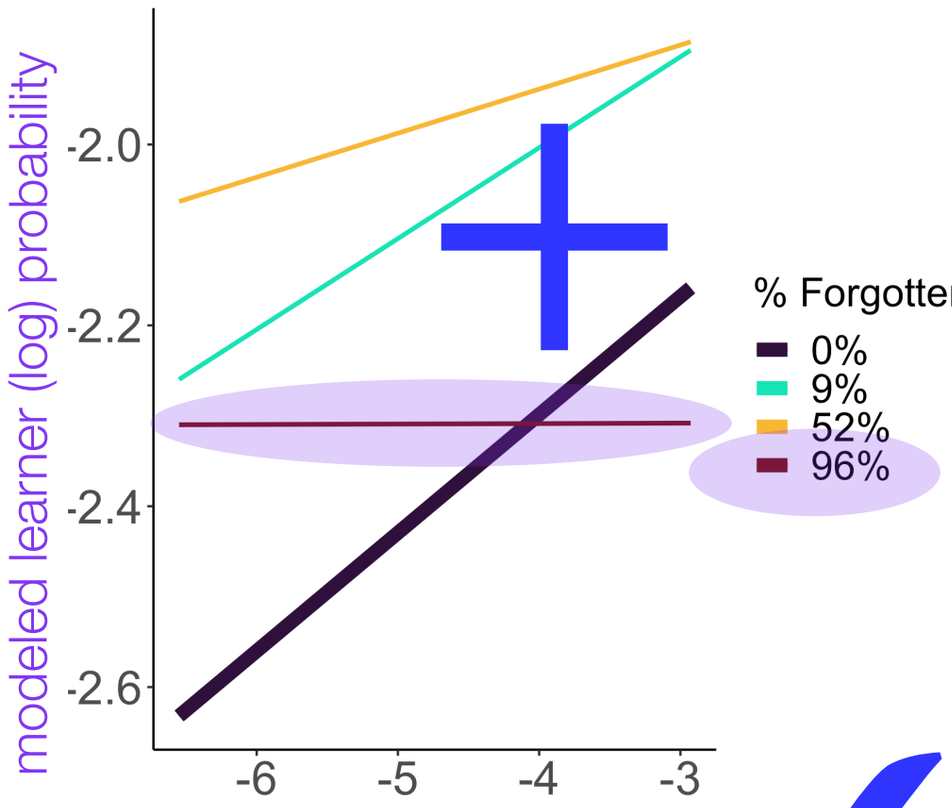


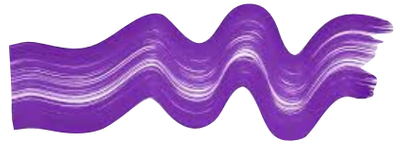





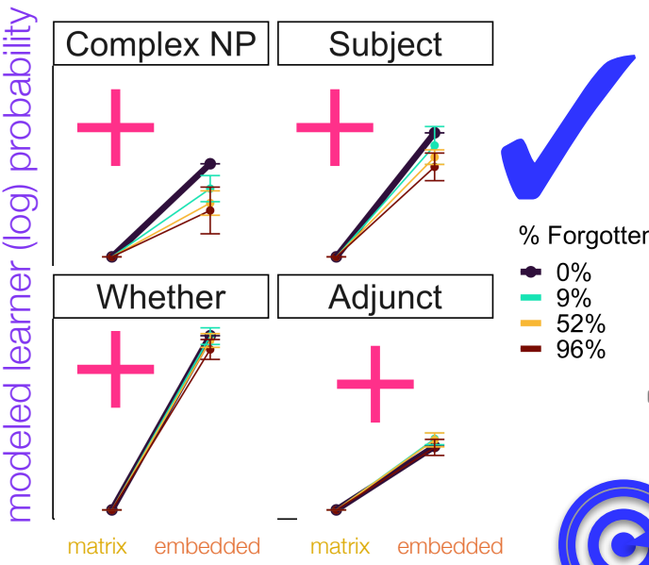

*Liu et al. 2019, 2022*

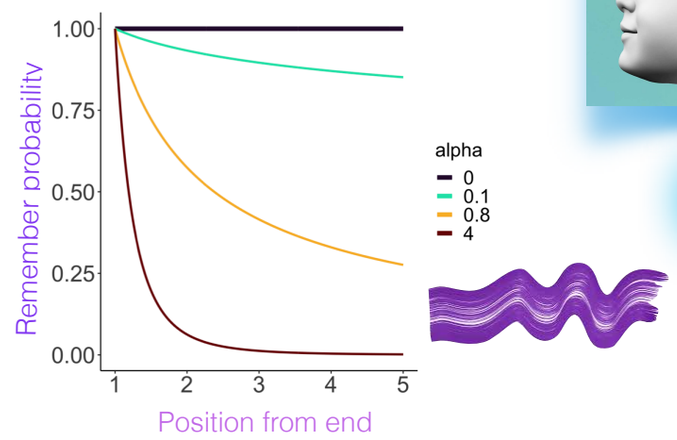


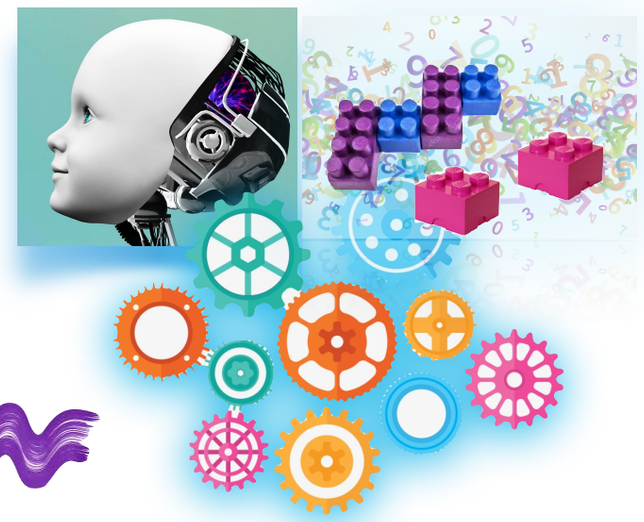
This happens, except when there's an extraordinary amount of forgetting.

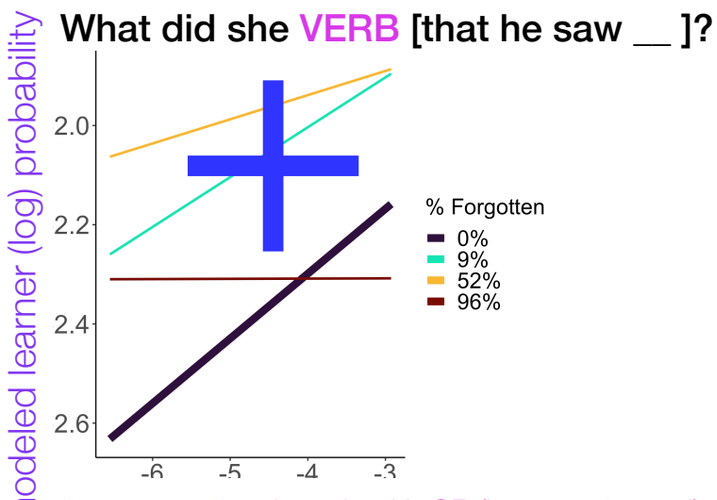
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory



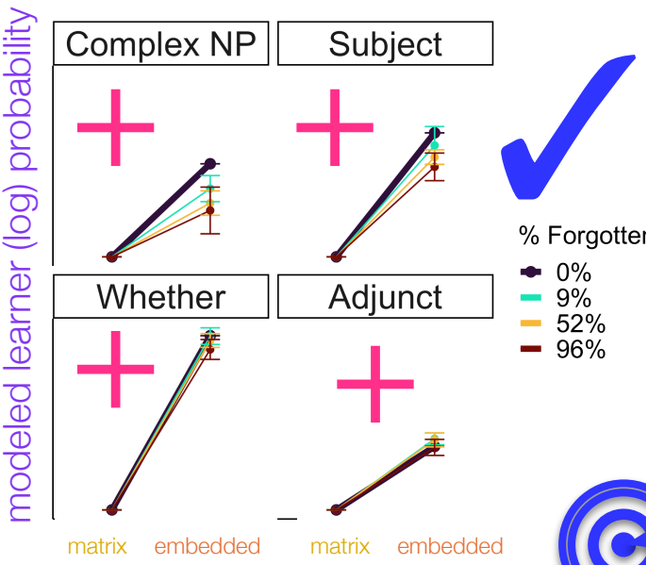





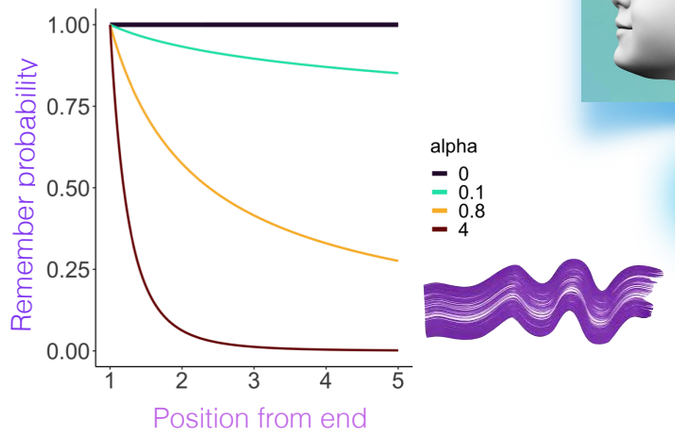

*Complex NP* + other *wh*-dependencies

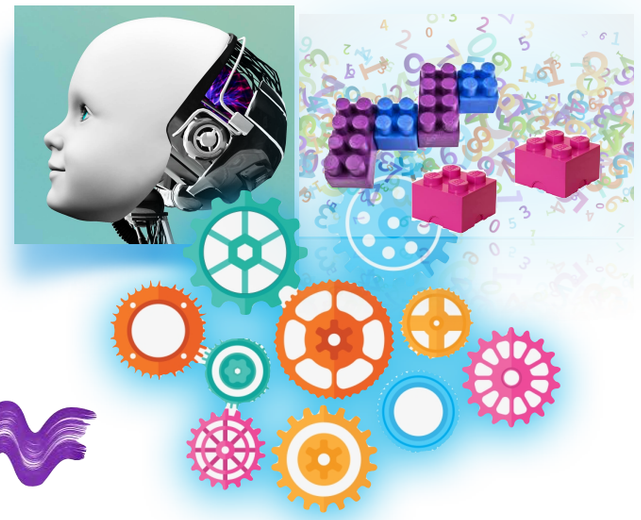

*De Villiers et al. 2008*

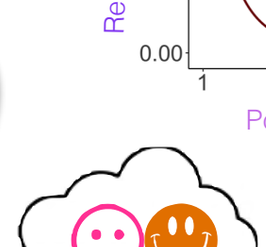
*Dickson et al. 2022, 2024, in prep.*

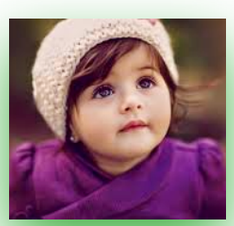


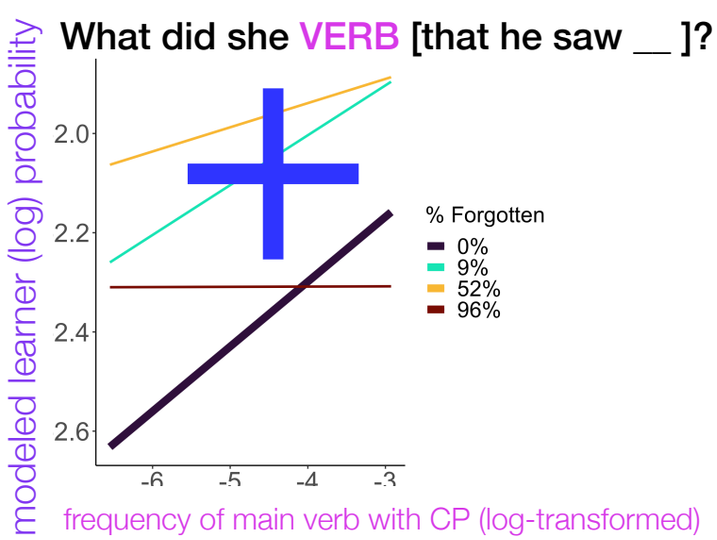

# Evaluating the theory



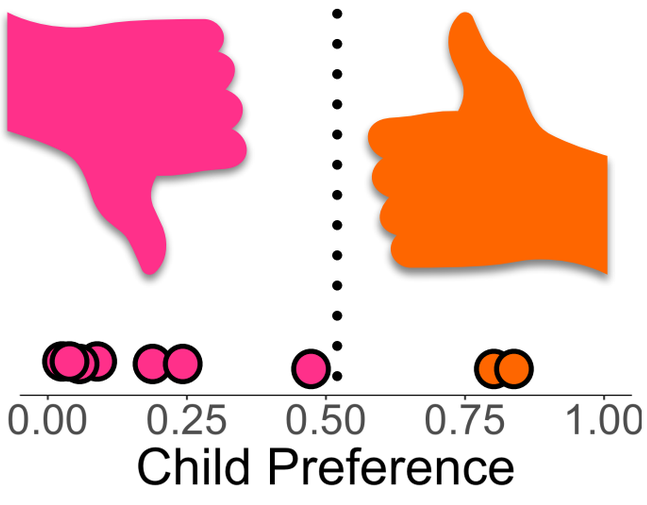










*Complex NP* + other *wh*-dependencies

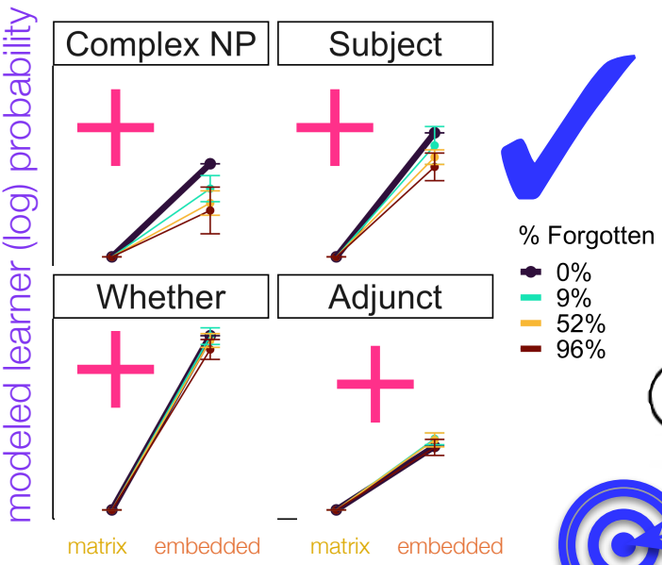


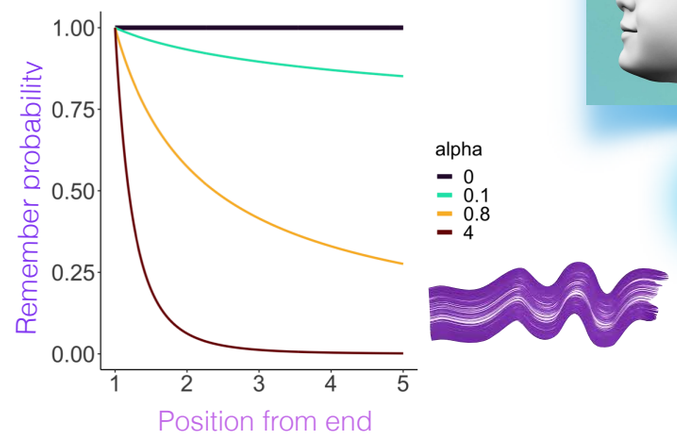

One way to think about these is as qualitative categories.

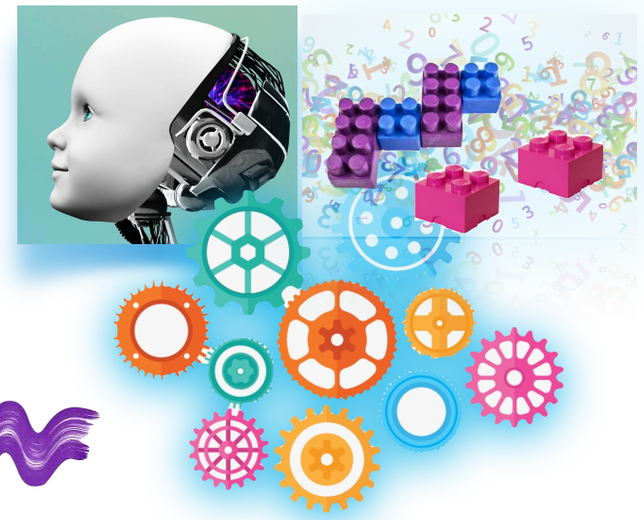
*De Villiers et al. 2008*

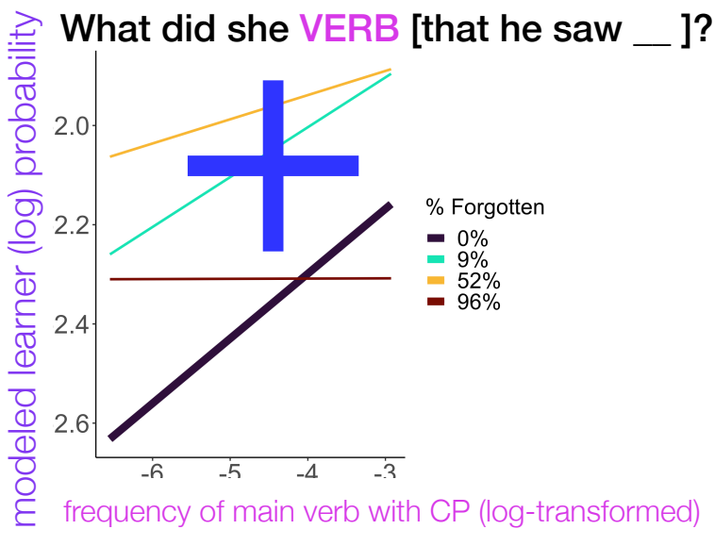
*Dickson et al. 2022, 2024, in prep.*

# Evaluating the theory



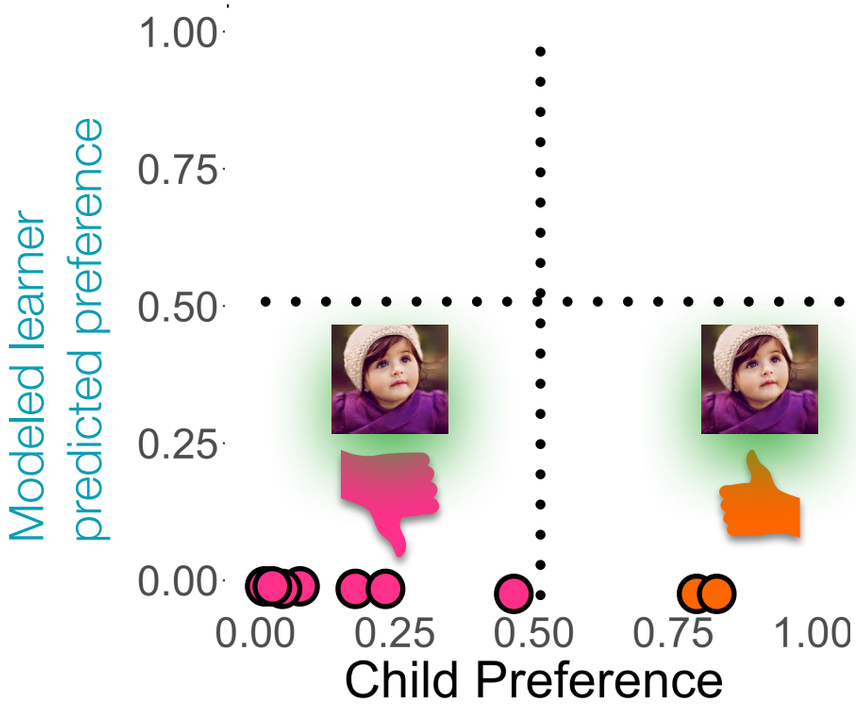








*Complex NP* + other *wh*-dependencies

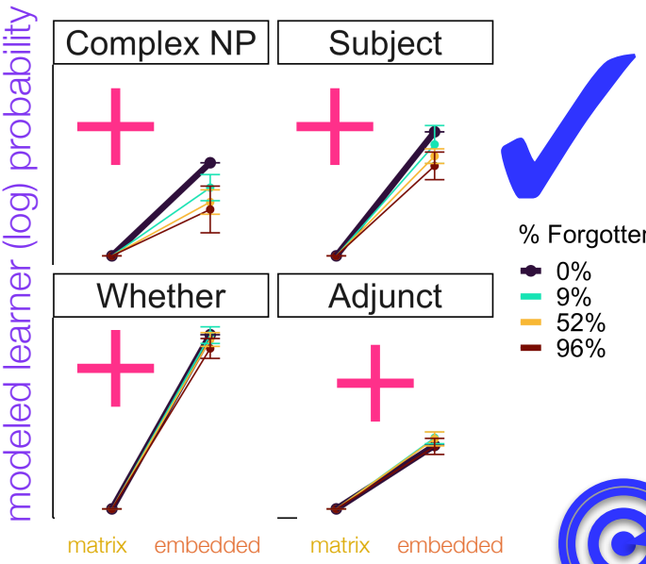



*De Villiers et al. 2008*

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

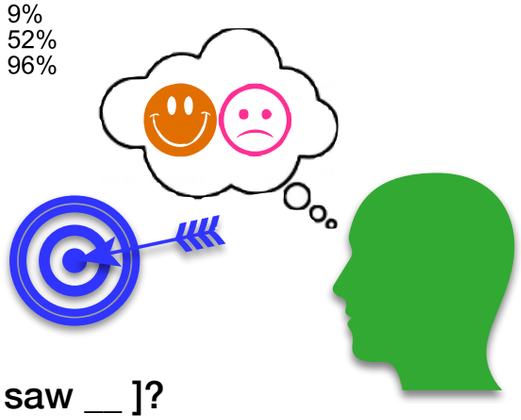
*Dickson et al. 2022, 2024, in prep.*

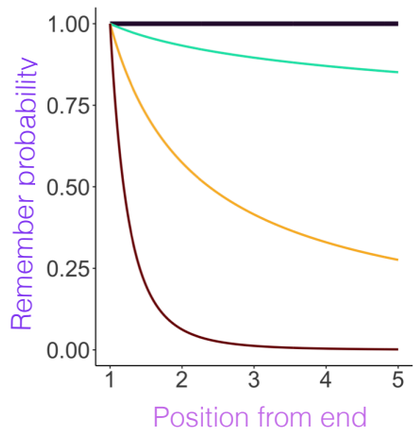
# Evaluating the theory

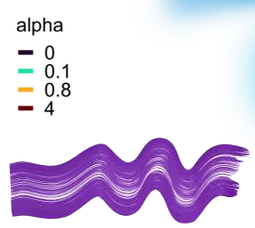


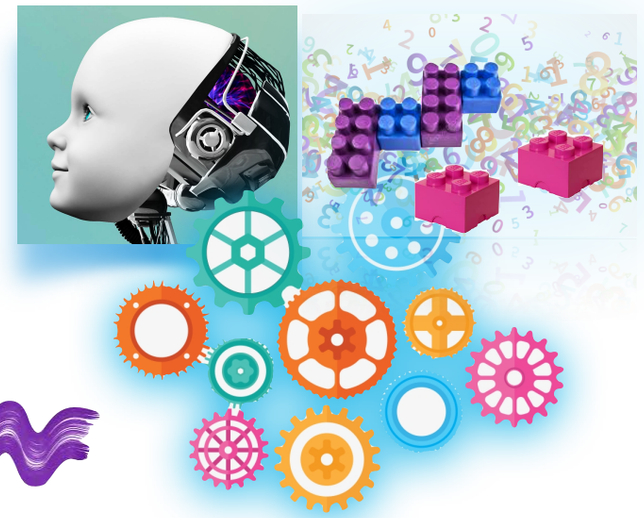


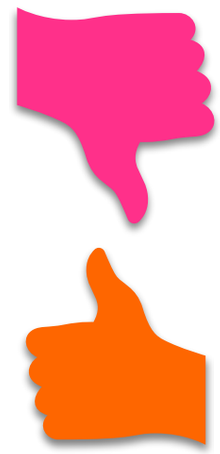




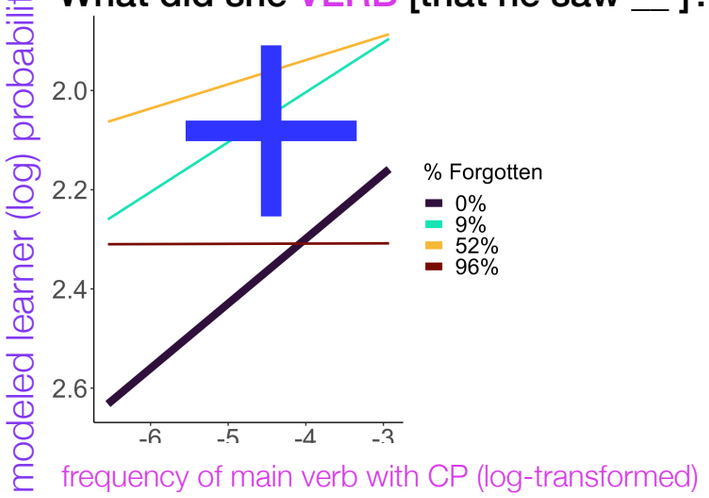


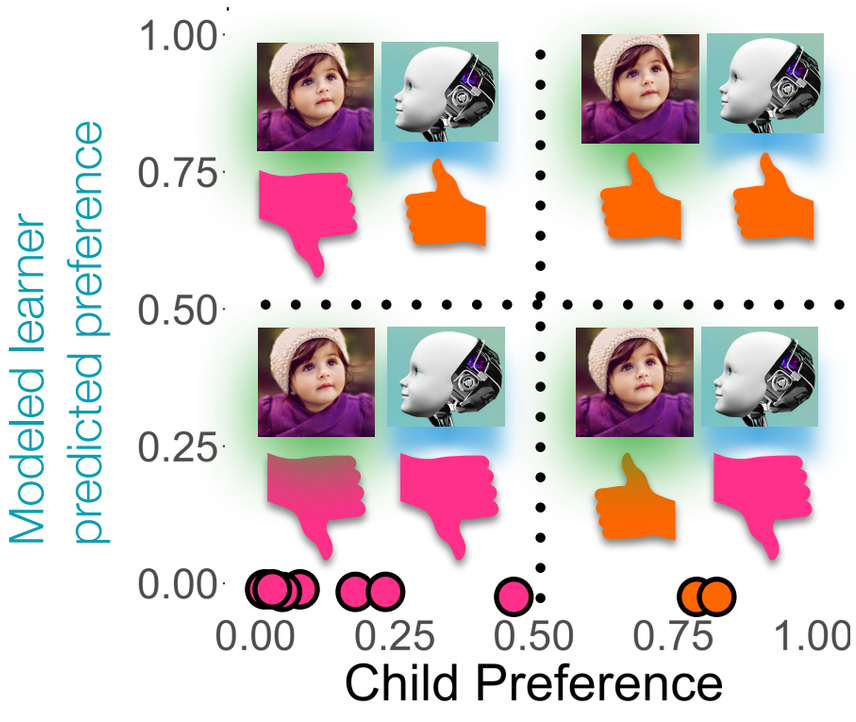





**What did she VERB [that he saw __ ]?**




*Complex NP* + other *wh*-dependencies

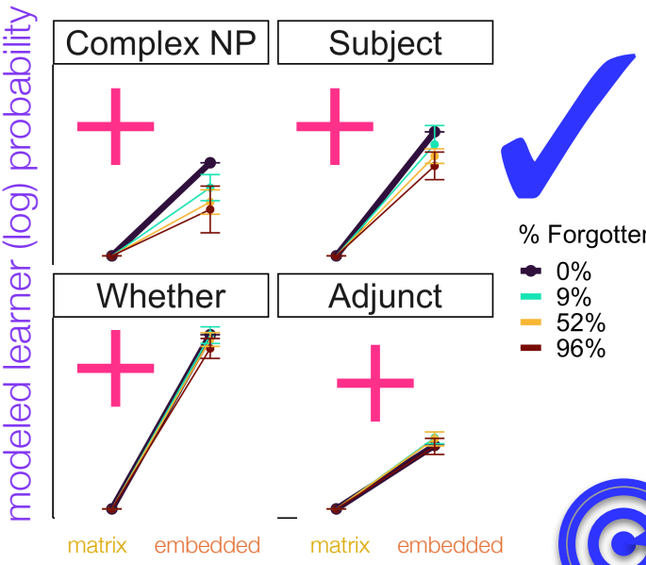


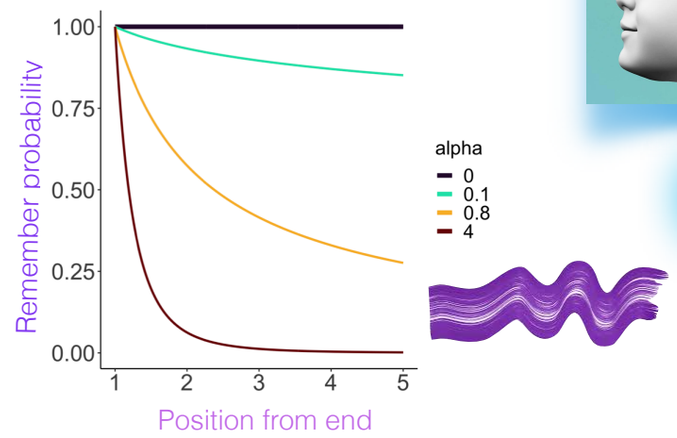

*De Villiers et al. 2008*

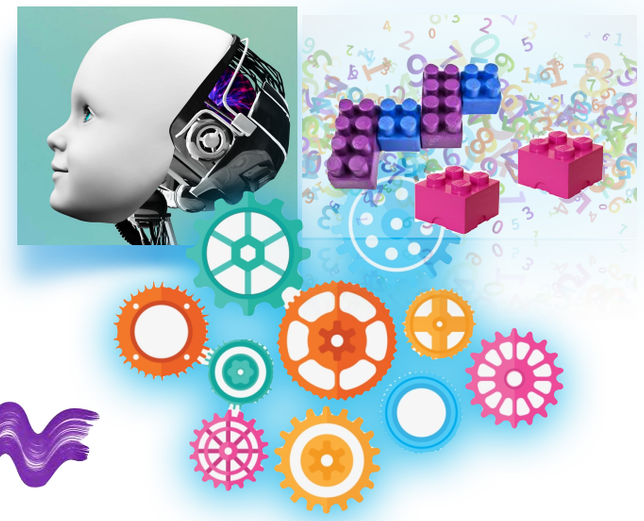
The different quadrants correspond to child vs. modeled learner qualitative predictions.

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

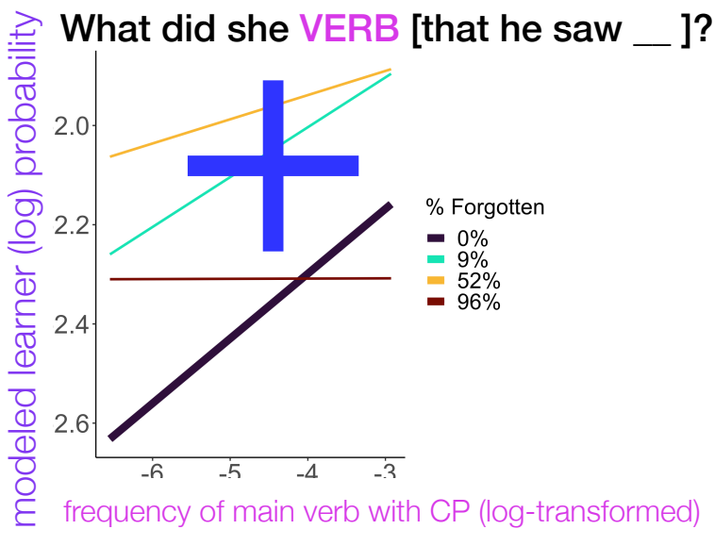
# Evaluating the theory



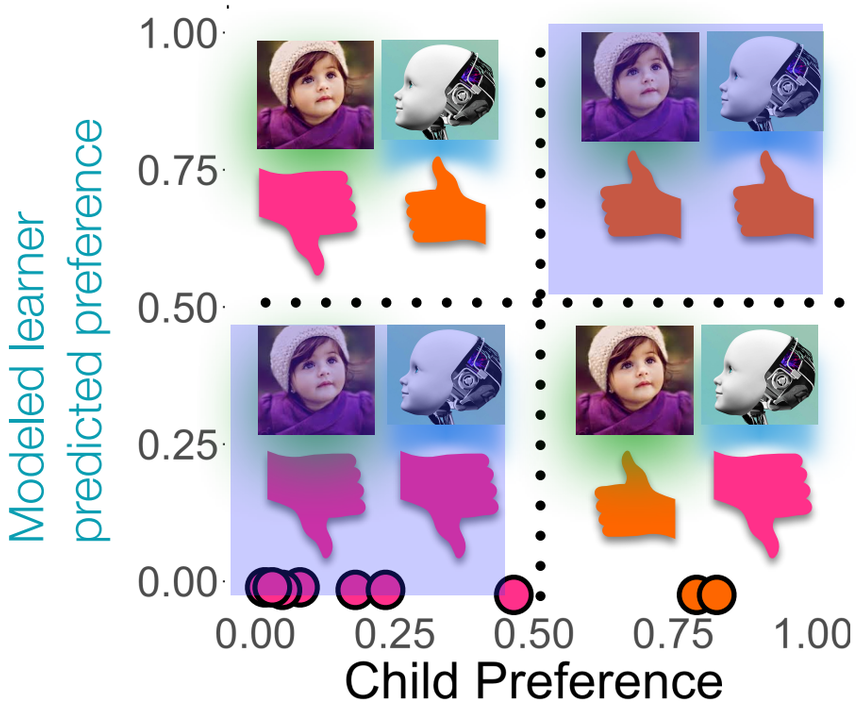








*Complex NP* + other *wh*-dependencies

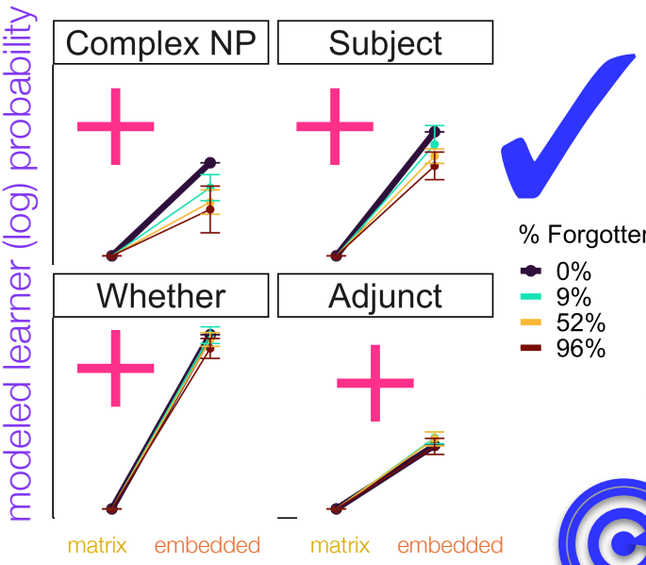


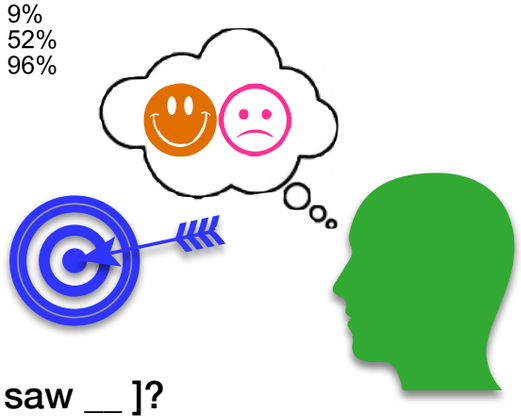

*De Villiers et al. 2008*

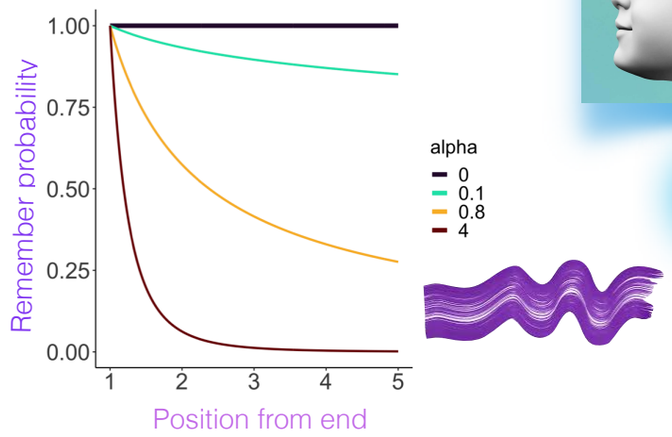
We'd like the modeled learner predictions to qualitatively match the observed child behavior.

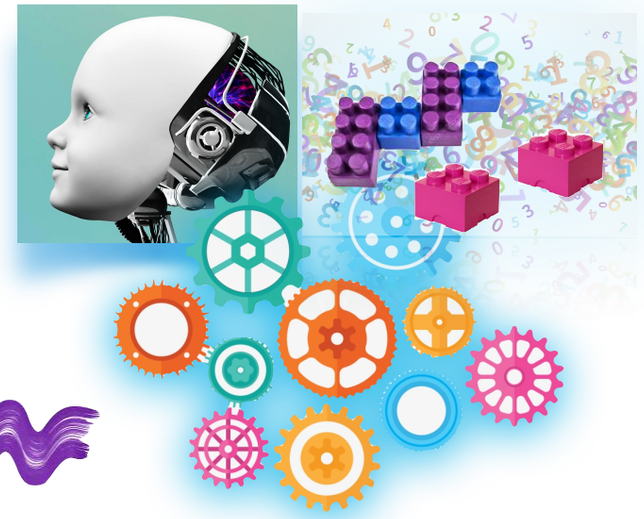
*Dickson et al. 2022, 2024, in prep.*

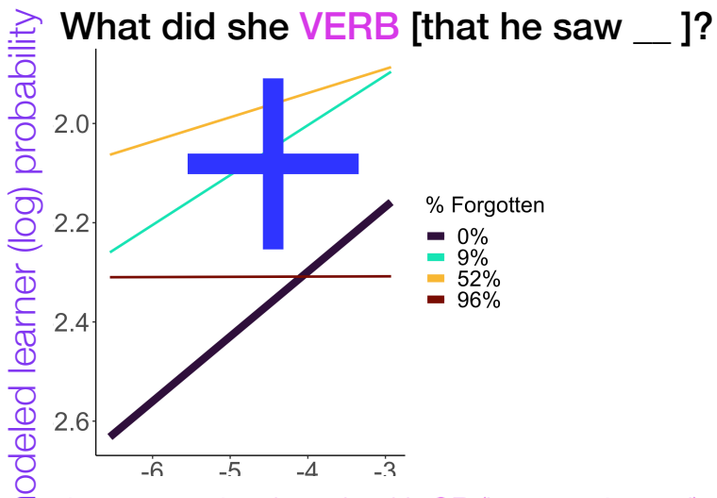
# Evaluating the theory



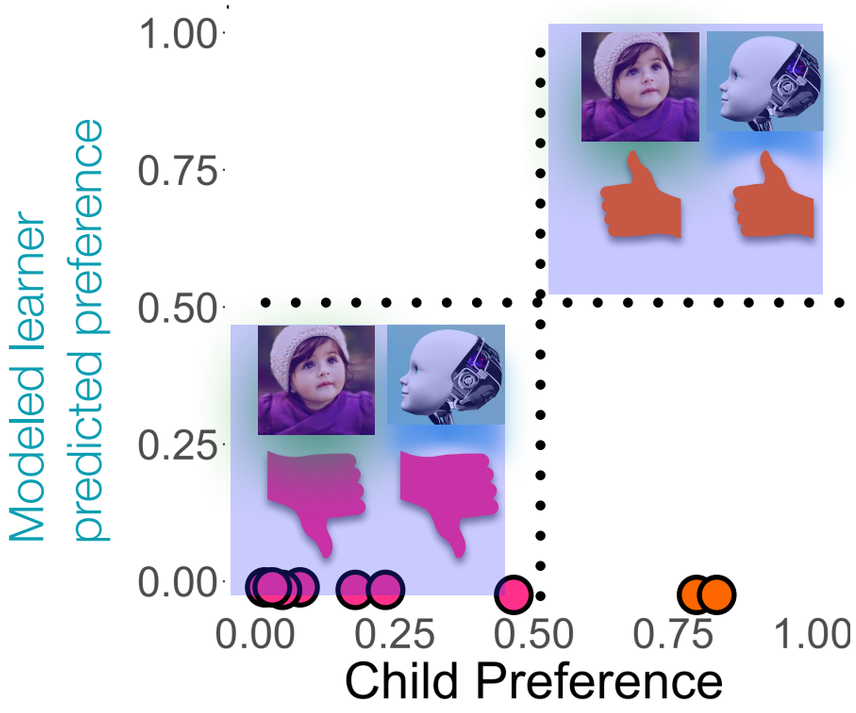





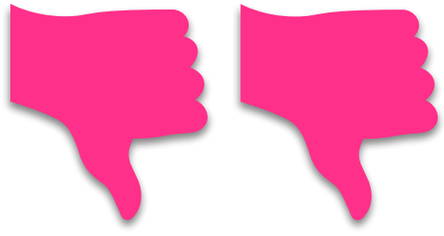



*Complex NP* + other *wh*-dependencies



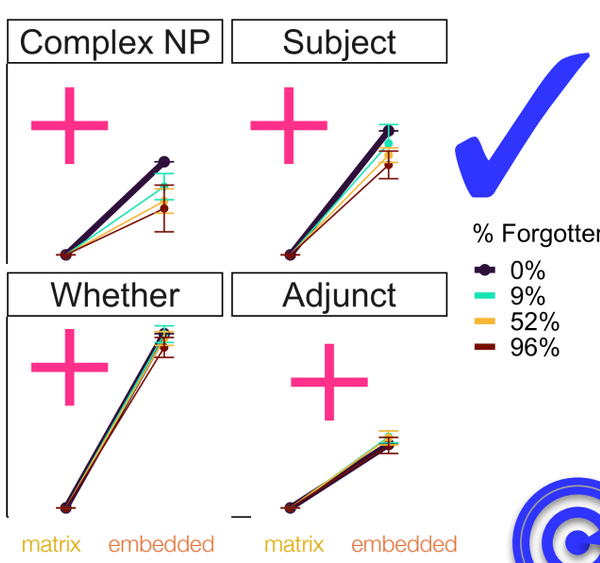

*De Villiers et al. 2008*

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



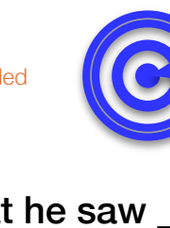
*Dickson et al. 2022, 2024, in prep.*

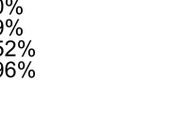
# Evaluating the theory



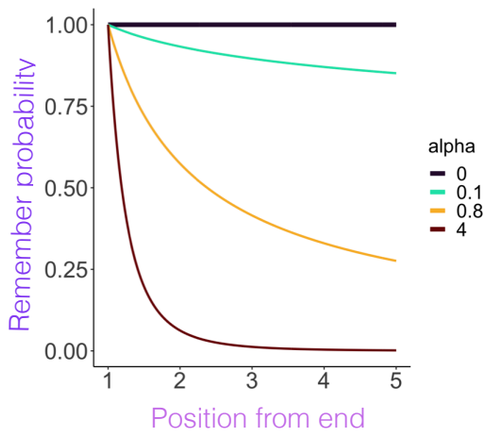


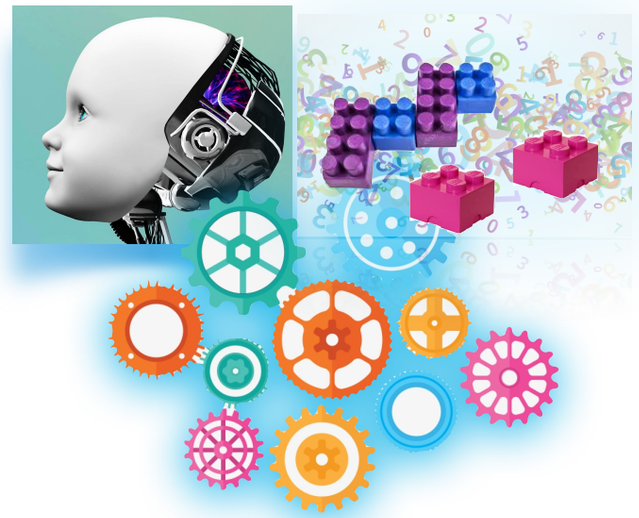


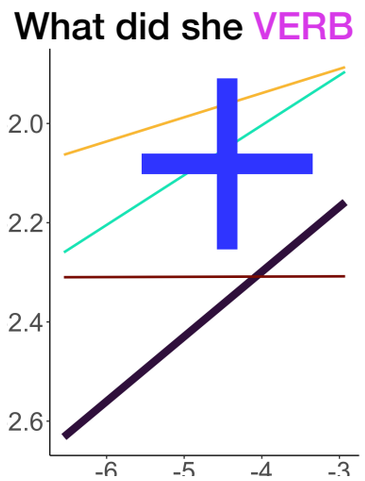





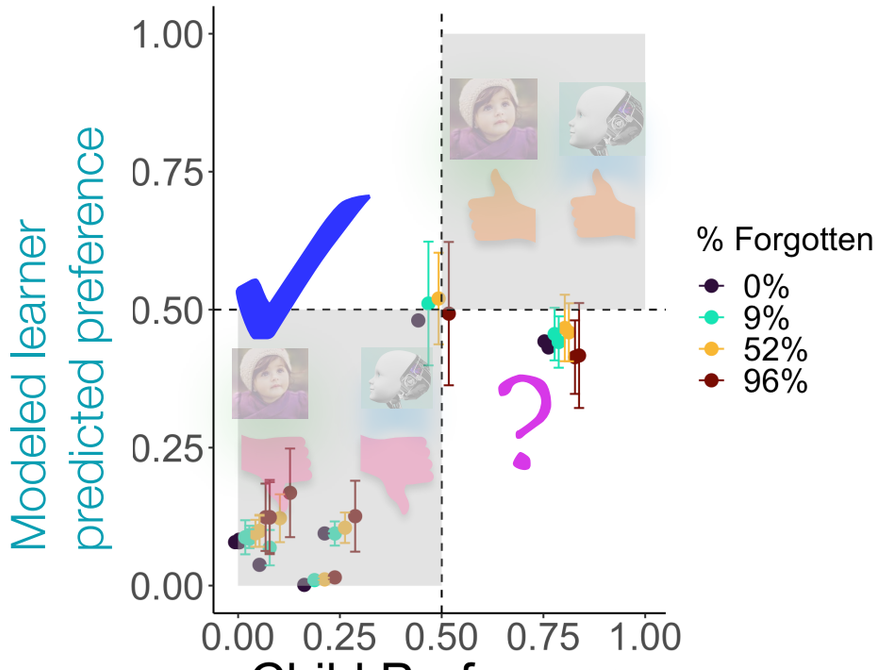
% Forgotten
0%
9%
52%
96%






*Complex NP* + other *wh*-dependencies



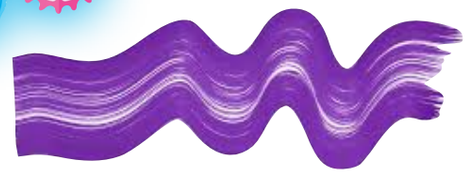

*De Villiers et al. 2008*

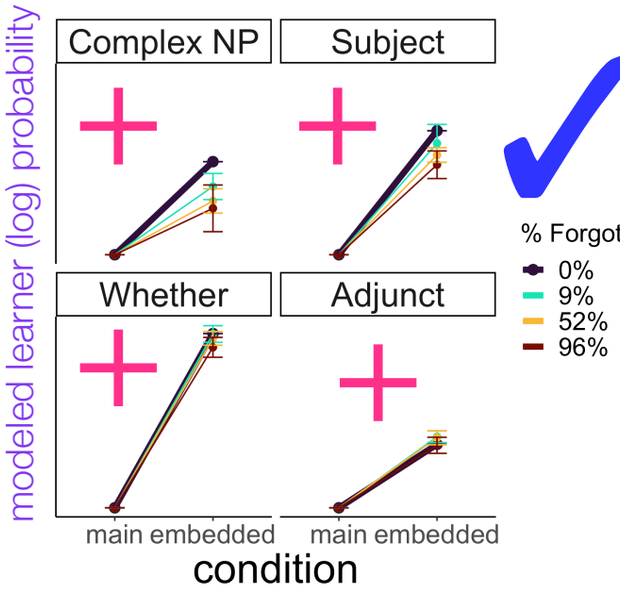
They mostly do, but there are some exceptions.

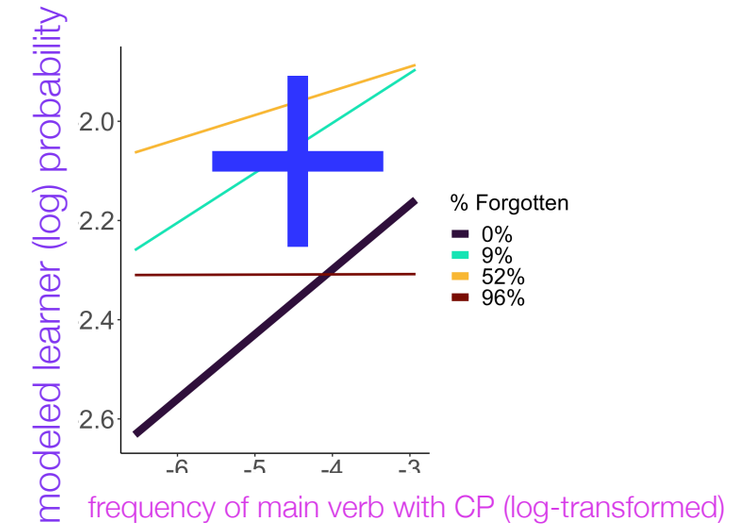
This is true, no matter how much forgetting there is.

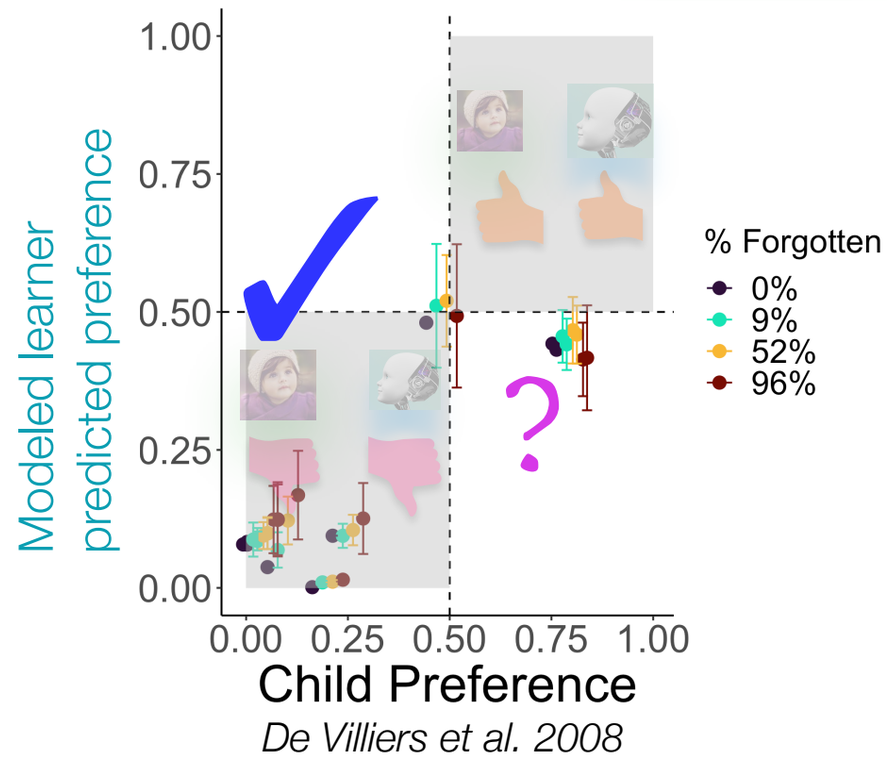
*Dickson et al. 2022, 2024, in prep.*

Takeaway: Modeled learners implementing this learning theory can generate most of the observed target behavior patterns, even with human(-like) memory limitations.





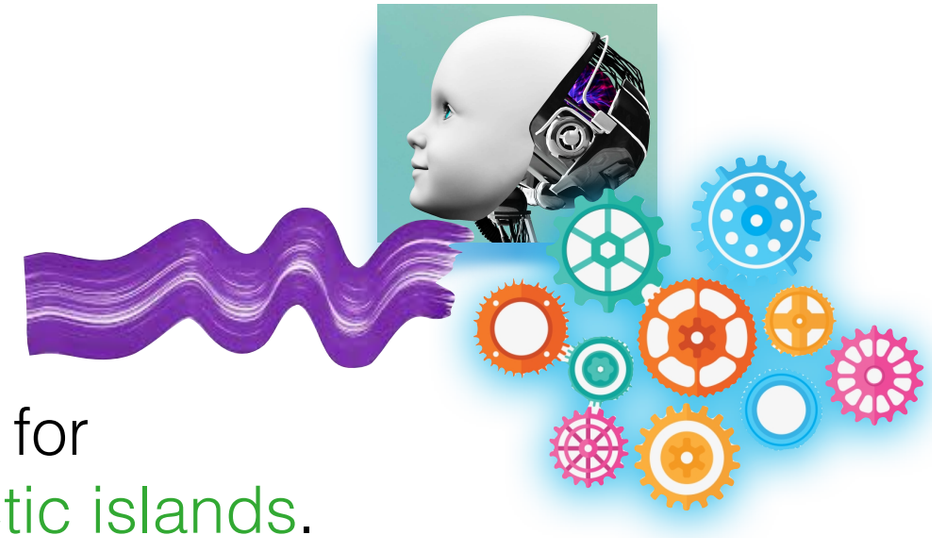



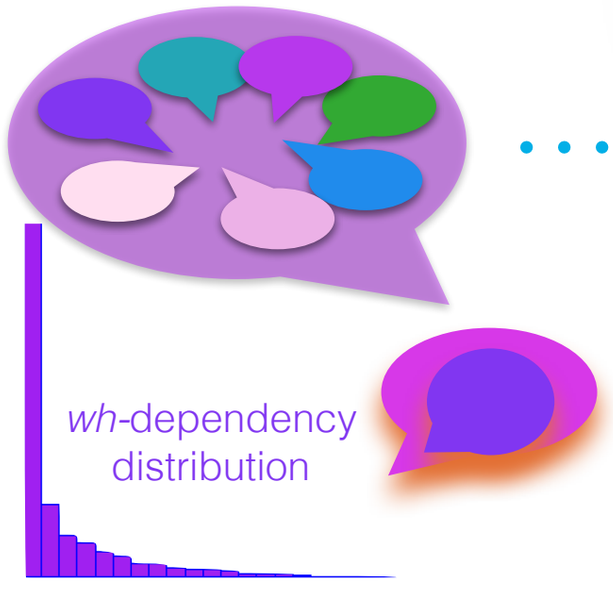

*De Villiers et al. 2008*

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

Learn the right building blocks



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

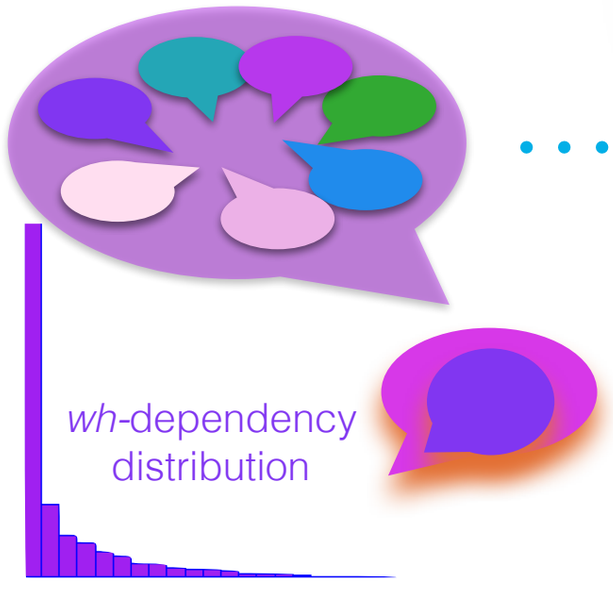




*Dickson et al. 2022, 2024, in prep.*

Learn the right building blocks

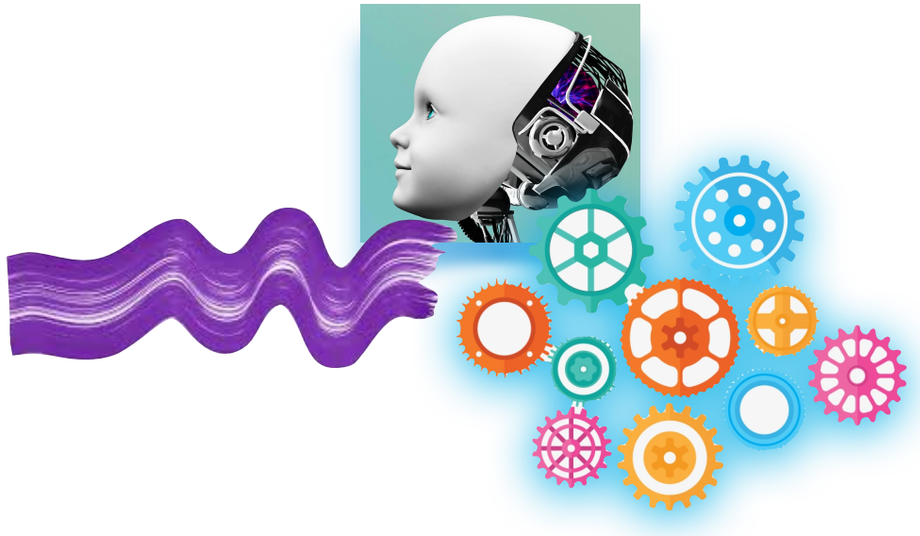
Key idea: Learning about the building blocks of *wh*-dependencies leads to knowledge about syntactic islands.



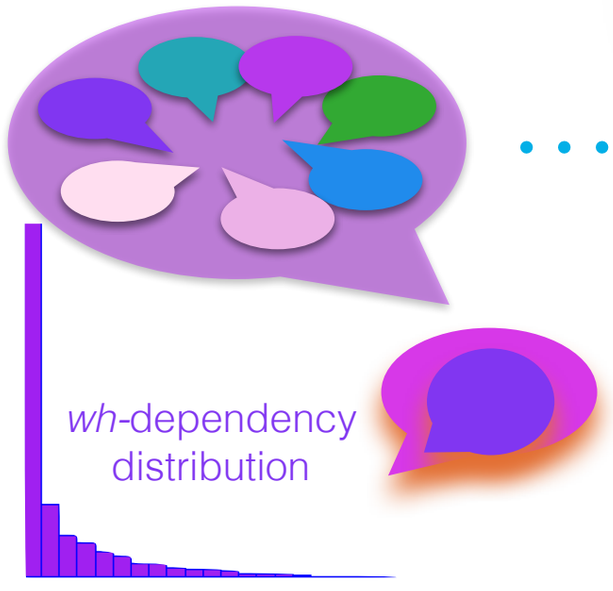


*Dickson et al. 2022, 2024, in prep.*

Learn the right building blocks



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





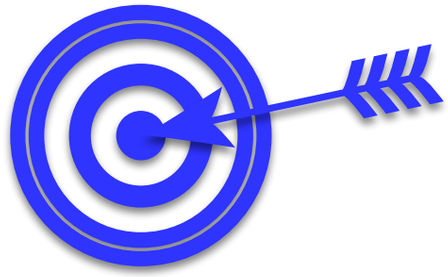
*Dickson et al. 2022, 2024, in prep.*

# The big picture

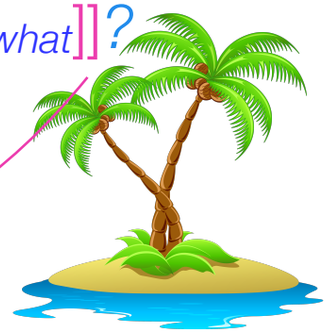

One way to succeed at learning about constraints on *wh*-dependencies (syntactic islands) is to learn them indirectly.

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

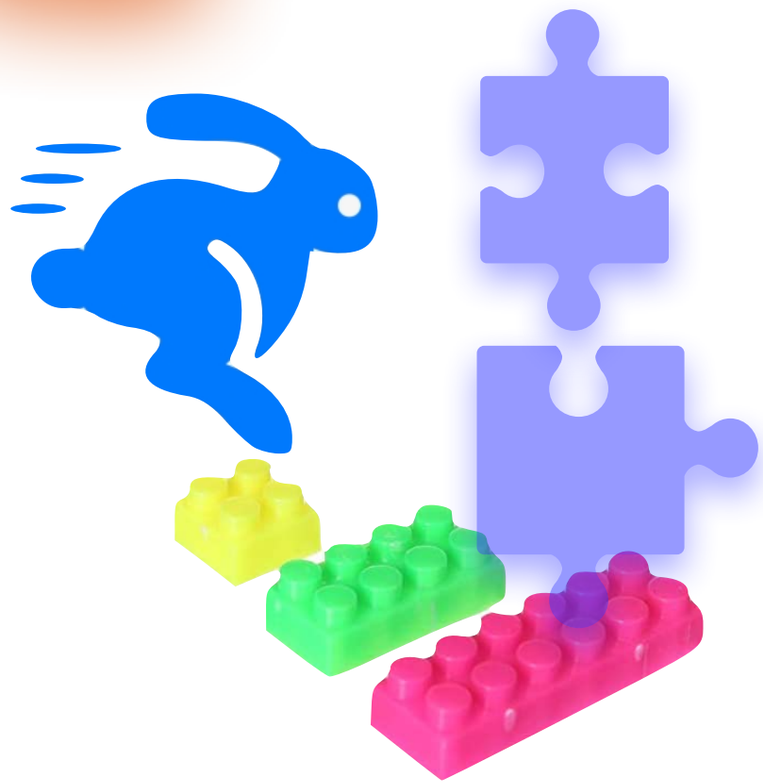
# The big picture



*What* [ [ __*what*]]?

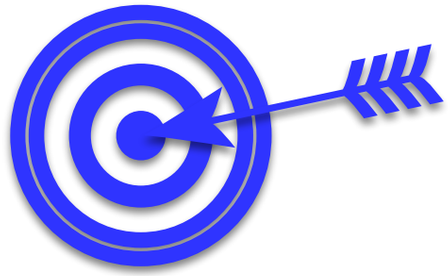


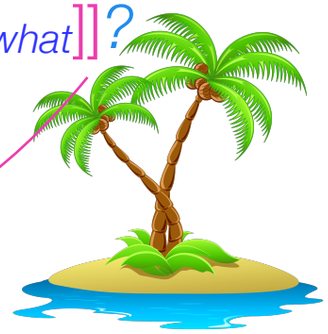
In particular, learn how to represent *wh*-dependencies efficiently using pieces that can be used to represent any *wh*-dependency.



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

# The big picture





*What* [[ __*what*]]?

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

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

# Thank you!

Niels Dickson

Jon Sprouse

Richard Futrell

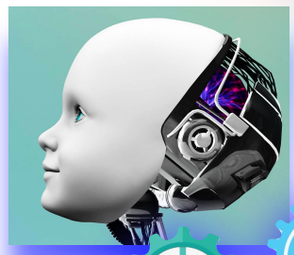
Alandi Bates

BUCLD 2018 UCSD Linguistics 2020 ForMA Group 2020

UMD Linguistics 2020 BUCLD 2021 SCiL 2022

UArizona Linguistics 2022 UChicago LEAP 2022

UPenn 2023 UCI QuantLang Collective



Computation of Language Laboratory

UC Irvine


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