Investigating child language acquisition using computational cognitive modeling: **How to understand kids using math**

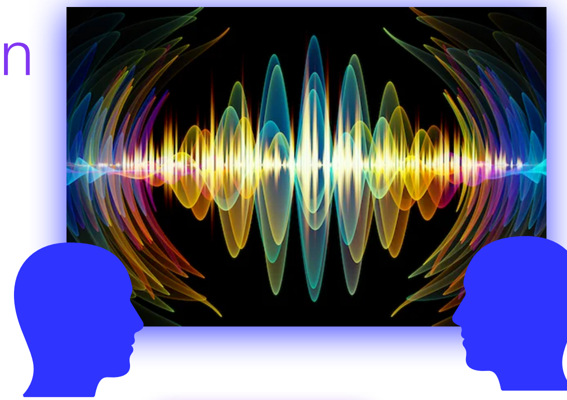
Lisa Pearl
University of California, Irvine

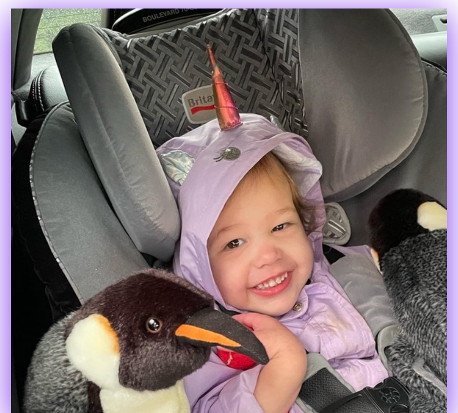
Computation of Language Laboratory
UC Irvine

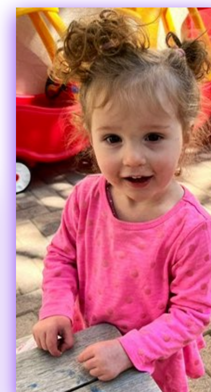
Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lpearl@uci.edu

October 21, 2024
Department of Linguistics
USC

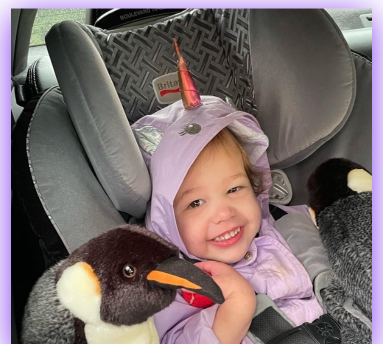
(Briefly) about child language acquisition

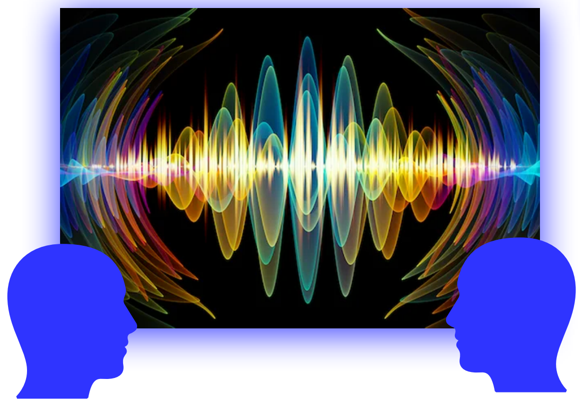






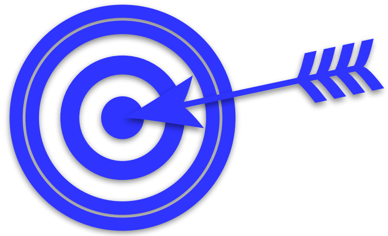
child language acquisition:
extraction of language information
by young children

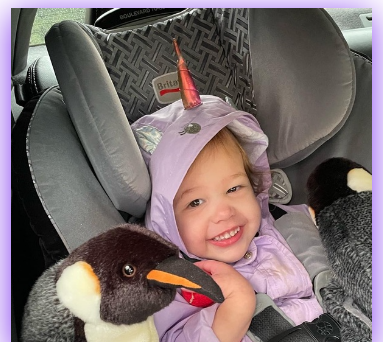


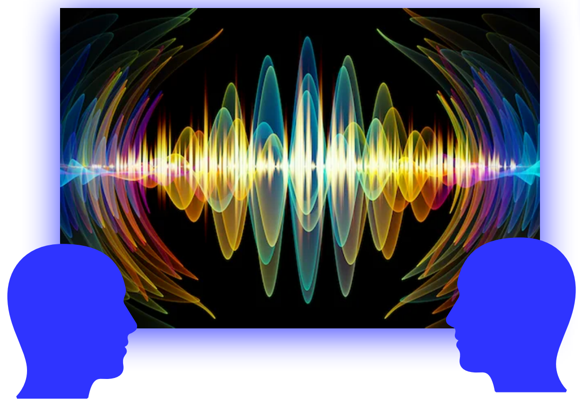


child language acquisition

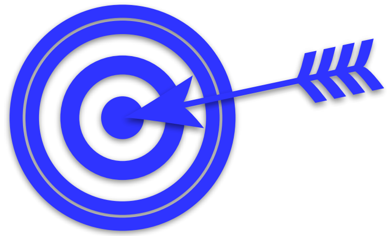
One question:
What kind of language information exactly is the target?

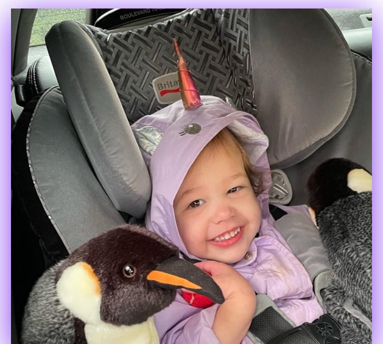


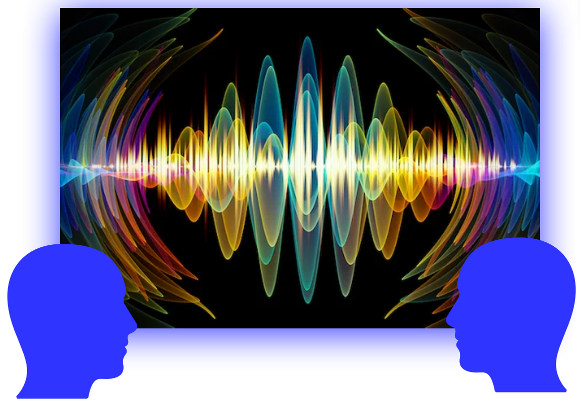




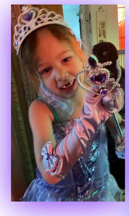
One example of target language information:
knowledge about *wh*-dependencies

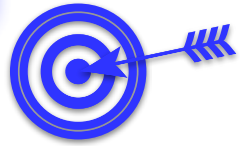






child language acquisition

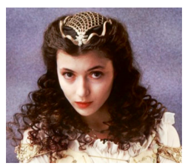


knowledge about *wh*-dependencies

for instance, which *wh*-dependencies are strongly dispreferred vs. preferred

This kitty was bought as a present for someone.

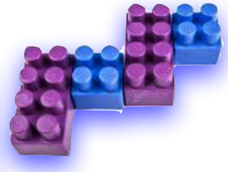


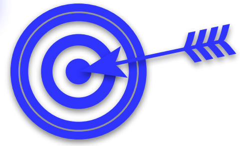
Lily thinks this kitty is pretty.


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


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

child language acquisition

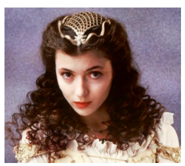




knowledge about *wh*-dependencies

for instance, which *wh*-dependencies are strongly dispreferred vs. preferred

The kitty's antics make Lily laugh.



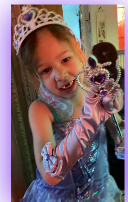
She wants to get him in order to pet him.

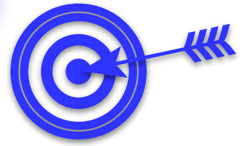
What did Lily get __what in order to pet __what?

What did Lily get __what and then pet __what?

What did Lily get __what before petting __what?

child language acquisition



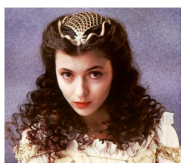



knowledge about *wh*-dependencies

for instance, which *wh*-dependencies are strongly dispreferred vs. preferred

This puppy is also really adorable.





Lily wants to pet him too, but she has to get the kitty first.

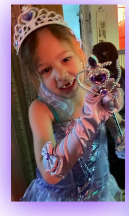

What … __what … __what?

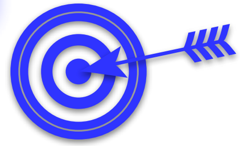
What did Lily get the kitty in order to pet __what?

What did Lily get the kitty and then pet __what?

What did Lily get the kitty before petting __what?

child language acquisition



knowledge about *wh*-dependencies


What … __what … __what?

What … __what?

The frequency of lexical items can also affect the judged acceptability of *wh*-dependencies.


Elizabeth knows that Jack saw something.
She feels the need to tell someone else about it.

She's a bit unhappy about the situation.

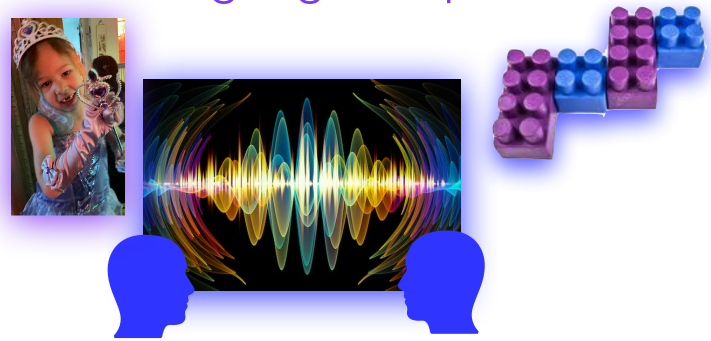
What did Elizabeth say that Jack saw __what?

more frequent

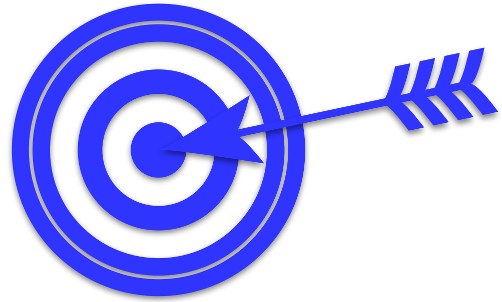
What did Elizabeth whine that Jack saw __what?

less frequent

child language acquisition



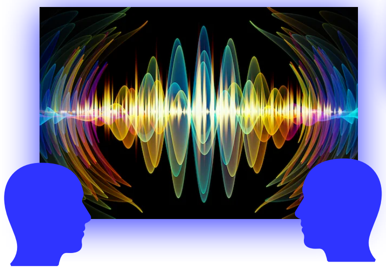
knowledge about *wh*-dependencies

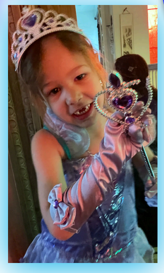


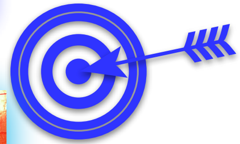
So how can children acquire this (and other) target knowledge?

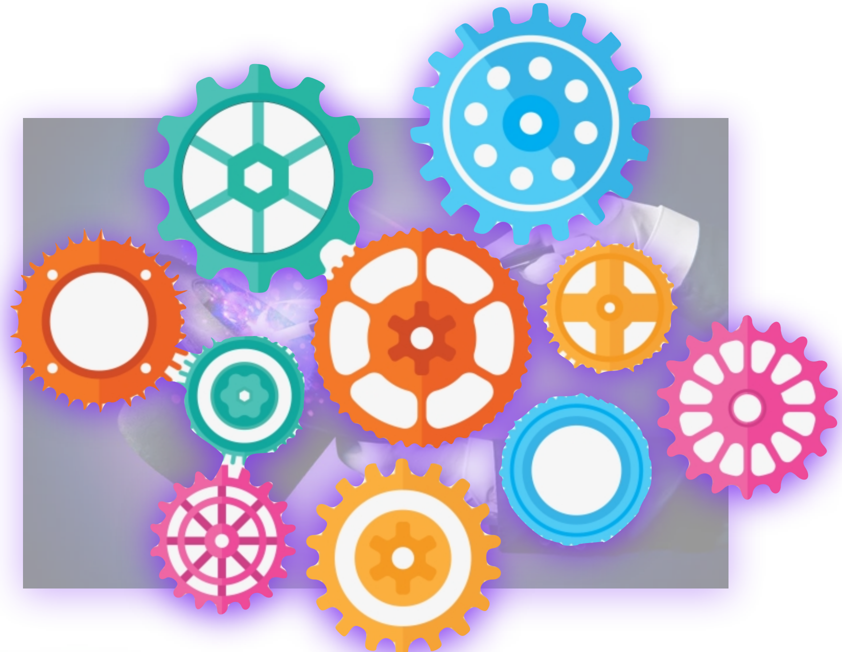
How does the magical process of acquisition work?

child language acquisition

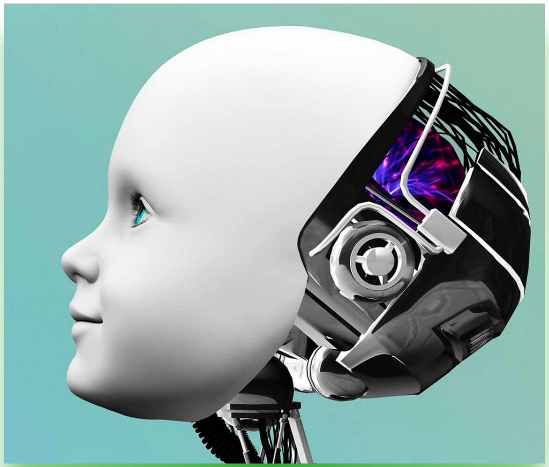




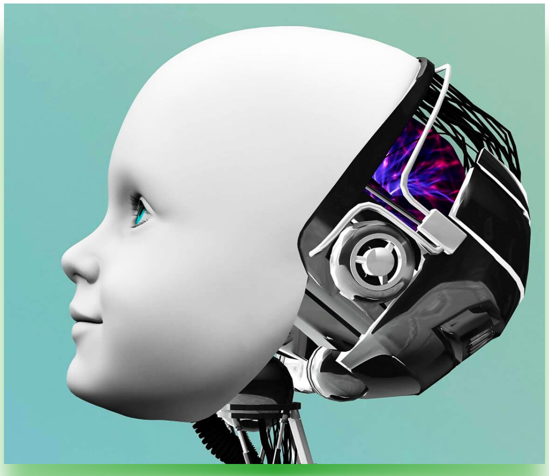


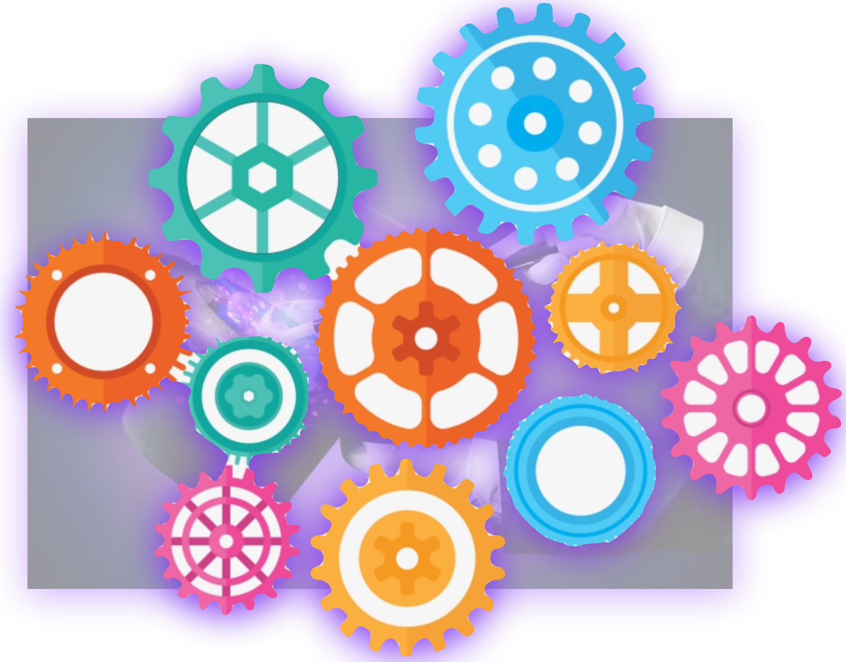


I primarily use quantitative methods like computational cognitive modeling to try to figure this out.



computational cognitive modeling

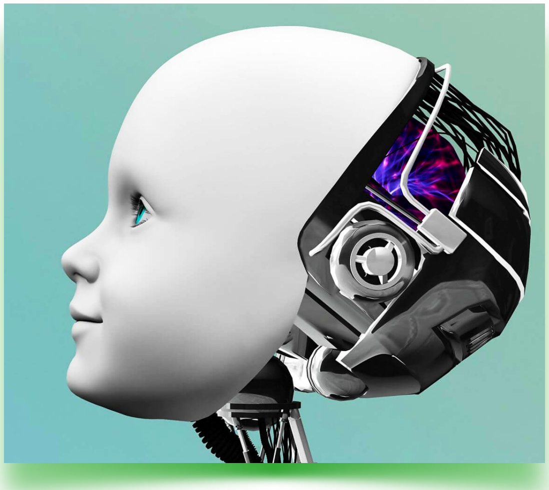


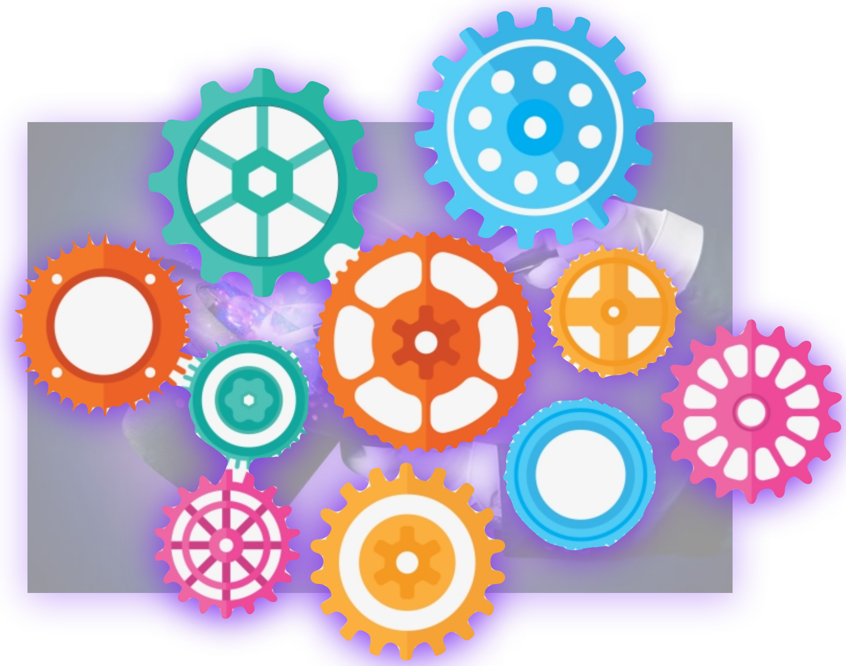


Computational cognitive modeling lets us explore theoretical ideas precisely, and evaluate how well any particular theory can explain empirical data on children's language acquisition.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

computational cognitive modeling

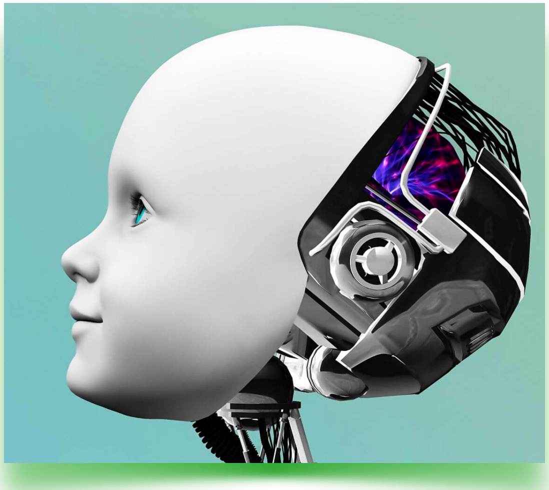


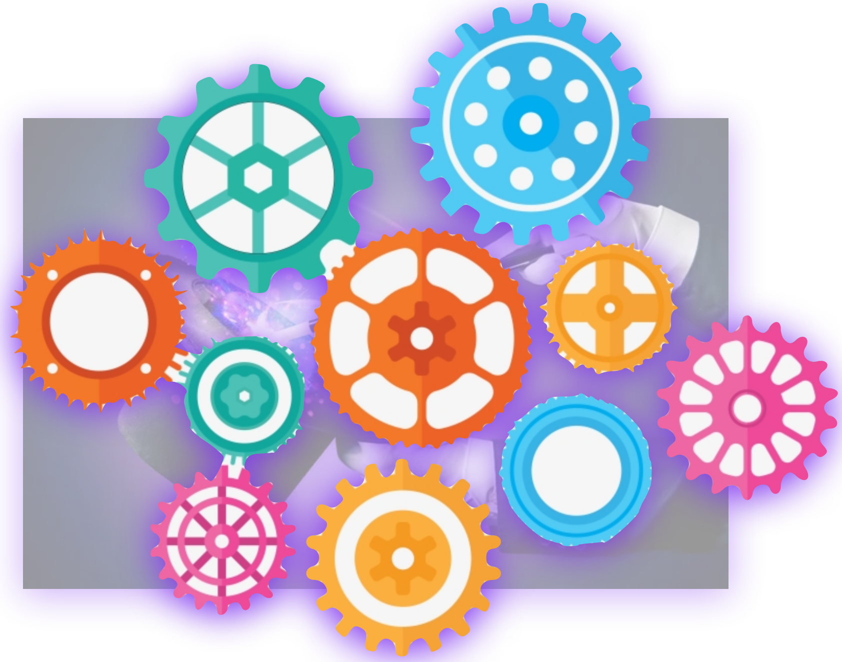


Math is at the heart of this tool.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

computational cognitive modeling

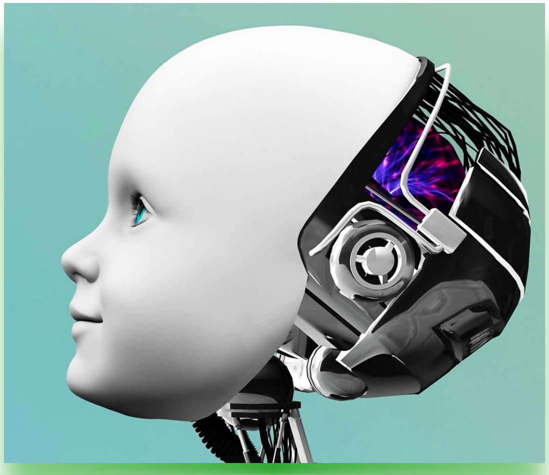


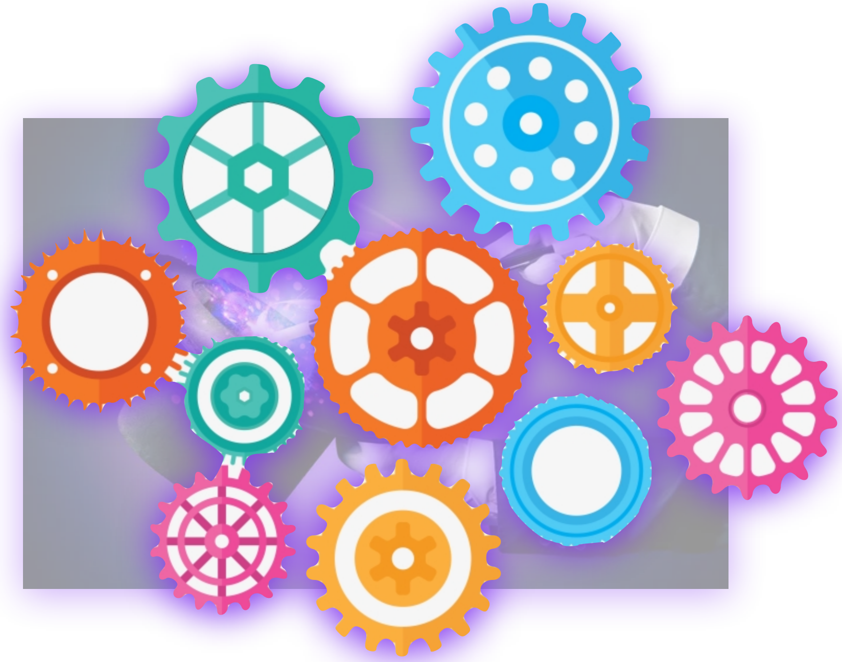


One main part: Counting things

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

computational cognitive modeling

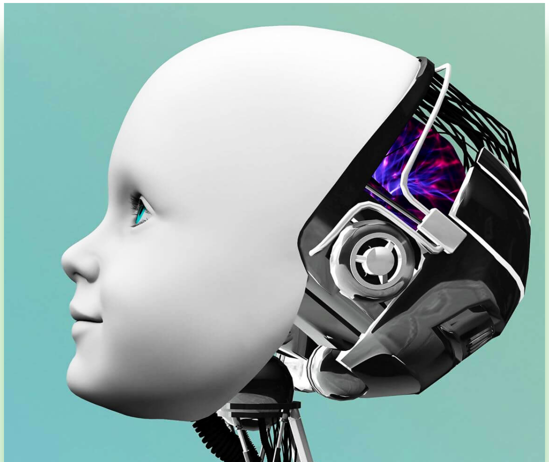


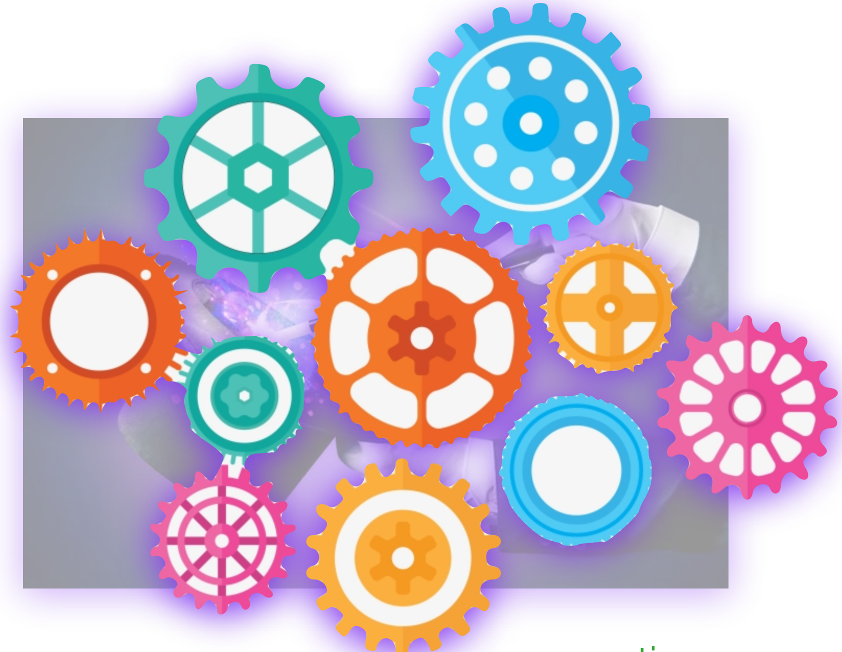


One main part: Counting things
(sometimes we count a lot of things)

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

computational cognitive modeling



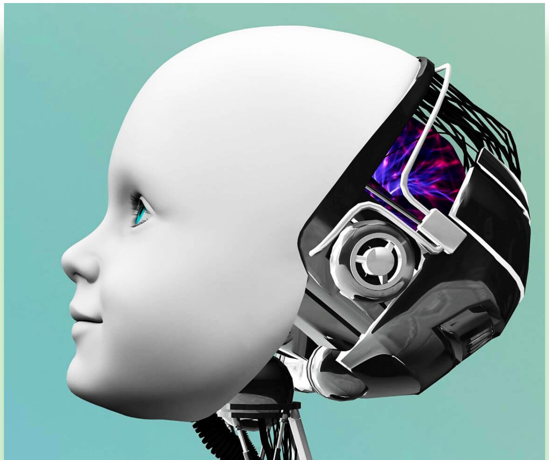


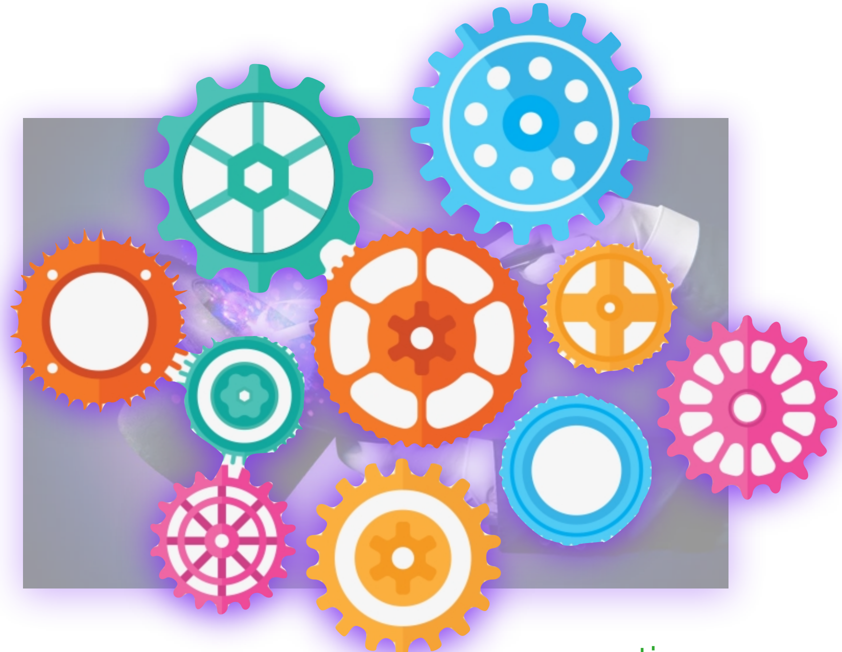
counting

Another part: principled reasoning based on those counts

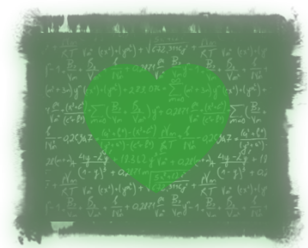
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

computational cognitive modeling





counting



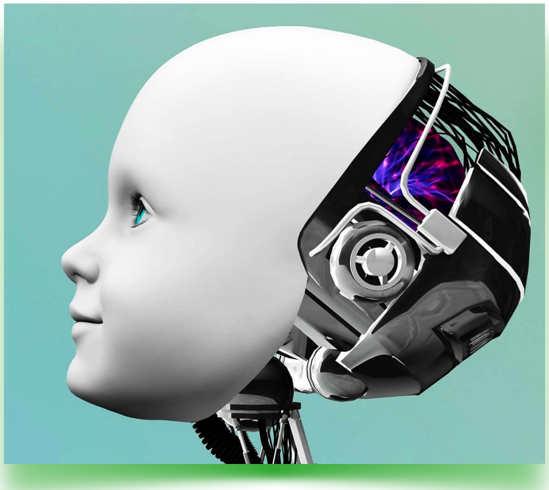
Another part: principled reasoning based on those counts

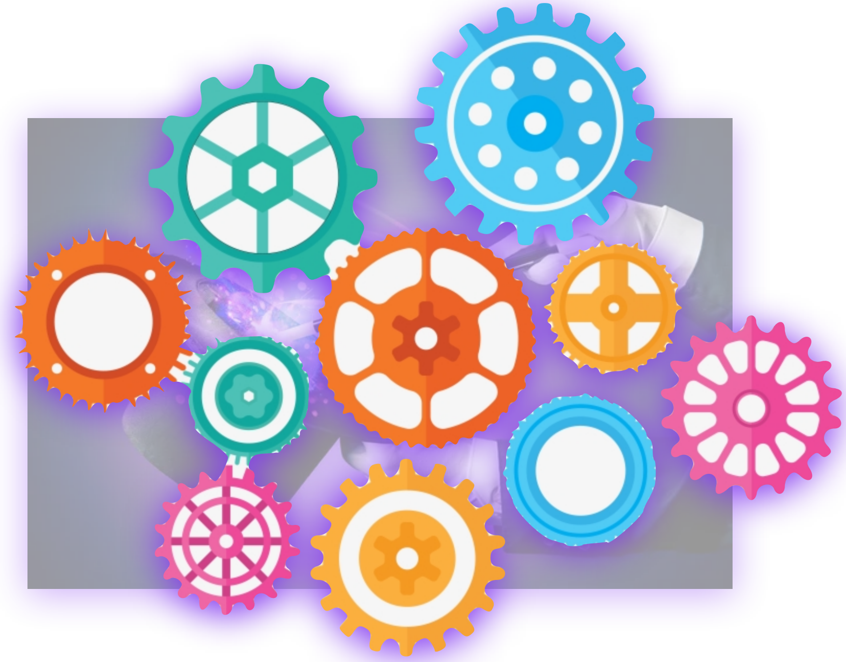
Bayesian inference

$$p(Generalization|Data) \propto p(Generalization) \cdot p(Data|Generalization)$$

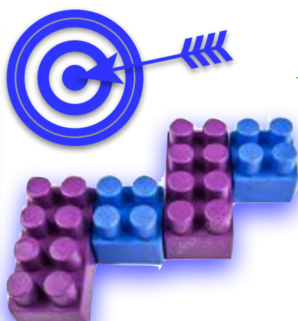
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

computational cognitive modeling





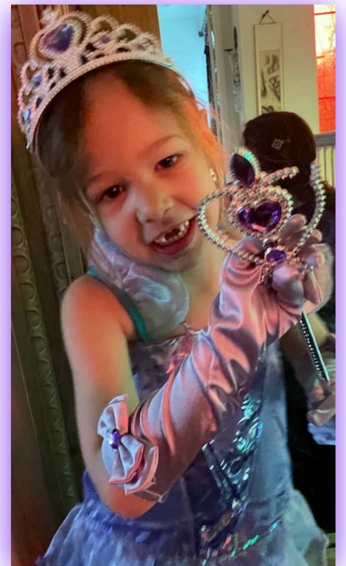
But what do we count and reason over? How do we connect that information to language acquisition?

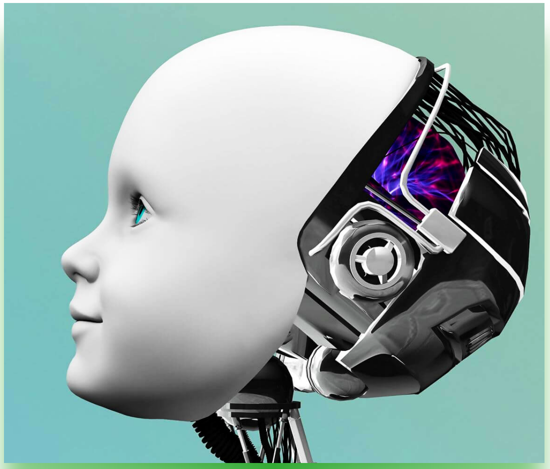


Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

child language acquisition

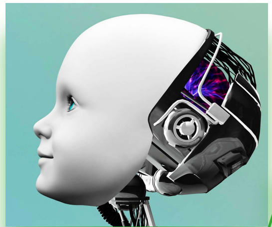
computational cognitive modeling



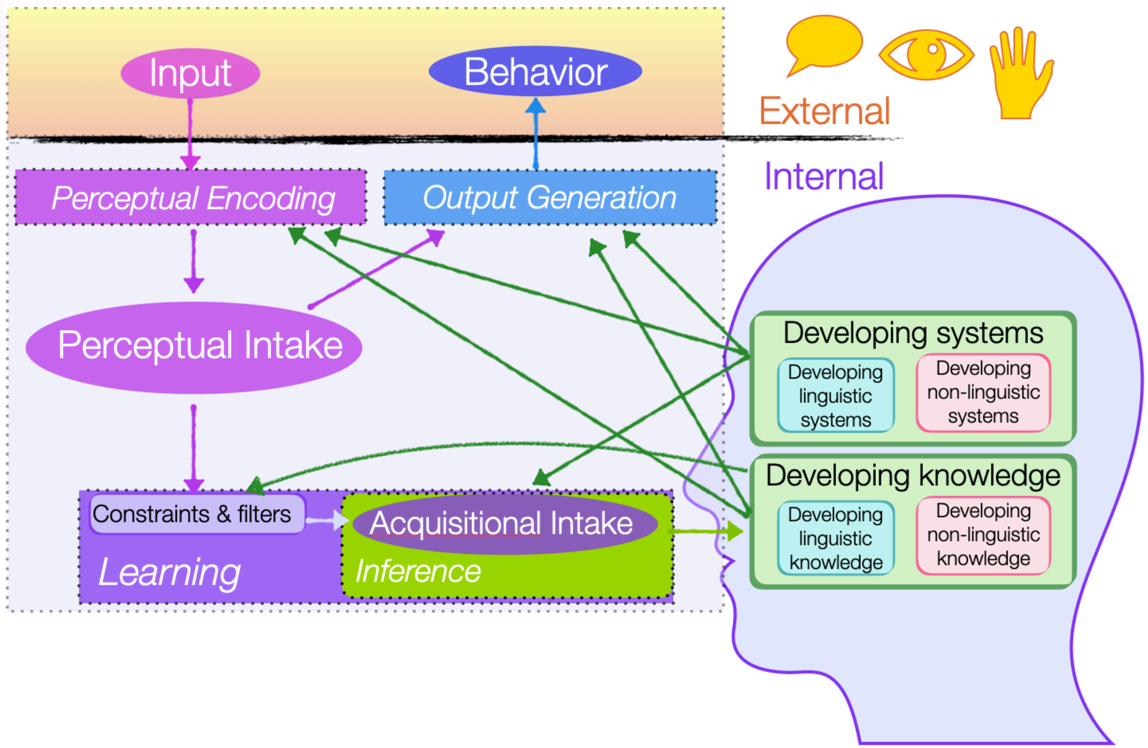


We can use computational cognitive modeling to encode a child's acquisition process very precisely.

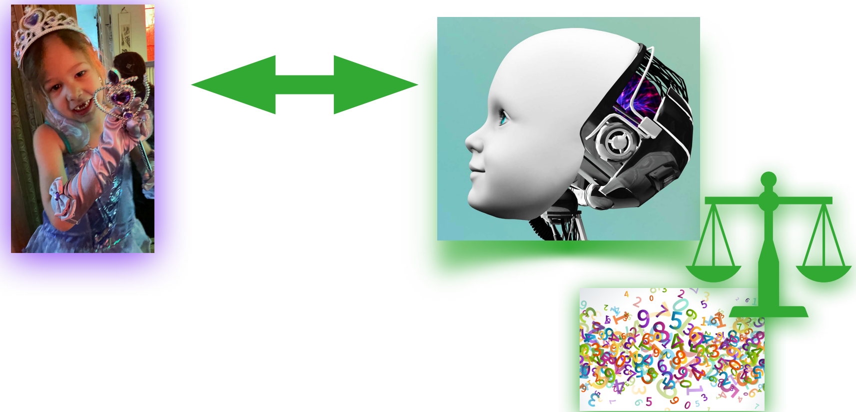
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review



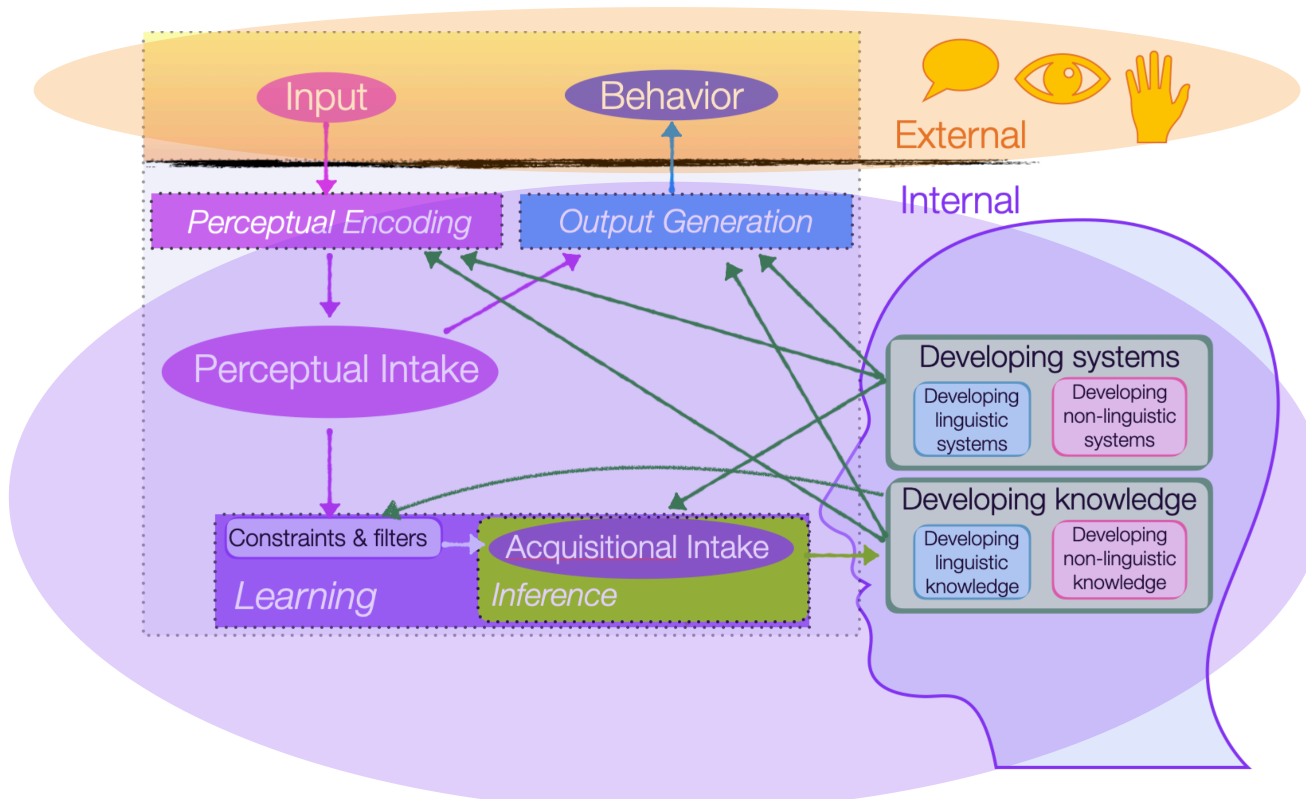
One way to encode a child's acquisition process.



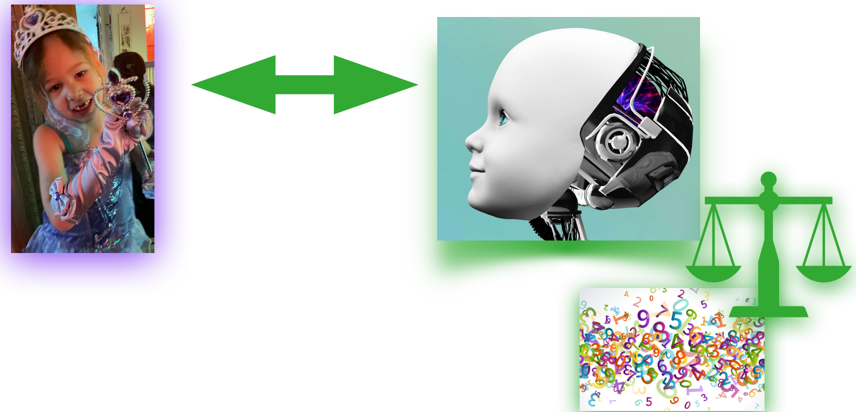
Pearl 2023 JoCL



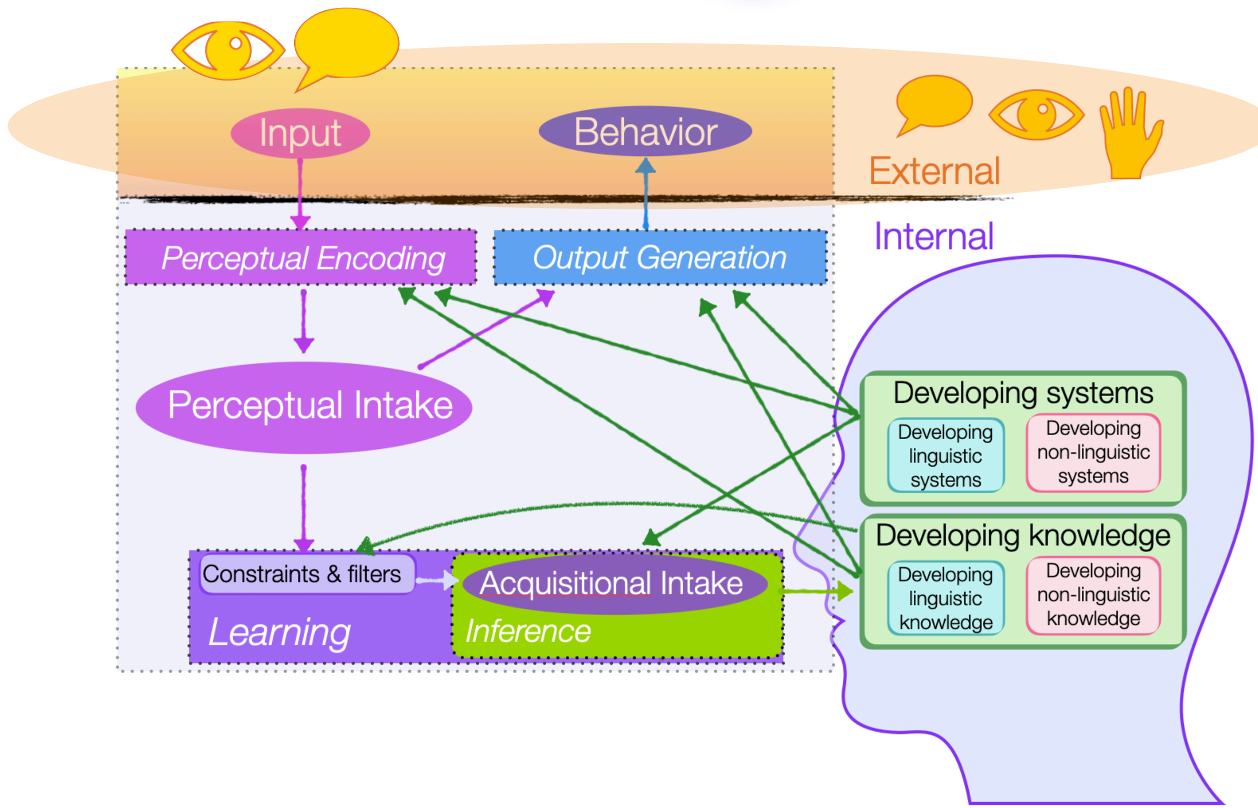
A basic distinction:
What's external vs.
internal to the child



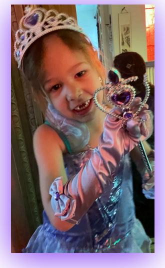
Pearl 2023 JoCL

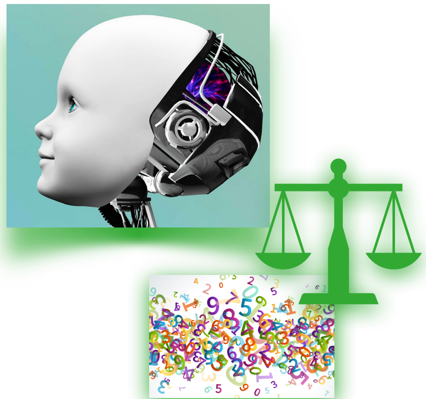


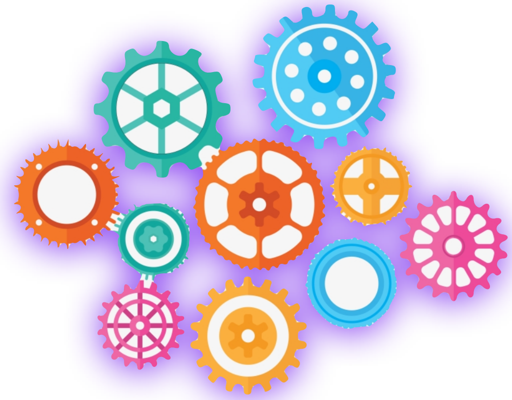
We can observe external things like the input signal to the child.



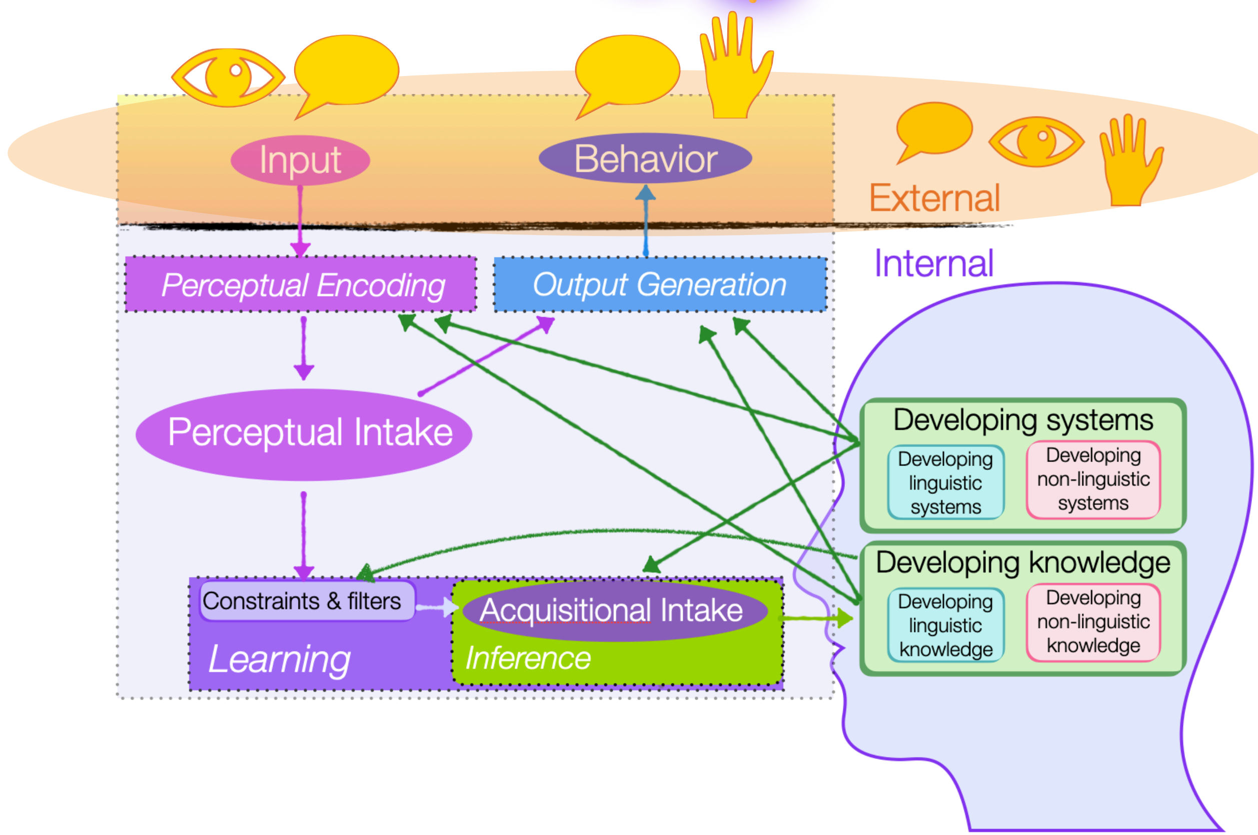
Pearl 2023 JoCL



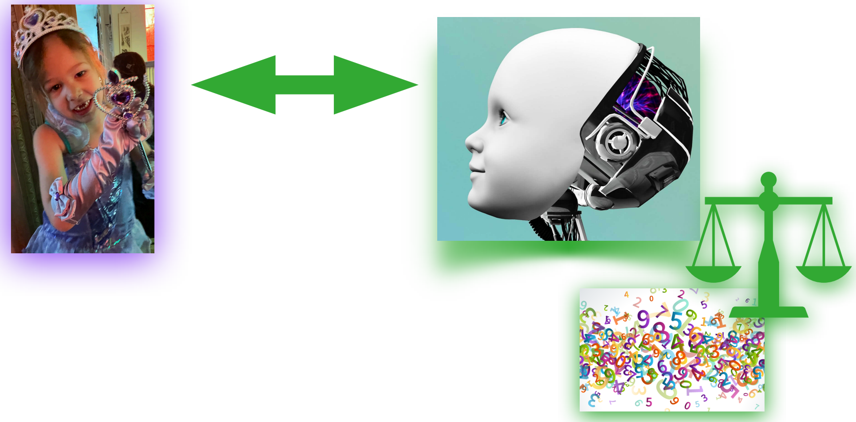




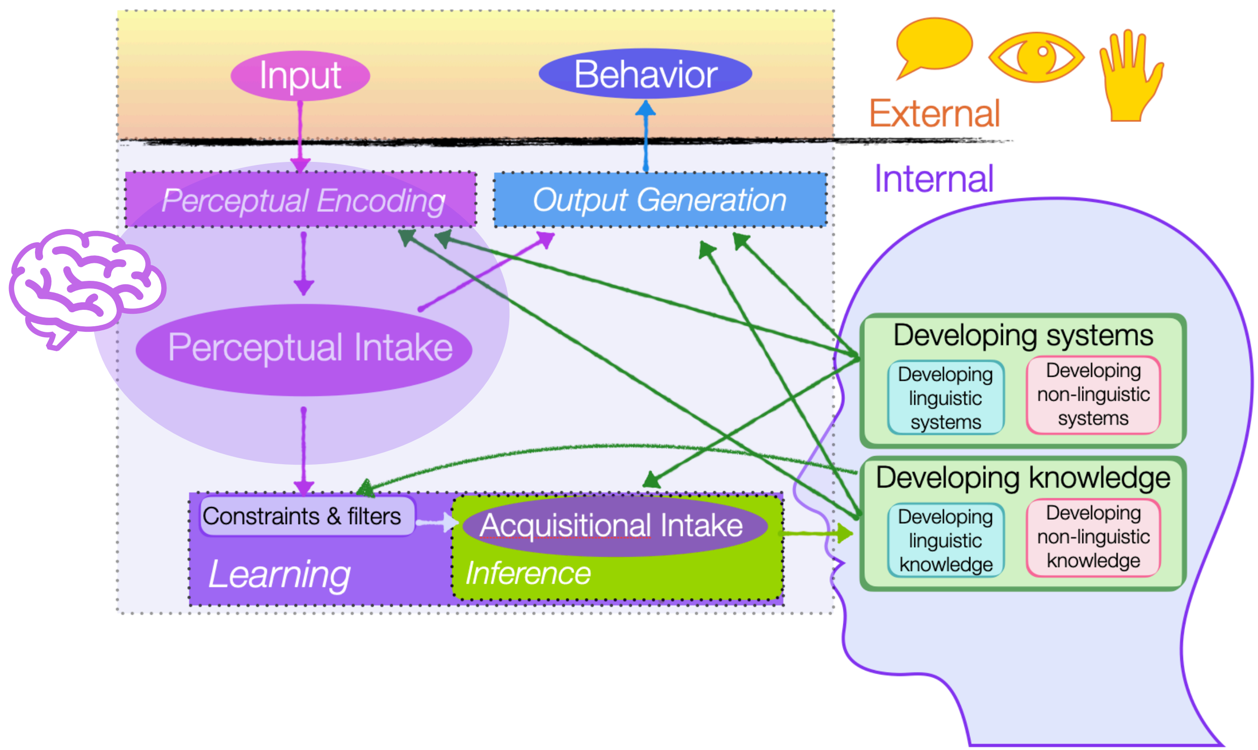
We can observe external things like the behavior the child demonstrates.



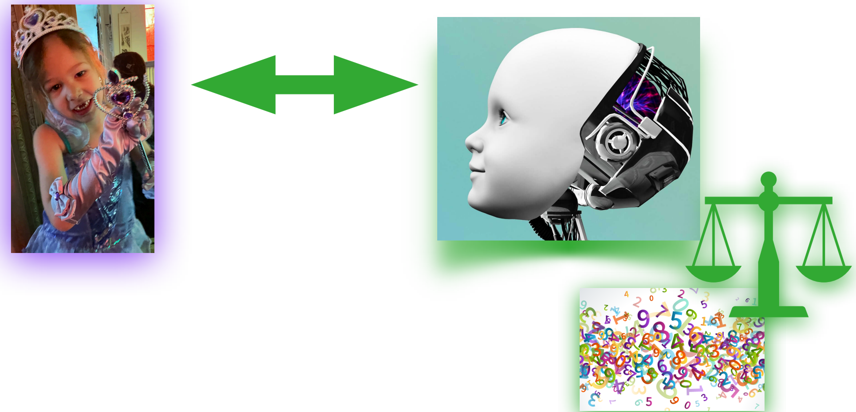
Pearl 2023 JoCL



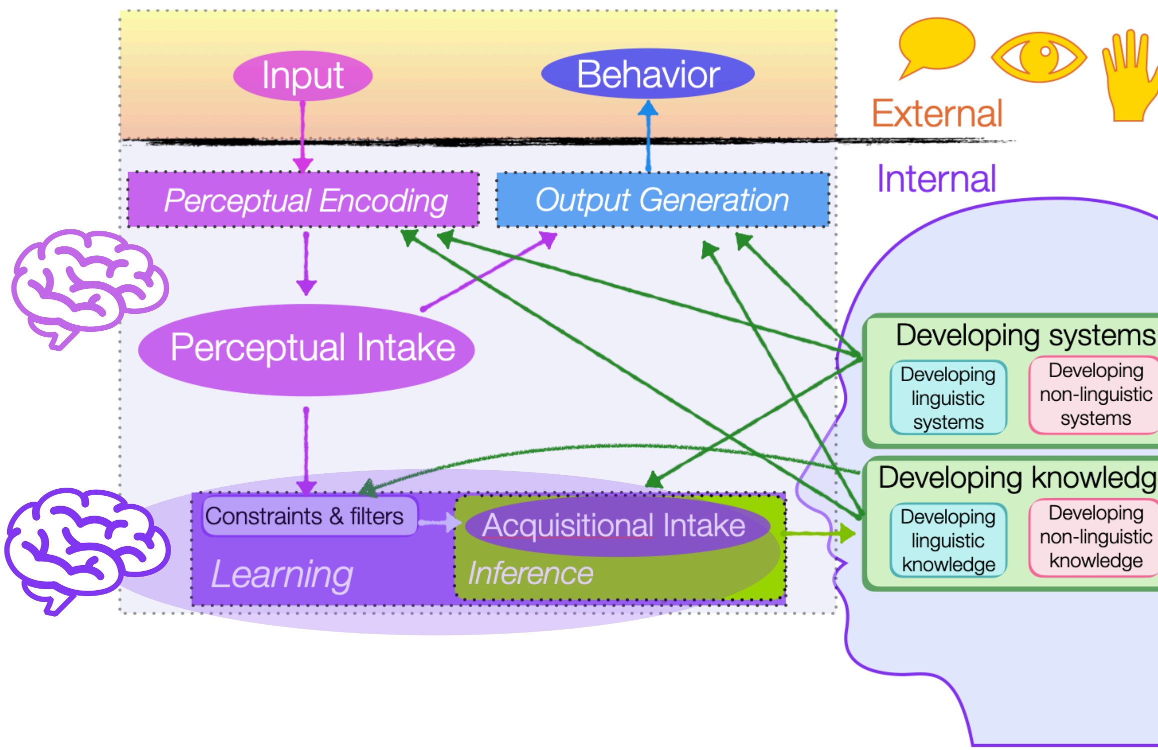
We can't observe internal things like what the child can perceive from the input signal.



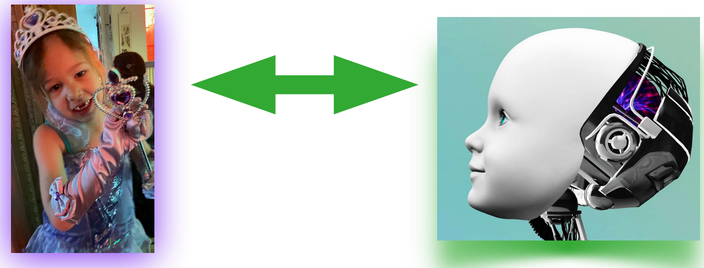
Pearl 2023 JoCL



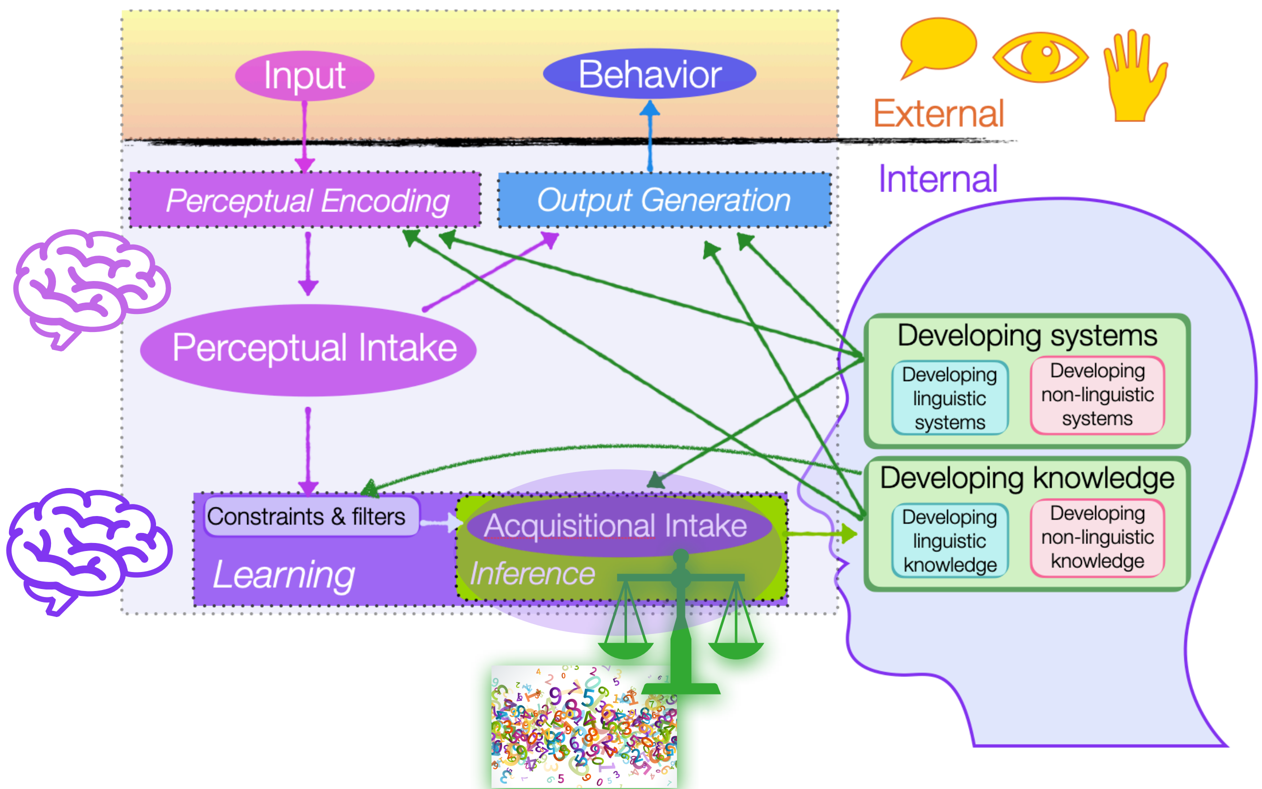
We can't observe internal things like what additional filtering there might be before the child learns from that intake.



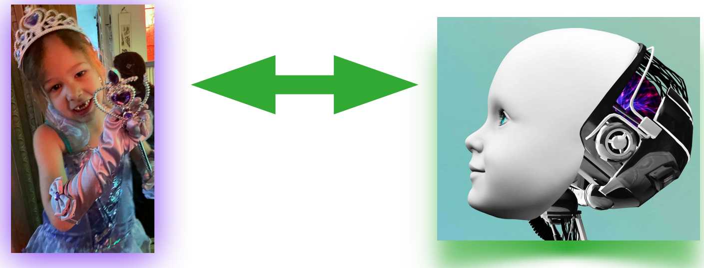
Pearl 2023 JoCL



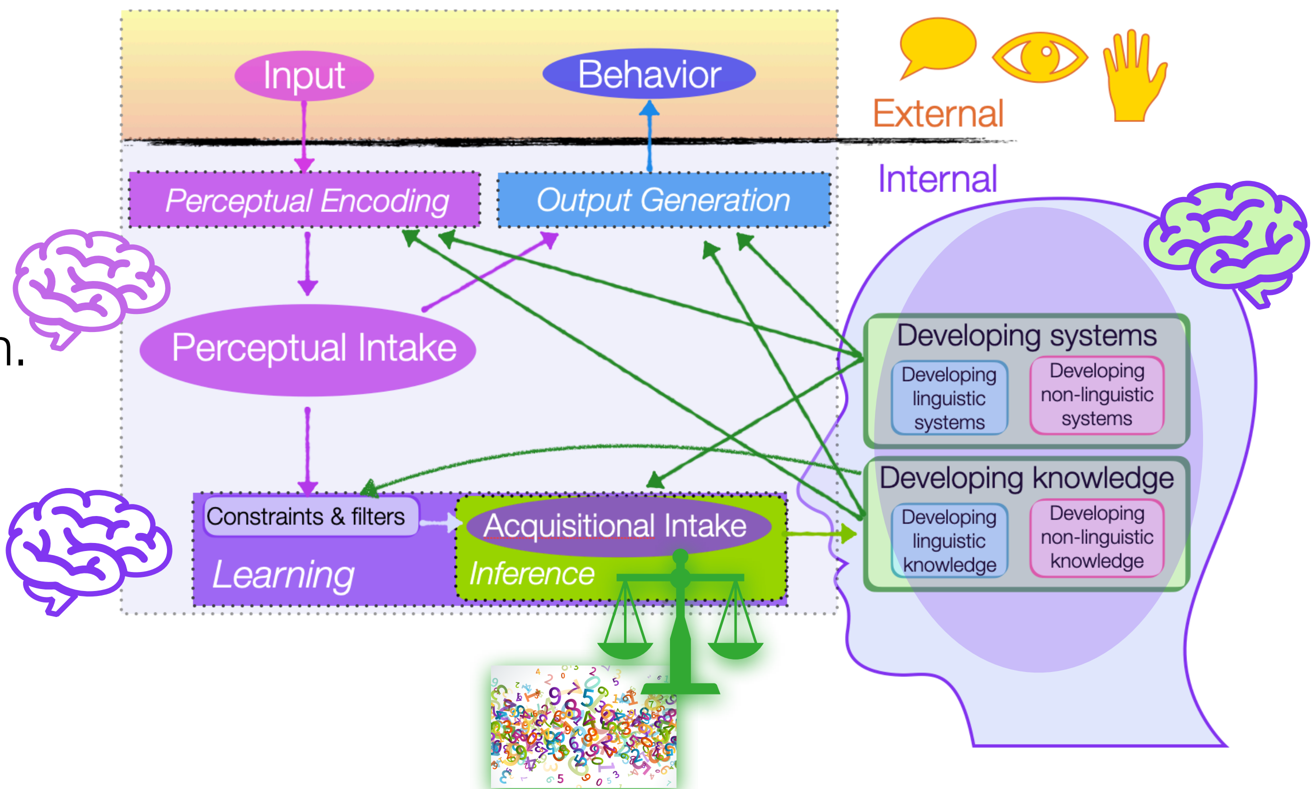
We can't observe internal things like how the inference process works.



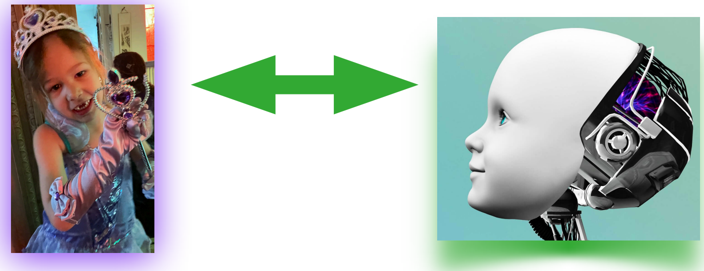
Pearl 2023 JoCL



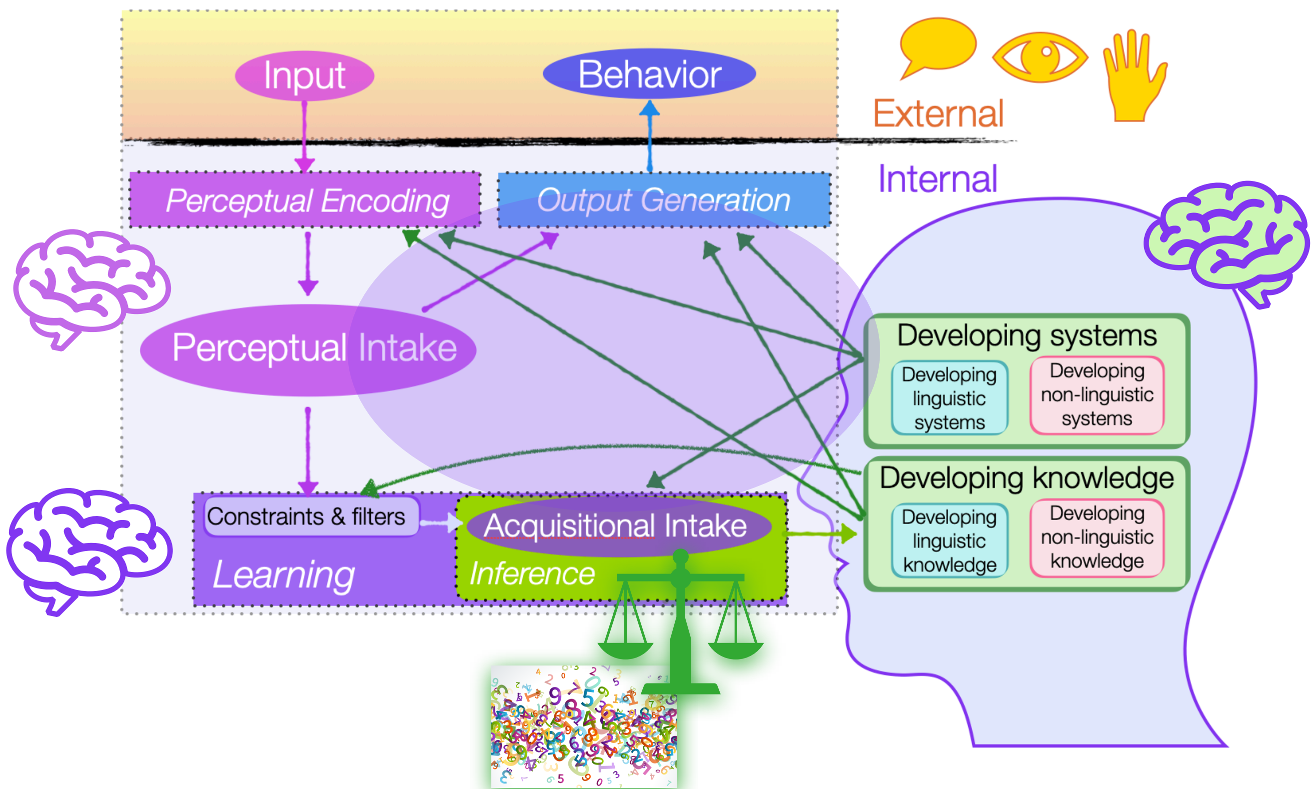
We can't observe internal things like what state the child's developing knowledge and developing cognitive systems are in.



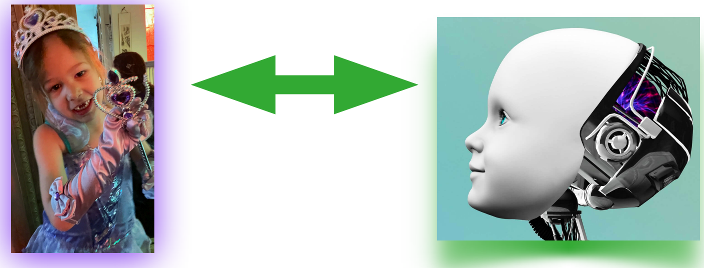
Pearl 2023 JoCL



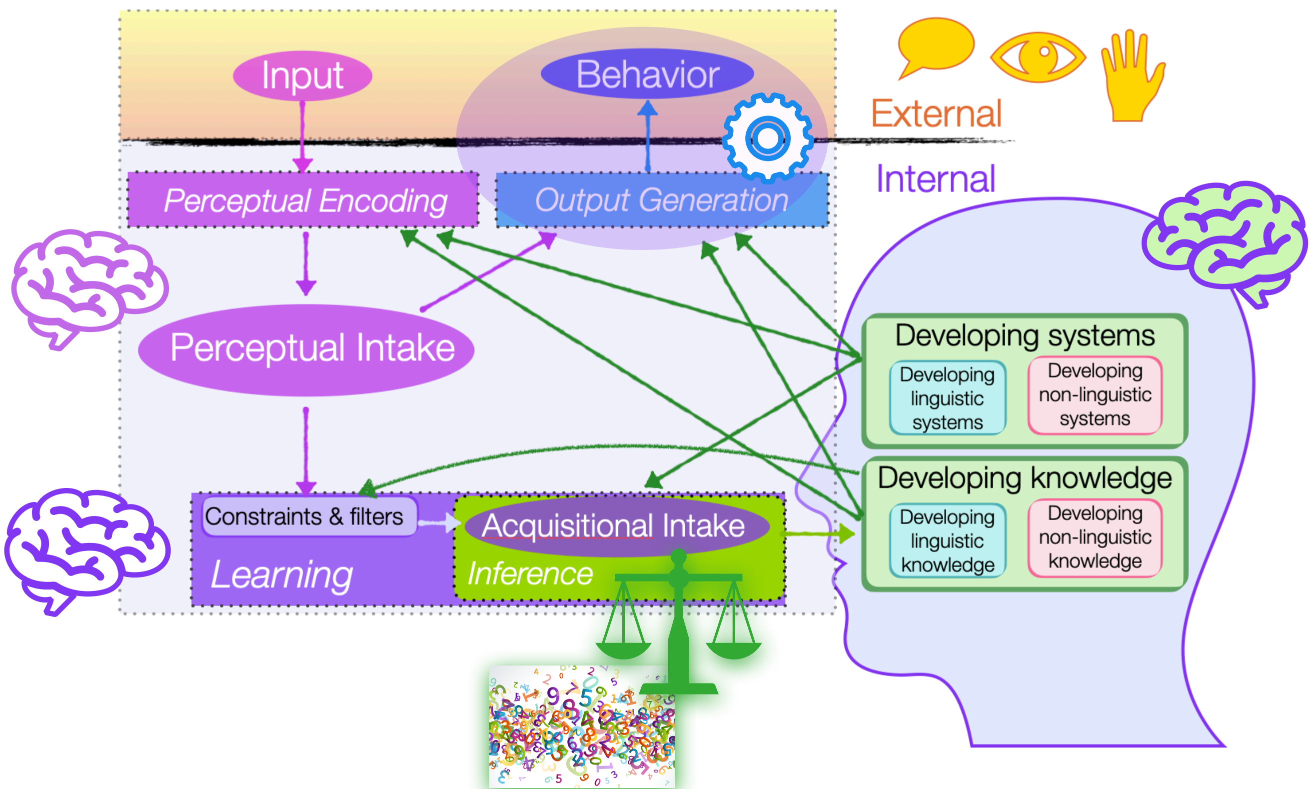
We can’t observe internal things like how these developing components impact everything else internal.



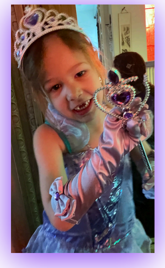
Pearl 2023 JoCL

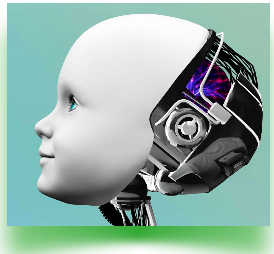


We can't observe internal things like how the child generates the output we eventually can observe.

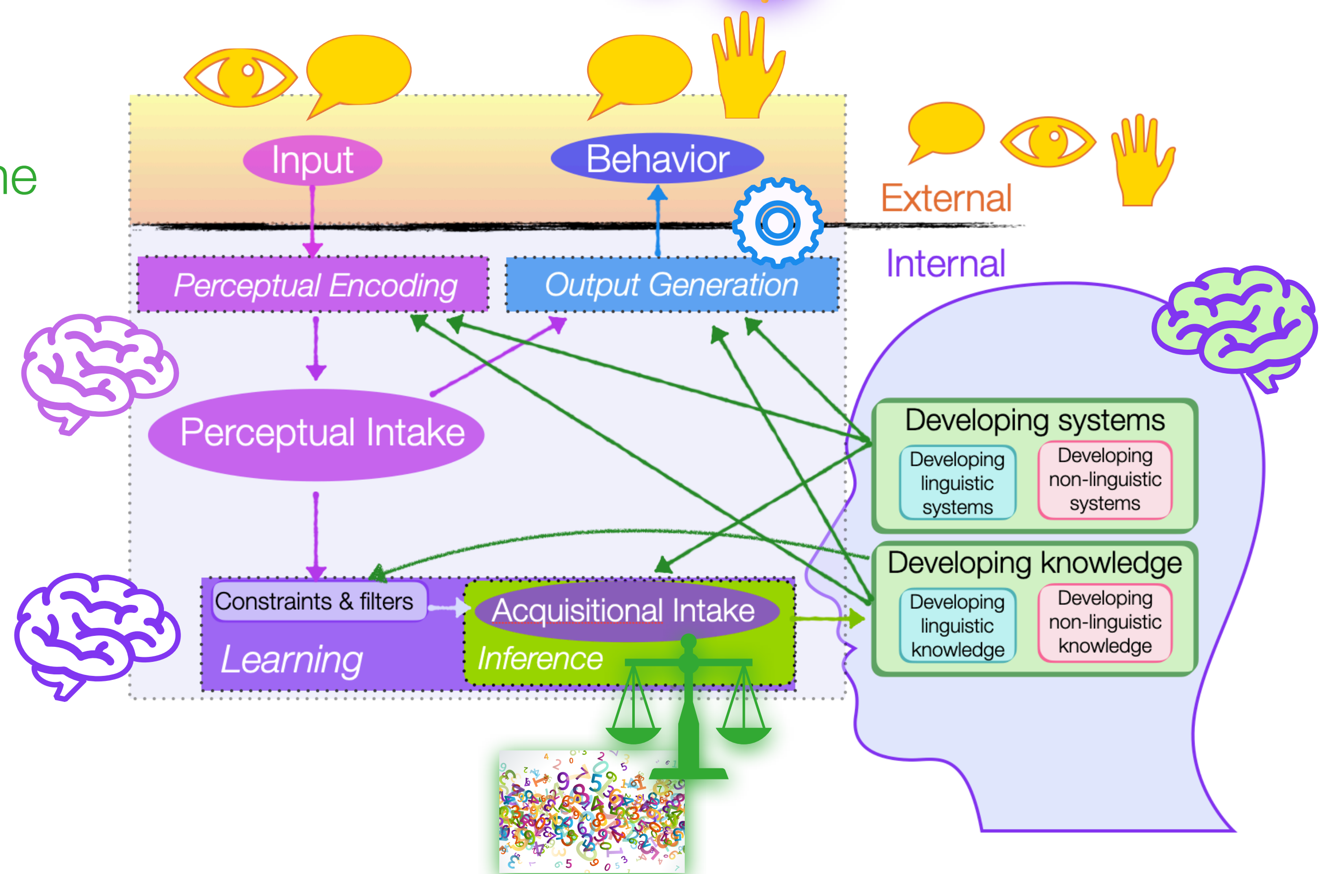


Pearl 2023 JoCL

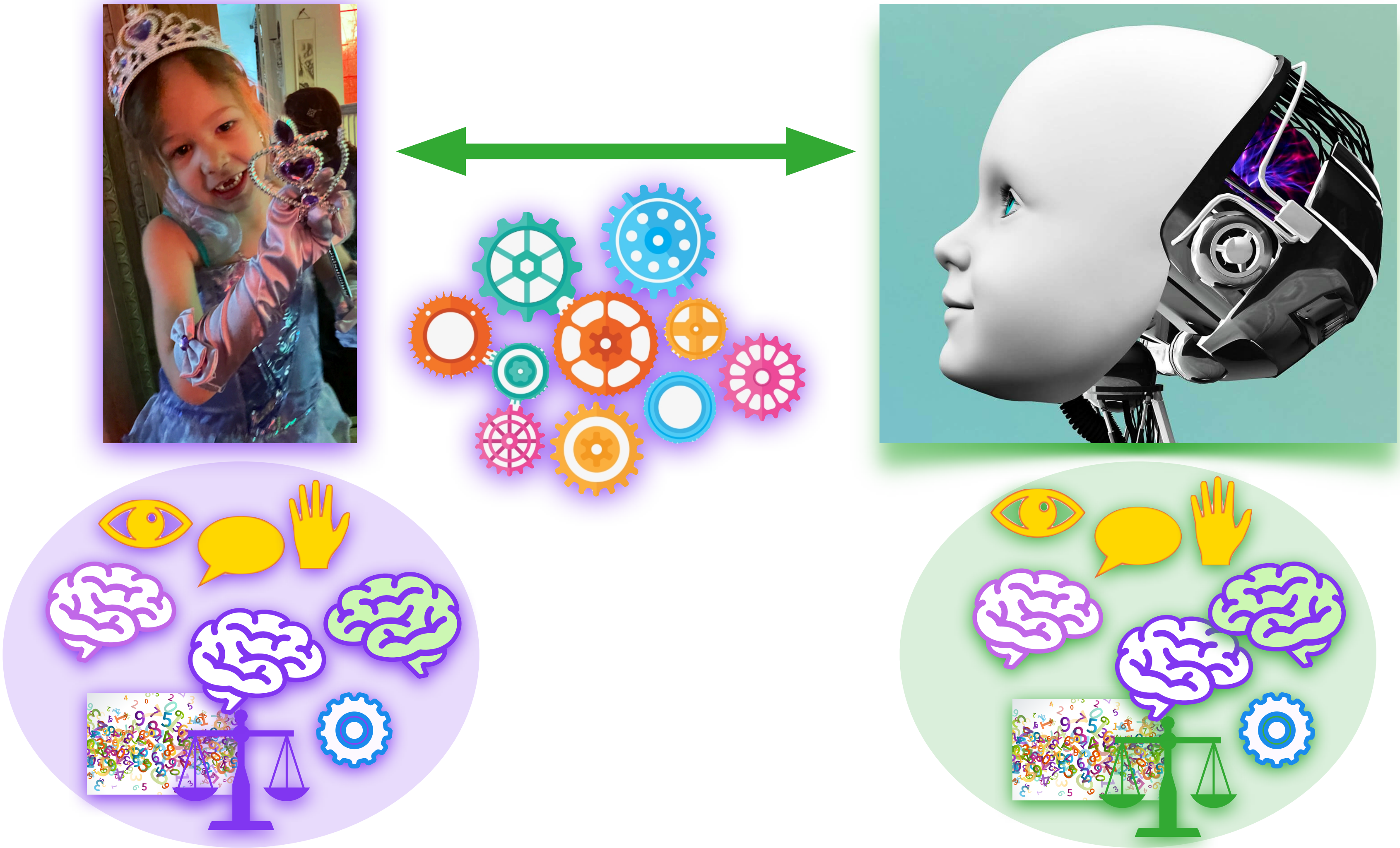




All these parts are important for making the process of acquisition very precise.

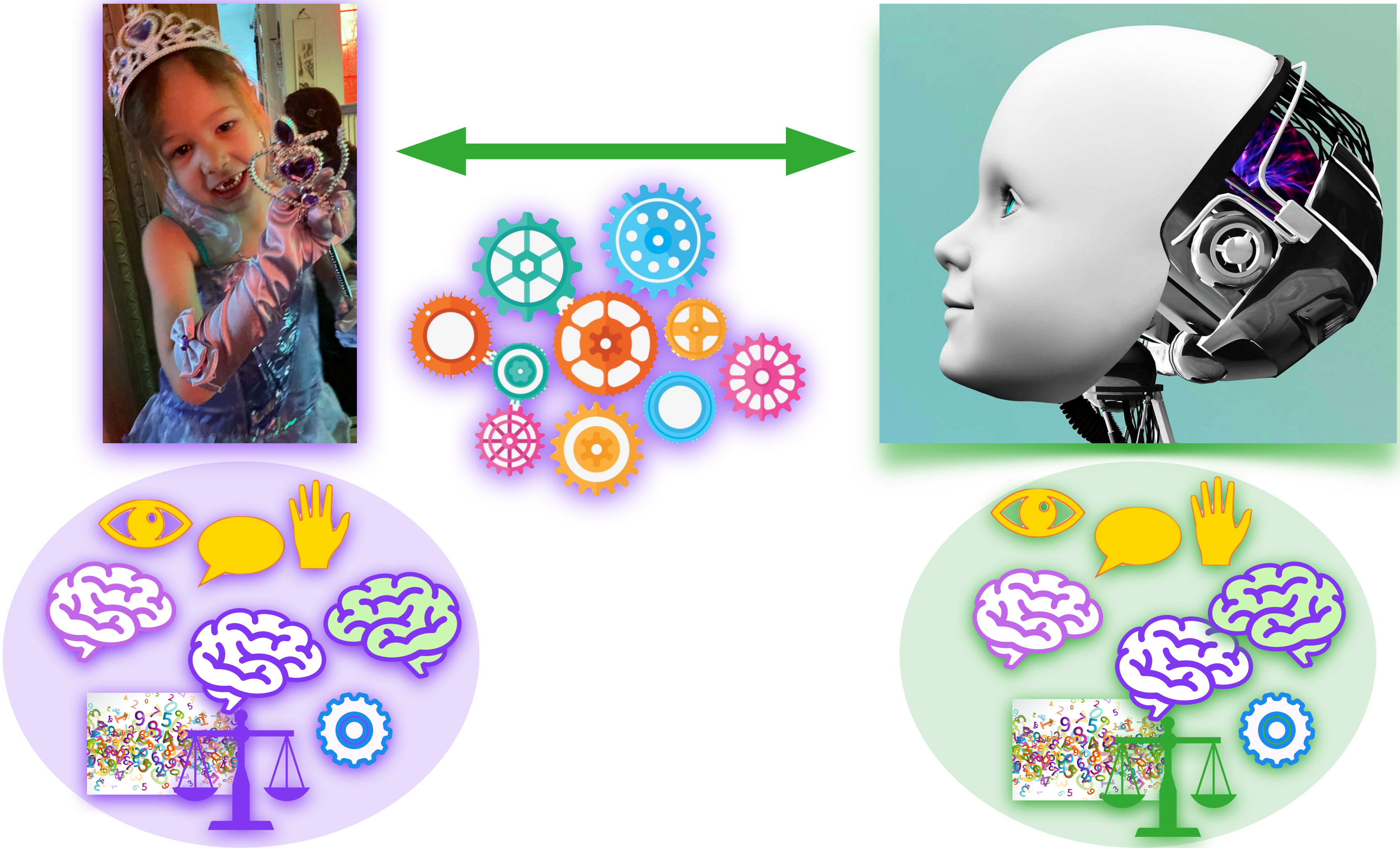


Pearl 2023 JoCL



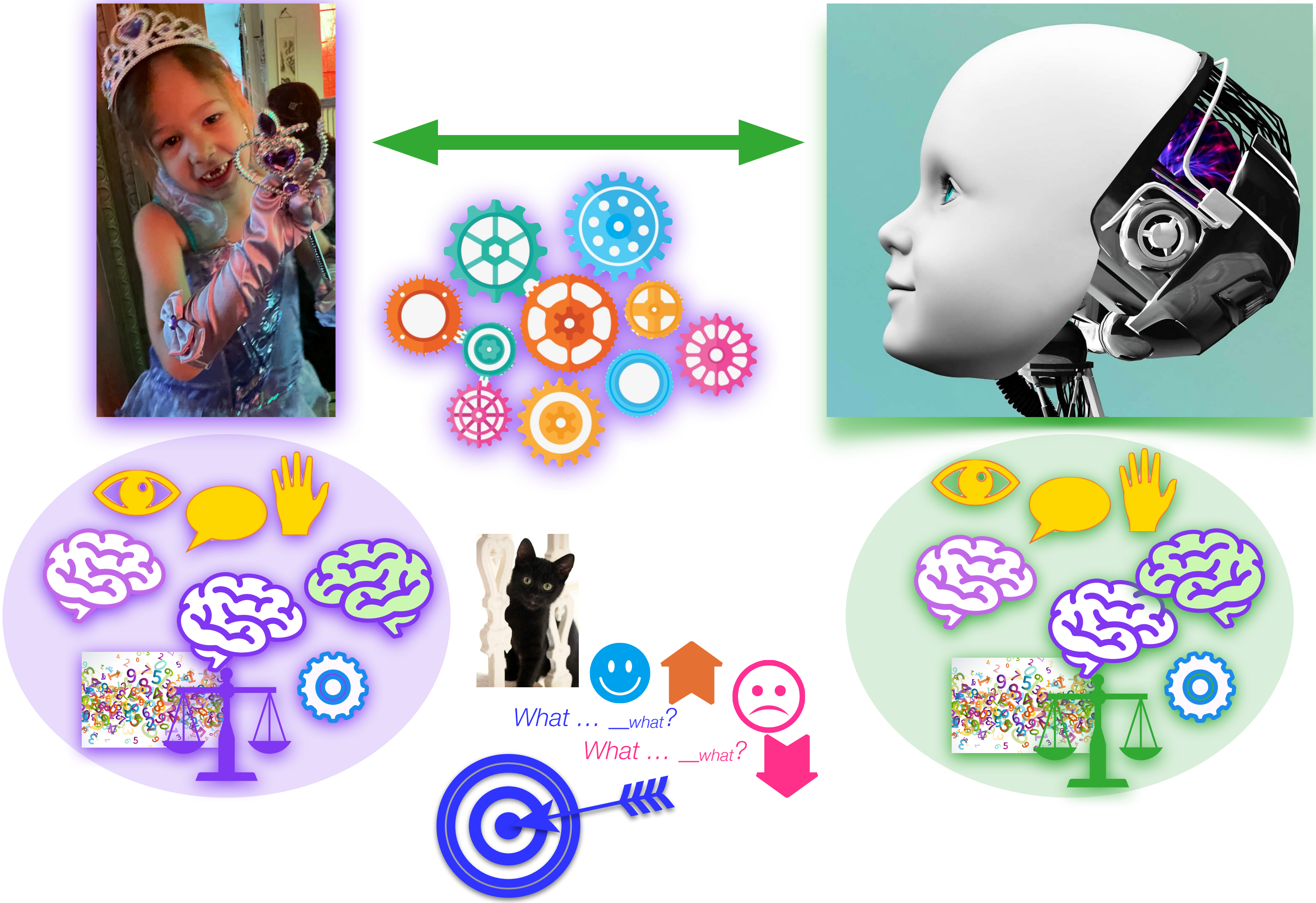
In a computational cognitive model, we try to implement all these parts of the acquisition process to **be the same** as how they work in children.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review



For some parts, we draw on available empirical data. For other parts, we rely on the acquisition theory we're trying to evaluate.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review



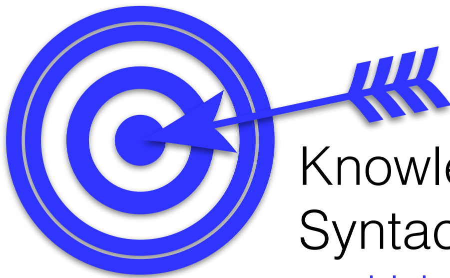
Let’s see how this works for the acquisition of some knowledge about *wh*-dependencies.

Knowledge:
Syntactic islands

…which involve constraints over *wh*-dependencies.

What … __what?

What … __what?



Knowledge:
Syntactic islands

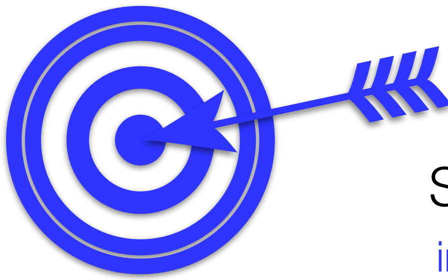
…which involve constraints over *wh*-dependencies.

This kitty was bought as a present for someone.

Lily thinks this kitty is pretty.

"Who does Lily think the kitty for __$_{who}$ is pretty?"

What does Lily think __$_{what}$ is pretty, and who does she think it's for __$_{who}$?



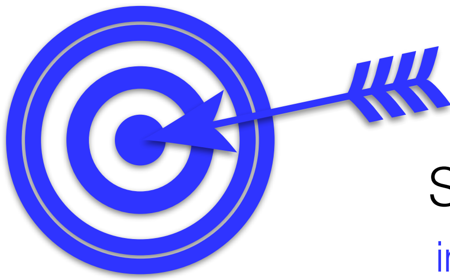
Syntactic islands

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

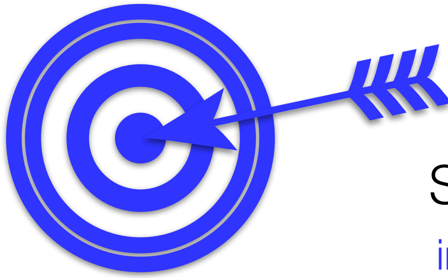
involve *wh*-dependencies.

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.

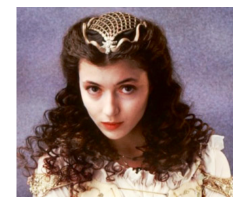


Syntactic islands

involve *wh*-dependencies.

What's going on here?

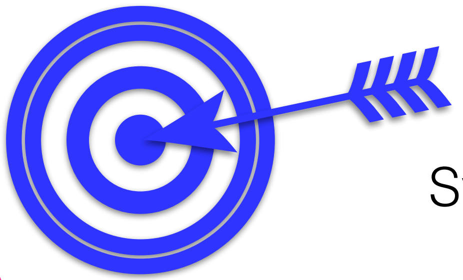
There's a dependency between the *wh*-word *who* and where it's understood (the gap)



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", which is latent structure the mind imposes on this sequence of words. A *wh*-dependency can't cross this structure. (metaphor: it can't "get out of" an island) [Ross 1967]



Syntactic islands

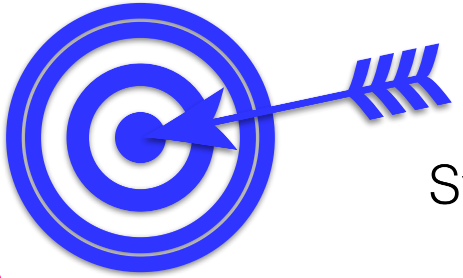
Ross 1967

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

Ross 1967

Who does Lily think the kitty for __$_{who}$ is pretty? Subject island

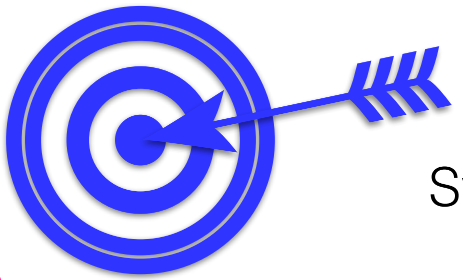
What did Jack make the claim that he bought __$_{what}$? Complex NP island

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

Ross 1967

Who does Lily think the kitty for__$_{\text{who}}$ is pretty? Subject island

What did Jack make the claim that he bought __$_{\text{what}}$? Complex NP island

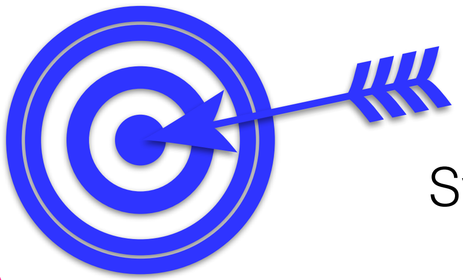
What did Elizabeth wonder whether Jack bought __$_{\text{what}}$? Whether island

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth worried it was something dangerous.

What did Elizabeth worry if Jack bought __$_{\text{what}}$?



Syntactic islands

Who does Lily think the kitty for__$_{who}$ is pretty? Subject island

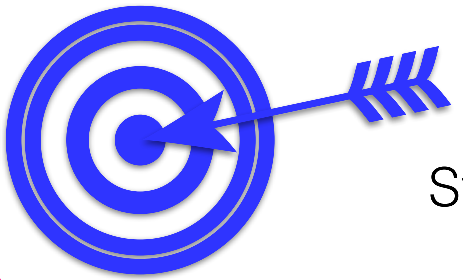
What did Jack make the claim that he bought __$_{what}$? Complex NP island

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

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

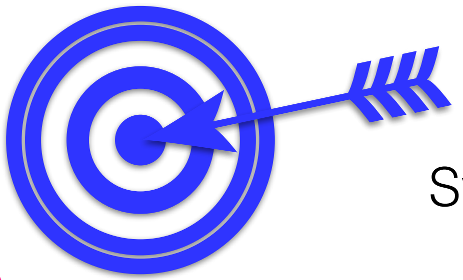
What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

Important: It's not about the length of the dependency.
(Chomsky 1965, Ross 1967)

What did Elizabeth think __what? ✔

Elizabeth



Syntactic islands

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

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

What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

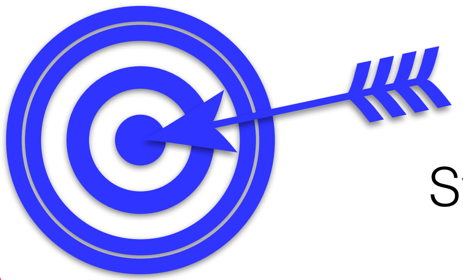
Elizabeth

Important: It's not about the length of the dependency.
(Chomsky 1965, Ross 1967)

Jack

What did Elizabeth think Jack said __what?

✔



Syntactic islands

Elizabeth

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

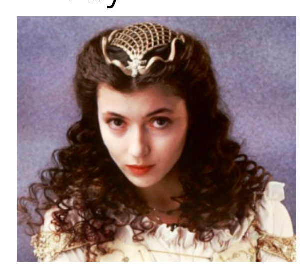
Jack

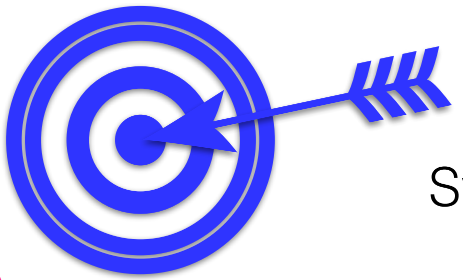
Important: It's not about the length of the dependency.

(Chomsky 1965, Ross 1967)

What did Elizabeth think Jack said Lily saw __what?

Lily





Syntactic islands

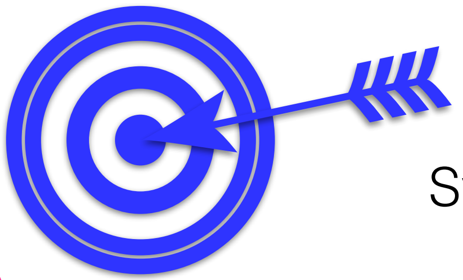
Who does Lily think the kitty for __$_{who}$ is pretty? Subject island

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

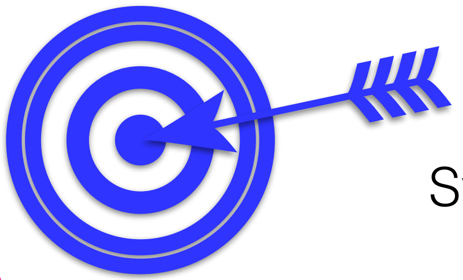
Who does Lily think the kitty for __who is pretty? Subject island

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

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

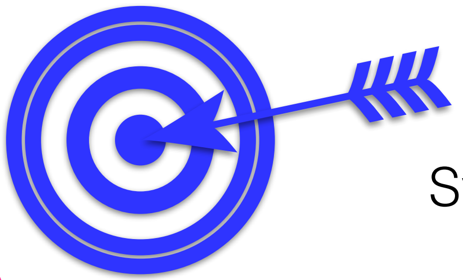
Recall:
The frequency of a lexical item can also affect adult acceptability judgments of potential syntactic islands.

What did Elizabeth say that Jack saw __what?

more frequent

What did Elizabeth whine that Jack saw __what?

less frequent



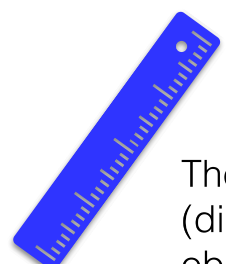
Syntactic islands

Who does Lily think the kitty for $\_\_{who}$ is pretty? Subject island

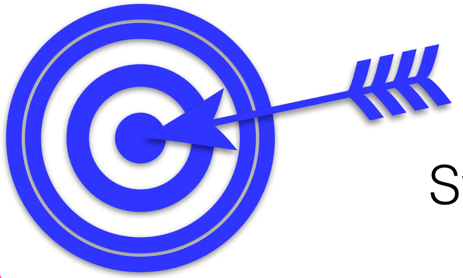
What did Jack make the claim that he bought $\_\_{what}$? Complex NP island

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.



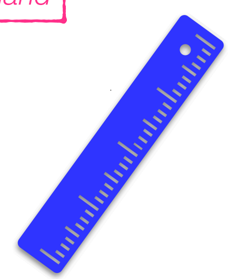
Syntactic islands

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

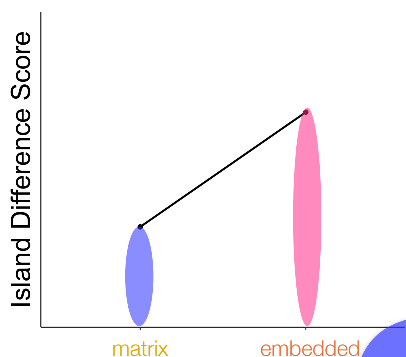
What did Jack make the claim that he bought __what? Complex NP island

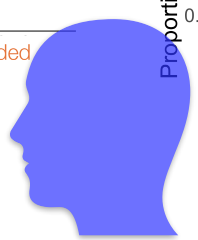
What did Elizabeth wonder whether Jack bought __what? Whether island

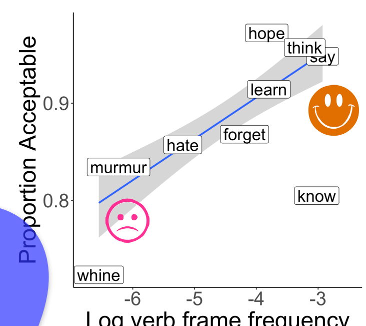
What did Elizabeth worry if Jack bought __what? Adjunct island

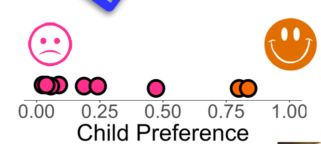


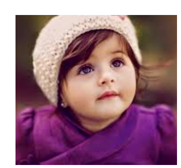
measurable observable behavior

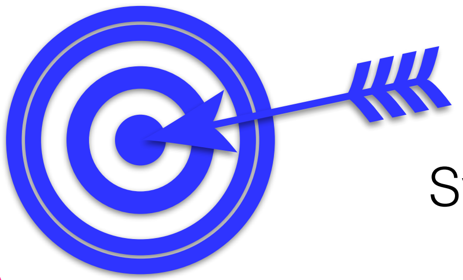




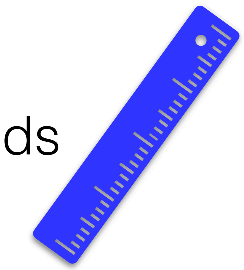


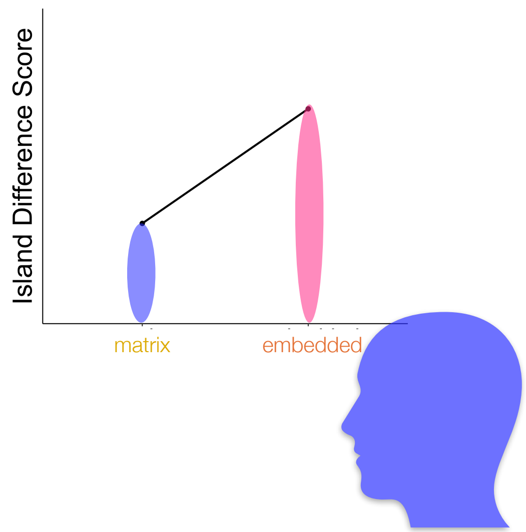


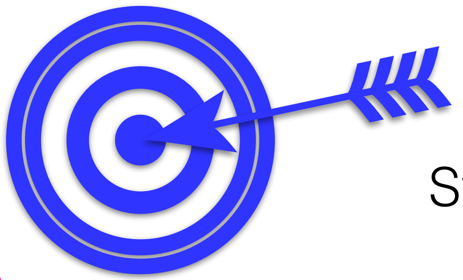




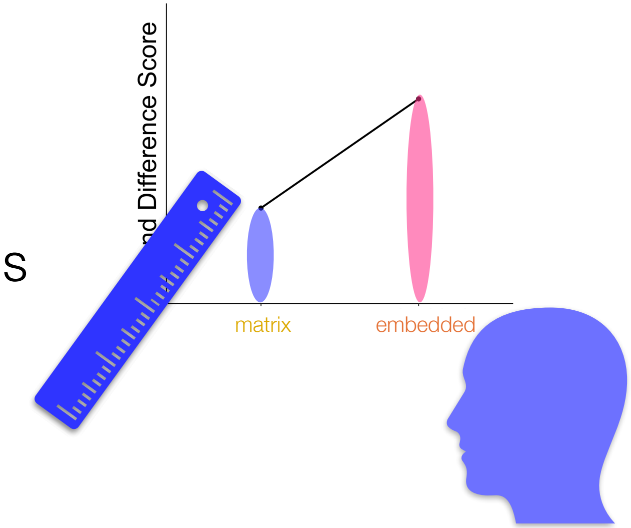
Syntactic islands







Syntactic islands



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*?

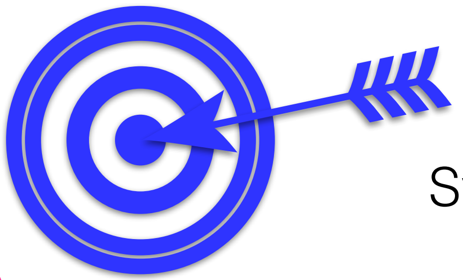
Who [CP… __*who*]?

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

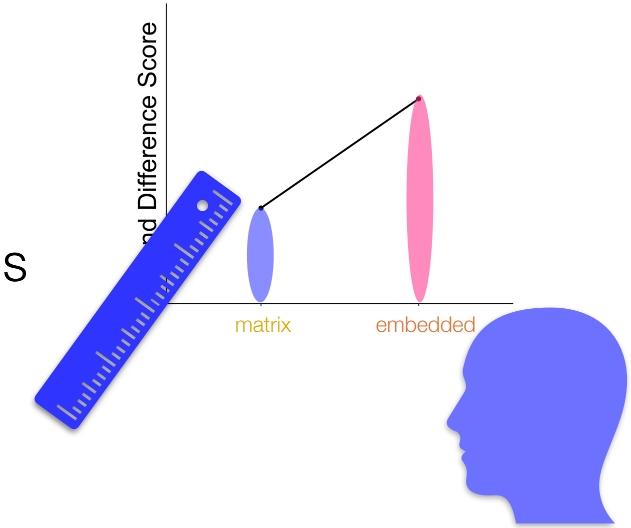
Who [non-island]?

Who [island]?

Sprouse et al. 2012



Syntactic islands



Adult knowledge as measured by acceptability judgment behavior

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

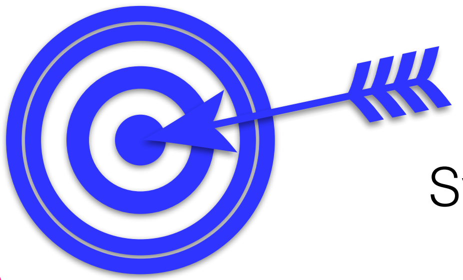
X

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

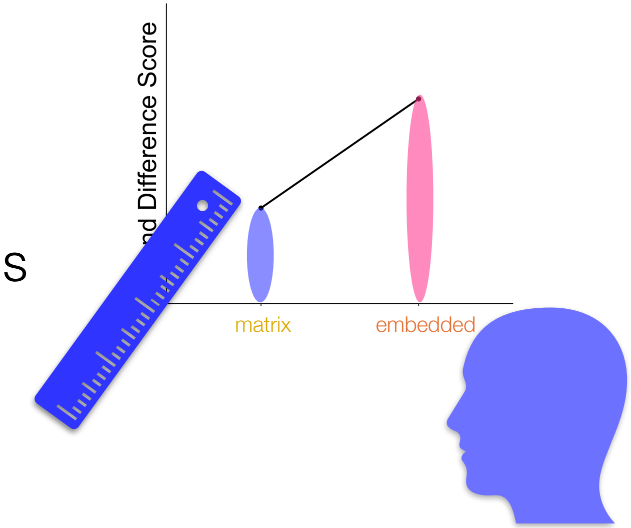
Subject island stimuli

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



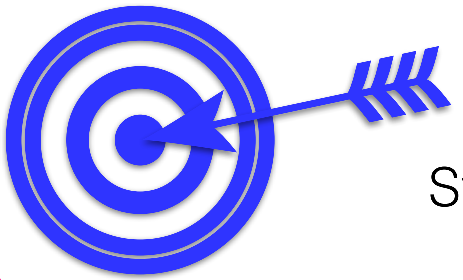
Adult knowledge as measured by acceptability judgment behavior

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

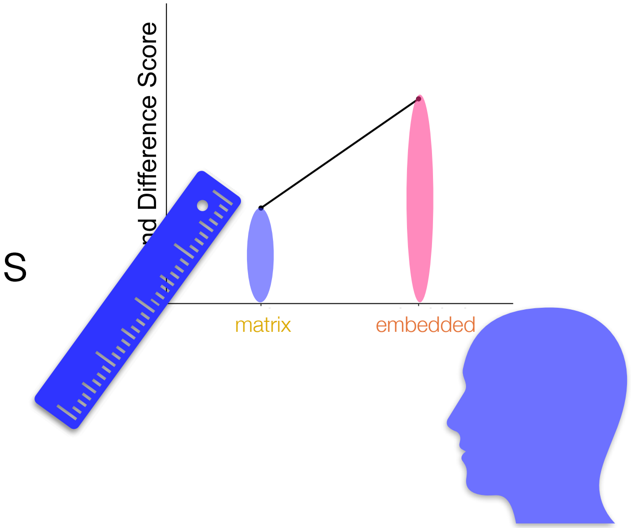
Whether island stimuli

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 knowledge as measured by acceptability judgment behavior

length of dependency
(matrix vs. embedded)

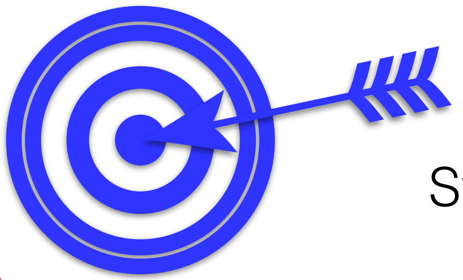
X

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

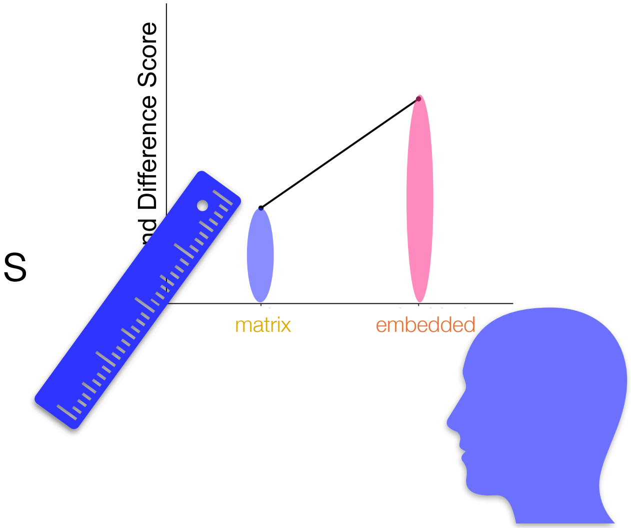
Adjunct island stimuli

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



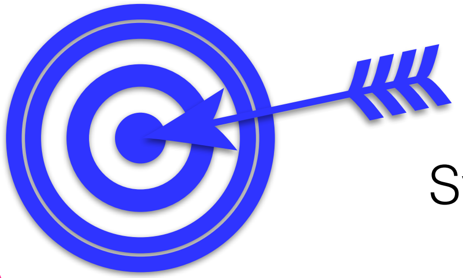
Adult knowledge as measured by acceptability judgment behavior

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

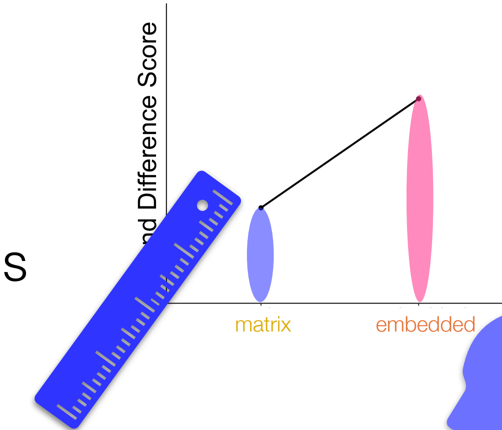
Complex NP island stimuli

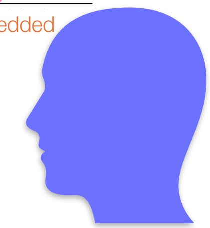
Who __ claimed [that Lily forgot the necklace]? matrix | non-island
What did the teacher claim [that Lily forgot __]? embedded | non-island
Who __ made [the claim that Lily forgot the necklace]? matrix | island
*What did the teacher make [the claim that Lily forgot __]? embedded | island

Sprouse et al. 2012



Syntactic islands





acceptability judgment behavior

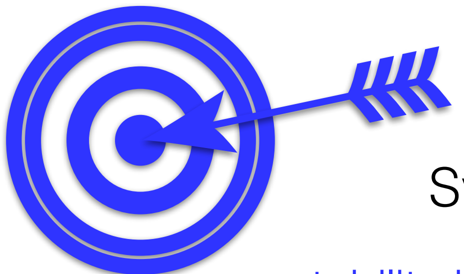
length of dependency
(matrix vs. embedded)

X

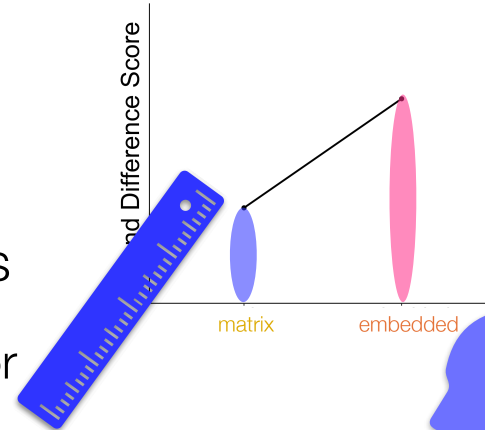
presence of an island structure
(non-island vs. island)

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.

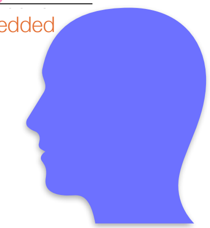
Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior



length of dependency
(matrix vs. embedded)

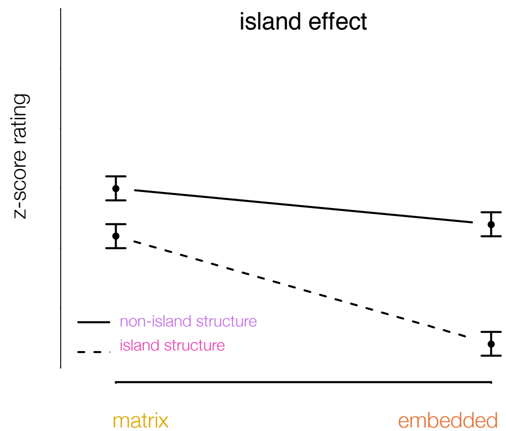
X

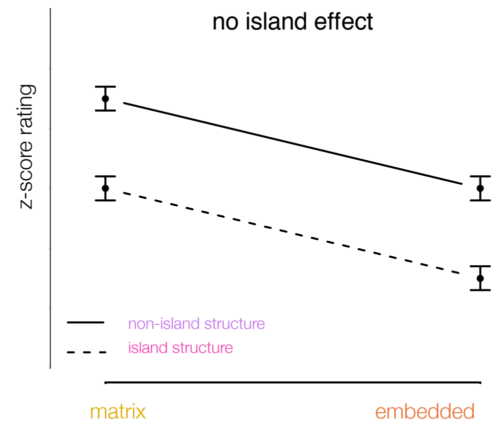
presence of an island structure
(non-island vs. island)

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

Who [non-island]?

Who [island]?

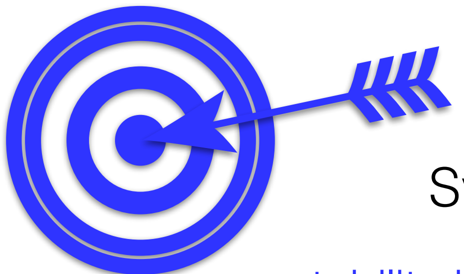




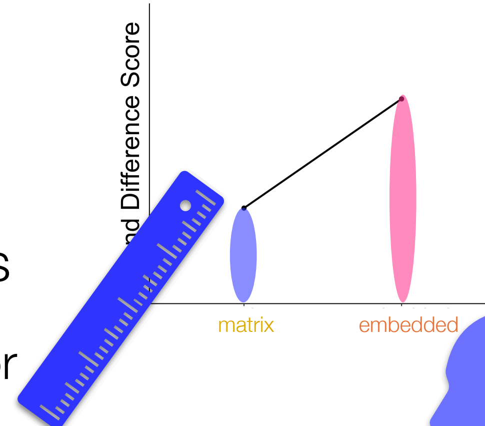
Who __*who*?

Who [CP… __*who*]?

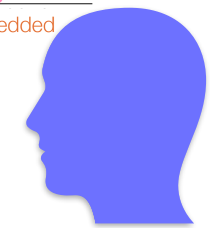
Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior



length of dependency
(matrix vs. embedded)

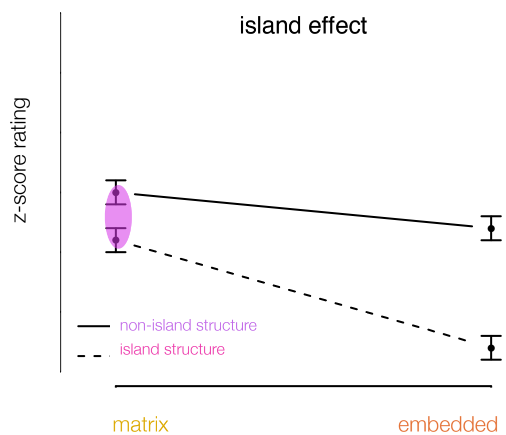
X

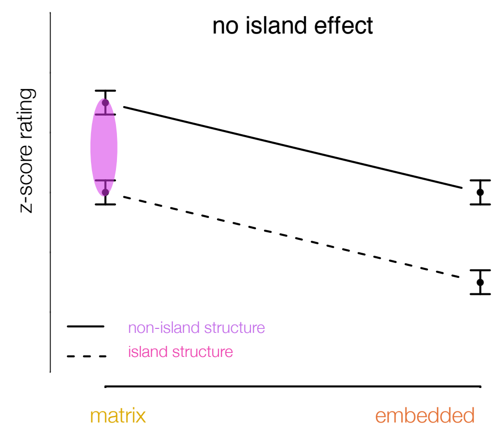
presence of an island structure
(non-island vs. island)

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

Who [non-island]?

Who [island]?

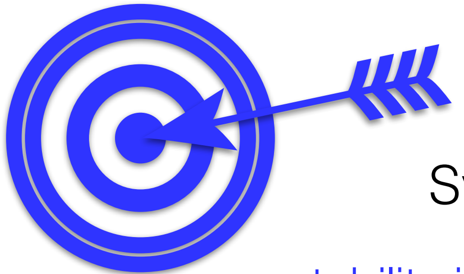




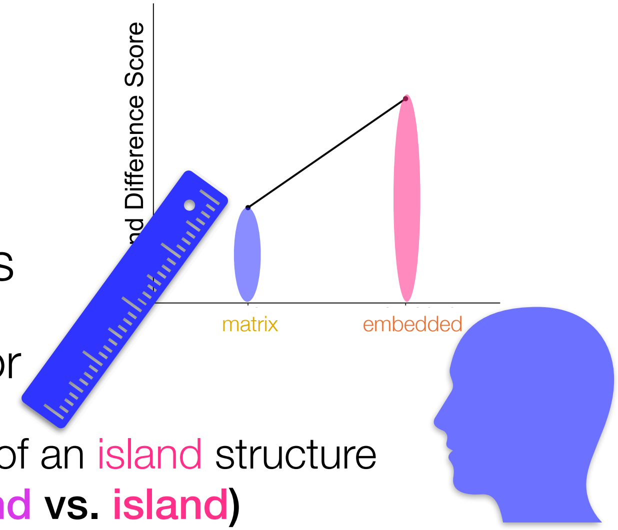
Who *__who*?

Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior

length of dependency
(matrix vs. embedded)

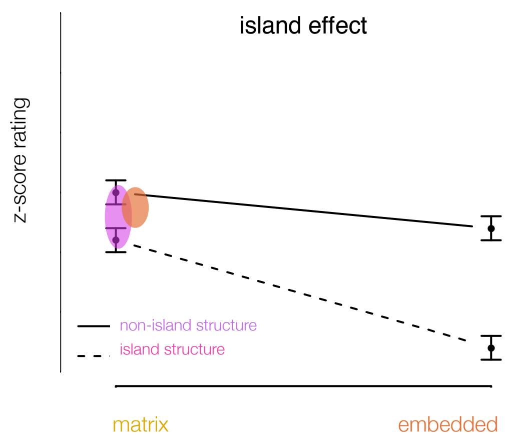
X

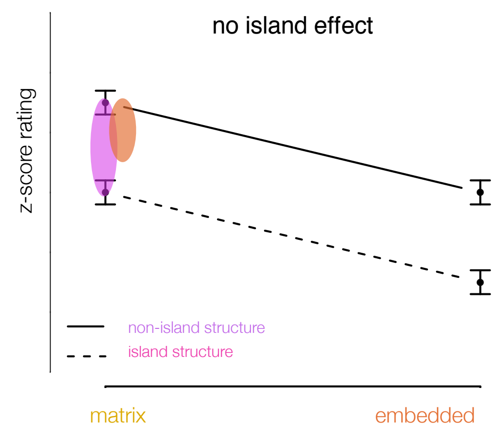
presence of an island structure
(non-island vs. island)

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

Who [non-island]?

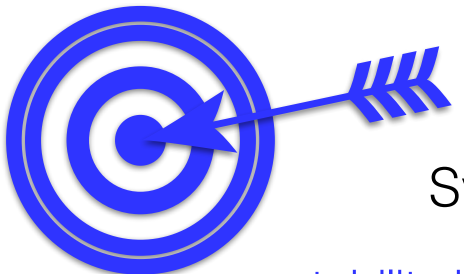
Who [island]?



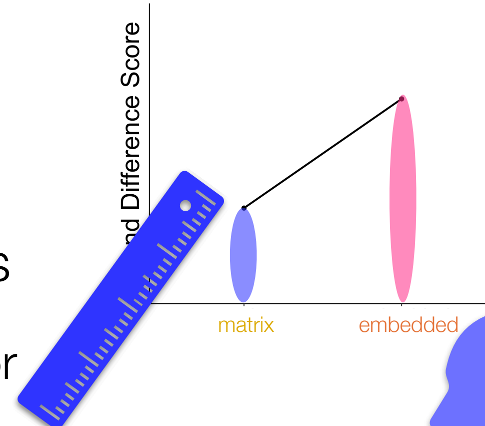


Who *__who*?   *Who* [CP… *__who*]?

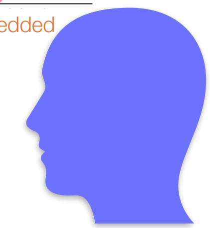
Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior



length of dependency
(matrix vs. embedded)

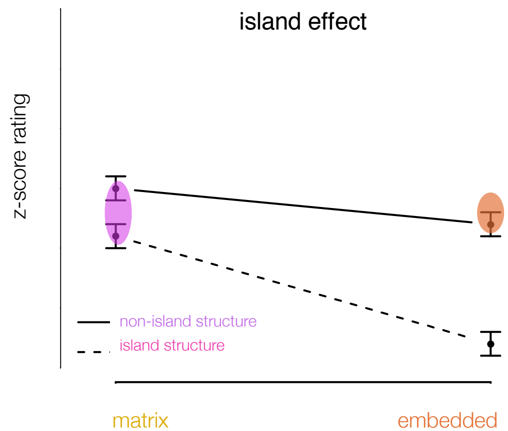
X

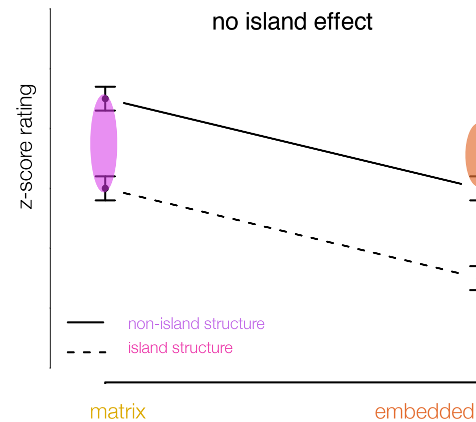
presence of an island structure
(non-island vs. island)

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

Who [non-island]?

Who [island]?

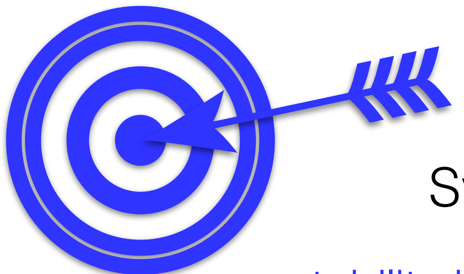




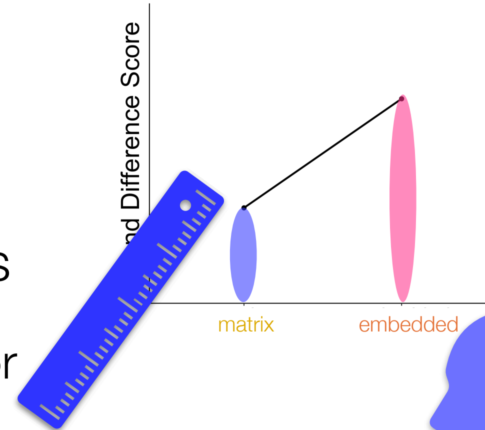
Who *__who*?

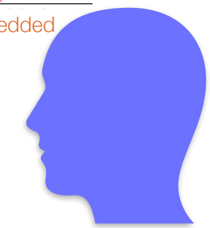
Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands





acceptability judgment behavior

length of dependency
(matrix vs. embedded)

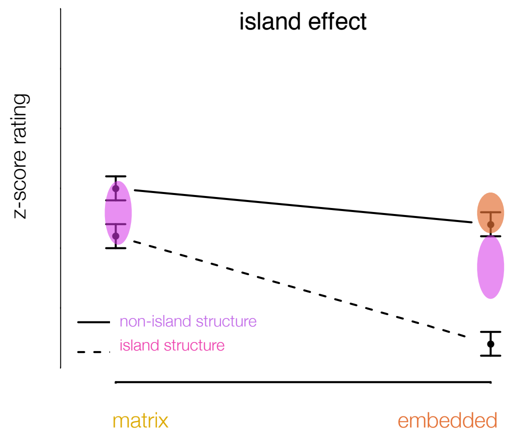
X

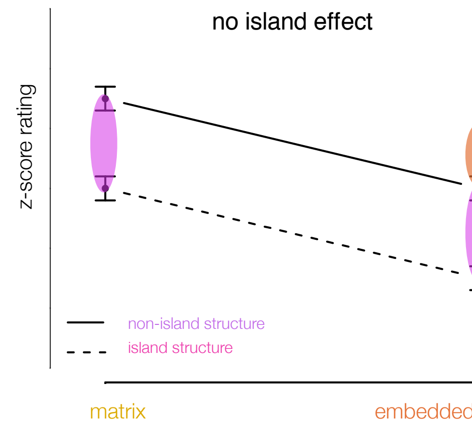
presence of an island structure
(non-island vs. island)

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

Who [non-island]?

Who [island]?

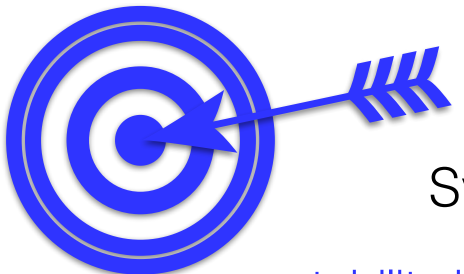




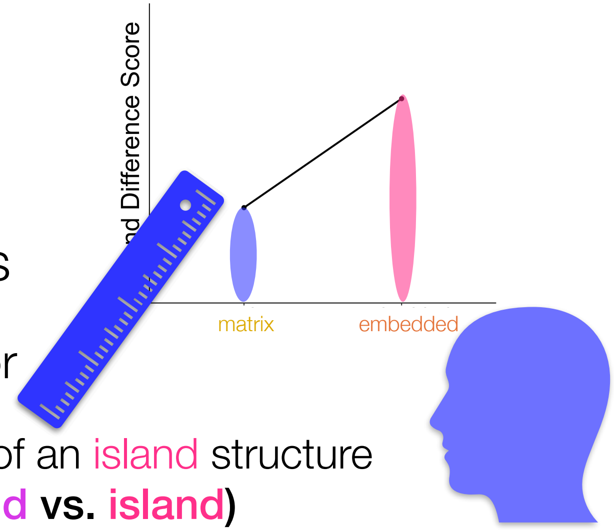
Who *__who*?

Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior

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

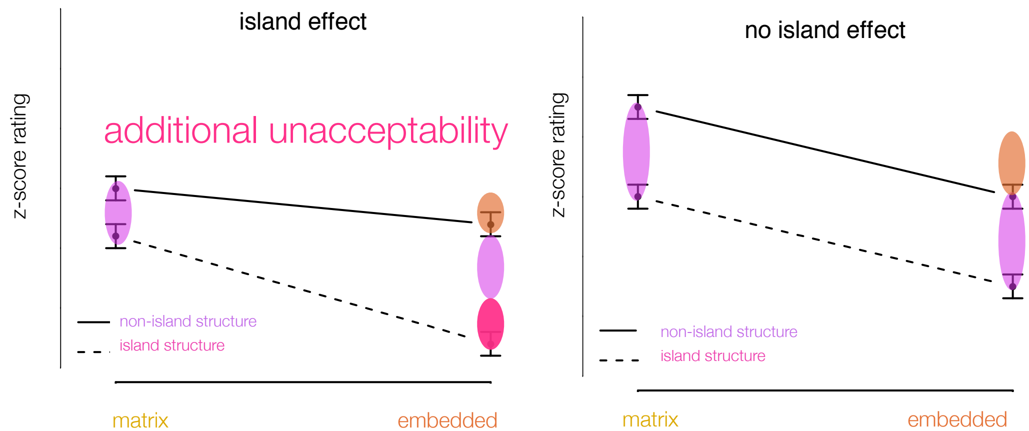
X

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

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

Who [non-island]?

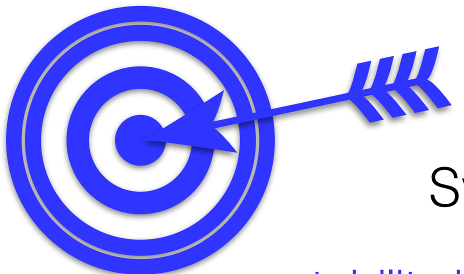
Who [island]?



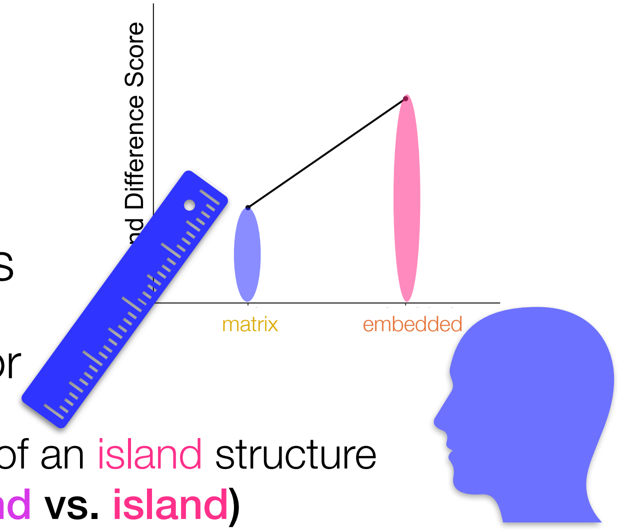
Who *__who*?

Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior

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

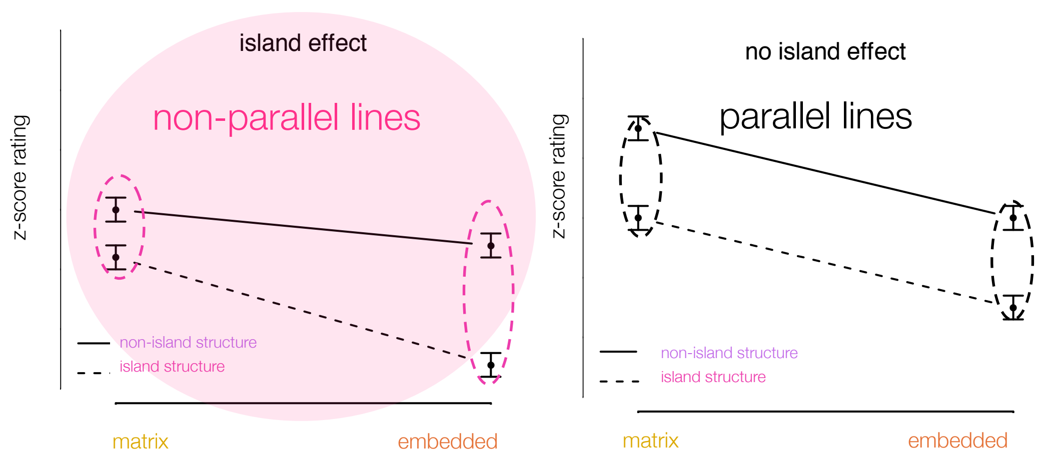
X

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

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

Who [non-island]?

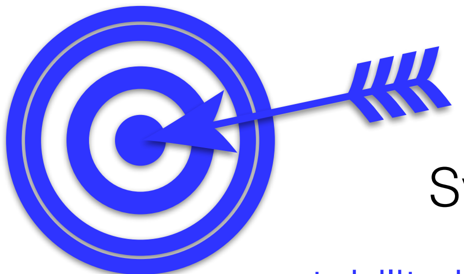
Who [island]?



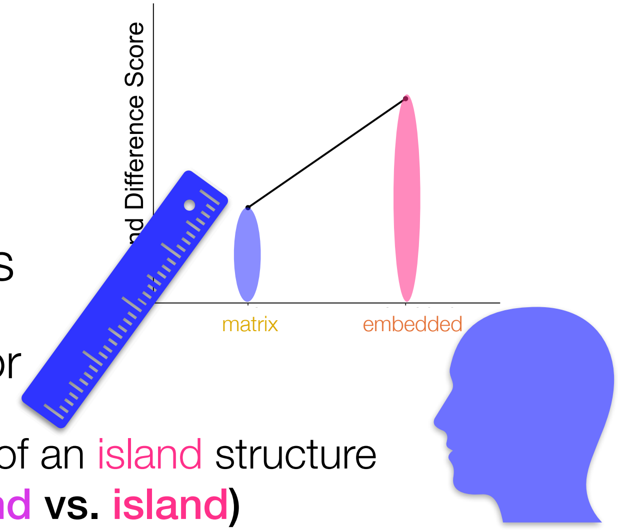
Who *__who*?

Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior

length of dependency
(matrix vs. embedded)

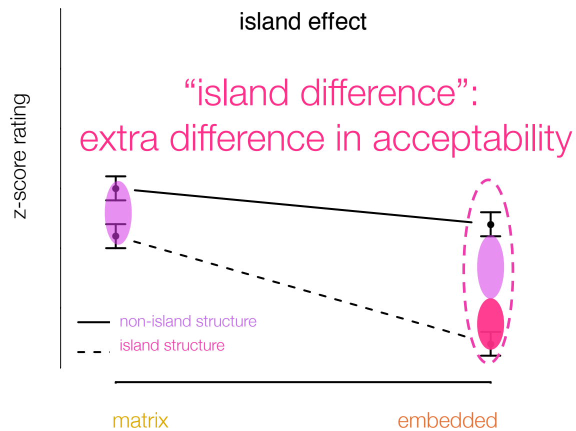
X

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

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

Who [non-island]?

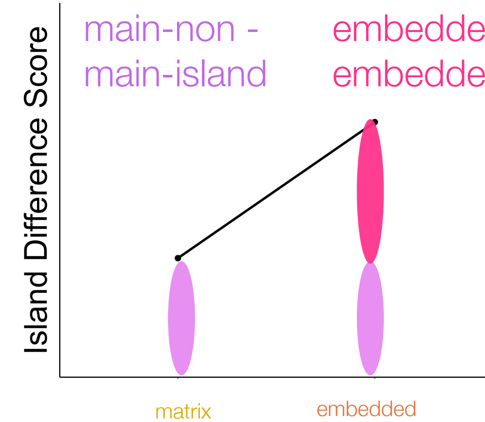
Who [island]?



"island difference":

main-non - main-island

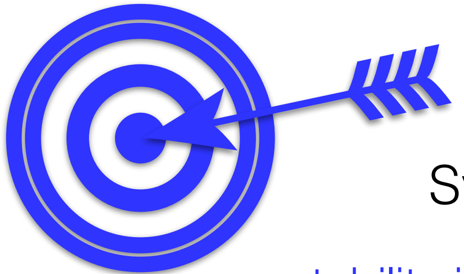
embedded-non - embedded-island



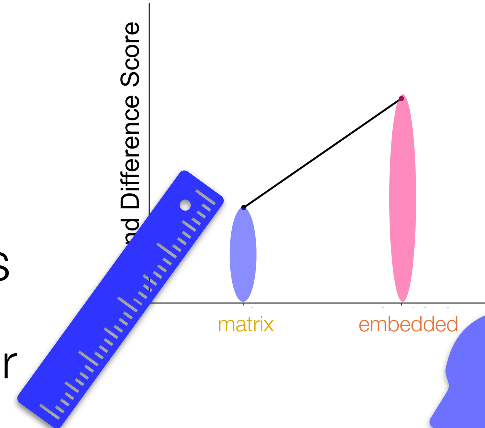
Who __*who*?

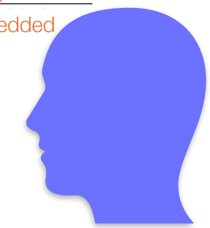
Who [CP… __*who*]?

Sprouse et al. 2012



Syntactic islands





acceptability judgment behavior

length of dependency
(matrix vs. embedded)

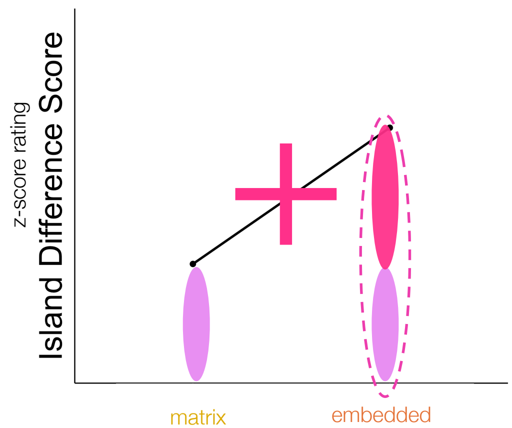
X

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

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

Who [non-island]?

Who [island]?

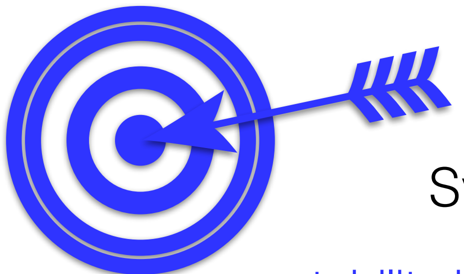


"island difference":
superadditivity = positive difference
(positive slope)

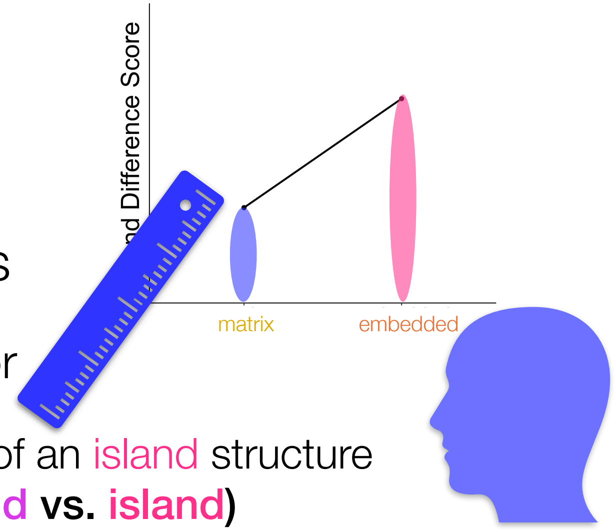
Who *__who*?

Who [CP… *__who*]?

Sprouse et al. 2012



Syntactic islands



acceptability judgment behavior

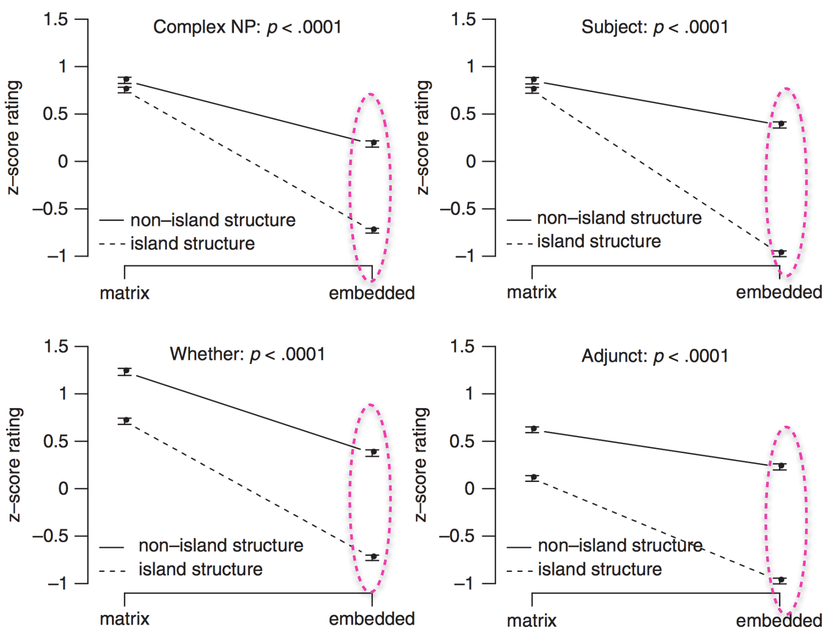
length of dependency
(matrix vs. embedded)

X

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

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

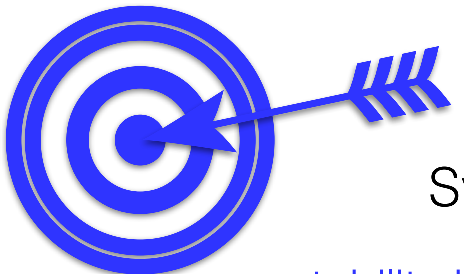
Sprouse et al. (2012): acceptability judgments from 173 adult subjects



(non-parallel lines)

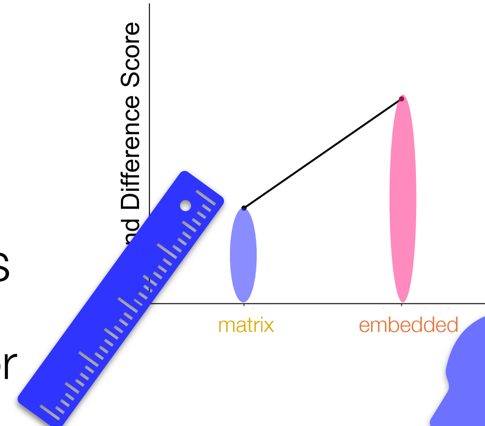
superadditivity for all four island types

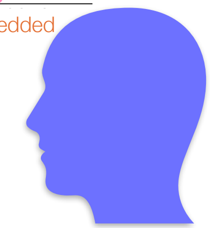
Sprouse et al. 2012



Syntactic islands

acceptability judgment behavior





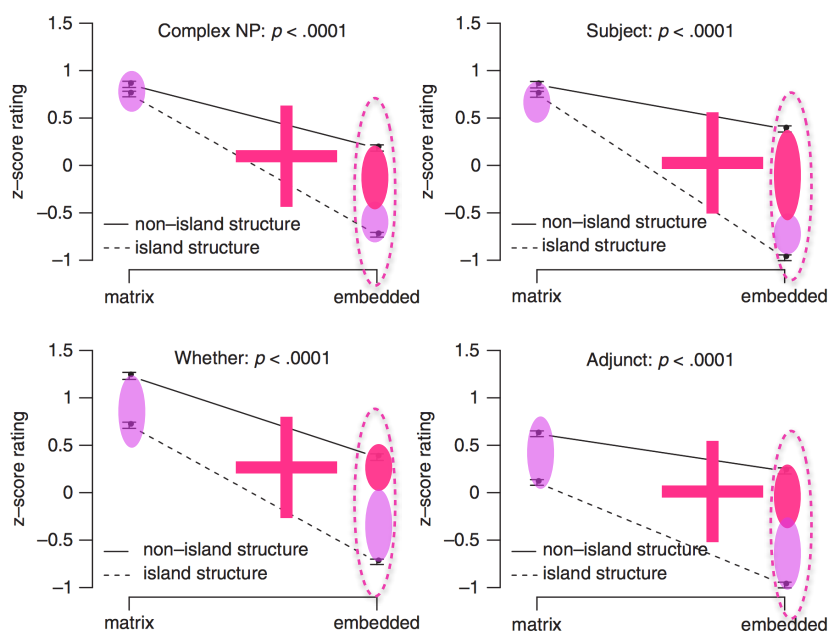
length of dependency
(matrix vs. embedded)

X

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

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

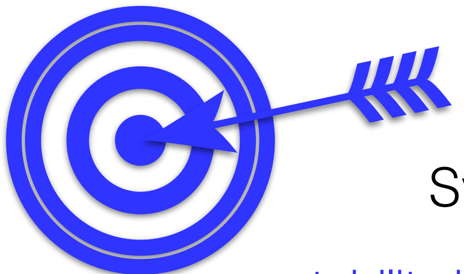
Sprouse et al. (2012): acceptability judgments from 173 adult subjects



(positive difference)

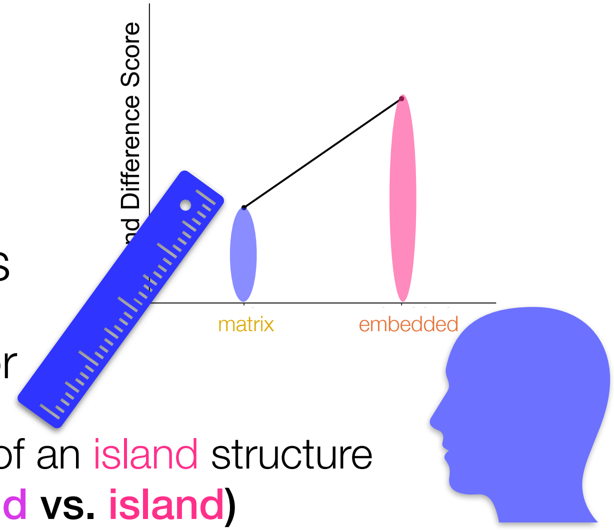
superadditivity for all four island types

Sprouse et al. 2012



Syntactic islands

acceptability judgment behavior



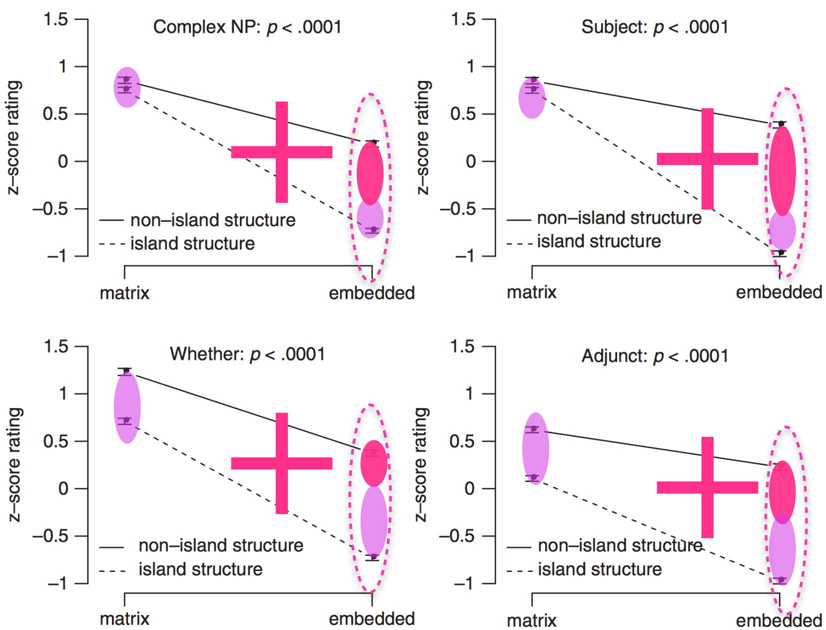
length of dependency
(matrix vs. embedded)

X

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

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

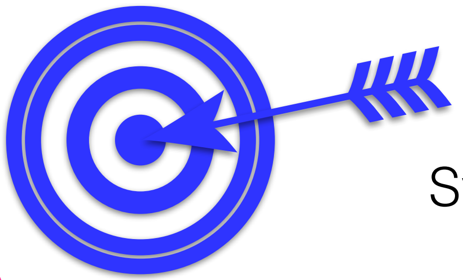
Sprouse et al. (2012): acceptability judgments from 173 adult subjects



superadditivity for all
four island types

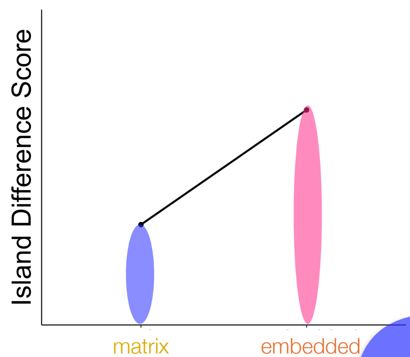
= knowledge that
dependencies
crossing these
island structures
are dispreferred.

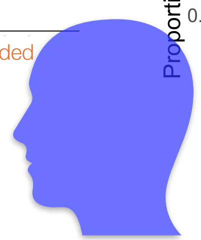
Sprouse et al. 2012

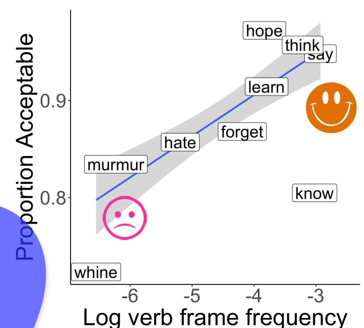


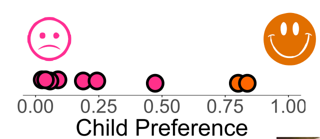
Syntactic islands

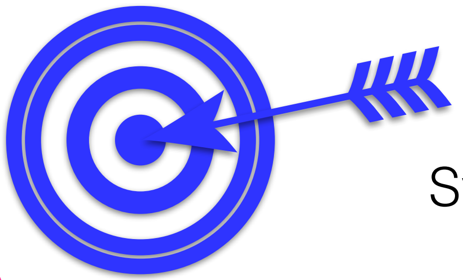
measurable observable behavior



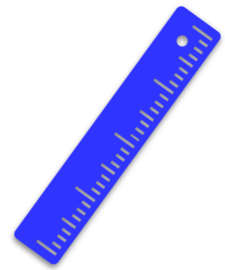


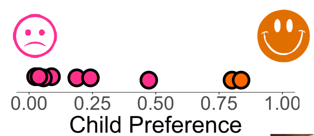


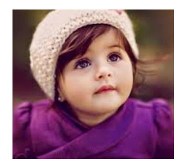


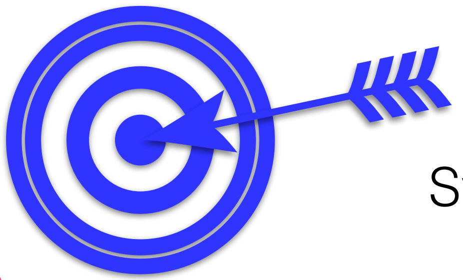


Syntactic islands

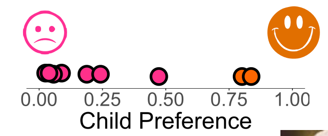


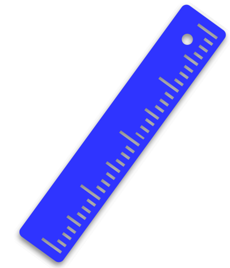






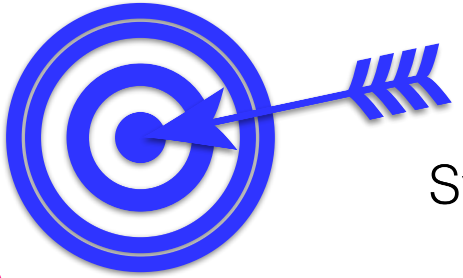
Syntactic islands

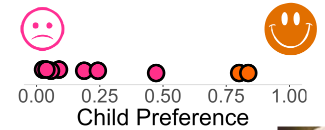


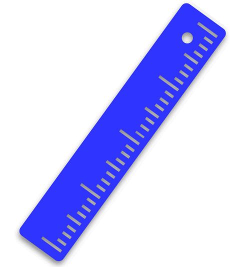


Child knowledge as measured by preferred interpretation behavior

De Villiers et al. 2008:
How do children prefer to interpret potentially ambiguous *wh*-questions?







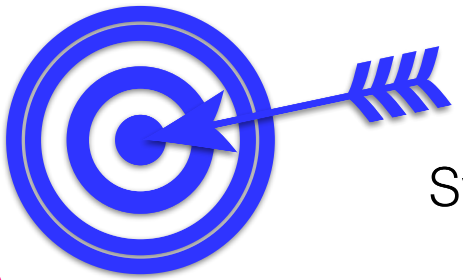
Syntactic islands

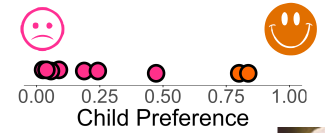
preferred interpretation behavior

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

context

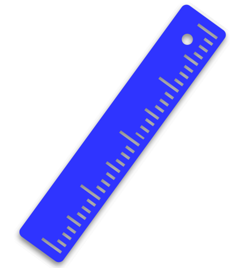
De Villiers et al. 2008





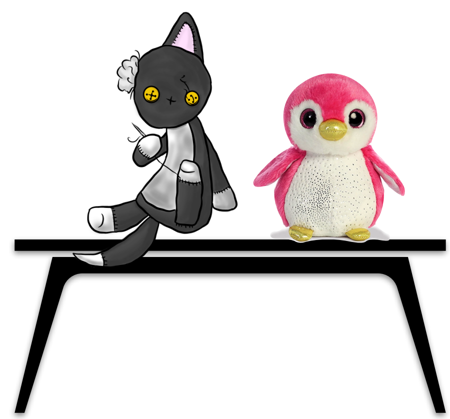
Syntactic islands

preferred interpretation behavior

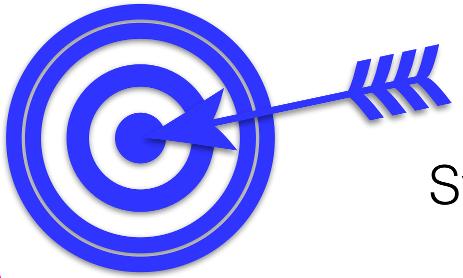


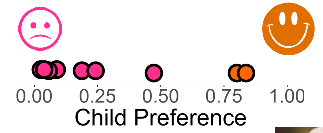
How do children prefer to interpret potentially ambiguous *wh*-questions?

context

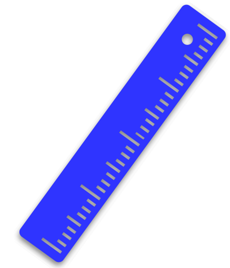


De Villiers et al. 2008





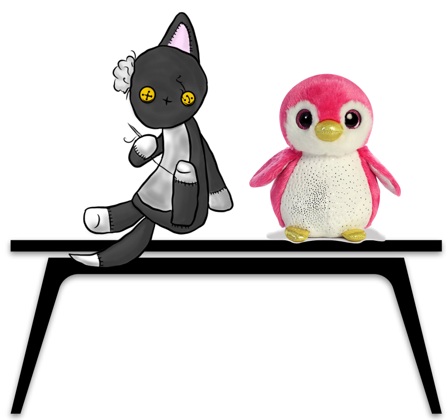
Syntactic islands



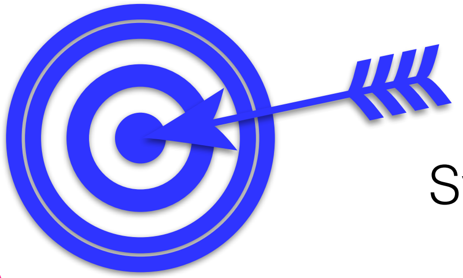
preferred interpretation behavior

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*?

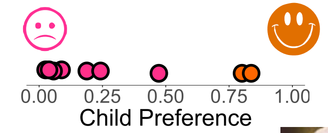


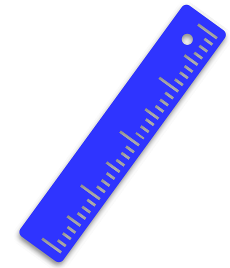
De Villiers et al. 2008



Syntactic islands

preferred interpretation behavior



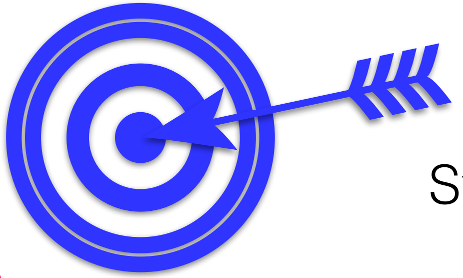


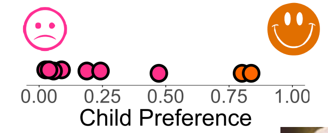
How do children prefer to interpret potentially ambiguous *wh*-questions?

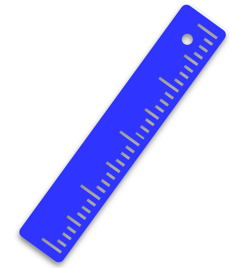
What did the boy [fix the cat *that was lying on the table* [with __*what*]]?

a needle and thread

De Villiers et al. 2008







Syntactic islands

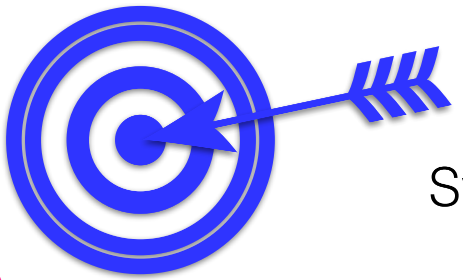
preferred interpretation behavior

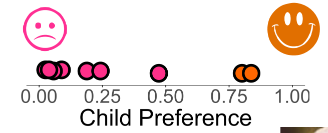
How do children prefer to interpret potentially ambiguous *wh*-questions?

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

a penguin

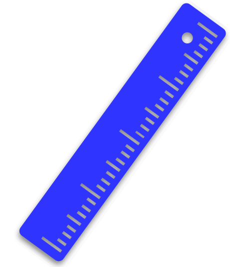
De Villiers et al. 2008





Syntactic islands

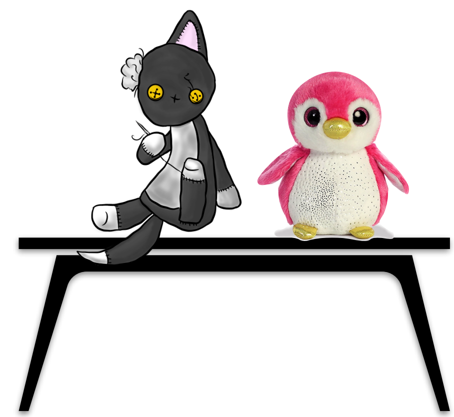
preferred interpretation behavior



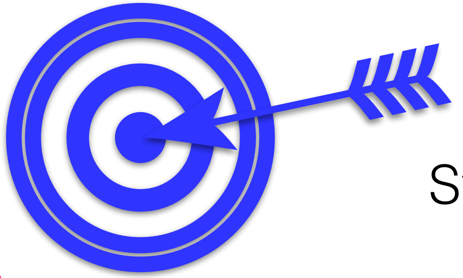
How do children prefer to interpret potentially ambiguous *wh*-questions?

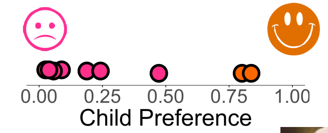
What did the boy fix the cat that was lying on the table with __*what*?

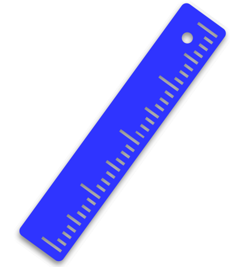
children strongly prefer this interpretation



De Villiers et al. 2008







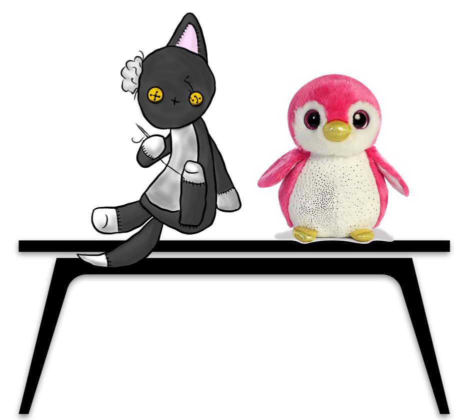
Syntactic islands

preferred interpretation behavior

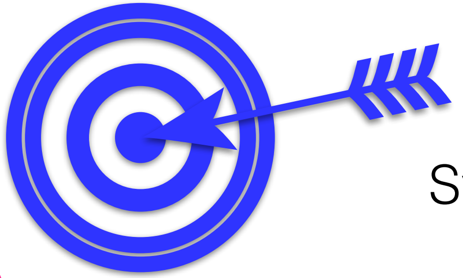
How do children prefer to interpret potentially ambiguous *wh*-questions?

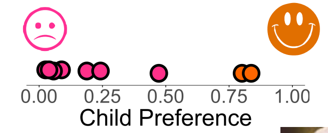
What did the boy fix the cat that was lying on the table with __*what*?

...and strongly disprefer this interpretation



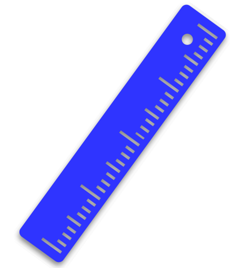
De Villiers et al. 2008





Syntactic islands

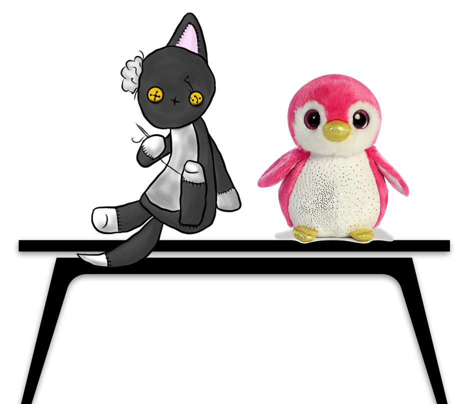
preferred interpretation behavior



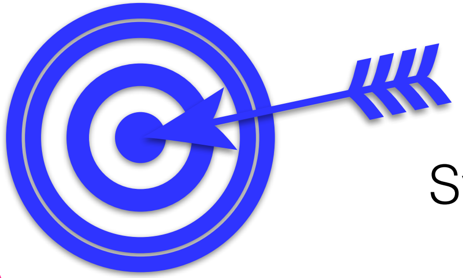
How do children prefer to interpret potentially ambiguous *wh*-questions?

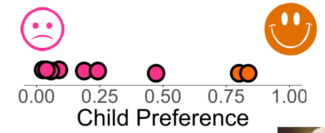
What did the boy [fix [the cat [that [was [lying [on [the table [with __*what*]]]]]]]]?

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



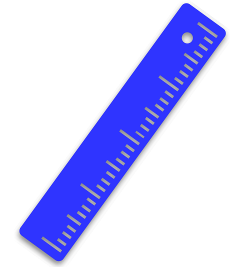
De Villiers et al. 2008





Syntactic islands

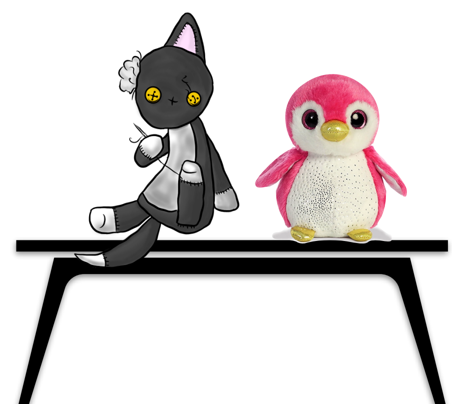
preferred interpretation behavior



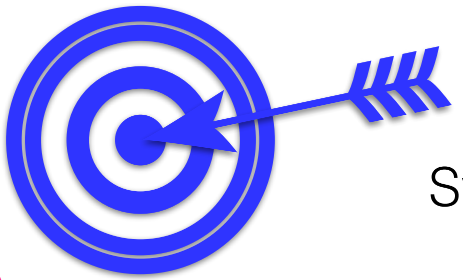
How do children prefer to interpret potentially ambiguous *wh*-questions?

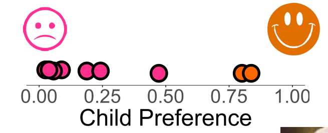
What did the boy [fix [NP the cat [that [was [lying [on [the table [with __*what*]]]]]]]]?

…which is a dependency that
crosses a Complex NP.

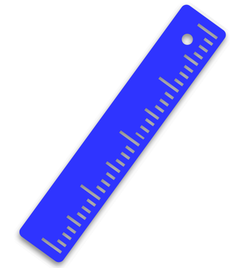


De Villiers et al. 2008





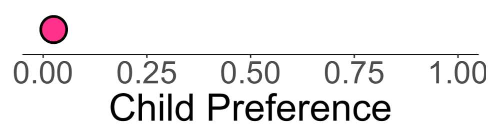
Syntactic islands

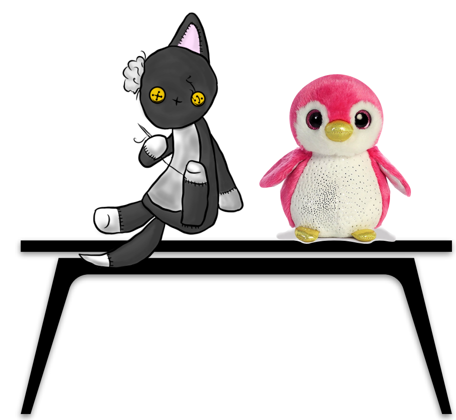


preferred interpretation behavior

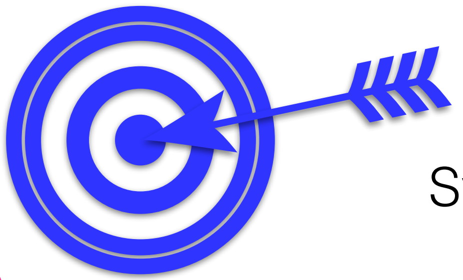
How do children prefer to interpret potentially ambiguous *wh*-questions?

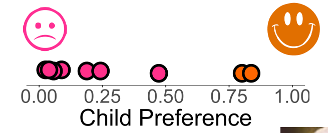
What did the boy [fix [$_{NP}$ the cat [that [was [lying [on [the table [with __*what*]]]]]]]]?

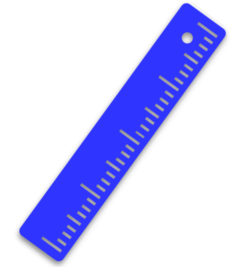




De Villiers et al. 2008







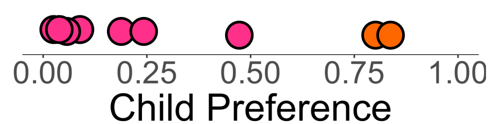
Syntactic islands

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 __?

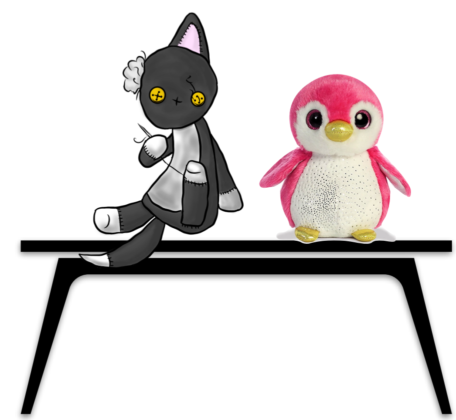
How did the mom learn what to bake __?

How did the girl ask where to ride __?

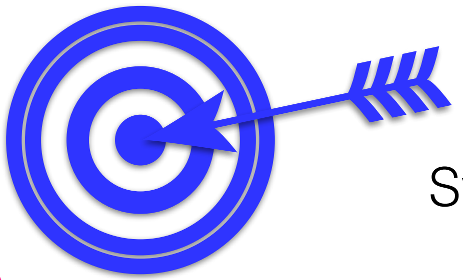
How did the boy who sneezed drink the milk ___?

What did the mother say she bought __?

Who did the policewoman help to call __?

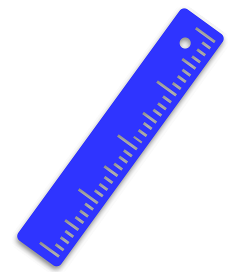


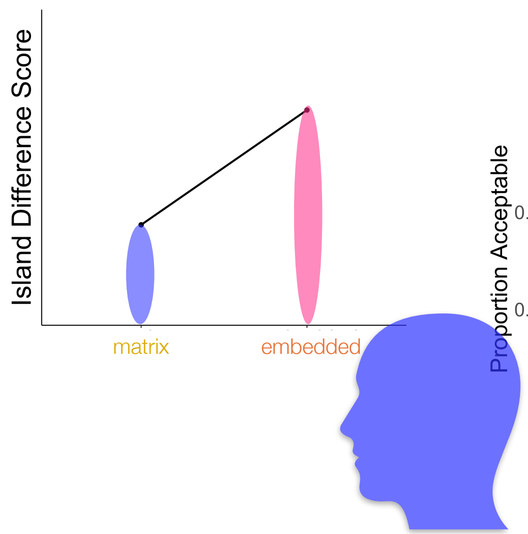
De Villiers et al. 2008

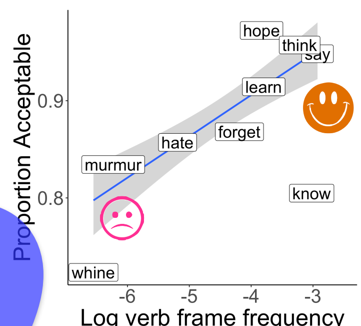


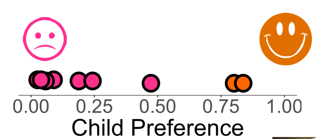
Syntactic islands

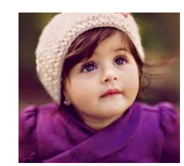
measurable observable behavior

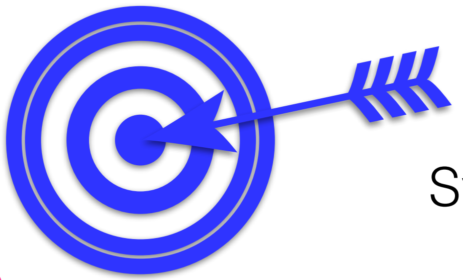






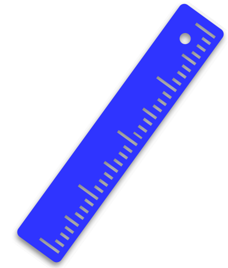


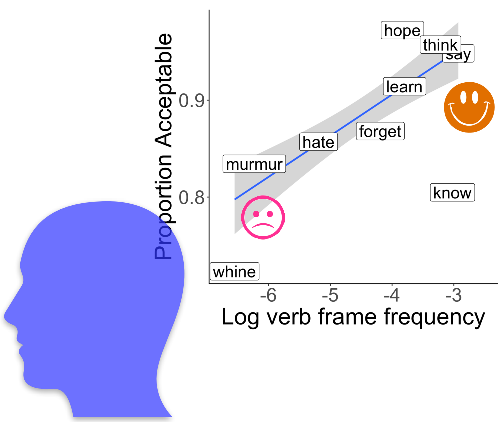


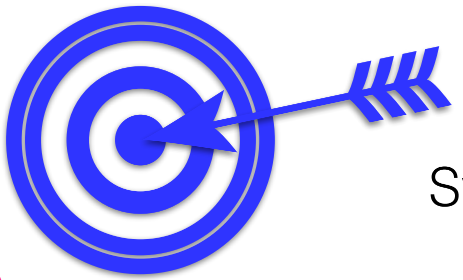


Syntactic islands

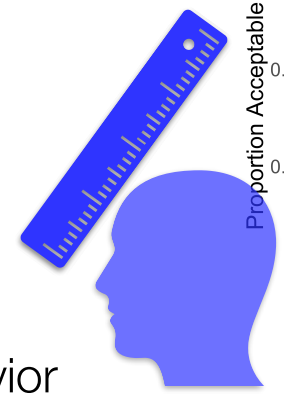
measurable observable behavior

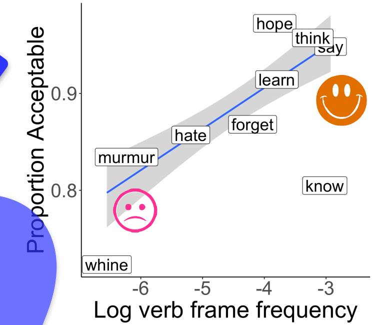






Syntactic islands





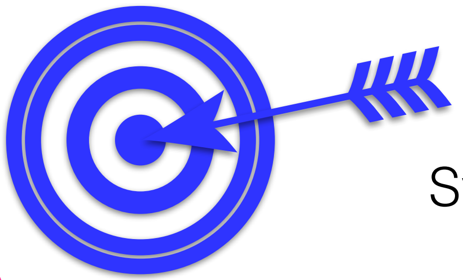
Adult acceptability judgment behavior

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

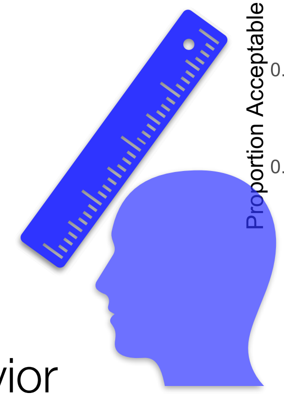
What did she think [that he saw __]?
What did she say [that he saw __]?

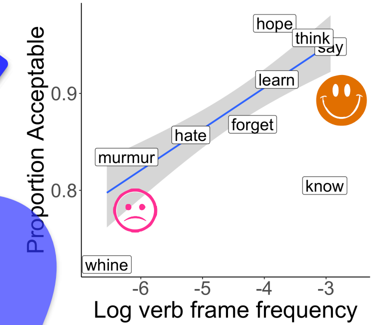
What did she whine [that he saw __]?
What did she mumble [that he saw __]?

Liu et al. 2019, 2022



Syntactic islands





Adult acceptability judgment behavior

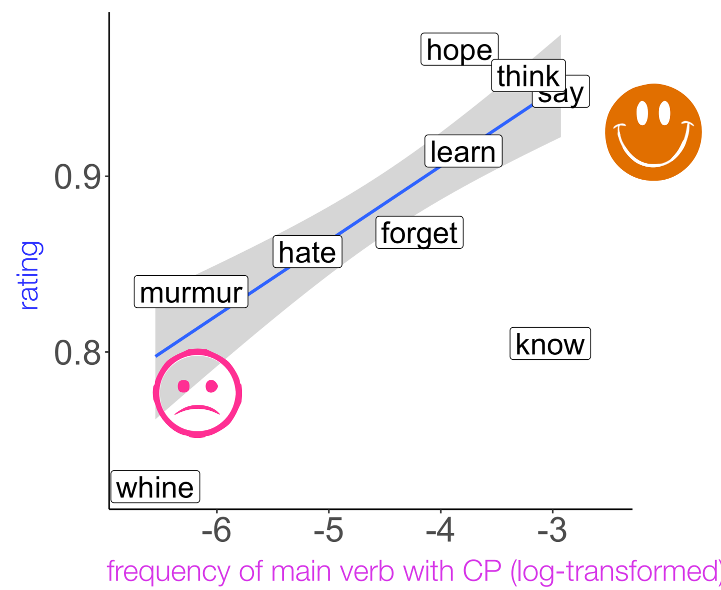
…can depend on the lexical item in the main verb

What did she VERB [that he saw __]?

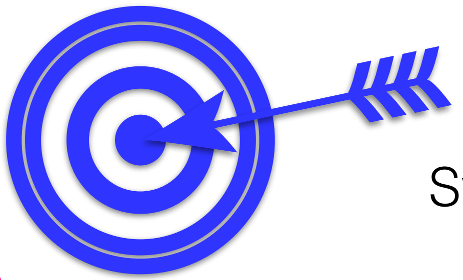
What did she think [that he saw __]?
What did she say [that he saw __]?

CP

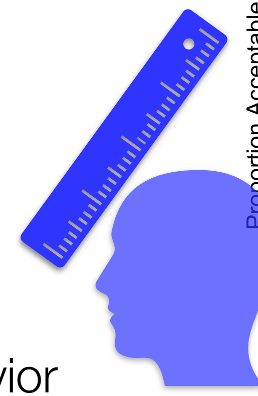
What did she whine [that he saw __]?
What did she mumble [that he saw __]?

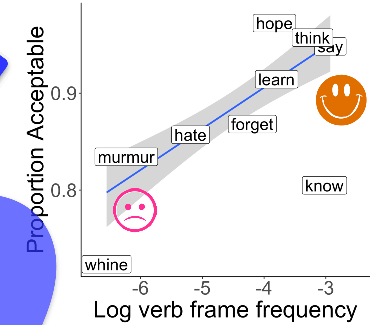


Liu et al. 2019, 2022



Syntactic islands



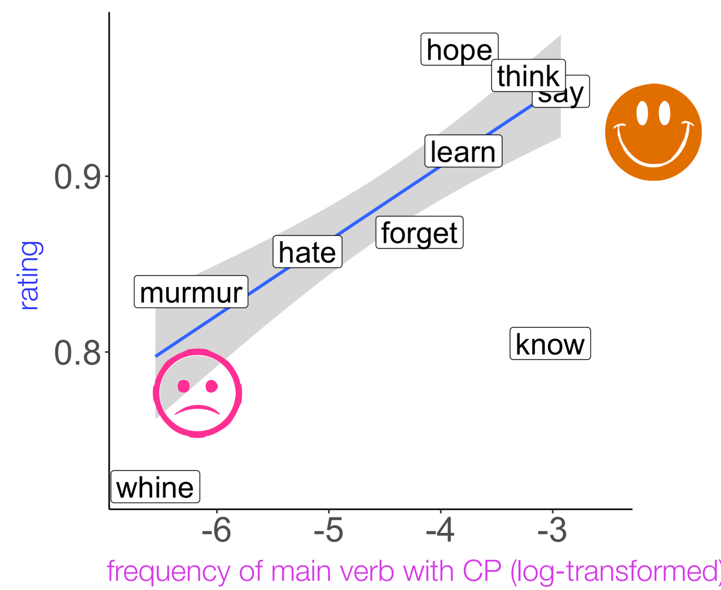


Adult acceptability judgment behavior

…can depend on the lexical item in the main verb

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

What did she VERB [that he saw __]?

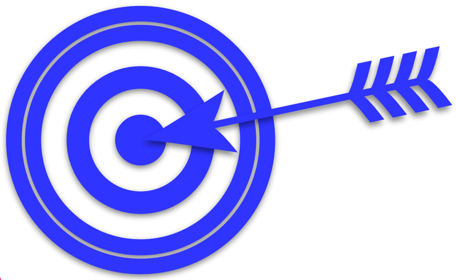


Liu et al. 2019, 2022

Syntactic islands

Adult & child judgments

= behavioral target outcome



What *__what?*

Subject island

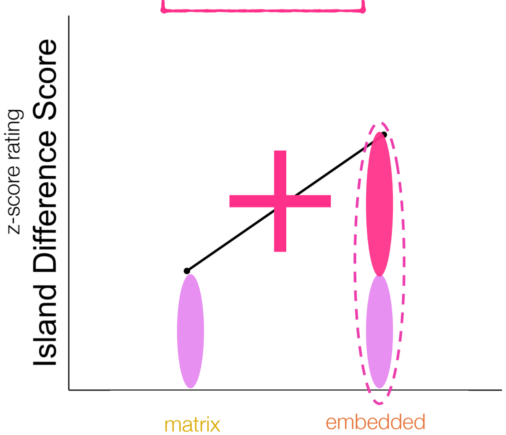
Complex NP island

Whether island

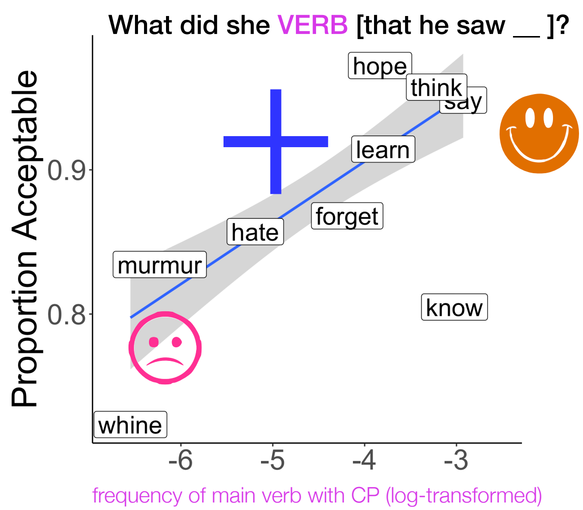
Adjunct island

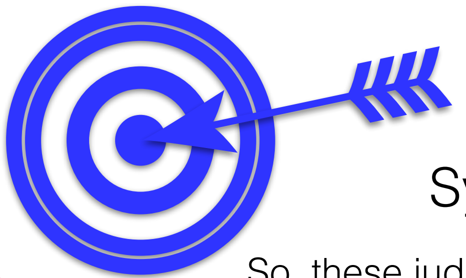
Complex NP + other *wh*-dependencies

De Villiers et al. 2008



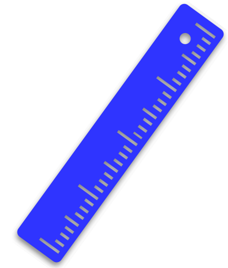
Sprouse et al. 2012

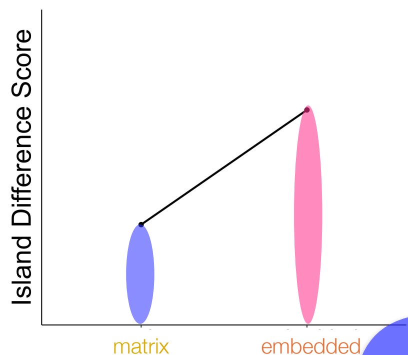


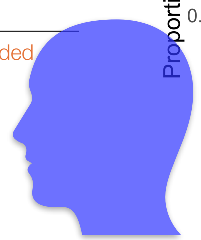


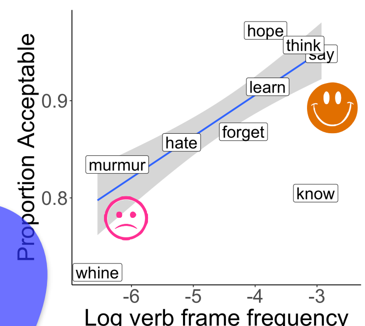
Syntactic islands

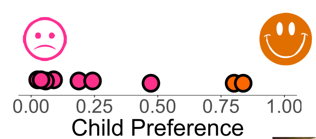
So, these judgments and (dis)preferences can serve as a target for successful acquisition — an outcome we can measure.

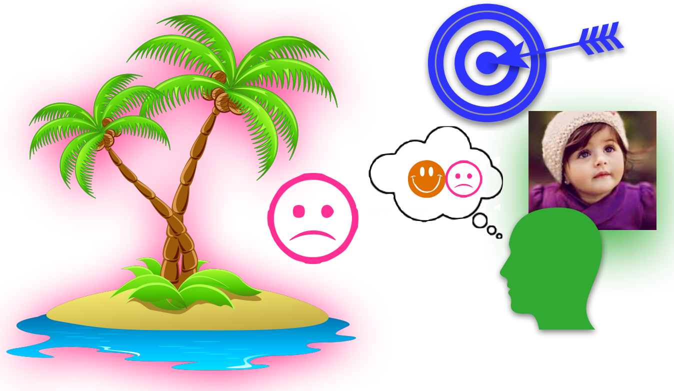








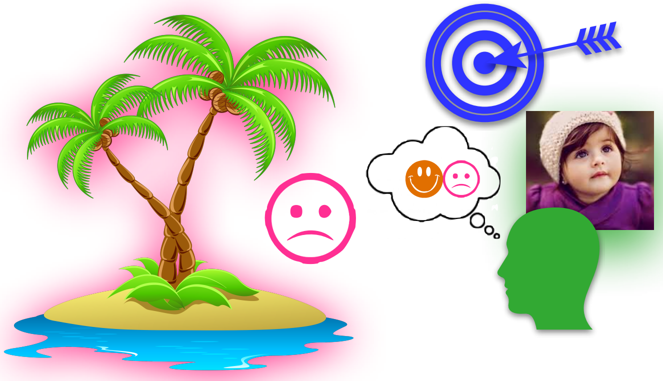




Syntactic islands

How long do children have to learn?

Syntactic islands



How long do children have to learn?

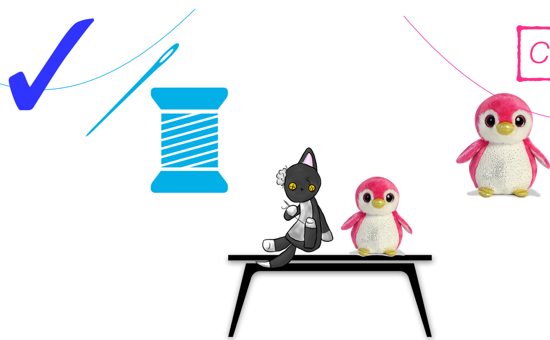
De Villiers et al. 2008:
Data from four-year-olds.

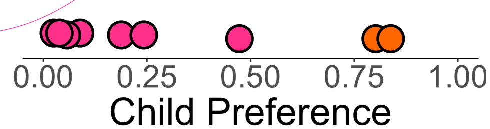
What __*what*?

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

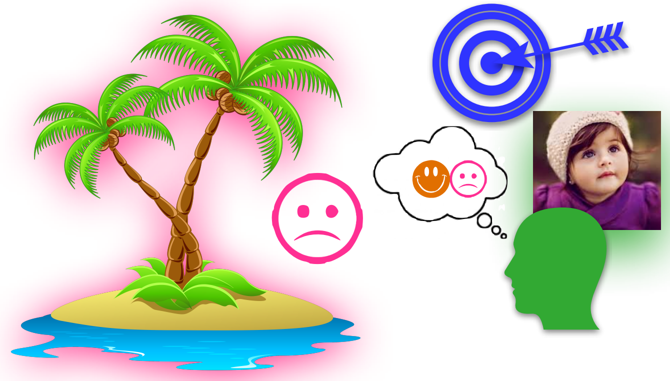
Complex NP

+ other *wh*-dependencies





Syntactic islands



How long do children have to learn?

So input through age four.

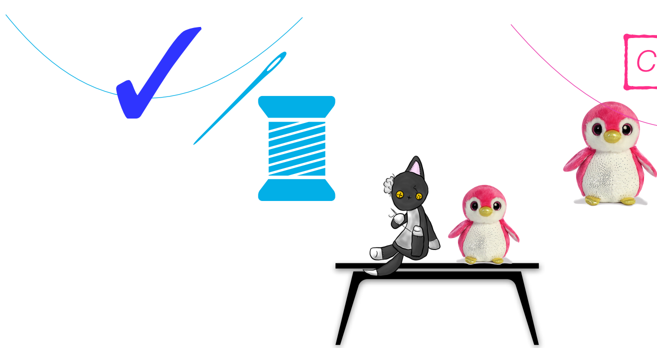
(<60 months)

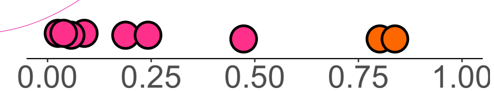
What *__what*?

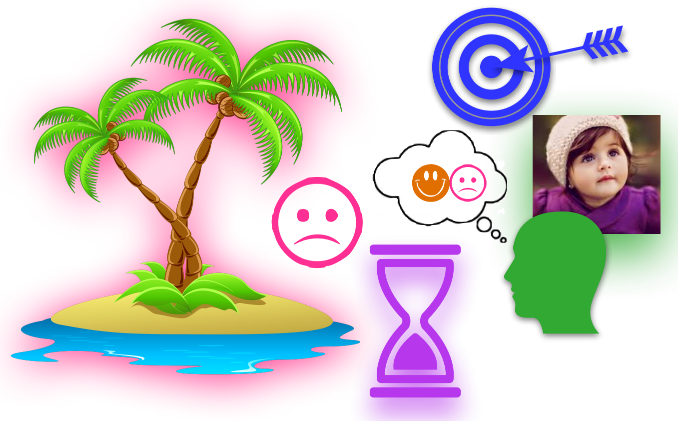
What [NP [CP *__what*]]?

Complex NP

+ other *wh*-dependencies

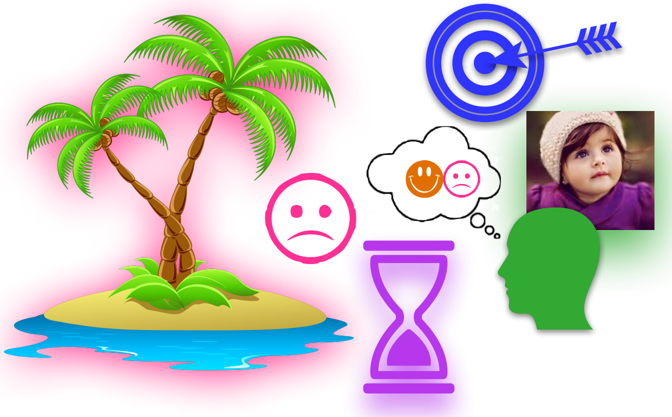






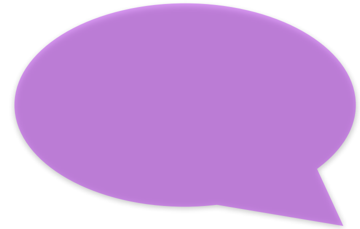
Syntactic islands

What input do children get?



Syntactic islands

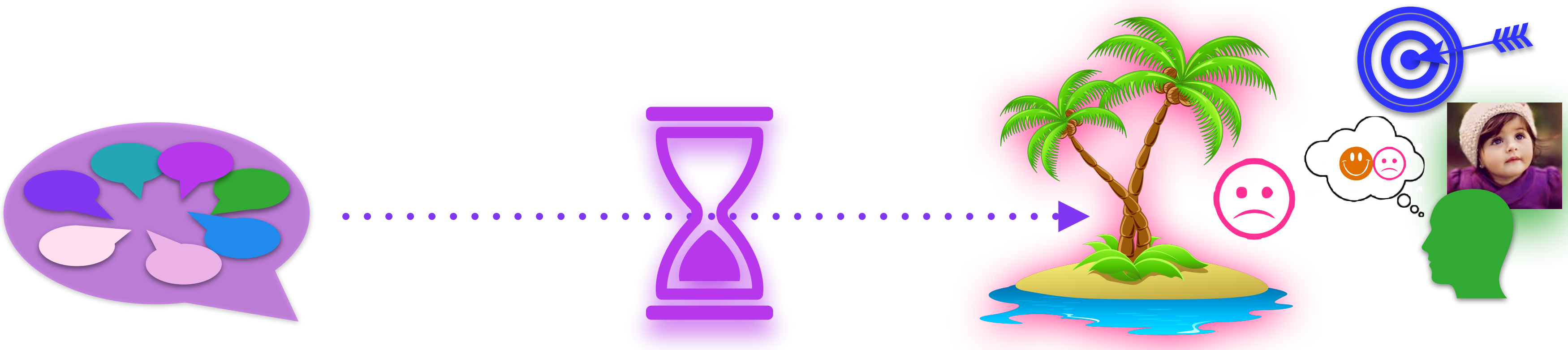
What input do children get?



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

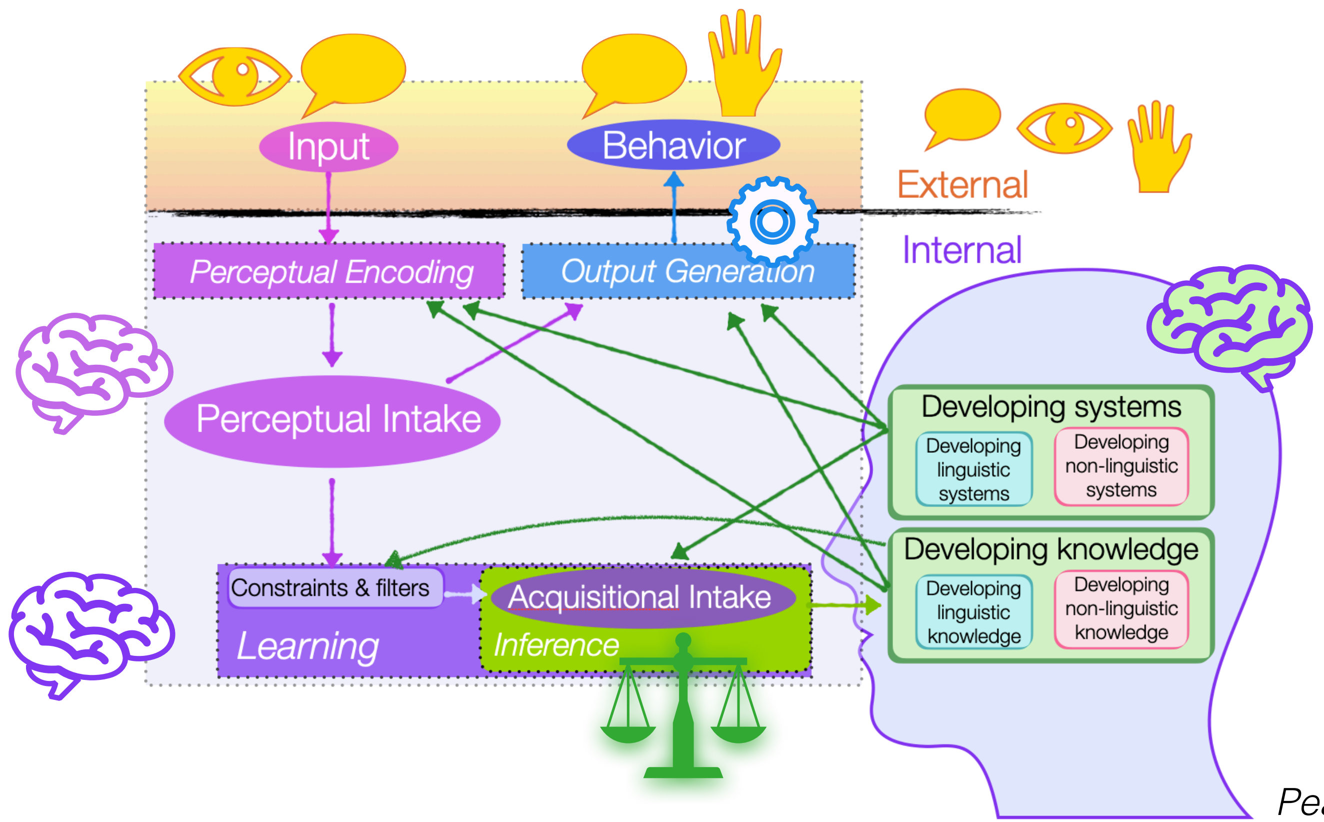
Syntactic islands

This is the acquisition problem



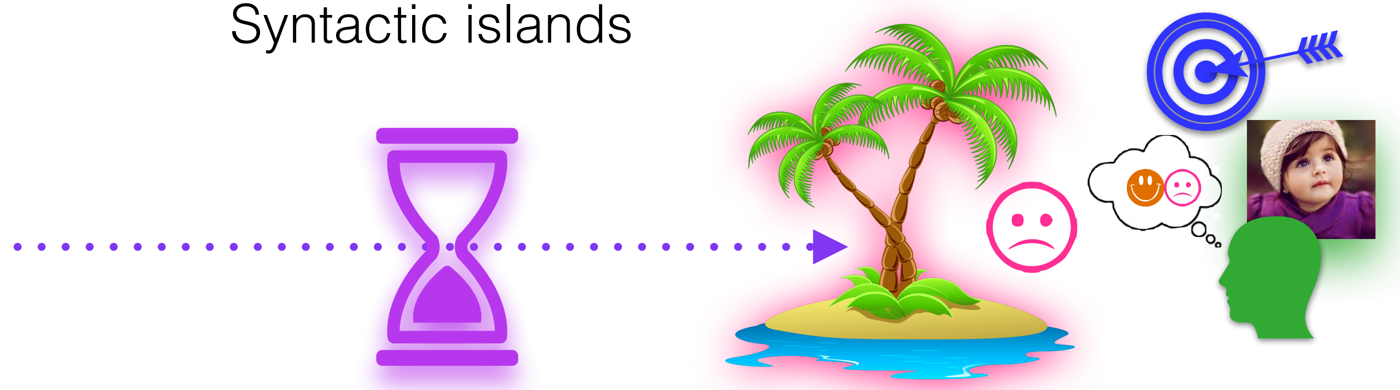
Syntactic islands

Let's map this to how we characterized the acquisition process before

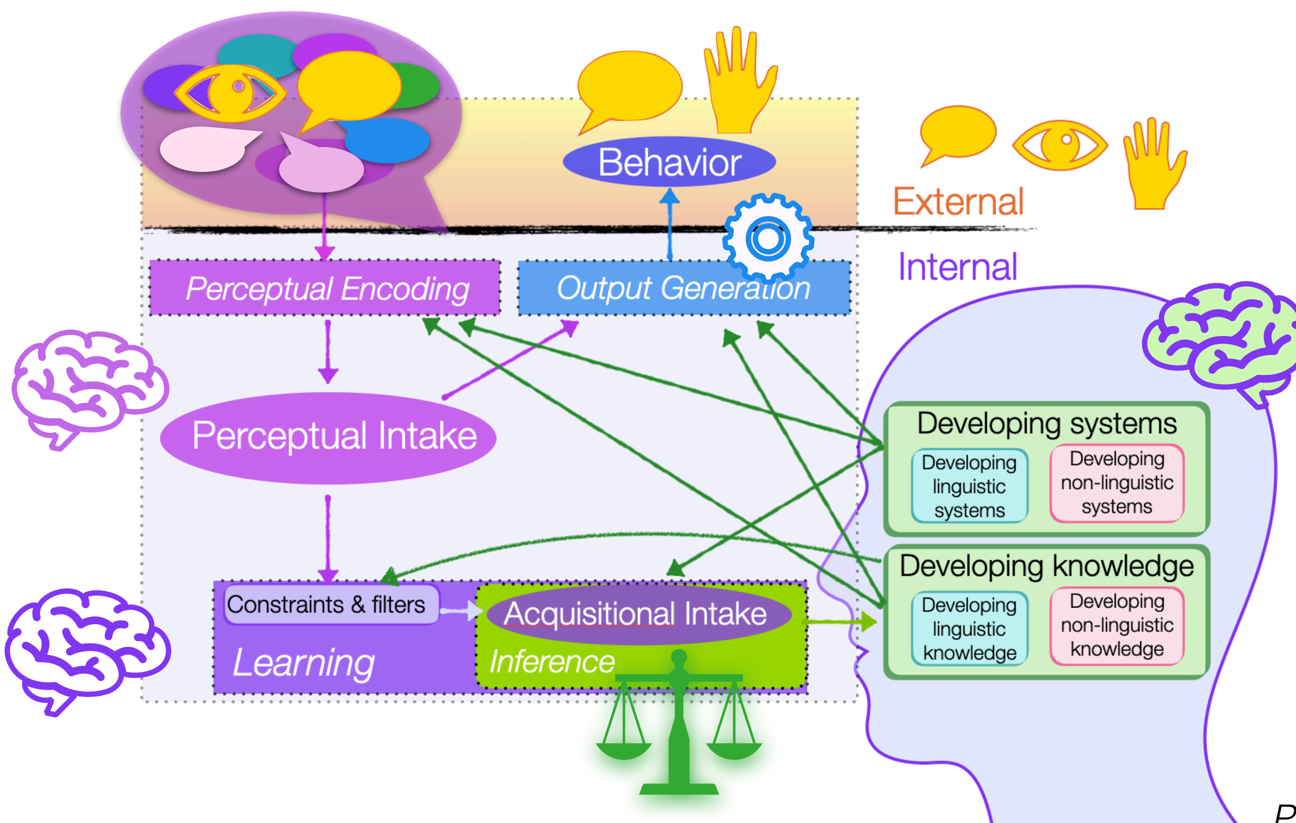


Pearl 2023 JoCL

Syntactic islands



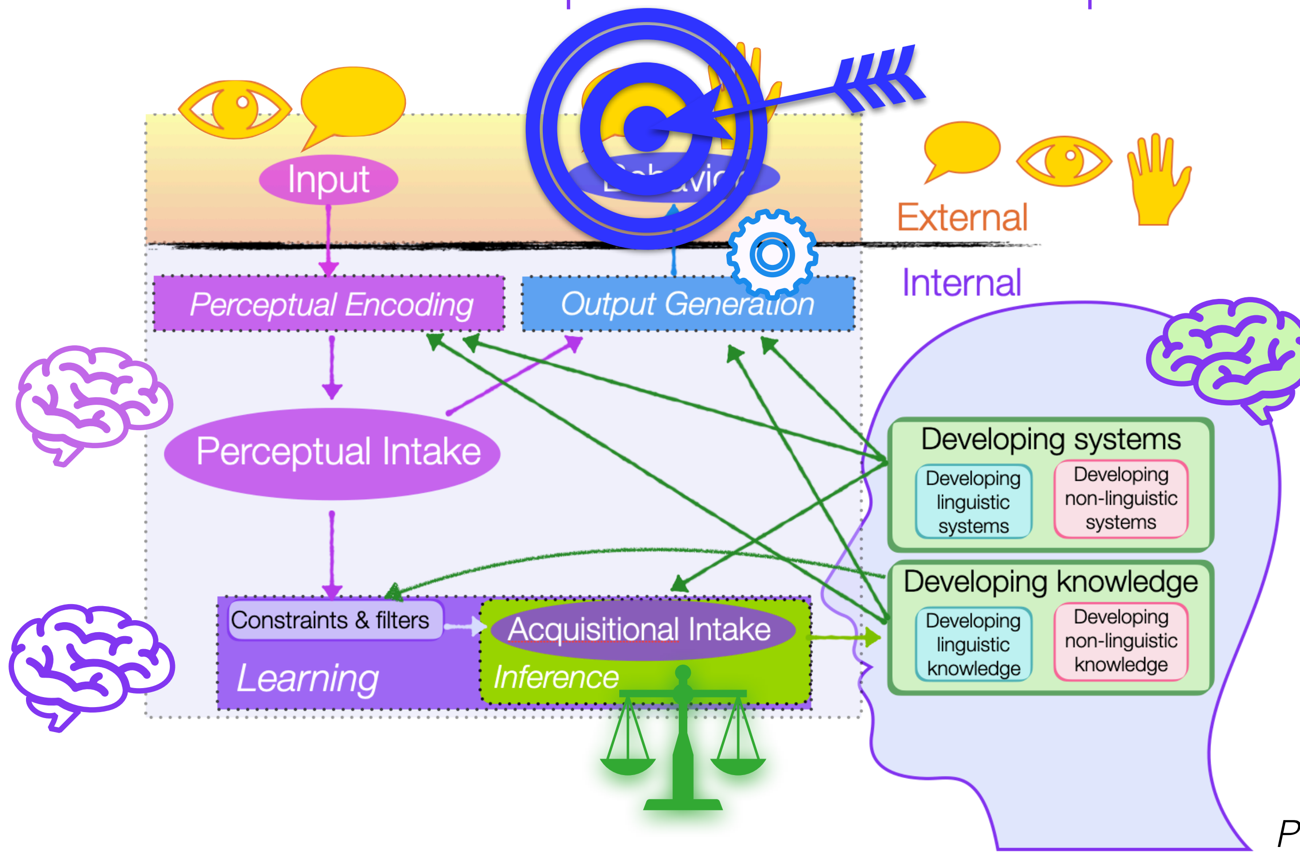
We can observe the input for *wh*-dependencies



Pearl 2023 JoCL

Syntactic islands

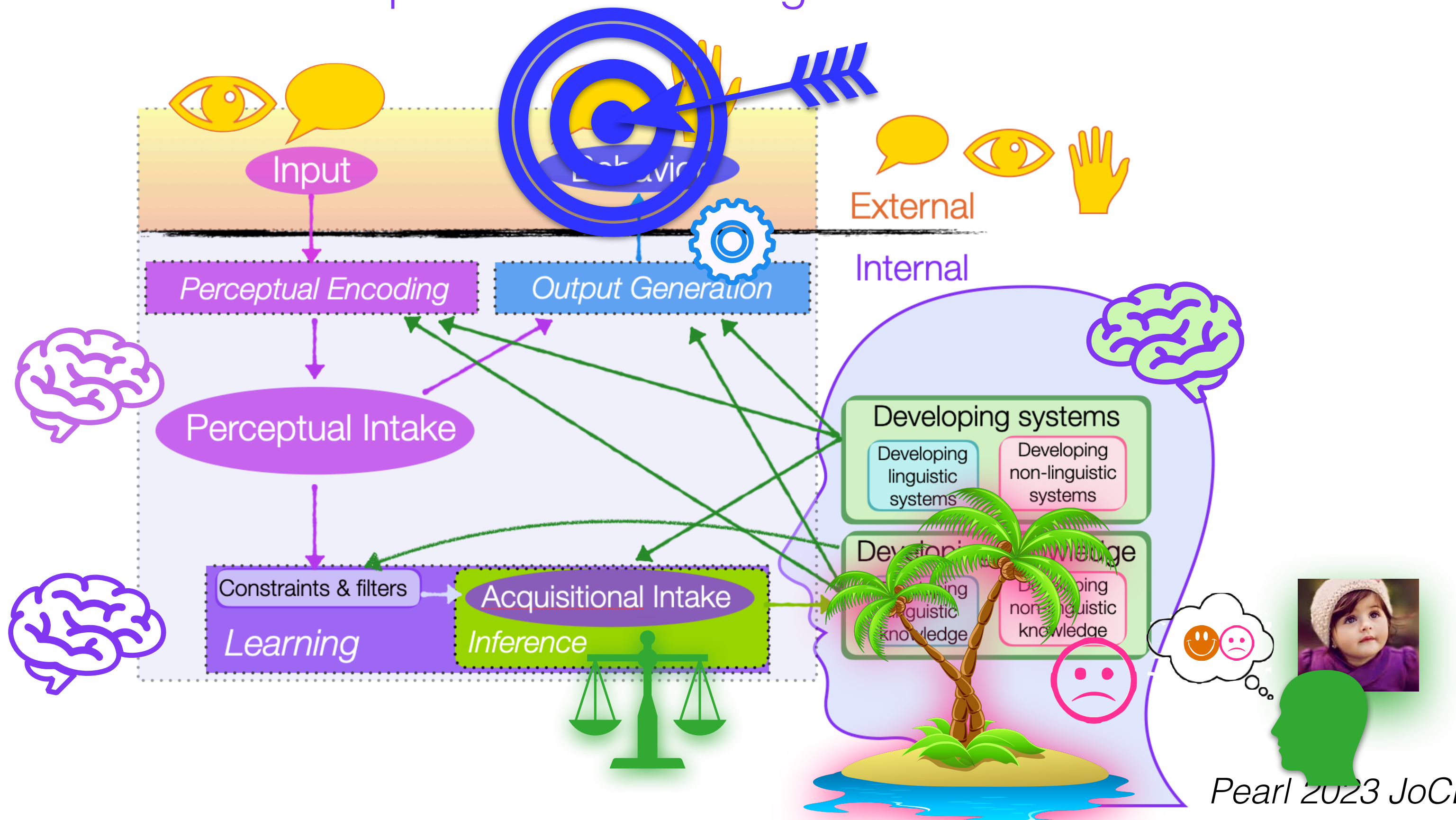
We can observe the output behavior for *wh*-dependencies



Pearl 2023 JoCL

Syntactic islands

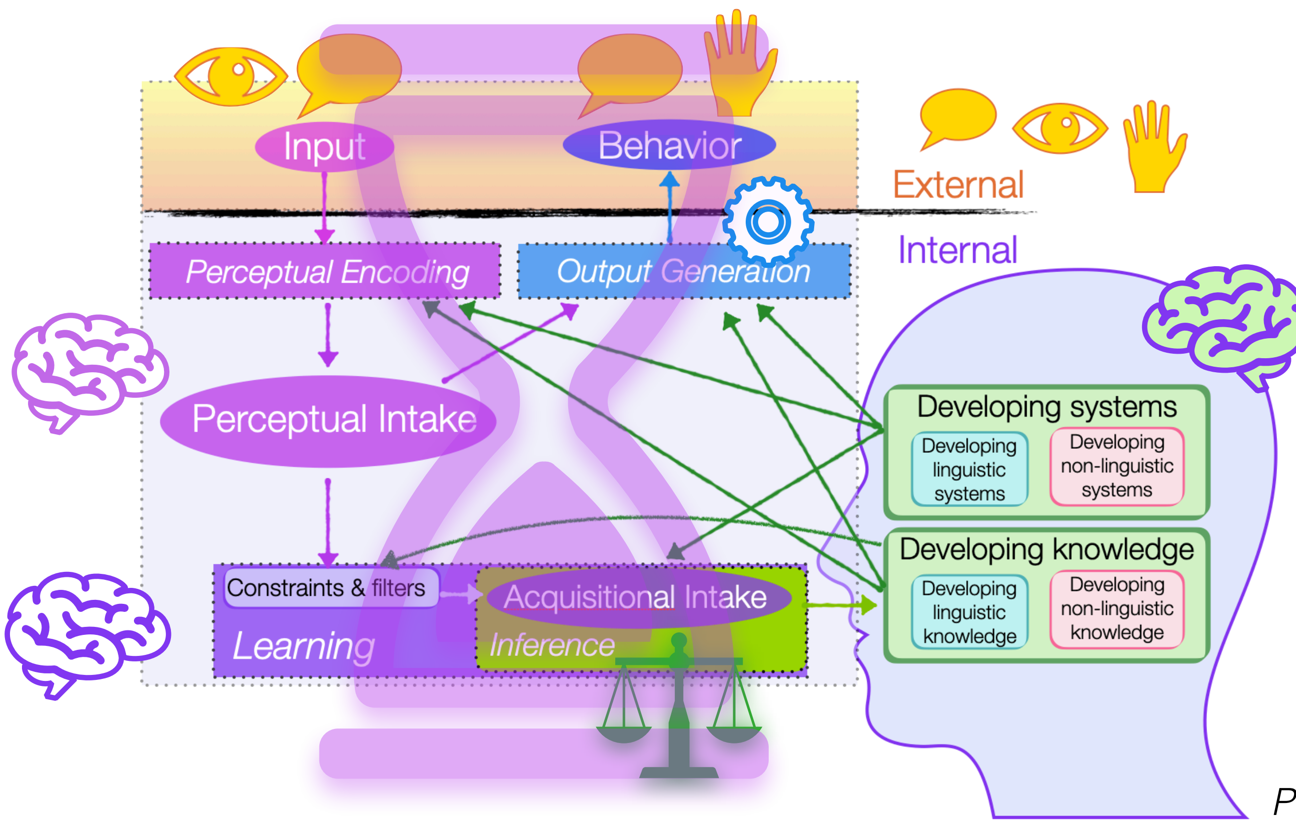
…which corresponds to knowledge we can't observe.



Pearl 2023 JoCL

Syntactic islands

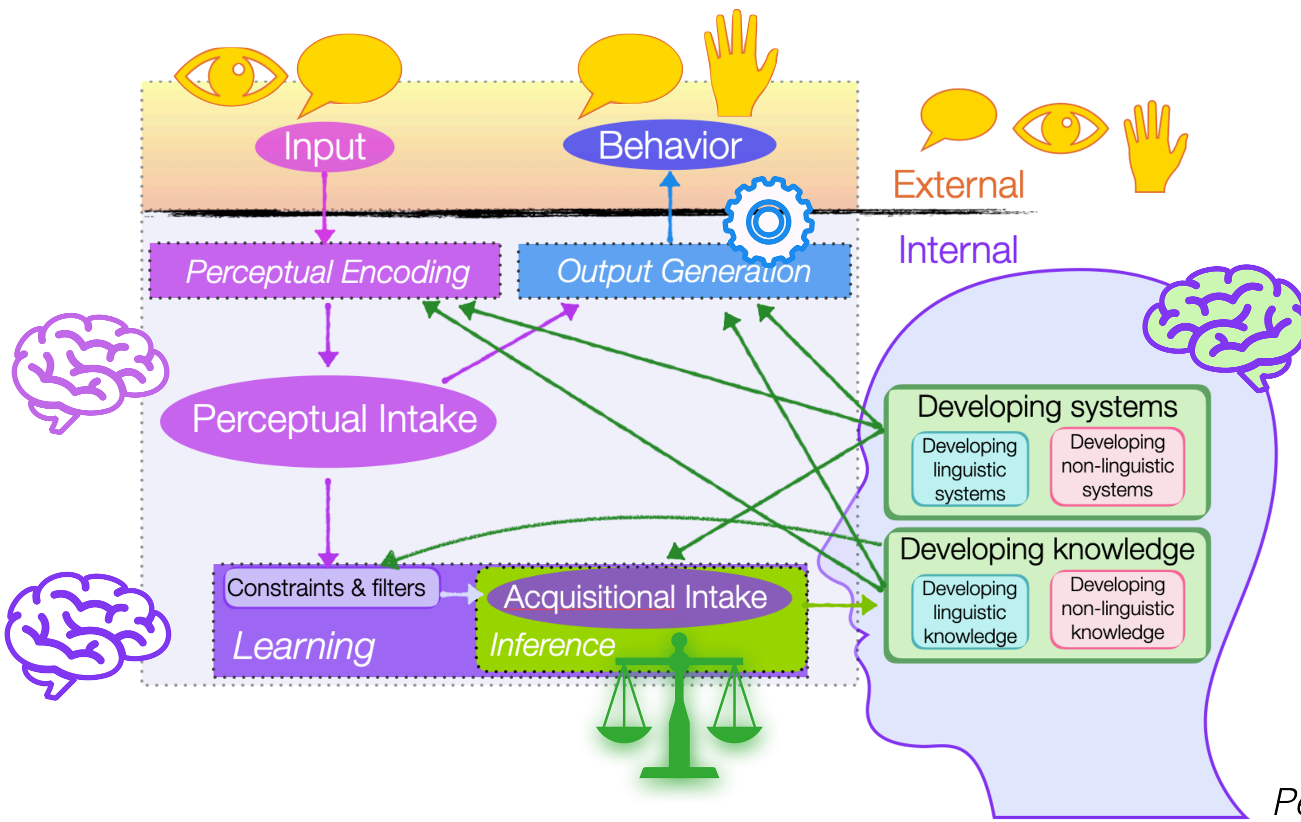
This whole process occurs over time.



Pearl 2023 JoCL

Syntactic islands

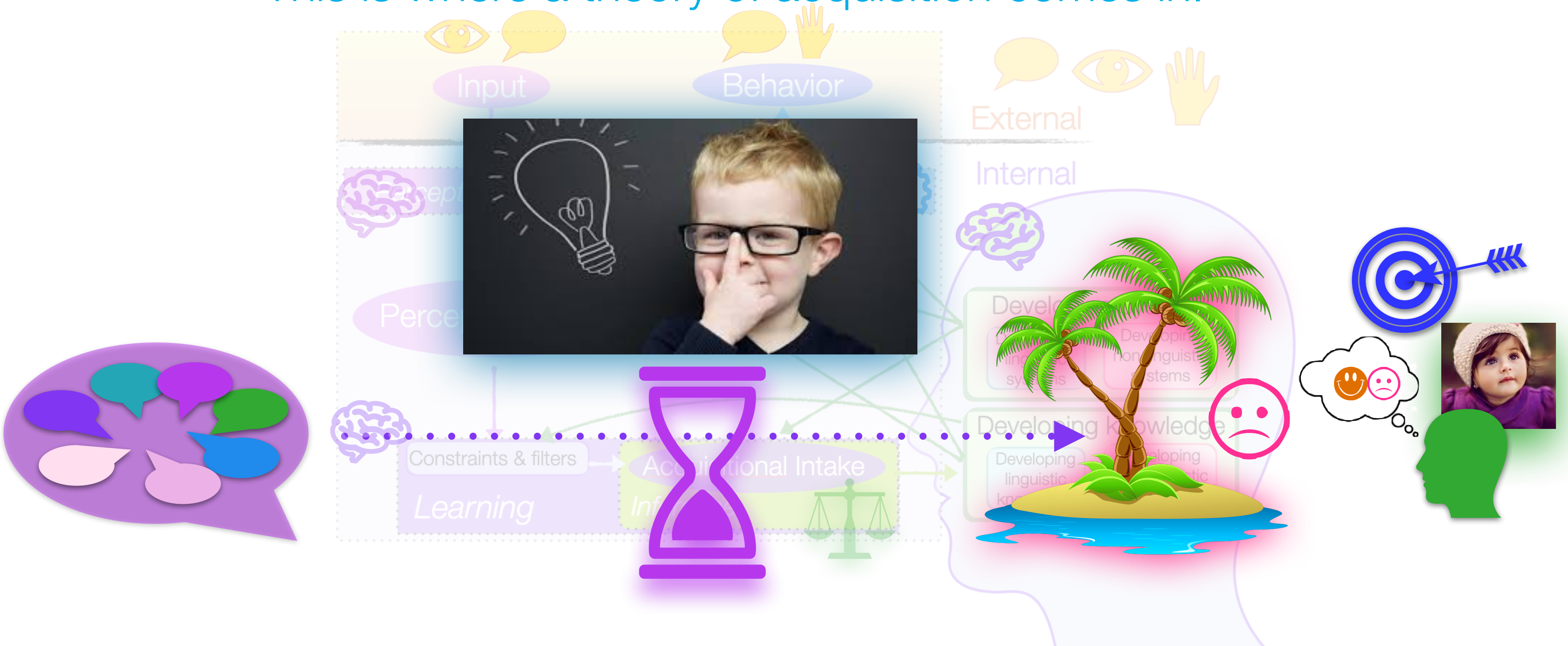
So how exactly does it work?



Pearl 2023 JoCL

Syntactic islands

This is where a theory of acquisition comes in.

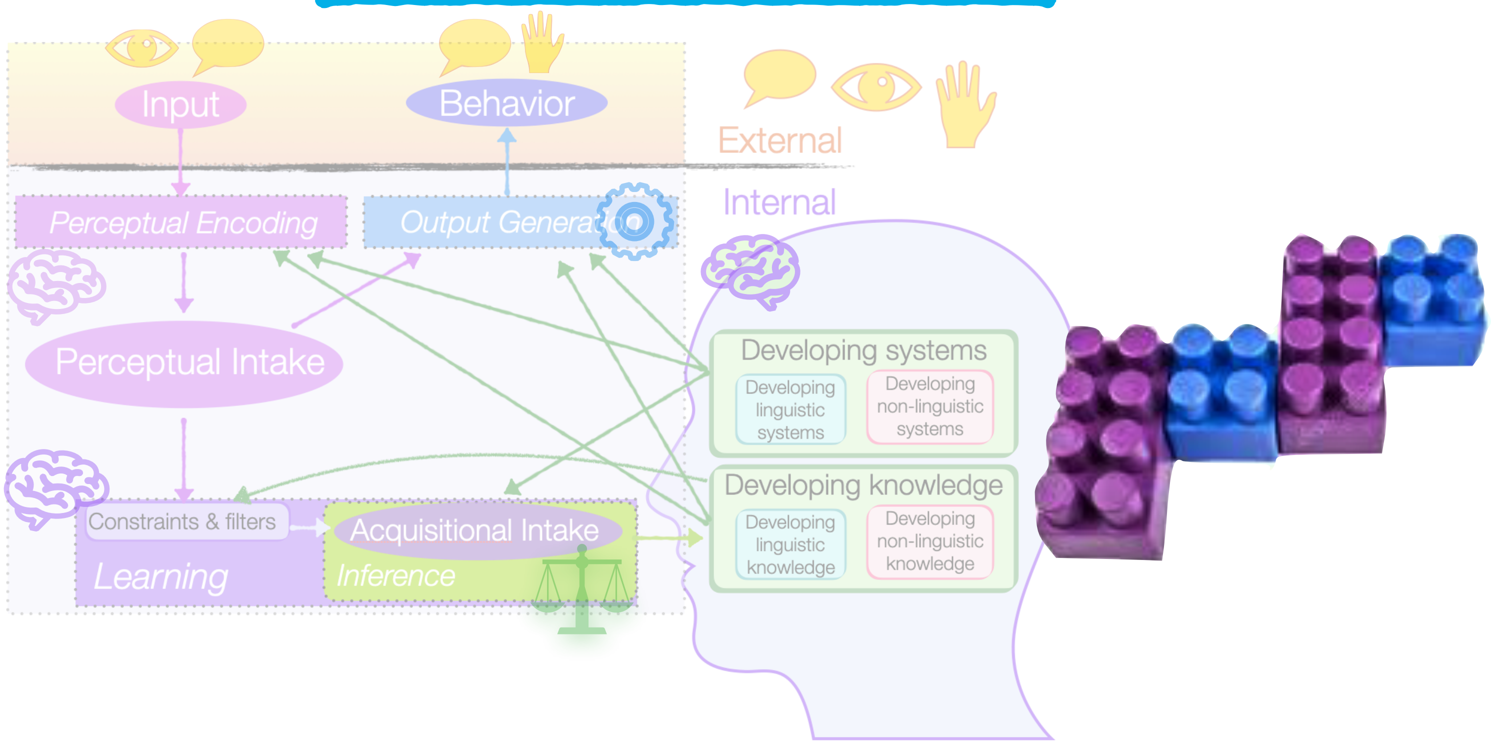


Syntactic islands

one learning theory for *wh*-dependencies

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

Learn the right building blocks



Learn the right building blocks

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

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

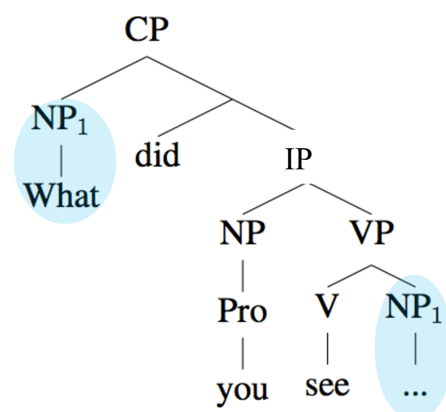
Learn the right building blocks

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

building blocks of *wh*-dependencies

Dependencies represented as a sequence of container nodes

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



Learn the right building blocks

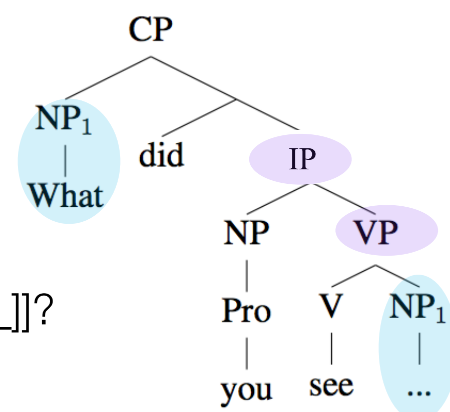
Pearl & Sprouse 2013
Dickson, Pearl, & Futrell 2022, 2024, in prep.

building blocks of *wh*-dependencies

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*



Learn the right building blocks

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

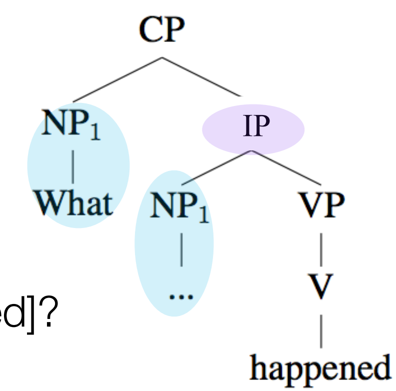
building blocks of *wh*-dependencies

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*



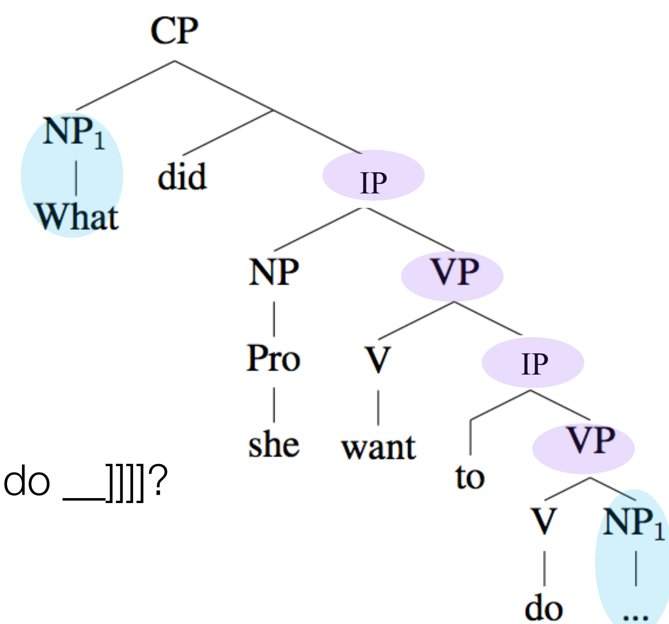
Learn the right building blocks

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

building blocks of *wh*-dependencies

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*

Learn the right building blocks

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

building blocks of *wh*-dependencies

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

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

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

(Much) less acceptable dependencies have low probability segments

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

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

Learn the right building blocks

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

building blocks of *wh*-dependencies

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

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

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

(Much) less acceptable dependencies have low probability segments

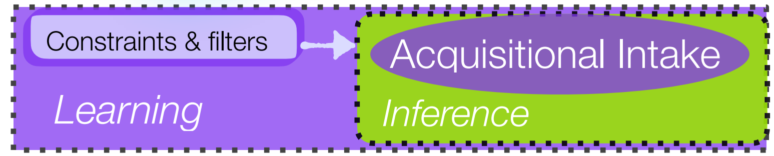
[$_{CP}$ *Who* *did* [$_{IP}$ *Lily* [$_{VP}$ *think* [$_{CP\text{-}that}$ [$_{IP}$ [$_{NP}$ *the kitty* [$_{PP}$ *for* __]] *was pretty* ?]]]]

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

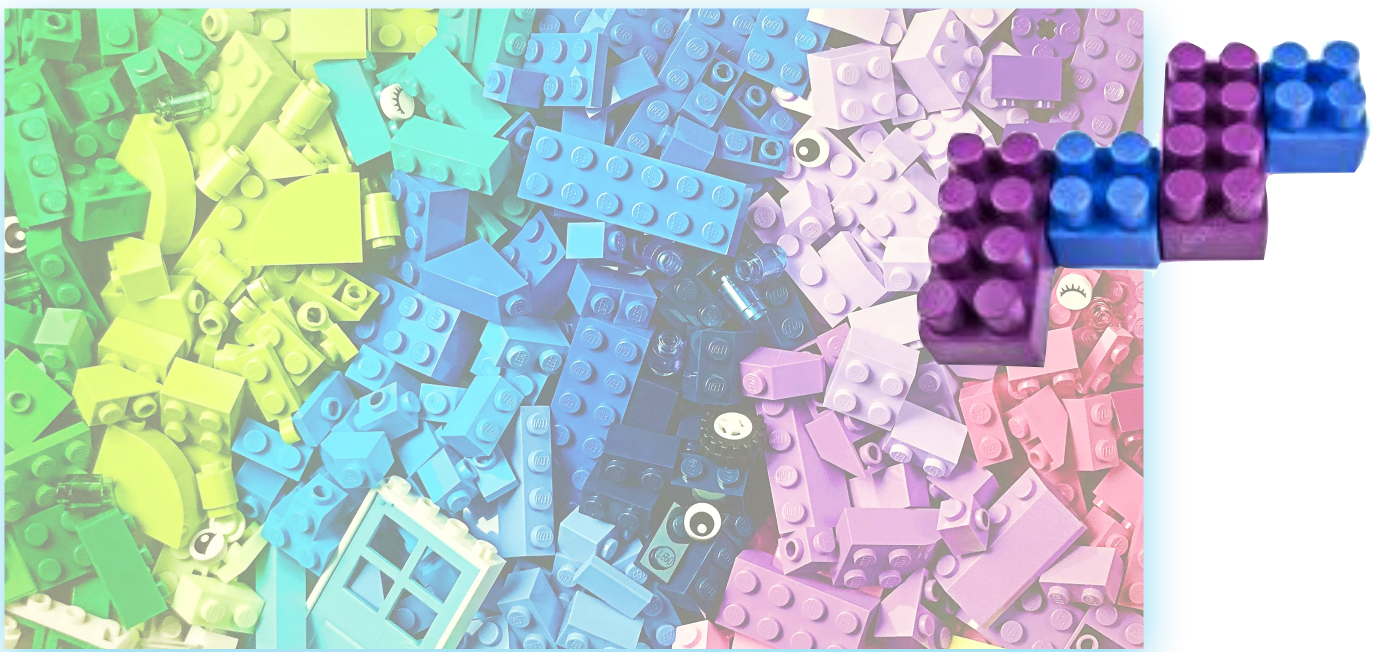
So if children break these dependencies into smaller building blocks, they can identify if a dependency has bad segments (made up of one or more low probability building blocks).

Learn the right building blocks

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



Theory: The child tries to learn
what the "best" building blocks are

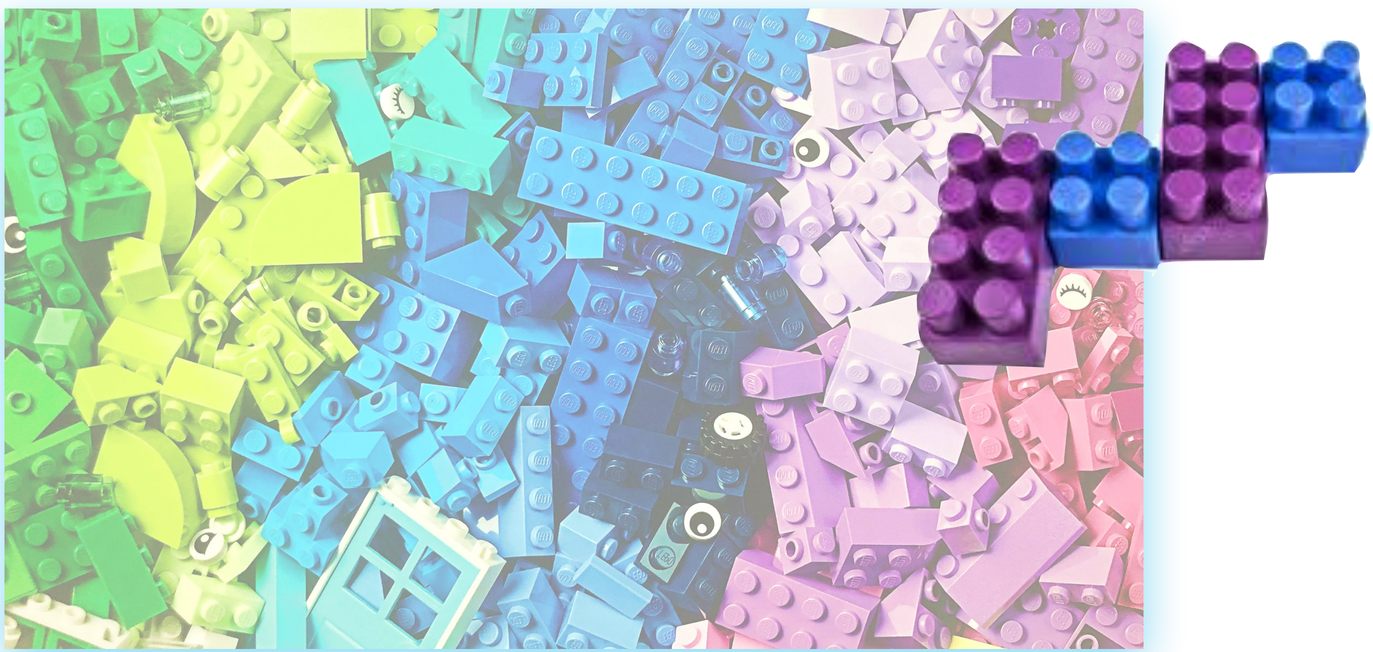


Learn the right building blocks

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

Guiding intuition:

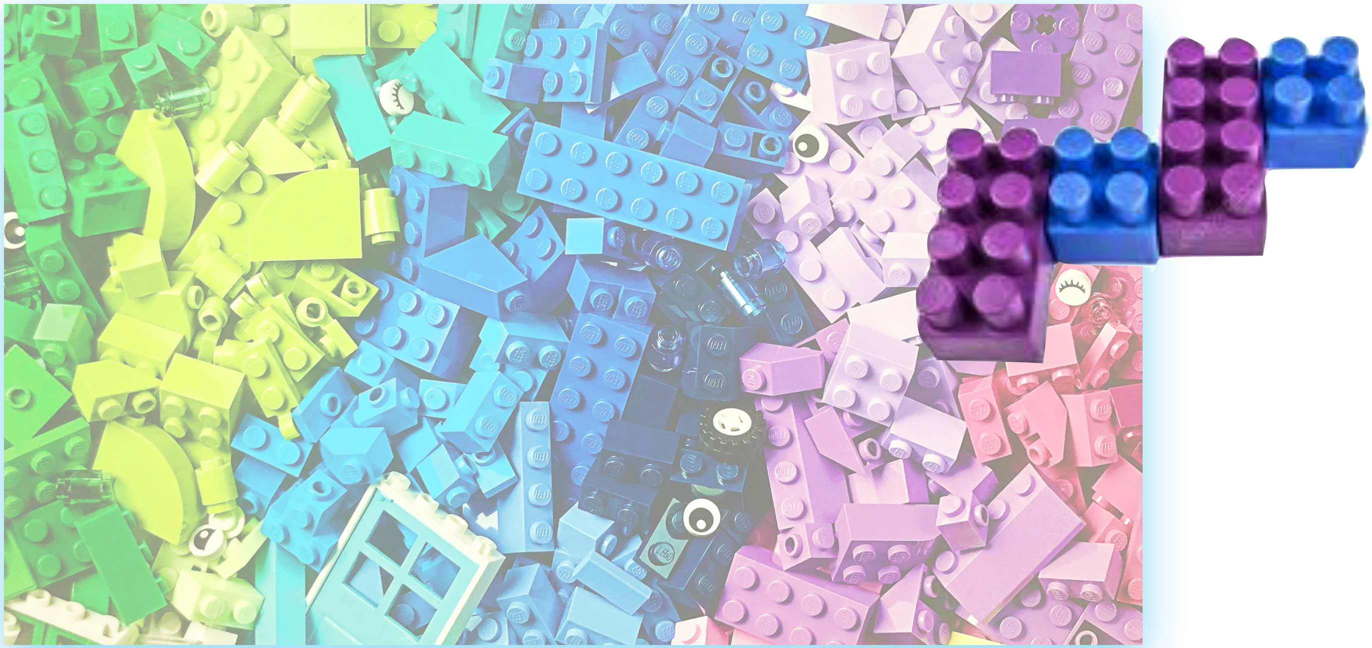
the “best” building blocks are the most “efficient” ones.



Learn the right building blocks

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

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

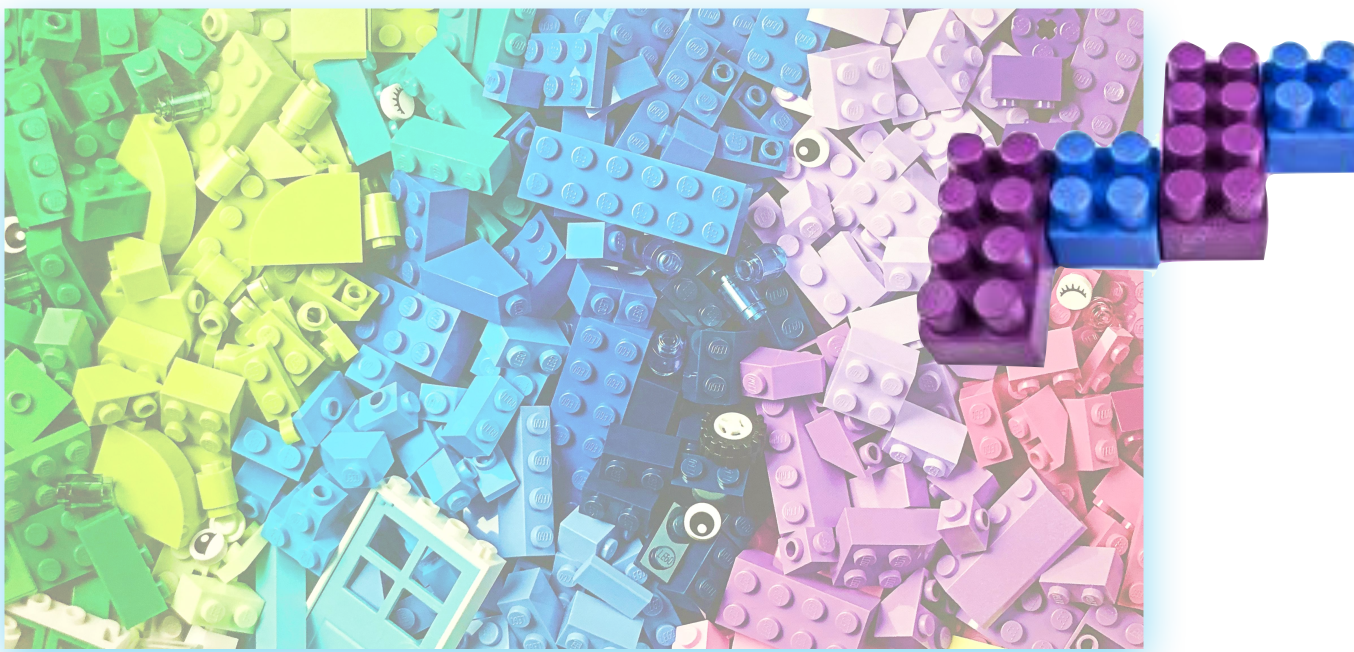


Learn the right building blocks

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

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).



Learn the right building blocks

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

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

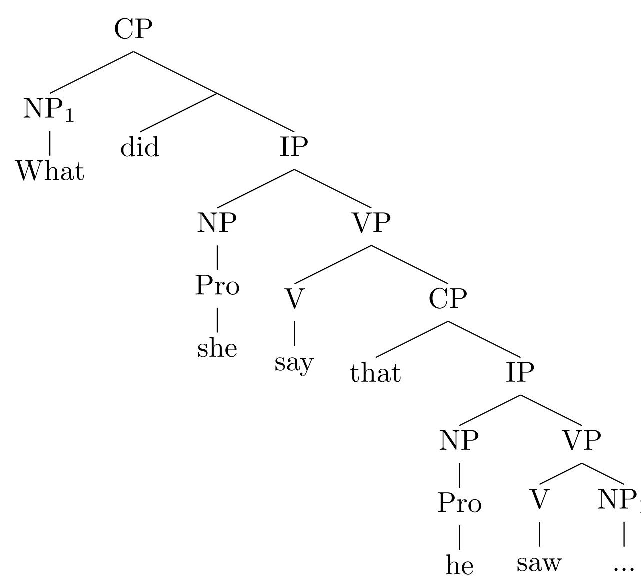
Learn the right building blocks

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

learning efficient building blocks

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

What did she say that he saw __ ?



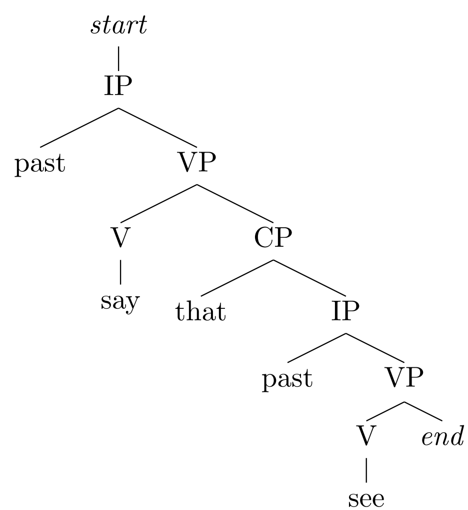
Learn the right building blocks

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

learning efficient building blocks

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

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



Learn the right building blocks

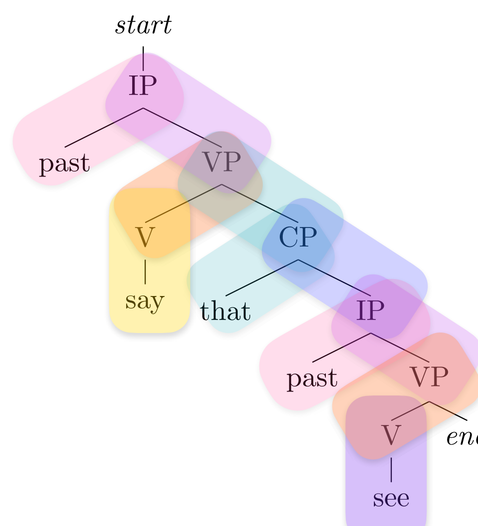
Dickson, Pearl, & Futrell 2022, 2024, in prep.

learning efficient building blocks

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

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

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



Learn the right building blocks

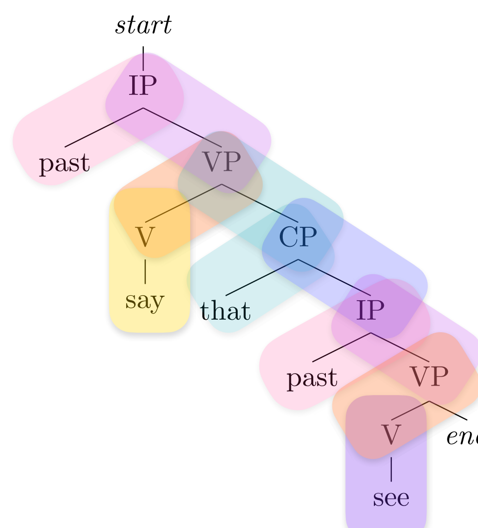
Dickson, Pearl, & Futrell 2022, 2024, in prep.

learning efficient building blocks

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

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

It may be slower to put together many small blocks.



Learn the right building blocks

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

learning efficient building blocks

a balance between

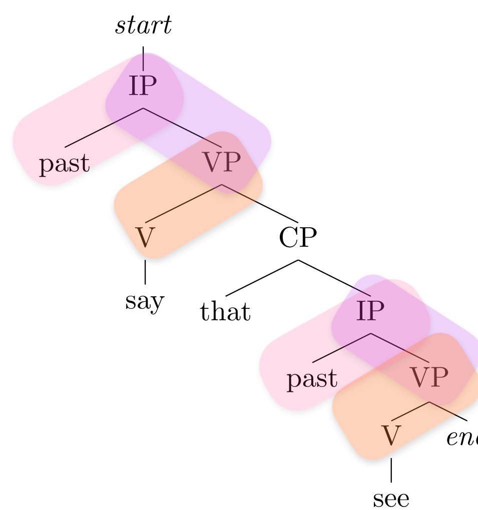
(1) how big they are

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

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

many smaller

slower because many



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

Learn the right building blocks

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

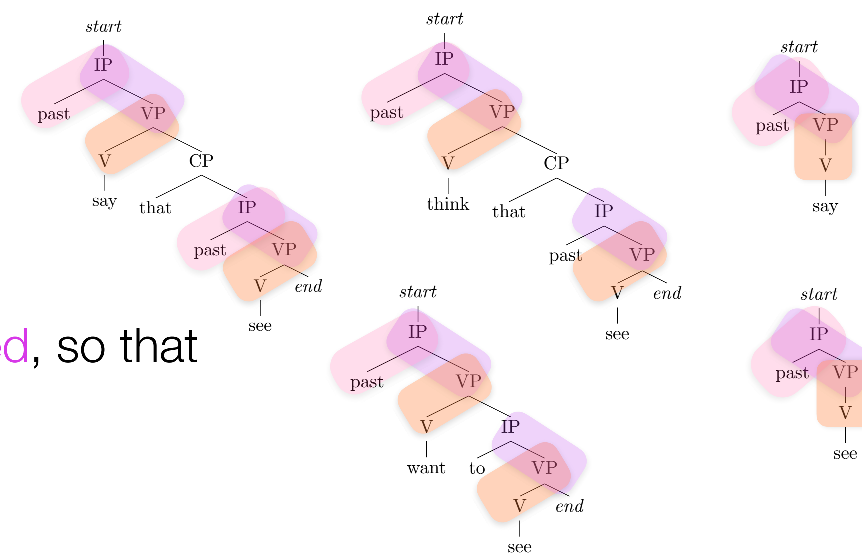
learning efficient building blocks

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

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

many smaller

slower because many



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

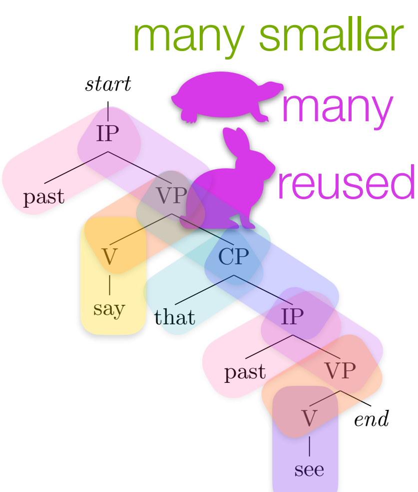
Learn the right building blocks

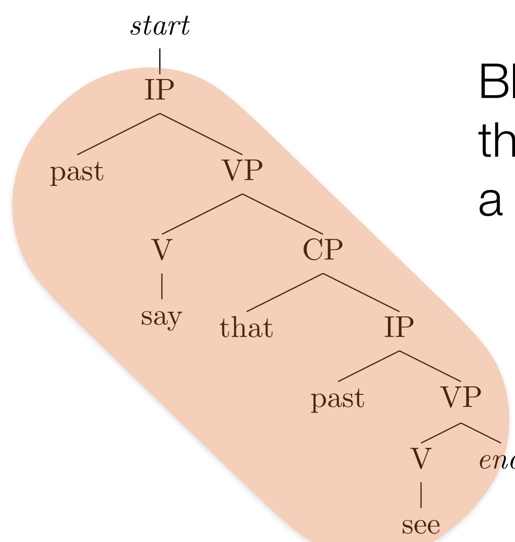
Dickson, Pearl, & Futrell 2022, 2024, in prep.

learning efficient building blocks

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

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





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

Learn the right building blocks

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

learning efficient building blocks

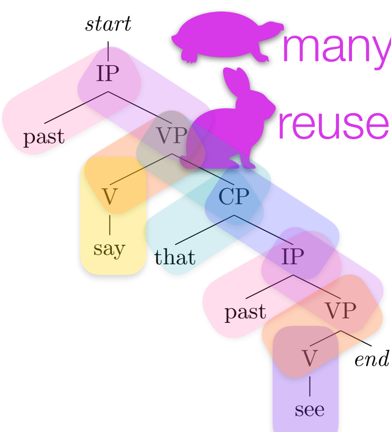
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

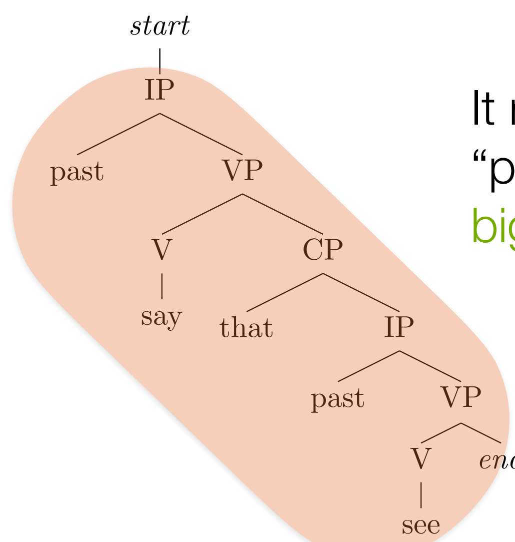
start-IP$_{\text{past}}$-VP$_{\text{say}}$-CP$_{\text{that}}$-IP$_{\text{past}}$-VP$_{\text{see}}$-*end*

many smaller

many

reused





It may be faster to "put together" one big block.

Learn the right building blocks

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

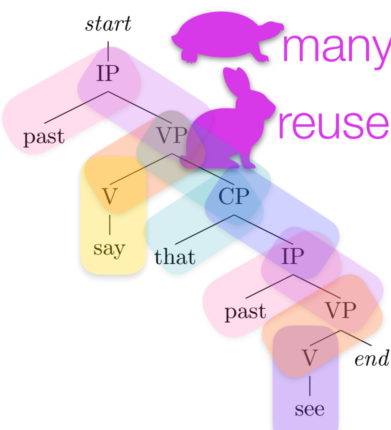
learning efficient building blocks

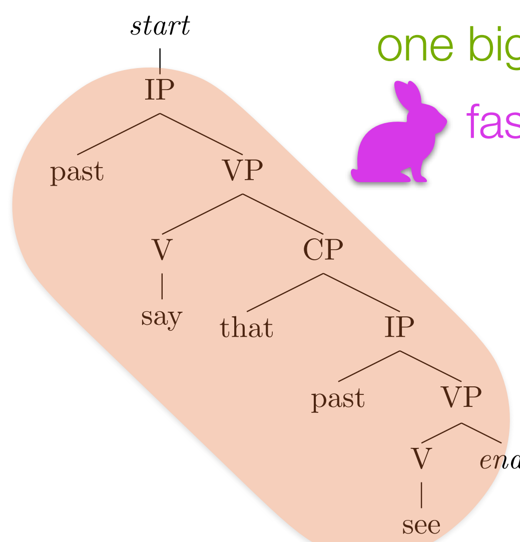
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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 to access if the block is used rarely.

Learn the right building blocks

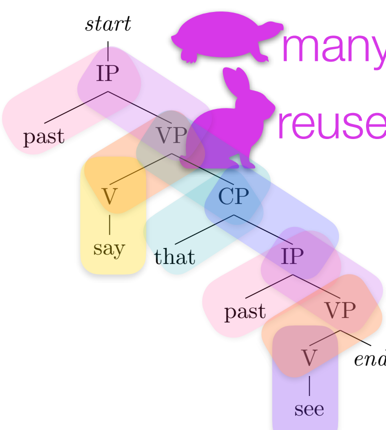
Dickson, Pearl, & Futrell 2022, 2024, in prep.

learning efficient building blocks

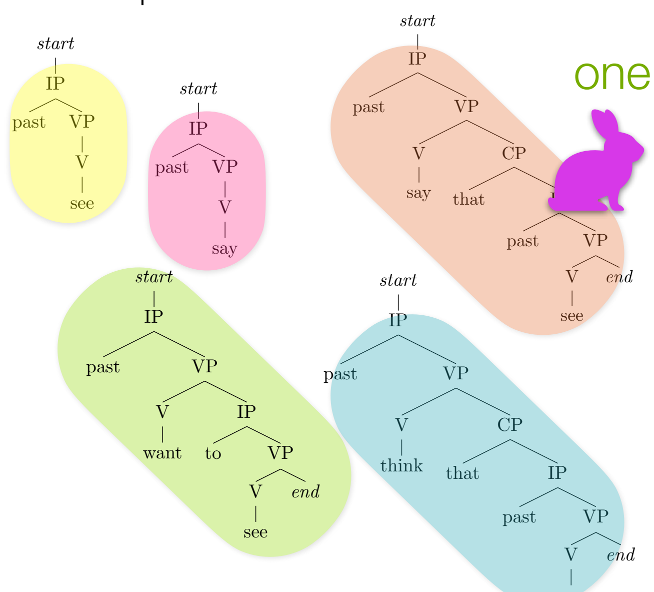
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

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 to access if the block is used rarely.

Learn the right building blocks

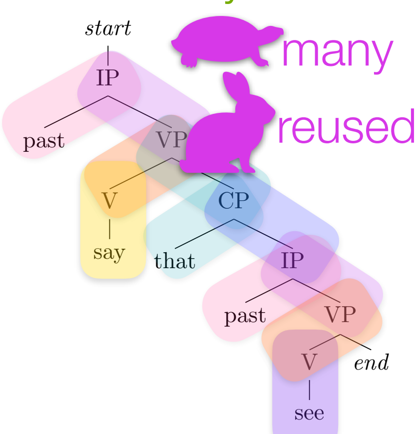
Dickson, Pearl, & Futrell 2022, 2024, in prep.

learning efficient building blocks

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

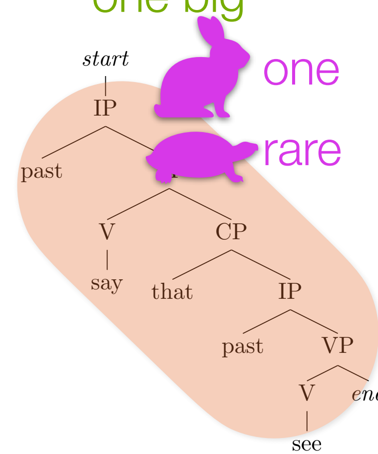
$\textit{start}\text{-IP}_{\text{past}}\text{-VP}_{\text{say}}\text{-CP}_{\text{that}}\text{-IP}_{\text{past}}\text{-VP}_{\text{see}}\text{-}\textit{end}$

many smaller



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

one big



Learn the right building blocks

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

learning efficient building blocks

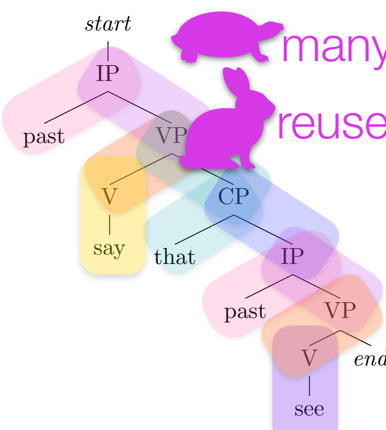
a balance between
(1) how big they are
(2) how fast they are to put together to make a *wh*-dependency

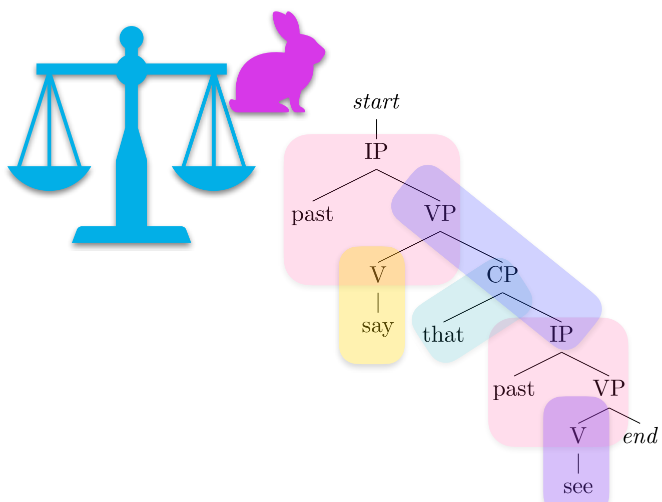
start-IP$_{\text{past}}$-VP$_{\text{say}}$-CP$_{\text{that}}$-IP$_{\text{past}}$-VP$_{\text{see}}$-*end*

many smaller

many

reused

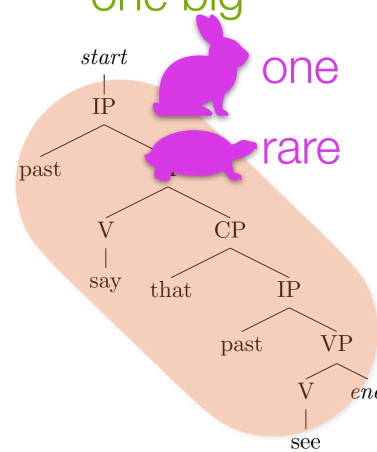




one big

one

rare



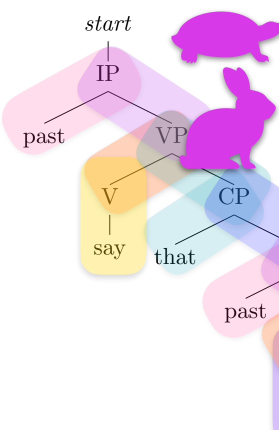
Learn the right building blocks

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

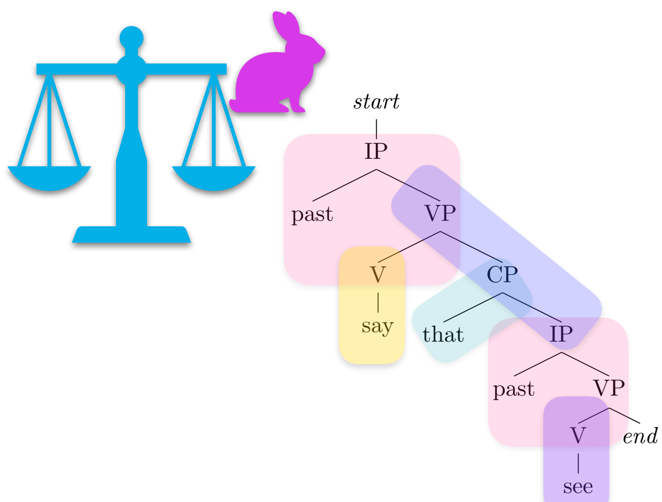
learning efficient building blocks

How can children find the best balance?

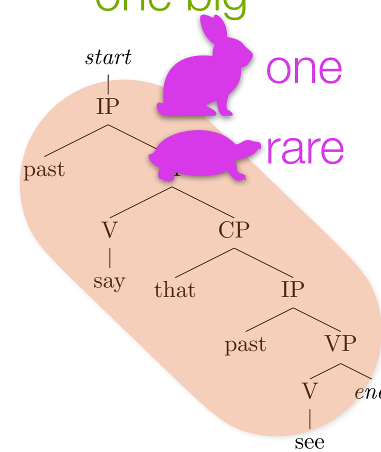
many smaller



many
reused



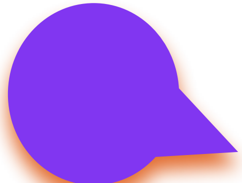
one big



one
rare

Learn the right building blocks

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



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

Learn the right building blocks

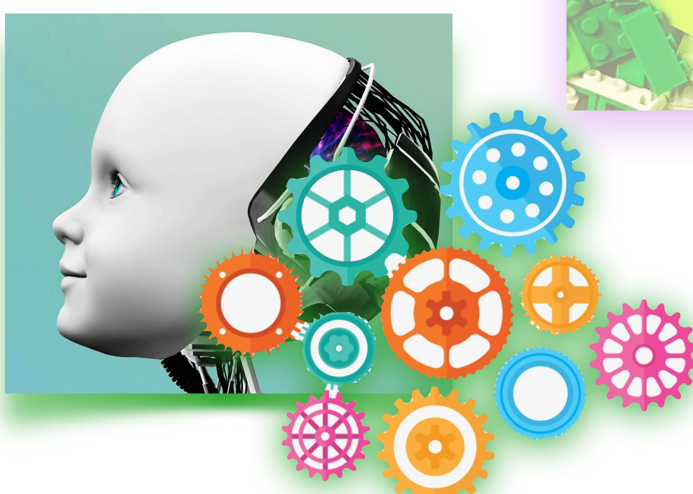
Dickson, Pearl, & Futrell 2022, 2024, in prep.

Note: There's a considerable body of evidence suggesting that young children are capable of Bayesian inference (3 years: Xu & Tenenbaum, 2007; 9 months: Gerken, 2006; Dewar & Xu, 2010; Gerken, 2010; 6 months: Denison, Reed, & Xu, 2011, among many others) — though they're likely approximating this mental computation as best they can.

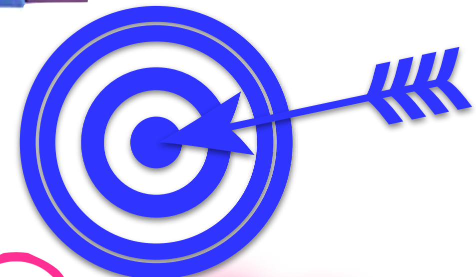
Learn the right building blocks

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

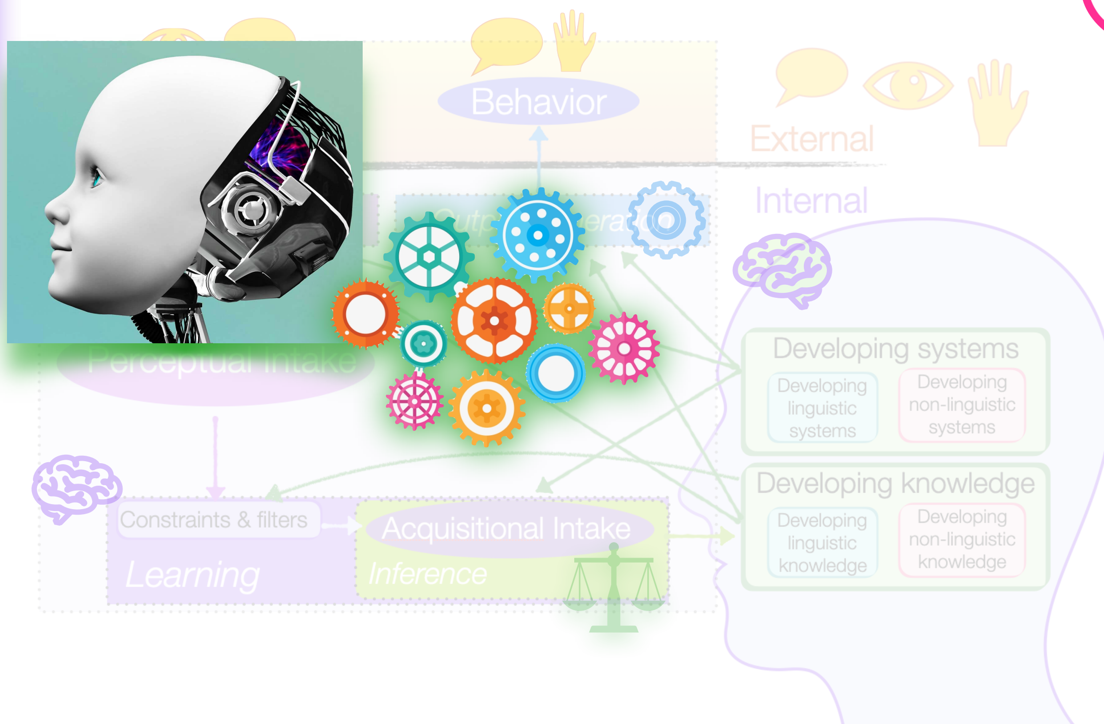
So that's what the modeled child will do

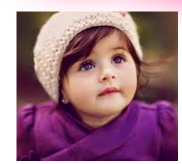


Dickson, Pearl, & Futrell 2022, 2024, in prep.



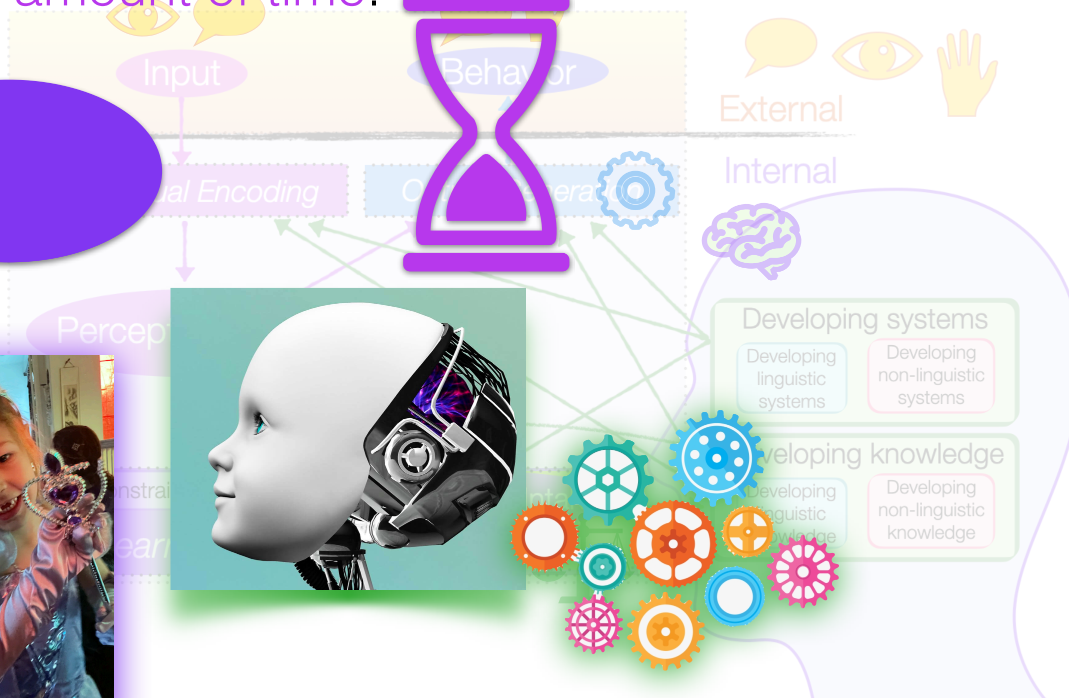
We'll see if this modeled child can learn the target knowledge that real children learn.

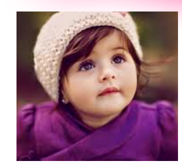




Dickson, Pearl, & Futrell 2022, 2024, in prep.

We'll have the modeled child learn from the same kind of input children encounter, for the same amount of time.





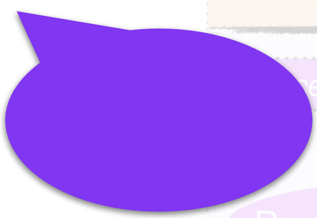
Dickson, Pearl, & Futrell 2022, 2024, in prep.

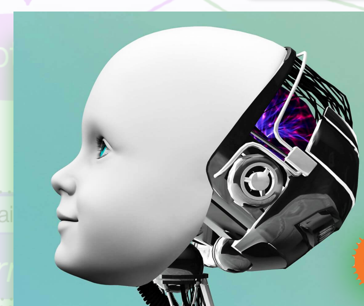
We'll have the modeled child learn from the same kind of input children encounter, for the same amount of time.

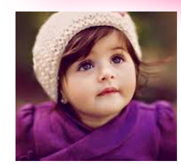
Estimated from the CHILDES Treebank (Pearl & Sprouse 2013)

≈10,442,258 min

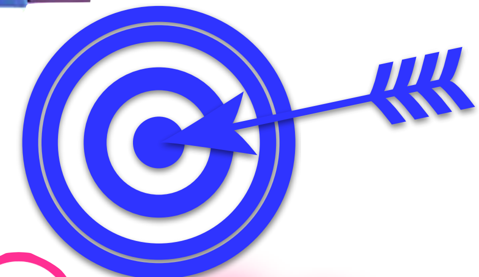
(derived from De Villiers et al. 2008, Perkins & Lidz 2021, & Davis et al. 2004)







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

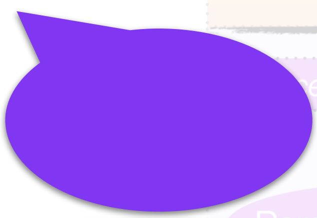


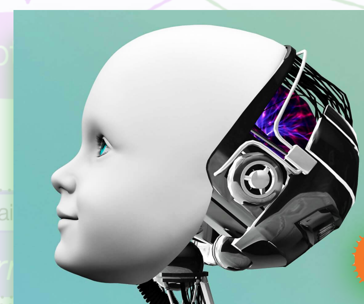
We'll have the modeled child learn from the same kind of input children encounter, for the same amount of time.

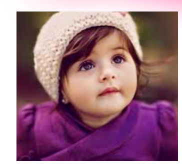
Estimated from the CHILDES Treebank (Pearl & Sprouse 2013)

≈2.15 million *wh*-dependencies

(derived from Hoff-Ginsberg 1998 and Rowe 2012)

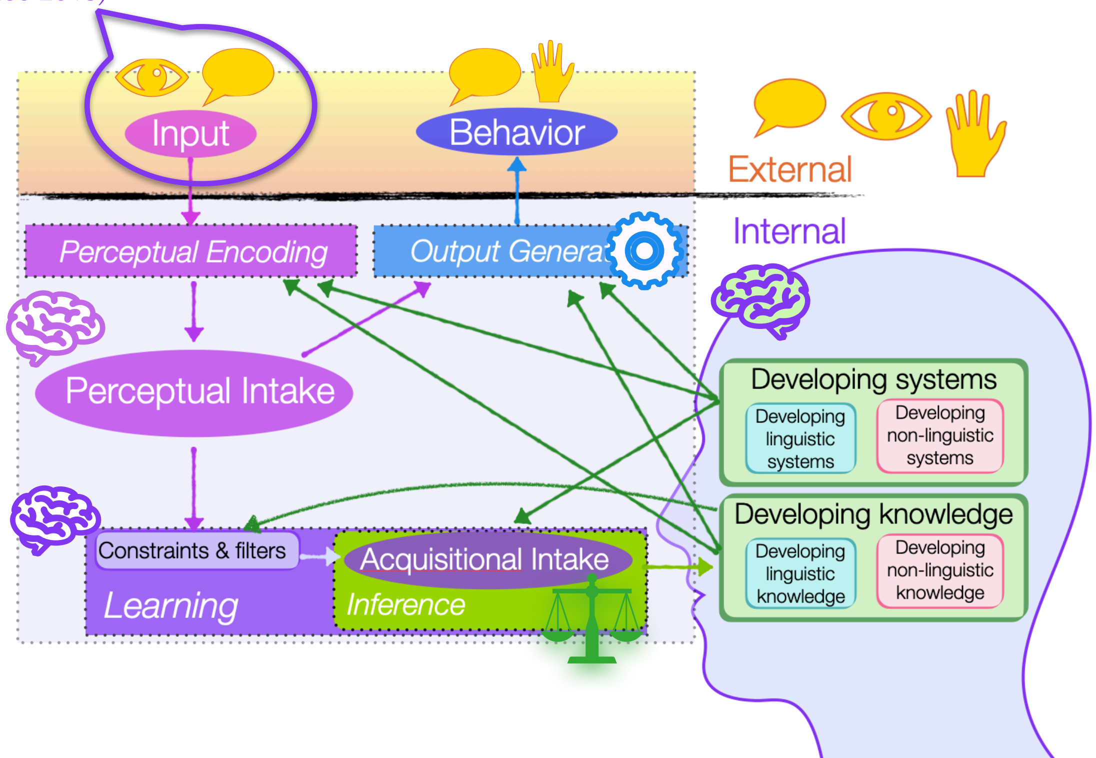






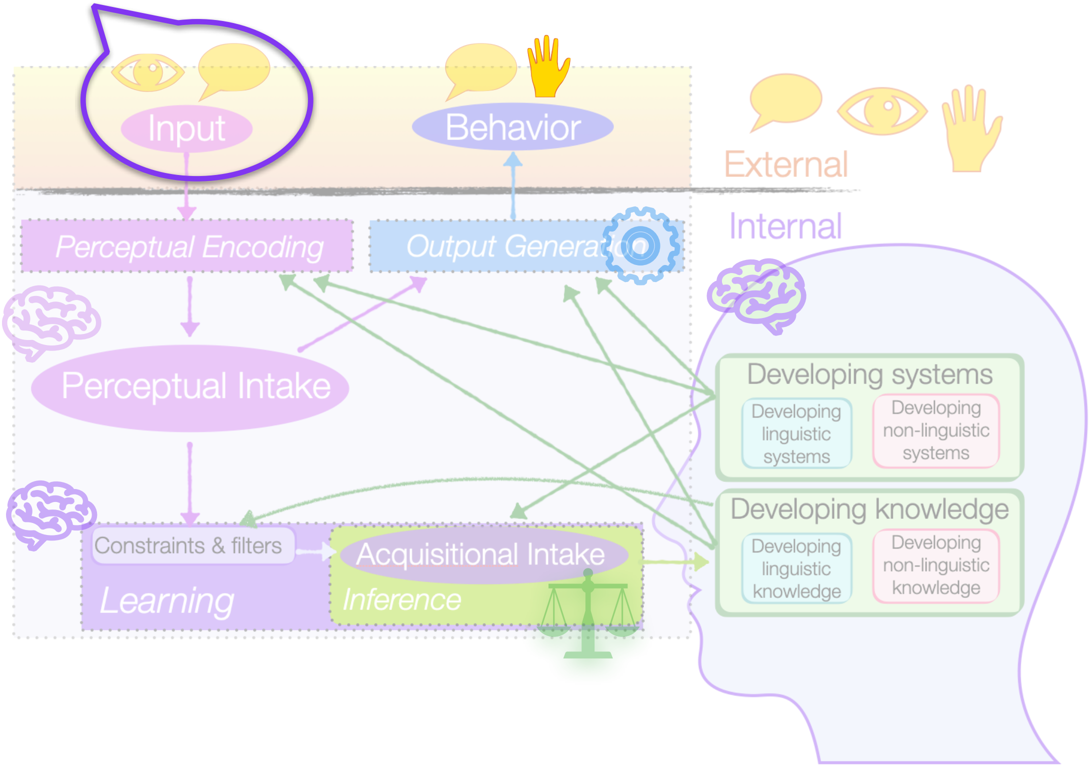
A more detailed look at the input

Estimated from the CHILDES Treebank (Pearl & Sprouse 2013)



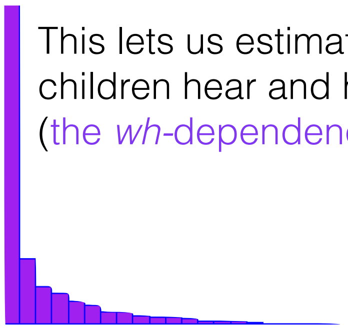
A more detailed look at the input

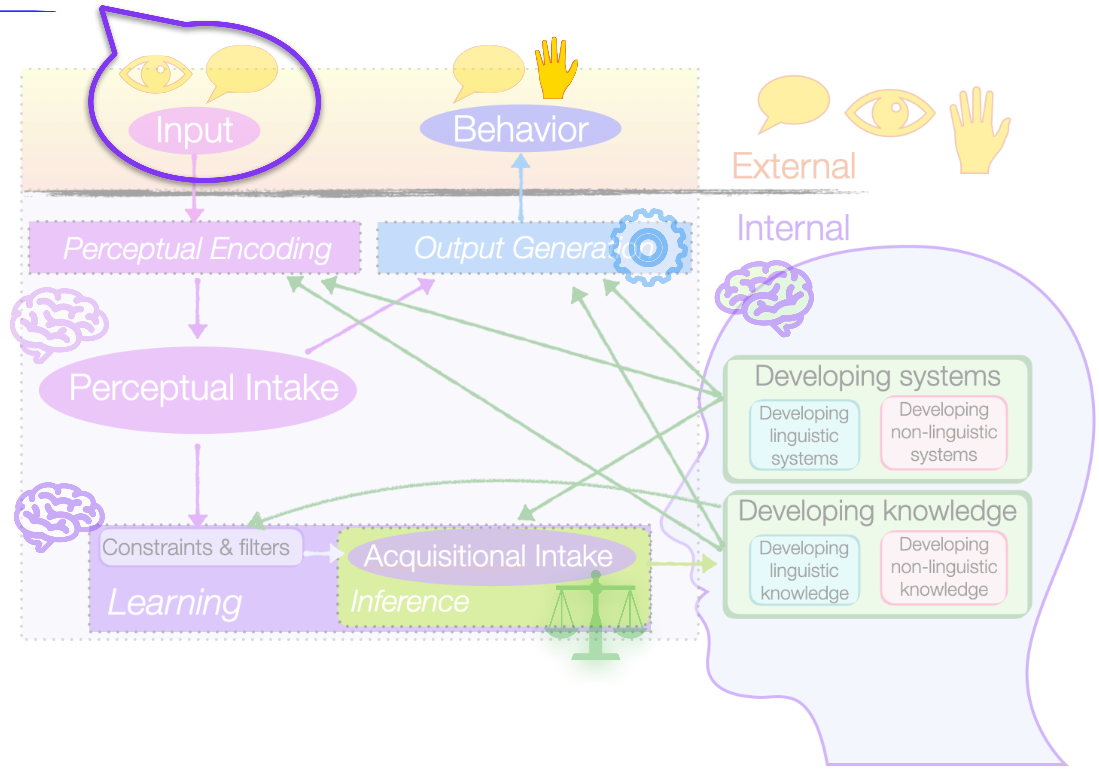
12.7K *wh*-dependencies from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 children between the ages of 1 and 5 years old.



A more detailed look at the input

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

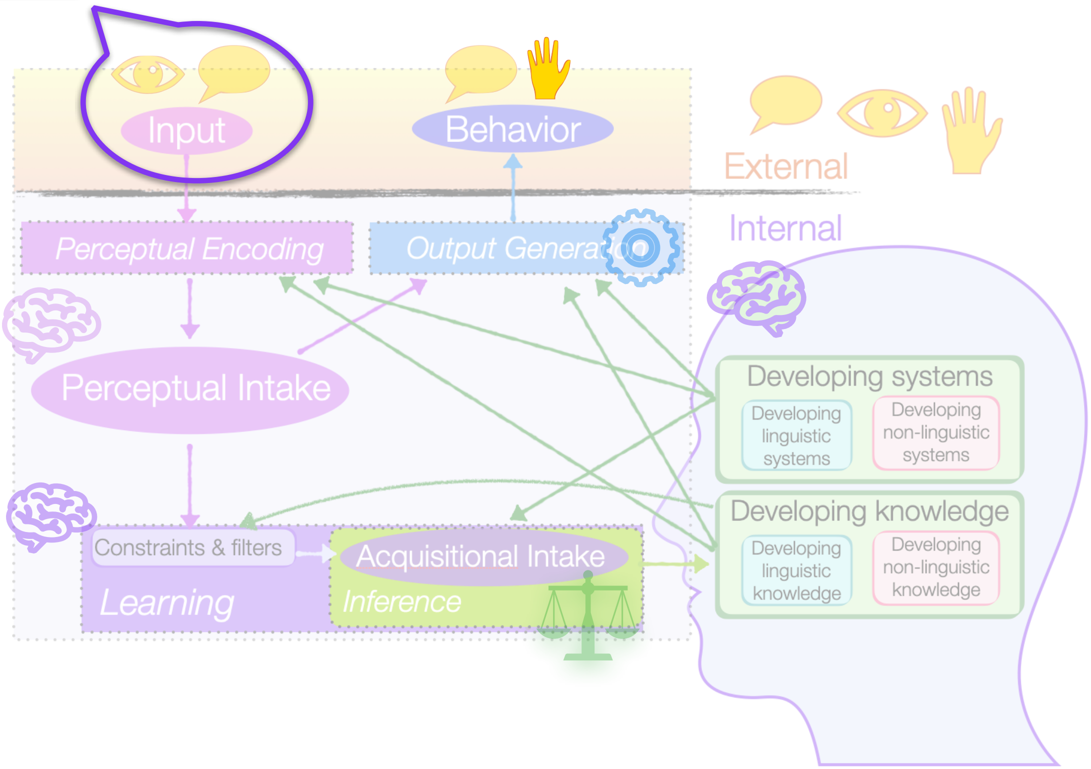




A more detailed look at the input

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

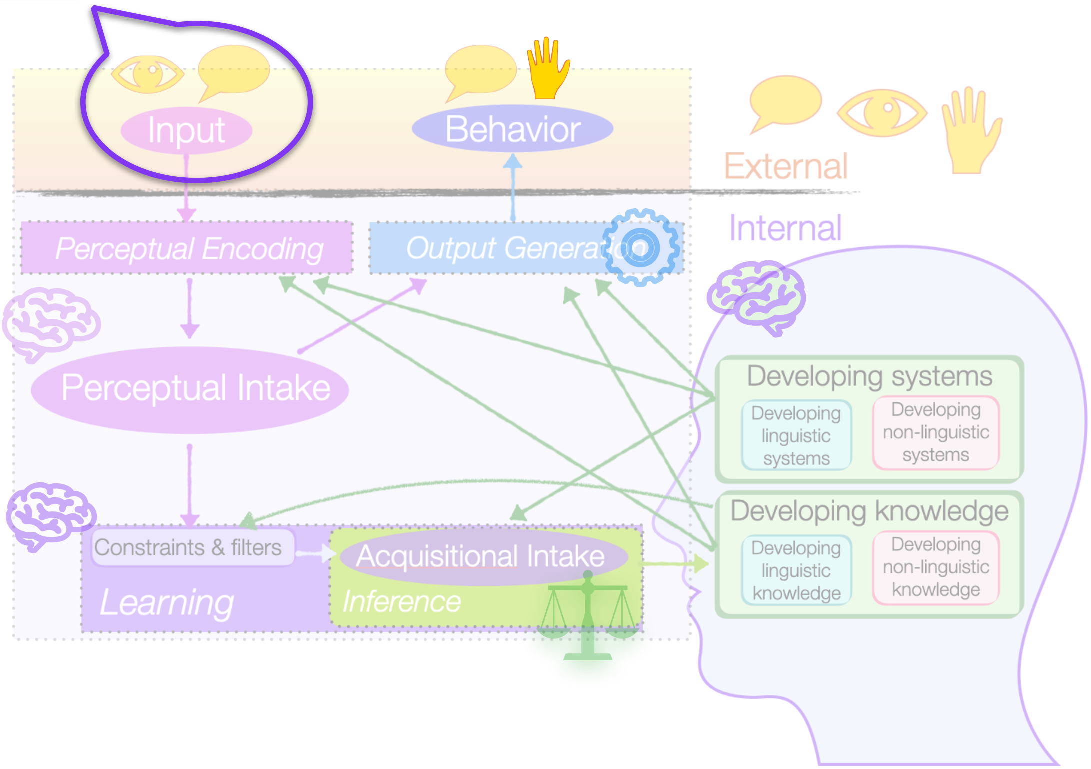
(<60 months)



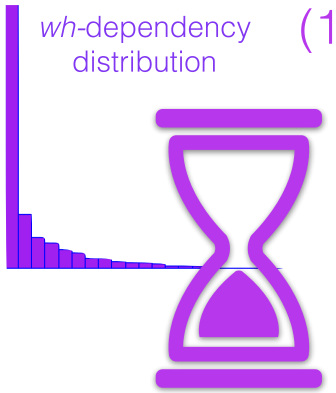
A more detailed look at the input

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.



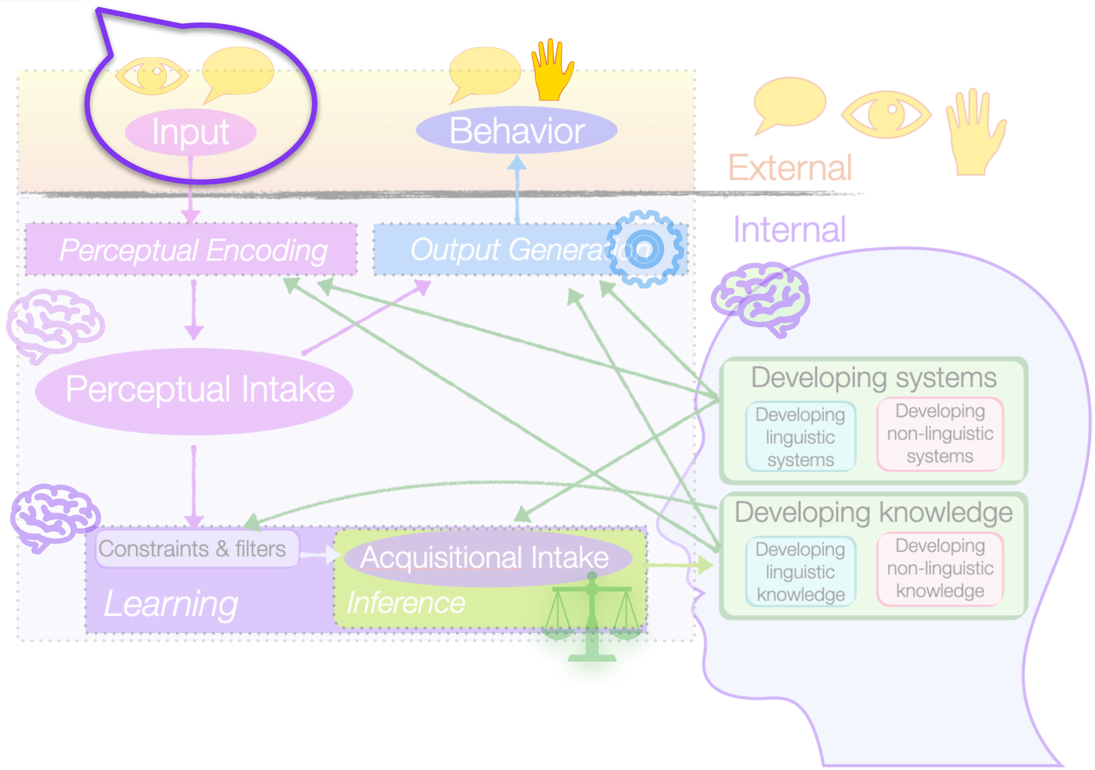
A more detailed look at the input



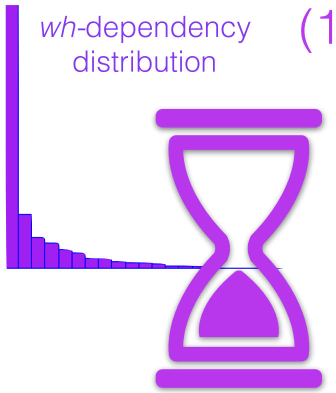
(18 months ≤ age < 60 months)

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.



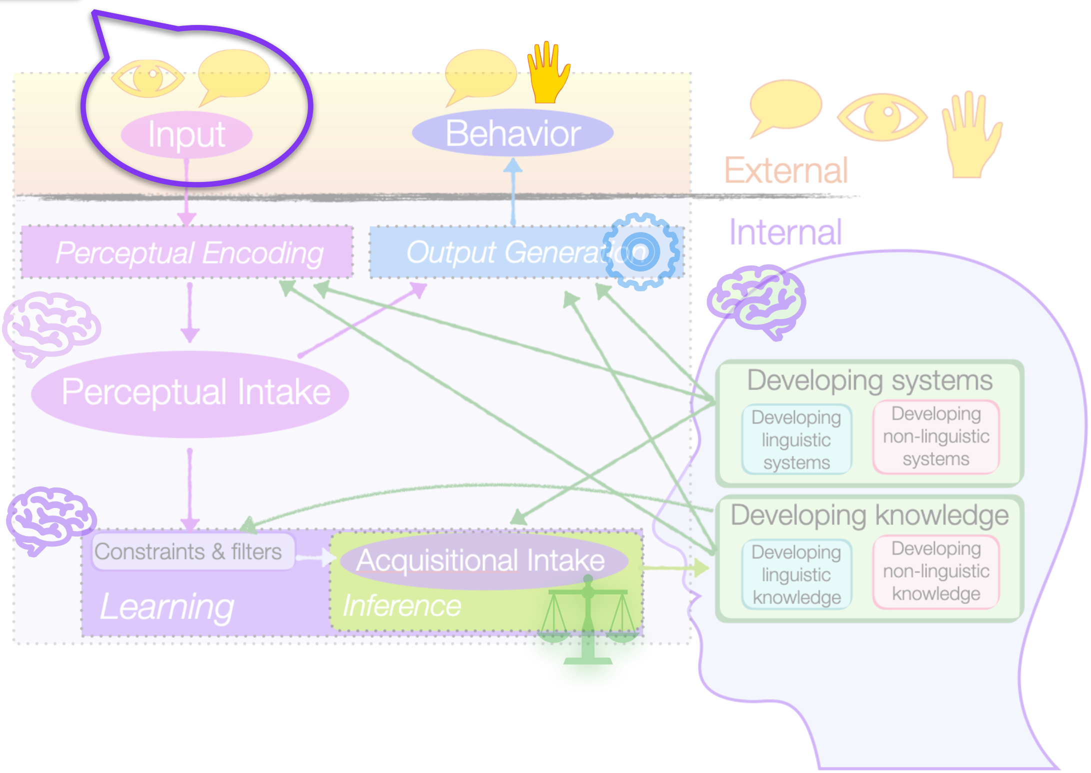
A more detailed look at the input



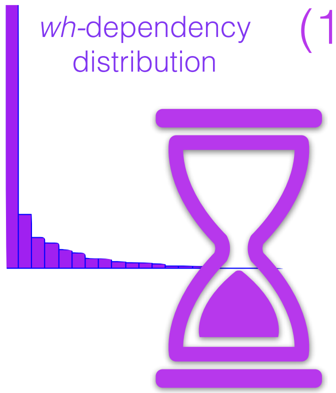
(18 months ≤ age < 60 months)

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.



A more detailed look at the input

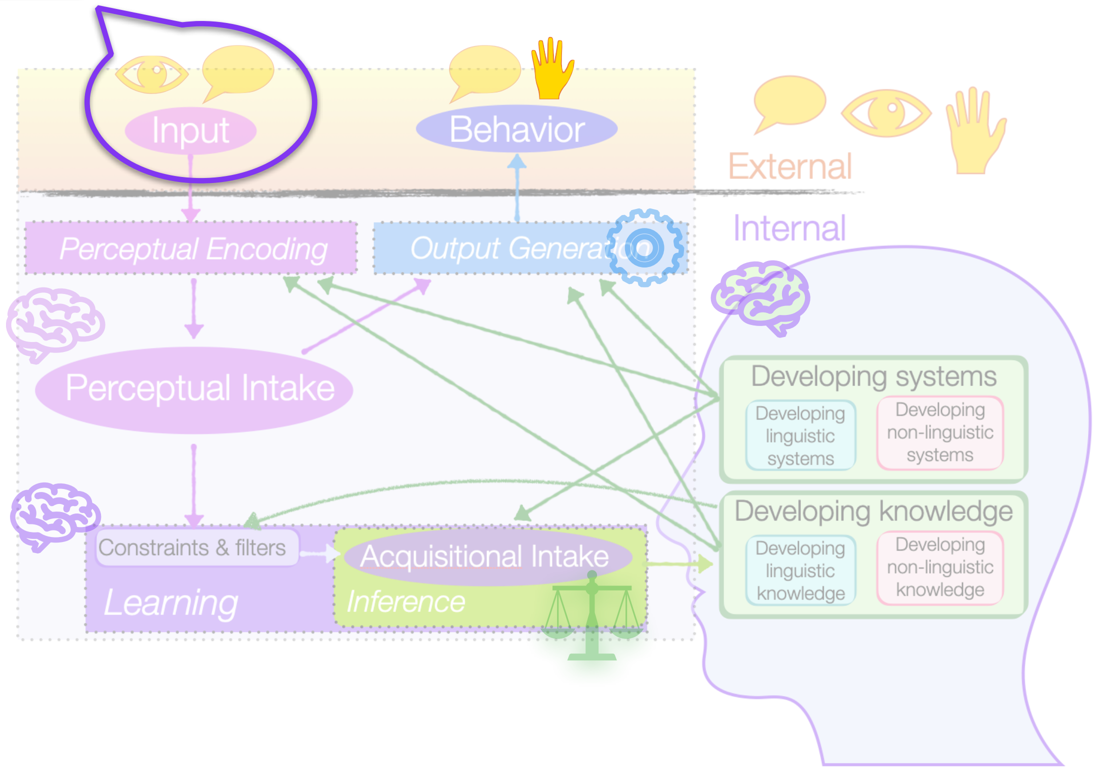


(18 months ≤ age < 60 months)

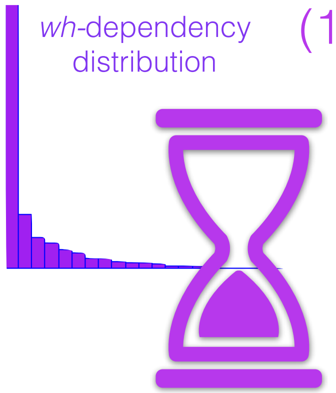
How many minutes:

≈10,442,258

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



A more detailed look at the input

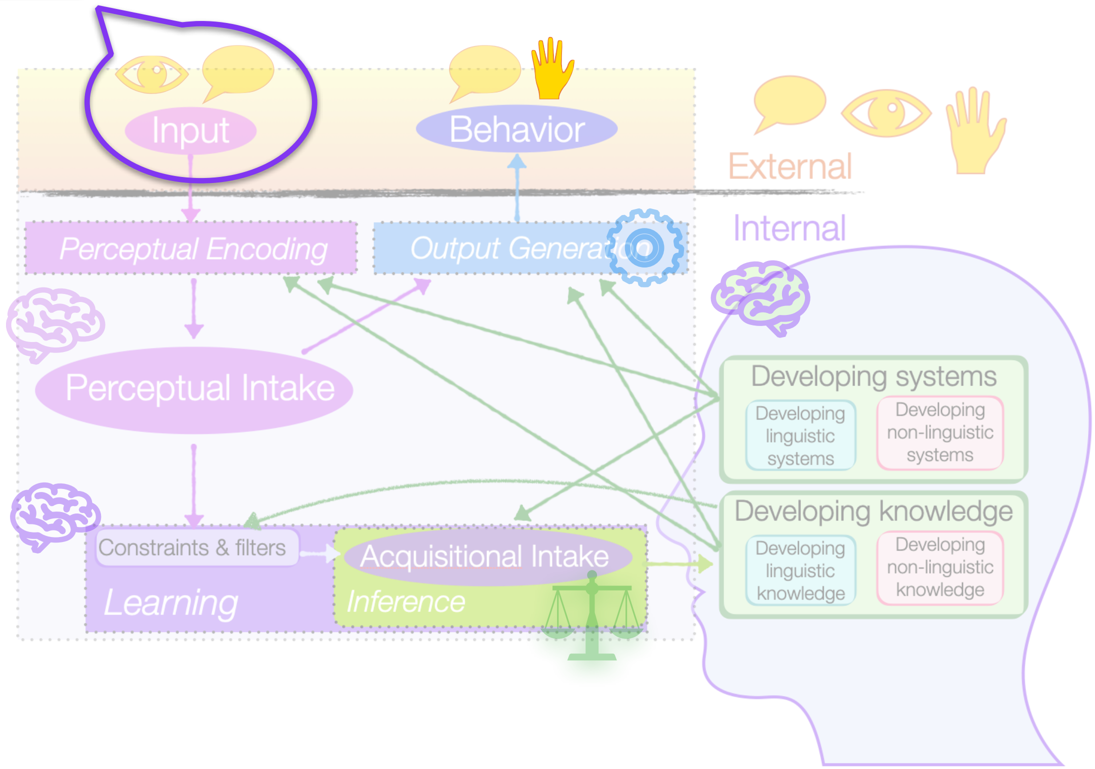


(18 months ≤ age < 60 months)

(≈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.



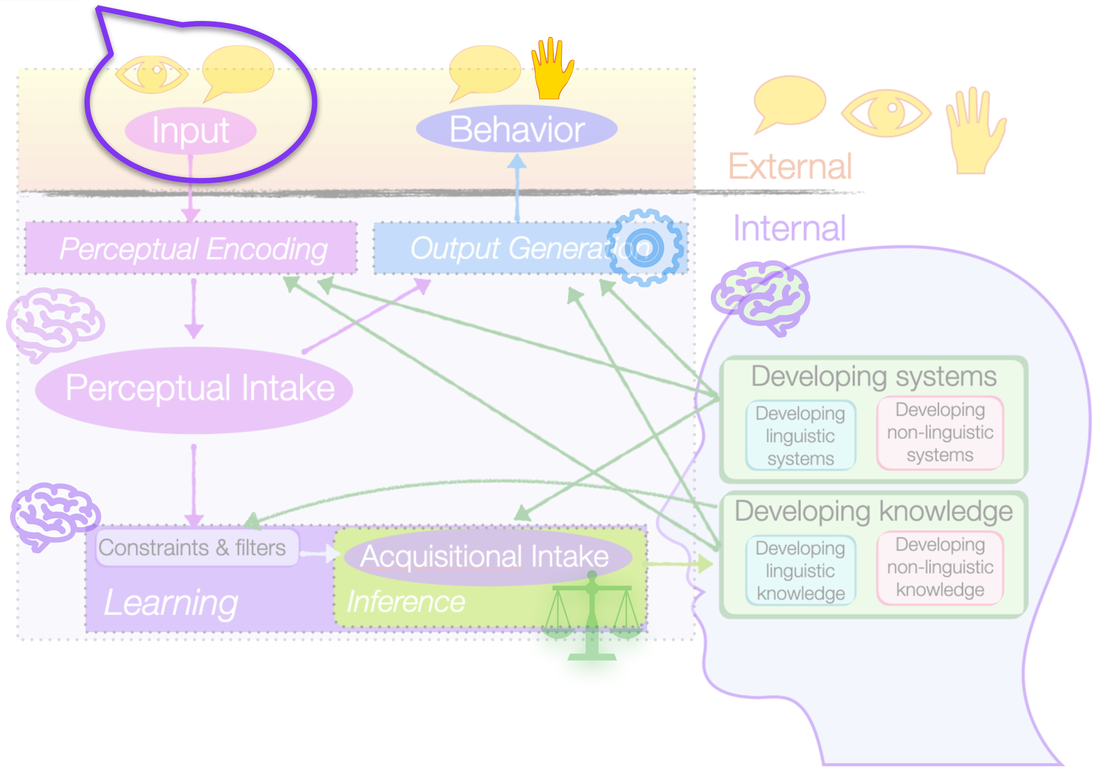
A more detailed look at the input

(18 months ≤ age < 60 months)

(≈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.



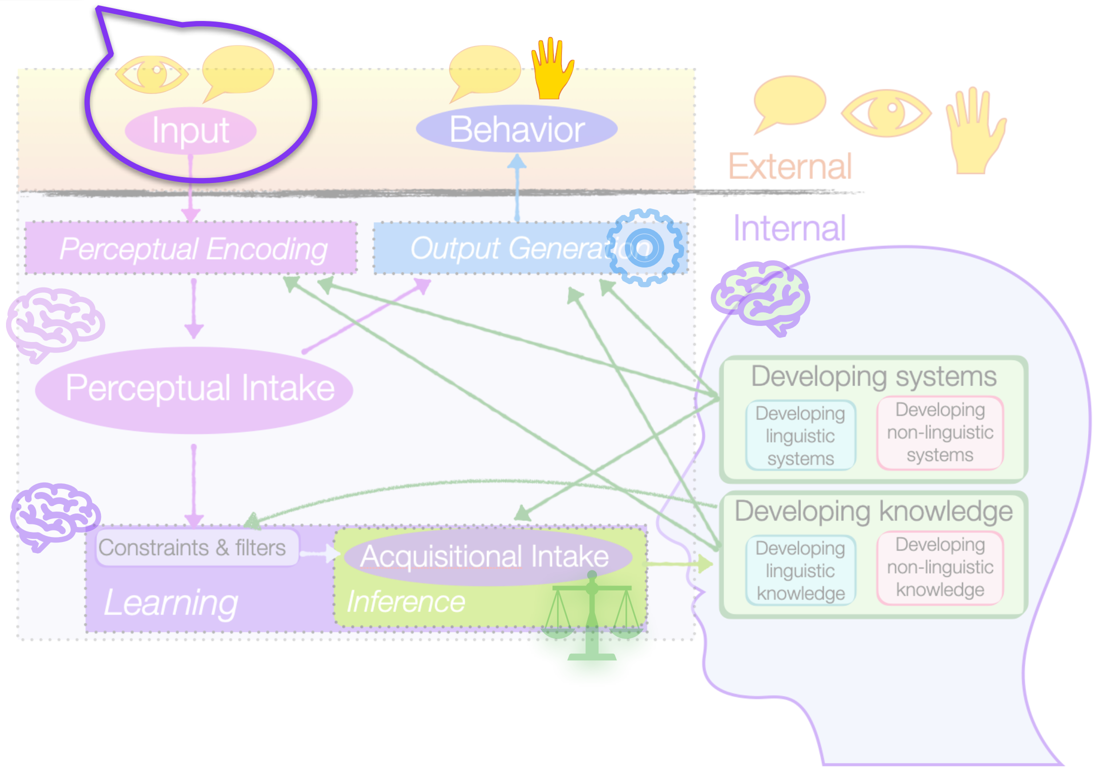
A more detailed look at the input

(18 months ≤ age < 60 months)

(≈10,442,258 minutes)

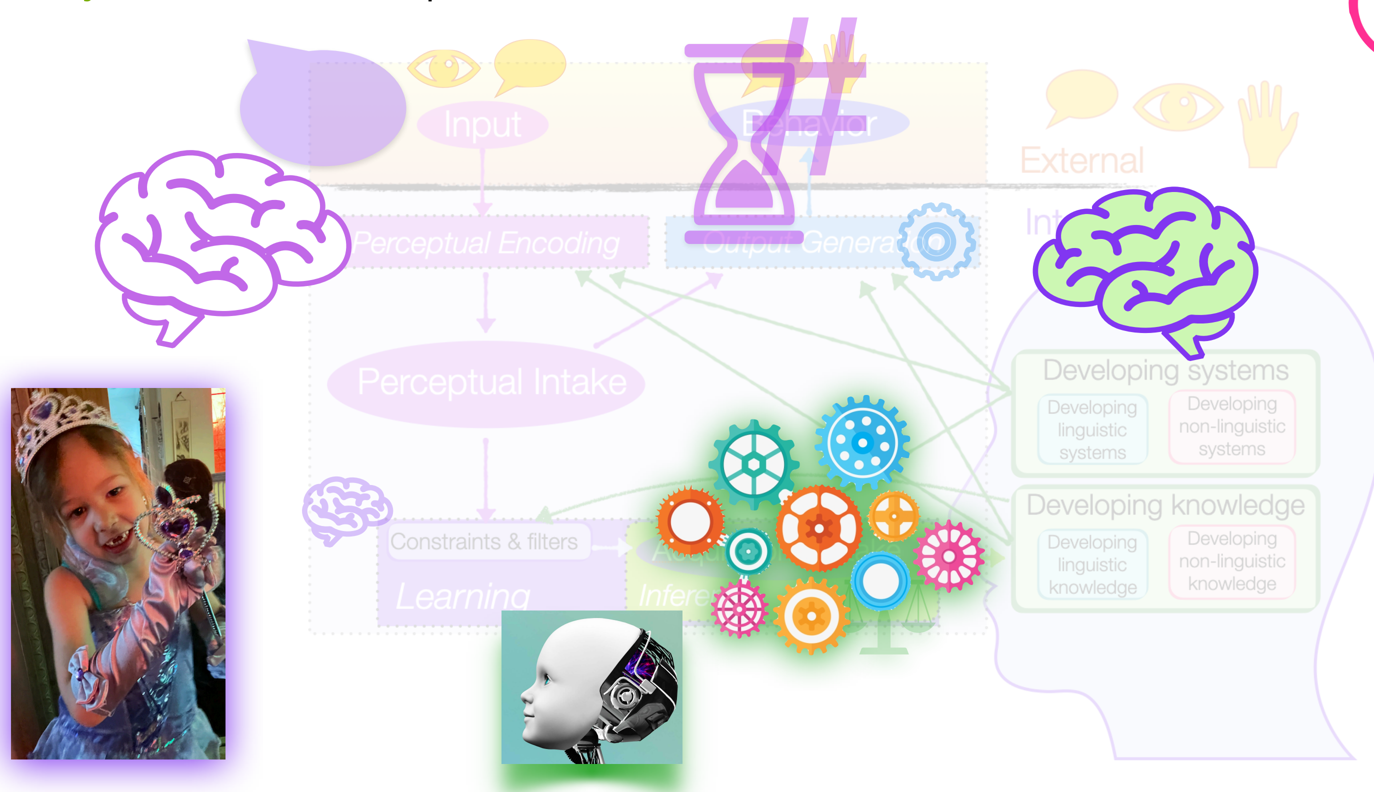
≈ 2.15 million *wh*-dependencies

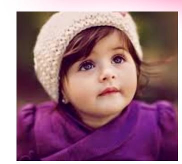
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.



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

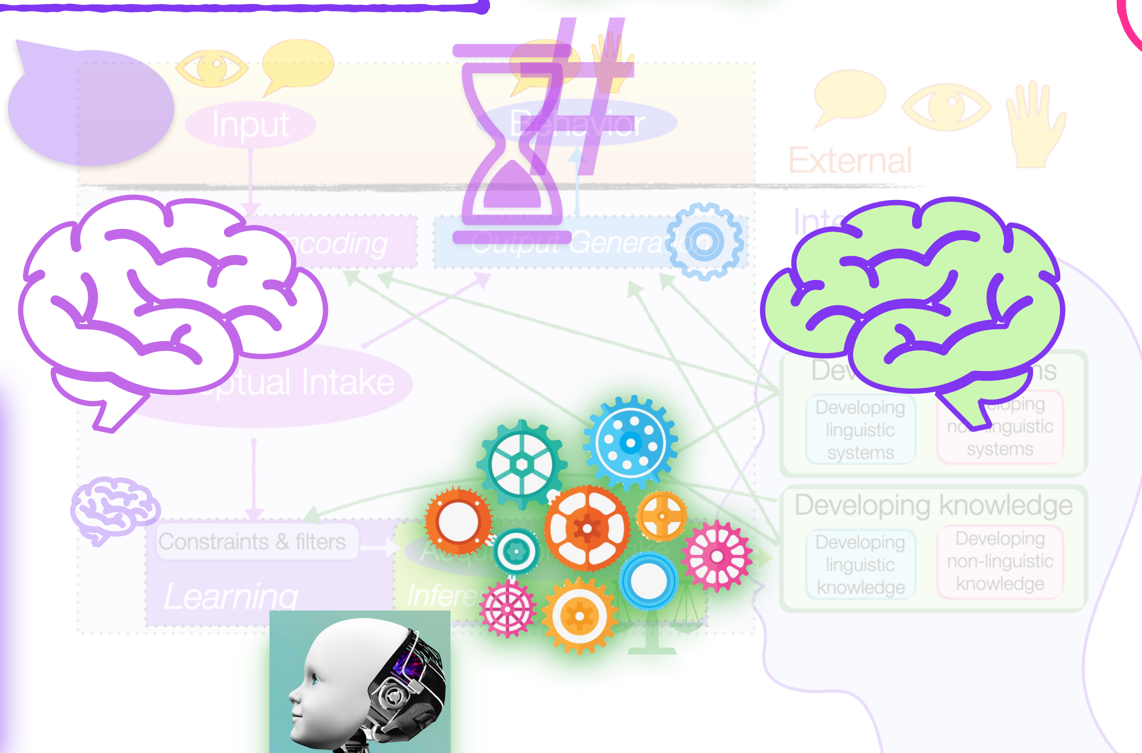
We also know a bit about the developing systems that impact the child's intake.

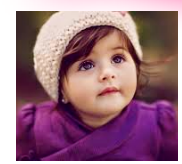




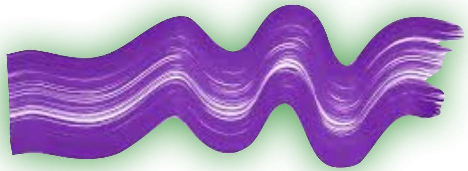
Dickson, Pearl, & Futrell 2022, 2024, in prep.

One known wrinkle:
Children's memory isn't perfect

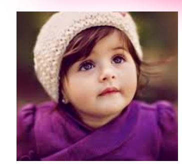


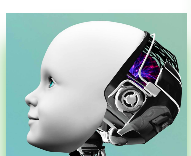


Dickson, Pearl, & Futrell 2022, 2024, in prep.



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

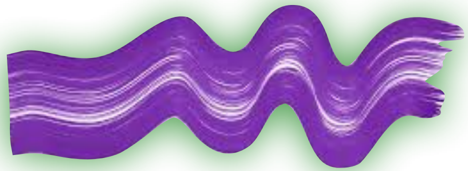




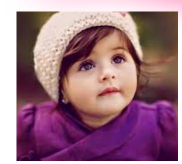




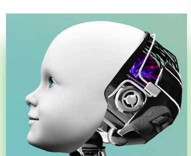

Dickson, Pearl, & Futrell 2022, 2024, in prep.

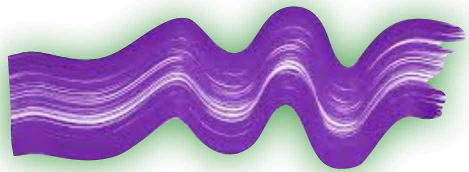

Children’s short term memory, along with related abilities like encoding information with context and maintaining attention, develops over time (Paris 1978, Gathercole et al. 2004, Fandakova et al. 2014).

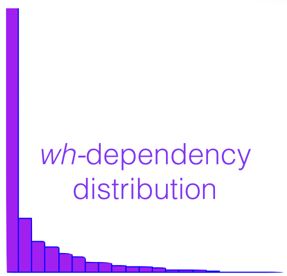




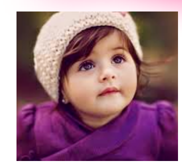


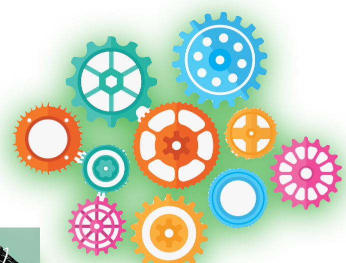
Dickson, Pearl, & Futrell 2022, 2024, in prep.

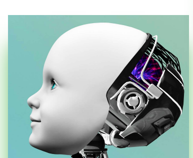




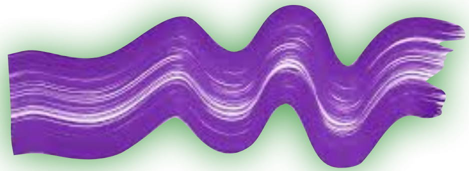
Upshot: Children might **not take in** all the *wh*-dependency distribution information in their input.

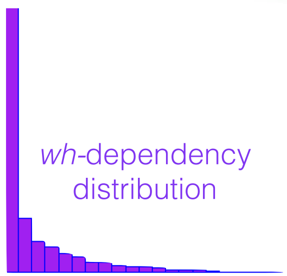




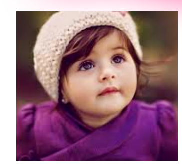


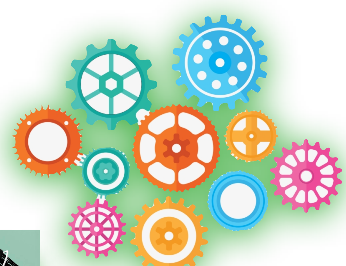
Dickson, Pearl, & Futrell 2022, 2024, in prep.

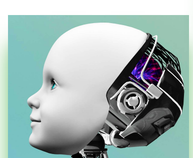


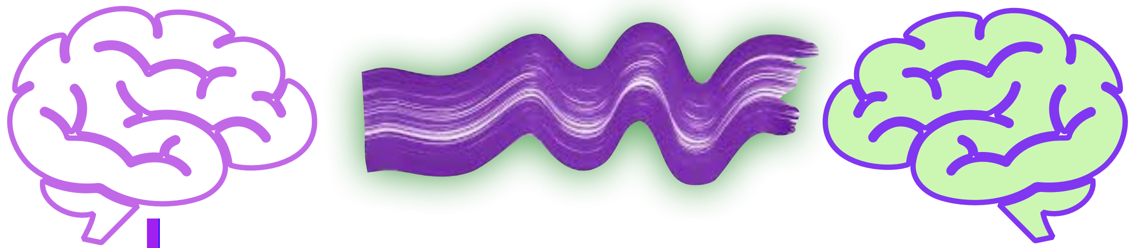


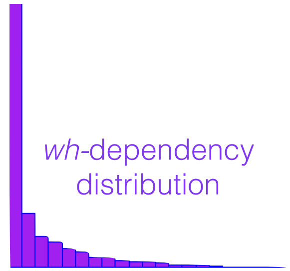
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.



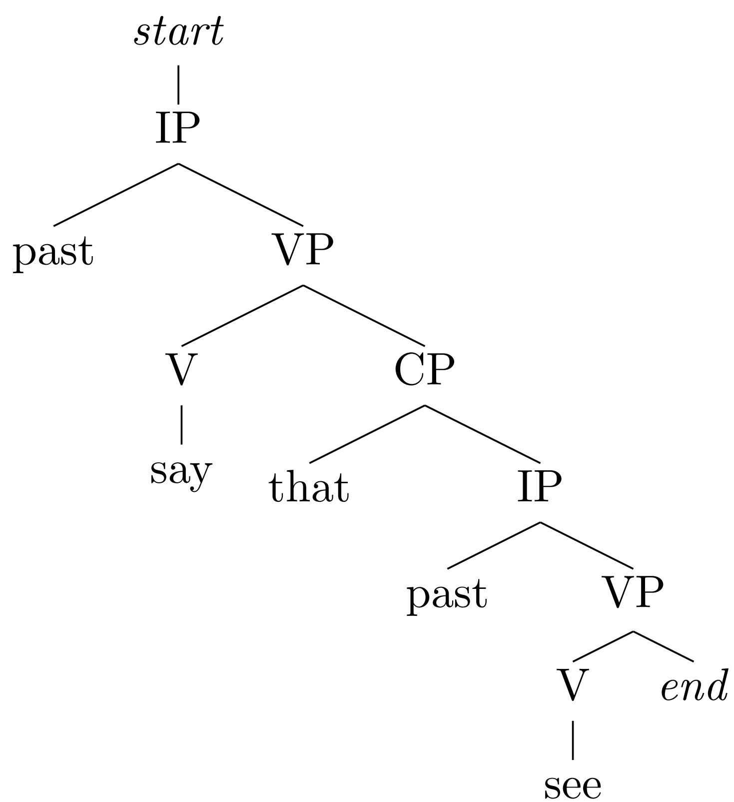


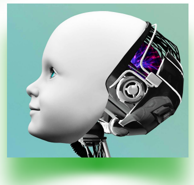




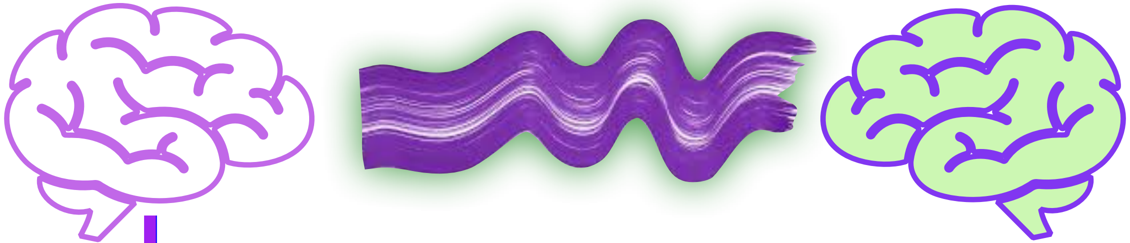


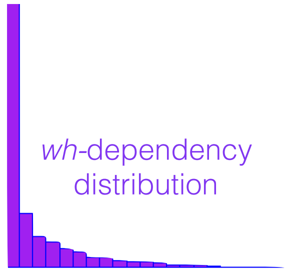
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.





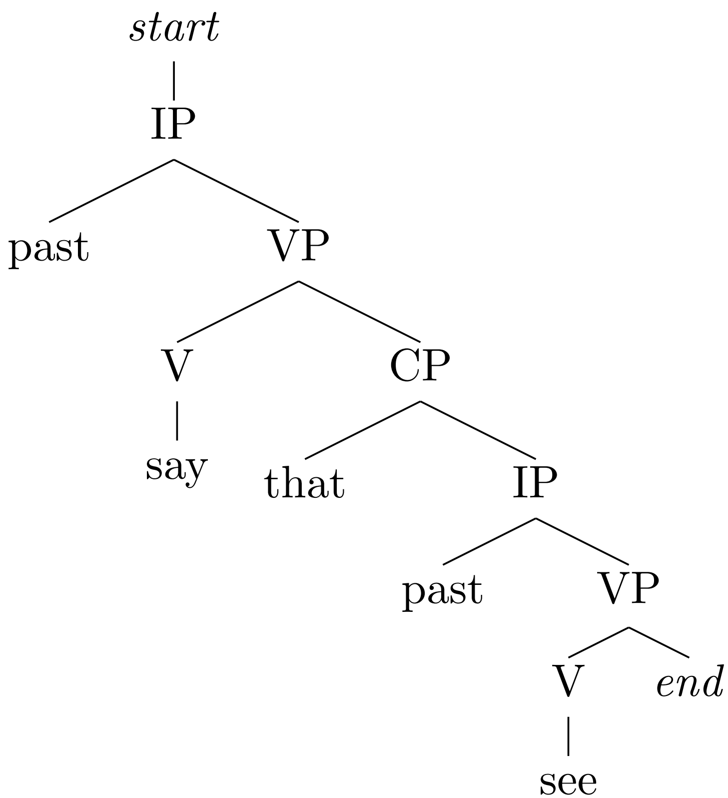
Dickson, Pearl, & Futrell 2024, in prep.

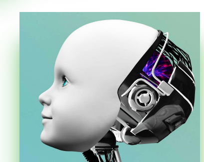




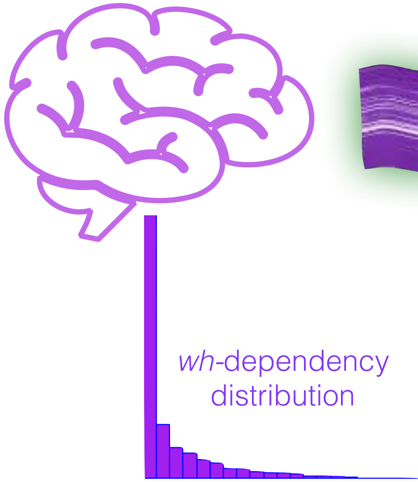
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

perfect intake



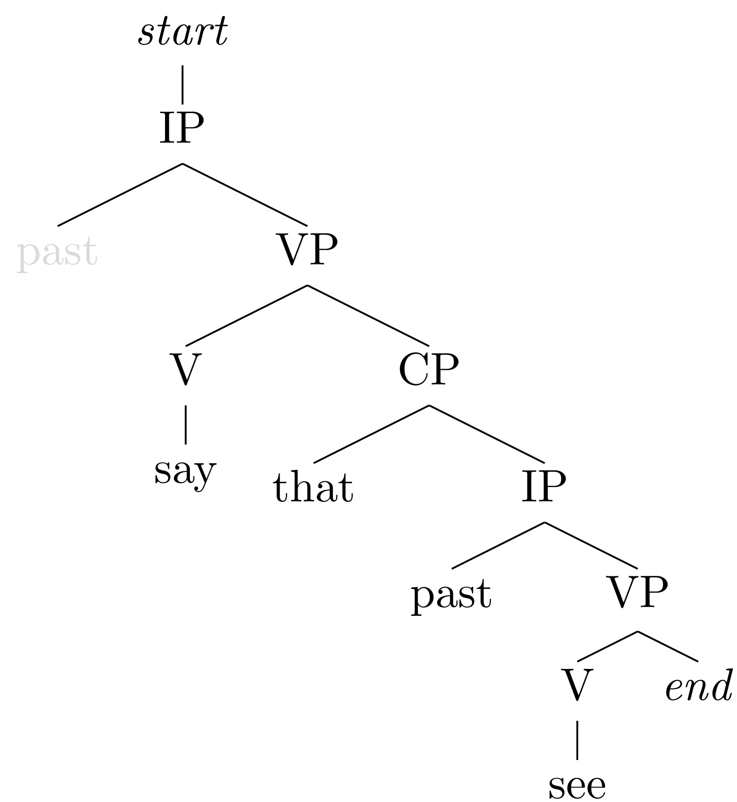


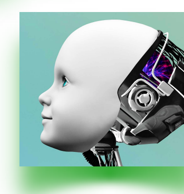
Dickson, Pearl, & Futrell 2024, in prep.



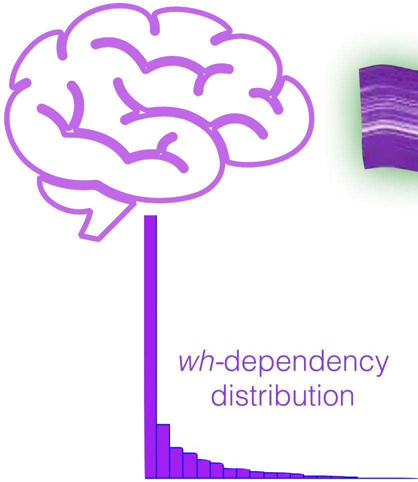
Learner intake:
Some parts of any particular
wh-dependency may be
forgotten in the moment.

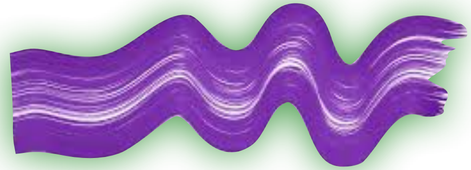
imperfect intake
(a little forgetting)





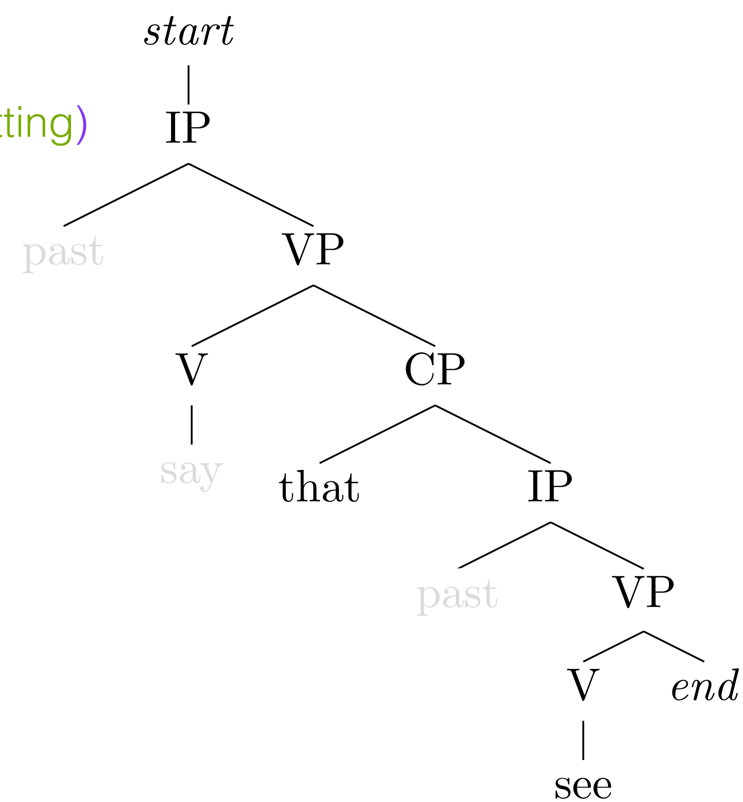
Dickson, Pearl, & Futrell 2024, in prep.

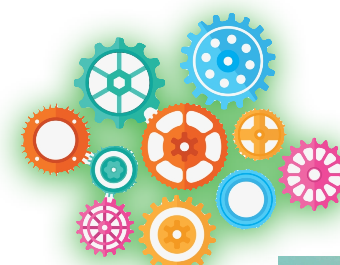


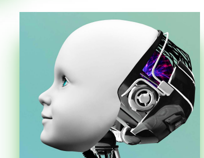


Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

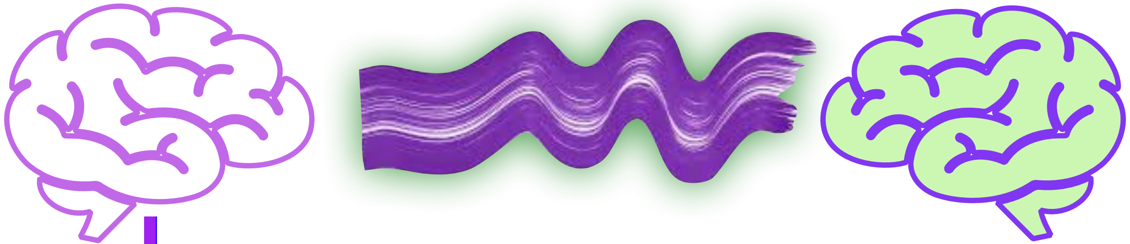
imperfect intake
(a little more forgetting)

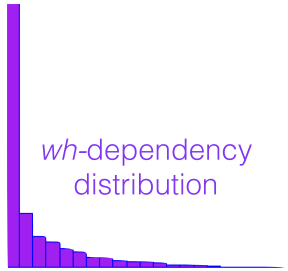




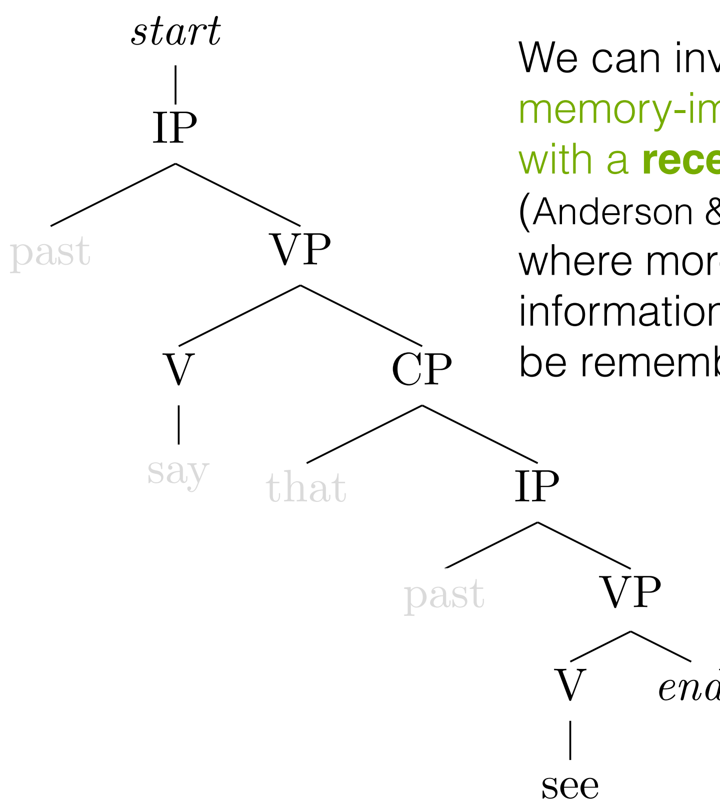


Dickson, Pearl, & Futrell 2024, in prep.

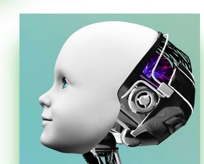




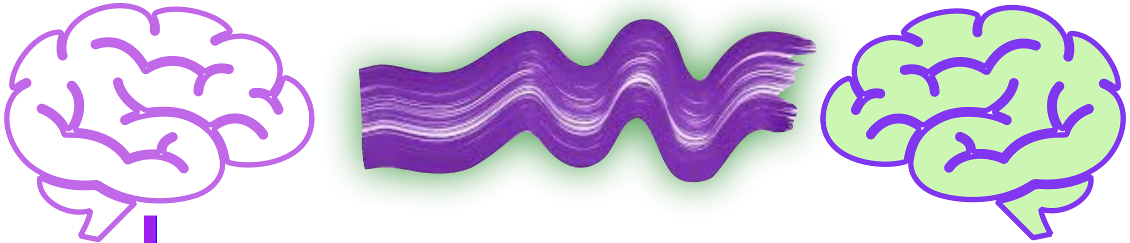
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

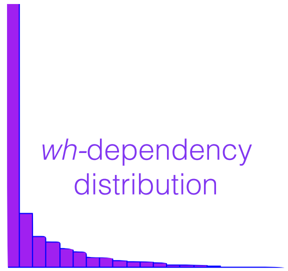


We can investigate memory-impacted learners with a **recency effect** (Anderson & Milson 1989), where more recent information is more likely to be remembered.



Dickson, Pearl, & Futrell 2024, in prep.

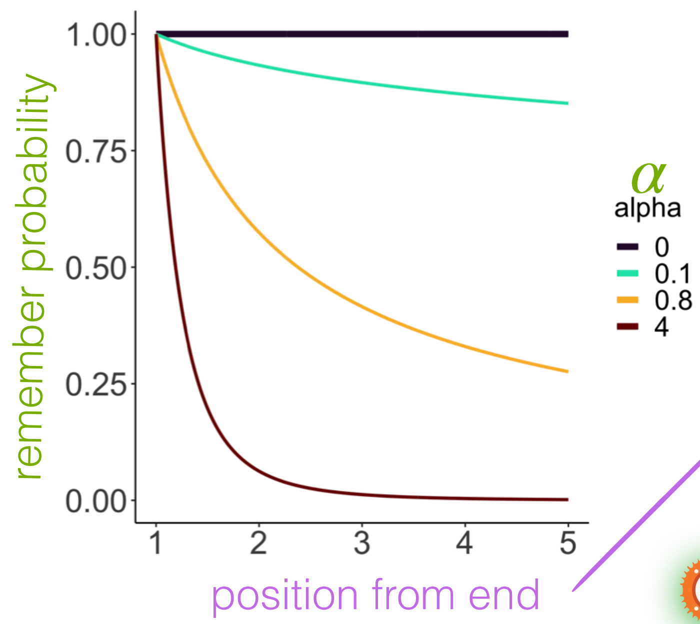


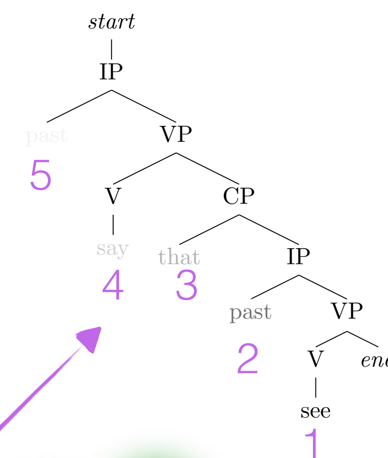


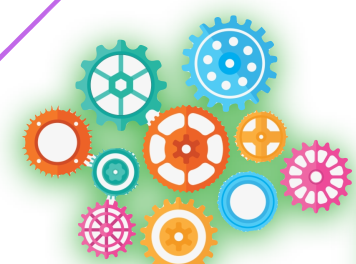
Learner intake:
Some parts of any particular *wh*-dependency may be forgotten in the moment.

Anderson & Milson 1989

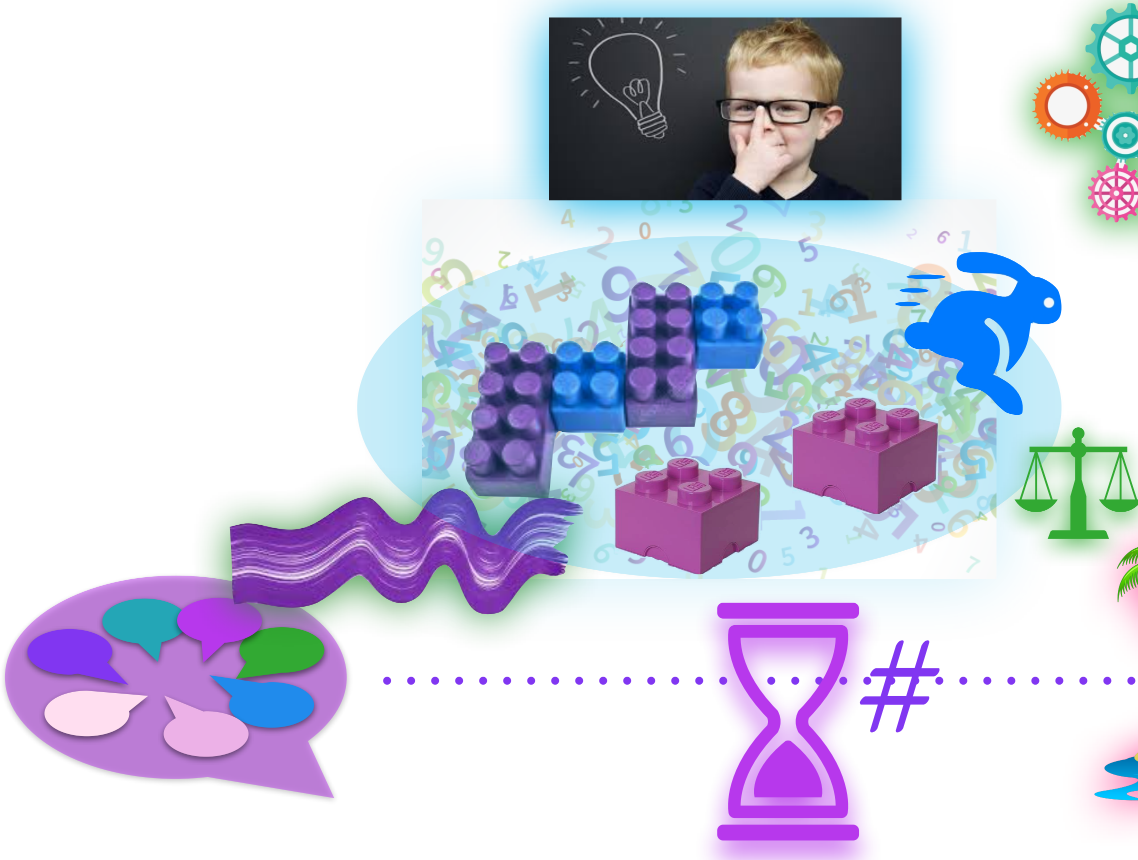
$$\frac{1}{(position + 1)^{\alpha}}$$

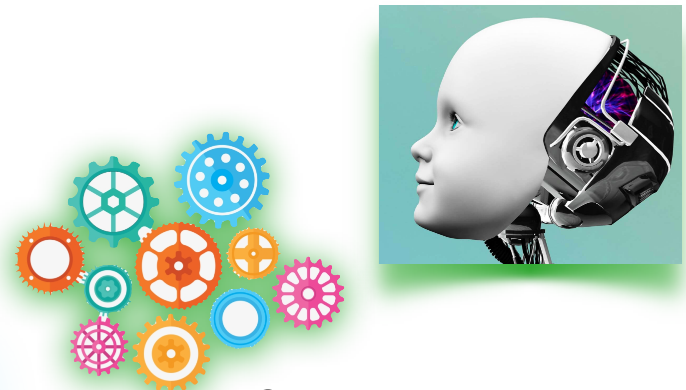




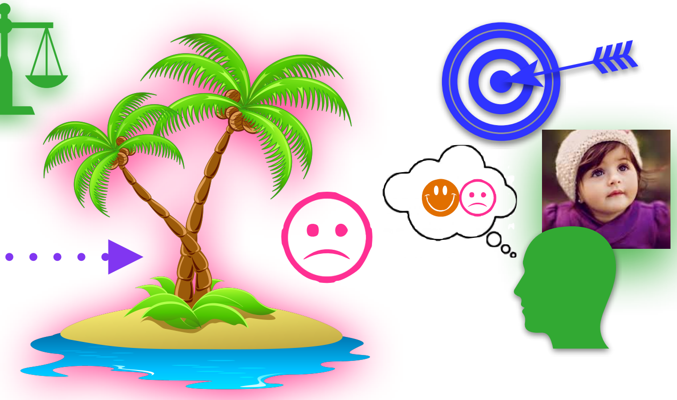


Dickson, Pearl, & Futrell 2024, in prep.



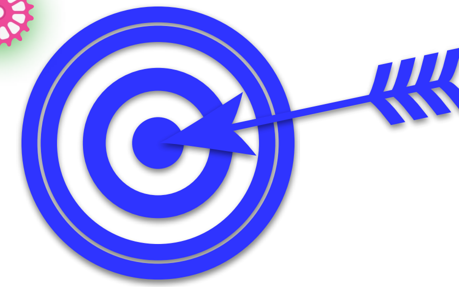


Can the modeled learner produce the appropriate observable behavior?

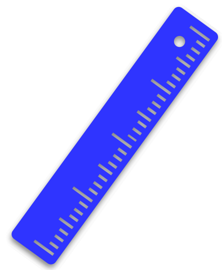


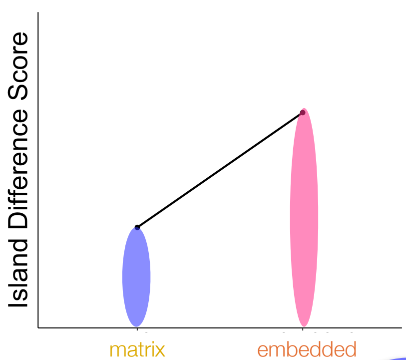
Pearl & Sprouse 2013, Bates & Pearl 2019, Pearl & Bates 2022
Dickson et al. 2022, 2024, in prep.

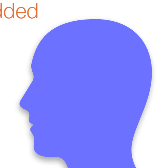
Dickson, Pearl, & Futrell 2022, 2024, in prep.

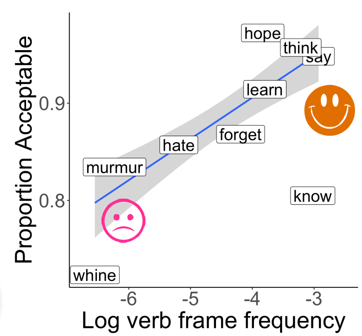


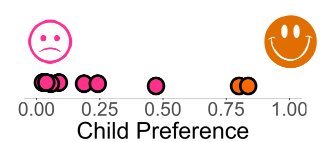
A reminder of the measurable target behavior patterns

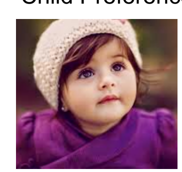




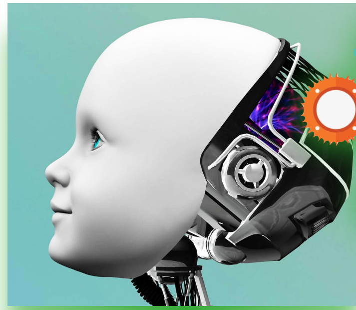


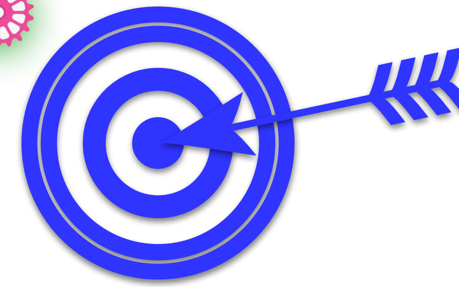




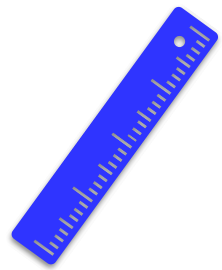


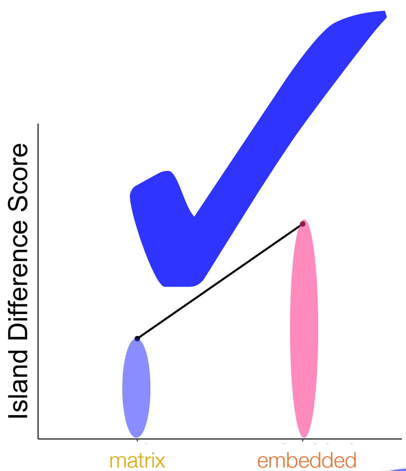
Dickson, Pearl, & Futrell 2022, 2024, in prep.

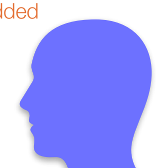


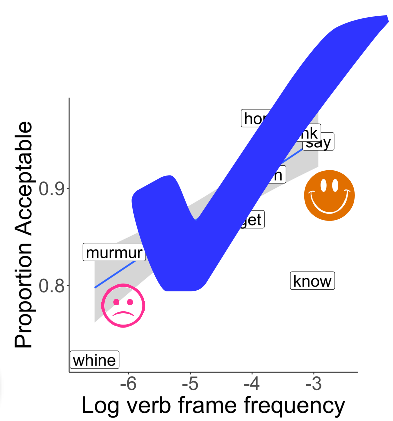


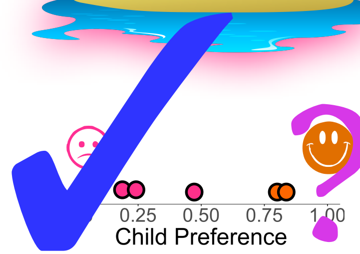
Our modeled child can generate almost all of these target patterns.



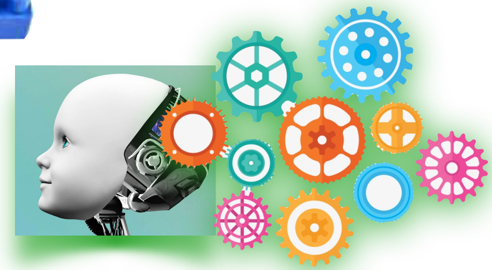


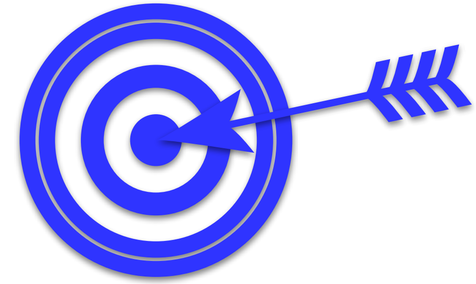






Dickson, Pearl, & Futrell 2022, 2024, in prep.





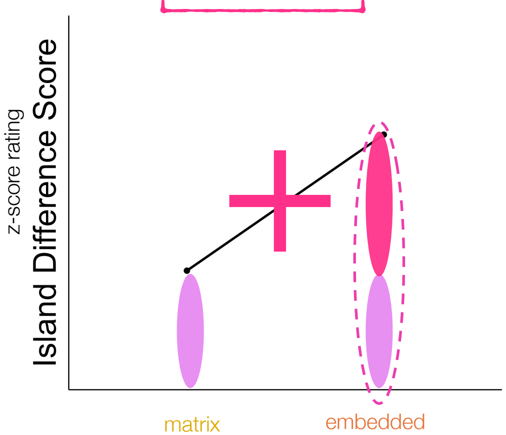
A closer look at these target patterns.

Subject island

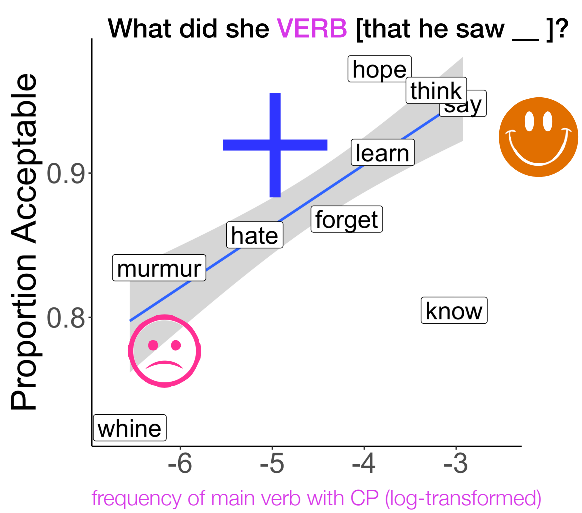
Complex NP island

Whether island

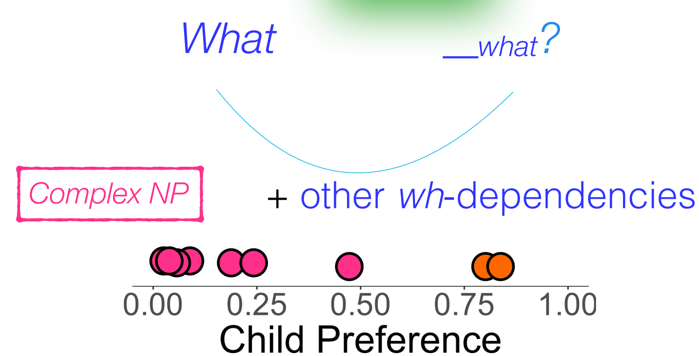
Adjunct island



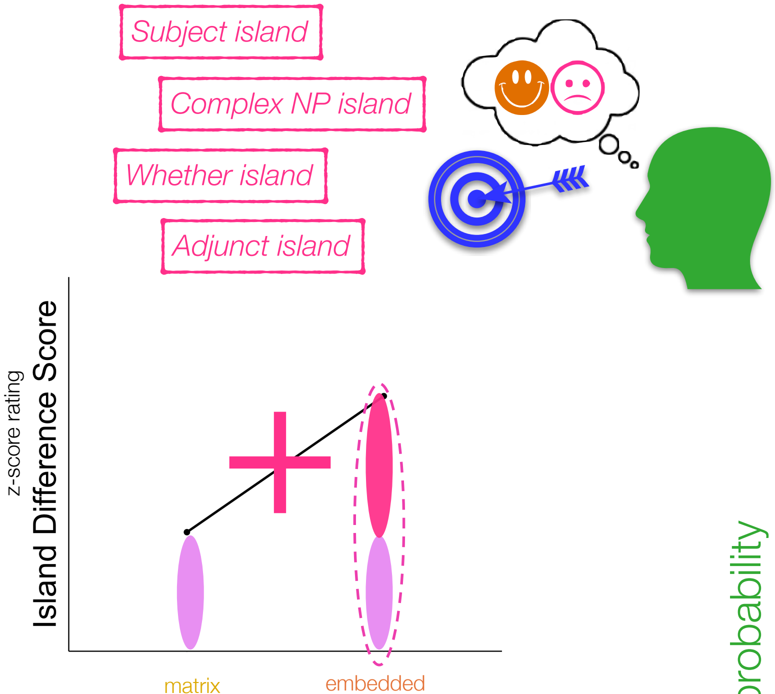
Sprouse et al. 2012



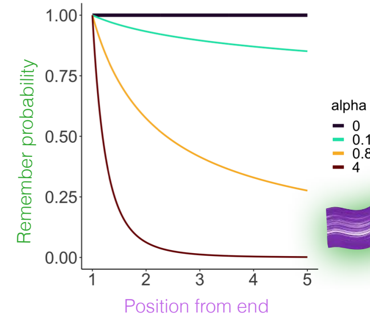
Liu et al. 2019, 2022

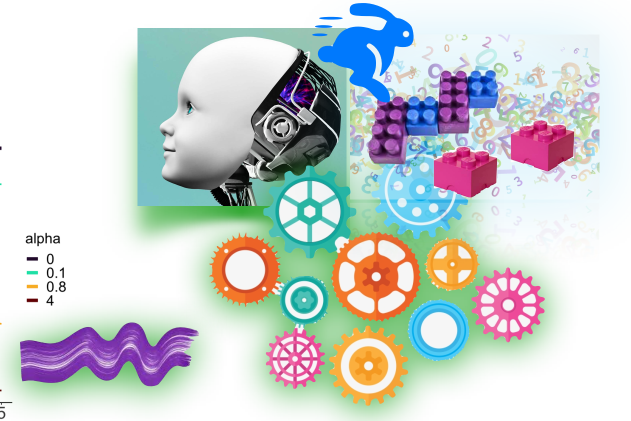


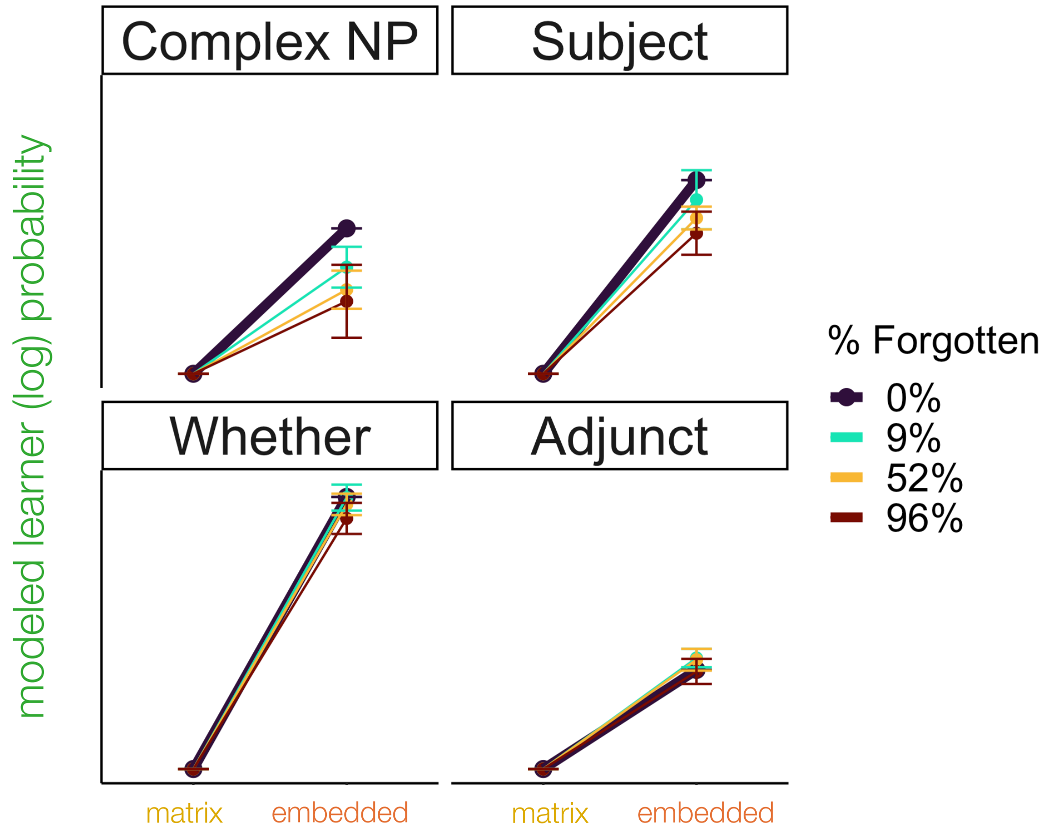
De Villiers et al. 2008



Sprouse et al. 2012







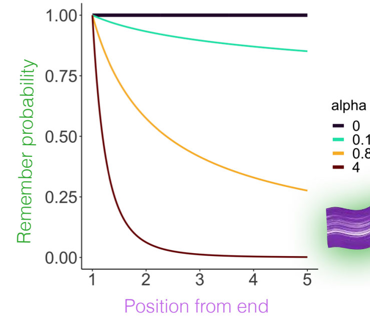
Dickson et al. 2022, 2024, in prep.

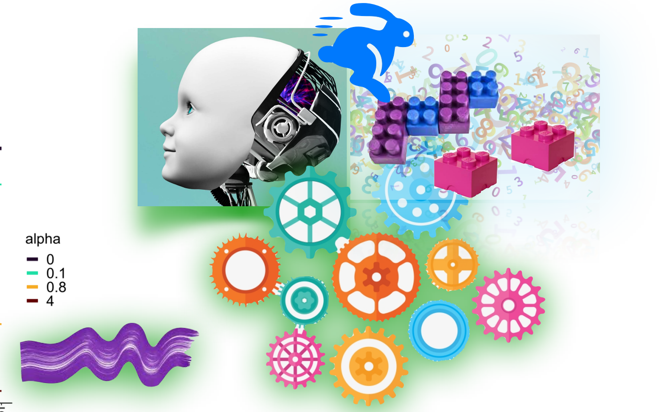
Subject island

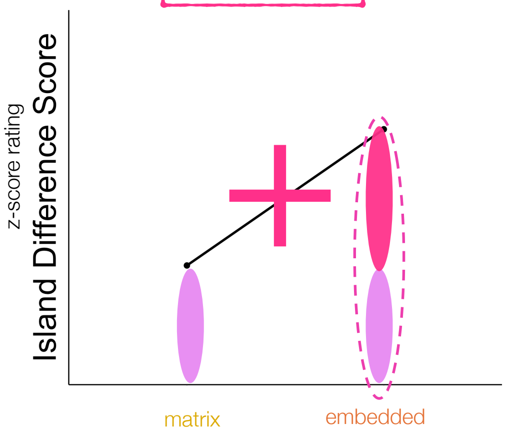
Complex NP island

Whether island

Adjunct island

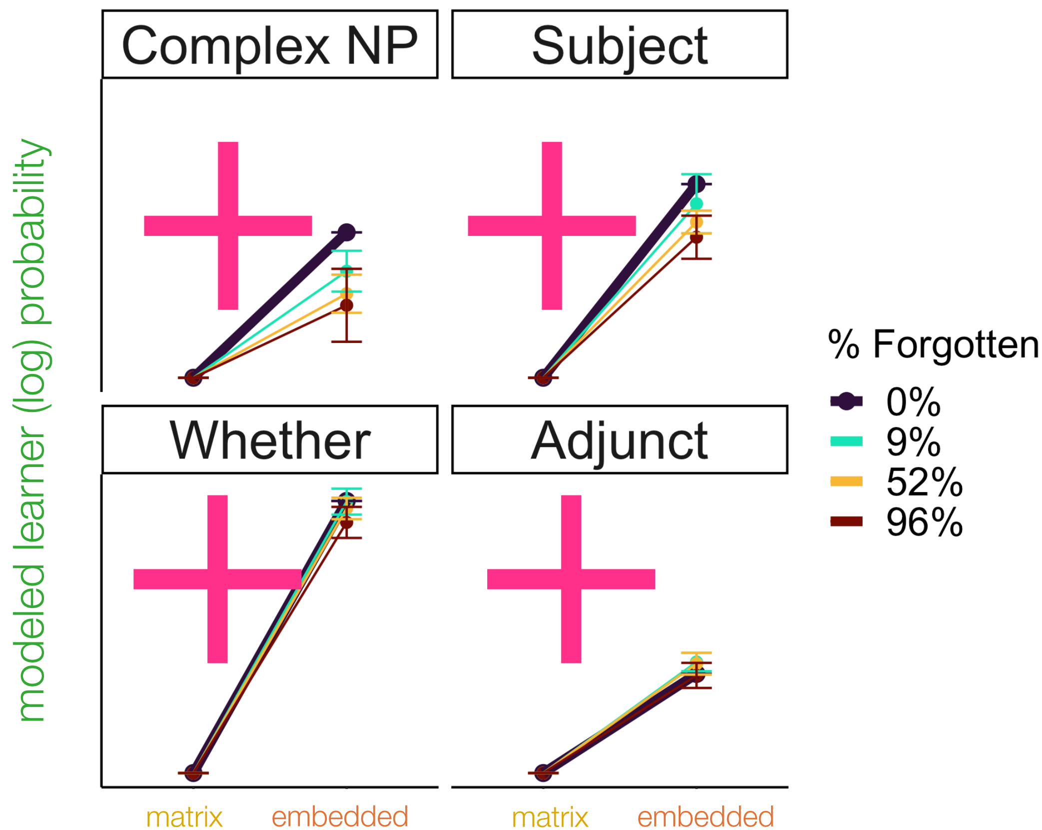






Sprouse et al. 2012

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



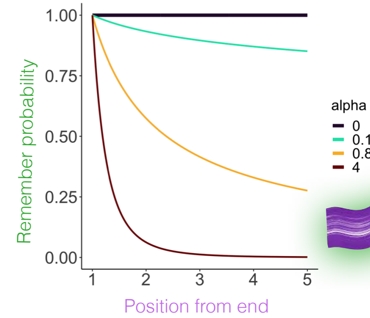
Dickson et al. 2022, 2024, in prep.

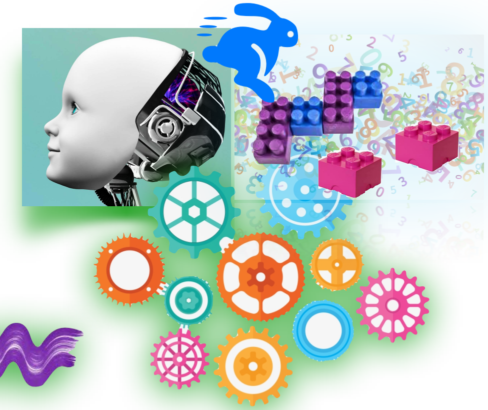
Subject island

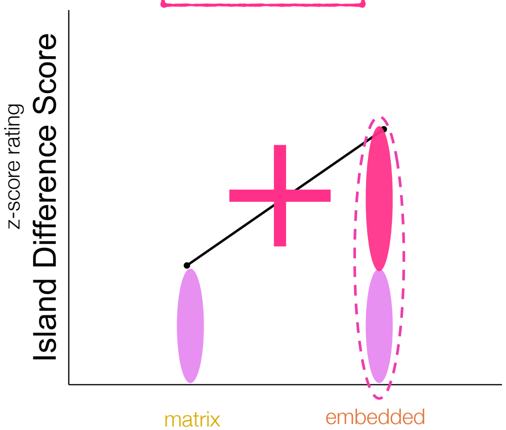
Complex NP island

Whether island

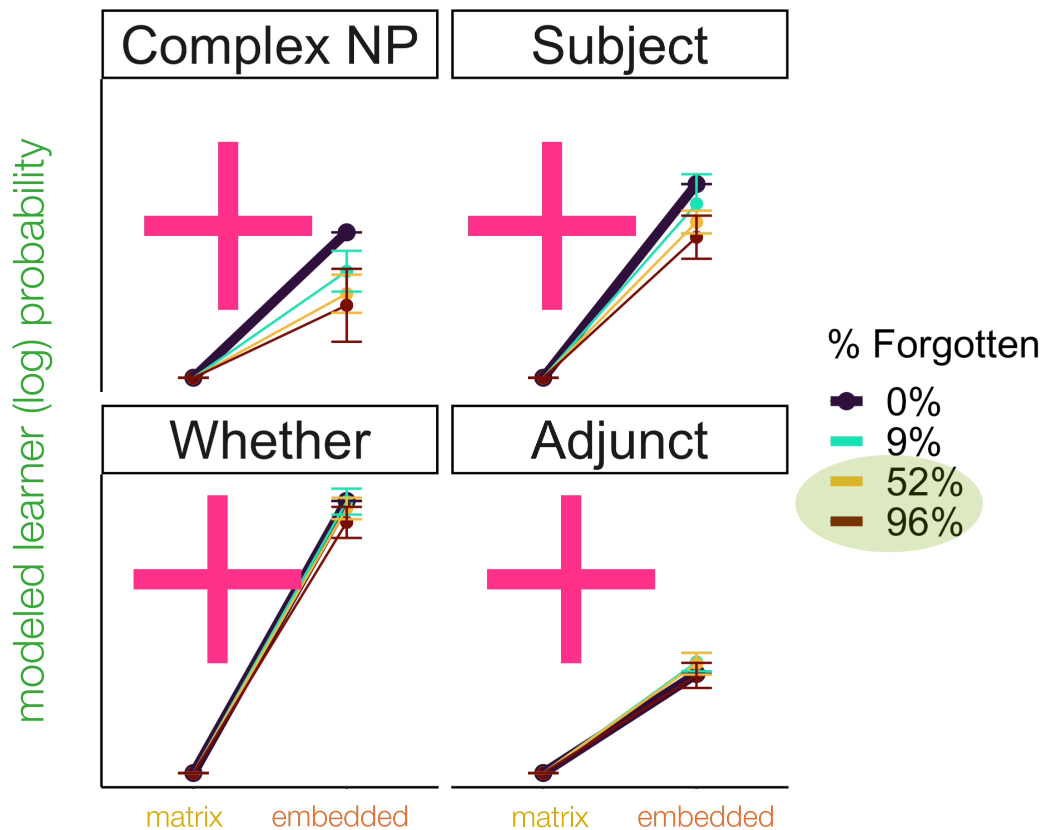
Adjunct island







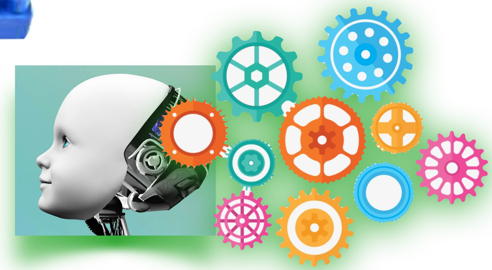
Sprouse et al. 2012

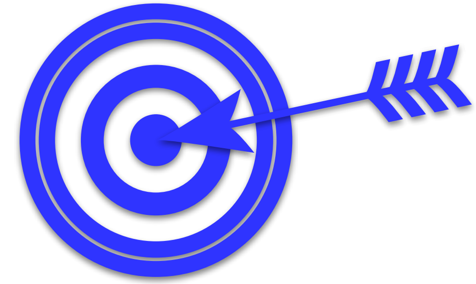


This happens even if they forget a whole lot.

Dickson et al. 2022, 2024, in prep.

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





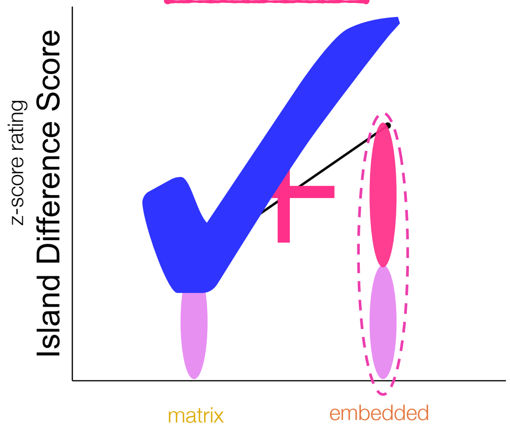
A closer look at these target patterns.

Subject island

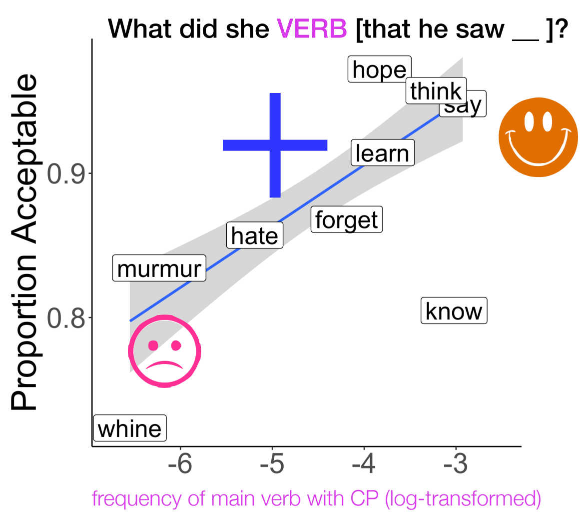
Complex NP island

Whether island

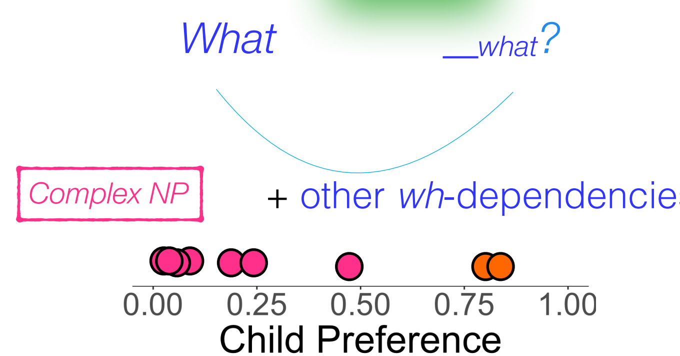
Adjunct island



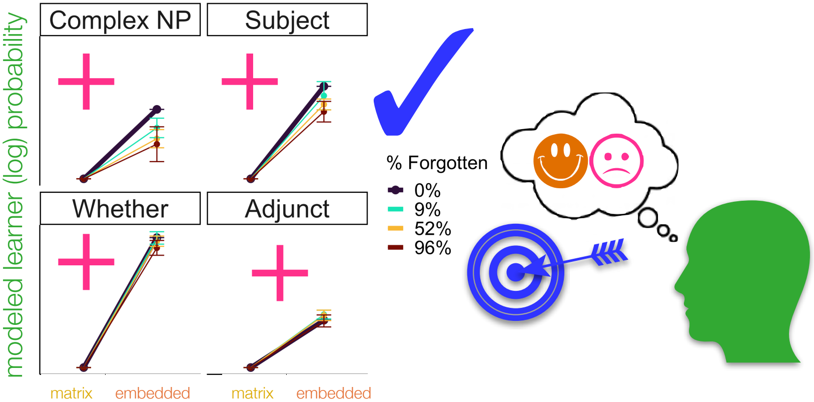
Sprouse et al. 2012

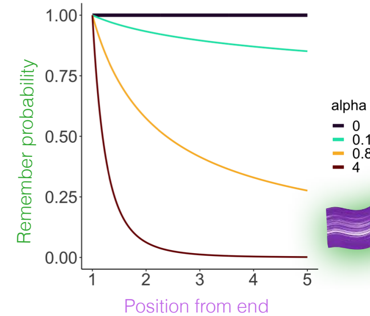


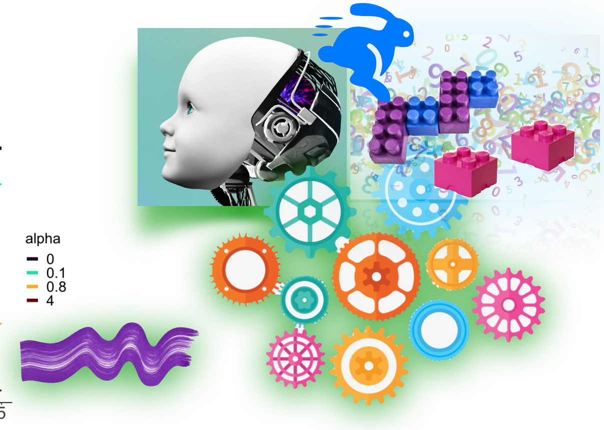
Liu et al. 2019, 2022

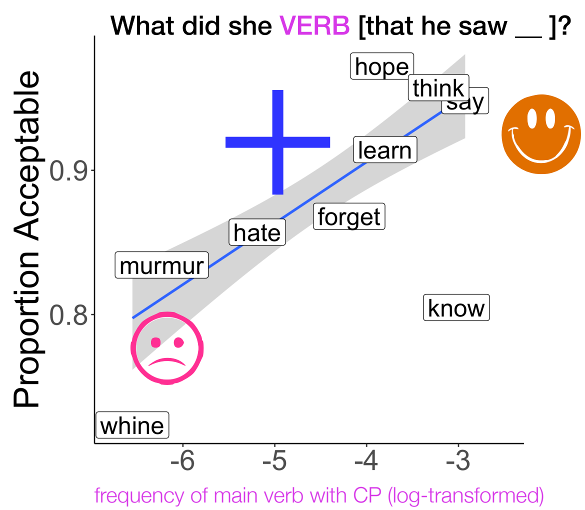


De Villiers et al. 2008



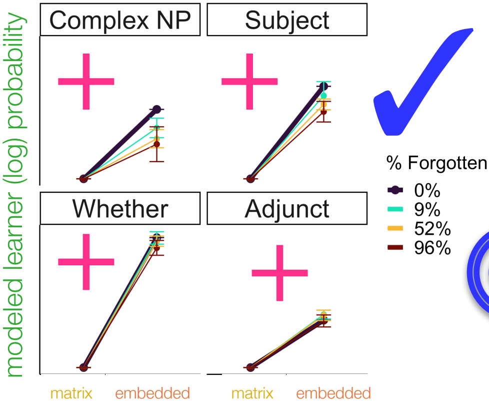


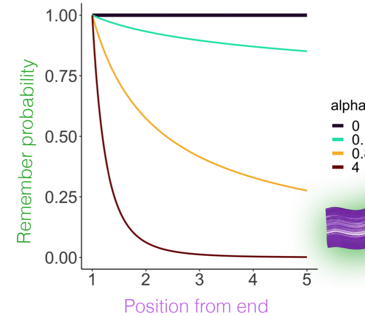


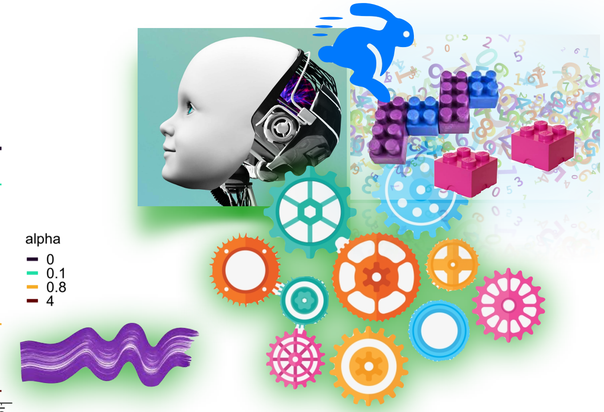


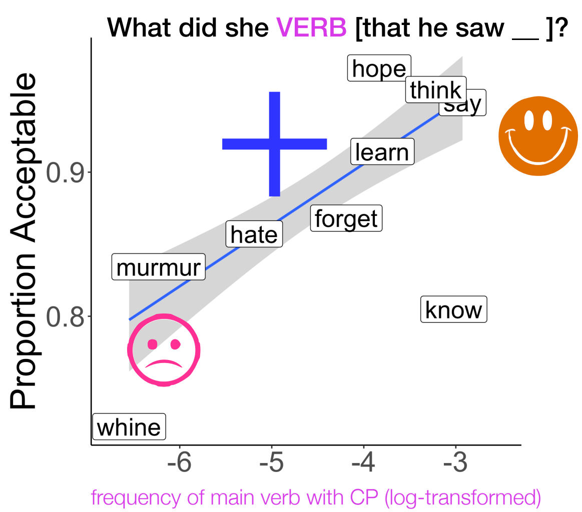
Liu et al. 2019, 2022

Dickson et al. 2022, 2024, in prep.

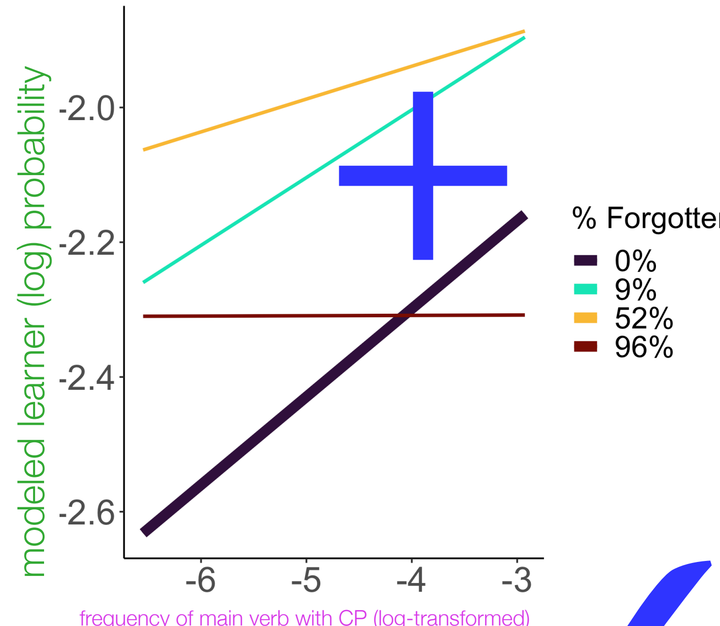






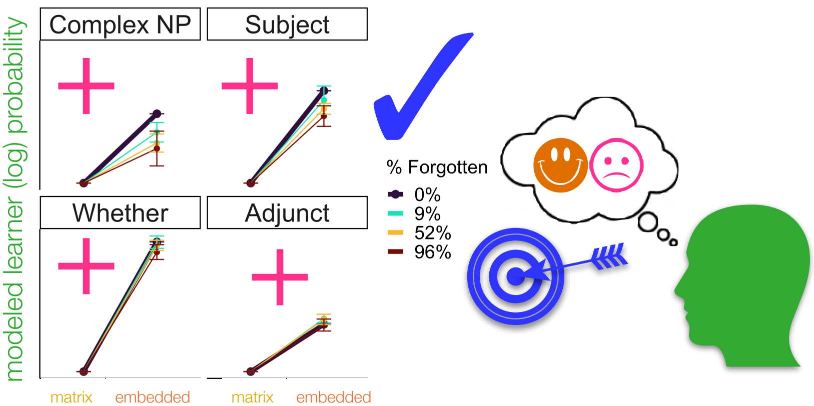


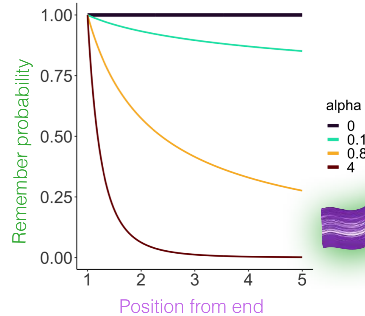
Liu et al. 2019, 2022

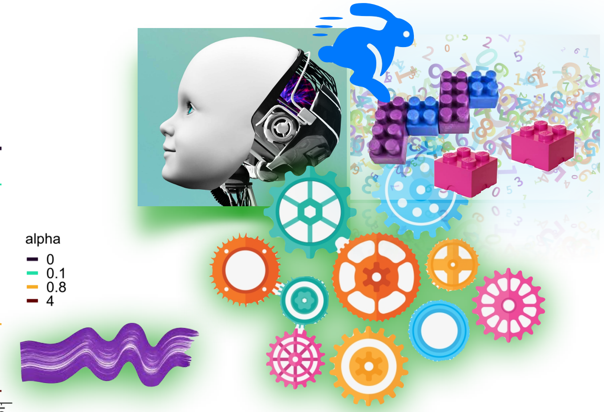


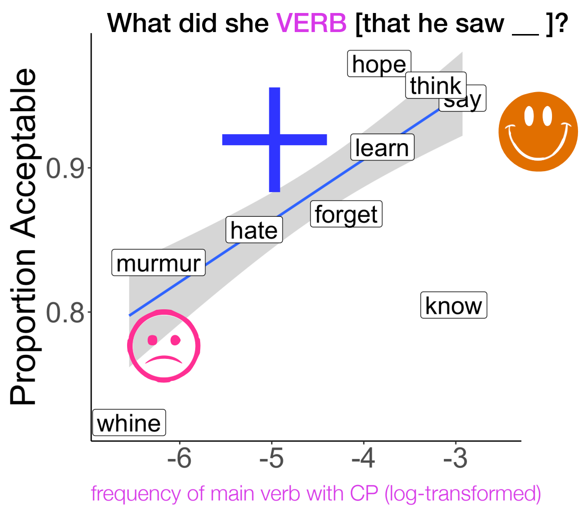
Most modeled learners can predict the observed pattern (positive correlation with frequency).

Dickson et al. 2022, 2024, in prep.

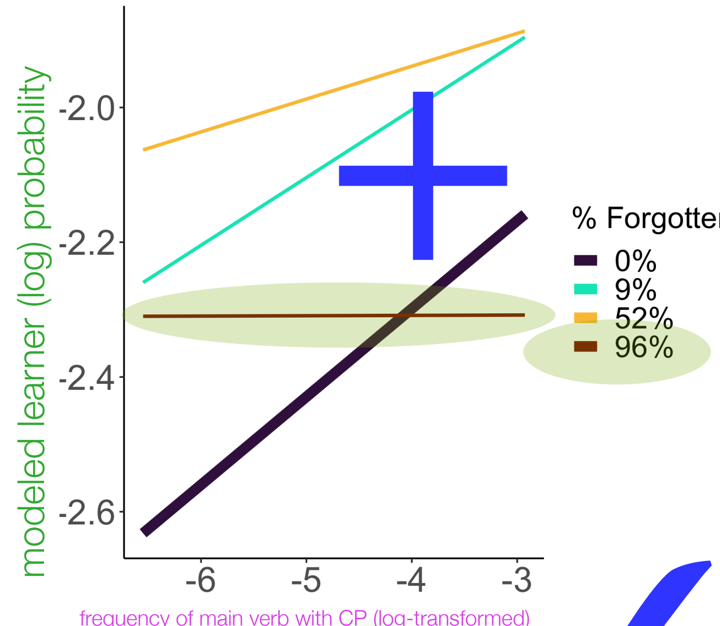


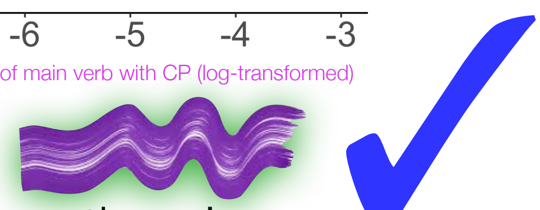






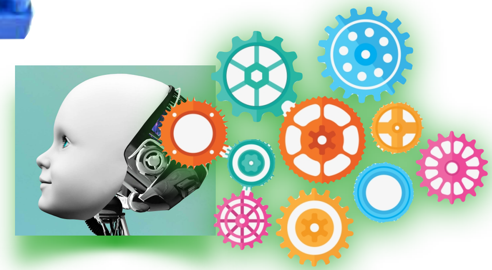
Liu et al. 2019, 2022



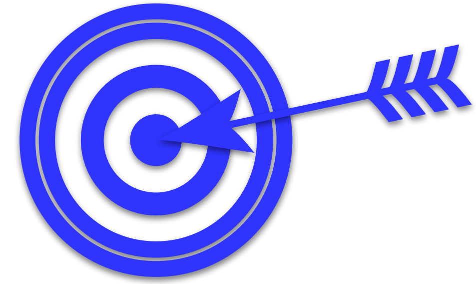


This happens, except when there's an extraordinary amount of forgetting.

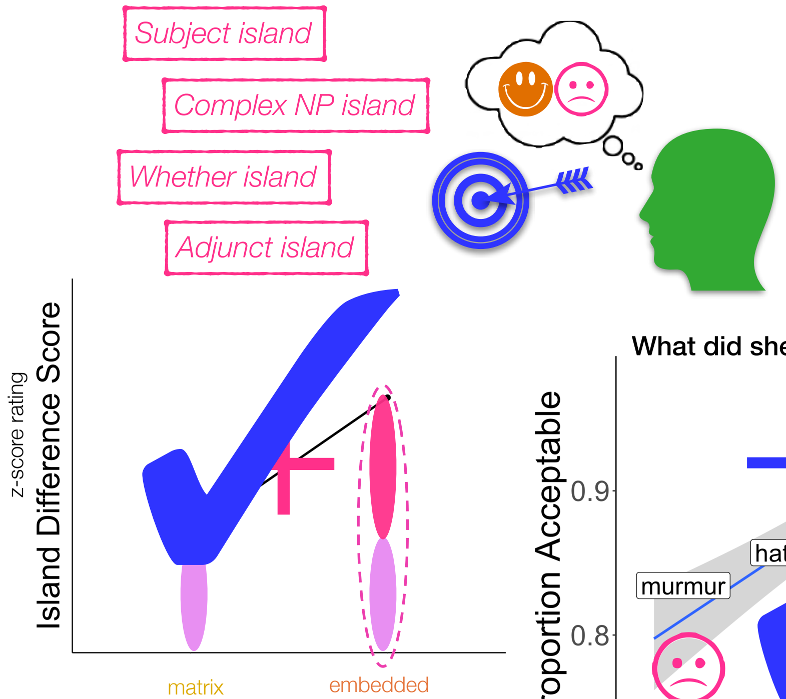
Dickson et al. 2022, 2024, in prep.



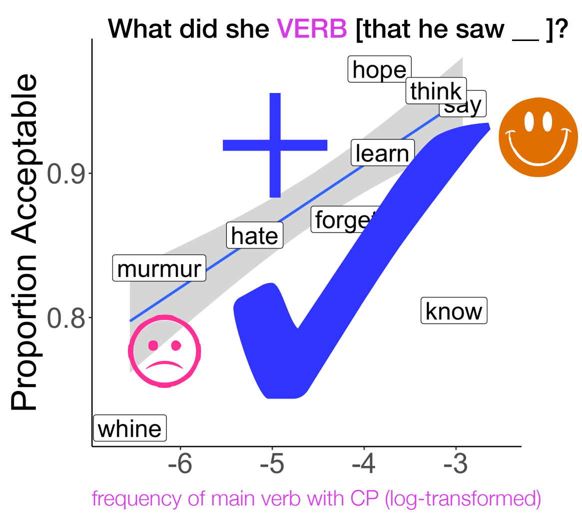
Dickson, Pearl, & Futrell 2022, 2024, in prep.



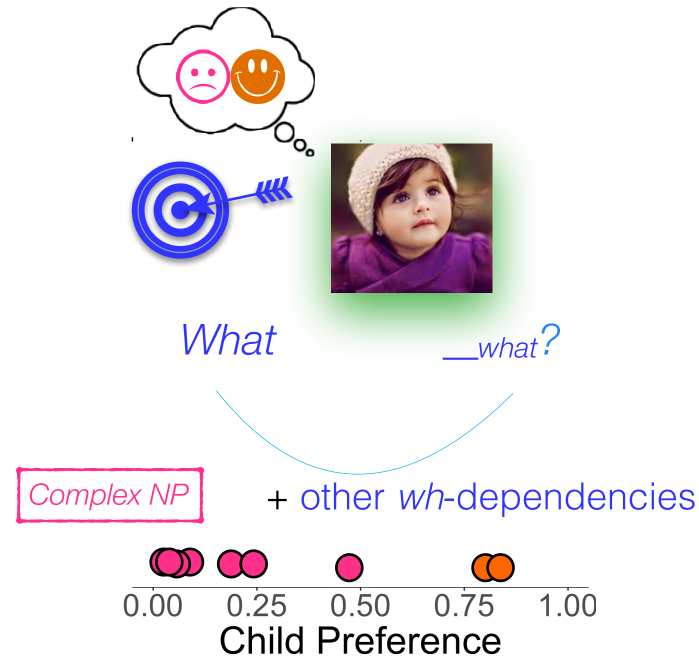
A closer look at these target patterns.



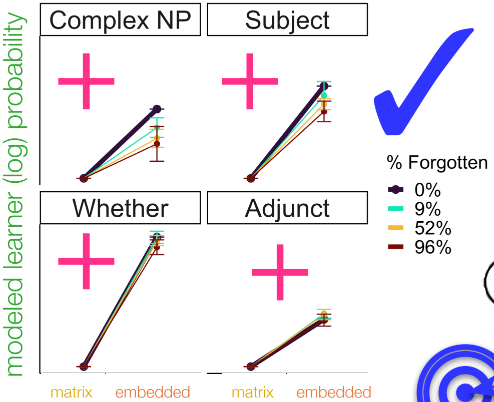
Sprouse et al. 2012

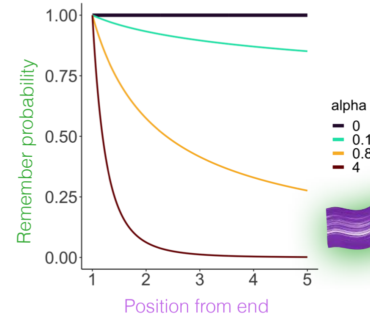


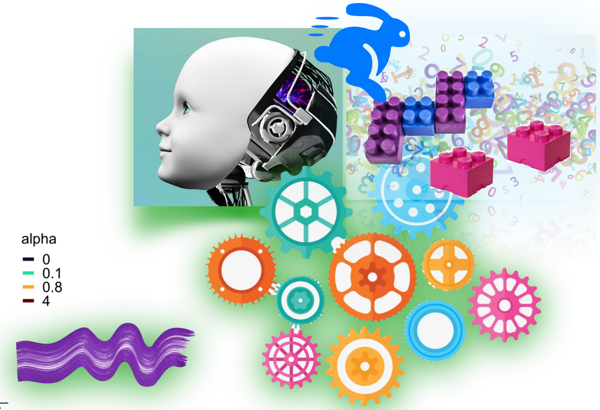
Liu et al. 2019, 2022

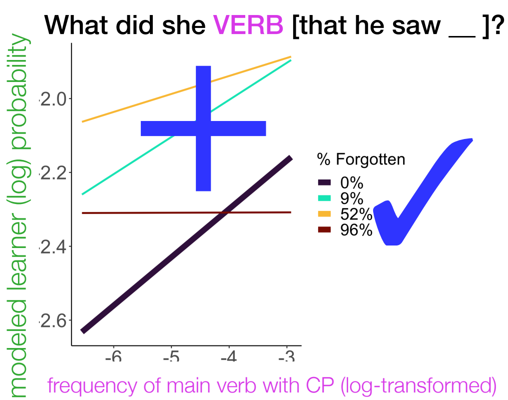


De Villiers et al. 2008

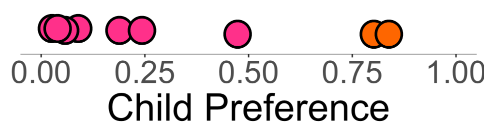






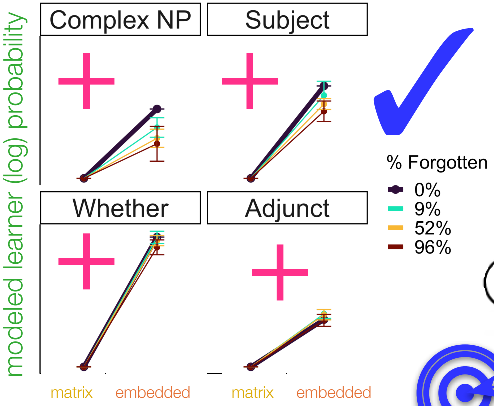


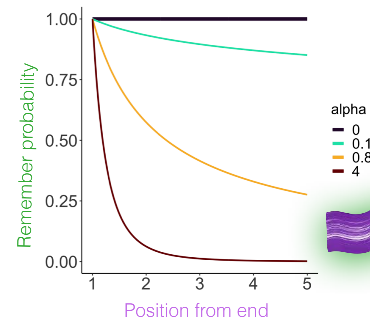
Complex NP + other *wh*-dependencies

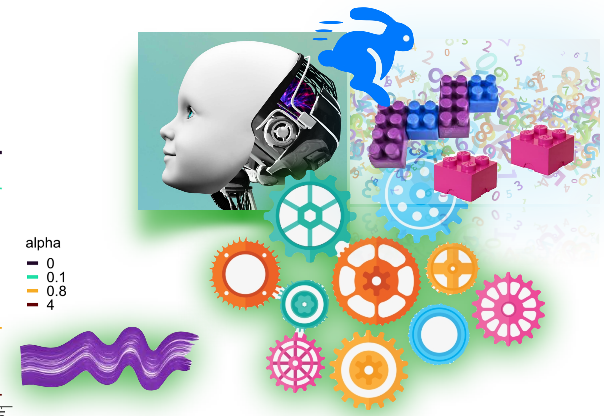


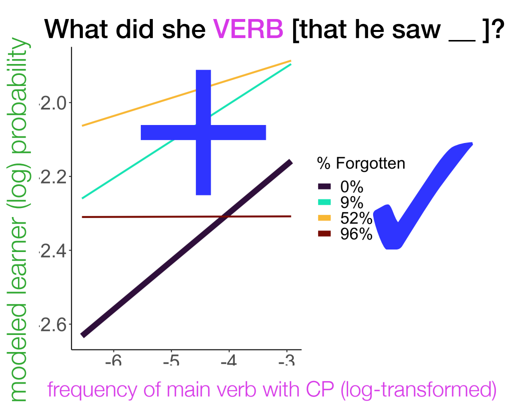
De Villiers et al. 2008

Dickson et al. 2022, 2024, in prep.

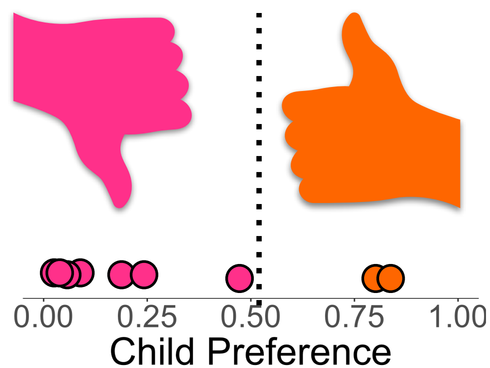








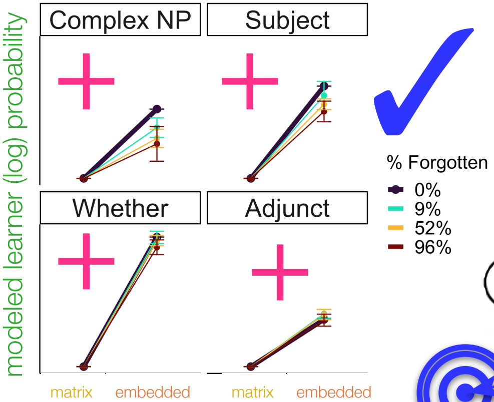
Complex NP + other *wh*-dependencies

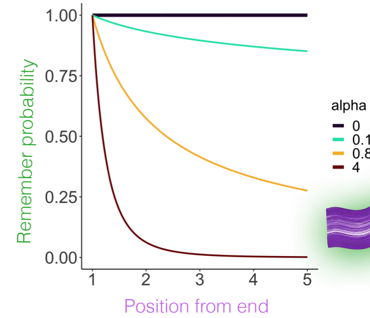


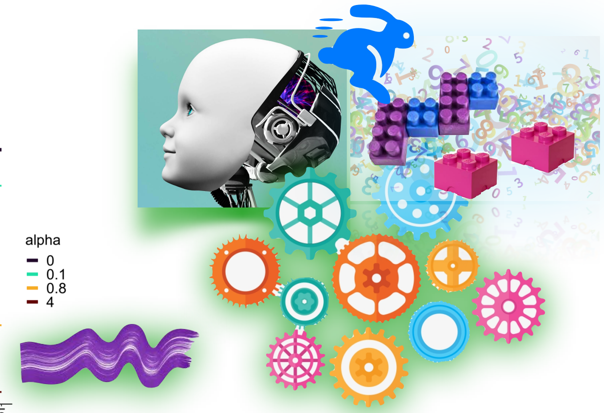
One way to think about these is as qualitative categories.

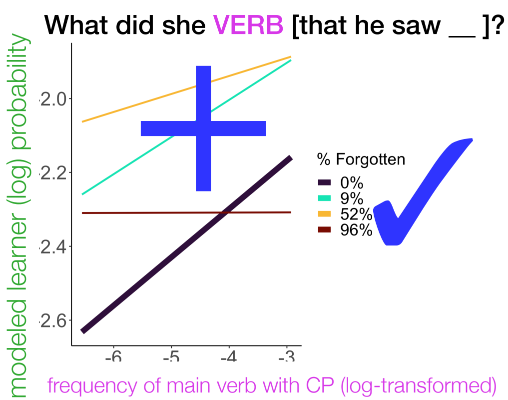
De Villiers et al. 2008

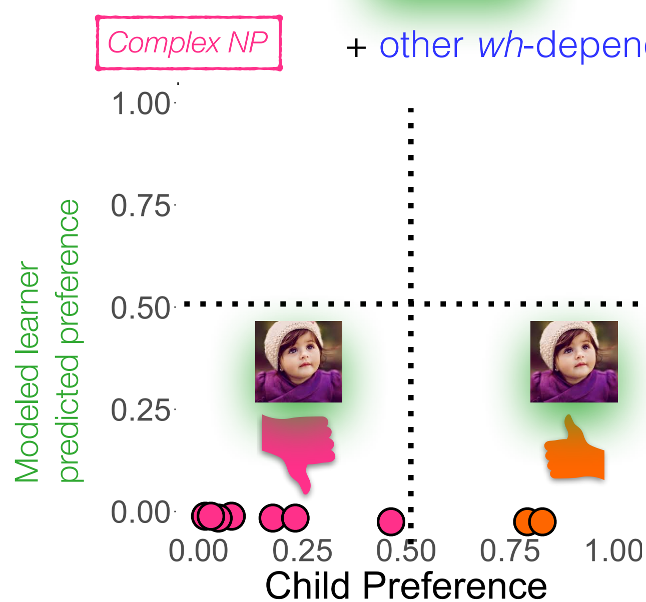
Dickson et al. 2022, 2024, in prep.









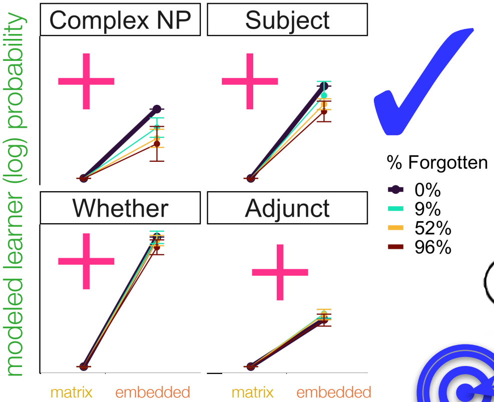


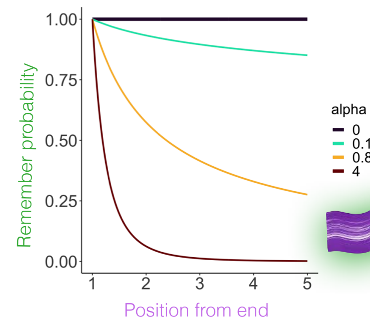
+ other *wh*-dependencies

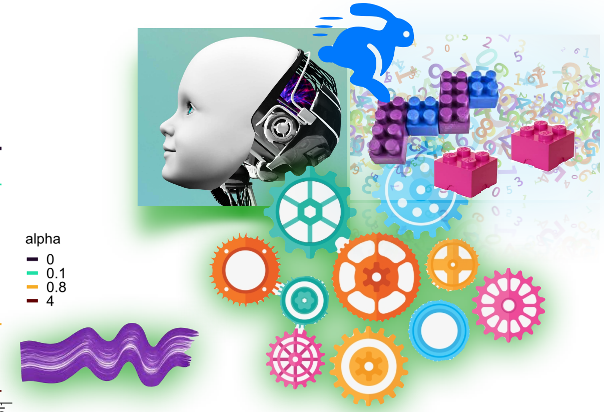
De Villiers et al. 2008

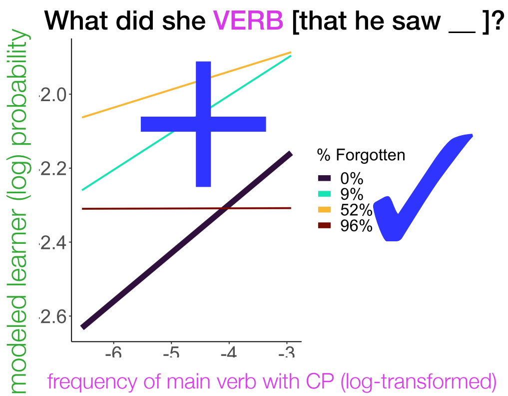
Then, we want to try to connect these observed qualitative categories to modeled learner predicted judgments.

Dickson et al. 2022, 2024, in prep.

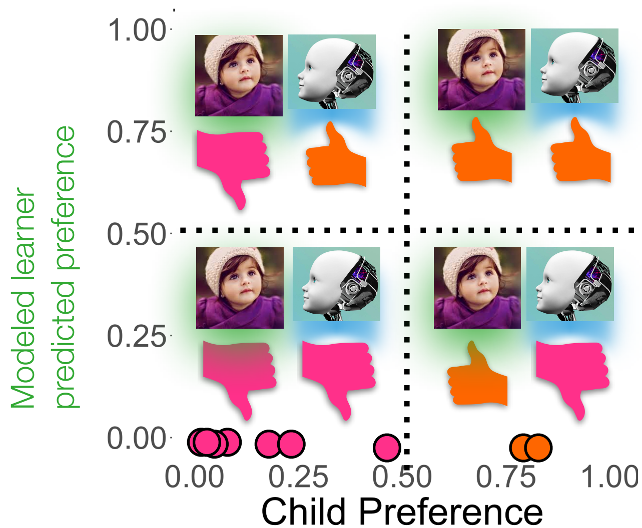




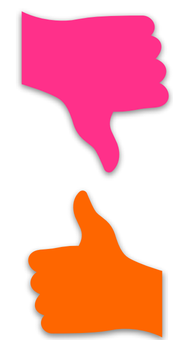




Complex NP + other *wh*-dependencies

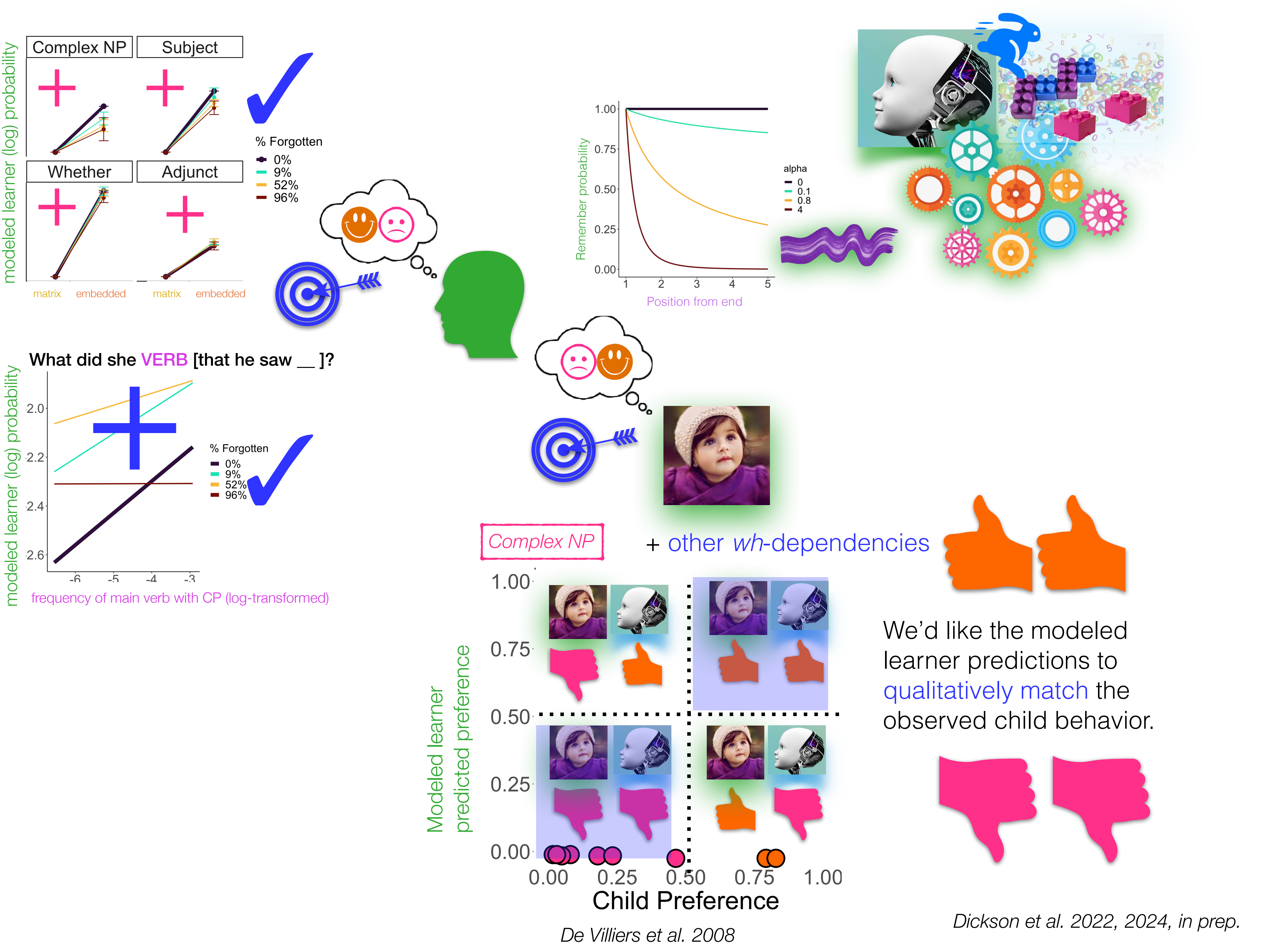


De Villiers et al. 2008

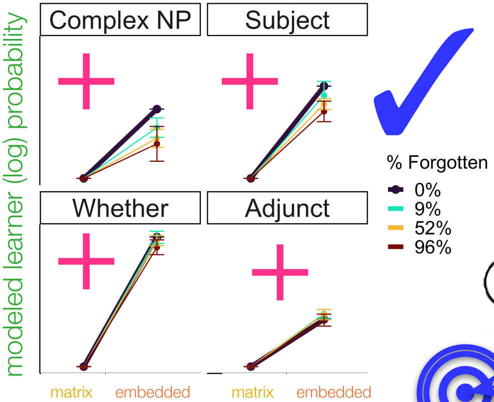


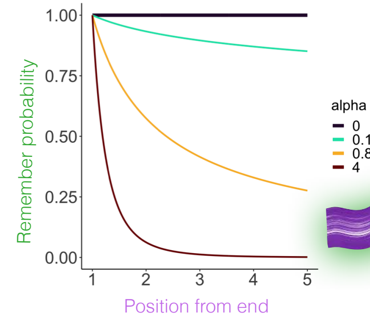
The different quadrants correspond to child vs. modeled learner qualitative predictions.

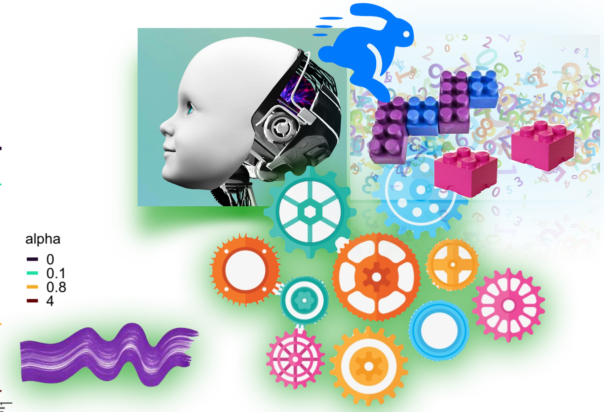
Dickson et al. 2022, 2024, in prep.

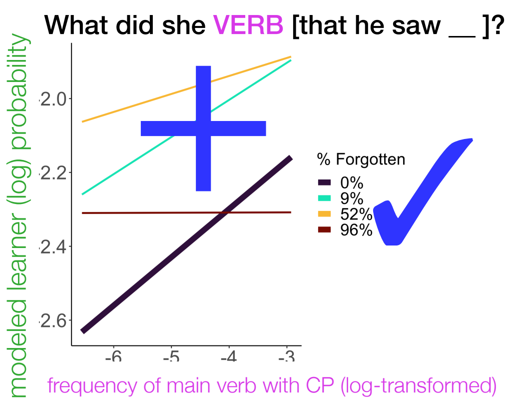


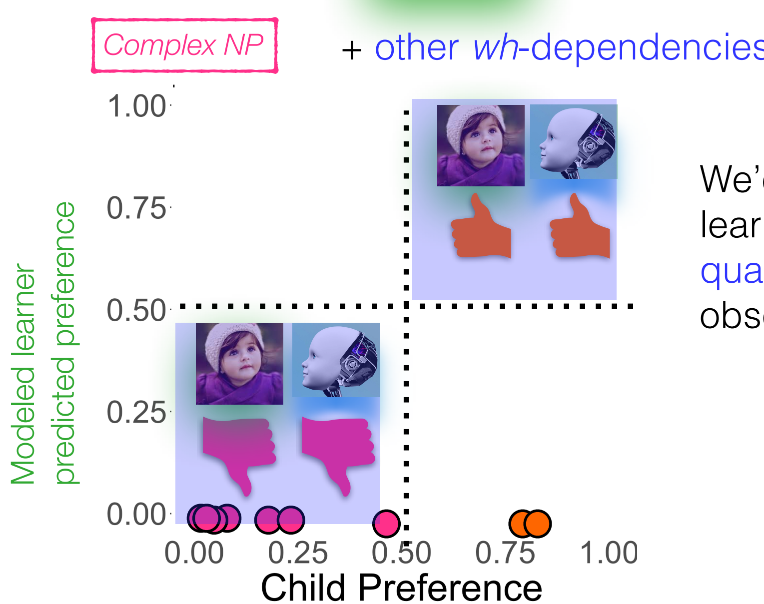
Complex NP
Subject
Whether
Adjunct
modeled learner (log) probability
matrix
embedded
matrix
embedded
% Forgotten
0%
9%
52%
96%
1.00
0.75
0.50
0.25
0.00
Remember probability
Position from end
1
2
3
4
5
alpha
0
0.1
0.8
4
What did she VERB [that he saw __]?
modeled learner (log) probability
-2.0
-2.2
-2.4
-2.6
-6
-5
-4
-3
frequency of main verb with CP (log-transformed)
% Forgotten
0%
9%
52%
96%
Complex NP
+ other *wh*-dependencies
Modeled learner predicted preference
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Child Preference
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.





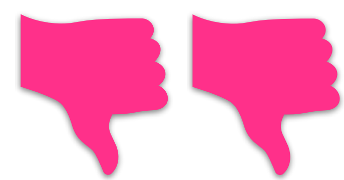




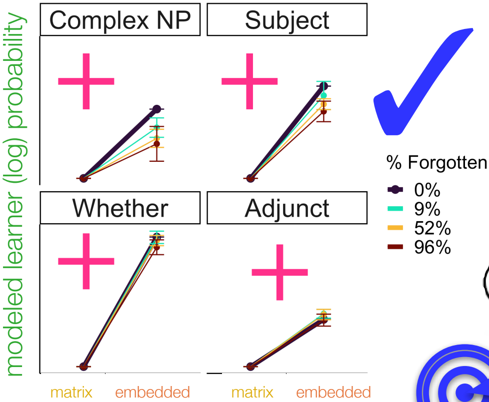


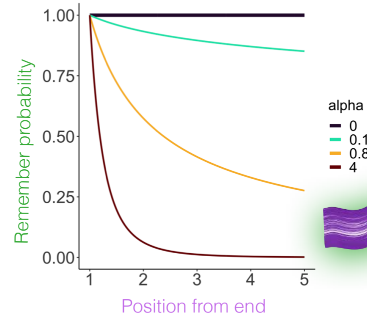
De Villiers et al. 2008

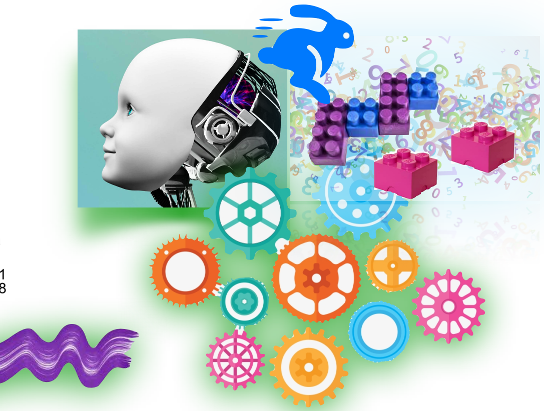
We'd like the modeled learner predictions to qualitatively match the observed child behavior.

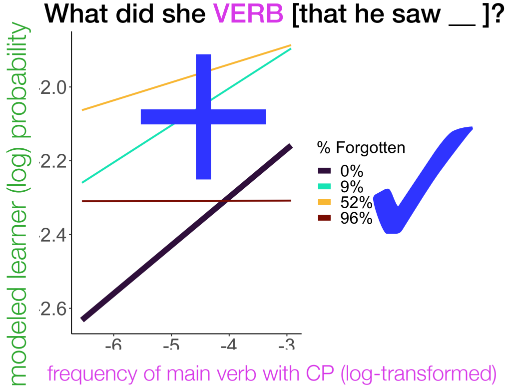


Dickson et al. 2022, 2024, in prep.

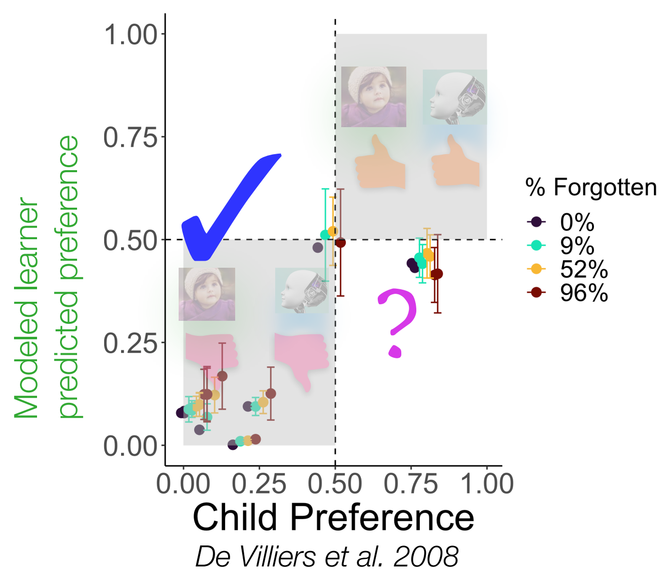








Complex NP + other *wh*-dependencies

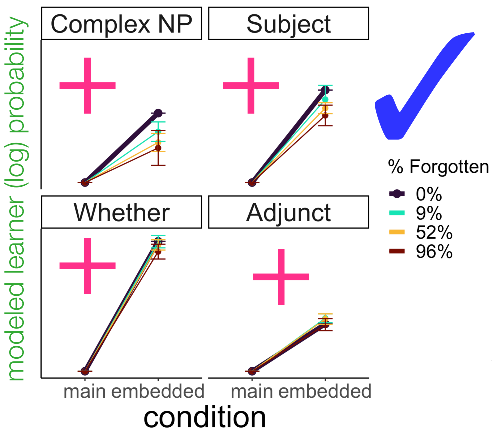


They mostly do, but there are some exceptions.

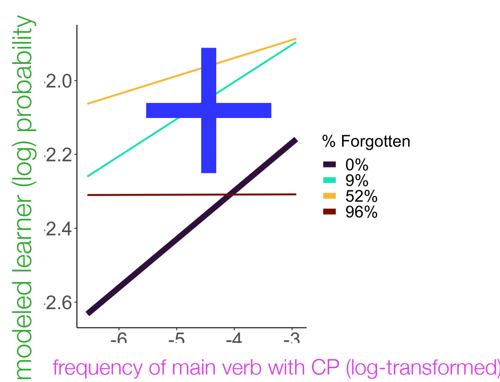
This is true, no matter how much forgetting there is.

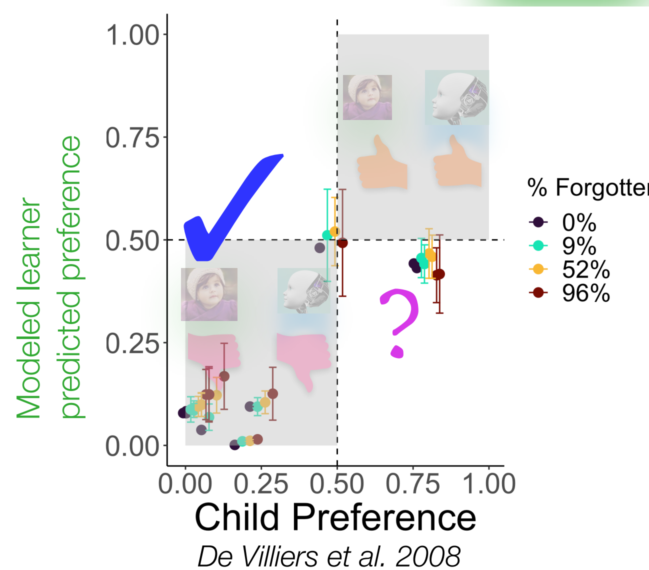
Dickson et al. 2022, 2024, in prep.

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



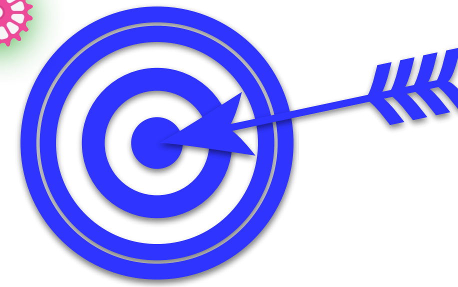
What did she VERB [that he saw __]?



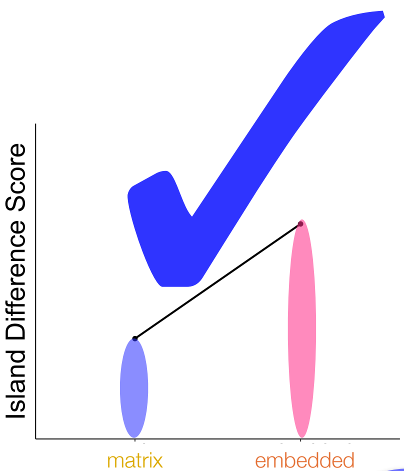


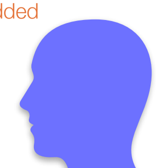
Dickson et al. 2022, 2024, in prep.

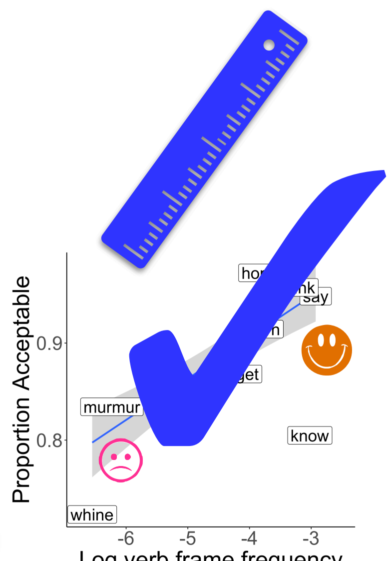
Dickson, Pearl, & Futrell 2022, 2024, in prep.

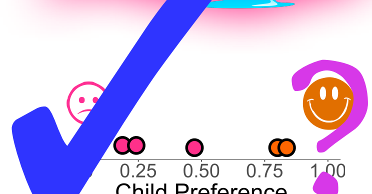


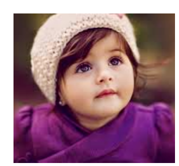
This means the modeled child, who encoded the “efficient building blocks” learning theory, was able to generate almost all of children’s target behavior.

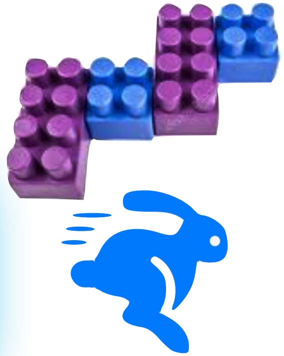




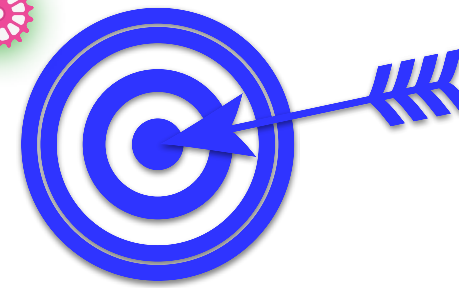




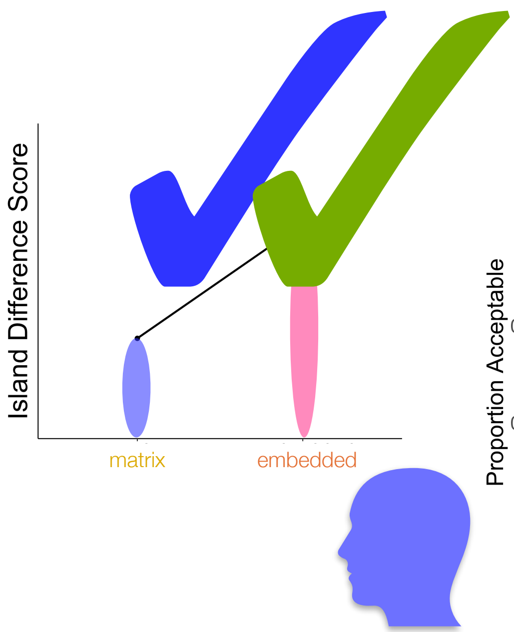


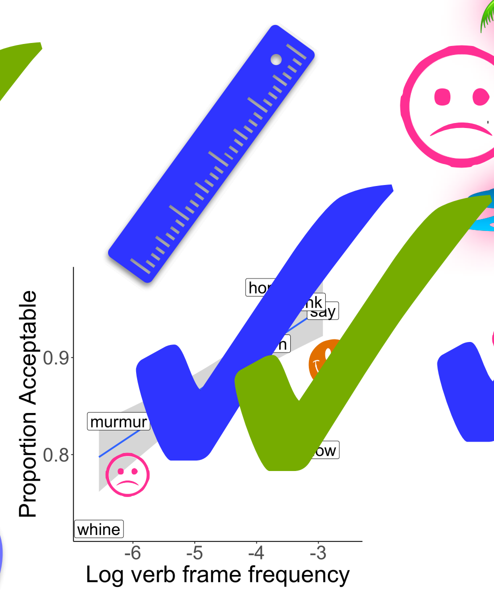


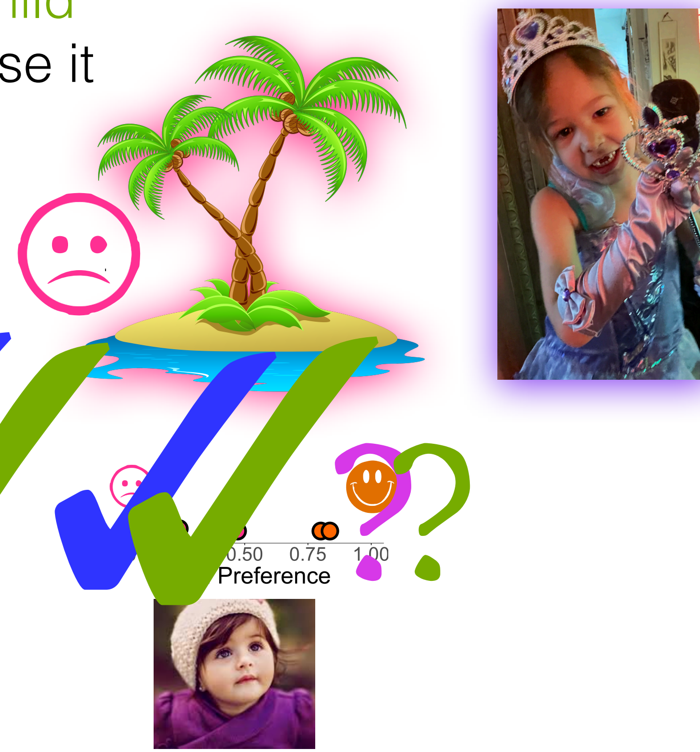
Dickson, Pearl, & Futrell 2022, 2024, in prep.



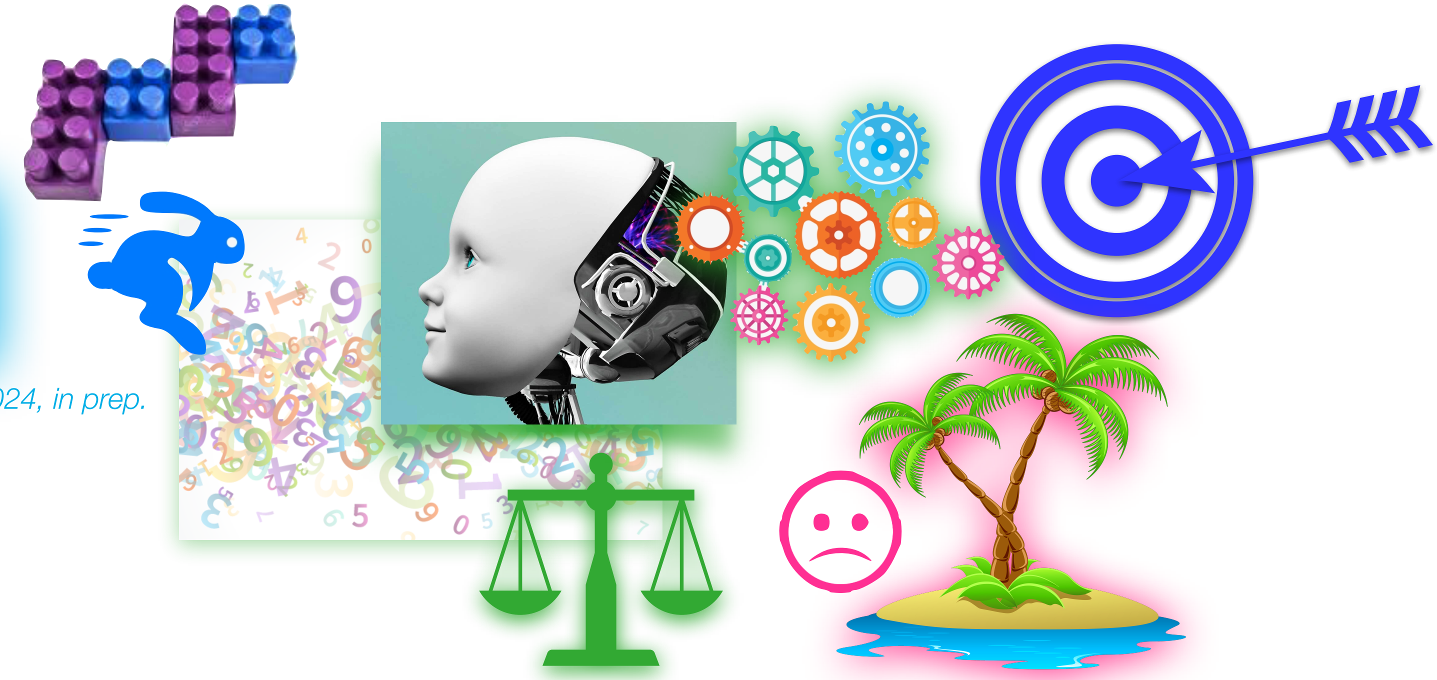
And this is also true even if the modeled child has child-like memory constraints that cause it to forget some of its intake.





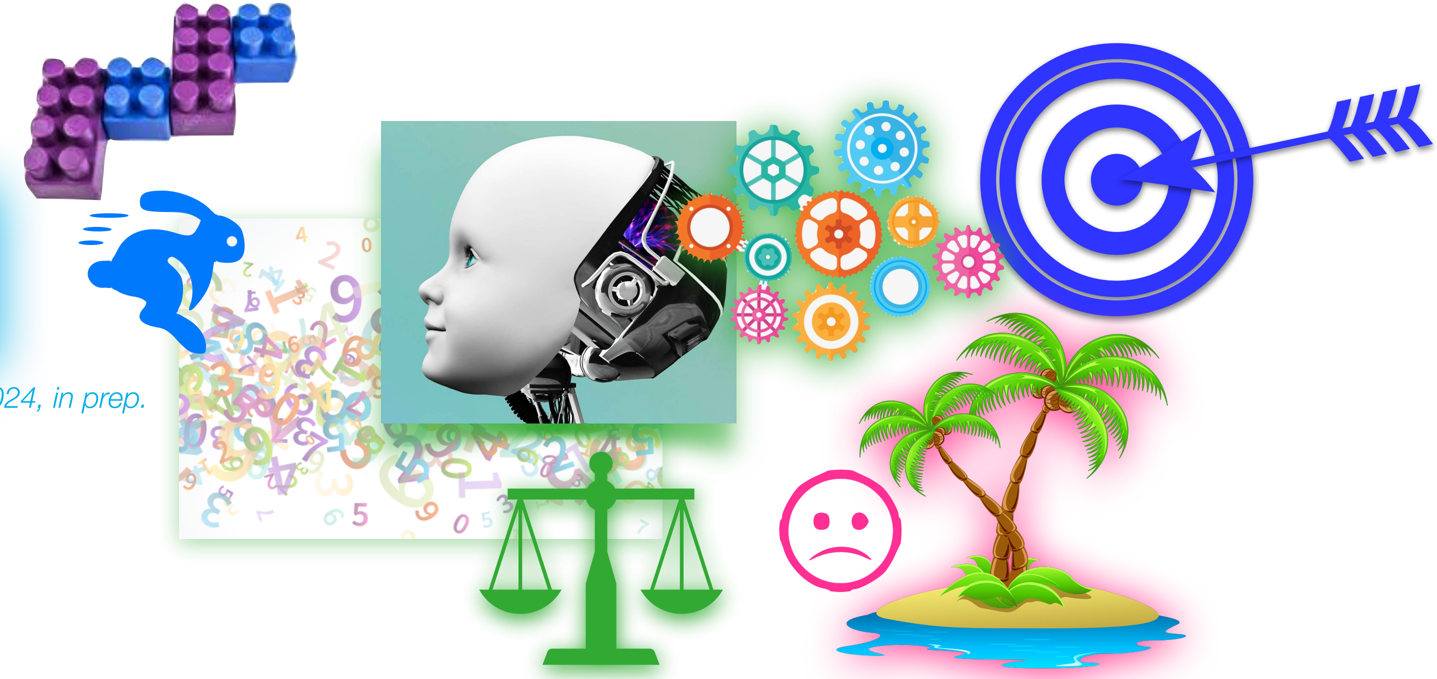


Dickson, Pearl, & Futrell 2022, 2024, in prep.

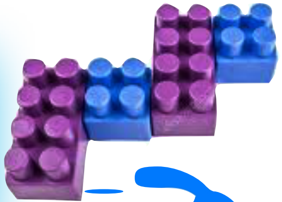


Takeaway: This learning theory (implemented by the modeled child using math) is pretty good at capturing children's target behavior.

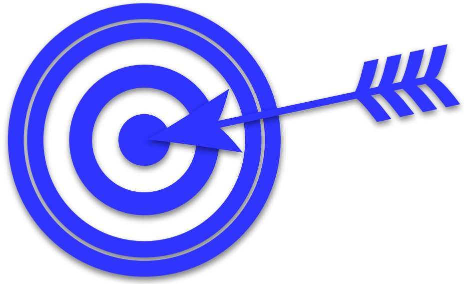
Dickson, Pearl, & Futrell 2022, 2024, in prep.



Takeaway: So, it may be a pretty good explanation for how children acquire (some) syntactic island knowledge from their input.

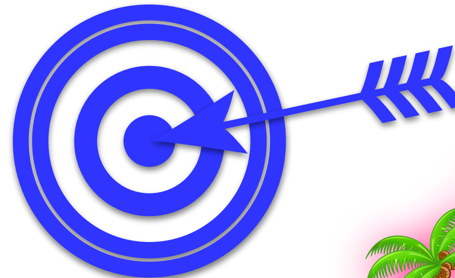


Dickson, Pearl, & Futrell 2022, 2024, in prep.



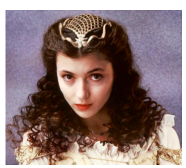
What about other syntactic island knowledge?
And other knowledge about *wh*-dependencies?

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



Some other *wh*-dependency knowledge

The kitty's antics make Lily laugh.

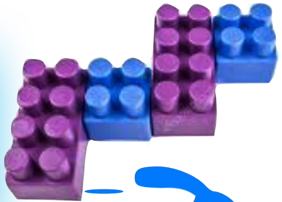


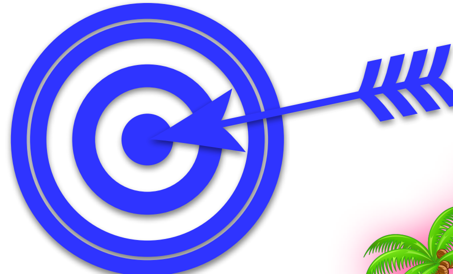
She wants to get him in order to pet him.

What did Lily get __$_{what}$ in order to pet __$_{what}$?

What did Lily get __$_{what}$ and then pet __$_{what}$?

What did Lily get __$_{what}$ before petting __$_{what}$?

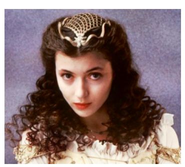




Some other *wh*-dependency knowledge

This puppy is also really adorable.

What …__what … __what?

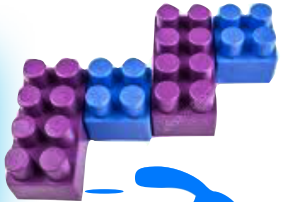


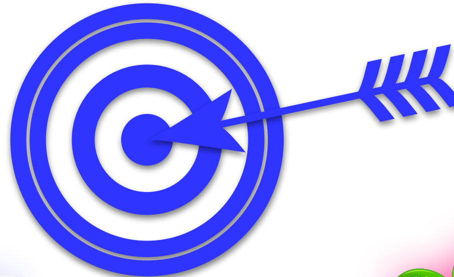
Lily wants to pet him too,
but she has to get the kitty first.

What did Lily get the kitty in order to pet __what?

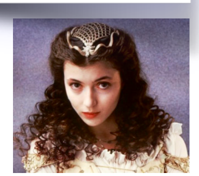
What did Lily get the kitty and then pet __what?

What did Lily get the kitty before petting __what?




How do children learn that these dependencies with multiple gaps are much more acceptable?




What did Lily get __what in order to pet __what?

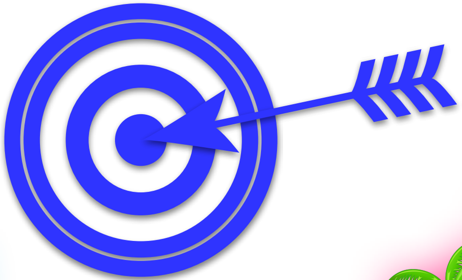
What did Lily get __what and then pet __what?

What did Lily get __what before petting __what?

What did Lily get the kitty in order to pet __what?

What did Lily get the kitty and then pet __what?

What did Lily get the kitty before petting __what?



Could this learning theory about the right building blocks explain how?

What did Lily get __what in order to pet __what?

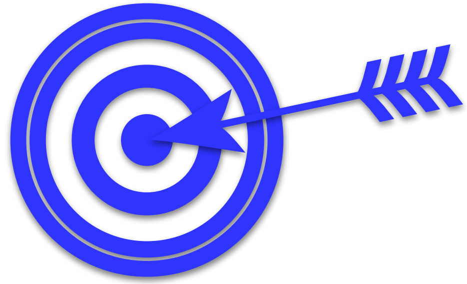
What did Lily get __what and then pet __what?

What did Lily get __what before petting __what?

What did Lily get the kitty in order to pet __what?

What did Lily get the kitty and then pet __what?

What did Lily get the kitty before petting __what?

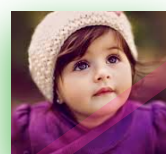


Some child preferences we know about *wh*-dependencies
(which we can mostly explain)

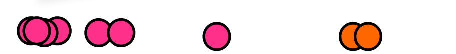
What did the boy [fix [$_{NP}$ the cat [that [was [lying [on [the table [with ___$_{what}$]]]]]]]]?

vs.

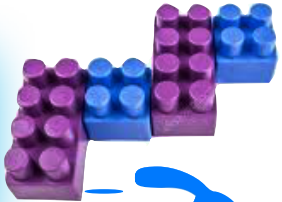
What did the boy [$_{VP}$ fix
[the cat that was lying on the table]
with ___$_{what}$]]?

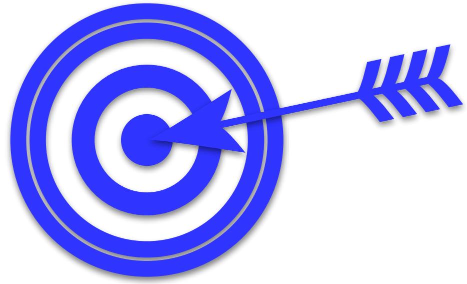


Complex NP + other *wh*-dependencies



De Villiers et al. 2008

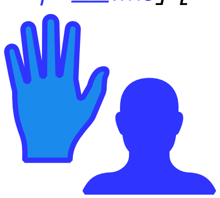



Some child preferences we know about *wh*-dependencies
(which we can mostly explain)

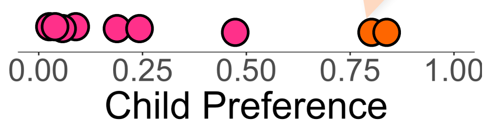
Who did [the policewoman [help [to [call __who]]]]?

vs.

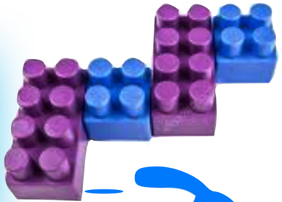
Who did [the policewoman [help __who] [to call]]?

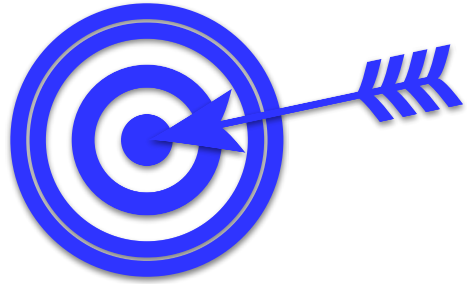


Complex NP + other *wh*-dependencies



De Villiers et al. 2008



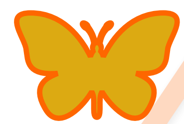
Some child preferences we know about *wh*-dependencies
(still to explain)

Where did [Lizzie [say [that [she was going to catch butterflies __where]]]]?

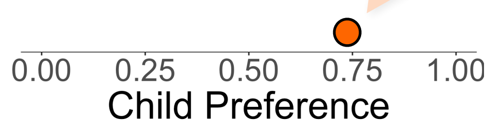
vs.

Where did [Lizzie say __where [that [she was going to catch butterflies]]]?

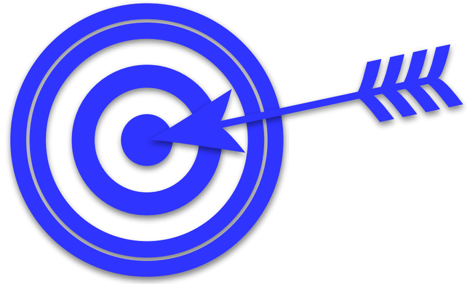




by age 5



Omaki al. 2014



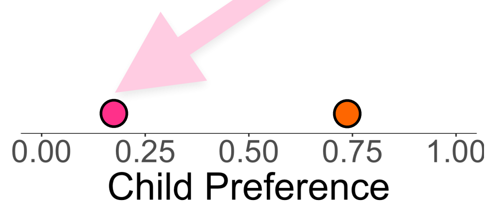
Some child preferences we know about *wh*-dependencies
(still to explain)

Where did [Lizzie [say to someone [that [she was going to catch butterflies ___*where*]]]]?

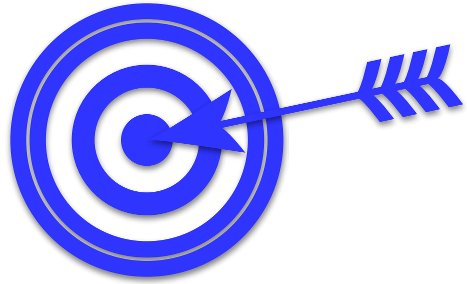
vs.

Where did [Lizzie [say to someone ___where that [she was going to catch butterflies]]]?

by age 5



Omaki al. 2014



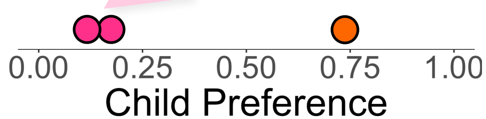
Some child preferences we know about *wh*-dependencies
(still to explain)

Where did [Lizzie [tell someone [that [she was going to catch butterflies ___*where*]]]]?

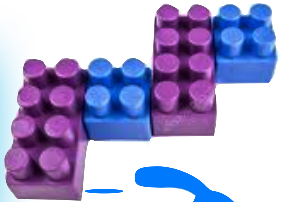
vs.

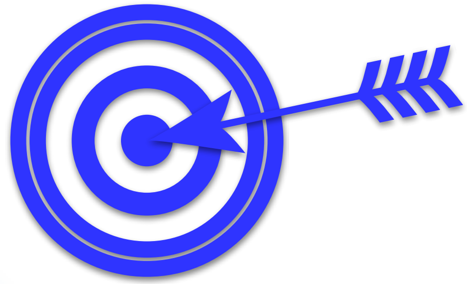
Where did [Lizzie [tell someone ___where that [she was going to catch butterflies]]]?

by age 5



Omaki al. 2014





Could this learning theory about the right building blocks explain how these preferences develop?

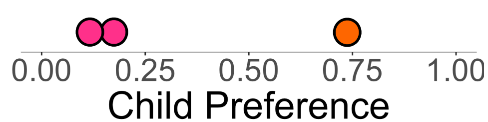
catch … __where

say __where

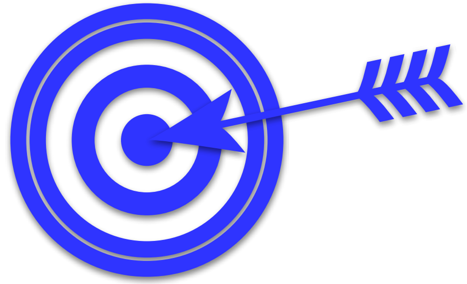
say to someone __where

tell someone __where

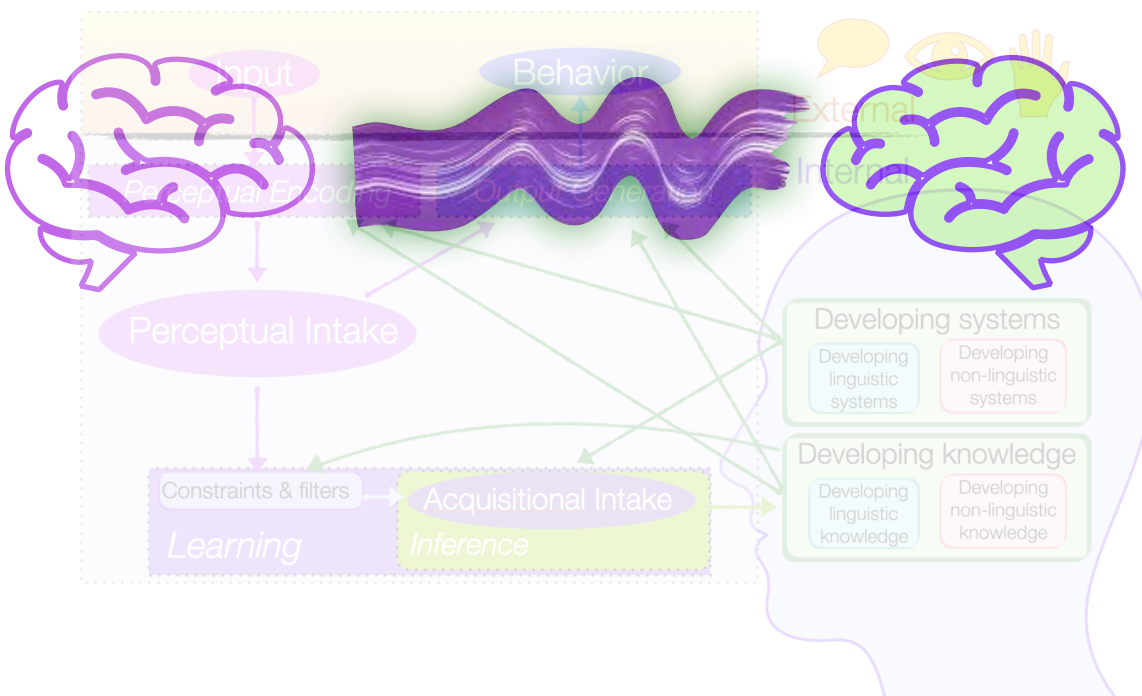
by age 5

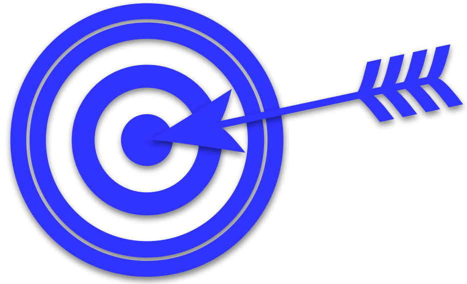


Omaki al. 2014

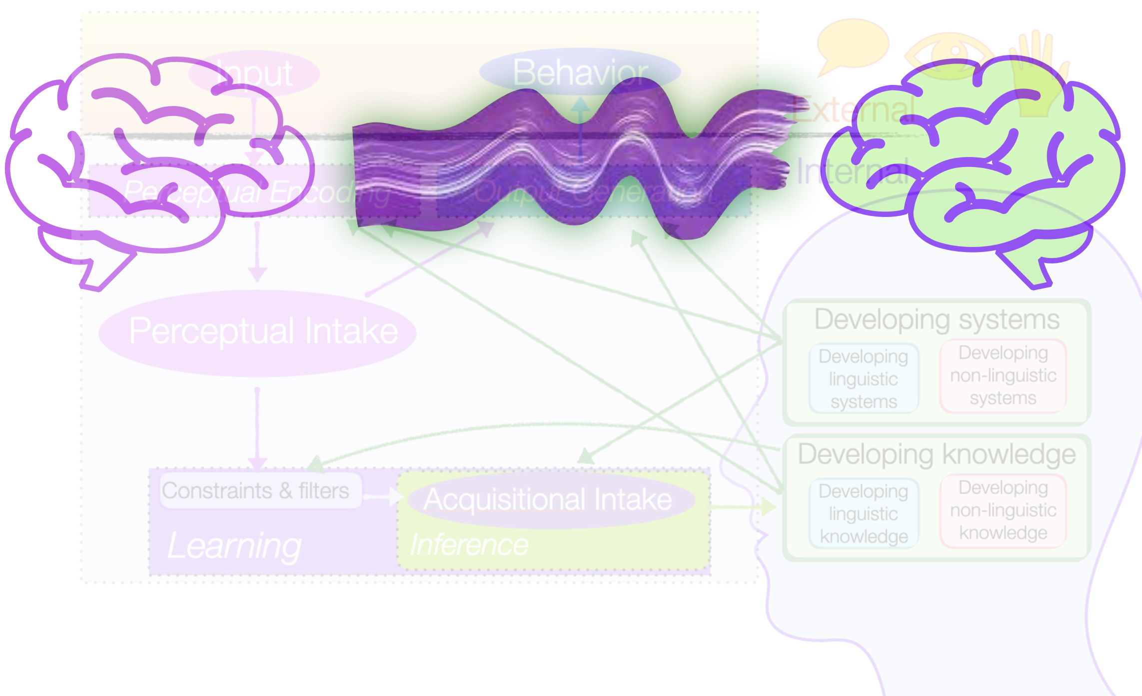


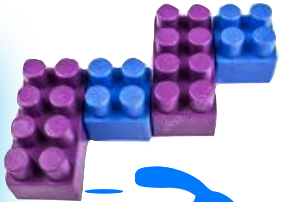
This learning theory was able to explain the development of a lot of target knowledge even when there were some (hopefully child-like) memory limitations.

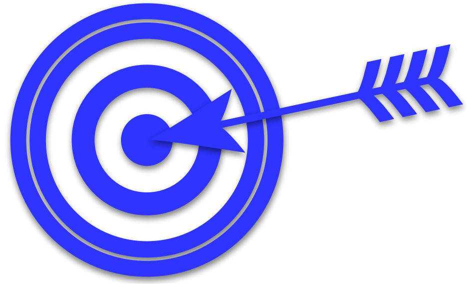




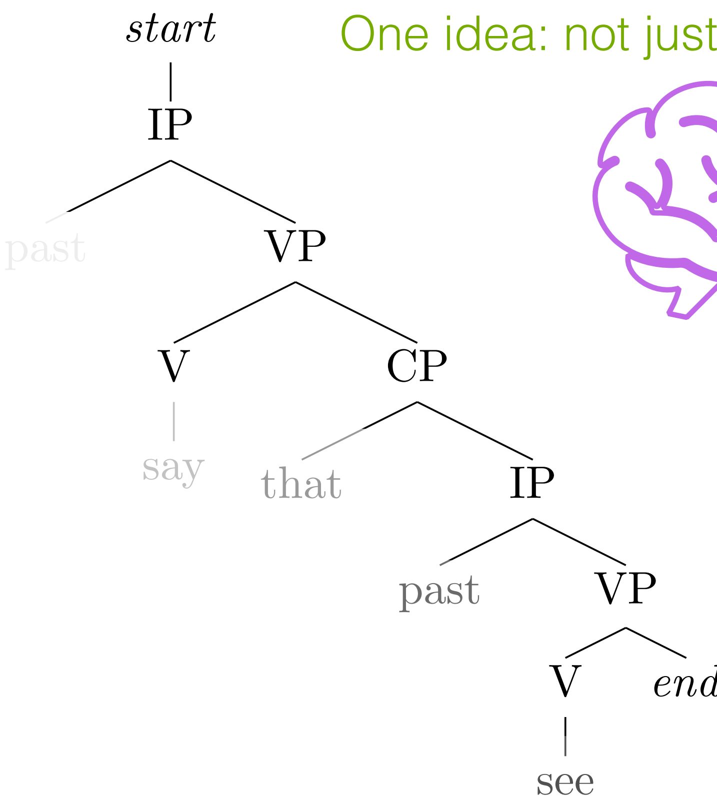
What happens if other abilities are also developing?



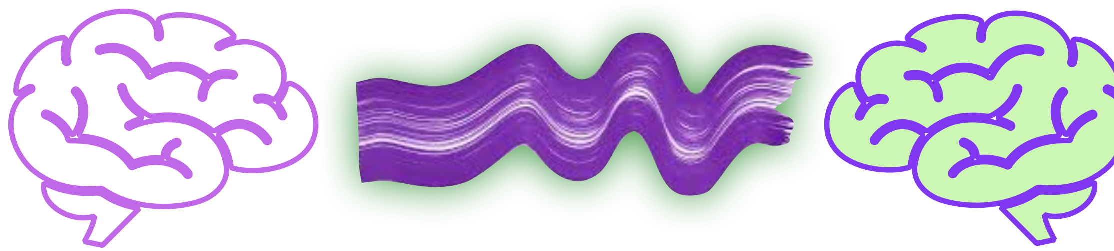



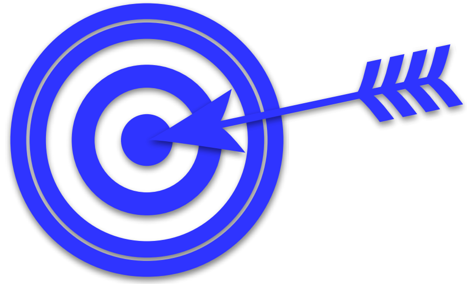


What if children's developing memory impacts their intake in other ways?

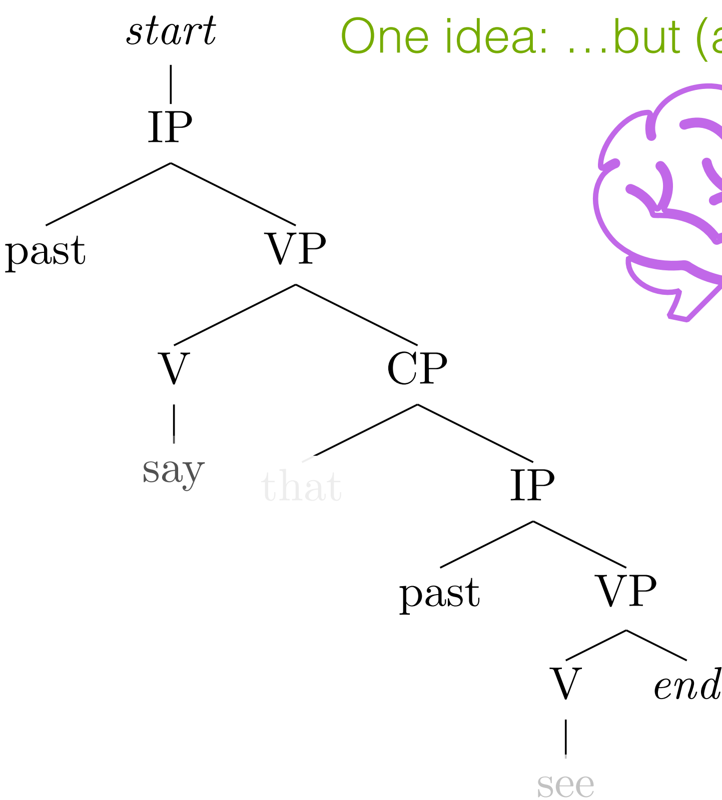


One idea: not just forgetting lexical items based on position…

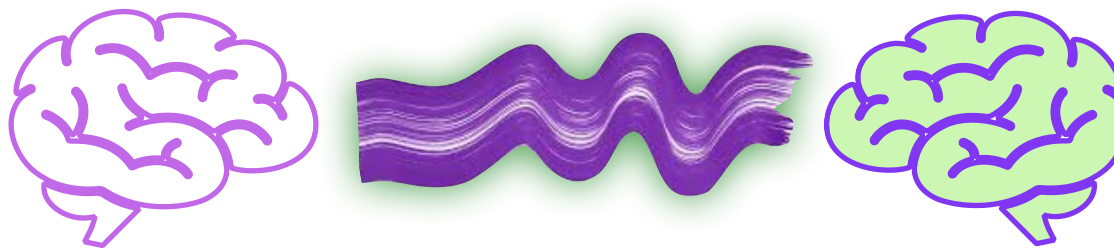


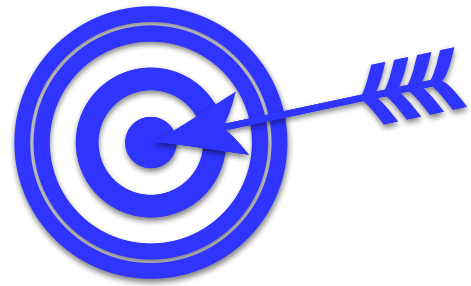


What if children's developing memory impacts their intake in other ways?

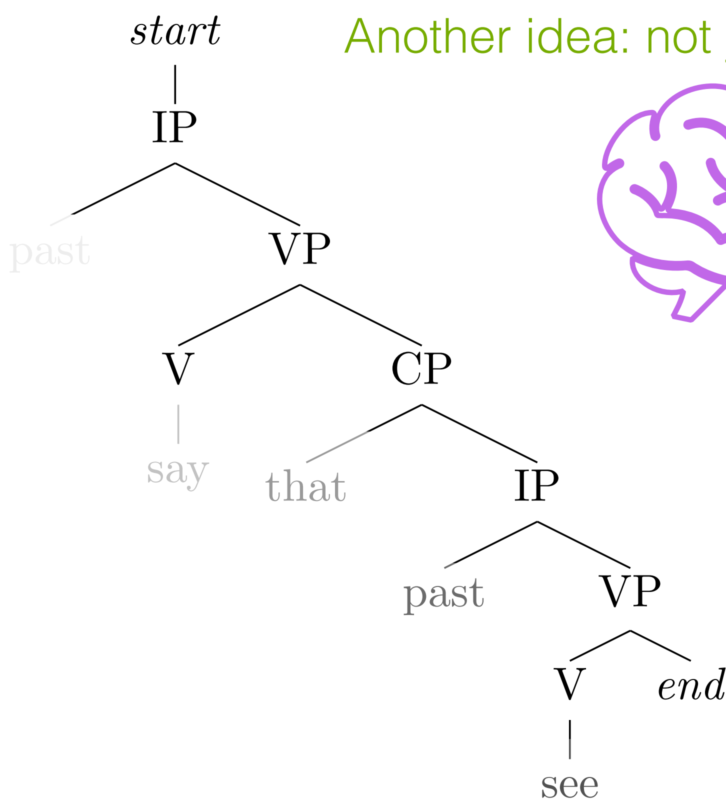


One idea: …but (also) based on frequency

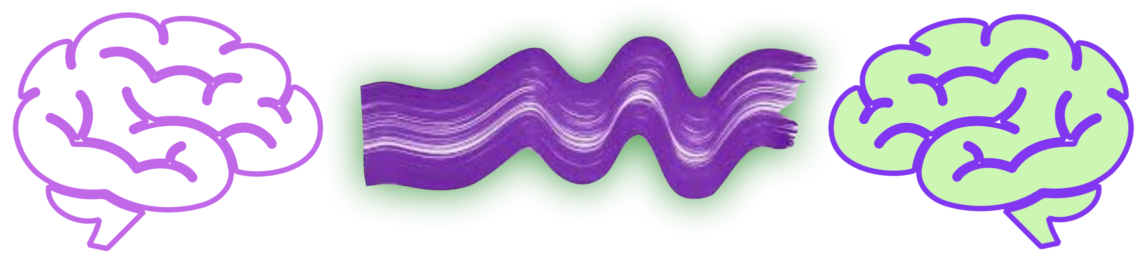


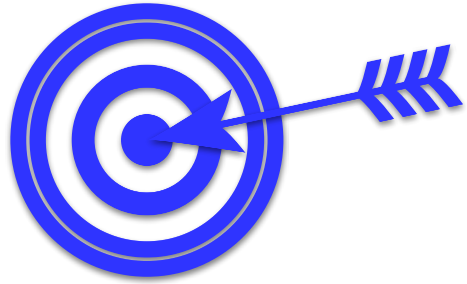


What if children's developing memory impacts their intake in other ways?

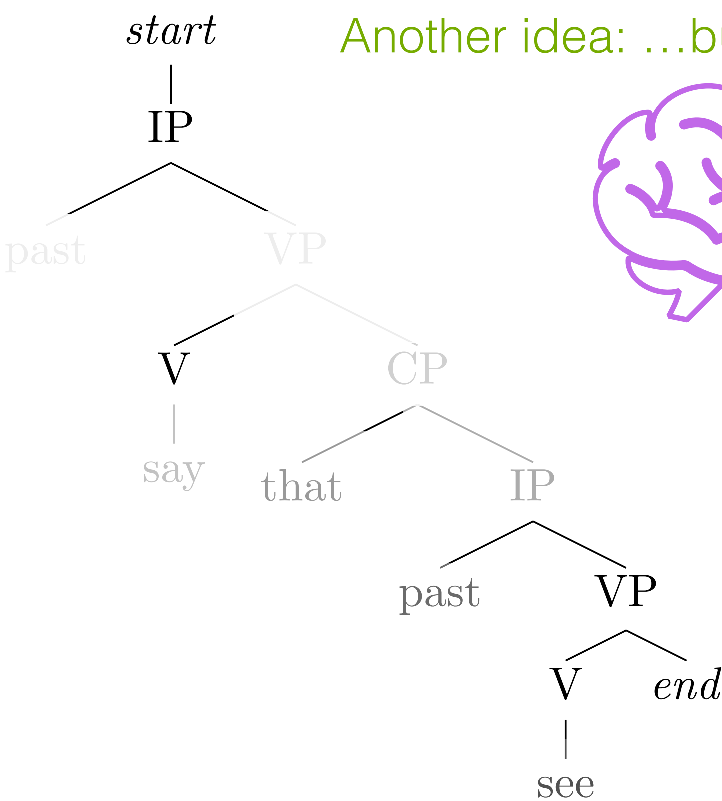


Another idea: not just forgetting lexical items…

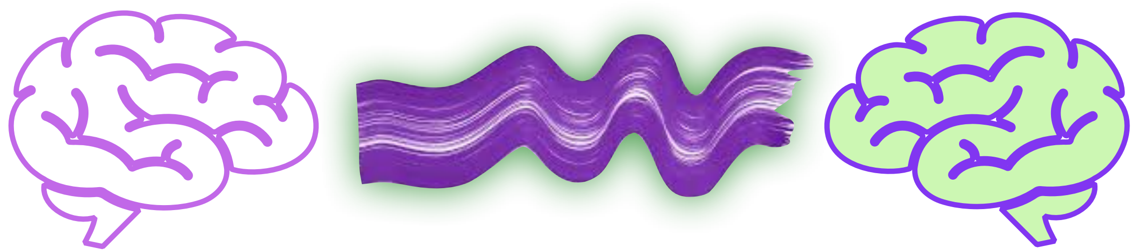


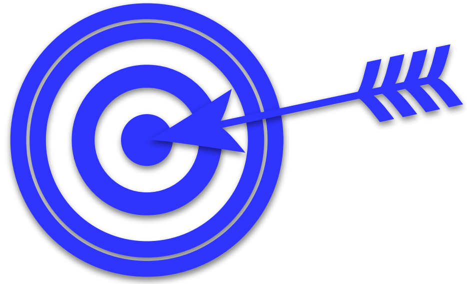


What if children's developing memory impacts their intake in other ways?

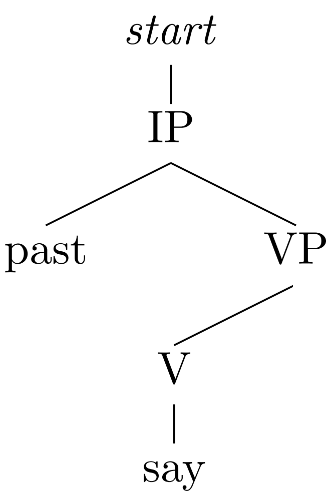


Another idea: …but also structure

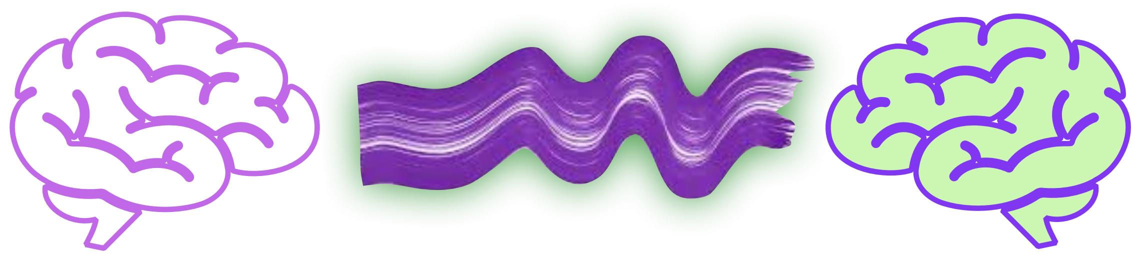




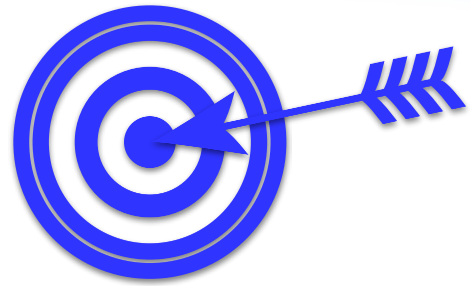
What if children's developing memory impacts their intake in other ways?

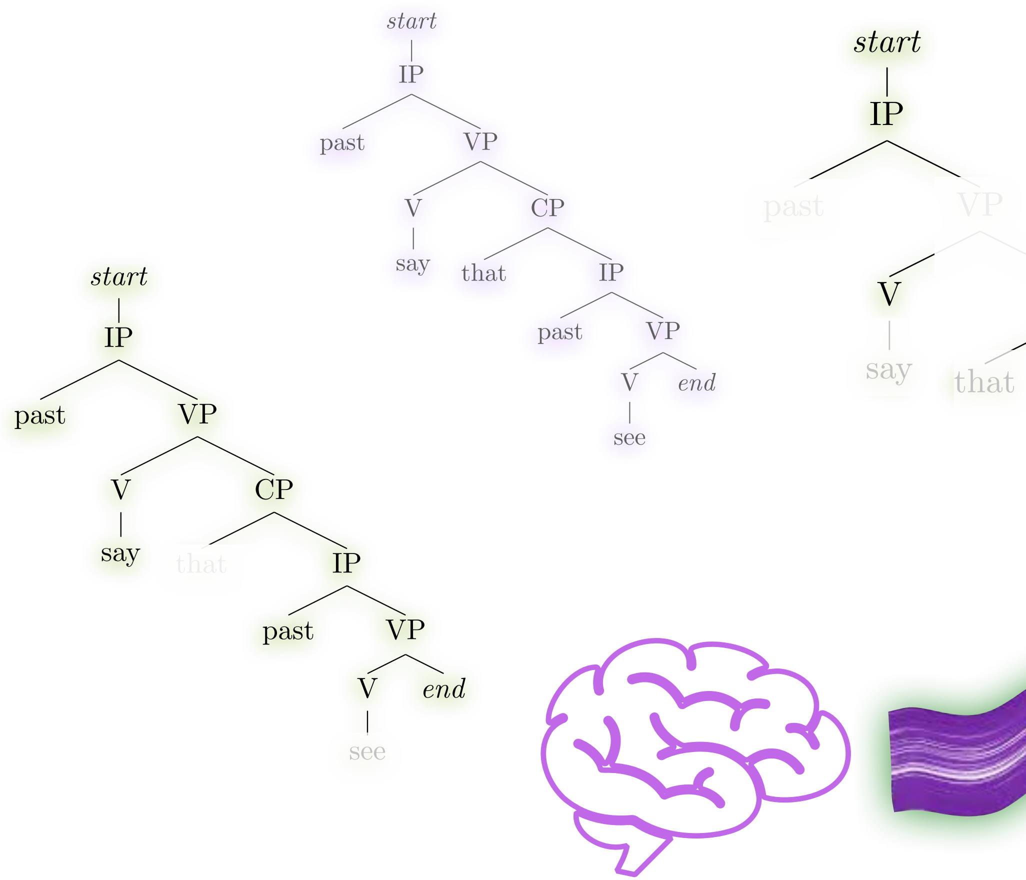
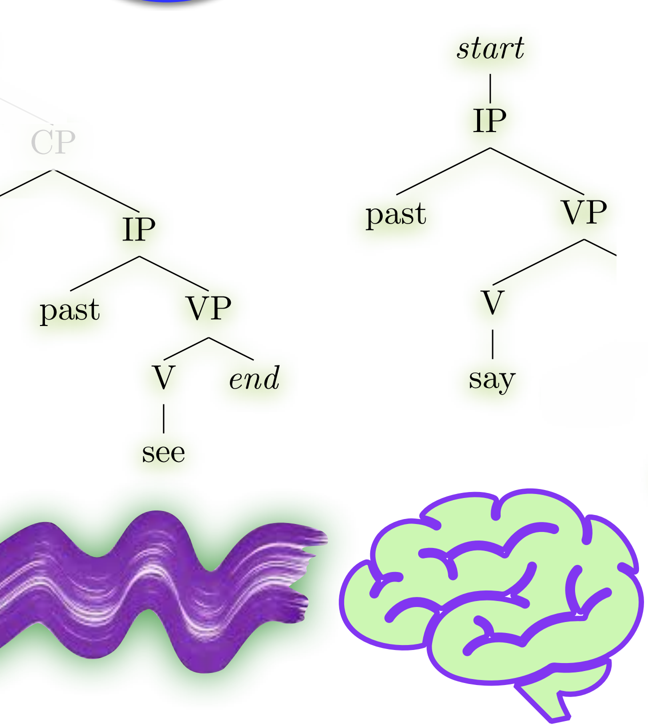


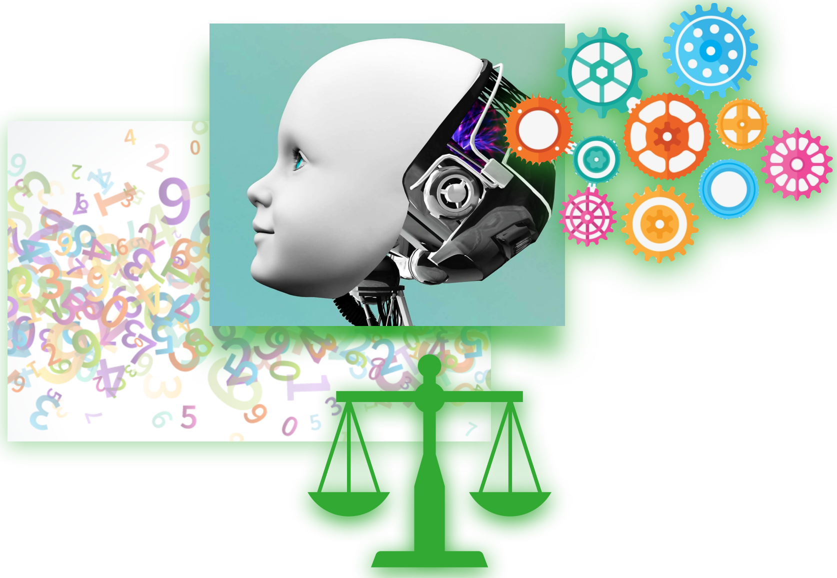
Another idea: mis-parsing the input and being unable to revise in time (Trueswell et al. 1999, Snedeker 2013, Omaki & Lidz 2015)



Could this learning theory about the right building blocks still explain the target behavior?

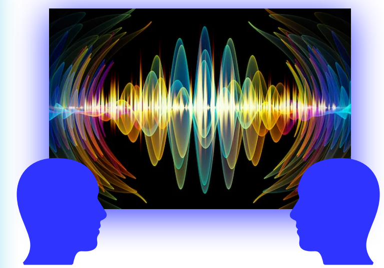


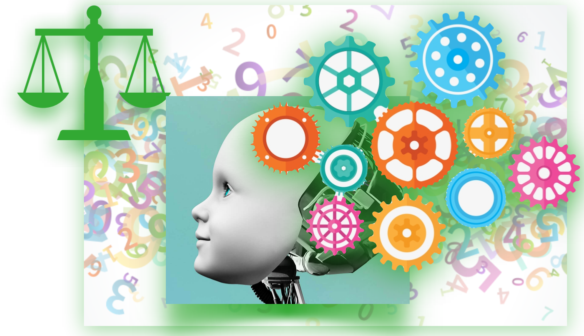





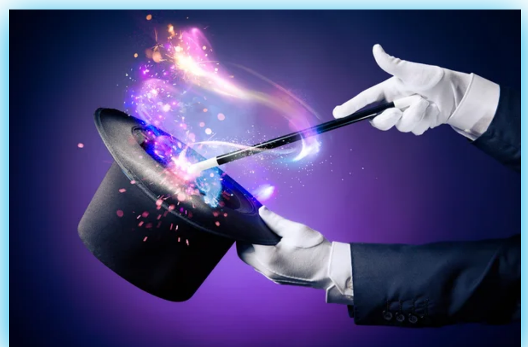
Some other discoveries from my research group, using computational cognitive modeling

Children transform the input signal they encounter in order to learn.

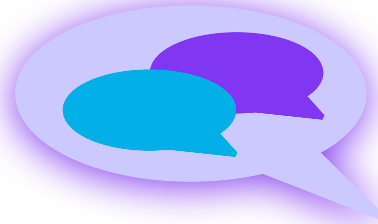


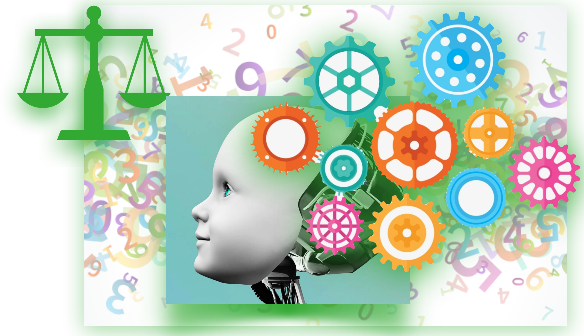


Sometimes acquisition success may result only when the data intake for acquisition is a selective subset of the available input.

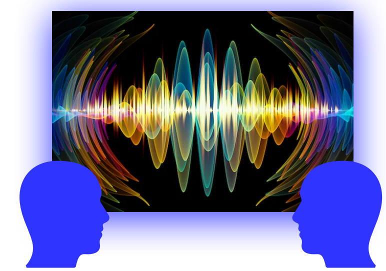


(*basic word order*: Pearl 2005a,b, 2007, Pearl and Weinberg 2007; *metrical stress*: Pearl 2007, 2008, 2009, 2011, Pearl, Ho and Detrano 2014, 2016, Pearl 2017; *English anaphoric one*: Pearl 2007, Pearl and Lidz 2009; *syntactic islands*: Pearl and Sprouse 2013a,b, Pearl 2014, Pearl and Sprouse 2015; *English passive*: Nguyen and Pearl 2019, 2021)





However, sometimes acquisition success may also occur by taking a broader perspective on what counts as relevant data.

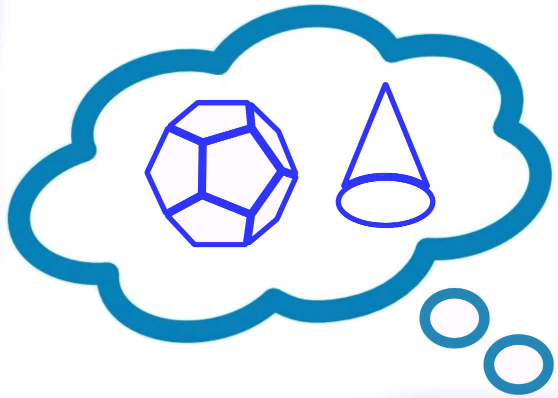


(Pearl 2023a; *English anaphoric one*: Pearl and Mis 2011, 2016; *syntactic islands*: Pearl and Sprouse 2013a,b, Pearl 2014, Pearl and Sprouse 2015, Pearl 2017, Bates and Pearl 2019, Pearl and Bates 2022; *English passive*: Nguyen and Pearl 2018, 2019, 2021).

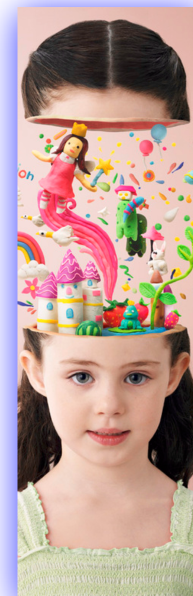
Children's knowledge during language development may be different than we thought.

To learn some types of linguistic knowledge, children may need to have language-specific knowledge built into their minds — however the exact form of this knowledge may be different than what we previously thought.

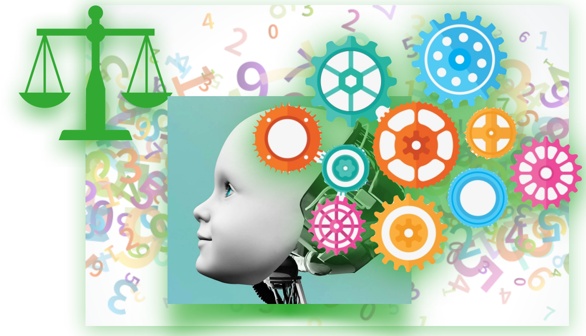
(*English anaphoric one*: Pearl and Lidz 2009, Pearl and Mis 2011, 2016; *syntactic islands*: Pearl and Sprouse 2013a,b, Pearl 2014, Pearl and Sprouse 2015, Dickson, Pearl and Futrell 2022, Dickson, Futrell, & Pearl 2024; *linking theories*: Pearl and Sprouse 2019, 2021)



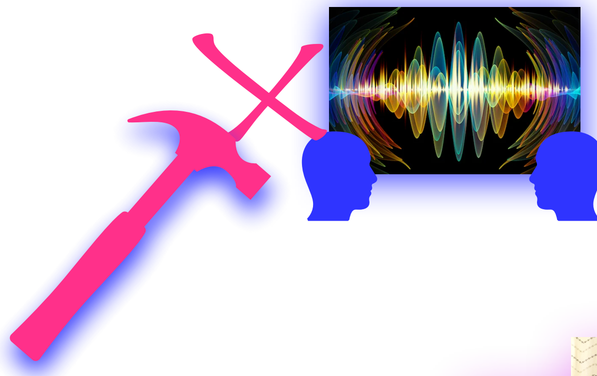
Children's linguistic knowledge may sometimes be far more well-developed than we realize earlier than we realize, with strong similarities between child and adult representations



(*closed-class syntactic categories*: Bates, Pearl and Braunwald 2018; *quantifier scope ambiguity resolution*: Savinelli, Scontras and Pearl 2017, 2018, Scontras and Pearl 2021; *adjective ordering preferences*: Bar-Sever, Lee, Scontras and Pearl 2018; *pronoun interpretation*: Forsythe and Pearl 2019, Pearl and Forsythe under review)

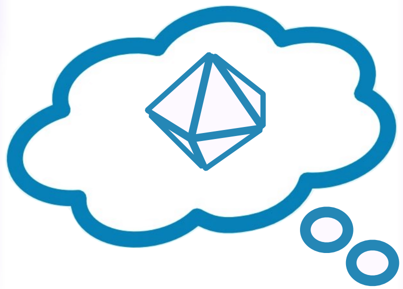


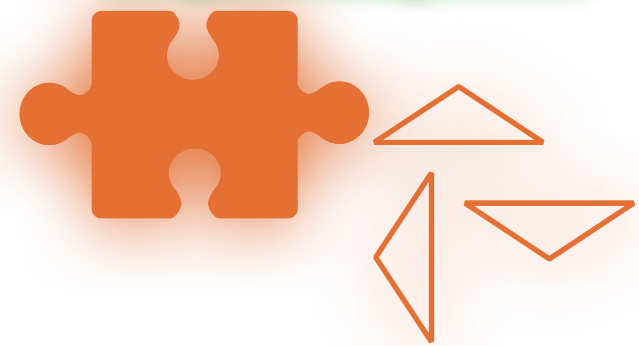
"Different" doesn't automatically mean "worse" when it comes to language development.



While there is certainly input variation across socio-economic status, the same language learning outcome could still occur, despite the variation. So, interventions targeted at "fixing" the input aren't likely to be effective.

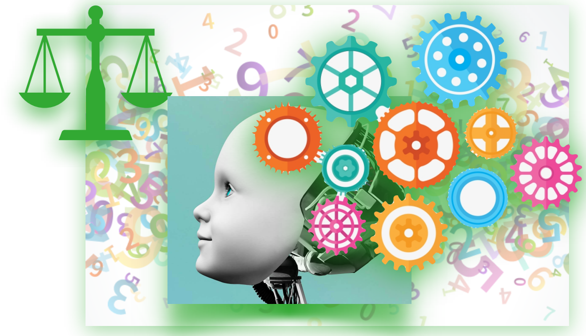
syntactic islands: Bates and Pearl 2019, Pearl and Bates 2022





Even if the representations of very young children may not match adult representations, they can still be “good enough” for helping other acquisition processes get started.

early speech segmentation: Phillips and Pearl 2012, 2014a,b, 2015a,b,c, Pearl and Phillips 2018; *early syntactic categorization*: Bar-Sever and Pearl 2016



Learners with processing constraints (like children) may sometimes learn better than learners with fewer limitations ("less is more").

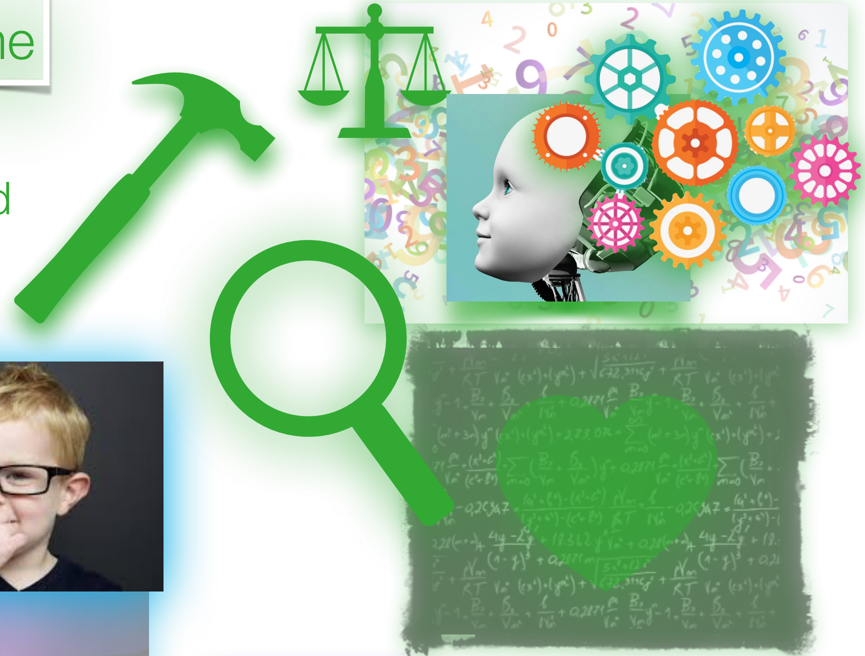
speech segmentation: Pearl, Goldwater and Steyvers 2010, 2011, Phillips and Pearl 2012, 2015c

Take home

Computational cognitive modeling is one way to use math to investigate the magical process of child language acquisition

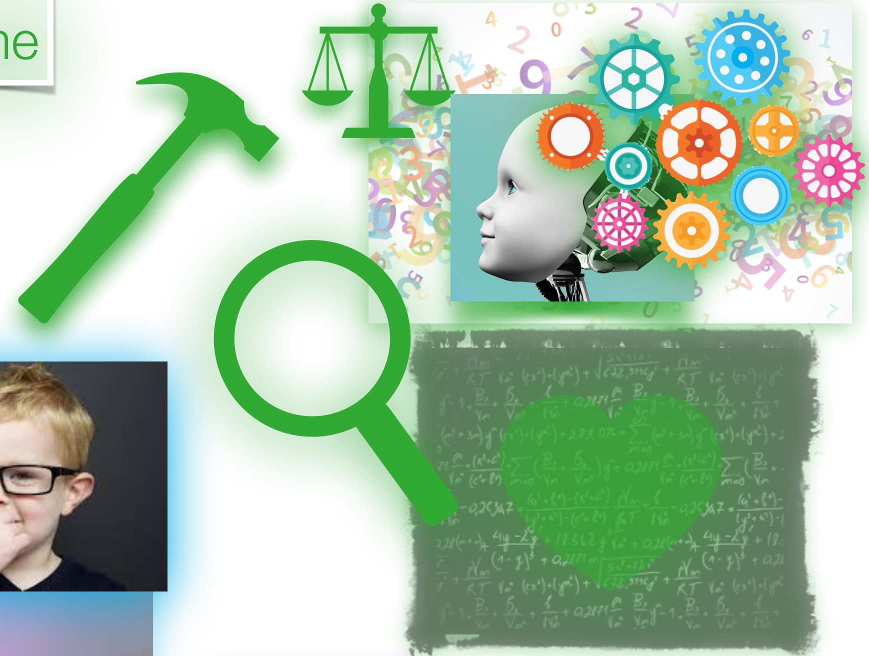
Take home

It allows us to implement learning theories concretely, evaluate them, and better understand how they (and potentially children) work.



Take home

So let's keep using this wonderful tool to investigate child language acquisition!



Thank you!

Niels Dickson | Richard Futrell | Jon Sprouse | 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 Pomona Acquisition Workshop 2024

CSU Fullerton Linguistics Students Association 2024

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