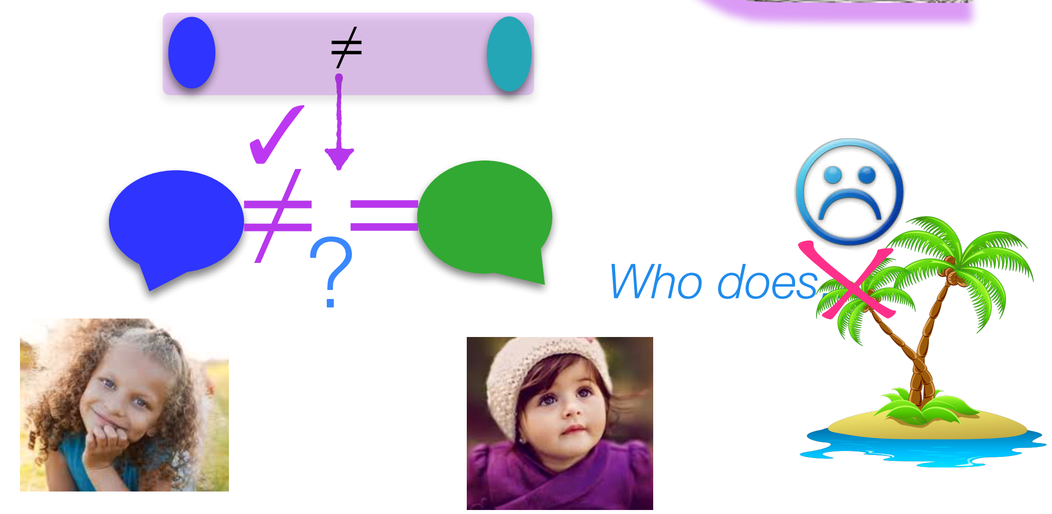
A new way to identify if variation in children's input is developmentally meaningful: A look at syntactic knowledge across socio-economic status

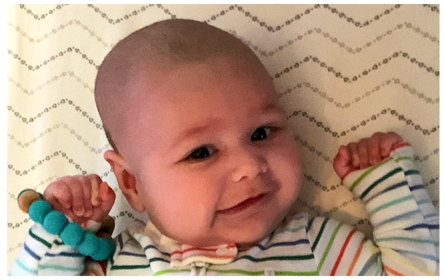
Lisa Pearl
University of California, Irvine

Computation of Language Laboratory
UC Irvine

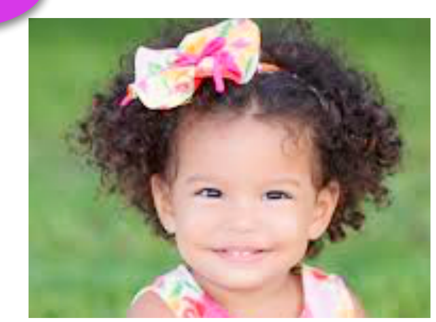
Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lpearl@uci.edu

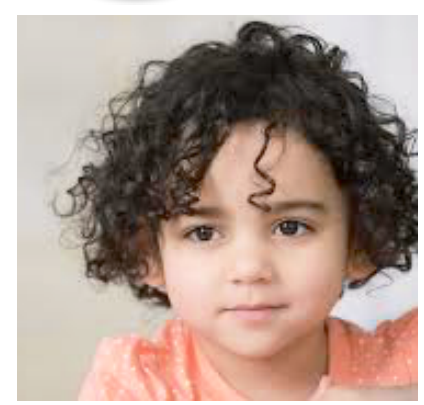


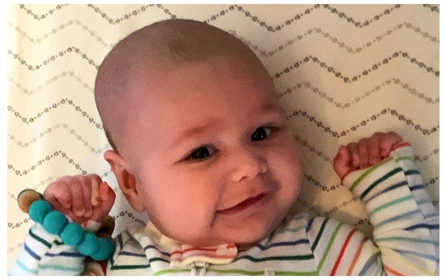
November 4, 2021
EyeLands Lab
Norwegian University of Science and Technology



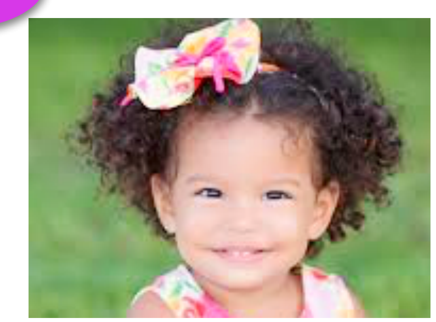
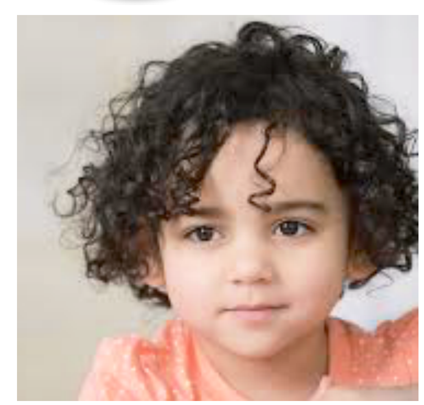
There’s lots of variation
in children’s input



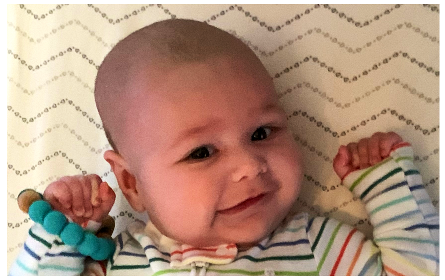



Developmentally-meaningful variation
impacts language development
in a measurable way

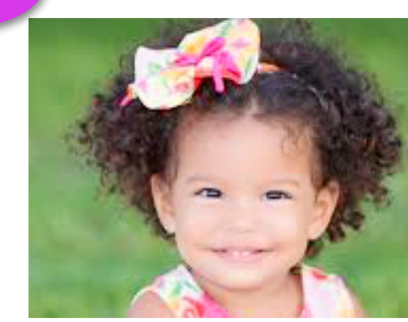




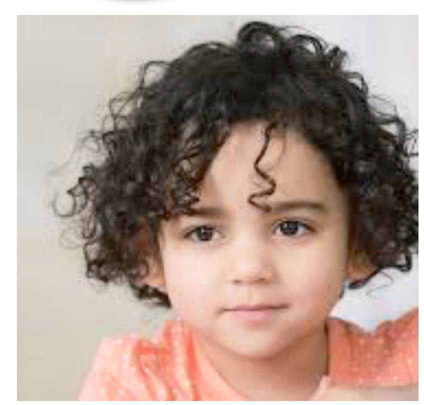




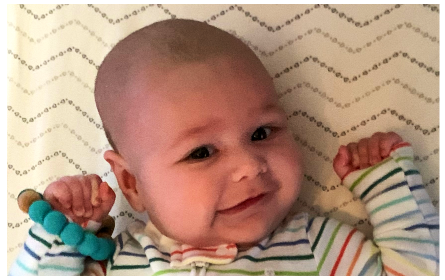

Developmentally-meaningful input deficits can lead to language delays









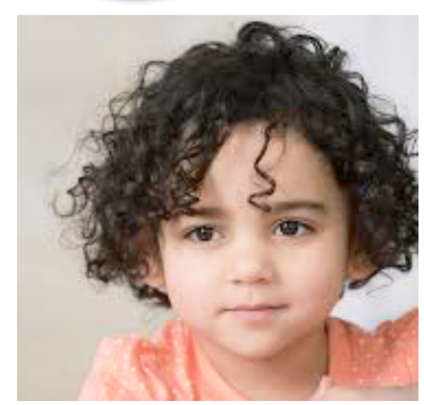
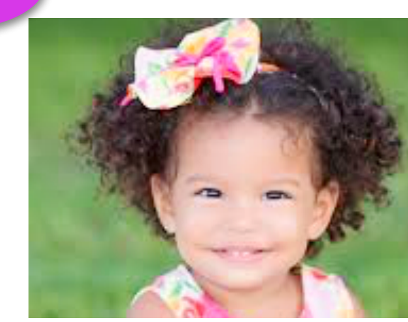


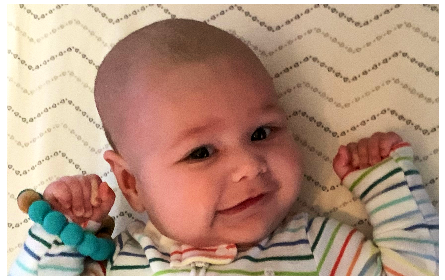


If there’s an input-based language delay and we know what the crucial input deficit is, we can intervene and fix that deficit.

Impactful interventions ✓

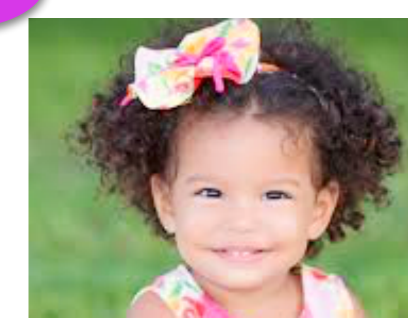


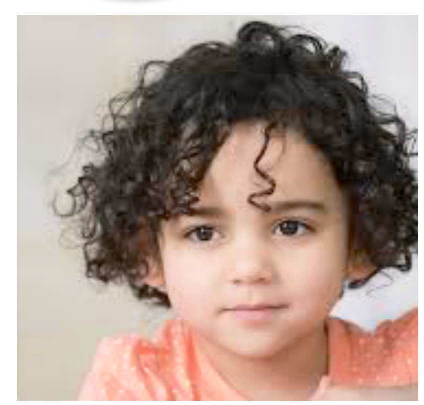



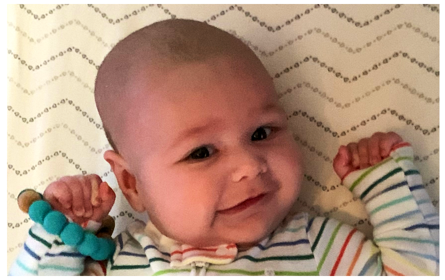




Important: If a language outcome difference isn't input-based, then "fixing" the input won't help.


Impactful interventions

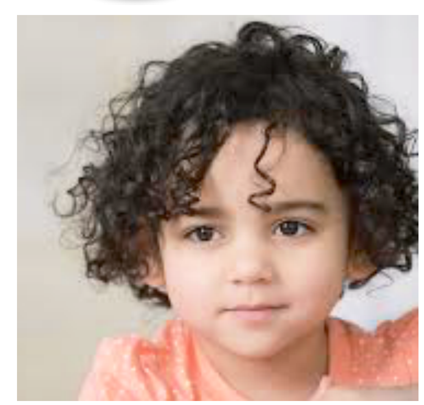



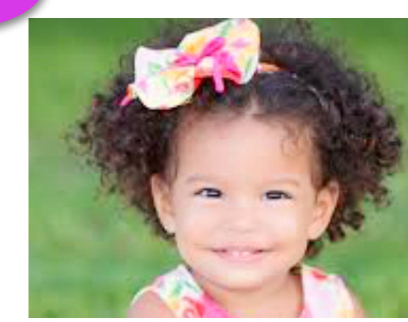




That's why it's important to know if the input is (at least partially) causing the language developmental issue… or not.

Impactful interventions





Input-based language delays appear across socio-economic status (SES). Lower-SES children are often behind their higher-SES peers.

Low-SES language input can differ from high-SES input in both overall quantity of speech and the quality of that speech (Hart & Risley 1995, Huttenlocher et al. 2010, Rowe 2012, Schwab & Lew-Williams 2016, Rowe et al. 2017).

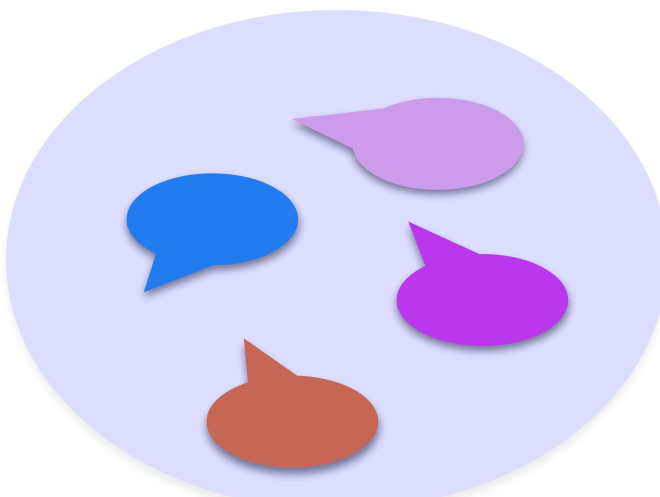
Quality can be measured by different aspects of the input, like diversity of vocabulary …

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

Quality can be measured by different aspects of the input, like diversity of vocabulary, diversity of syntactic constructions, and frequency of decontextualized speech.

The kitty wasn't there

Because we're going tomorrow



We saw her yesterday, didn't we?

The penguins should be at the zoo

Because the penguins were being fed.

The kitty wasn't there

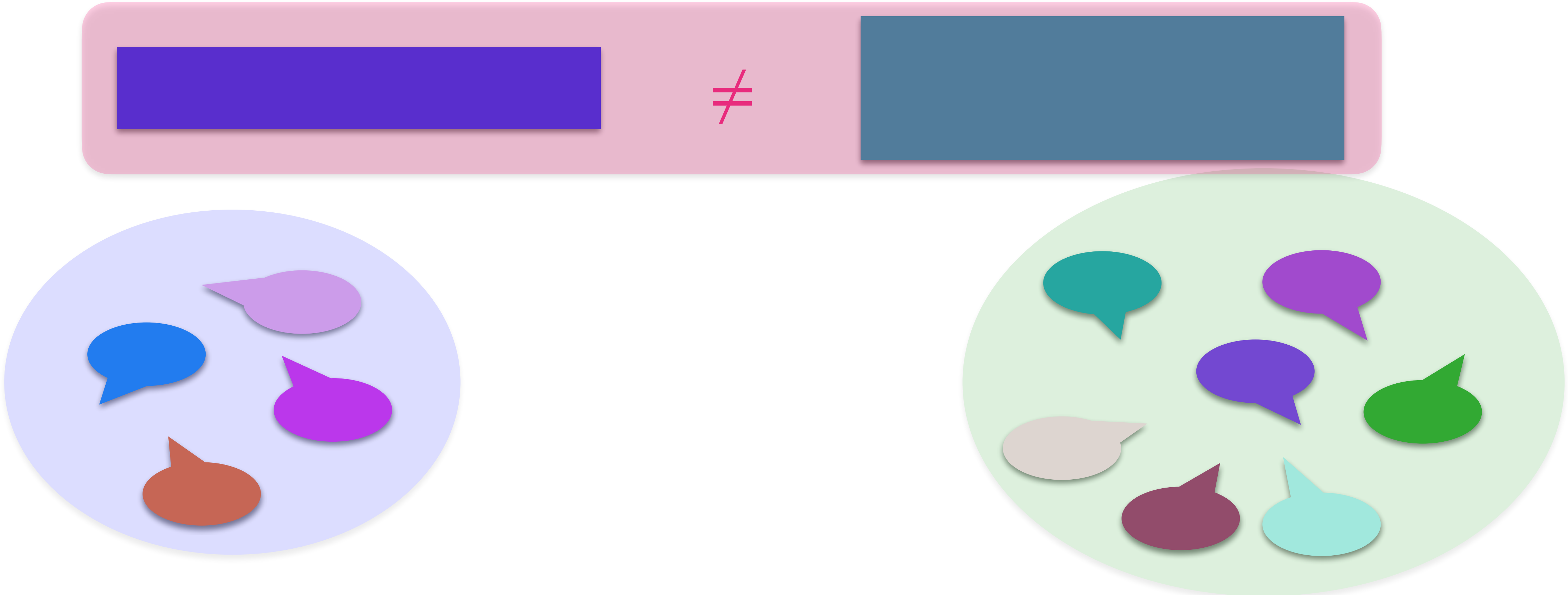
Because we're going tomorrow

We'll see the kitty on Friday

How can we tell if any particular input difference is developmentally meaningful (that is, it impacts language development)?

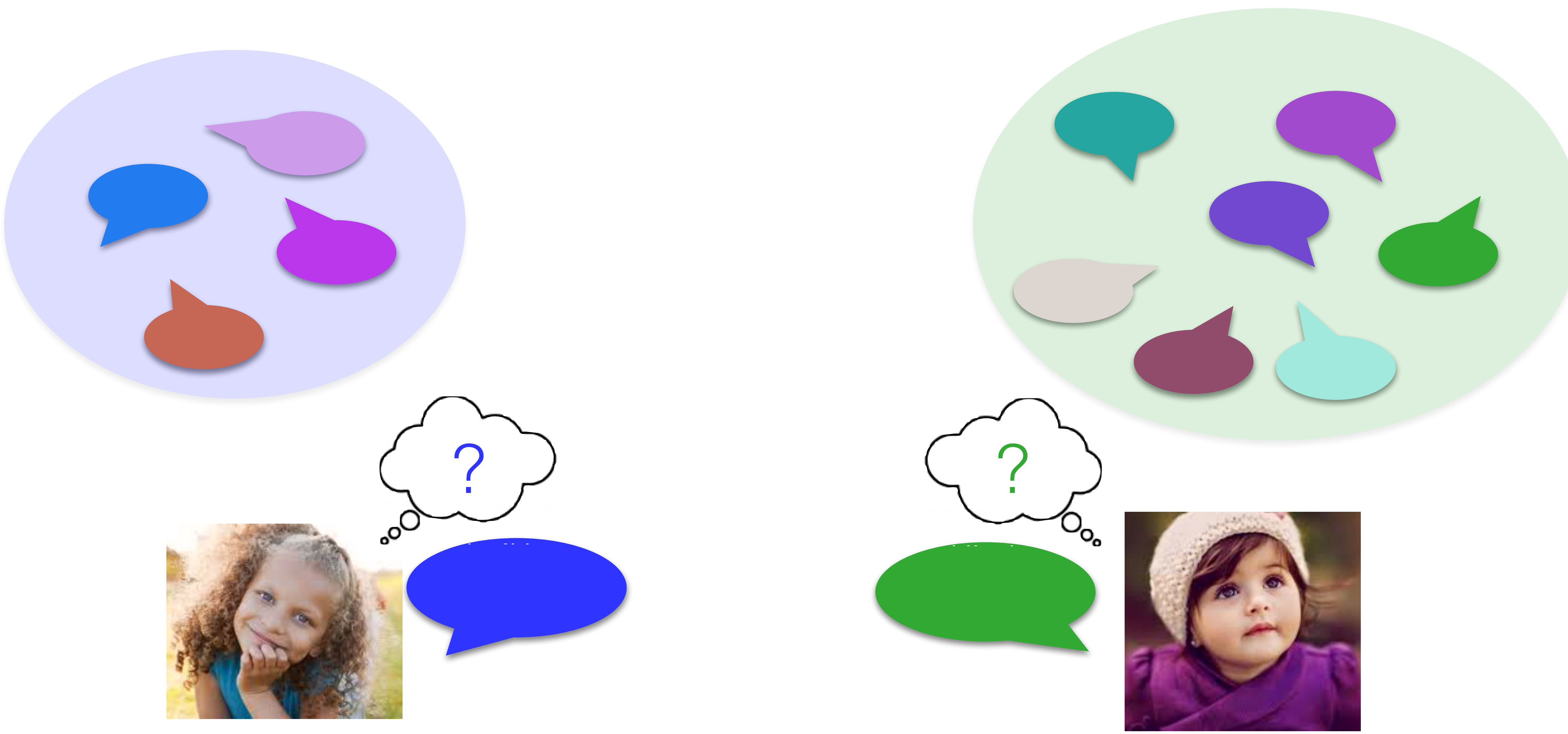
One (standard) way:

- Notice that there's a difference



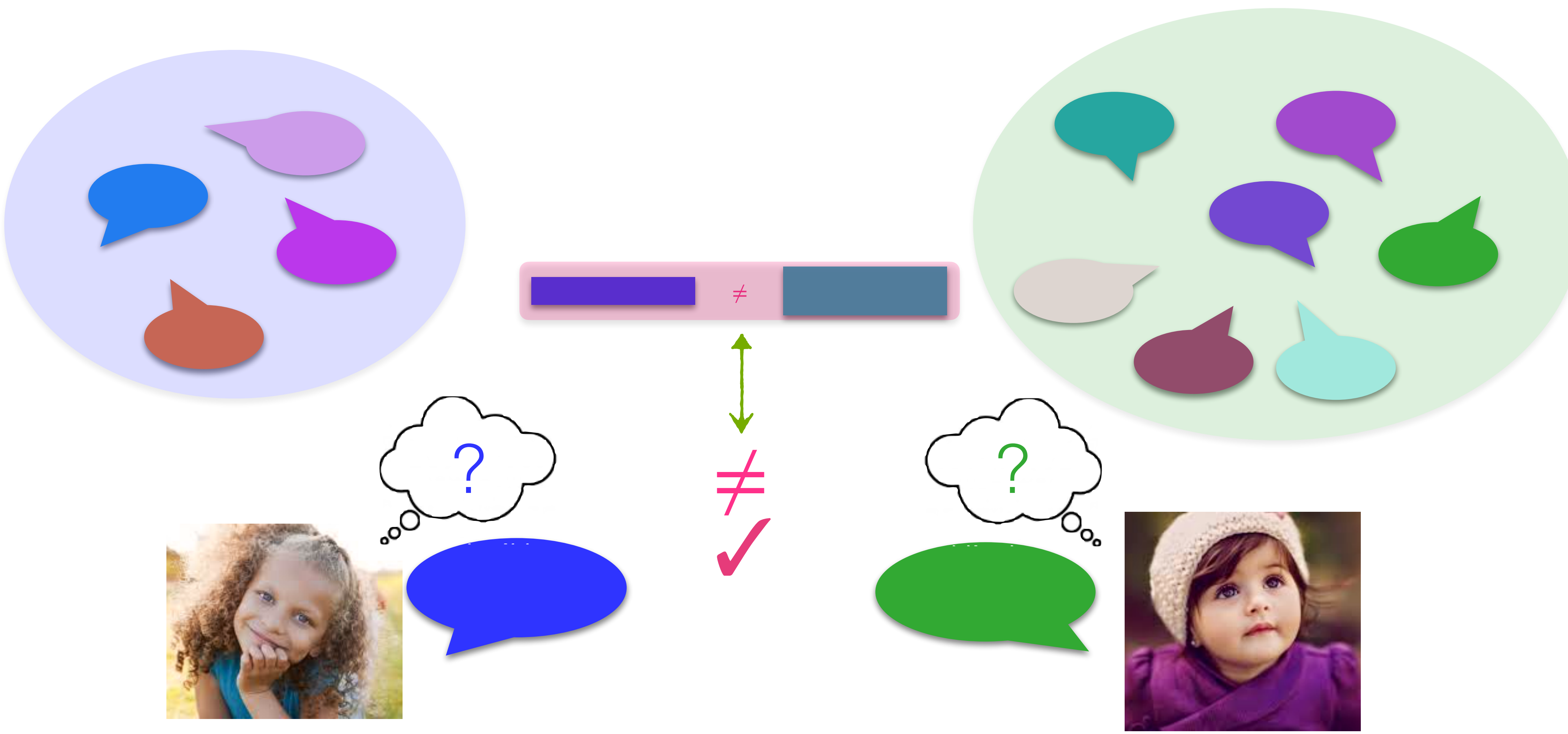
One (standard) way:

- Notice that there's a difference
- Measure language acquisition outcomes



One (standard) way:

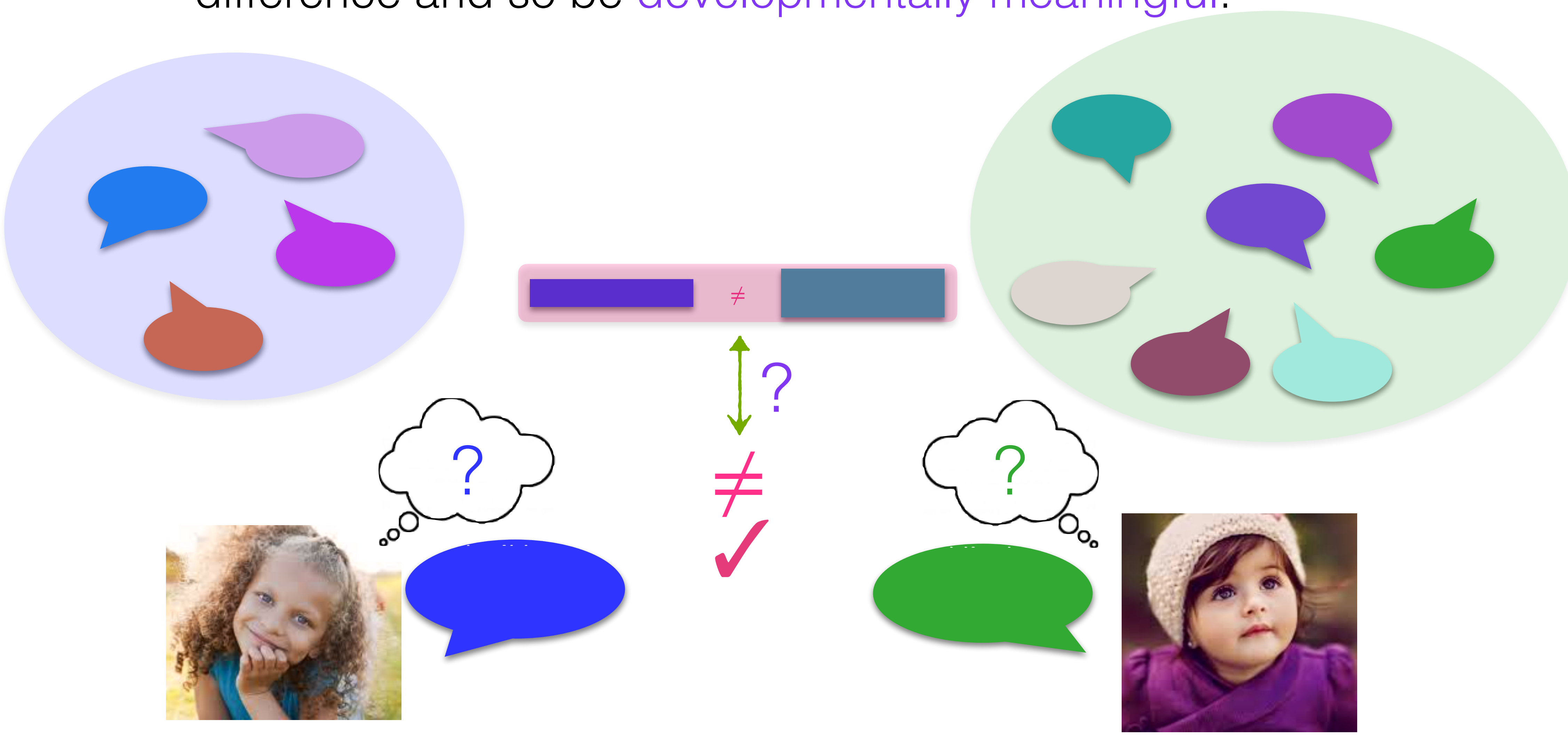
- Notice that there’s a difference
- Measure language acquisition outcomes
- See if that input difference correlates with any outcome differences



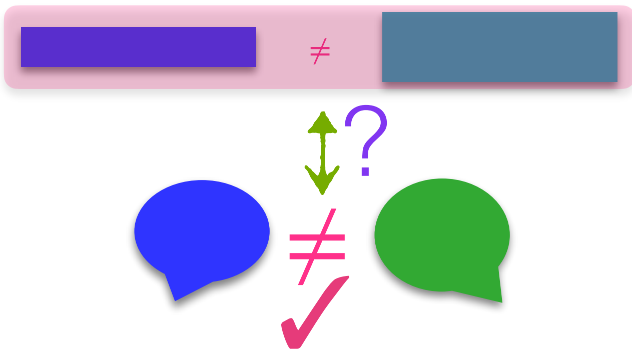
One (standard) way:

- Notice that there's a difference
- Measure language acquisition outcomes
- See if that input difference correlates with any outcome differences

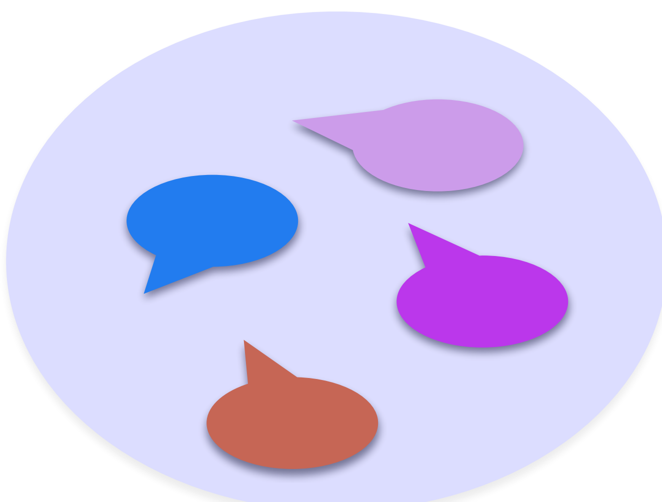
If so, then the input difference *might* cause the outcome difference and so be developmentally meaningful.

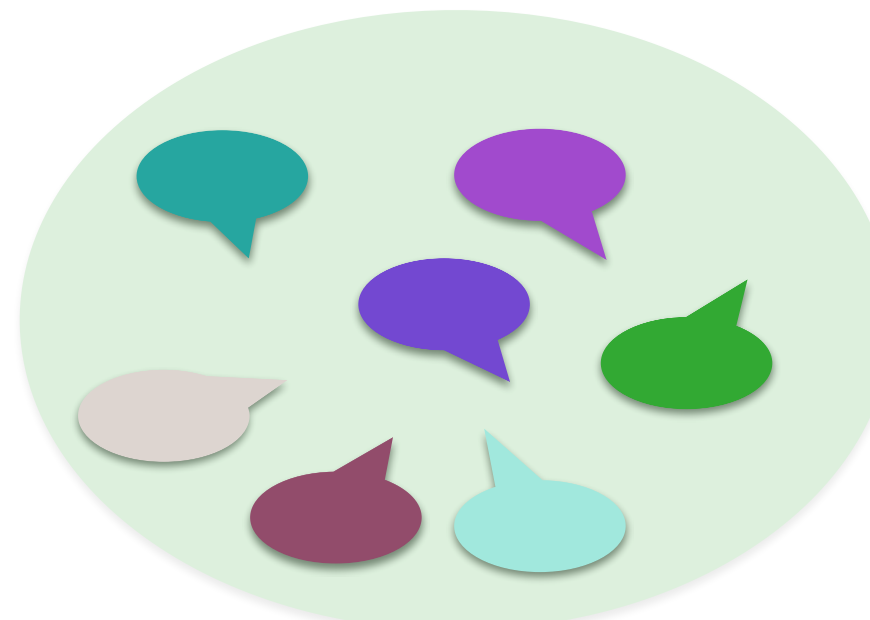


One (standard) way

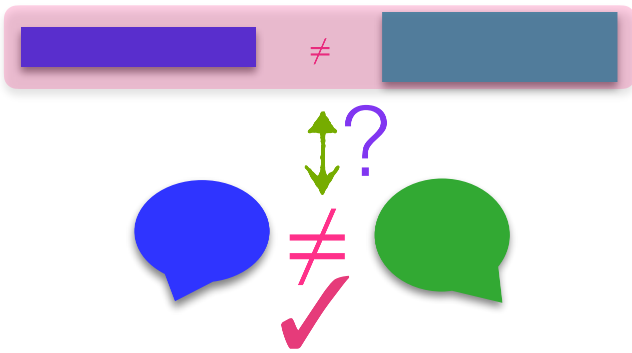


A new (complementary) way uses developmental computational modeling.



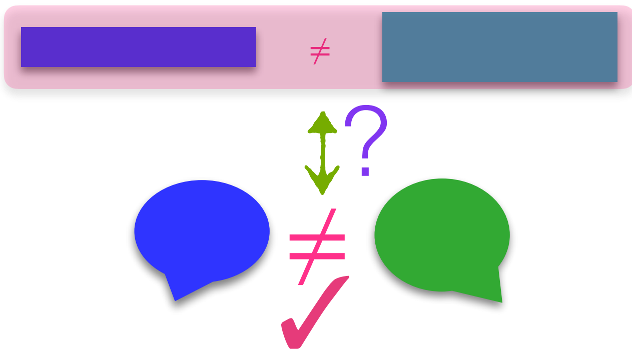


One (standard) way



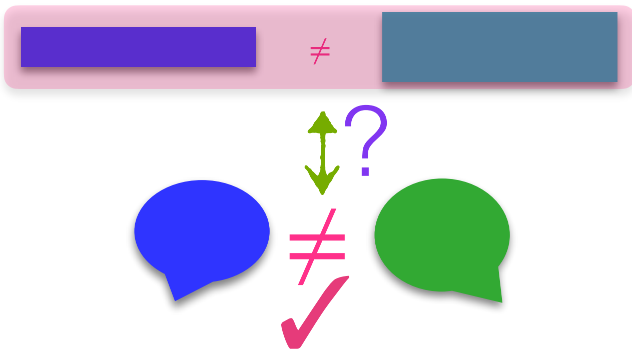
A developmental computational model implements a specific learning theory …

One (standard) way

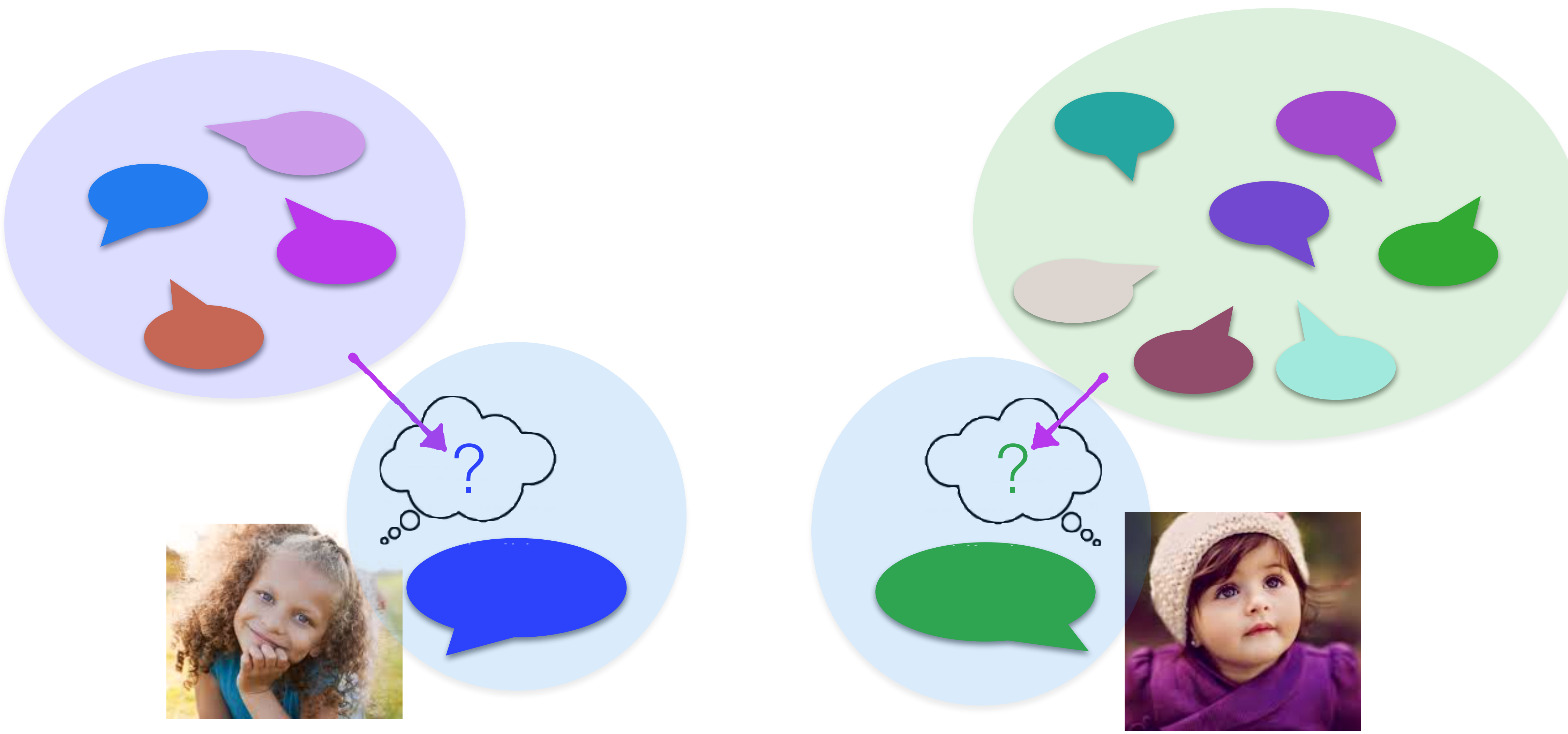


A developmental computational model implements a specific learning theory about how children use their input …

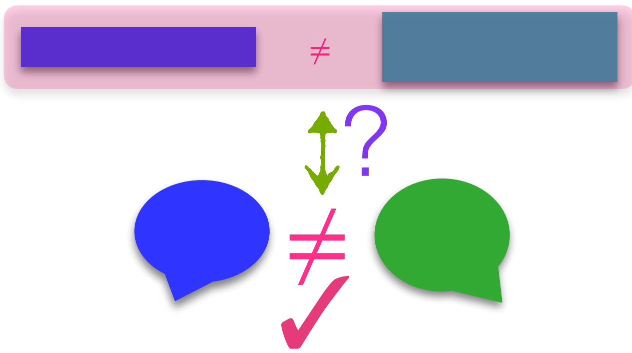
One (standard) way



A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.

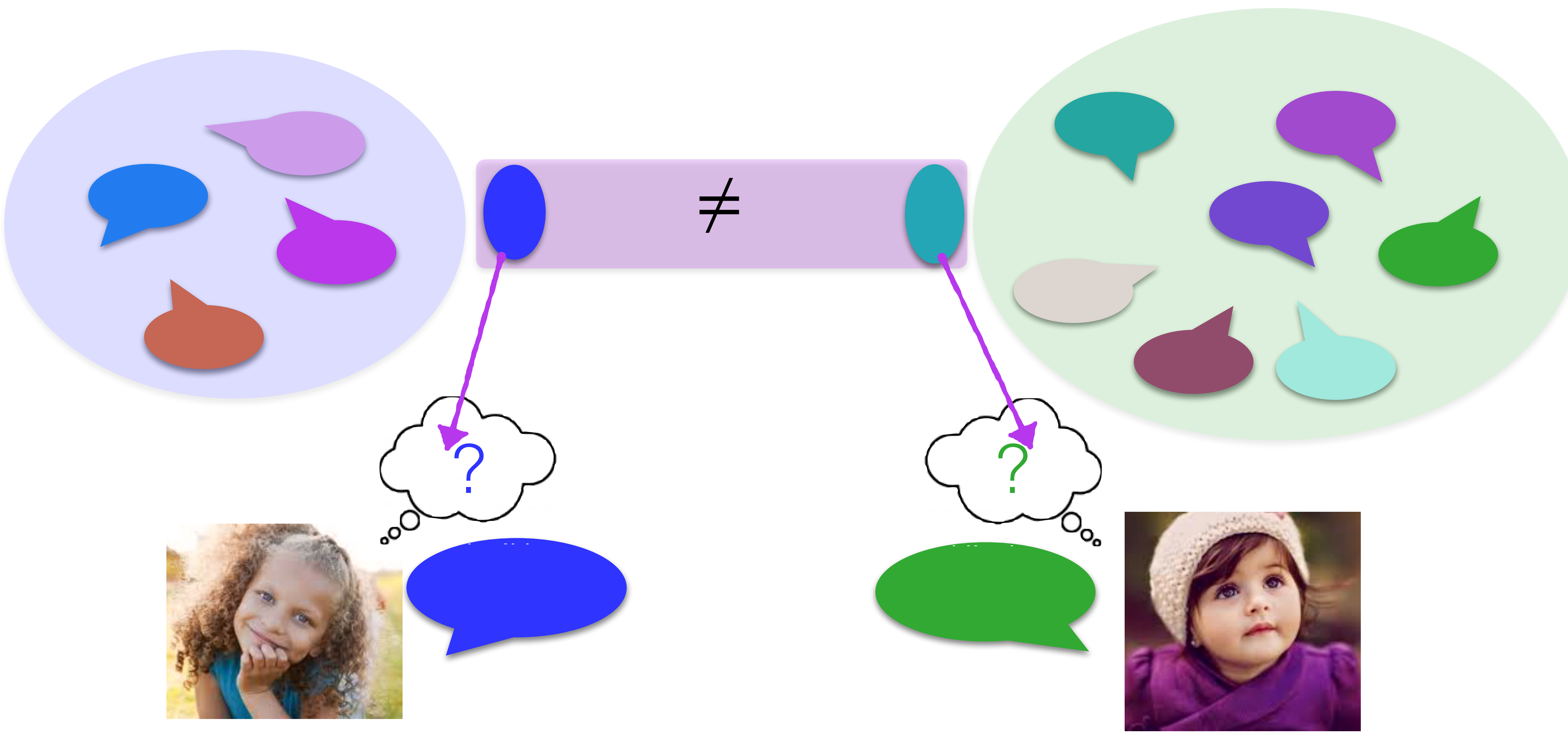


One (standard) way

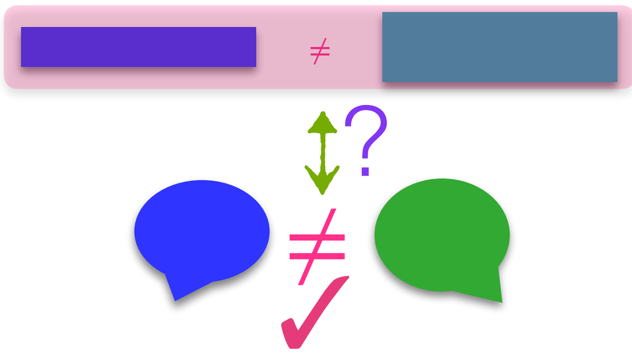


A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.

Important: the learning theory implemented by the model specifies what aspect of the input matters.

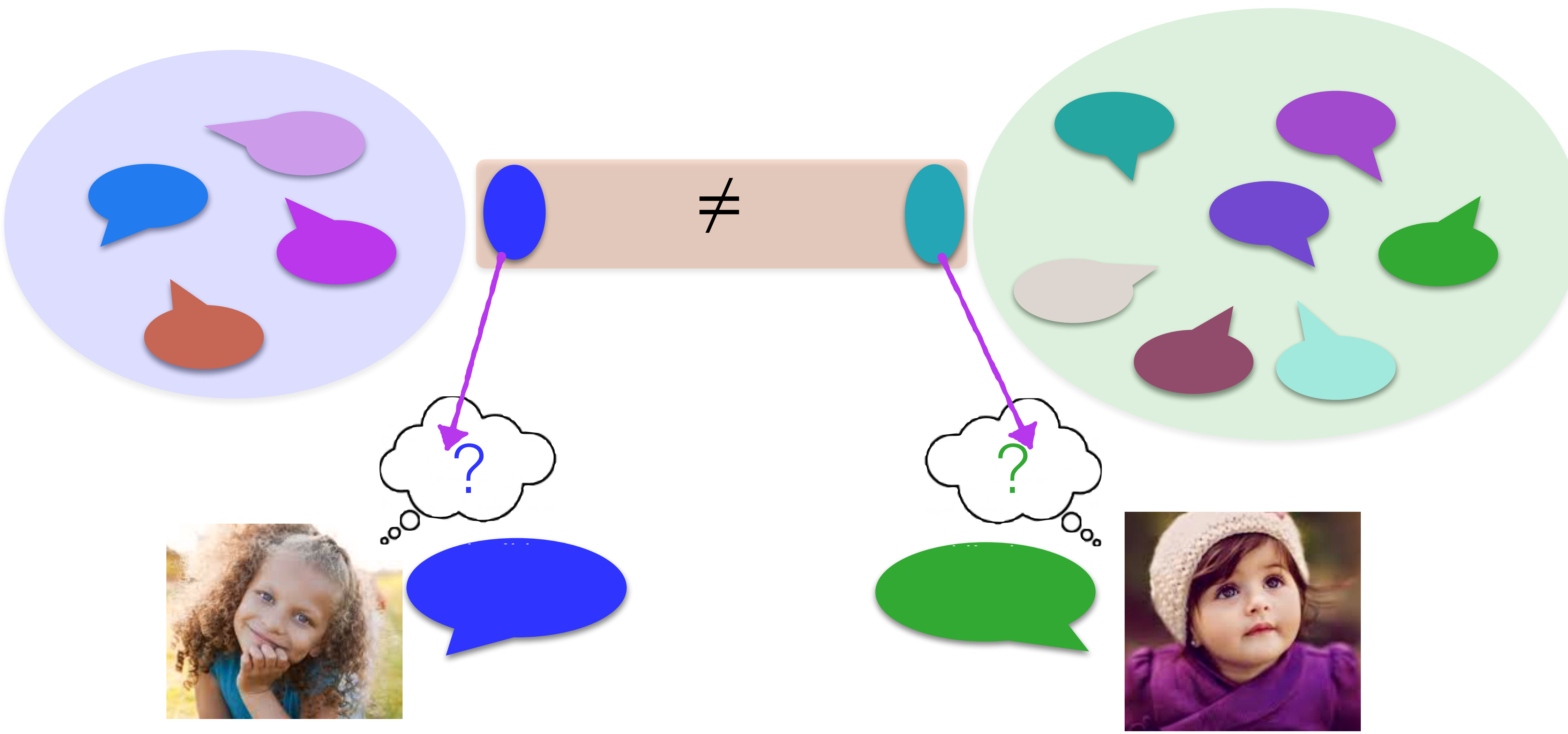


One (standard) way

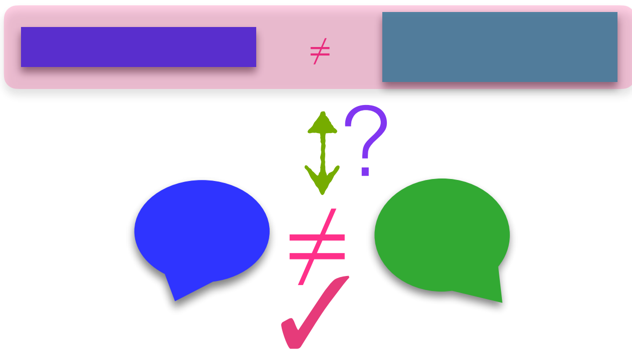


A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.

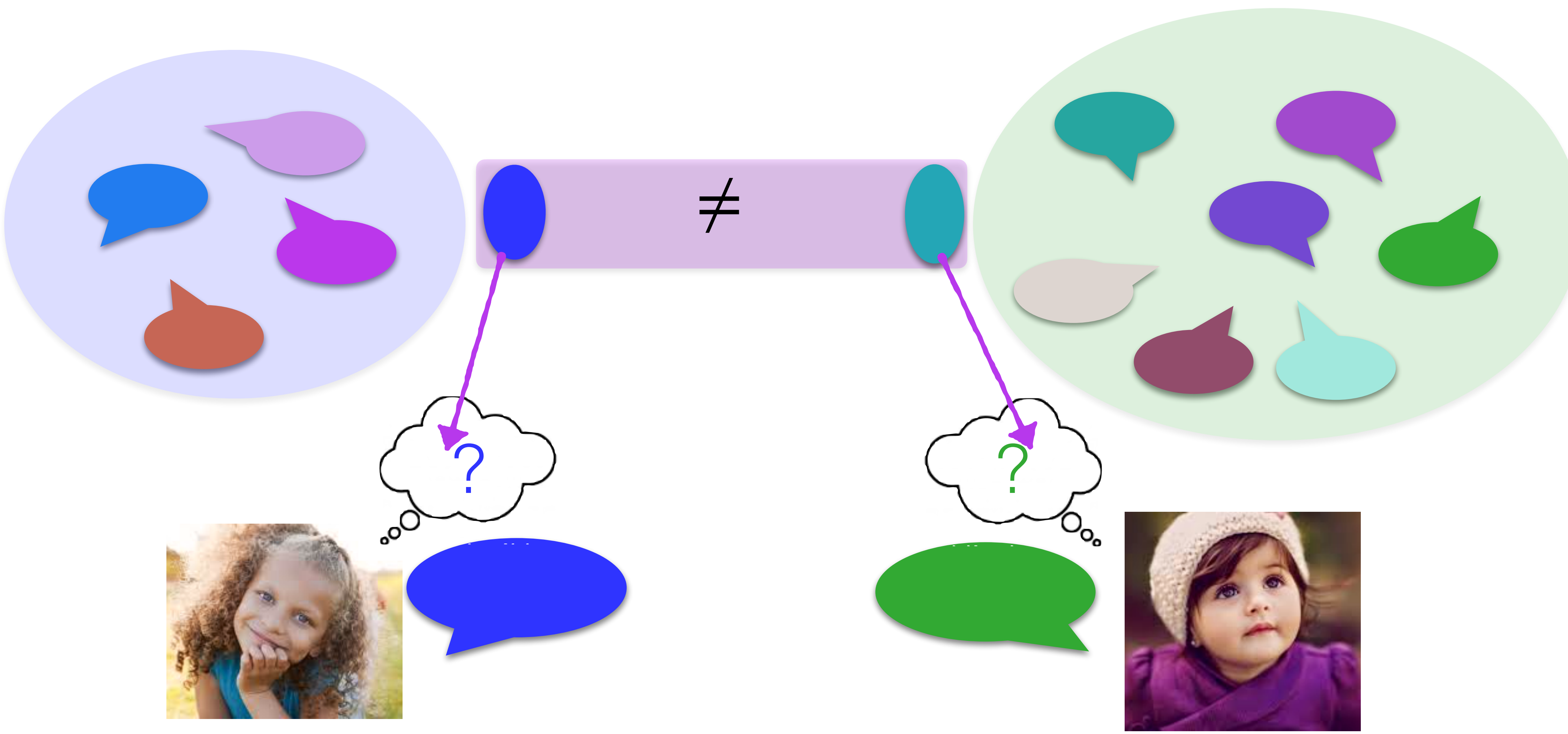
If we know what input part matters, we can target that part for intervention if needed.



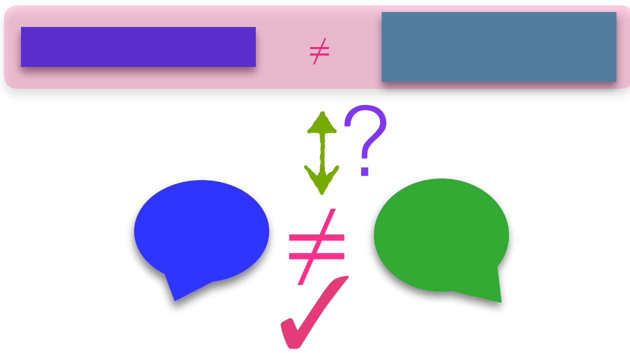
One (standard) way



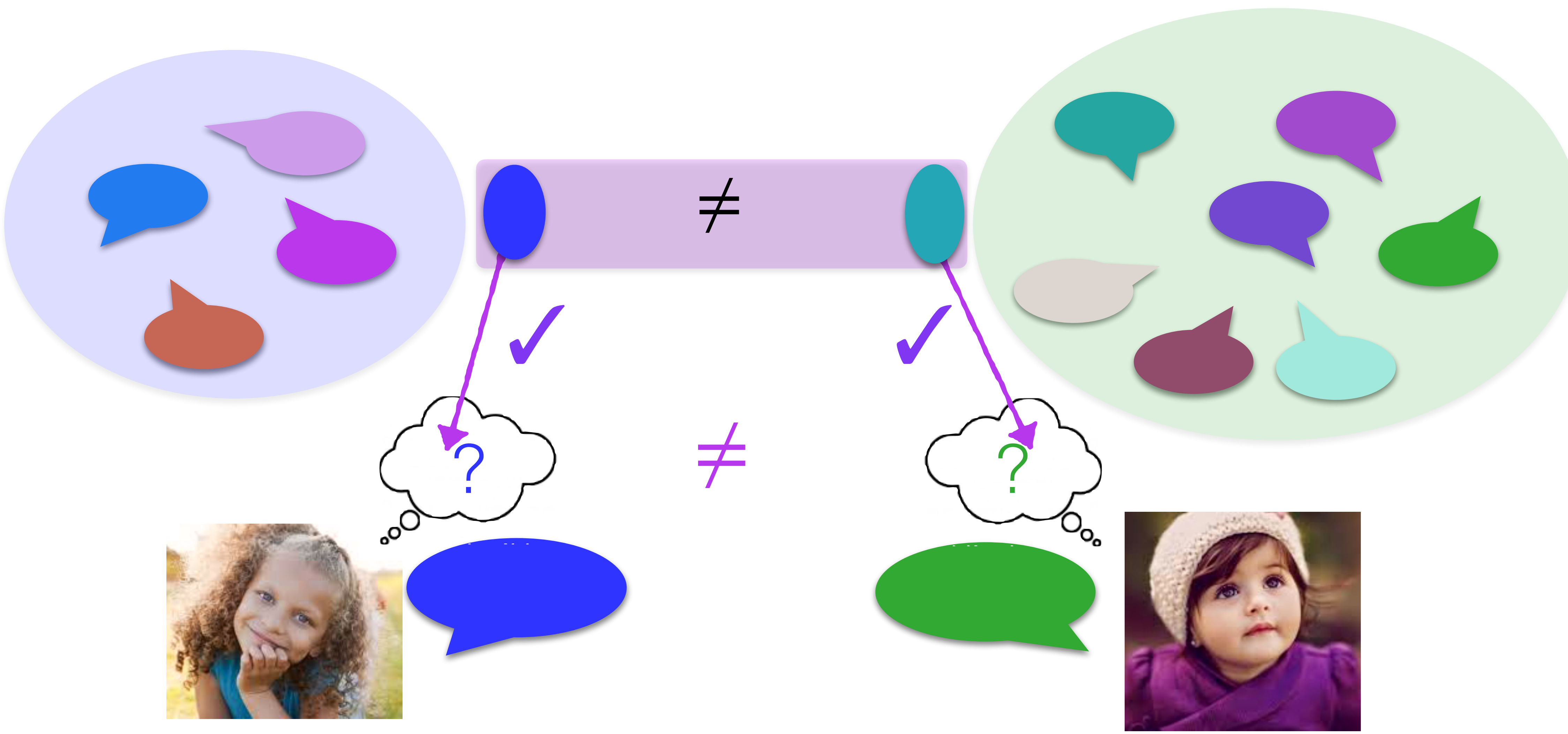
So, a developmental computational model can predict the language outcome on the basis of the input.



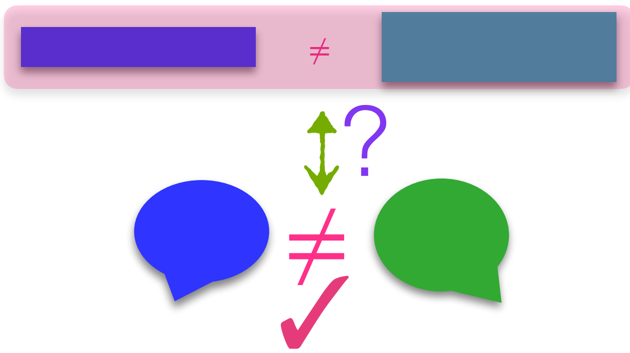
One (standard) way



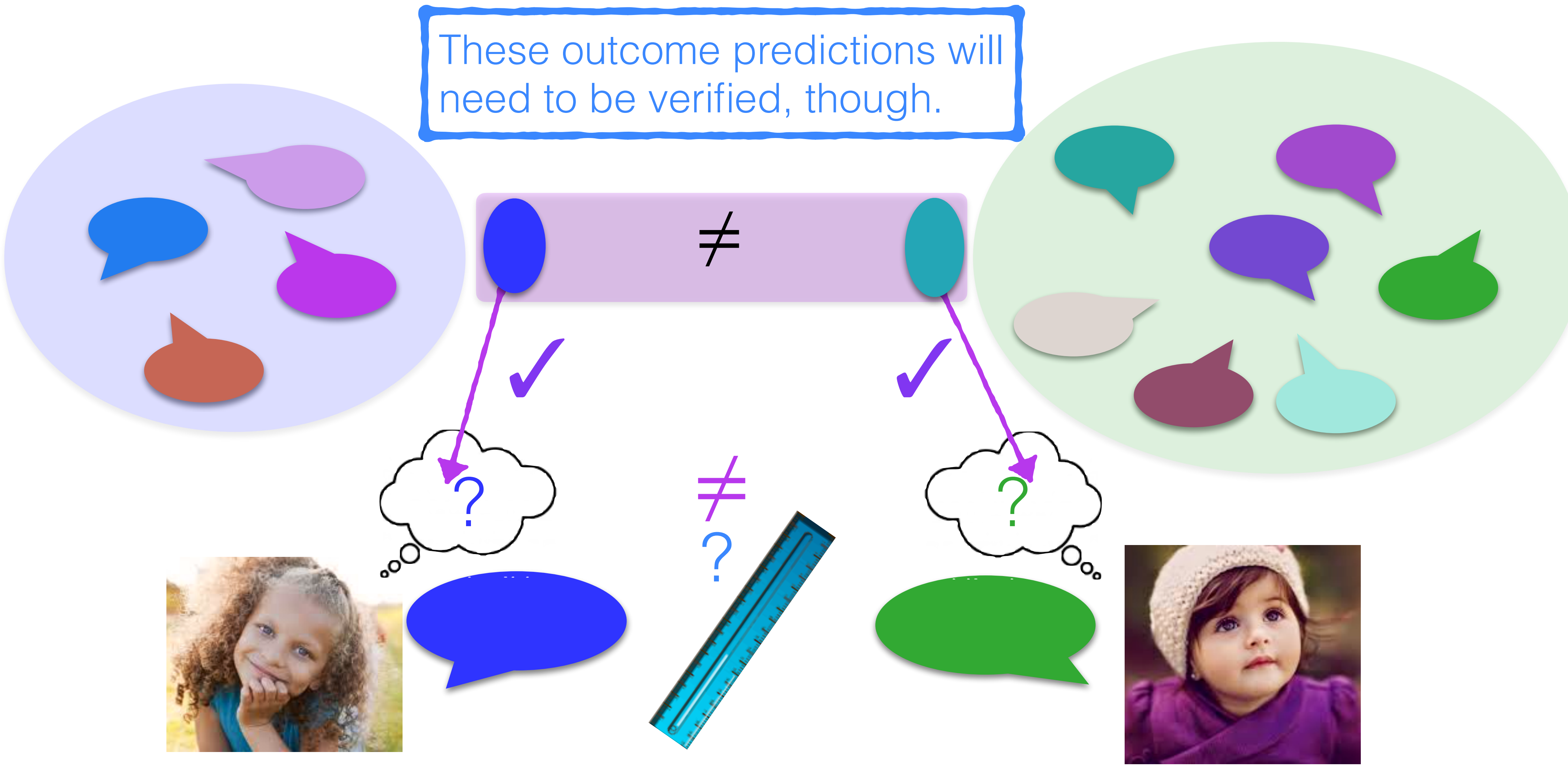
If the predicted outcomes differ, then it's because the input difference caused that outcome difference. So, the input difference is predicted to be developmentally meaningful.



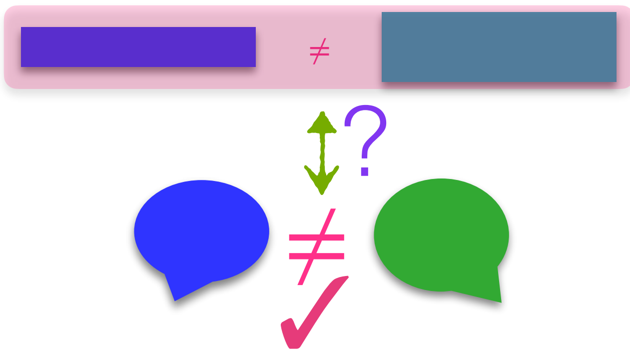
One (standard) way



If the predicted outcomes differ, then it’s because the input difference caused that outcome difference. So, the input difference is predicted to be developmentally meaningful.

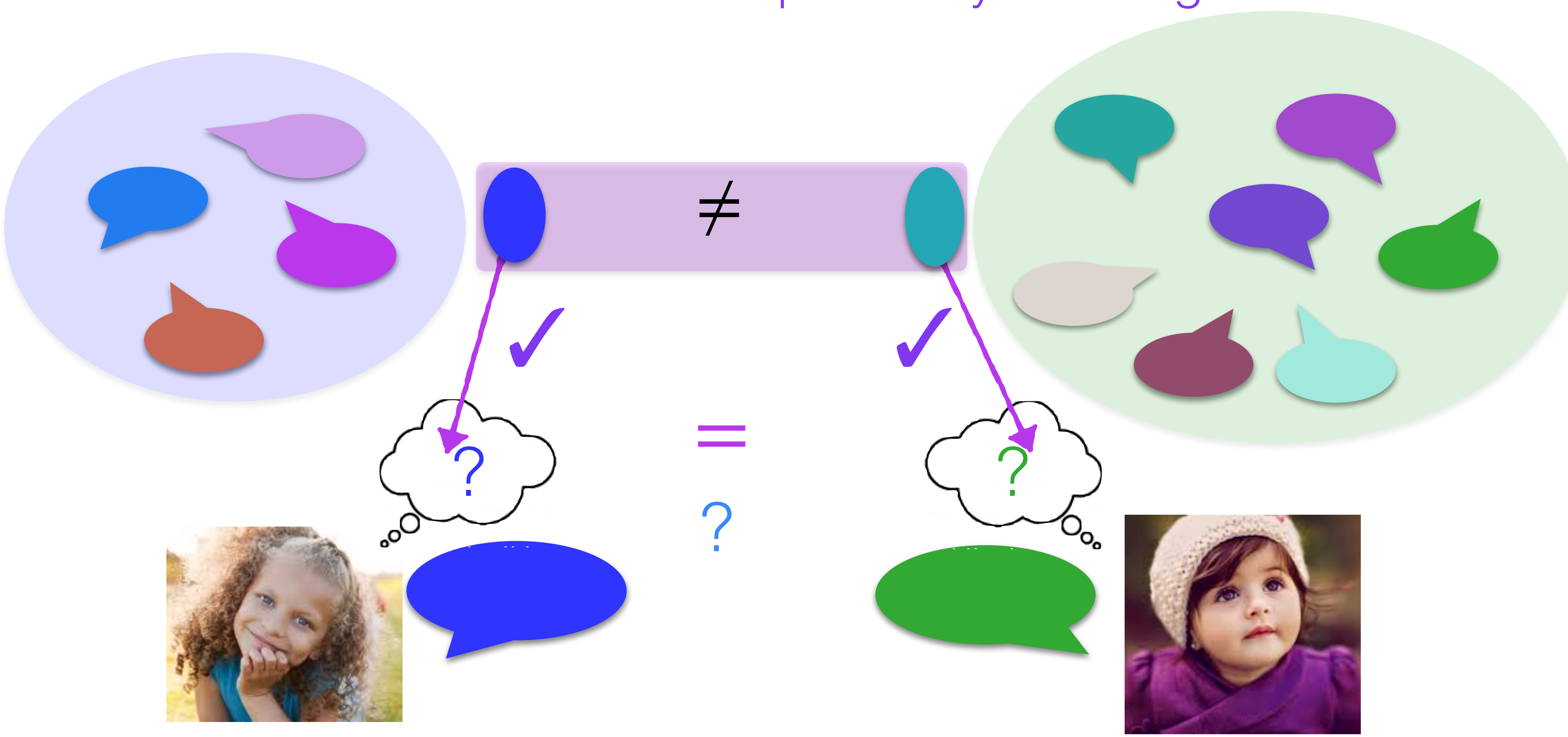


One (standard) way



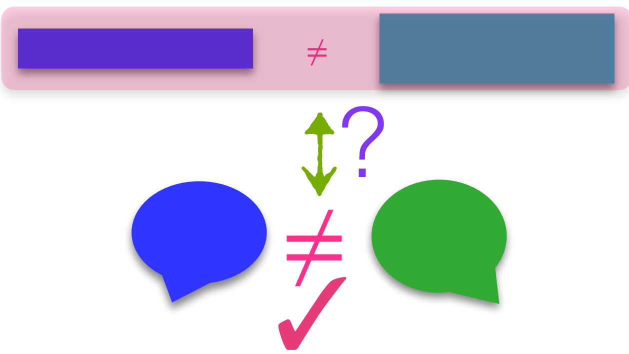
Bonus: Because the learning theory in the model is causal, we can predict if the input should cause similar outcomes, too.

In that case, the input difference isn't developmentally meaningful.

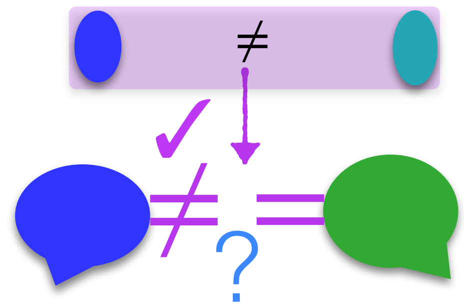


Detecting if input differences are developmentally meaningful

One (standard) way

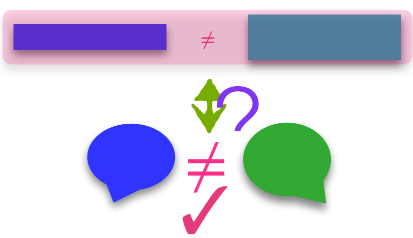


A new (complementary) way



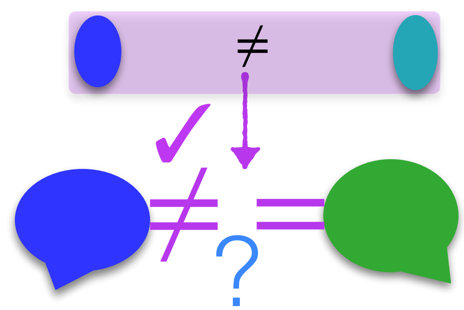
Detecting if input differences are developmentally meaningful

One (standard) way



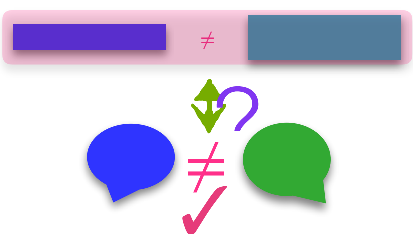
Today's focus

A new (complementary) way



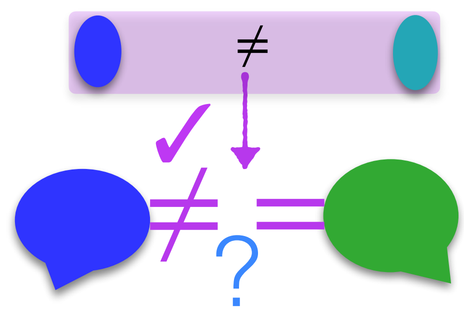
Detecting if input differences are developmentally meaningful

One (standard) way



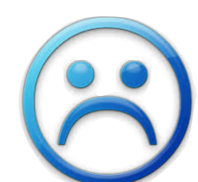
Today's focus

A new (complementary) way

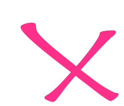


Case study:

Syntactic island acquisition

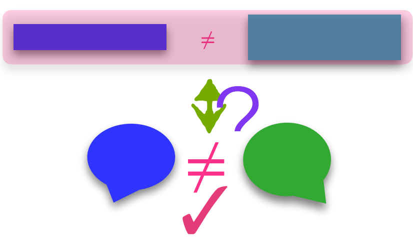


Who does…



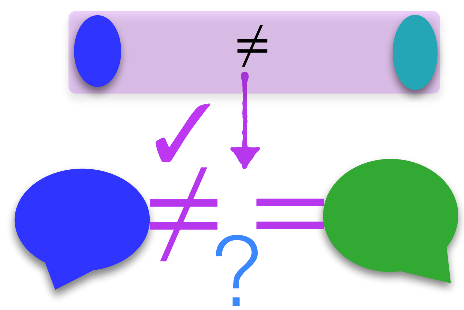
Detecting if input differences are developmentally meaningful

One (standard) way



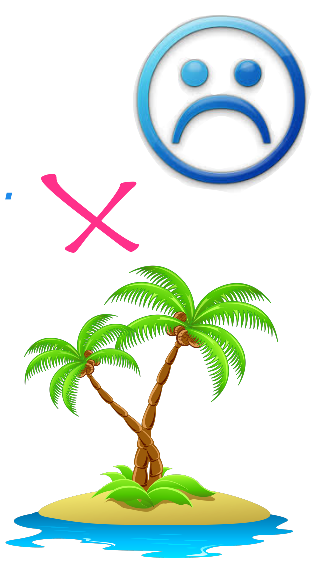
Today's focus

A new (complementary) way



Case study:
Syntactic island acquisition

Who does…



Why? It's higher-order syntactic knowledge where we don't know much about developmentally-meaningful input differences that may exist across SES.

Who does 

Syntactic island constraints

involve *wh*-dependencies.

This kitty was bought as a present for someone.

Lily thinks this kitty is pretty.

What's going on here?

Who does Lily think the kitty for is pretty? 

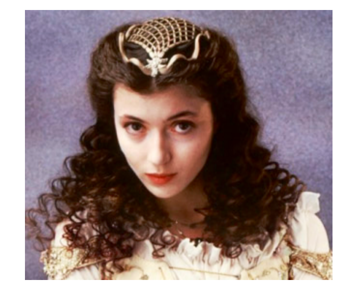
What does Lily think is pretty, and who does she think it's for? 

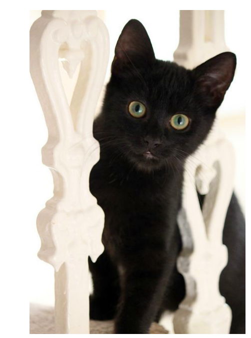
Syntactic island constraints

involve *wh*-dependencies.

What's going on here?

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





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

This dependency is not allowed in English.

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

Who does

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?

syntactic island (Ross 1967)

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

Subject island

Who does

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?

syntactic island (Ross 1967)

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

Jack is somewhat tricksy.

He claimed he bought something.

What did Jack make the claim that he bought __what?

Who does 

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?

syntactic island (Ross 1967)

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

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

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth wondered if he actually did and what it was.

What did Elizabeth wonder whether Jack bought __$_{what}$?

Who does

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?

syntactic island (Ross 1967)

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

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

What did Elizabeth wonder whether Jack bought __what? Whether island

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth worried it was something dangerous.

What did Elizabeth worry if Jack bought __what?

Who does

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?  syntactic island (Ross 1967)

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

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

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

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

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

(Chomsky 1965, Ross 1967)

Who does 

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?  syntactic island (Ross 1967)

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

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

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

What did Elizabeth worry if Jack bought $\_\_{what}$*?* Adjunct island

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

What did Elizabeth think $\_\_{what}$*?*

✔

Elizabeth

Who does 

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?  syntactic island (Ross 1967)

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

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

What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

Elizabeth

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

Jack

What did Elizabeth think Jack said __what?

✔

Who does 

Syntactic island constraints

involve *wh*-dependencies.

What's going on here?  syntactic island (Ross 1967)

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

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

What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

Elizabeth

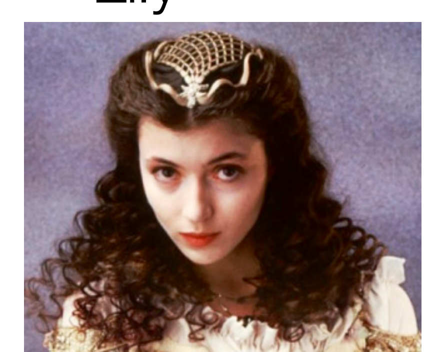
Jack

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

What did Elizabeth think Jack said Lily saw __what?

✔

Lily



Who does

Syntactic island constraints

involve *wh*-dependencies.

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

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

What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

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

These judgments are a measurable observable behavior that signal the successful acquisition of syntactic island knowledge.

Syntactic island constraints

involve *wh*-dependencies.

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

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

What did Elizabeth wonder whether Jack bought __what? Whether island

What did Elizabeth worry if Jack bought __what? Adjunct island

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

So, these judgments can serve as a target for successful acquisition — an outcome we can measure.

Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. 2012: magnitude estimation judgments

- factorial definition controlling for two salient properties of island-crossing dependencies

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

Who __*who*?

Who [CP… __*who*]?

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

Who [non-island]?

Who [island]?

Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

X

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

Complex NP island stimuli

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

Sprouse et al. 2012

Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

Complex NP island stimuli

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

Also, 4-year-olds across SES strongly disprefer *wh*-dependencies like this one (De Villiers et al. 2008).

Sprouse et al. 2012

Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

Subject island stimuli

Stimulus	Condition
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 island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

Whether island stimuli

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

Sprouse et al. 2012

Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

Adjunct island stimuli

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

Sprouse et al. 2012

Syntactic island constraints

High-SES adult judgments

= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

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

Who does

Syntactic island constraints

High-SES adult judgments
= behavioral target outcome

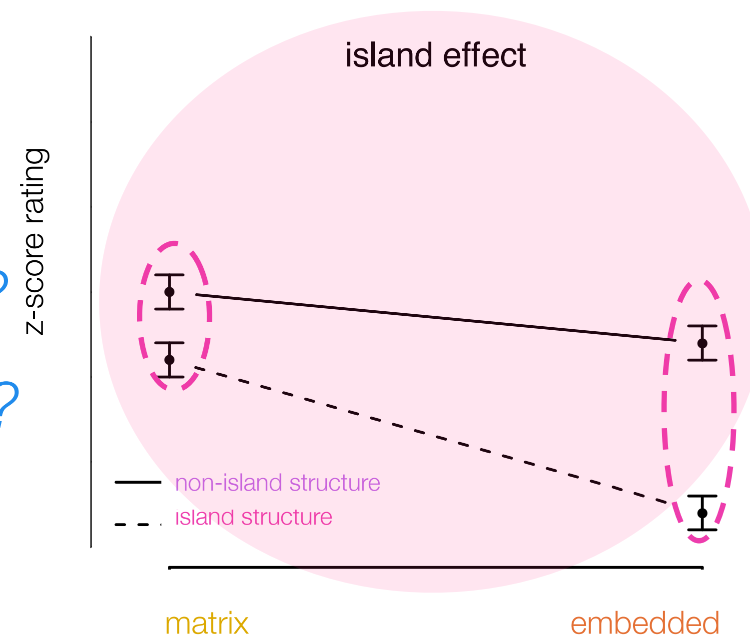
Adult knowledge as measured by acceptability judgment behavior

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

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

Who [non-island]?

Who [island]?



additional unacceptability

Who *__who*?

Who [CP… *__who*]?

Sprouse et al. 2012

Syntactic island constraints

High-SES adult judgments

= behavioral target outcome

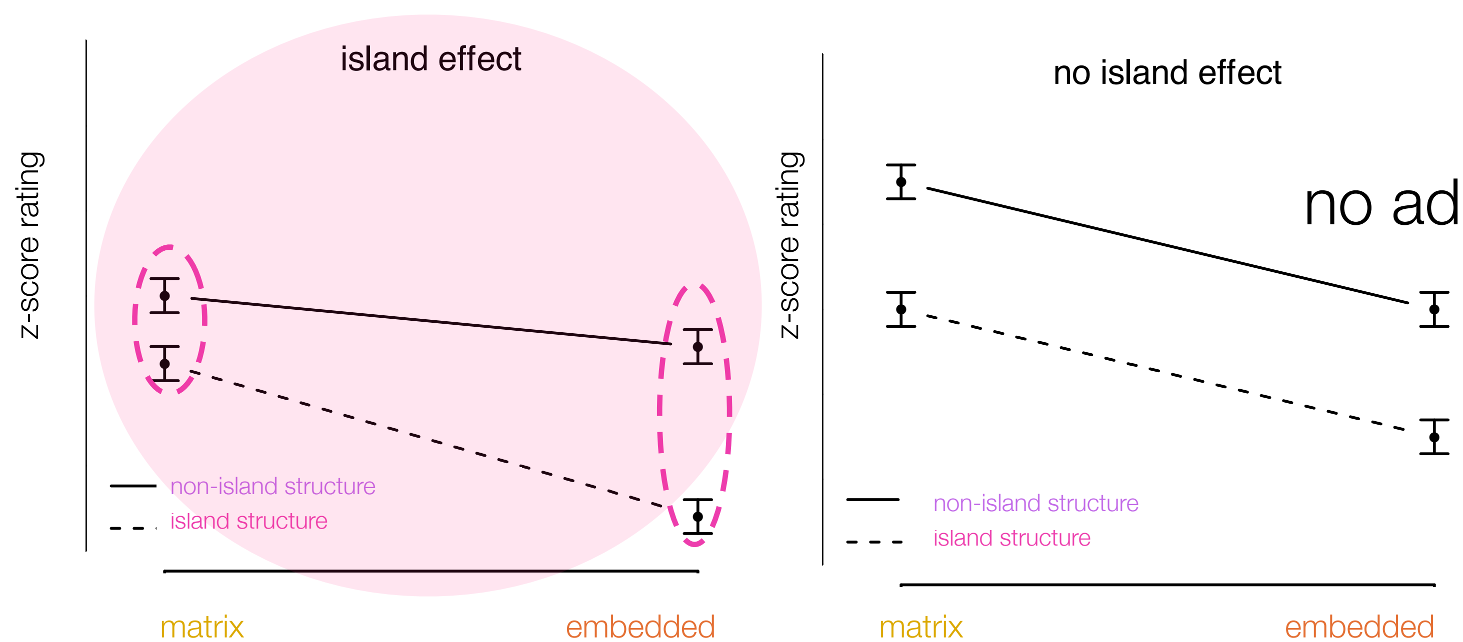
Adult knowledge as measured by acceptability judgment behavior

length of dependency
(matrix vs. embedded)

X

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

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



no additional unacceptability

Sprouse et al. 2012

Syntactic island constraints

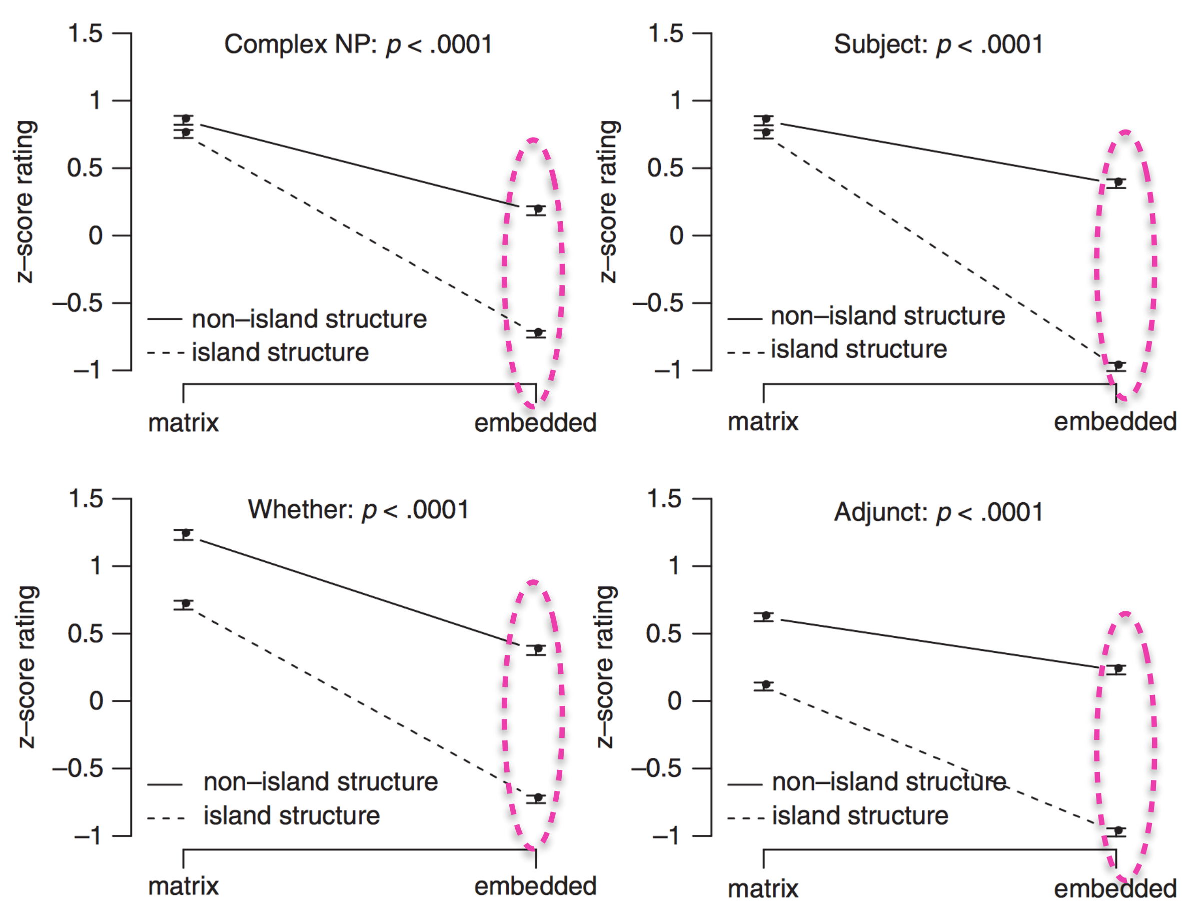
High-SES adult judgments
= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

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

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

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



✓ superadditivity for all four island types

Sprouse et al. 2012

Syntactic island constraints

High-SES adult judgments

= behavioral target outcome

Adult knowledge as measured by acceptability judgment behavior

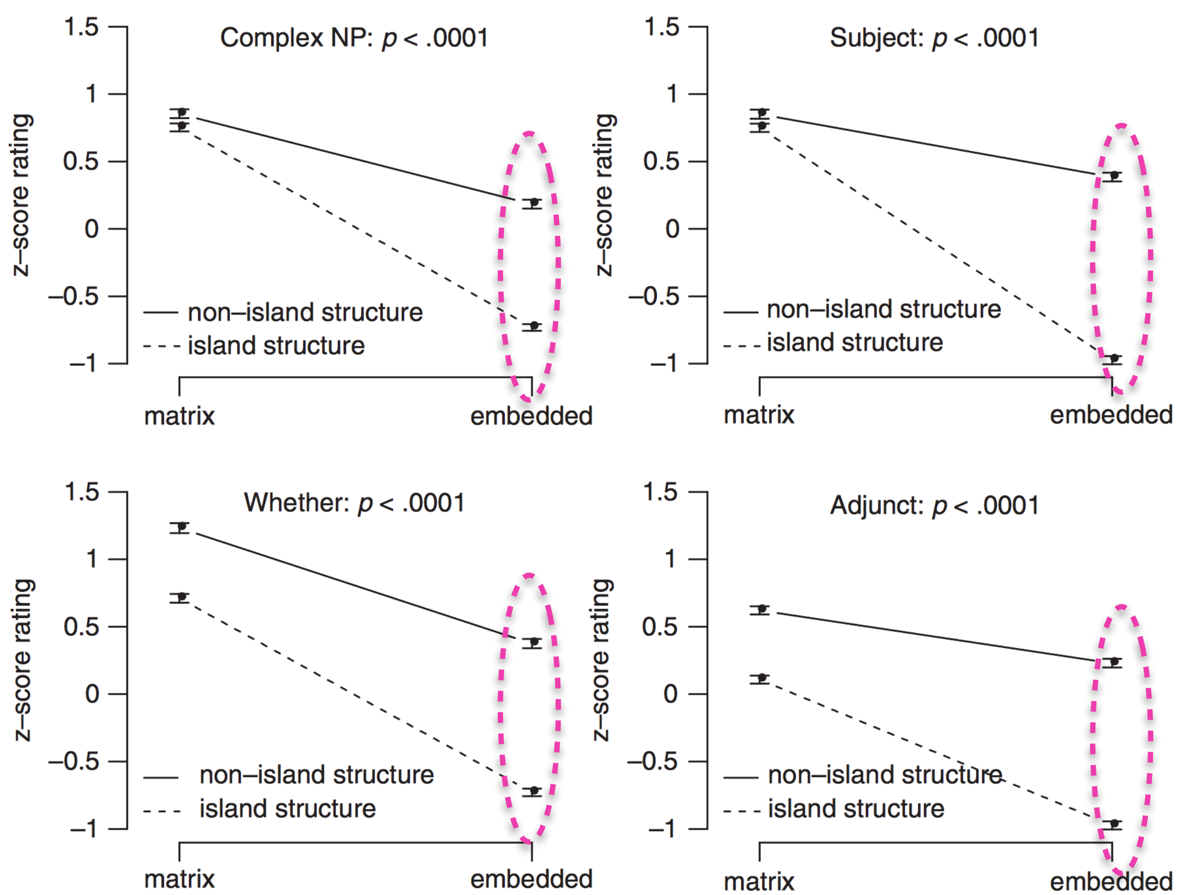
length of dependency
(matrix vs. embedded)

X

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

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

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

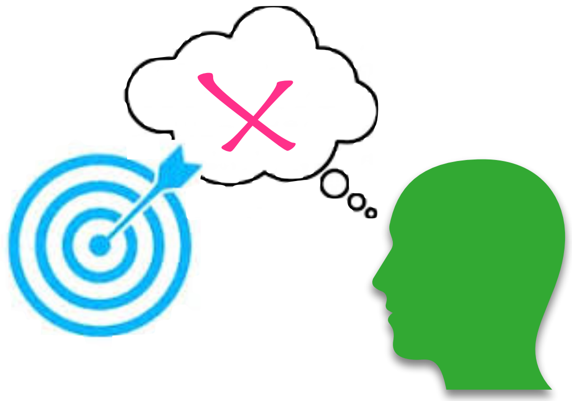


superadditivity for all four island types

= knowledge that dependencies can't cross these island structures.

Sprouse et al. 2012

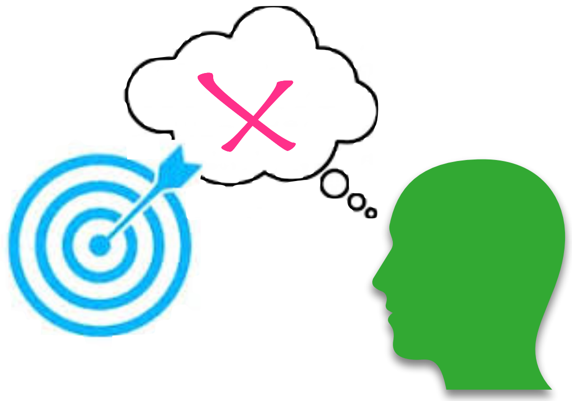
Who does



Okay, so what's the relevant input for learning this target knowledge?

Who does

That depends on how we think children learn it.



That depends on how we think children learn it.

Pearl & Sprouse 2013 intuition for a strategy:

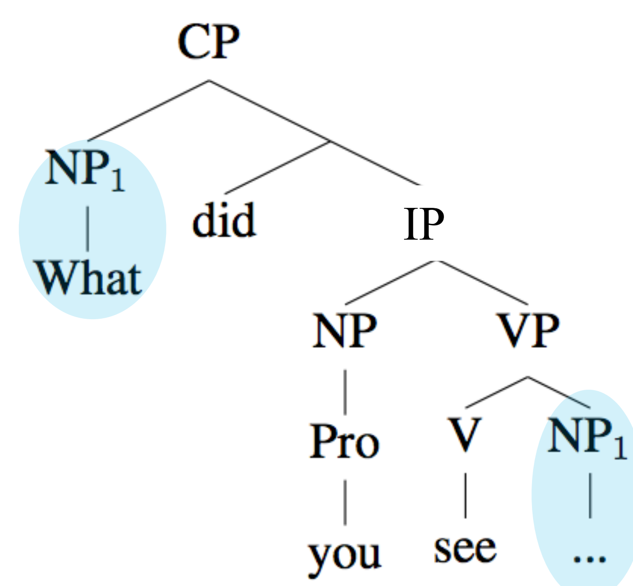
- Learn what you can from the dependencies you do actually observe in the input
- Apply it to make a judgment about the dependencies you haven't seen before, like those crossing syntactic islands (and maybe other longer dependencies, too).

The learning theory = a concrete learning strategy (Pearl & Sprouse 2013): View *wh*-dependencies in terms of their building blocks and track those building blocks in the input.

A strategy for learning syntactic islands

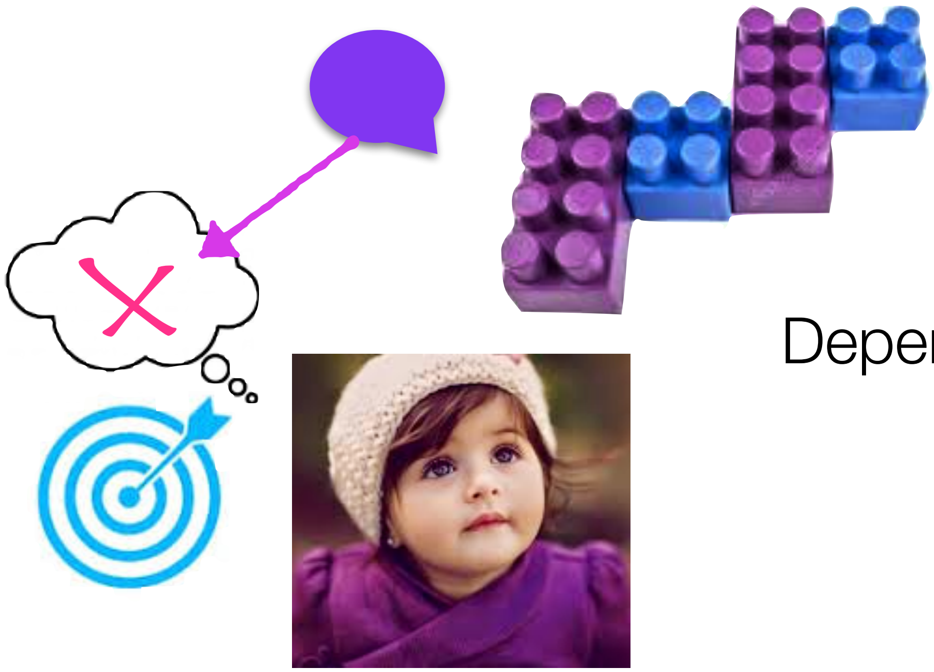
Dependencies represented as a sequence of container nodes

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



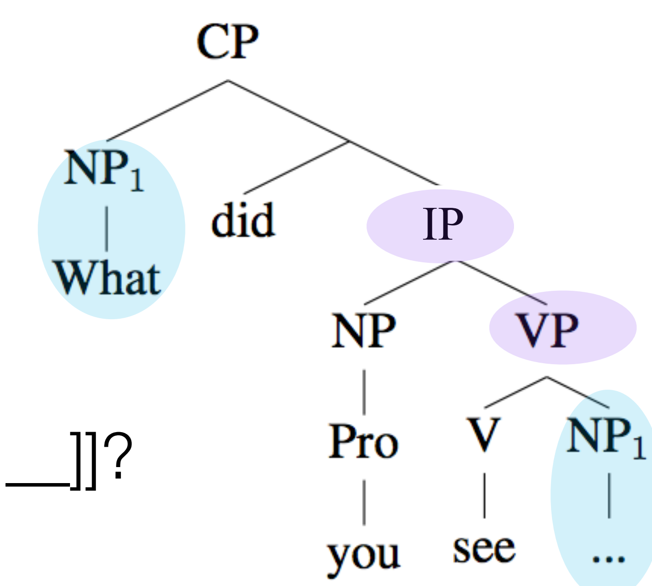
Pearl & Sprouse 2013

A strategy for learning syntactic islands



Dependencies represented as a sequence of container nodes

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



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

Pearl & Sprouse 2013

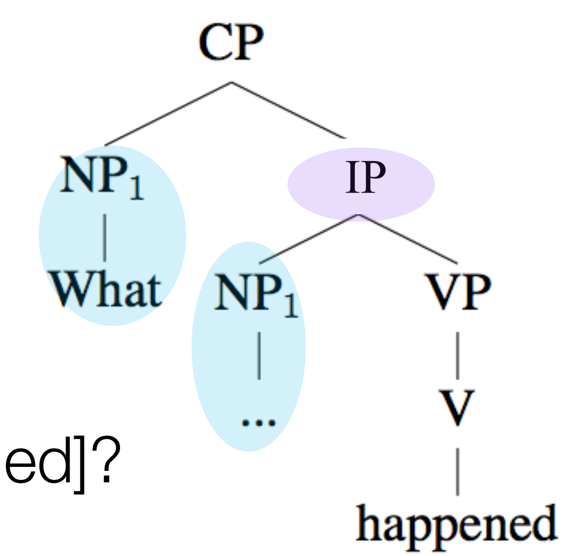
A strategy for learning syntactic islands

Who does

Dependencies represented as a sequence of container nodes

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

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



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

Pearl & Sprouse 2013

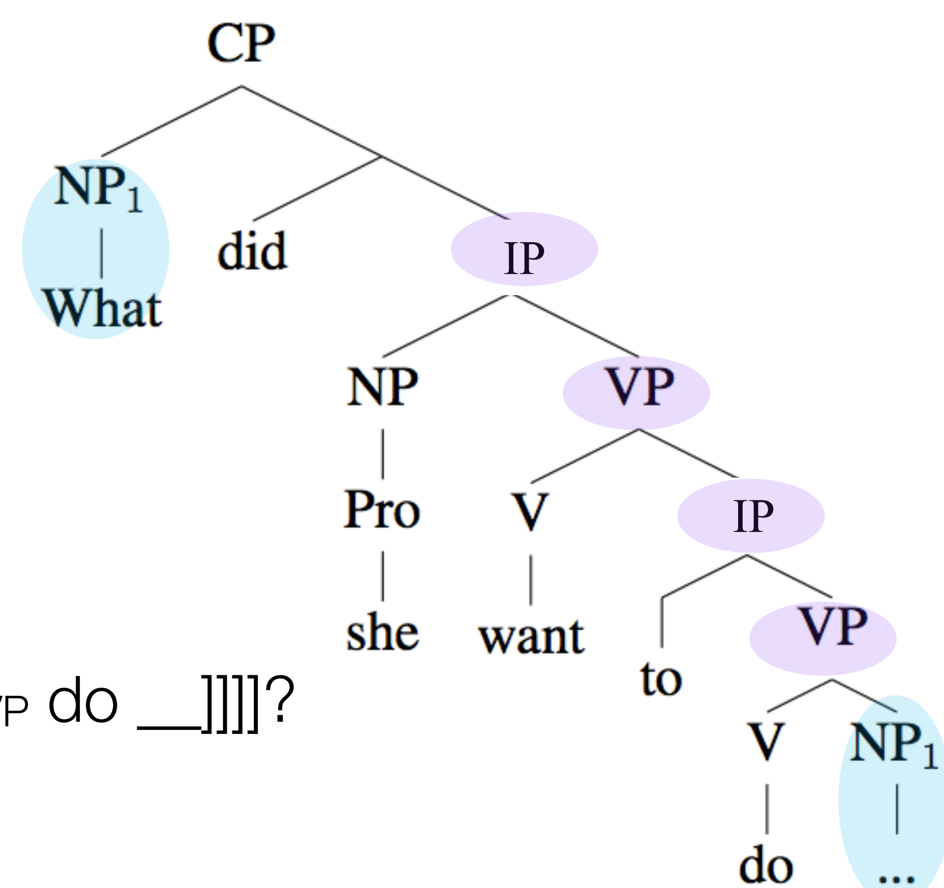
A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

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

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

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



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

Pearl & Sprouse 2013

A strategy for learning syntactic islands

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

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

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

(Much) less acceptable dependencies have low probability segments

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

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

Pearl & Sprouse 2013

A strategy for learning syntactic islands

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

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

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

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

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

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

Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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

$\boxed{\text{syntactic trigrams}}$

The building blocks: trigrams of container nodes

Pearl & Sprouse 2013

Who does 

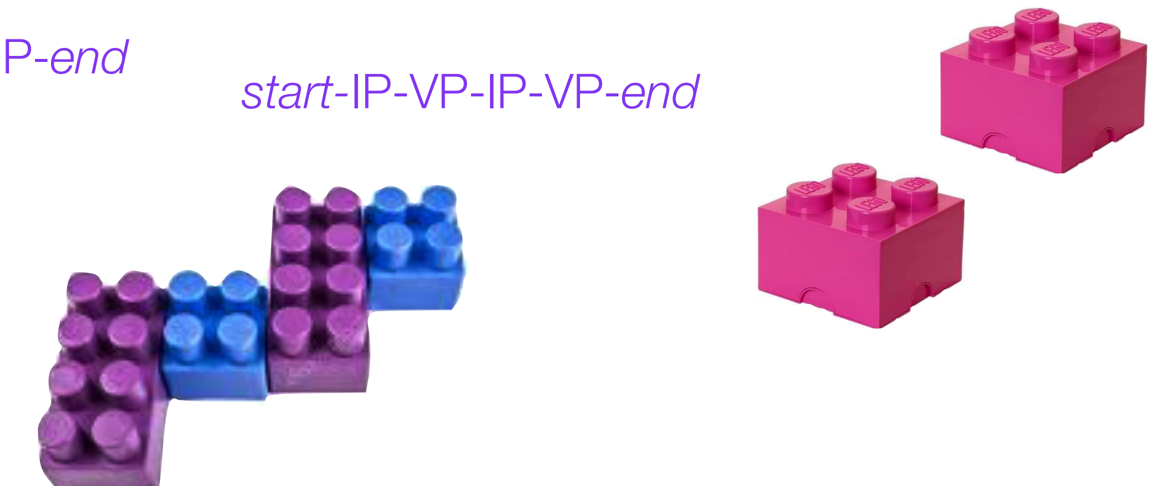
A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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



syntactic trigrams

The building blocks: trigrams of container nodes

start-IP-VP-*end*

start-IP-VP

IP-VP-*end*

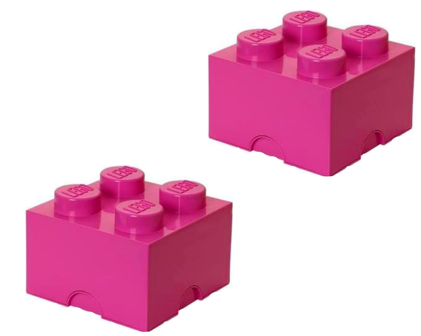
Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*



syntactic trigrams

The building blocks: trigrams of container nodes

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

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-*end*

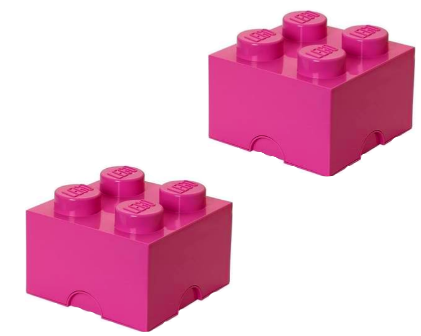
Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

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

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



syntactic trigrams

The building blocks: trigrams of container nodes

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-*end*

start-IP-*end*

start-IP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

Who does

start-IP-VP-*end*

start-IP-*end*

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

syntactic trigrams

The building blocks: trigrams of container nodes

IP-VP-IP

VP-IP-VP

IP-VP-*end*

start-IP-*end*

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

start-IP-VP

IP-VP-CP_{that}

VP-CP_{that}-IP

CP_{that}-IP-NP

IP-NP-PP

NP-PP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

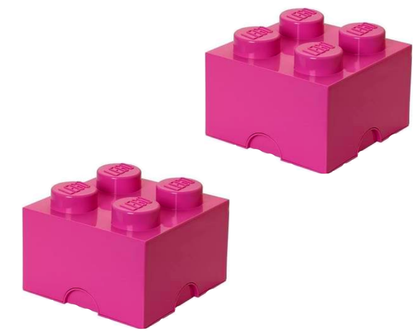
Who does

start-IP-VP-*end*

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

start-IP-*end*

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



The strategy: Track the relative frequency of the syntactic trigrams in your input

start-IP-VP

IP-VP-*end*

IP-VP-IP

VP-IP-VP

start-IP-*end*

IP-VP-CP$_{that}$

VP-PP-*end*

IP-VP-PP

VP-CP$_{that}$-IP

NP-PP-*end*

Pearl & Sprouse 2013

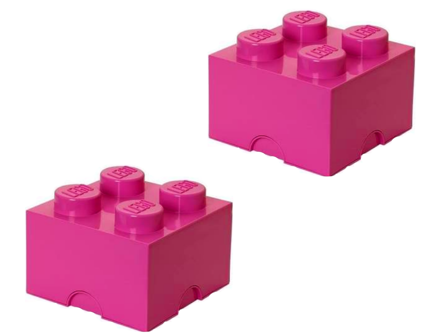
A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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



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

start-IP-VP

start-IP-*end*

IP-VP-*end*

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

IP-VP-IP

IP-VP-PP

VP-PP-*end*

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

VP-IP-VP

NP-PP-*end*

Pearl & Sprouse 2013

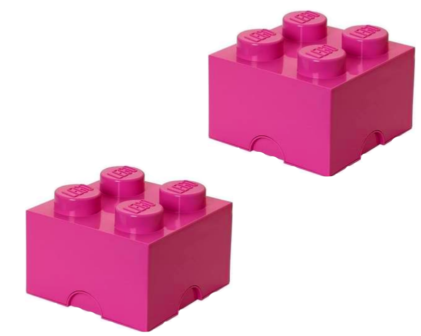
A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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



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

(And some never occur at all.)

start-IP-VP

start-IP-*end*

IP-VP-*end*

CP$_{that}$-IP-NP

IP-NP-PP

IP-VP-CP$_{that}$

IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

NP-PP-*end*

VP-CP$_{that}$-IP

Pearl & Sprouse 2013

Who does 

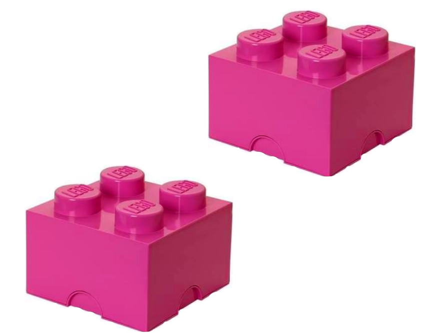
A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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



Relative syntactic trigram frequency:

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

start-IP-VP

start-IP-*end*

IP-VP-*end*

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

IP-VP-IP

IP-VP-PP

VP-PP-*end*

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

VP-IP-VP

NP-PP-*end*

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

IP-NP-PP

Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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

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

start-IP-VP

start-IP-*end*

IP-VP-*end*

CP$_{that}$-IP-NP

IP-NP-PP

IP-VP-CP$_{that}$

IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

NP-PP-*end*

VP-CP$_{that}$-IP

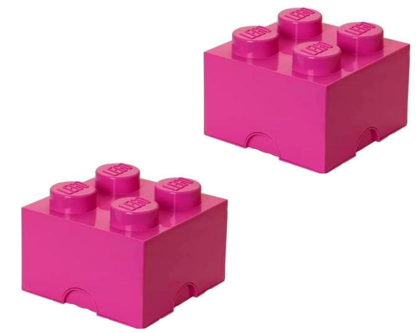
Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-*end*

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

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



start-IP-VP-*end*

start-IP-VP

IP-VP-*end*

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

start-IP-*end*

IP-VP-CP_{that}

IP-VP-IP

VP-CP_{that}-IP

IP-VP-PP

VP-IP-VP

VP-PP-*end*

NP-PP-*end*

CP_{that}-IP-NP

IP-NP-PP

Pearl & Sprouse 2013

Who does

A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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

start-IP-VP

IP-VP-IP

VP-IP-VP

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

IP-VP-*end*

start-IP-*end*

CP_{that}-IP-NP

IP-VP-CP_{that}

IP-VP-PP

VP-PP-*end*

IP-NP-PP

NP-PP-*end*

VP-CP_{that}-IP

Pearl & Sprouse 2013

Who does

A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-*end*

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

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

start-IP-VP

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

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

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

IP-NP-PP

NP-PP-*end*

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

start-IP-*end*

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-IP-VP

Pearl & Sprouse 2013

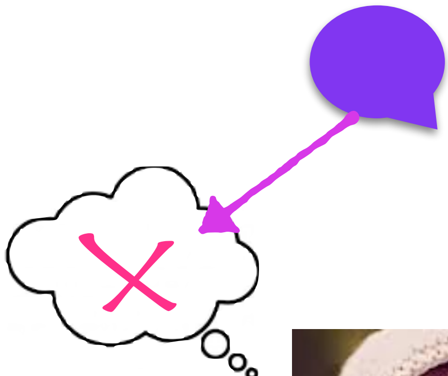
A strategy for learning syntactic islands

start-IP-VP-*end*

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

start-IP-*end*

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



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

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

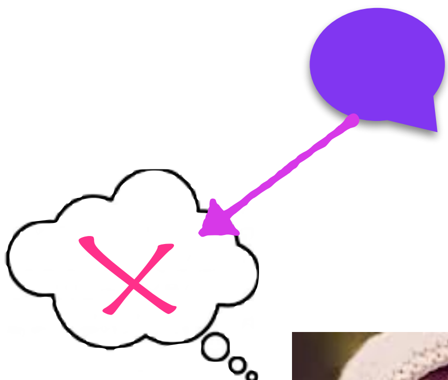
Pearl & Sprouse 2013

Who does 

A strategy for learning syntactic islands

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

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



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

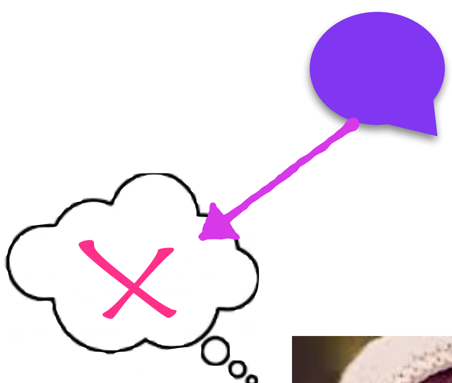
Lower probability dependencies are predicted to be less acceptable (dispreferred), compared to higher probability dependencies.

Pearl & Sprouse 2013

A strategy for learning syntactic islands

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

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



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

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

Complex NP island stimuli

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

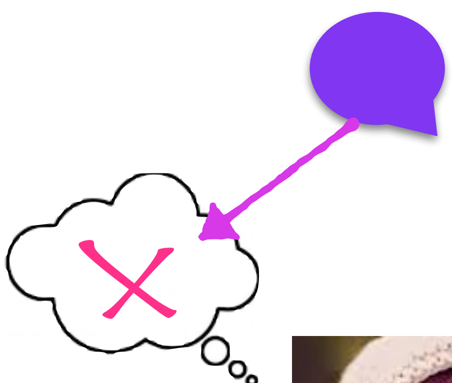
Pearl & Sprouse 2013

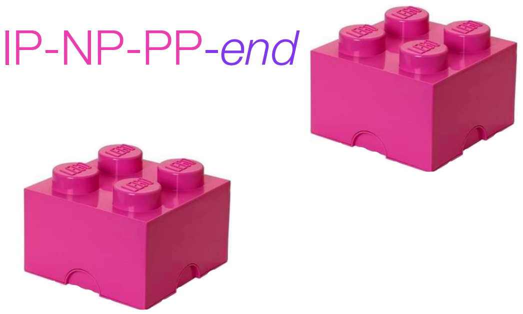
Who does

A strategy for learning syntactic islands

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

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





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

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

- can be transformed into container node sequences

Complex NP island stimuli

start-IP-end	matrix	non-island
start-IP-VP-CP$_{that}$-IP-VP-end	embedded	non-island
start-IP-end	matrix	island
start-IP-VP-NP-CP$_{that}$-IP-VP-end	embedded	island

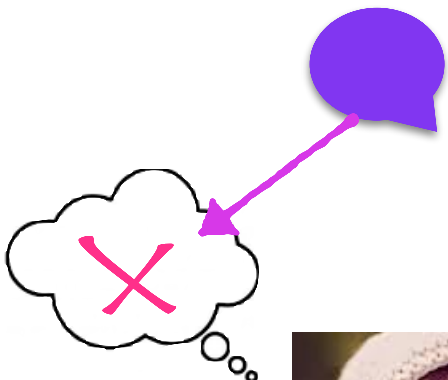
Pearl & Sprouse 2013

Who does

A strategy for learning syntactic islands

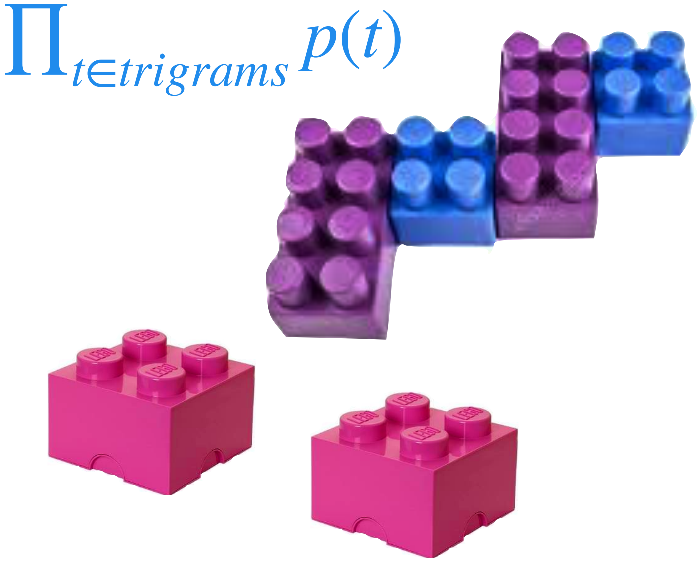
start-IP-VP-*end* *start*-IP-VP-CP$_{that}$-IP-NP-PP-*end*

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



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

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



Complex NP island stimuli

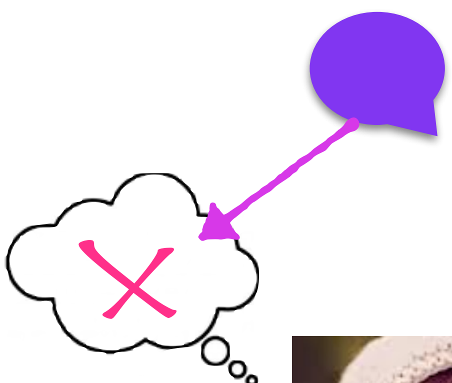
start-IP-end	matrix | non-island
start-IP-VP-CP$_{that}$-IP-VP-end	embedded | non-island
start-IP-end	matrix | island
start-IP-VP-NP-CP$_{that}$-IP-VP-end	embedded | island

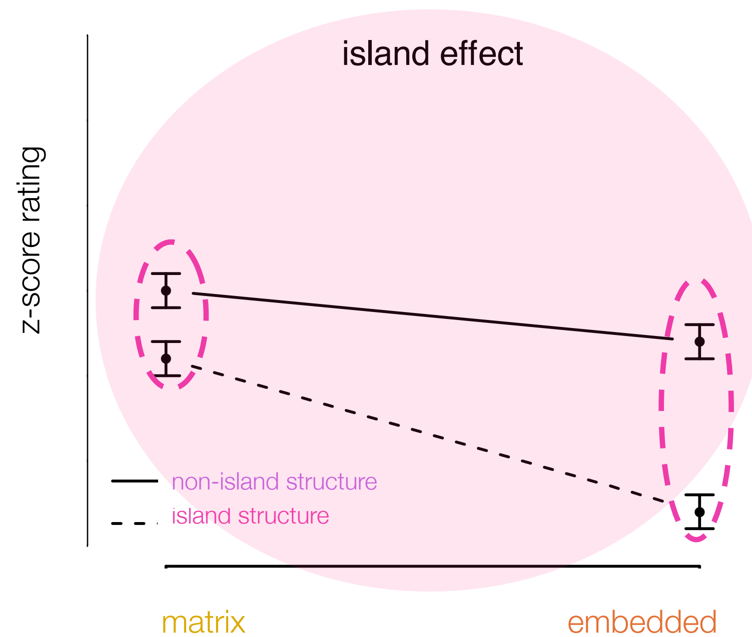
Pearl & Sprouse 2013

A strategy for learning syntactic islands

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

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



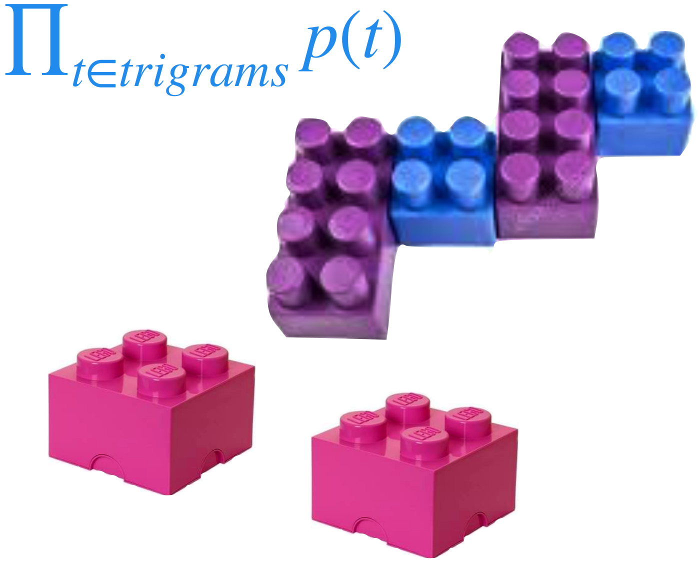


These probabilities can then be plotted to see if superadditivity is present in the predicted acceptability judgments.

Complex NP island stimuli

start-IP-end	matrix | non-island
start-IP-VP-CP$_{that}$-IP-VP-end	embedded | non-island
start-IP-end	matrix | island
start-IP-VP-NP-CP$_{that}$-IP-VP-end	embedded | island

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



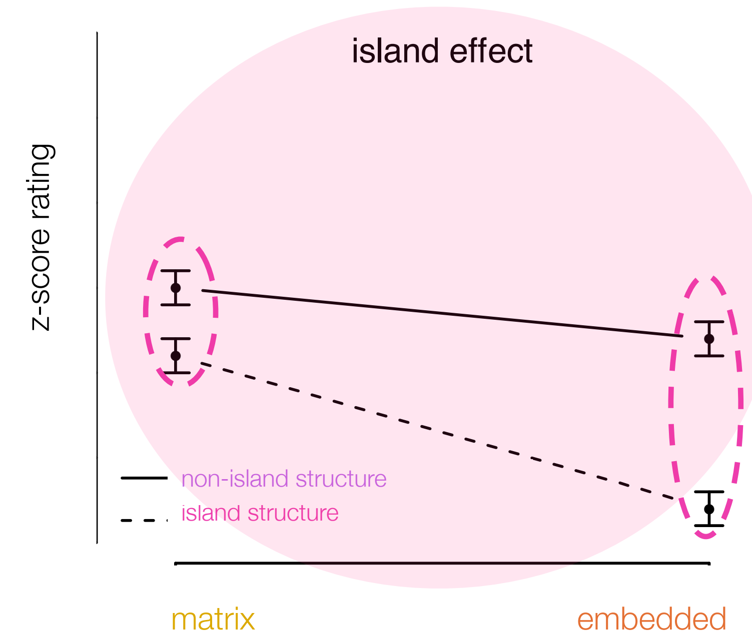
Pearl & Sprouse 2013

A strategy for learning syntactic islands

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

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

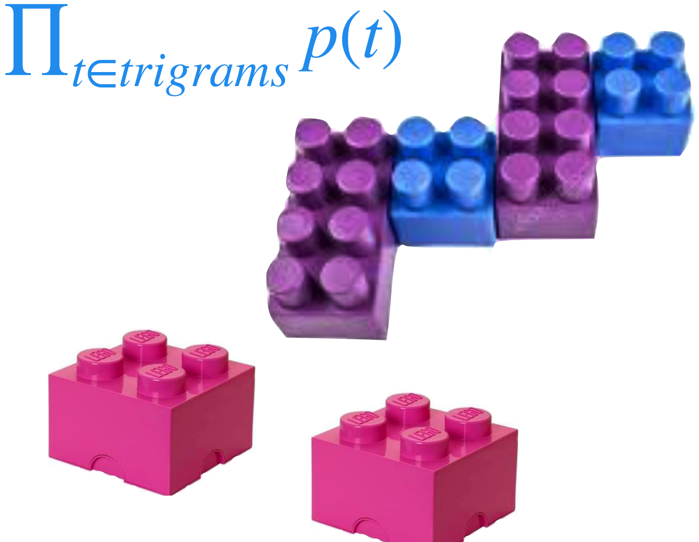
If so, then we predict the child to have syntactic island knowledge that allows the same judgment pattern as adults, learned from the building blocks in children's input.



Complex NP island stimuli

start-IP-end	matrix | non-island
start-IP-VP-CP$_{that}$-IP-VP-end	embedded | non-island
start-IP-end	matrix | island
start-IP-VP-NP-CP$_{that}$-IP-VP-end	embedded | island

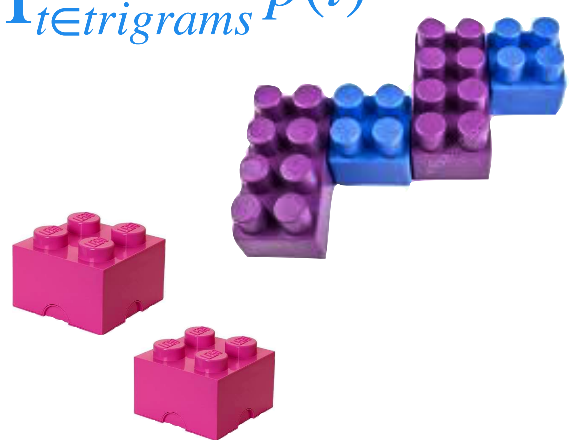
$$\prod_{t \in trigrams} p(t)$$

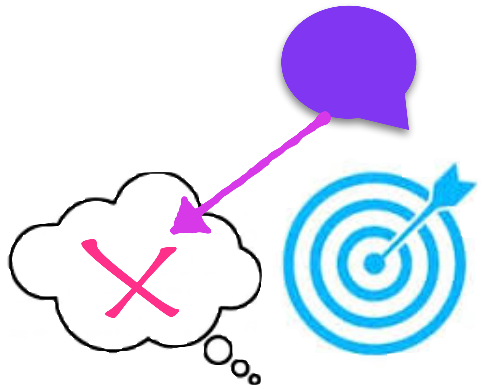


Pearl & Sprouse 2013

Who does

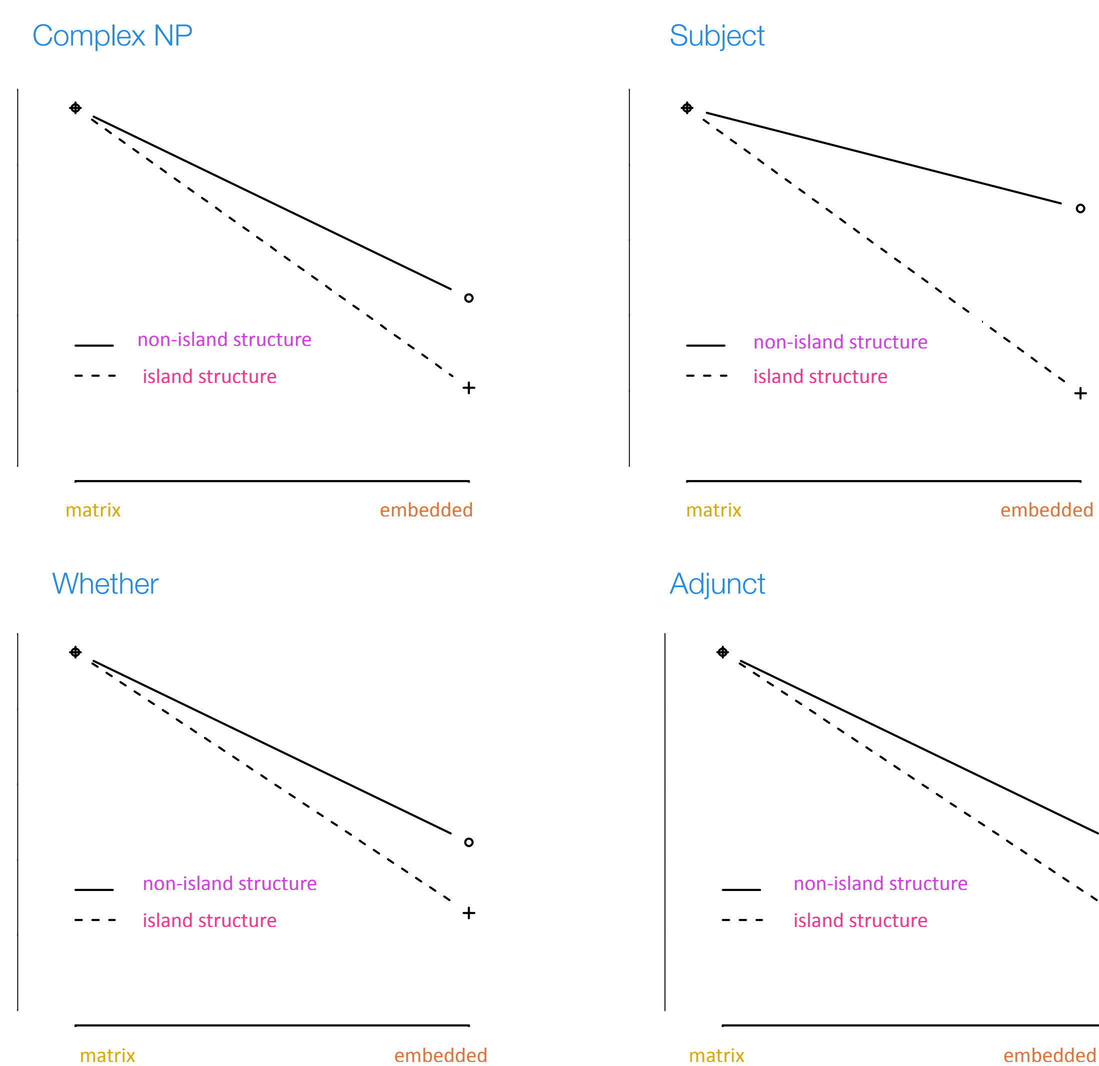
$\prod_{t \in trigrams} p(t)$





This strategy works for high-SES children's input

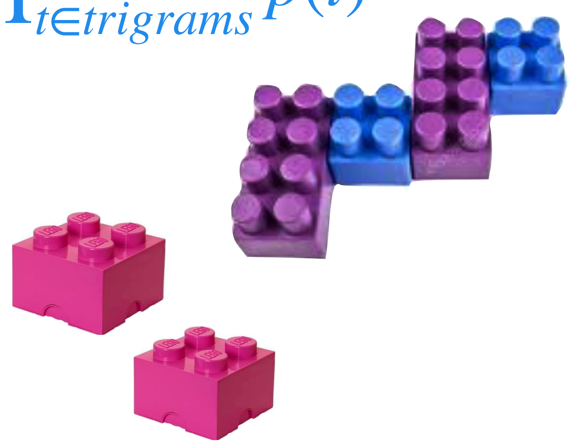
Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data (Pearl & Sprouse 2013).



Pearl & Sprouse 2013

This strategy works for high-SES children's input

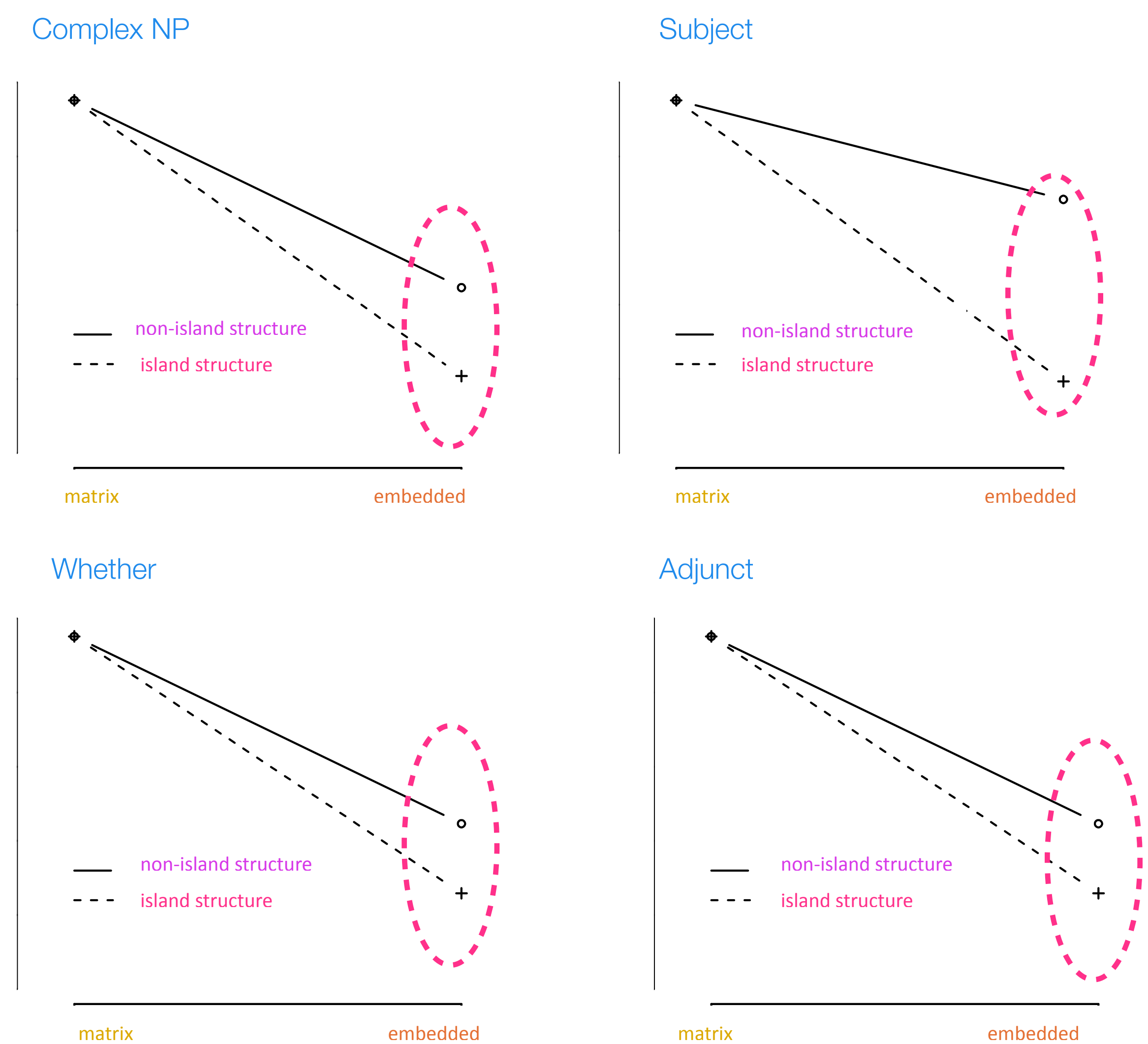
$\prod_{t \in trigrams} p(t)$



Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data (Pearl & Sprouse 2013).

Superadditivity for all four islands.

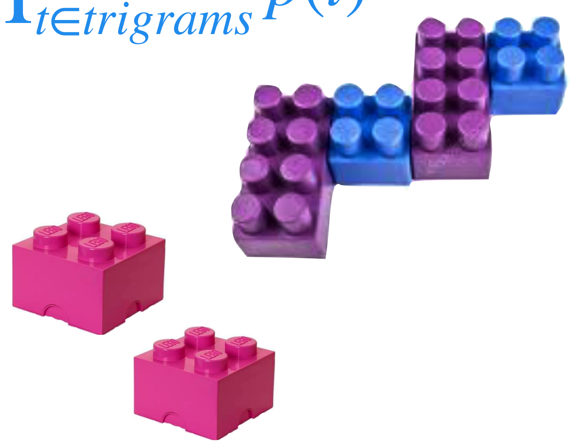
✓

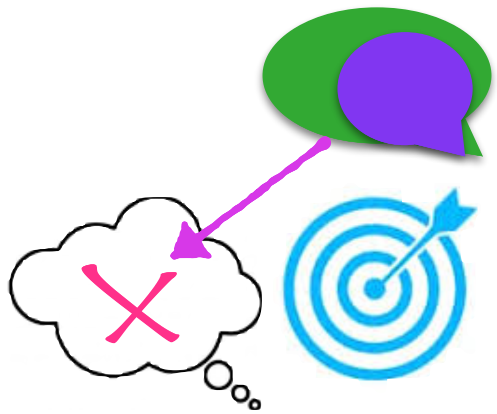


Pearl & Sprouse 2013

This strategy works for high-SES children's input

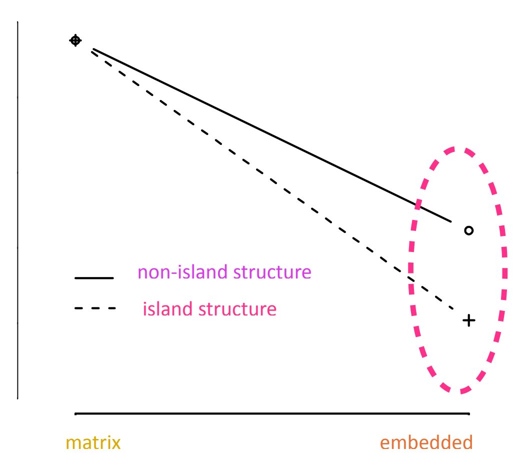
$\prod_{t \in trigrams} p(t)$



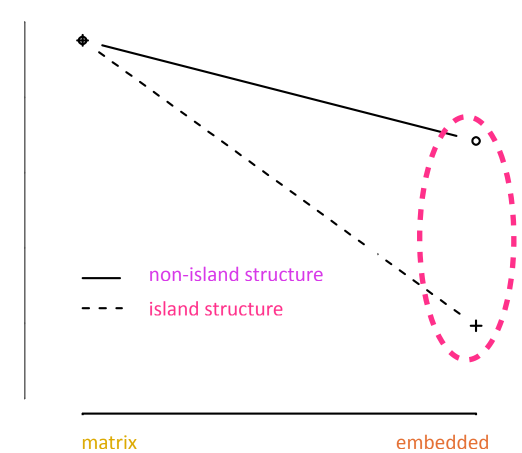


Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data (Pearl & Sprouse 2013).

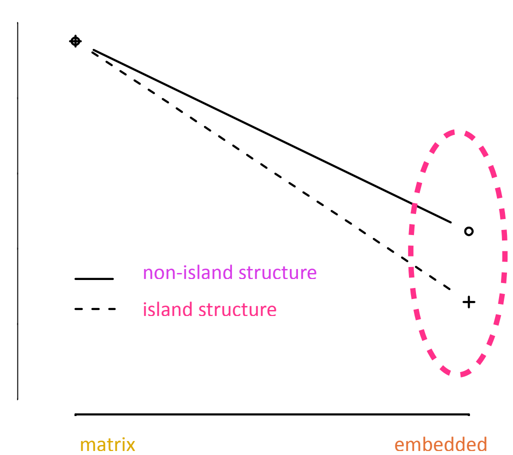
Complex NP



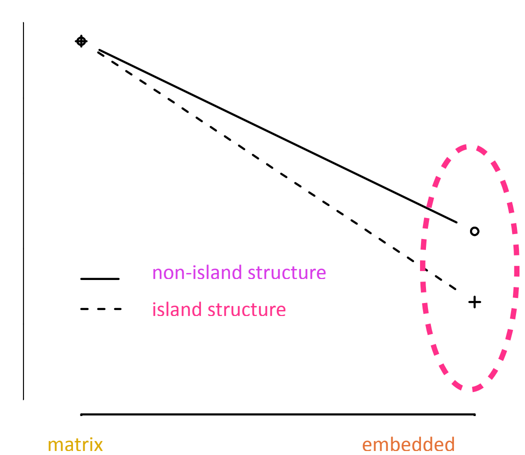
Subject



Whether



Adjunct

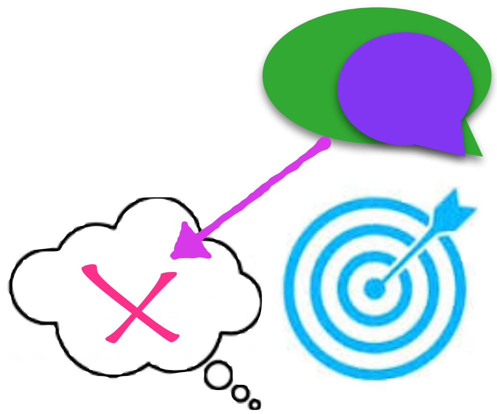


Implication:
High-SES child input
can support the acquisition of
syntactic islands,
using this learning strategy
that depends on
a certain part of the input.

Pearl & Sprouse 2013

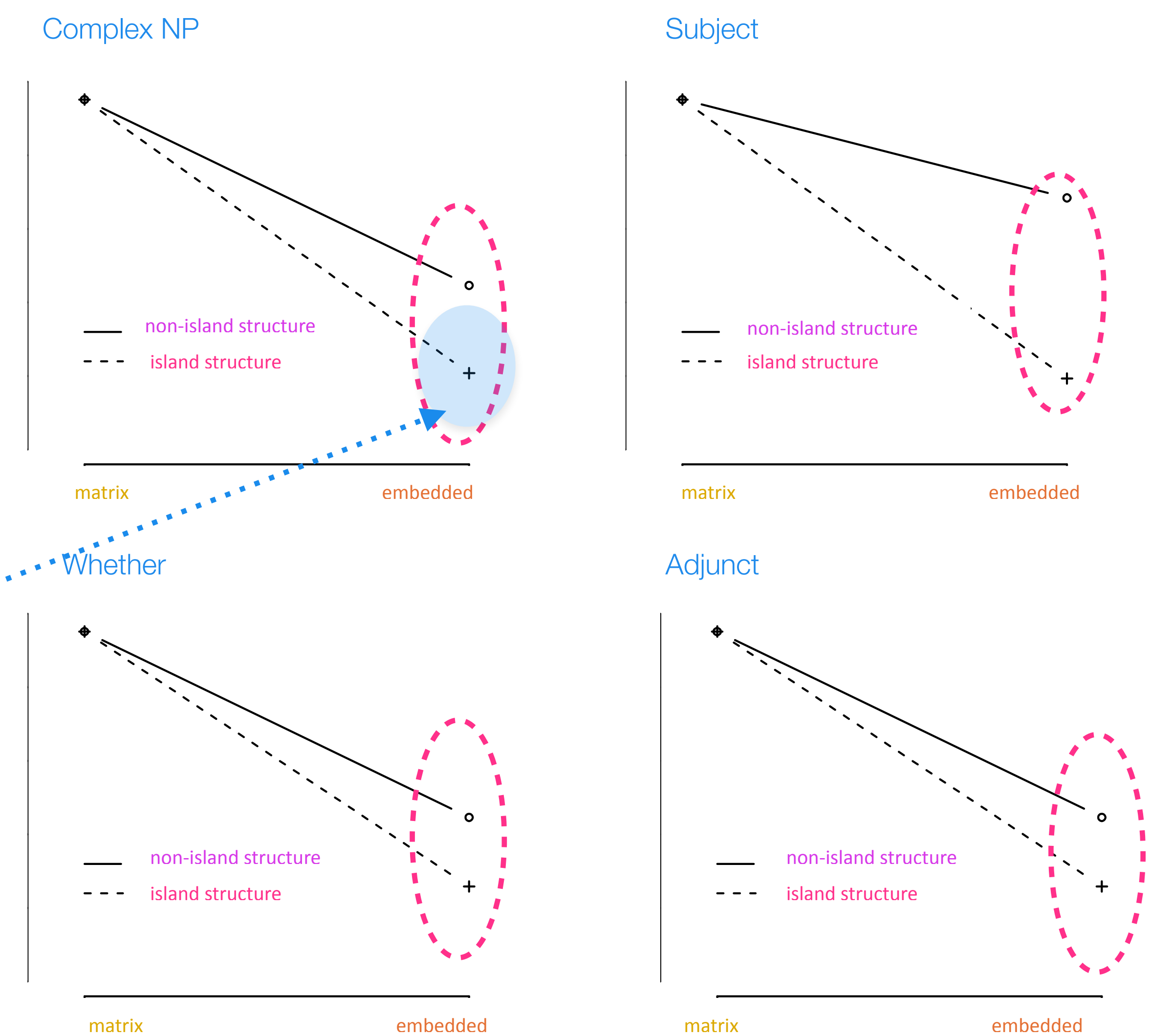
This strategy works for high-SES children's input

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



Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data (Pearl & Sprouse 2013).

Implication:
High-SES child input
can support the acquisition of
syntactic islands,
using this learning strategy
that depends on
a certain part of the input.



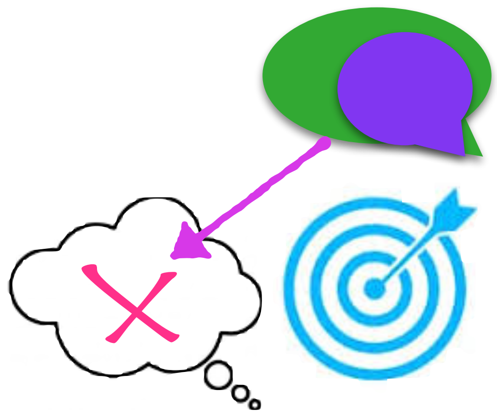
In line with high-SES 4-year-olds who disprefer the Complex NP island-crossing dependency (De Villiers et al. 2008).

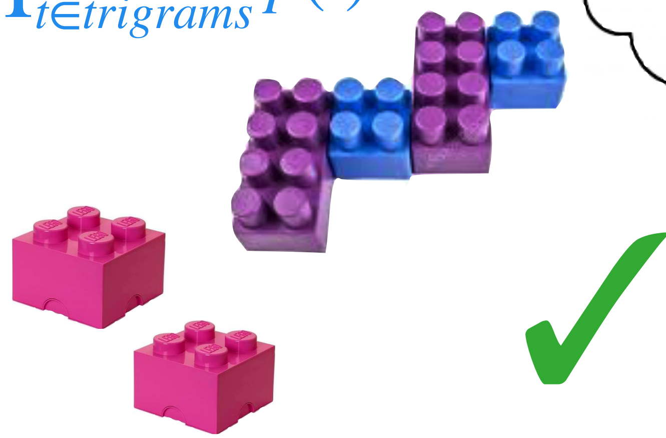
Pearl & Sprouse 2013

This strategy works for high-SES children's input

Who does

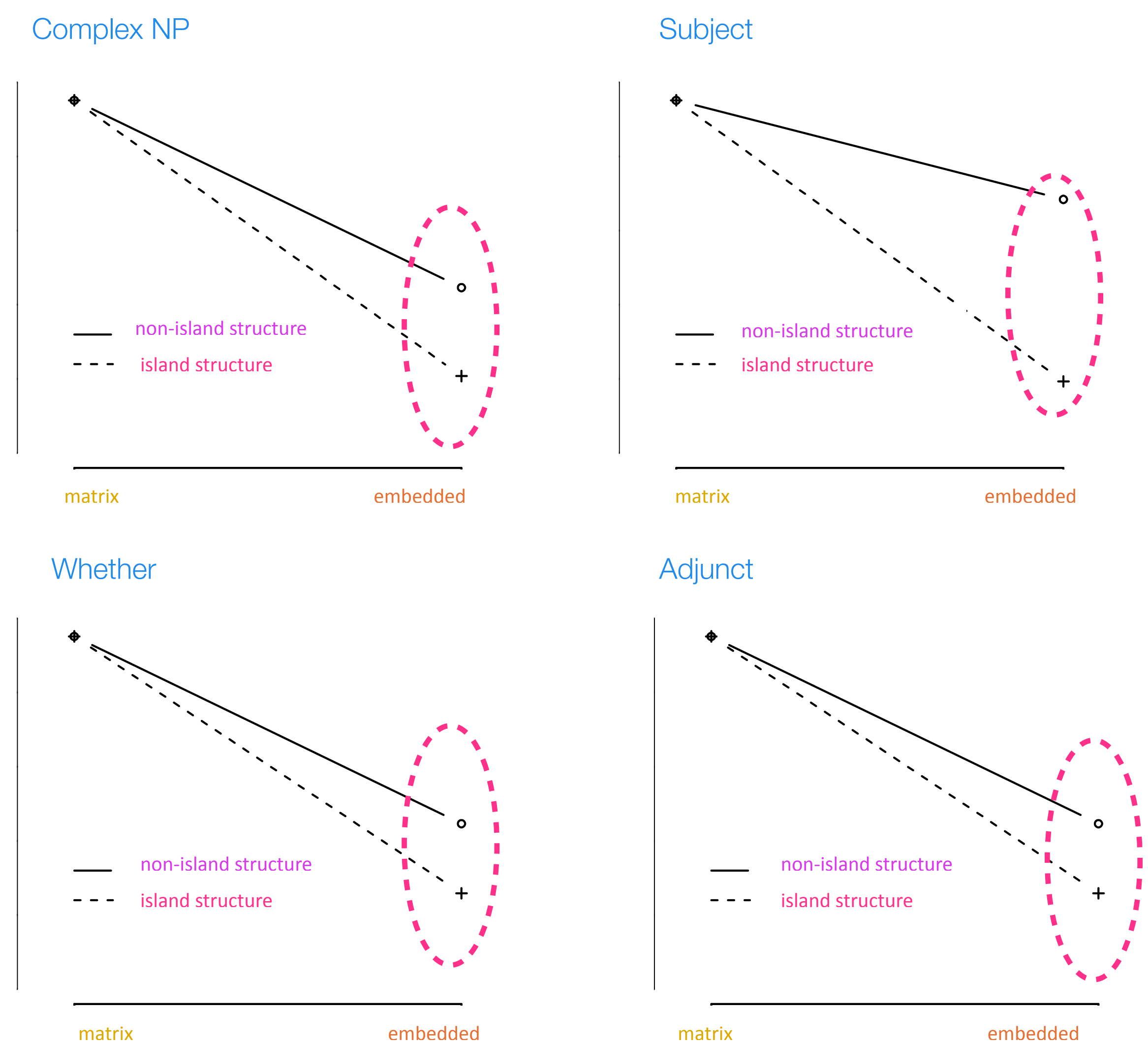
$\prod_{t \in trigrams} p(t)$





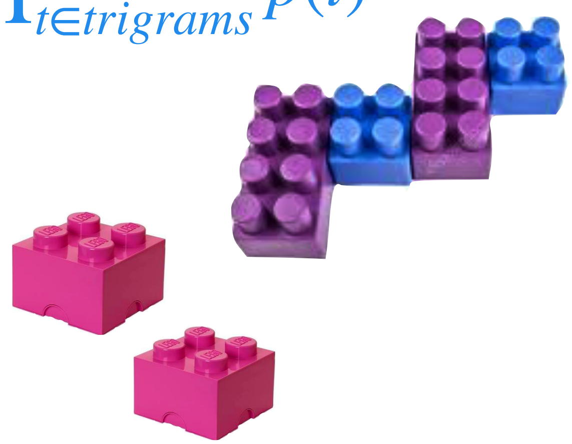
Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data (Pearl & Sprouse 2013).

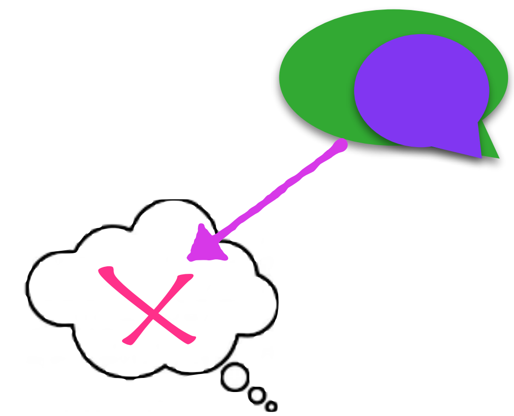
That input part is the *wh*-dependencies, and their building blocks (the syntactic trigrams).



Pearl & Sprouse 2013

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

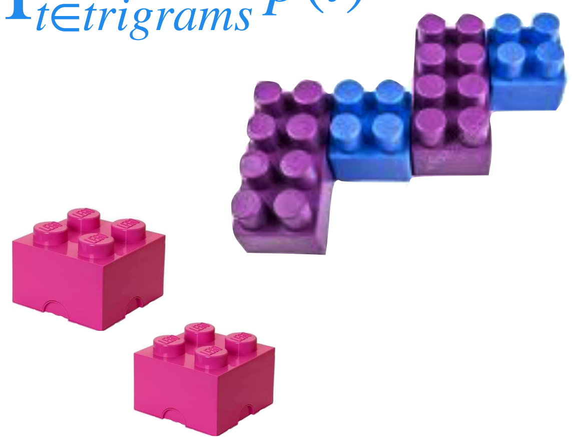


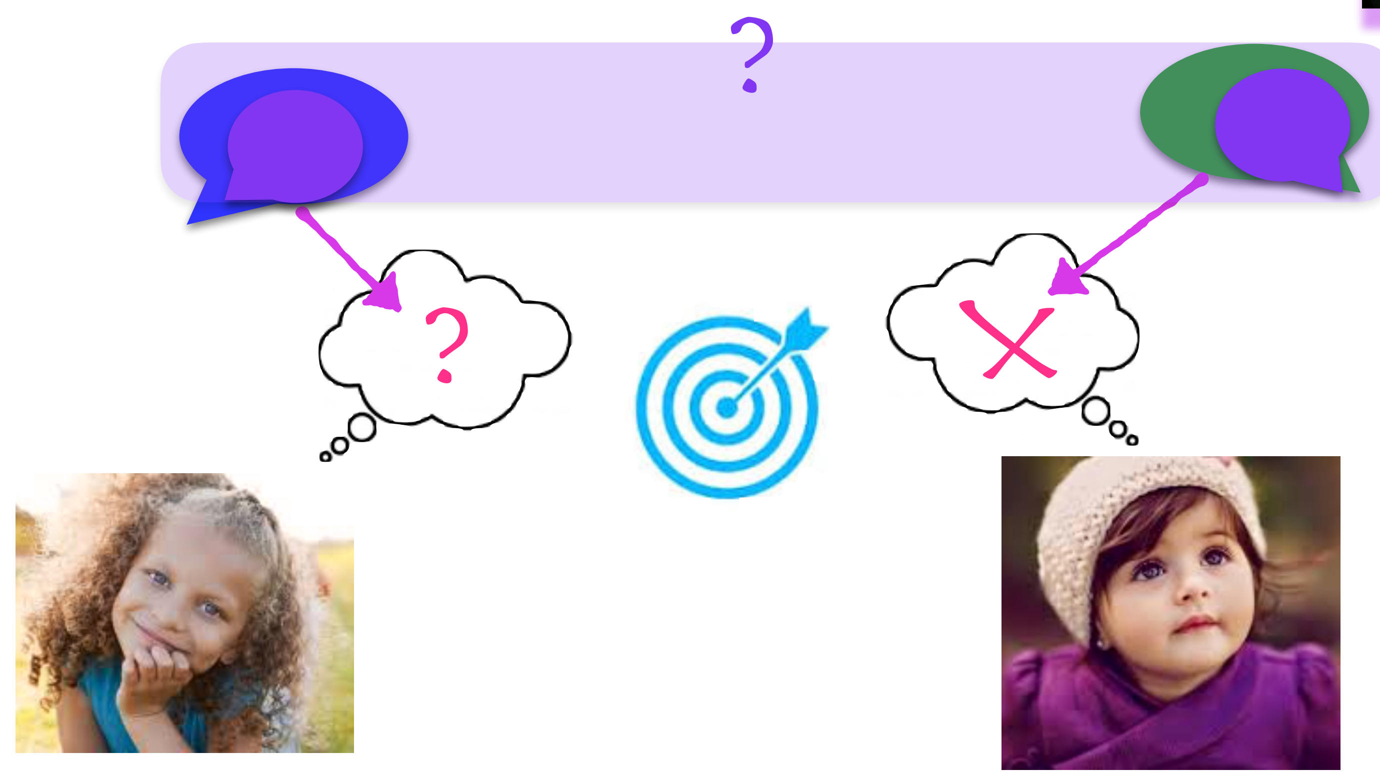


Are there meaningful differences across SES in this part of the input (the *wh*-dependencies and syntactic trigrams)?

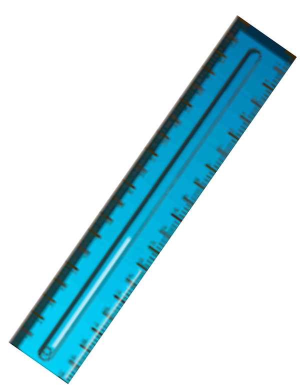
Bates & Pearl 2019, submitted

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





First, what are the differences? That is, how different does this input part look across SES?



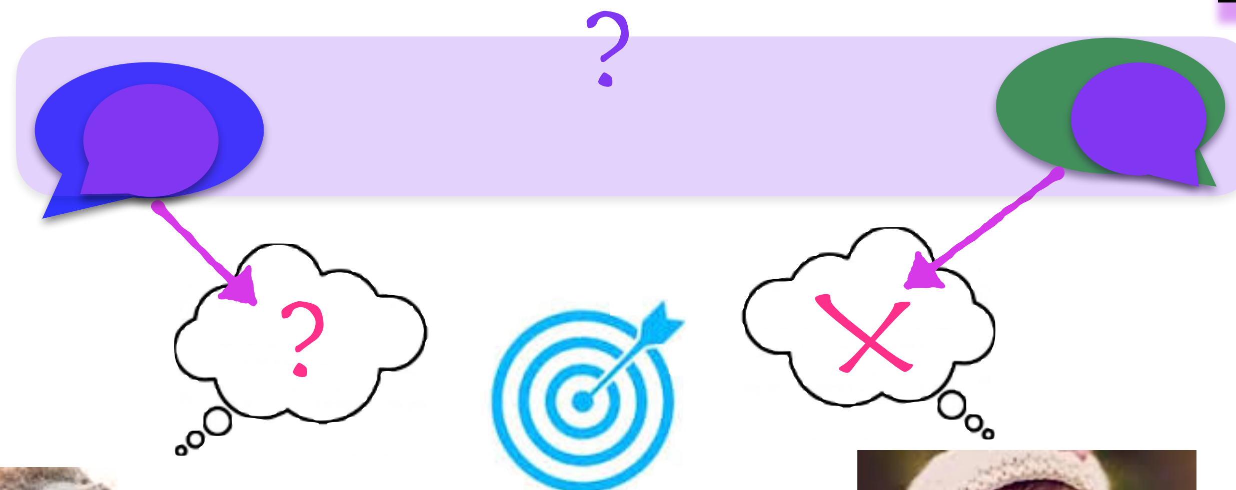
Let's measure the distribution of the relevant parts: the *wh*-dependencies and the syntactic trigrams.

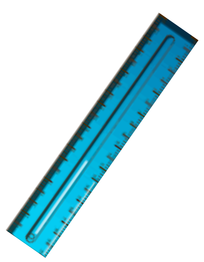
Bates & Pearl 2019, submitted

Who does

Measurable input differences

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





One way to measure differences in distribution:
the Jensen-Shannon divergence (JSDiv) (Endres & Schindelin 2003).

$$0 \leq JSDiv \leq 1$$

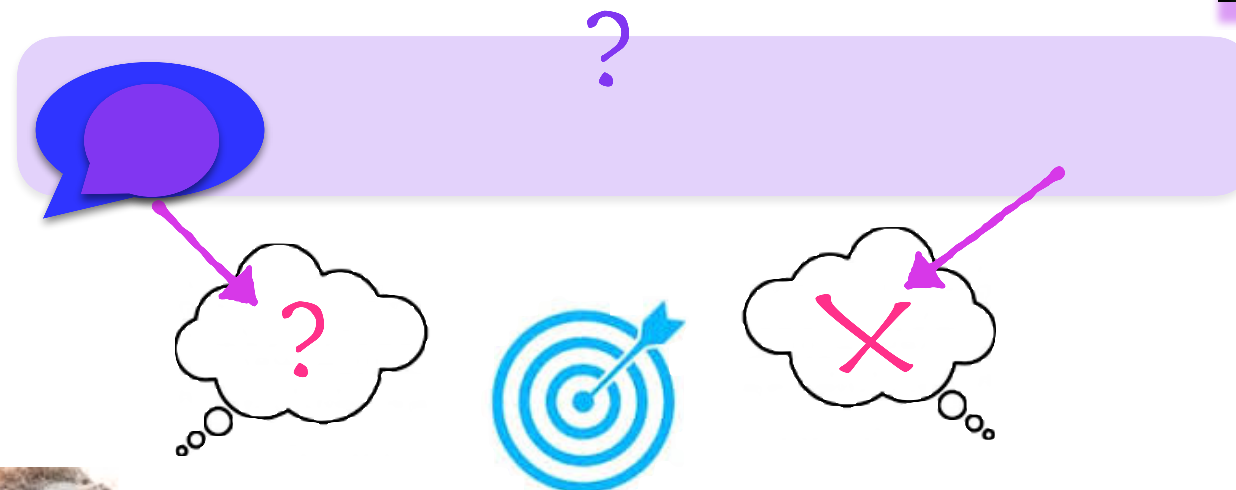
identical distributions | dissimilar distributions

$=$ | $\neq$

Bates & Pearl 2019, submitted

Measurable input differences

Who does



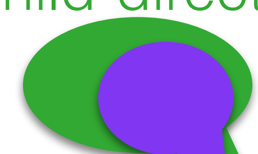
$\prod_{t \in trigrams} p(t)$

$=$ $\neq$

$0 \leq JSDiv \leq 1$

The input samples

High-SES
child-directed

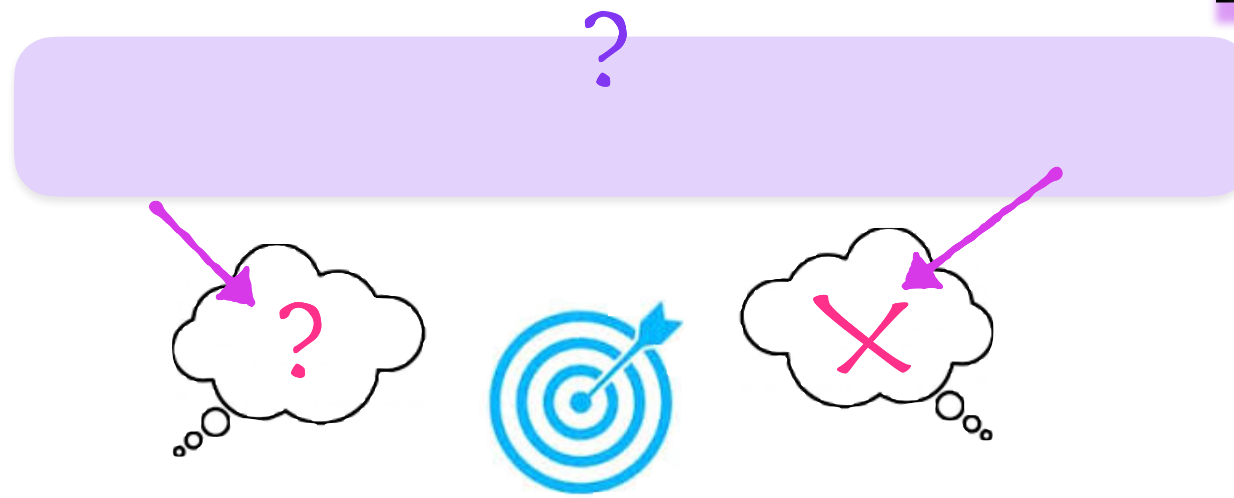


102K utterances (21K *wh*-dependencies) from the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 25 high-SES children between the ages of 1 and 5 years old.

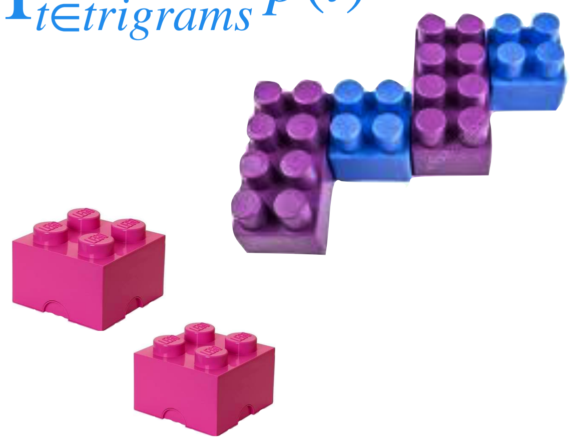
Bates & Pearl 2019, submitted

Measurable input differences

Who does



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

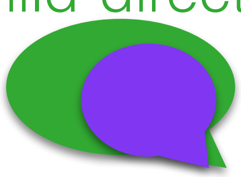


$=$ $\neq$

$0 \leq JSDiv \leq 1$

The input samples

High-SES
child-directed



21K *wh*-dependencies

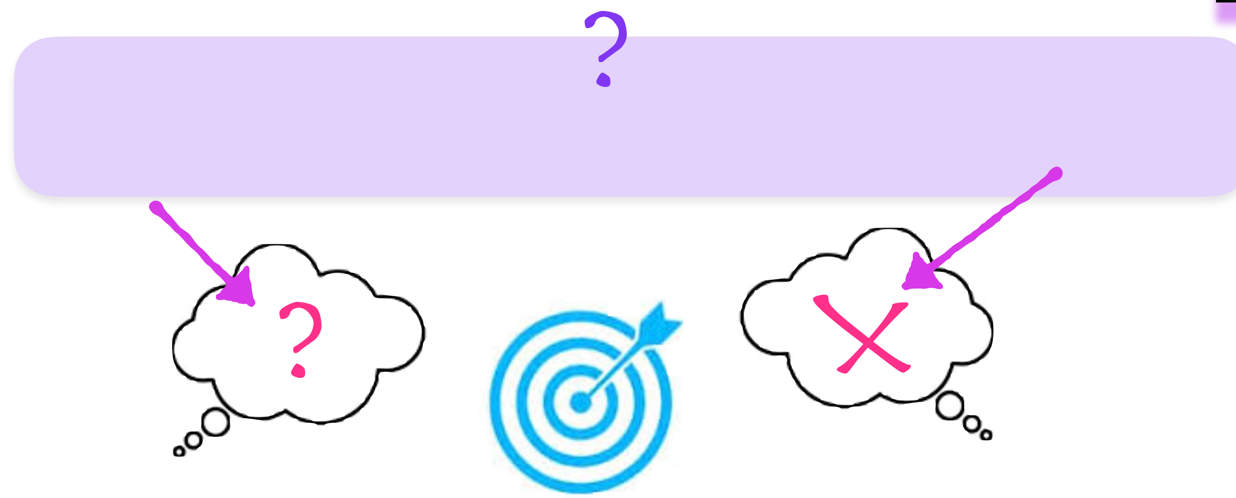
Low-SES
child-directed

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

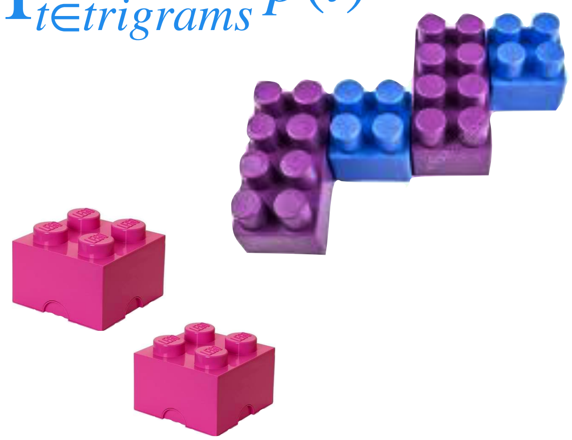
Bates & Pearl 2019, submitted

Measurable input differences

Who does



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

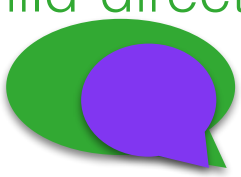


$=$ $\neq$

$$0 \leq JSDiv \leq 1$$

The input samples

High-SES child-directed



21K *wh*-dependencies

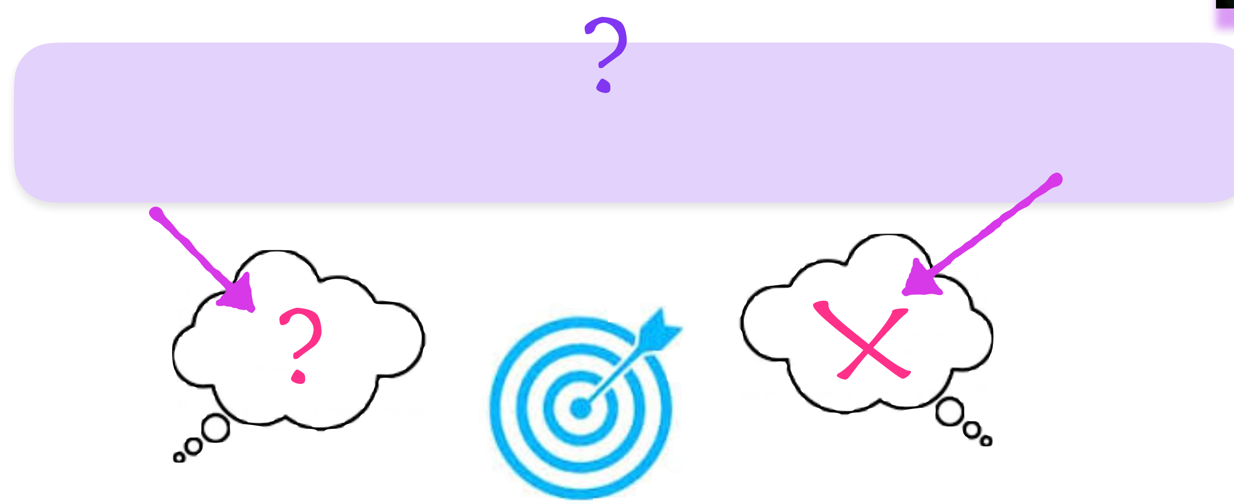
3.9K *wh*-dependencies

Low-SES child-directed

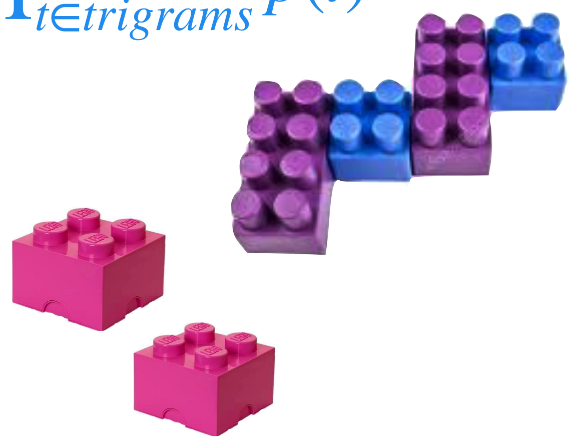
Note: SES was defined by the creators of the HSLLD corpus according to maternal education (6 years to some post-high school education) and annual income (70% reported < $20K/year).

Bates & Pearl 2019, submitted

Measurable input differences



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



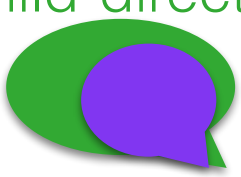
74.6K utterances (8.5K *wh*-dependencies) from the Switchboard corpus (Marcus et al. 1999) of adults speaking to each other over the phone.

$=$ $\neq$

$0 \leq JSDiv \leq 1$

The input samples

High-SES child-directed

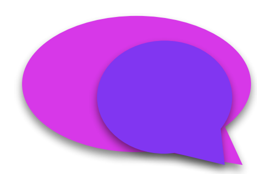


21K *wh*-dependencies

Low-SES child-directed

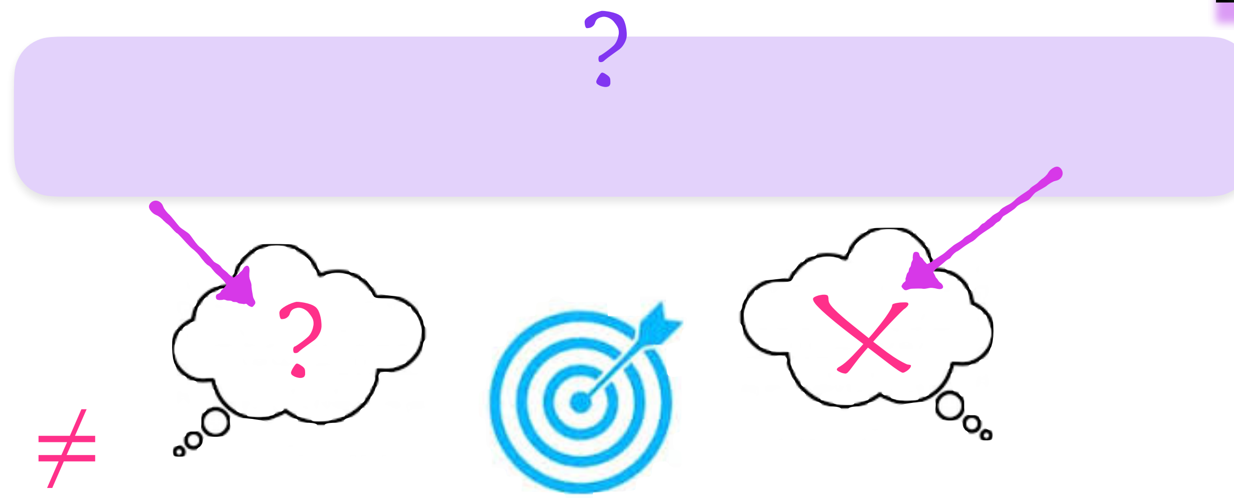
3.9K *wh*-dependencies

High-SES adult-directed

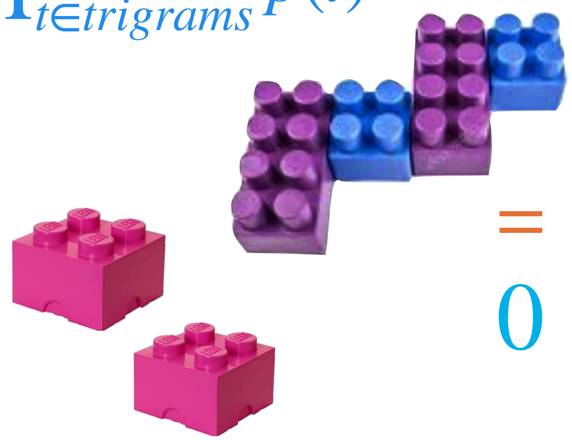


Bates & Pearl 2019, submitted

Measurable input differences



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



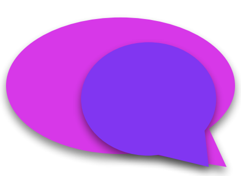
$=$ $\neq$

$$0 \leq JSDiv \leq 1$$

So what do we find?

In particular, is high-SES child-directed speech more like low-SES child-directed speech or more like high-SES adult-directed speech?

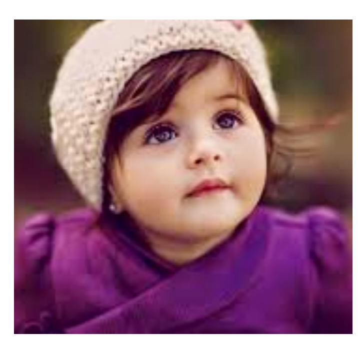
High-SES adult-directed



8.5K *wh*-dependencies

High-SES child-directed

21K *wh*-dependencies



Low-SES child-directed

3.9K *wh*-dependencies

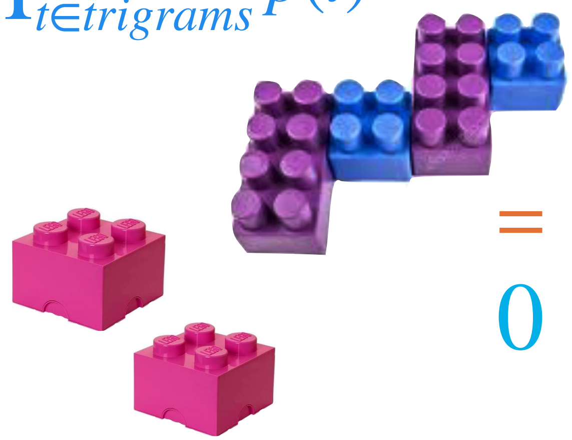
Bates & Pearl 2019, submitted

Measurable input differences

Who does

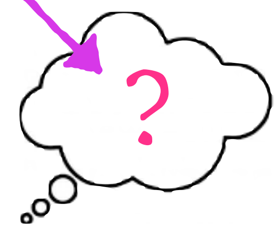
?

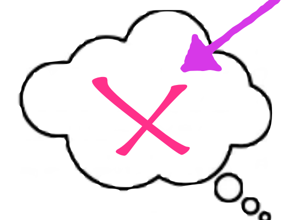
$\prod_{t \in trigrams} p(t)$



=

$\neq$

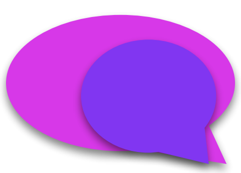




$0 \leq JSDiv \leq 1$

If high-SES child-directed speech is more like low-SES child-directed speech, then SES differences matter less than who the speech is directed at.

High-SES adult-directed

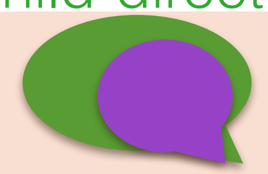


8.5K *wh*-dependencies

High-SES child-directed

SES differences

Low-SES child-directed



21K *wh*-dependencies

3.9K *wh*-dependencies

Bates & Pearl 2019, submitted

Measurable input differences

Who does 

?

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

=

≠

$0 \le JSDiv \le 1$

If high-SES child-directed speech is more like high-SES adult-directed speech, then SES differences matter more than who the speech is directed at.

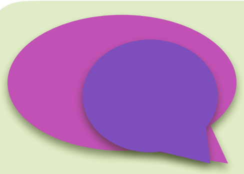
High-SES adult-directed

directed at who differences

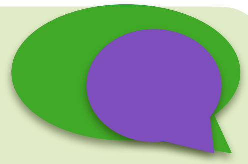
High-SES child-directed

SES differences

Low-SES child-directed

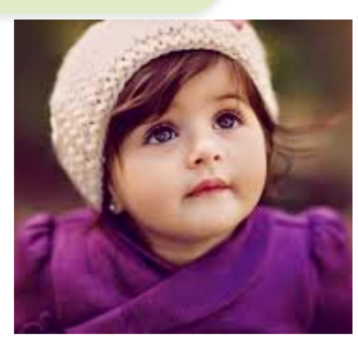


8.5K *wh*-dependencies



21K *wh*-dependencies

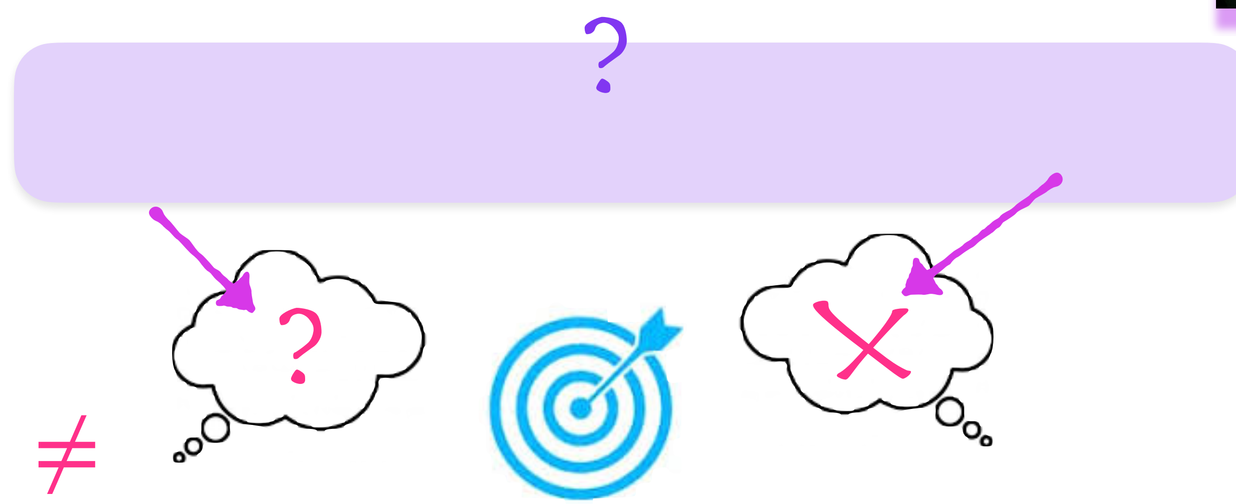
3.9K *wh*-dependencies



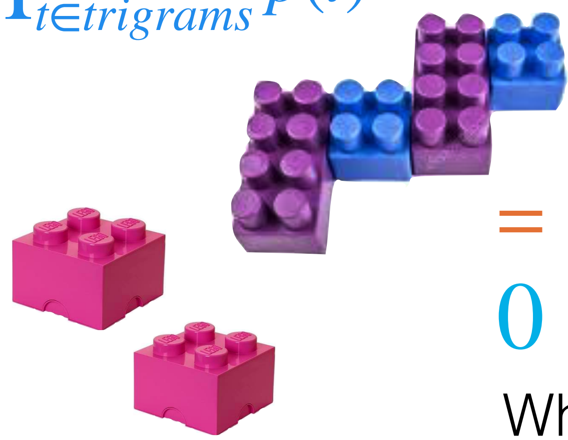
Bates & Pearl 2019, submitted

Who does

Measurable input differences



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



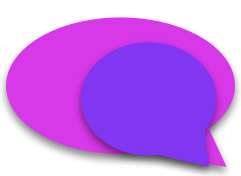
=

$\neq$

$0 \leq JSDiv \leq 1$

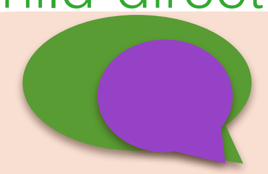
Whether we look at *wh*-dependencies or syntactic trigrams, we find the same pattern: high-SES and low-SES child-directed speech are more similar than high-SES child-directed and high-SES adult-directed speech.

High-SES
adult-directed



8.5K *wh*-dependencies

High-SES
child-directed



21K *wh*-dependencies

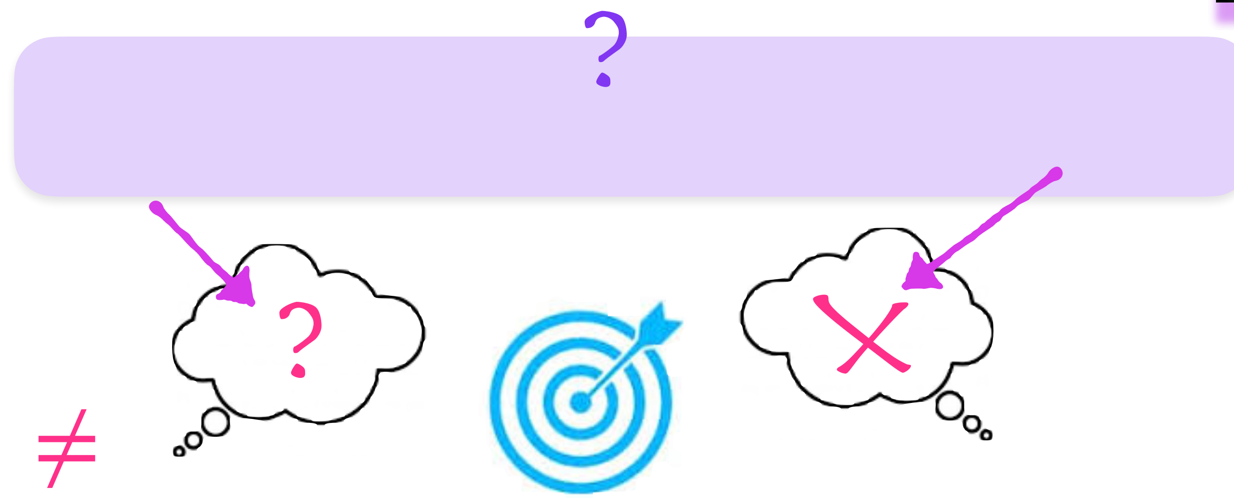
Low-SES
child-directed

3.9K *wh*-dependencies

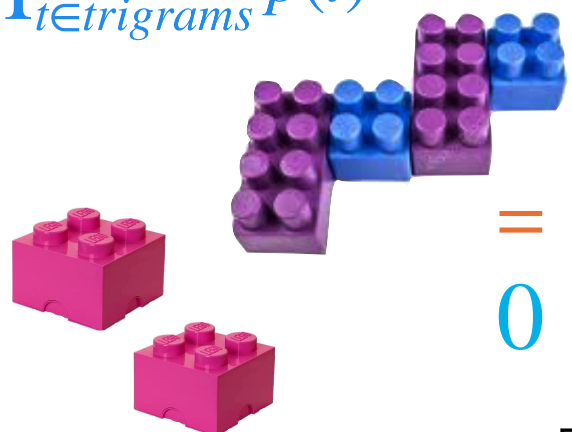
Bates & Pearl 2019, submitted

Who does

Measurable input differences



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



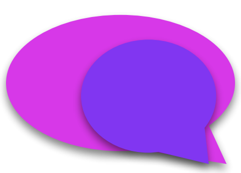
$=$ $\neq$

$$0 \leq JSDiv \leq 1$$

For *wh*-dependencies, high-SES child-directed speech is twice as similar to low-SES child-directed speech as it is to high-SES adult-directed speech.

High-SES adult-directed

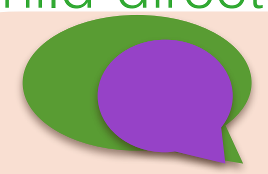
.00948



8.5K *wh*-dependencies

High-SES child-directed

.00445



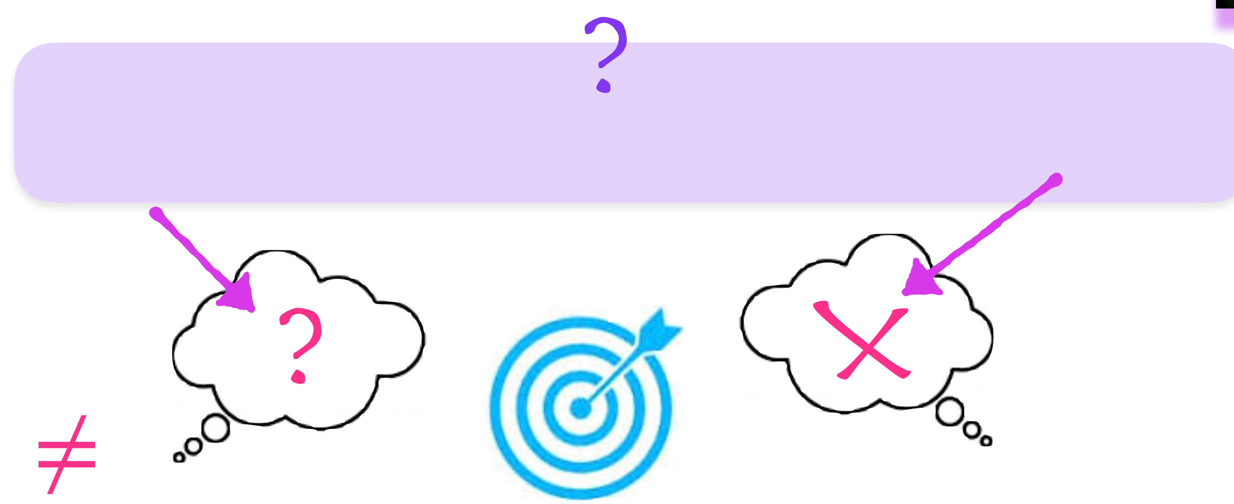
21K *wh*-dependencies

Low-SES child-directed

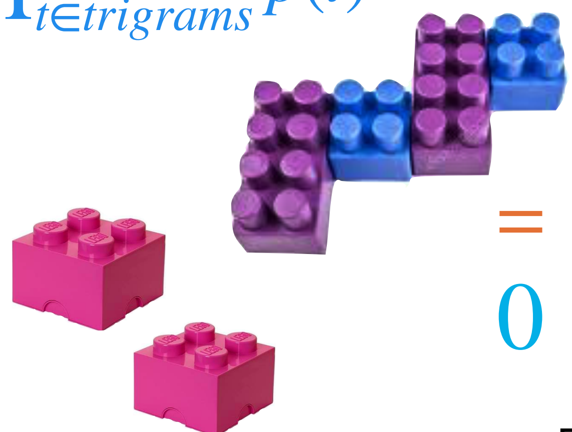
3.9K *wh*-dependencies

Bates & Pearl 2019, submitted

Measurable input differences



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



=

≠

$$0 \leq JSDiv \leq 1$$

For syntactic trigrams, high-SES child-directed speech is twice as similar to low-SES child-directed speech as it is to high-SES adult-directed speech.

High-SES adult-directed

.01825

High-SES child-directed

.00850

Low-SES child-directed

8.5K *wh*-dependencies

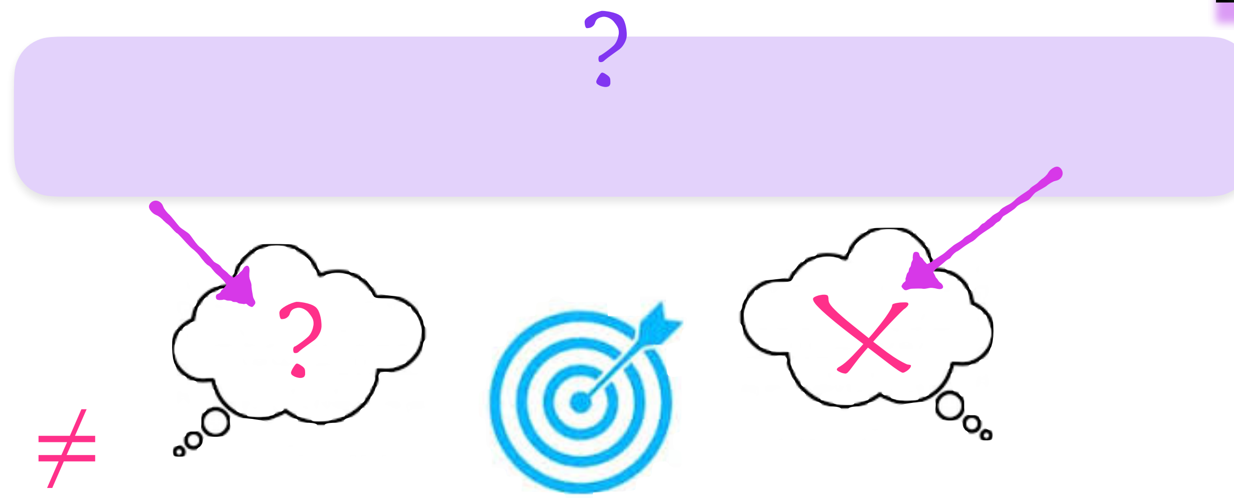
21K *wh*-dependencies

3.9K *wh*-dependencies

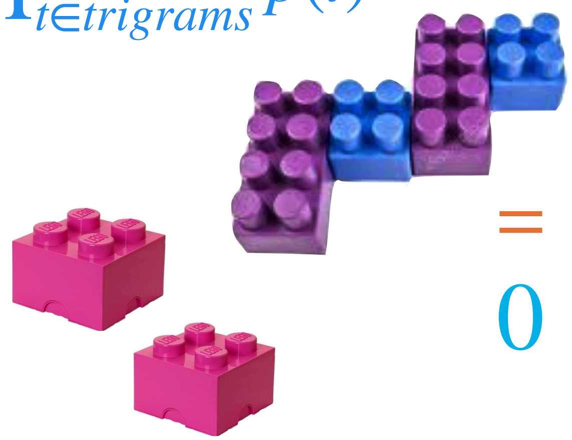
Bates & Pearl 2019, submitted

Measurable input differences

Who does 



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



=

$\neq$

$0 \leq JSDiv \leq 1$

Takeaway: This part of the input looks pretty similar across SES — more similar than child-directed vs. adult-directed speech within SES.

High-SES adult-directed

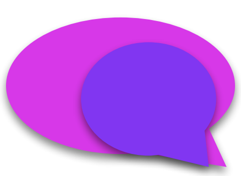
directed at who differences

High-SES child-directed

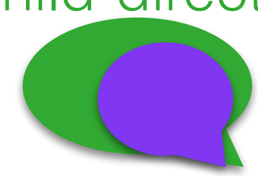
SES differences

Low-SES child-directed

$\approx$

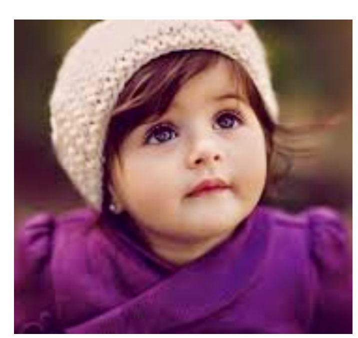


8.5K *wh*-dependencies



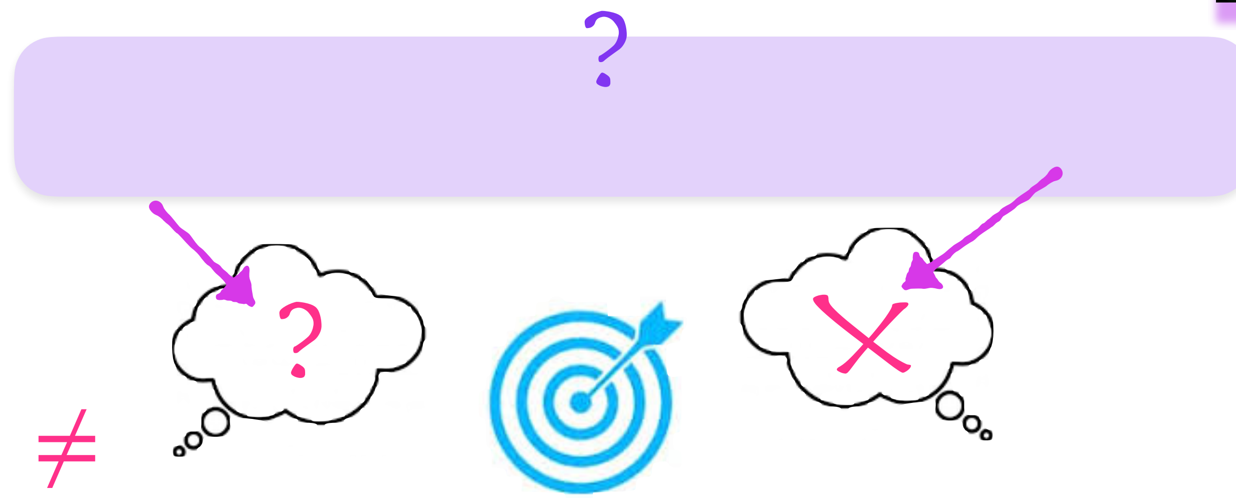
21K *wh*-dependencies

3.9K *wh*-dependencies

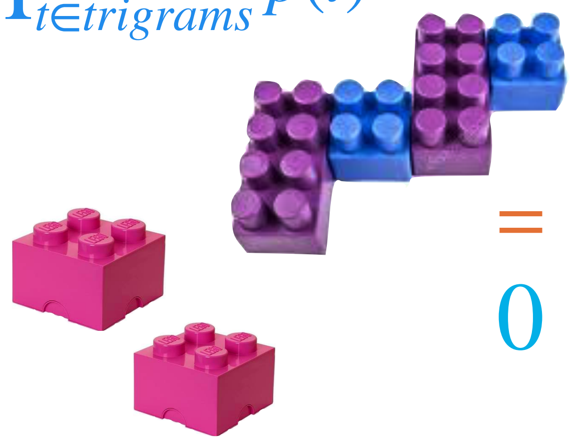


Bates & Pearl 2019, submitted

Measurable input differences



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



$=$ $\neq$

$$0 \leq JSDiv \leq 1$$

But this is just a (quantitative) way to describe the input differences…

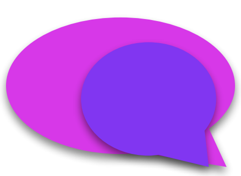
High-SES adult-directed

directed at who differences

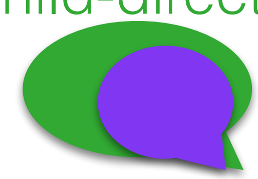
High-SES child-directed

SES differences

Low-SES child-directed



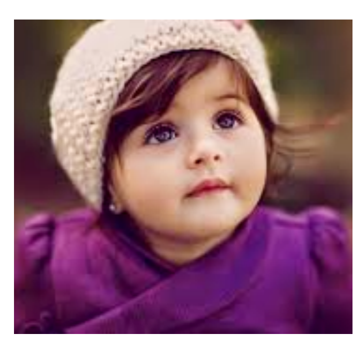
8.5K *wh*-dependencies



$\approx$

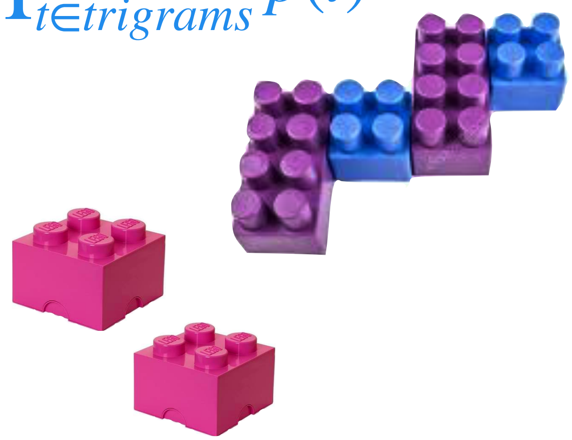
21K *wh*-dependencies

3.9K *wh*-dependencies

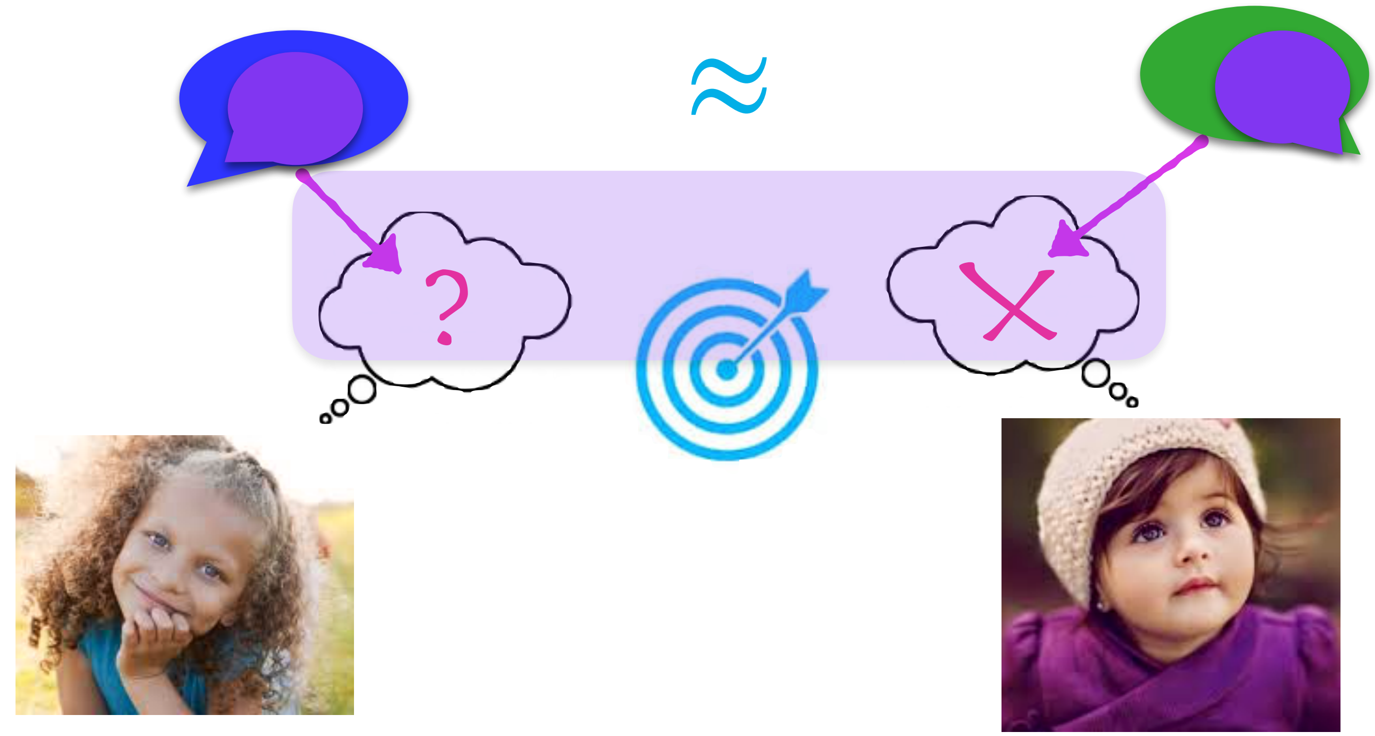


Bates & Pearl 2019, submitted

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

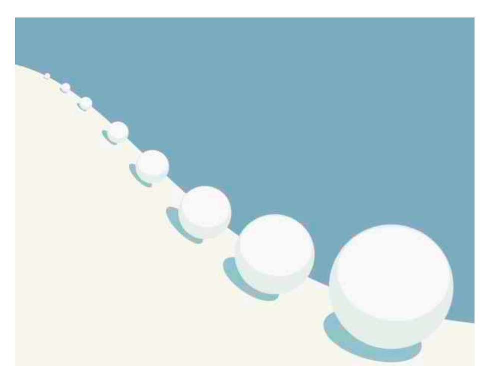


Meaningful input differences



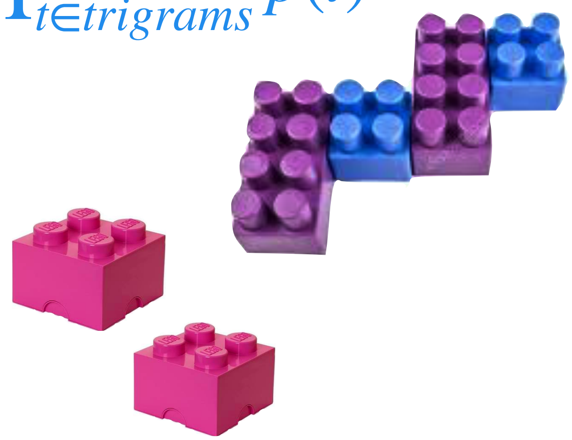
Does this part of the input act differently across SES? That is, are any differences (even if they're smaller) developmentally meaningful?

They might be — small differences in the input distribution might snowball into learning outcome differences.

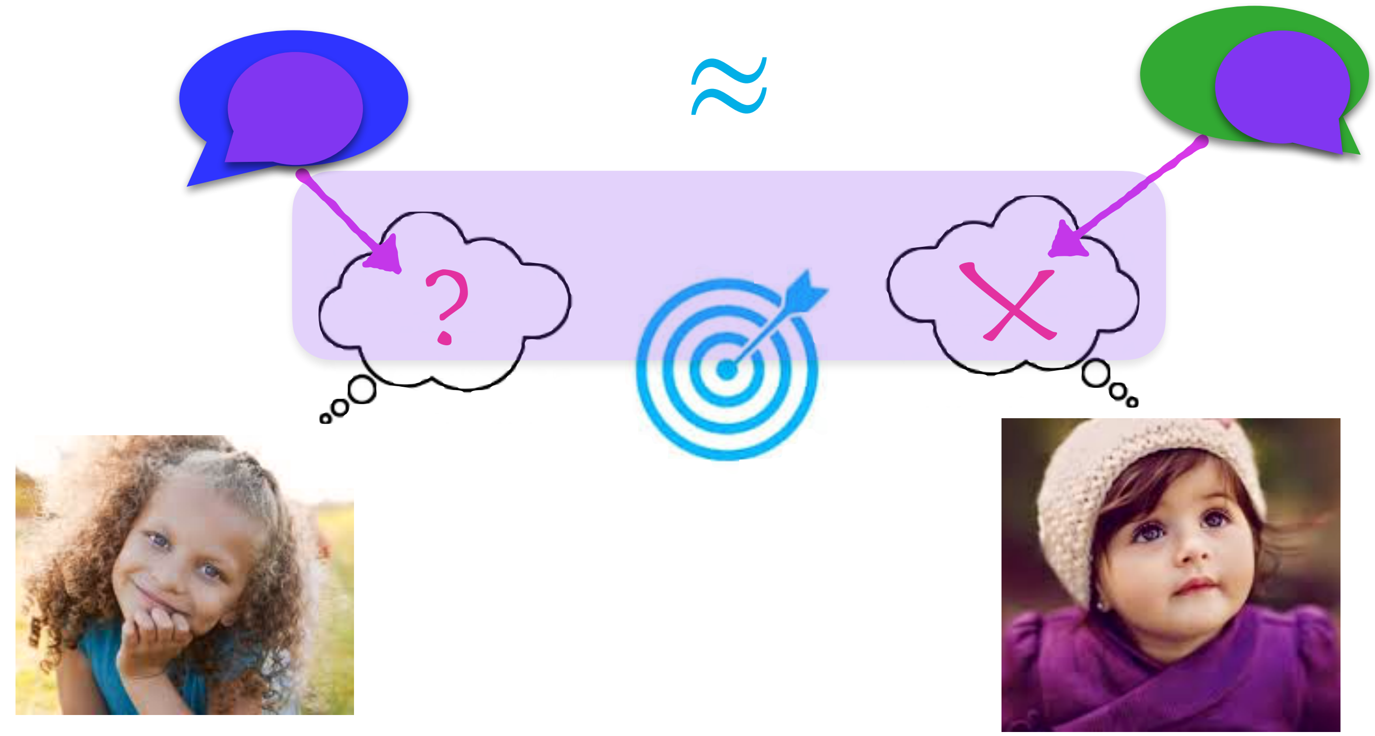


Bates & Pearl 2019, submitted

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



Meaningful input differences



Does this part of the input act differently across SES? That is, are any differences (even if they're smaller) developmentally meaningful?

They might be — small differences in the input distribution might snowball into learning outcome differences.

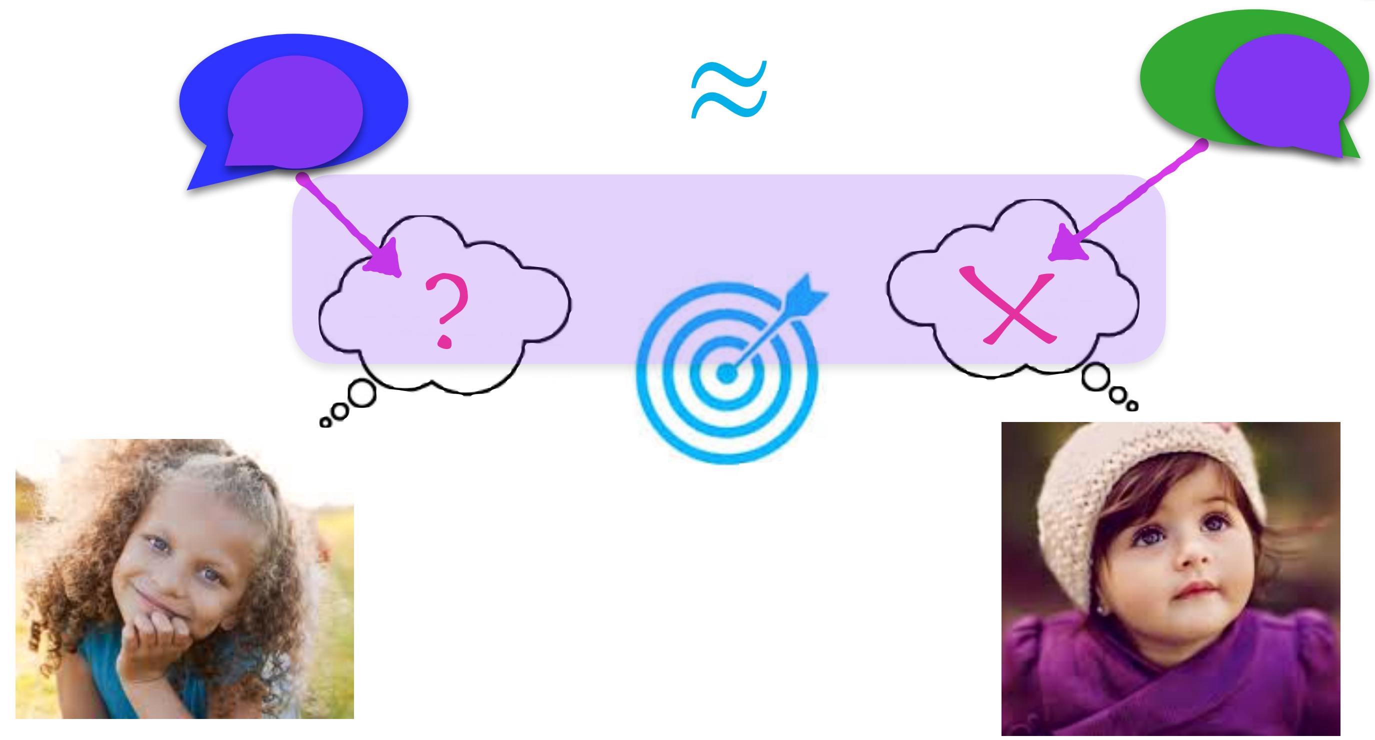
$\boxed{\textit{wh}\text{-dependencies}}$

76.7%	*start*-IP-VP-*end* What did Lily read $\_\__{what}$?	75.5%
10.3%	*start*-IP-*end* What $\_\__{what}$ happened?	12.8%

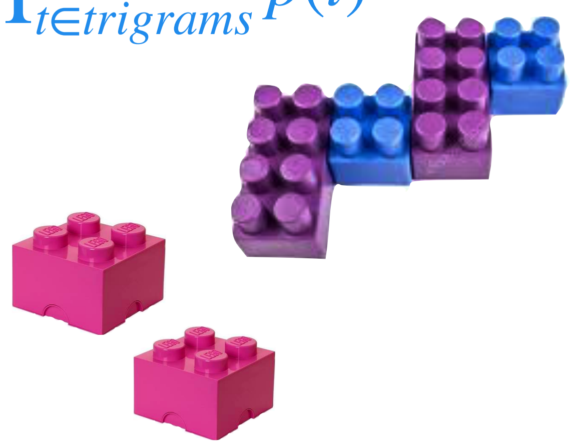
Bates & Pearl 2019, submitted

Meaningful input differences

Who does



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



Does this part of the input act differently across SES? That is, are any differences (even if they're smaller) developmentally meaningful?

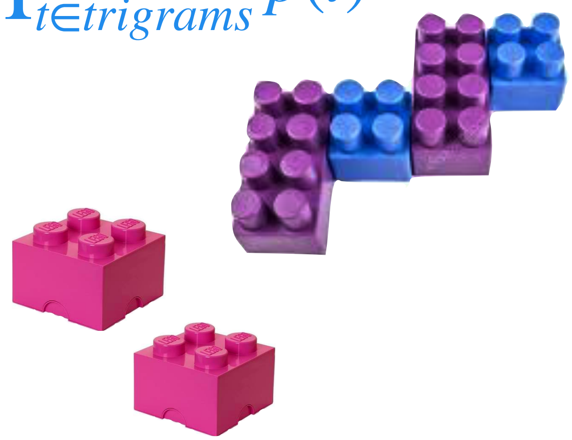
They might be — small differences in the input distribution might snowball into learning outcome differences.

syntactic trigrams

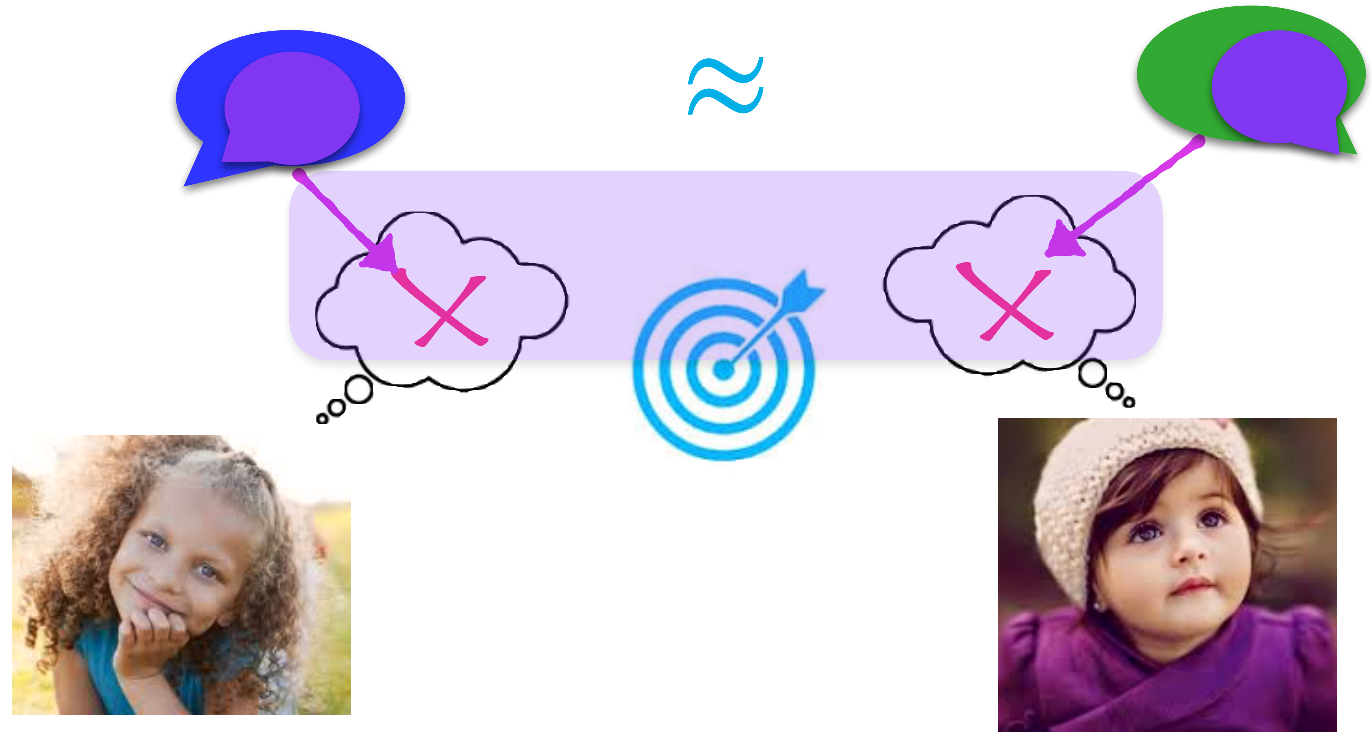
41.4%	*start*-IP-VP	41.8%
38.9%	IP-VP-*end*	40.0%
4.7%	start-IP-*end*	6.1%

Bates & Pearl 2019, submitted

Meaningful input differences



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

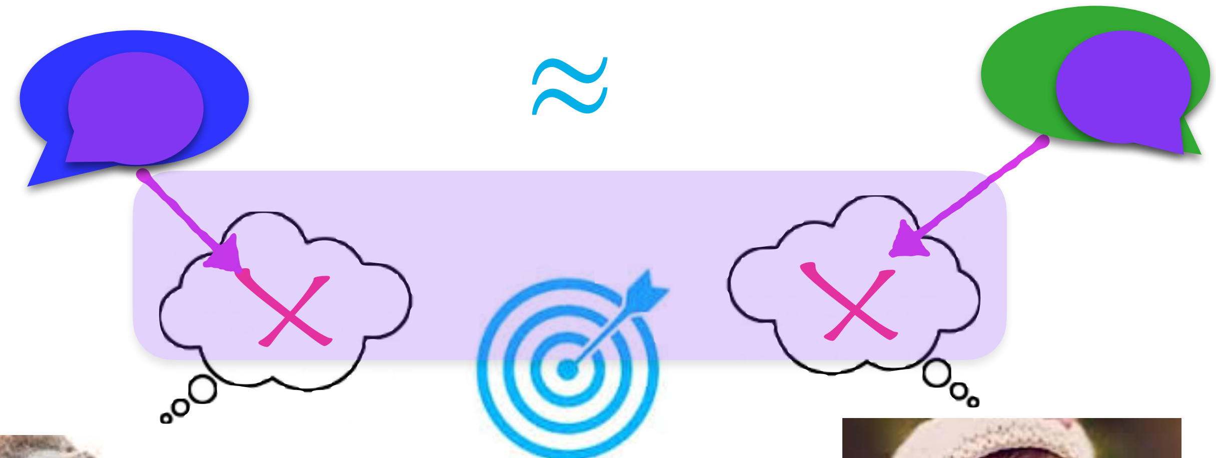


Does this part of the input act differently across SES? That is, are any differences (even if they're smaller) developmentally meaningful?

But…we already have some evidence (De Villiers et al. 2008) that low-SES children disprefer *wh*-dependencies that cross one of the islands we're looking at (Complex NP), just like high-SES children.

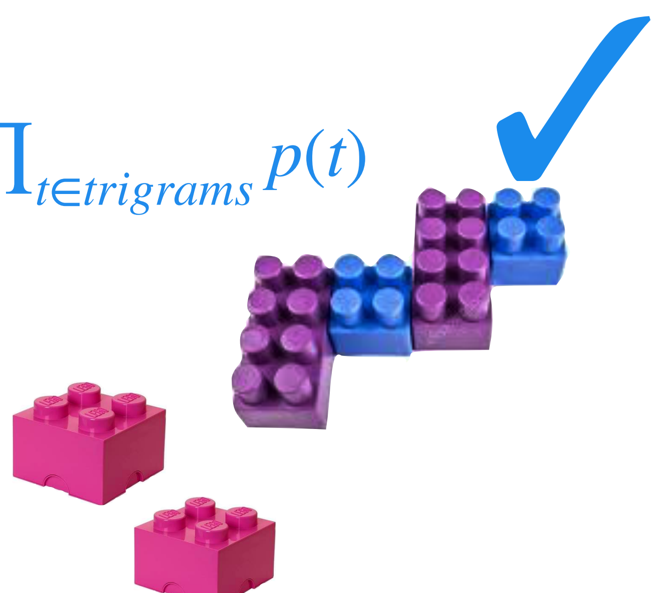
Bates & Pearl 2019, submitted

Meaningful input differences



Does this part of the input act differently across SES? That is, are any differences (even if they're smaller) developmentally meaningful?

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

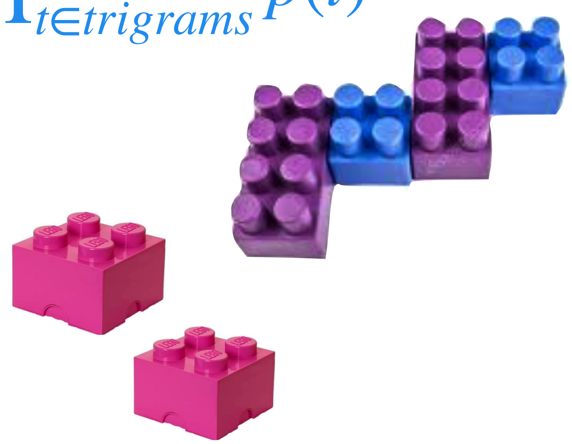


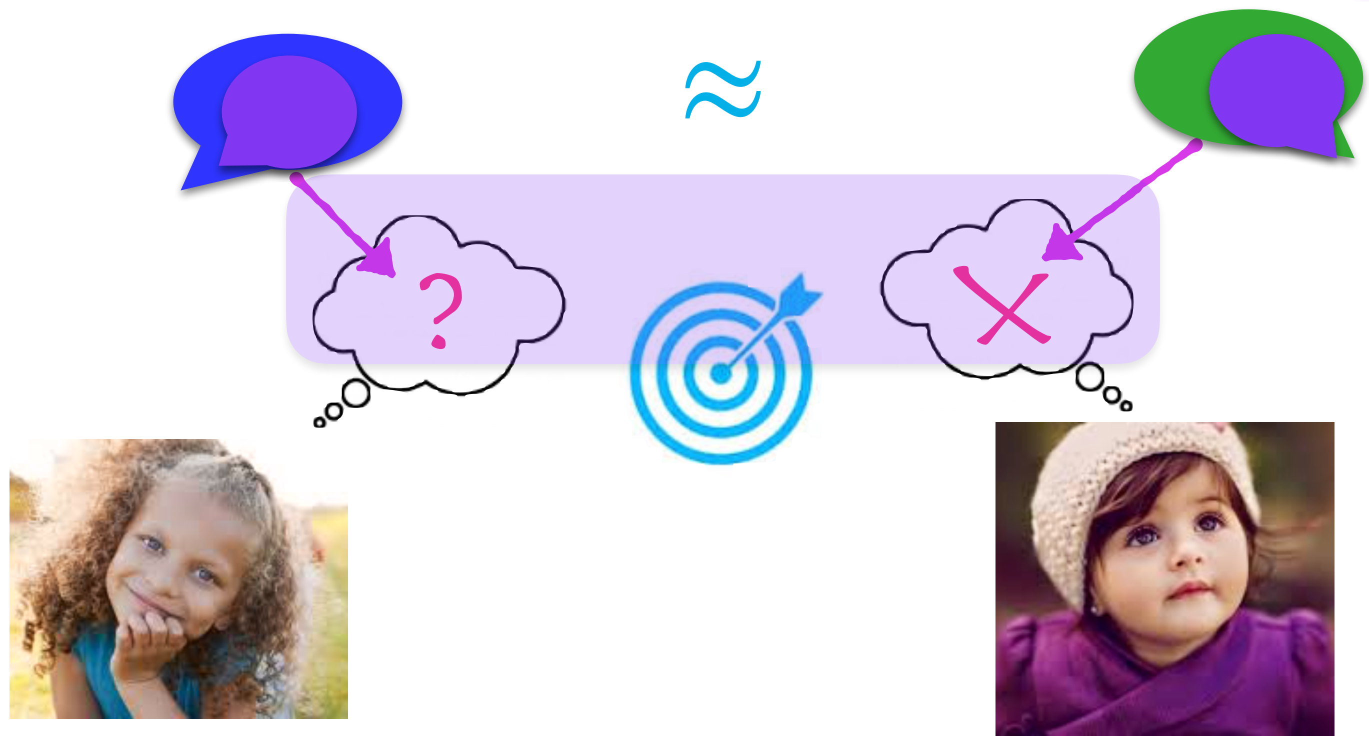
Checkpoint: So at least for Complex NP islands, if our learning theory is right, we shouldn't predict learning outcome differences across SES by age four. The input differences shouldn't be developmentally meaningful.

Bates & Pearl 2019, submitted

Meaningful input differences

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





Does this part of the input act differently across SES? That is, are any differences (even if they're smaller) developmentally meaningful?

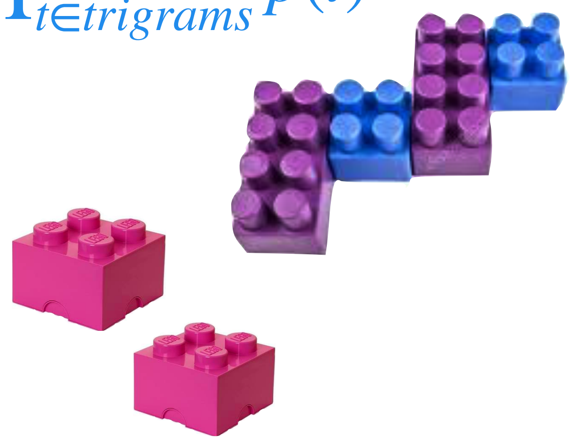
But what about the other three islands types (Subject, Whether, and Adjunct)? What do we predict for that knowledge at four years old? Are any input differences predicted to be developmentally meaningful for these islands?

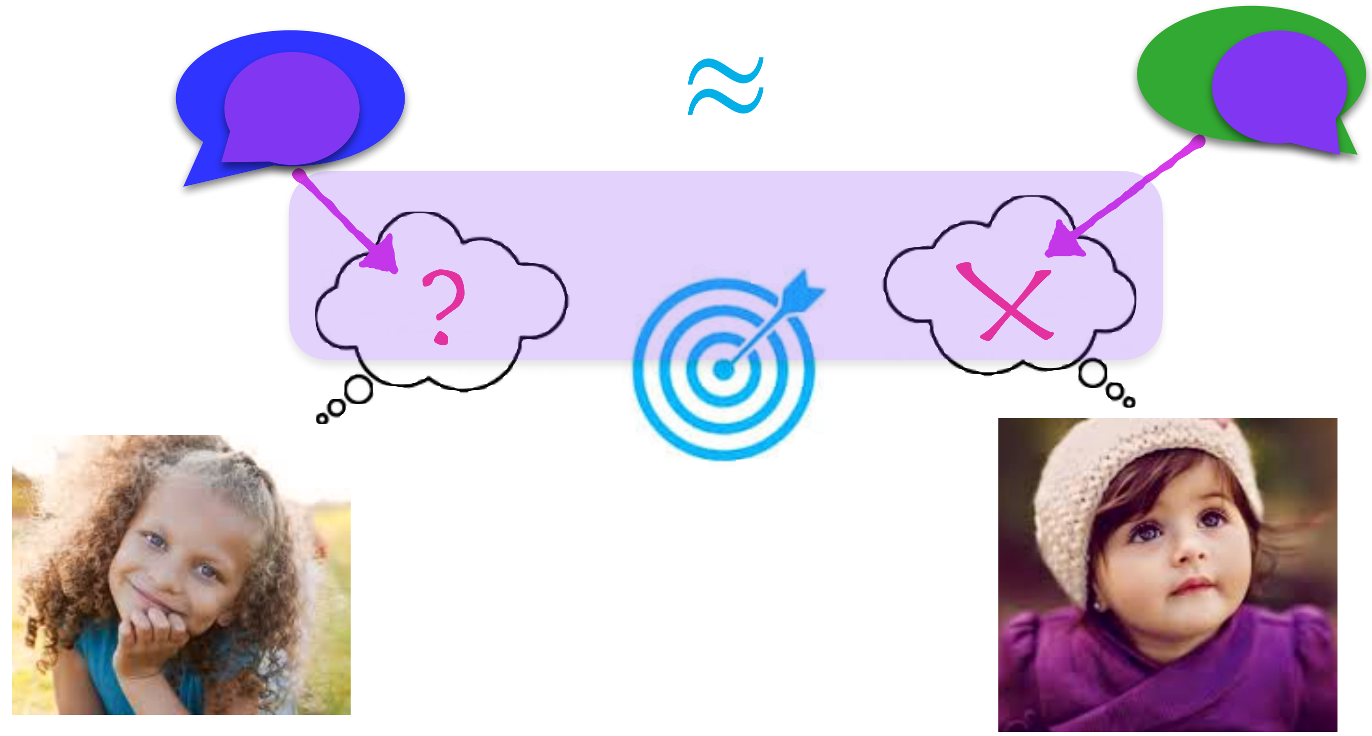
Bates & Pearl 2019, submitted

Meaningful input differences

Who does

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

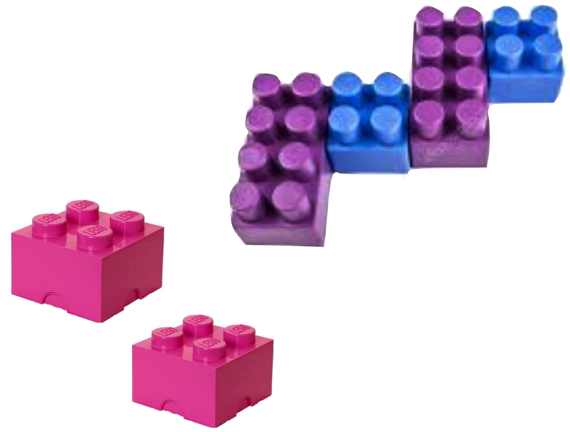




Let's use developmental computational modeling to find out.

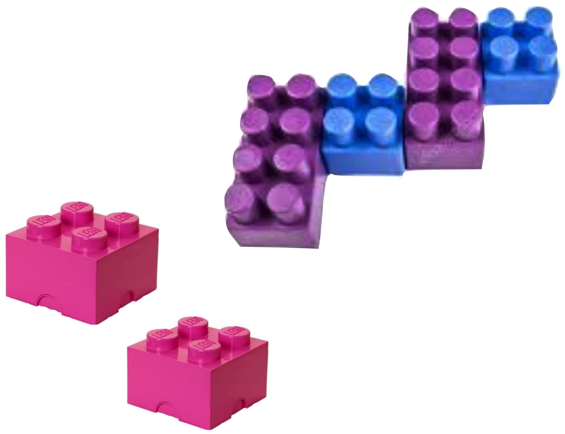
Bates & Pearl 2019, submitted

Who does 



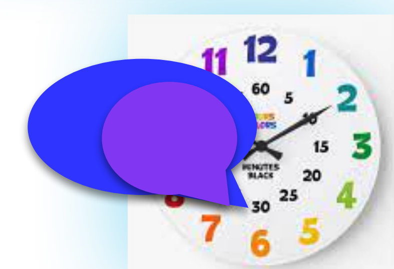
First, how much *wh*-dependency input do low-SES children hear?

Bates & Pearl 2019, submitted



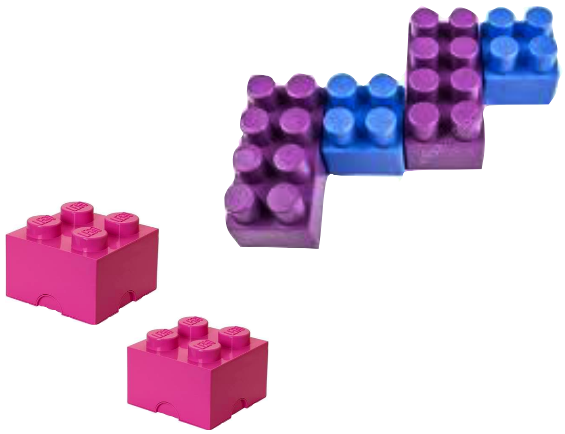
We can estimate how much by

- identifying the potential learning period for these syntactic islands (how much time)
- calculating how many *wh*-dependencies low-SES children would hear during this amount of time (how much input in that time).



Bates & Pearl 2019, submitted

Who does 



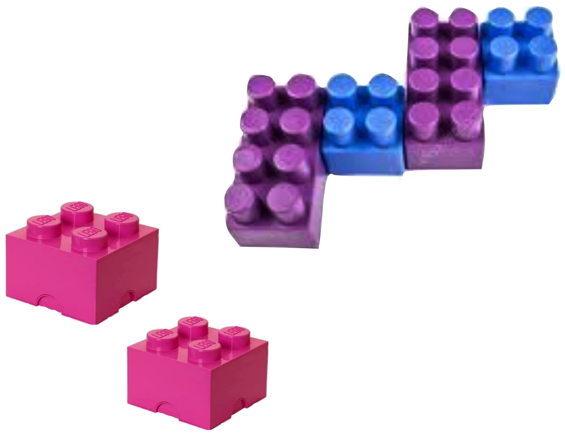
How much time

Children (both high-SES and low-SES) seem to know about Complex NP islands by four years old (<60 months).

End point: Just before that (59 months).

Bates & Pearl 2019, submitted

Who does

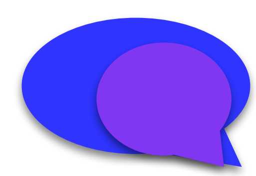


How much time

59 months

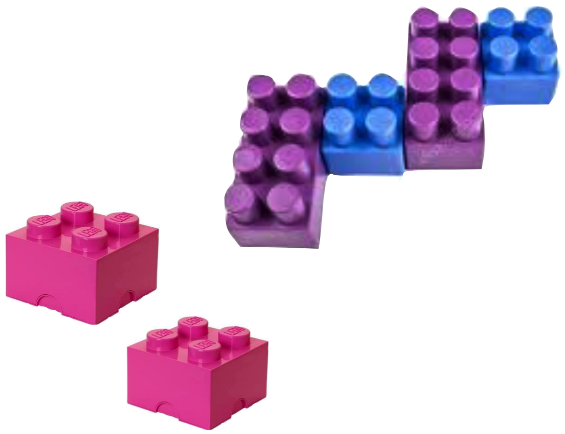
Children begin to represent the full structure of *wh*-dependencies (e.g., *wh*-questions and relative clauses) around 20 months: Seidl et al. 2003, Gagliardi et al. 2016, Perkins & Lidz 2020.

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



Bates & Pearl 2019, submitted

Who does



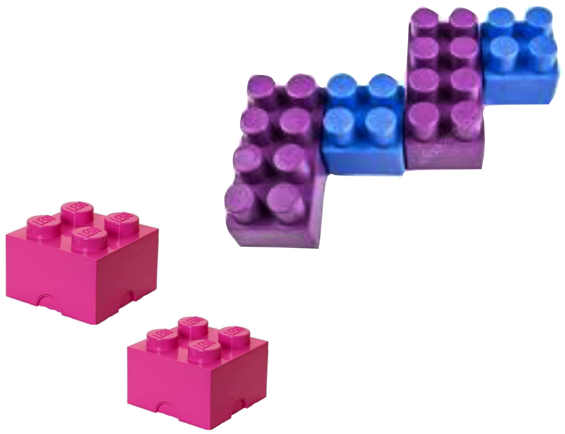
How much time

20 months ⟶ 59 months

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

Bates & Pearl 2019, submitted

Who does



How much time

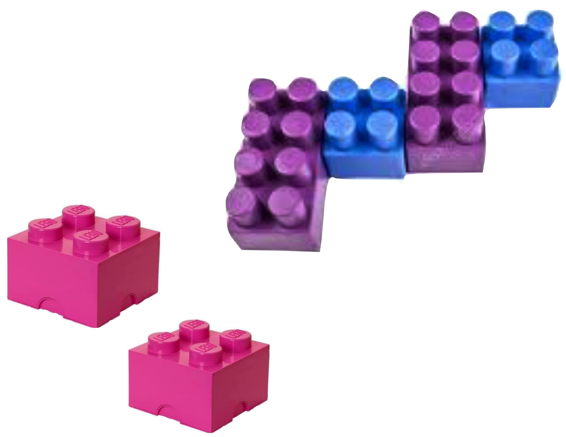
20 months ⟶ 59 months

How many hours

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

Bates & Pearl 2019, submitted

Who does



How much time

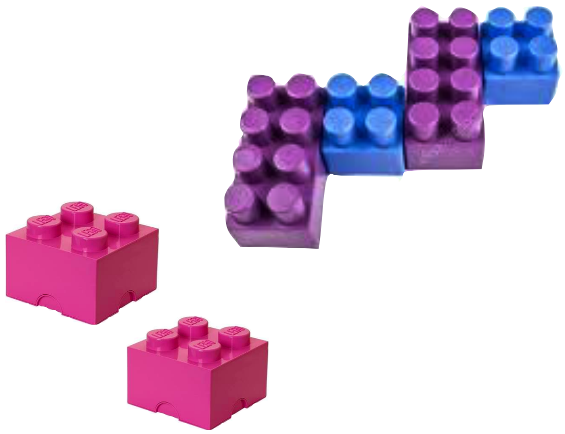
20 months → 59 months

≈ 14174 hours

How much input (in particular, how many *wh*-dependencies) do children encounter during that time?

Bates & Pearl 2019, submitted

Who does



How much time

20 months ⟶ 59 months

≈ 14174 hours

How many *wh*-dependencies

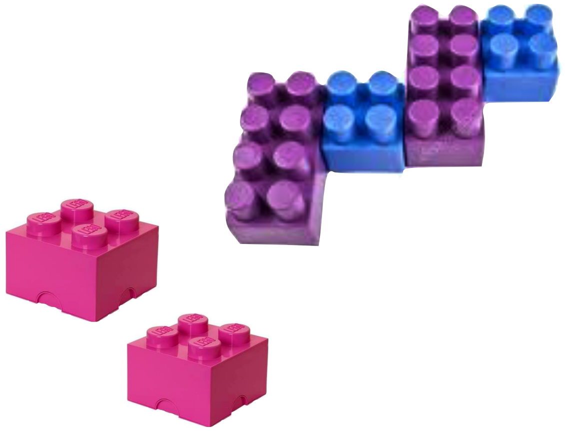
Hart & Risley 1995: Estimates of utterances per hour for high-SES and low-SES households

301

487

Bates & Pearl 2019, submitted

Who does



How much time

20 months → 59 months

≈ 14174 hours

How many *wh*-dependencies

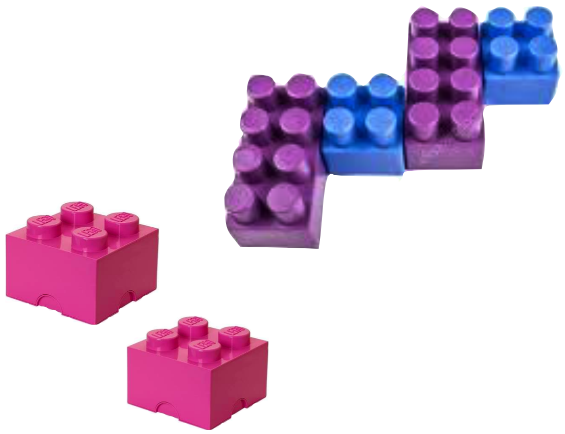
utterances per hour

We can use our own corpus samples to estimate the rate of *wh*-dependencies/utterance.

301

487

Bates & Pearl 2019, submitted



How much time

20 months ⟶ 59 months

≈ 14174 hours

How many *wh*-dependencies

utterances per hour

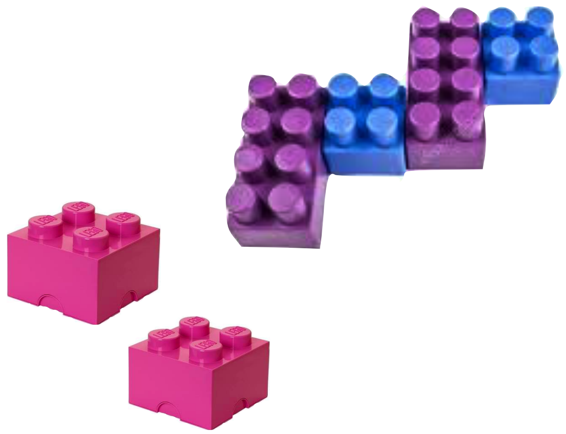
wh-dependencies/utterance

301 487

SES	hours	*	utt/hour	*	*wh*-dep/utt	=	total *wh*-dep
high-SES	14174	*	487	*	20932/101838	=	1,418,193
low-SES	14174	*	301	*	3904/31875	=	522,539

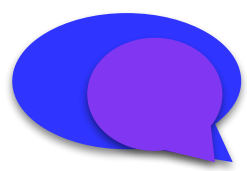
Bates & Pearl 2019, submitted

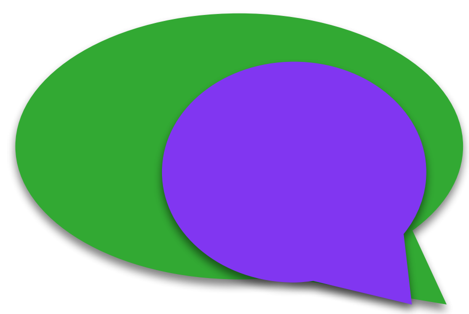
Who does



How many *wh*-dependencies

522,539 1,418,193

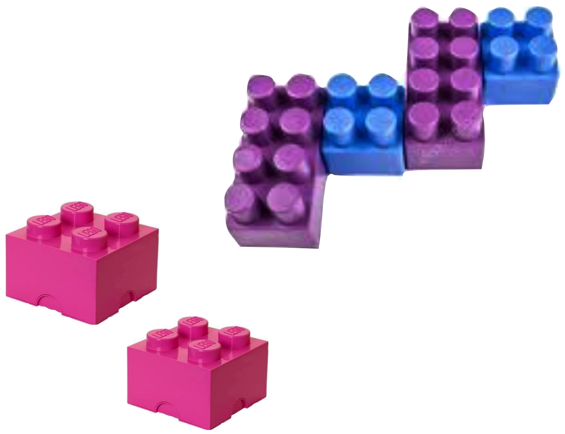




Low-SES children hear about a third the *wh*-dependencies that high-SES children do!

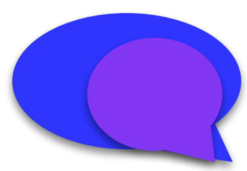
Bates & Pearl 2019, submitted

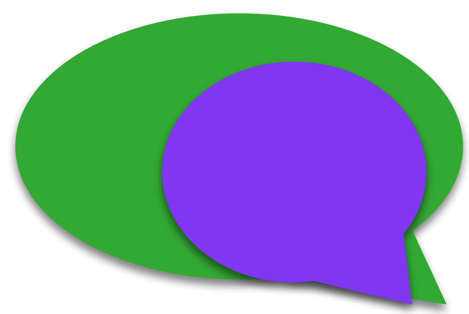
Who does



How many *wh*-dependencies

522,539 1,418,193

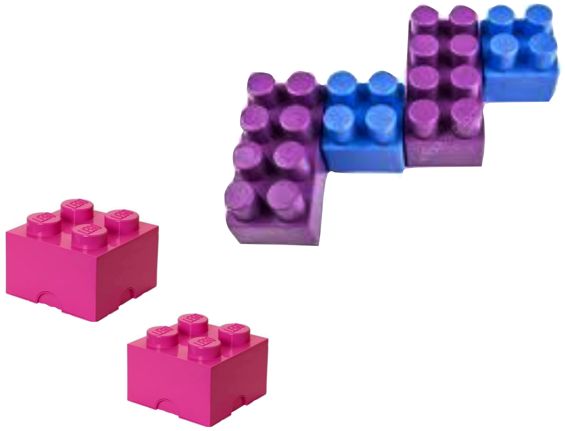




But does the learning theory implemented by our developmental model predict that quantitative difference to matter by age four?

Bates & Pearl 2019, submitted

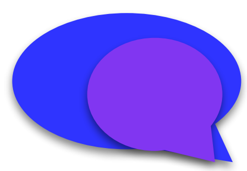
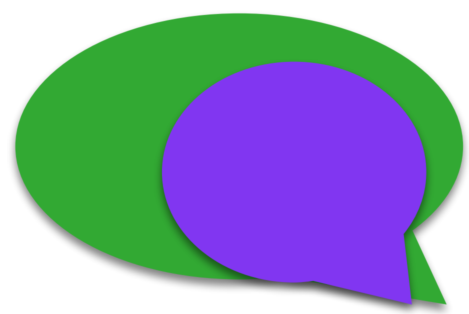
Who does 



How many *wh*-dependencies

522,539 1,418,193

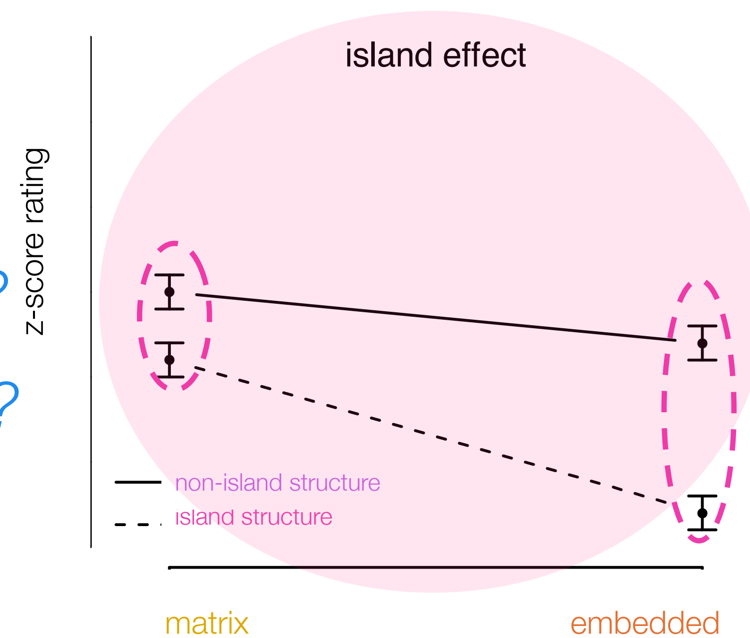
Let's look first at Complex NP islands (where it shouldn't matter if the learning theory is in fact right)

Bates & Pearl 2019, submitted

Reminder: Judgment behavior

Who does

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



Who [non-island]?

Who [island]?

Looking for superadditivity as the sign of syntactic islands knowledge

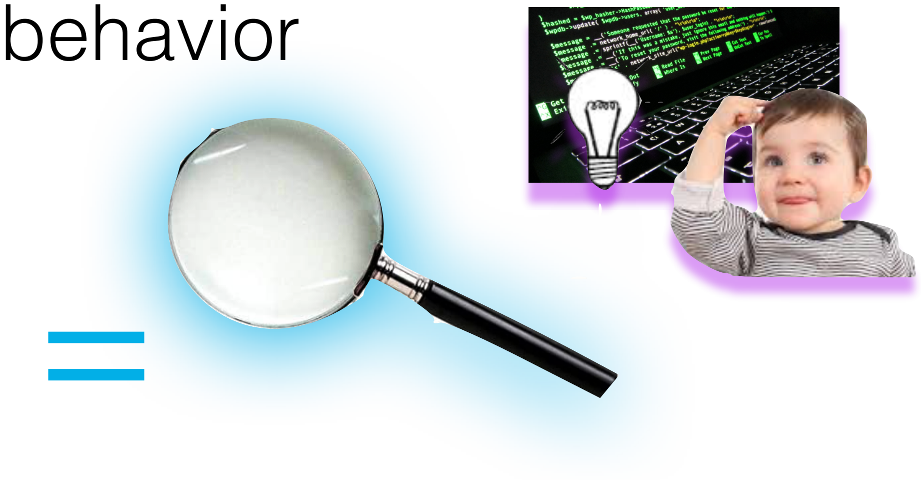
Who __who?

Who [CP… __who]?

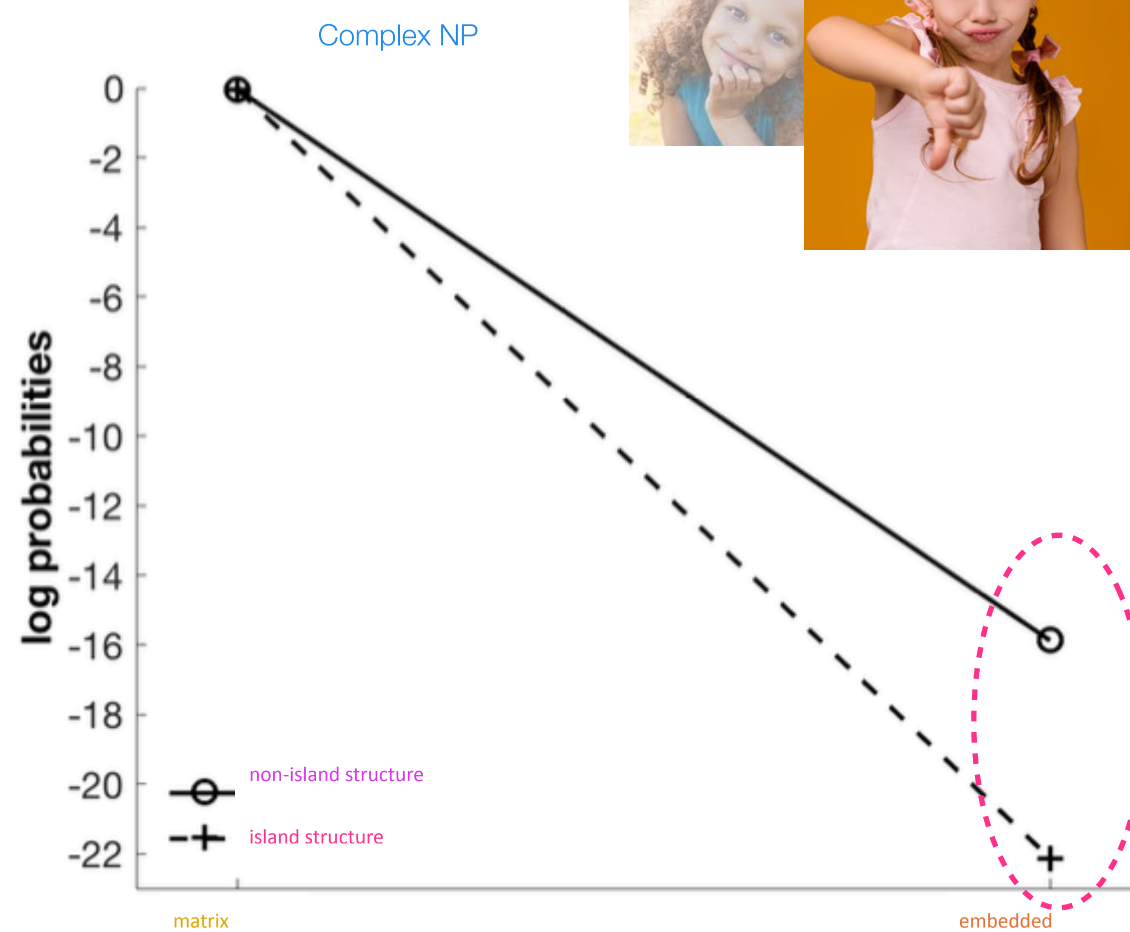
Bates & Pearl 2019, submitted

Predicted judgment behavior

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



Complex NP islands: Superadditivity predicted to be present for low-SES children too! (Low-SES children don't like dependencies crossing Complex NP islands either.)



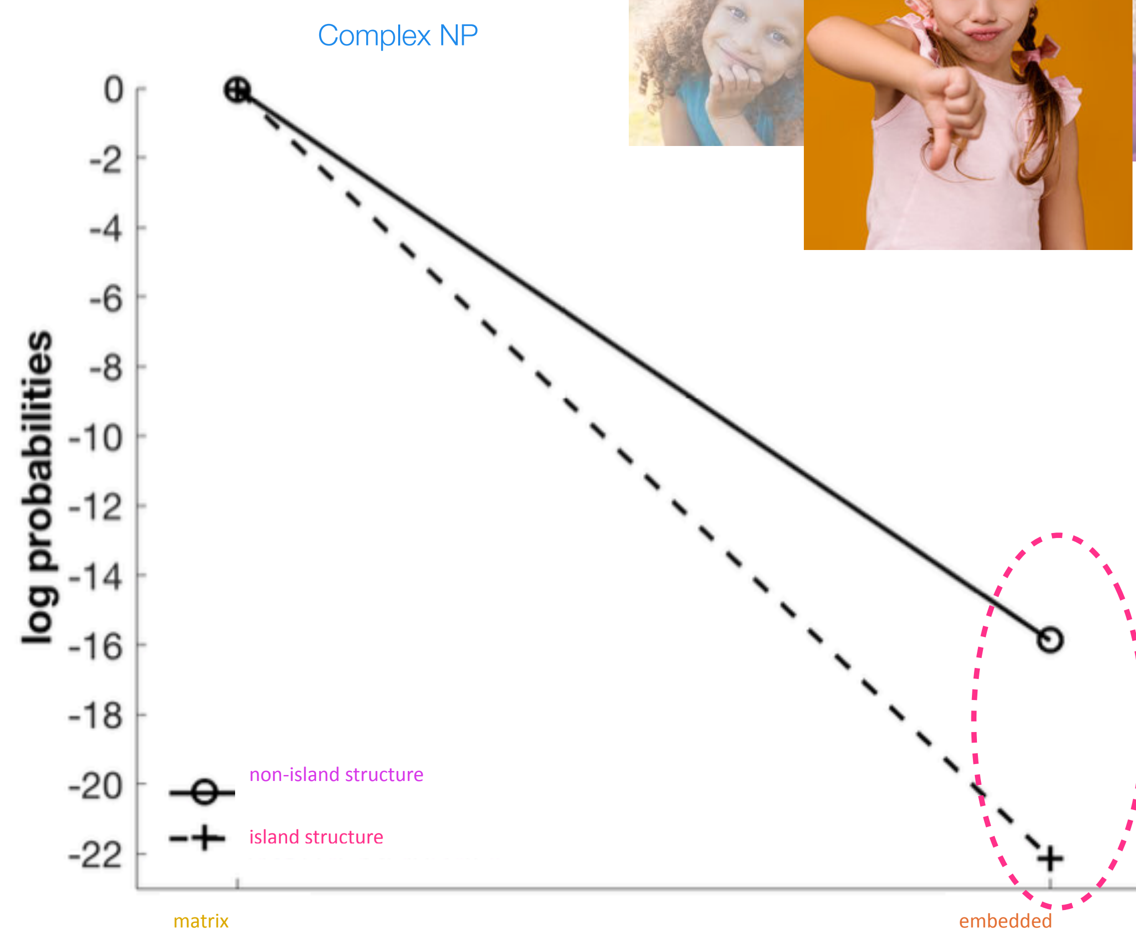
Bates & Pearl 2019, submitted

Predicted judgment behavior

Who does

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

Complex NP islands: Superadditivity predicted to be present for low-SES children too! (Low-SES children don't like dependencies crossing Complex NP islands either.)



No difference from high-SES behavior, despite measurable input differences. The learning theory has passed this checkpoint!

✓

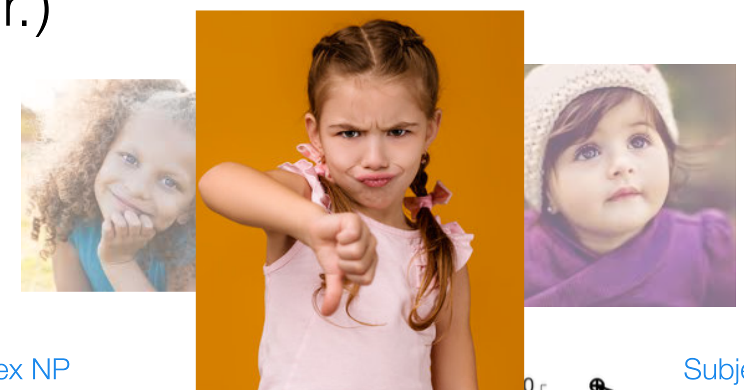
Now, what does the learning theory predict for the other three island types (Subject, Adjunct, and Whether)?

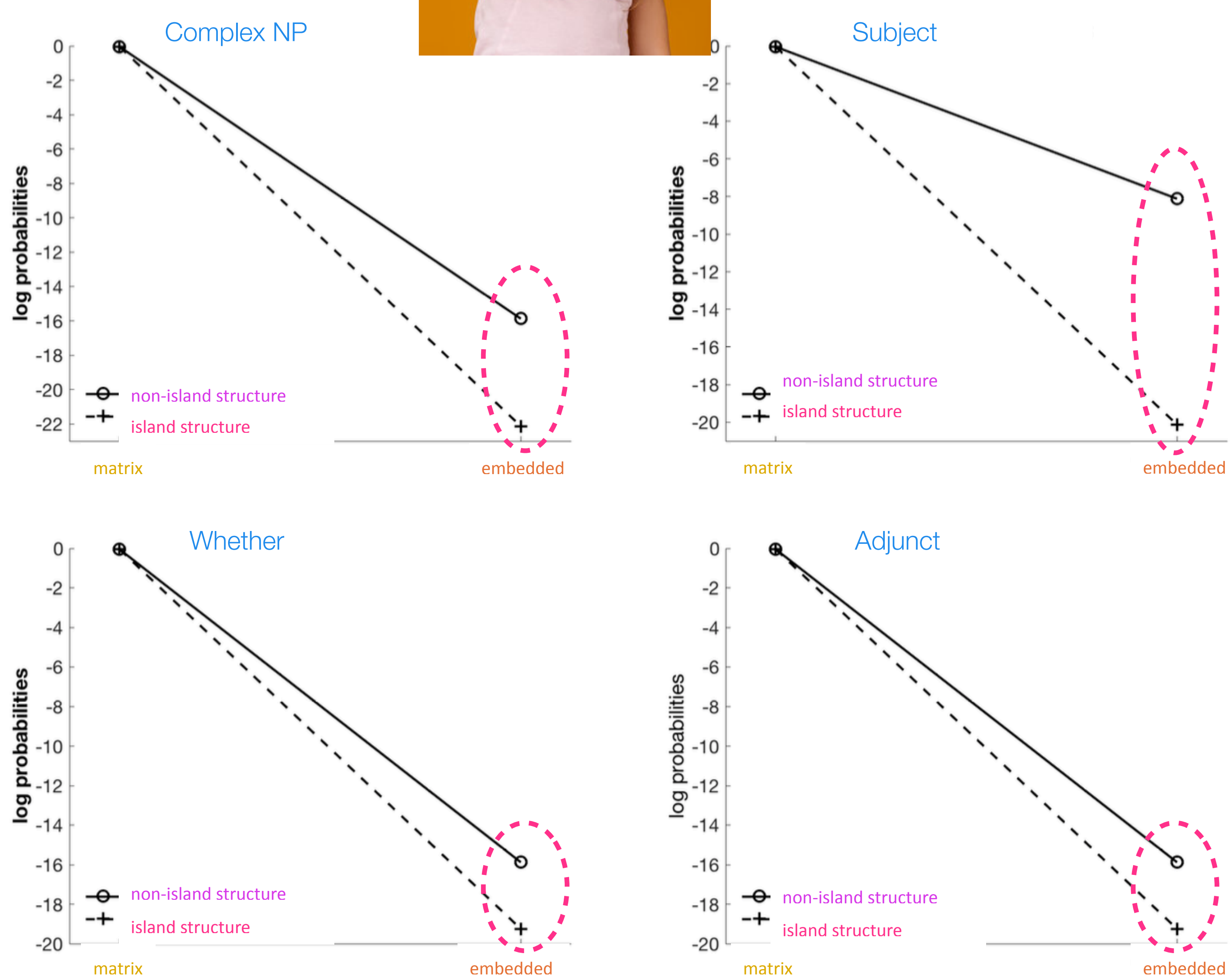
Bates & Pearl 2019, submitted

Meaningful input differences

The other islands: Superadditivity predicted to be present for low-SES children too! (Low-SES children shouldn't like dependencies crossing these islands either.)

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

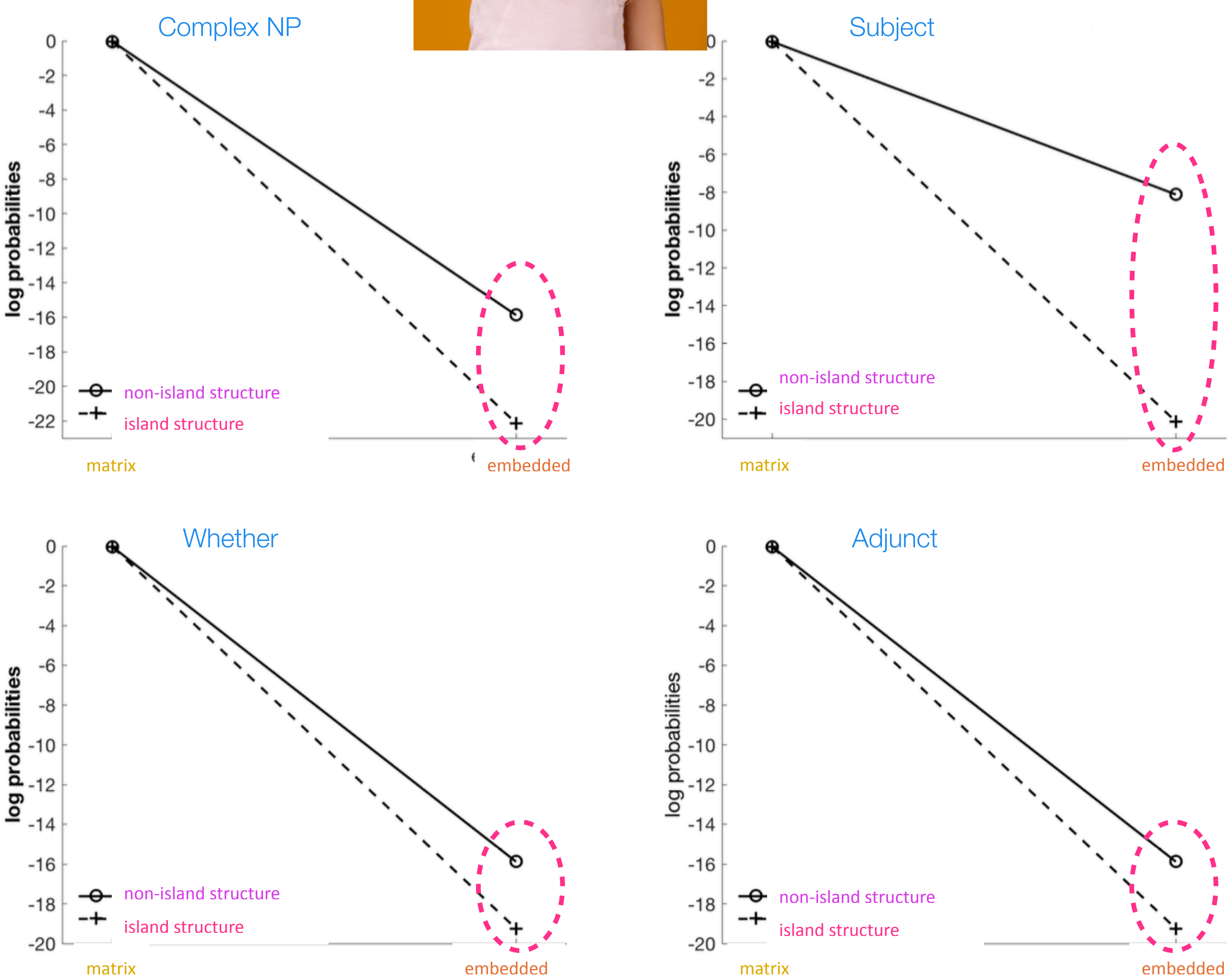




Bates & Pearl 2019, submitted

Meaningful input differences

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



This means low-SES input is predicted to support the same learning outcome knowledge (of these four syntactic islands) as high-SES input.

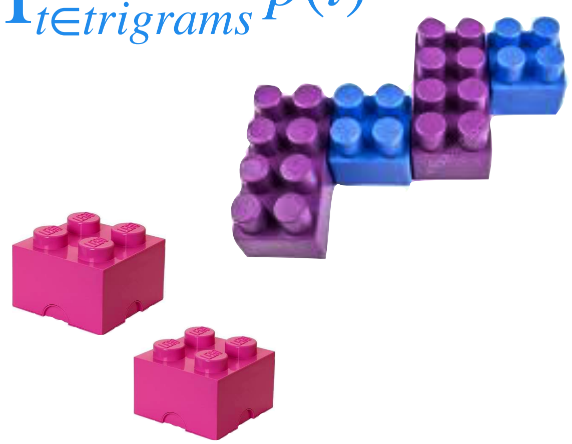
We know this is true for Complex NP islands. ✓

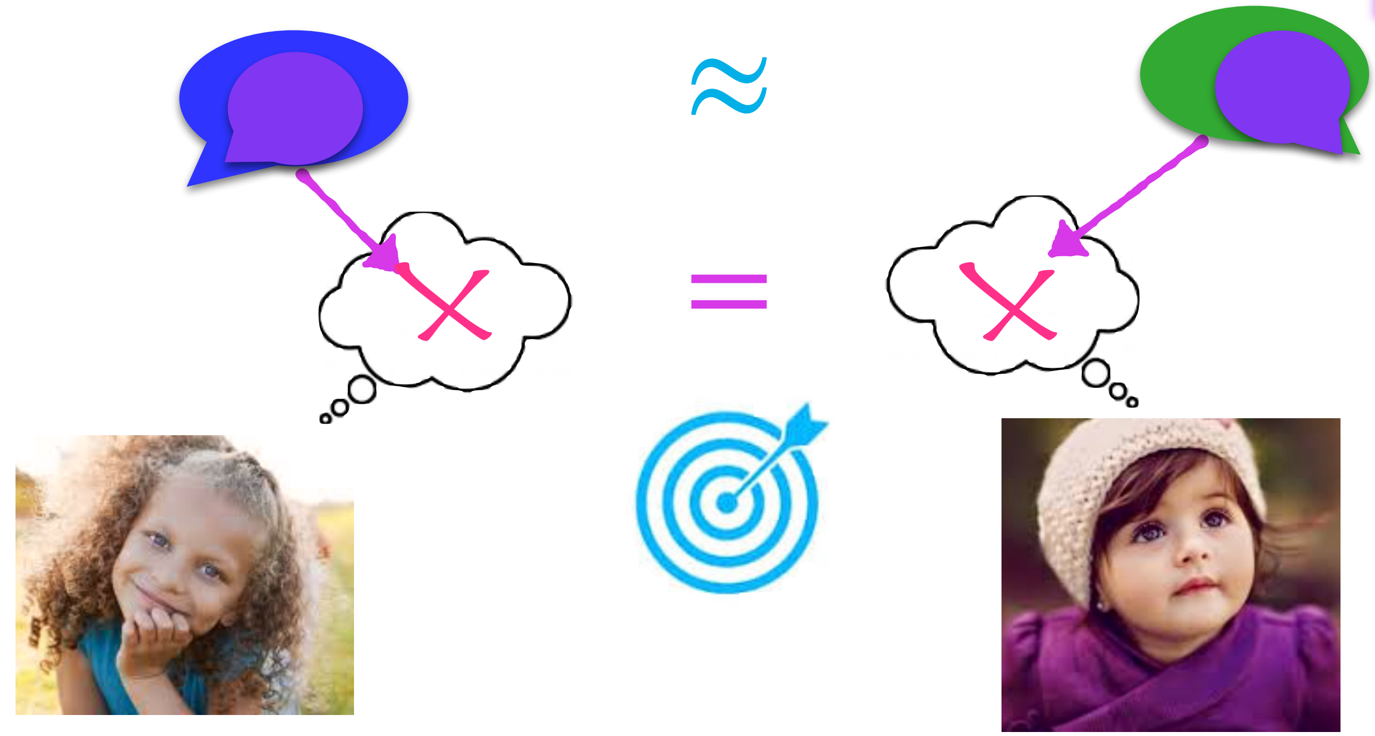
Future work: Collect child judgments for the other island types. ?

Bates & Pearl 2019, submitted

No meaningful input differences predicted

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



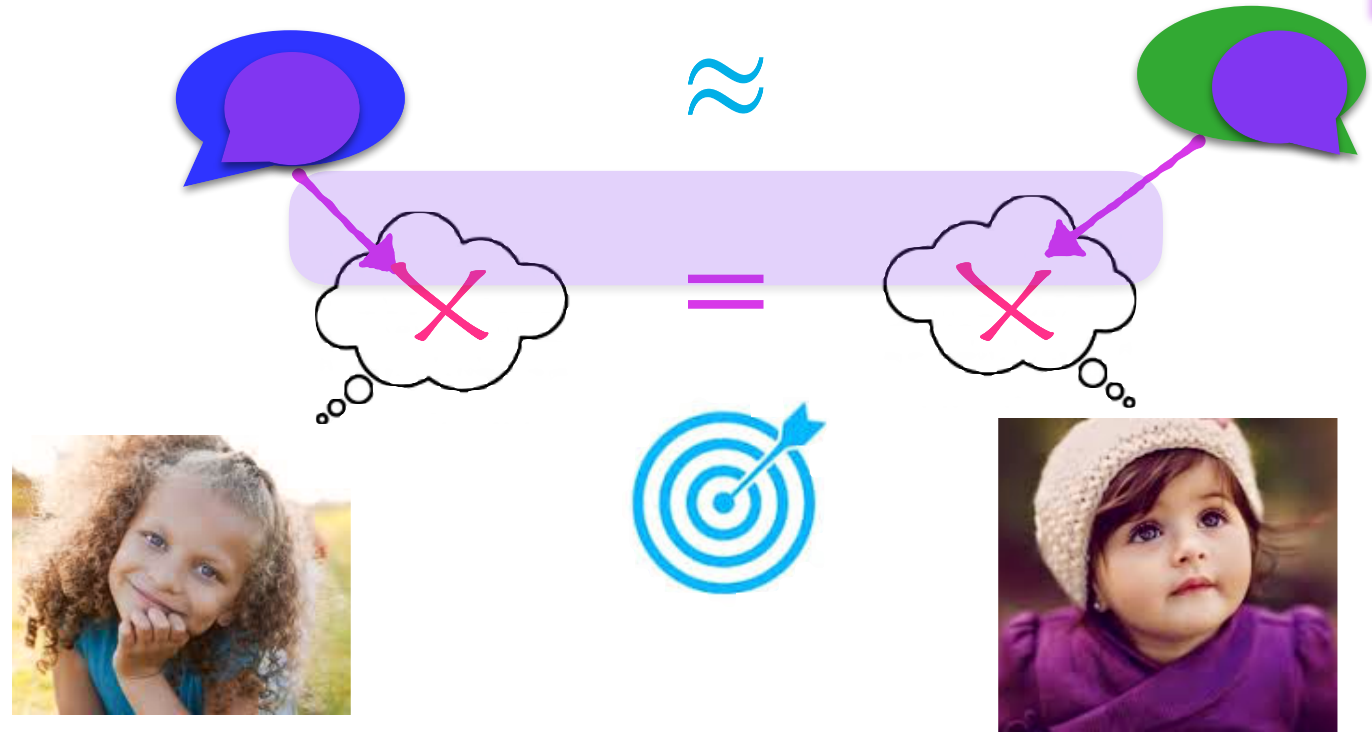


So, our developmental computational model predicts no meaningful input differences across SES when it comes to learning this syntactic island knowledge from this part of the input.

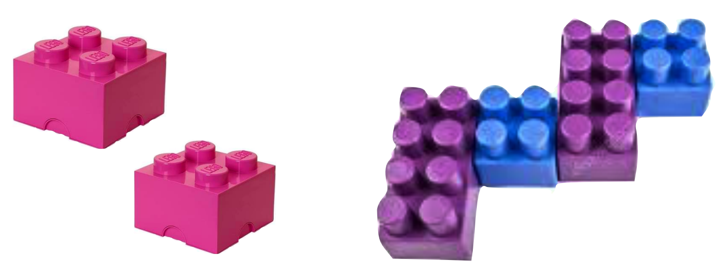
Bates & Pearl 2019, submitted

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

No meaningful input differences predicted

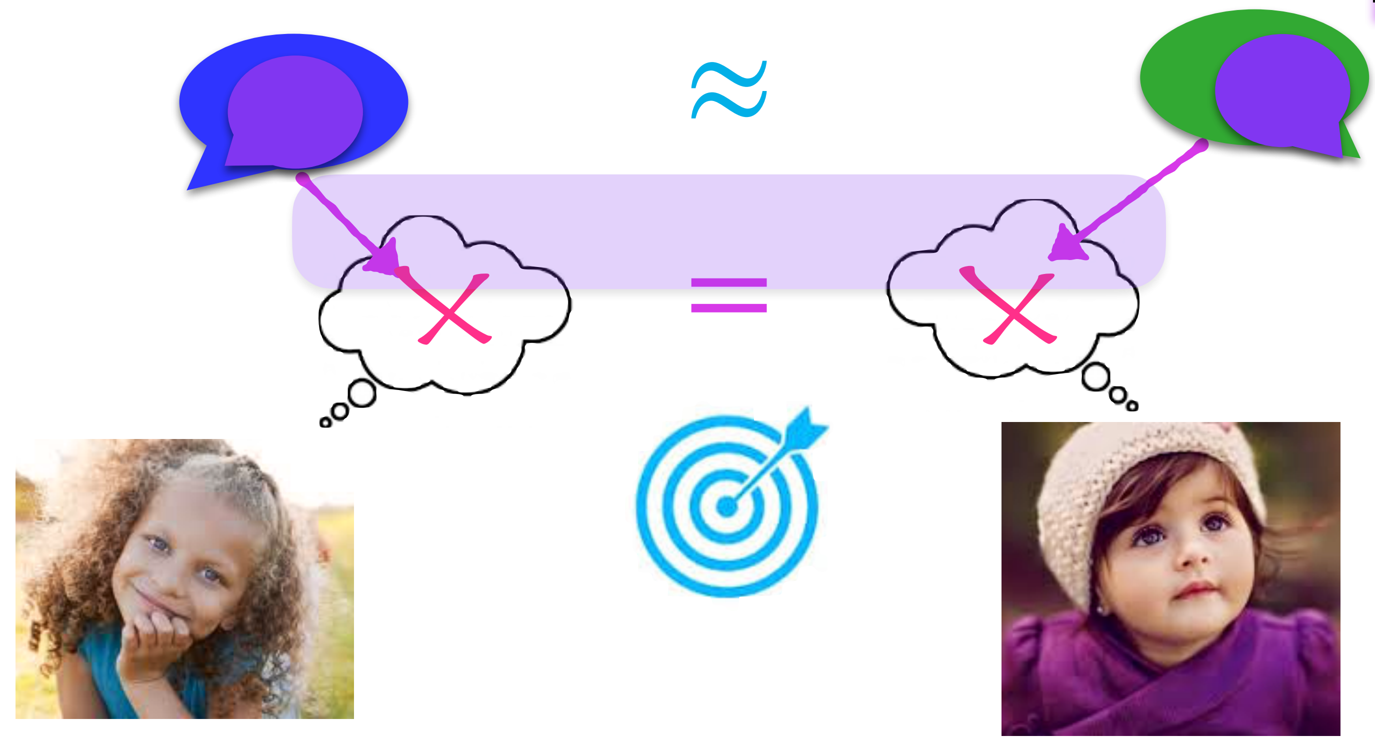


Useful: Because the learning theory specifies how the input is predicted to cause the knowledge to develop, we know which building blocks are particularly important.

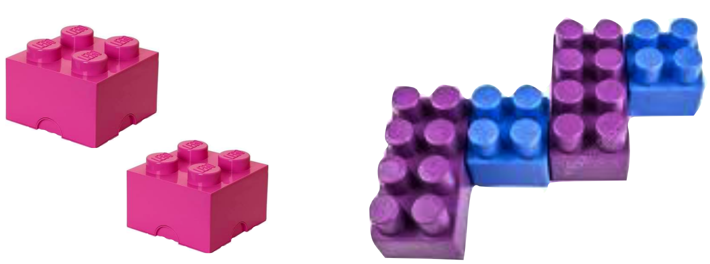


Bates & Pearl 2019, submitted

Key building blocks are there



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



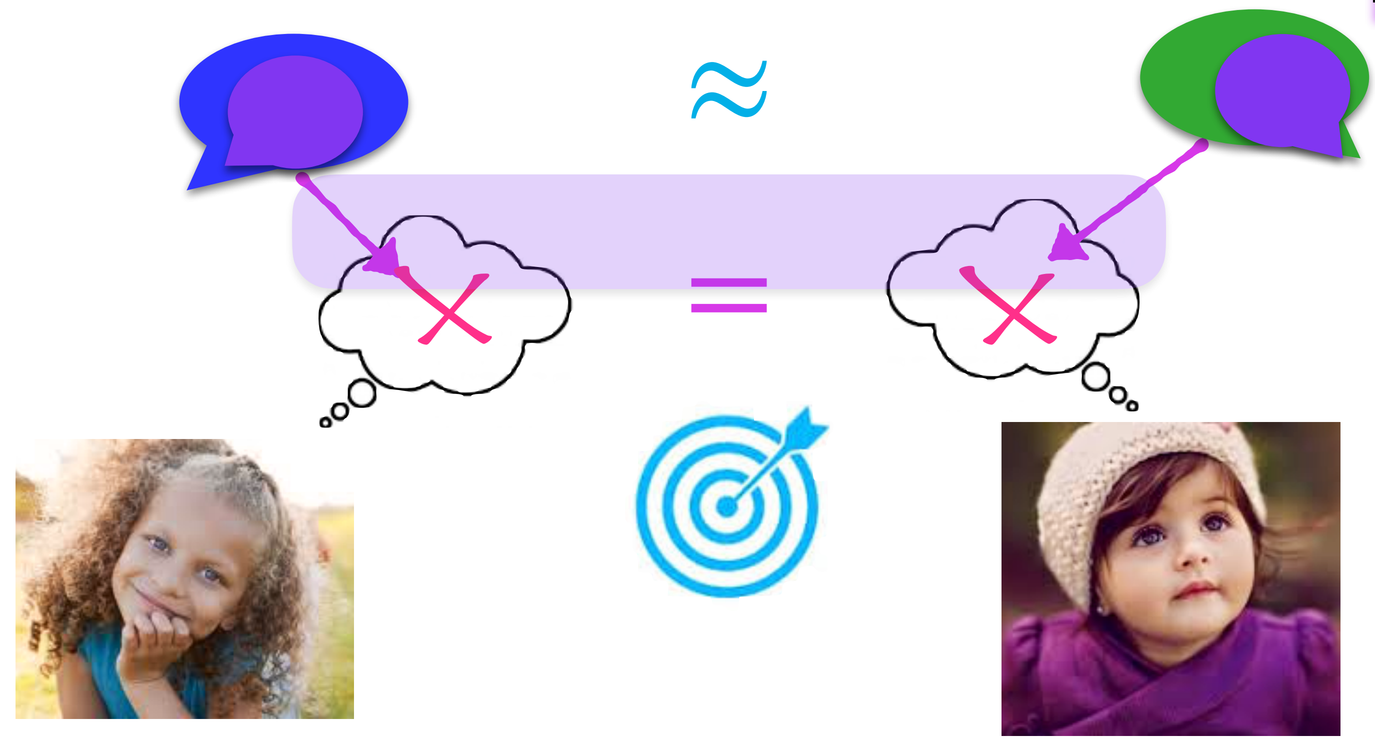
Key building blocks for success involve complementizer *that* (CP_{that}) - this is because two of the islands (Whether and Adjunct) only differ from grammatical dependencies by the complementizer used.

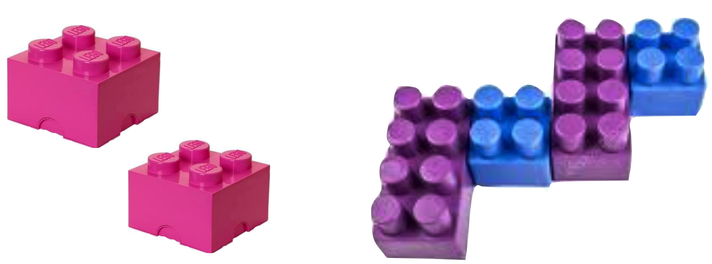
	What does the teacher think	[that Lily forgot __]?	embedded | non-island
whether	*What does the teacher wonder	[whether Lily forgot __]?	embedded | island
adjunct	*What does the teacher worry	[if Lily forgot __]?	embedded | island

Bates & Pearl 2019, submitted

Key building blocks are there

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





Key building blocks for success involve complementizer *that* (CP_{that}) - this is because two of the islands (Whether and Adjunct) only differ from grammatical dependencies by the complementizer used.

start-IP-VP-CP_{that}- IP-VP-*end* embedded | non-island

whether * *start*-IP-VP-$\mathrm{CP}_{whether}$-IP-VP-*end* embedded | island

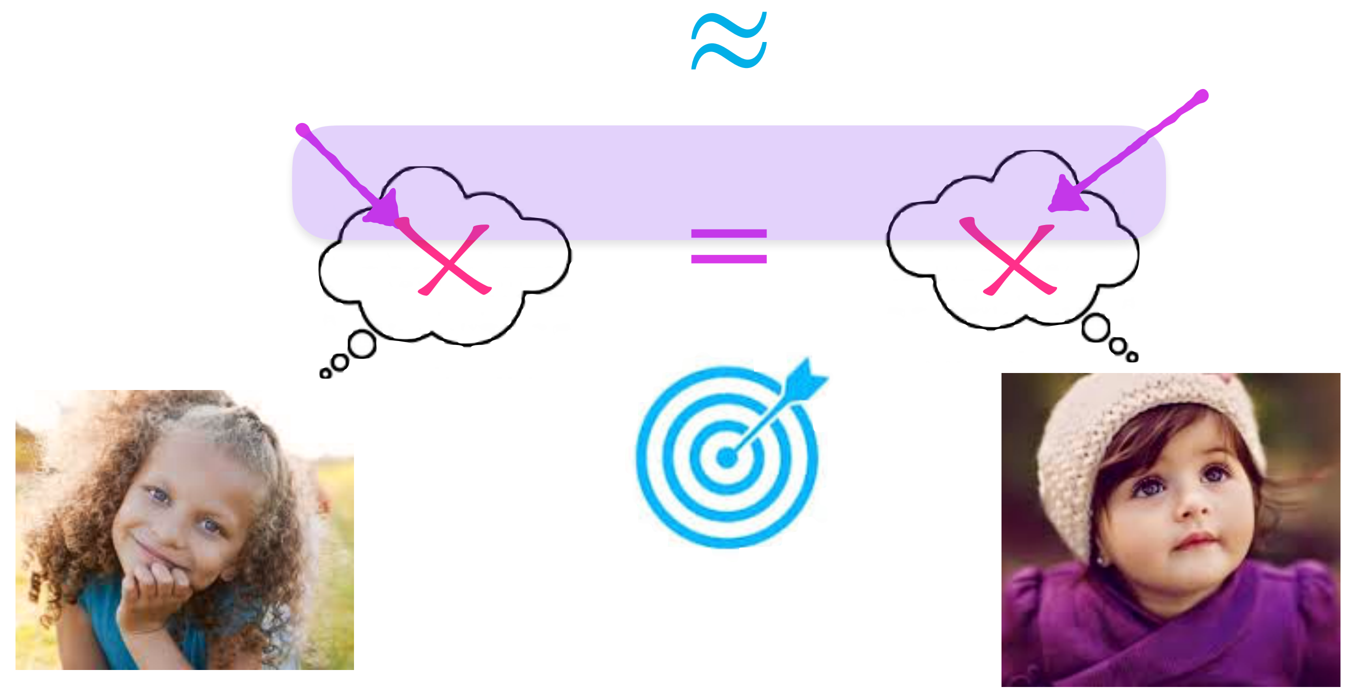
adjunct * *start*-IP-VP-CP_{if}- IP-VP-*end* embedded | island

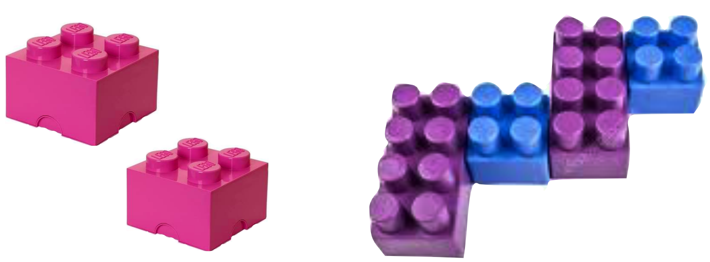
Bates & Pearl 2019, submitted

Key building blocks are there

Who does

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



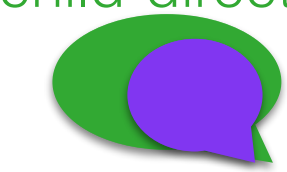


So, children need to encounter grammatical *wh*-dependencies that involve $\mathrm{CP}_{\mathrm{that}}$. These are actually pretty rare in child-directed speech.

Low-SES child-directed

2 instances of 3.9K (=.05%)

High-SES child-directed



2 instances of 21K (<.01%)

What do you think **that** $\_\__{\text{what}}$ happens?

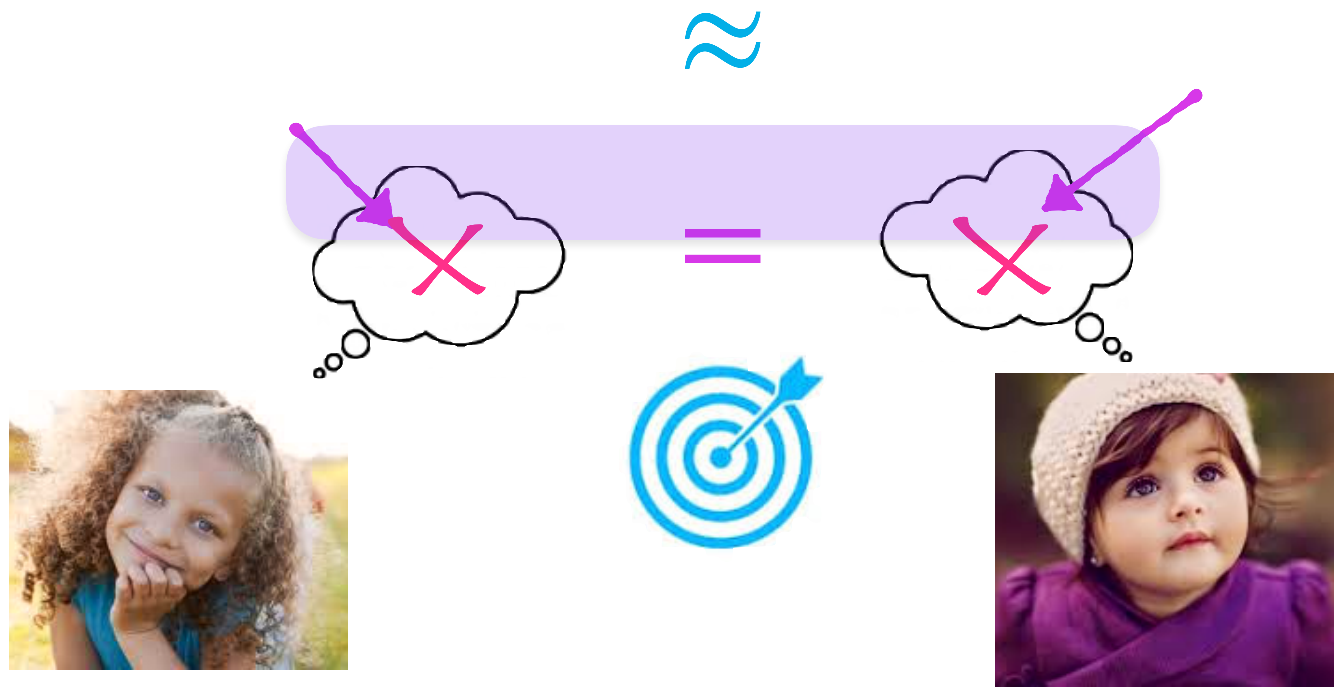
What do you think **that** Jack read $\_\__{\text{what}}$?

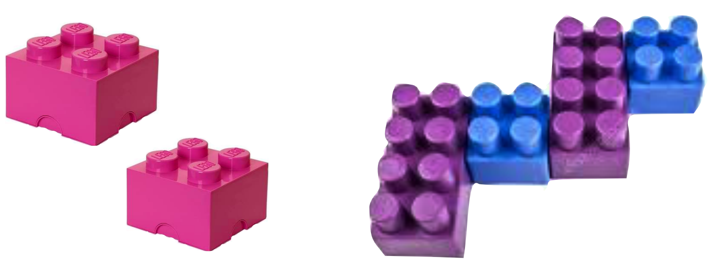
Bates & Pearl 2019, submitted

Who does

Key building blocks are there

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





But with enough input (over several years), even these rare cases are predicted to support learning.

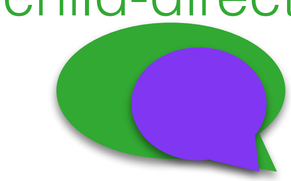
Low-SES child-directed

2 instances of 3.9K (=.05%)

522,539 * 2/3904 ≈ **268**

What do you think **that** __$_{what}$ happens?

High-SES child-directed



2 instances of 21K (<.01%)

1,418,193 * 2/20932 ≈ **136**

What do you think **that** Jack read __$_{what}$?

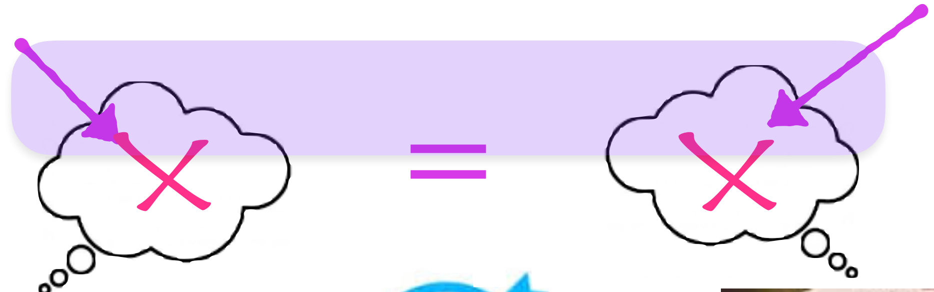
Bates & Pearl 2019, submitted

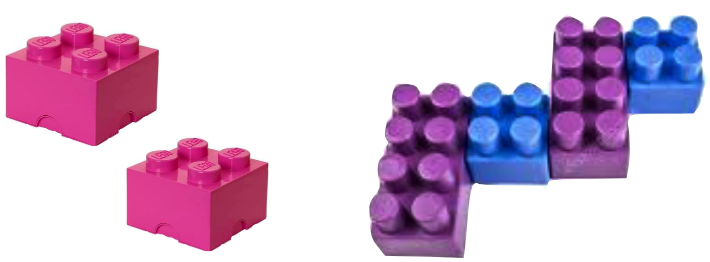
Who does

Key building blocks are there

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

≈





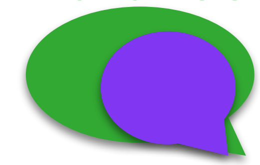
But with enough input (over several years), even these rare cases are predicted to support learning.

Low-SES child-directed

2 instances of
3.9K (=.05%)
8.45/month

What do you think **that** __$_{what}$ happens?

High-SES child-directed



2 instances of
21K (<.01%)
3.39/month

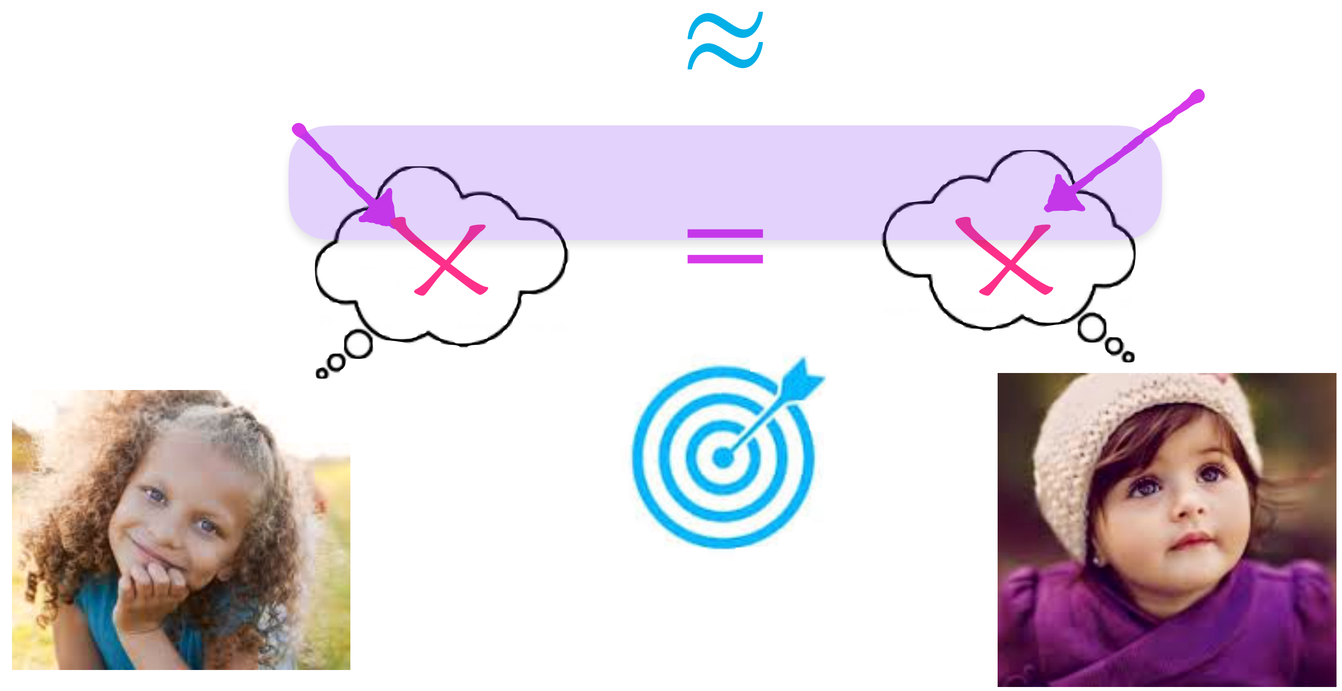
What do you think **that** Jack read __$_{what}$?

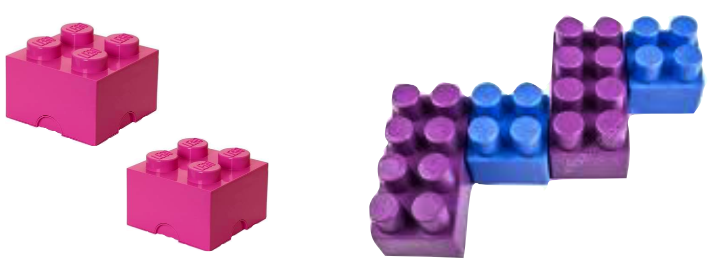
Bates & Pearl 2019, submitted

Key building blocks are there

Who does

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



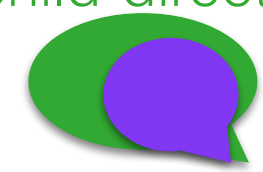


And in fact, if the samples are reasonably accurate, low-SES children actually see this building block more often.

Low-SES
child-directed

2 instances of
3.9K **(=.05%)**
8.45/month

High-SES
child-directed



2 instances of
21K **(<.01%)**
3.39/month

What do you think **that** __$_{what}$ happens?

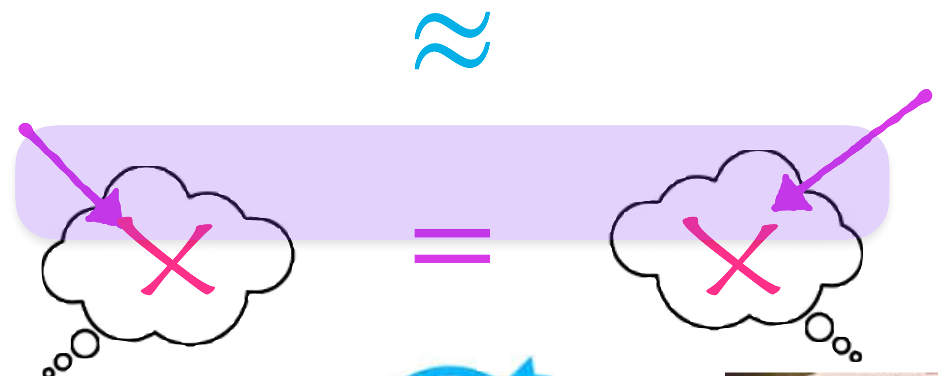
What do you think **that** Jack read __$_{what}$?

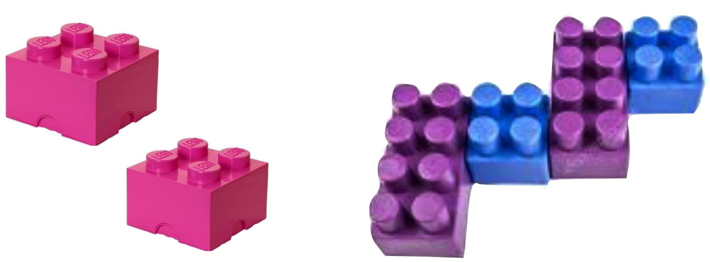
Bates & Pearl 2019, submitted

Who does

Key building blocks are there

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



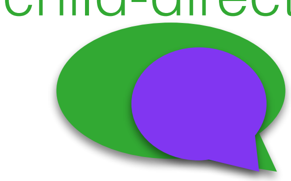


Interesting: The low-SES *wh*-dependency from the input sample with this building block is typically judged to be ungrammatical in the high-SES dialect (a *that*-trace violation).

Low-SES
child-directed

2 instances of
3.9K (=.05%)

High-SES
child-directed



2 instances of
21K (<.01%)

What do you think **that** $\_\__{what}$ happens?

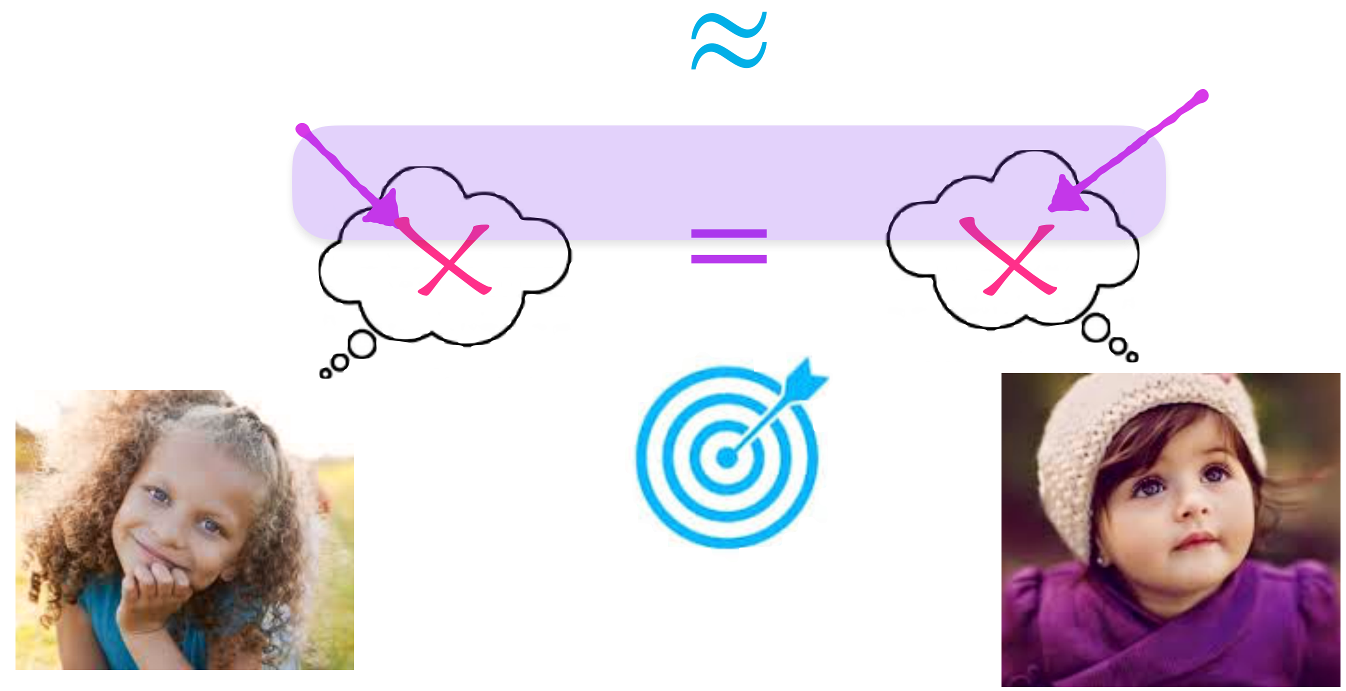
What do you think **that** Jack read $\_\__{what}$?

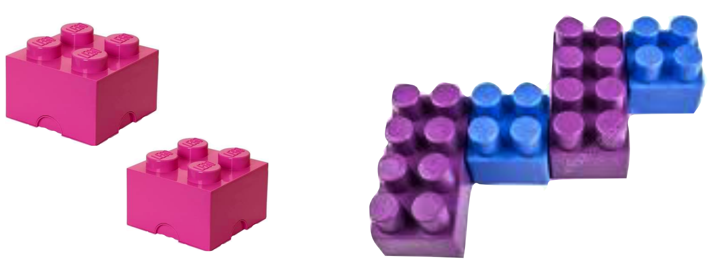
Bates & Pearl 2019, submitted

Who does

Key building blocks are there

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





Upshot: Low-SES children are predicted to achieve the same learning outcome as high-SES children by leveraging key building blocks from sources a high-SES child wouldn’t hear (because those sources are ungrammatical for high-SES speakers).

Low-SES
child-directed

High-SES
child-directed

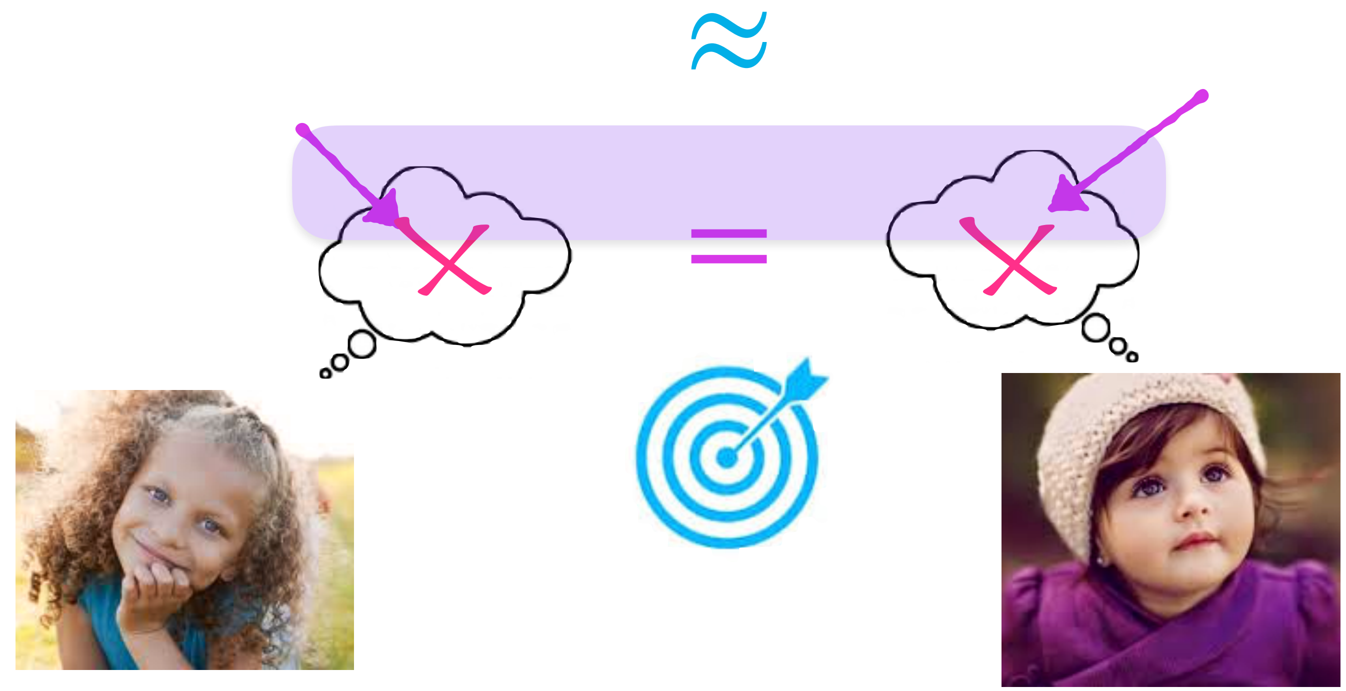
What do you think **that** $\_\__{what}$ happens?

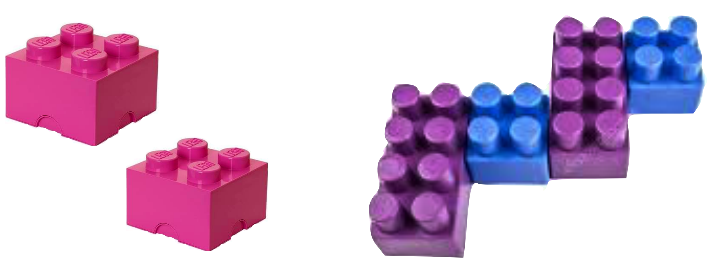
Bates & Pearl 2019, submitted

Who does

Key building blocks are there

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

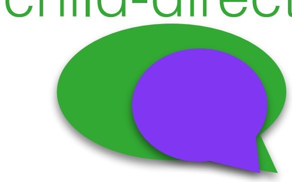




Takeaway: This is one reason why measurable differences in the input might not be developmentally meaningful. Under this learning theory, the building blocks may show up in different places, but they're still present in the input.

Low-SES child-directed

High-SES child-directed

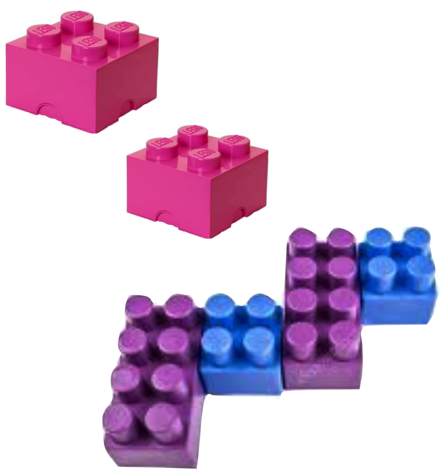


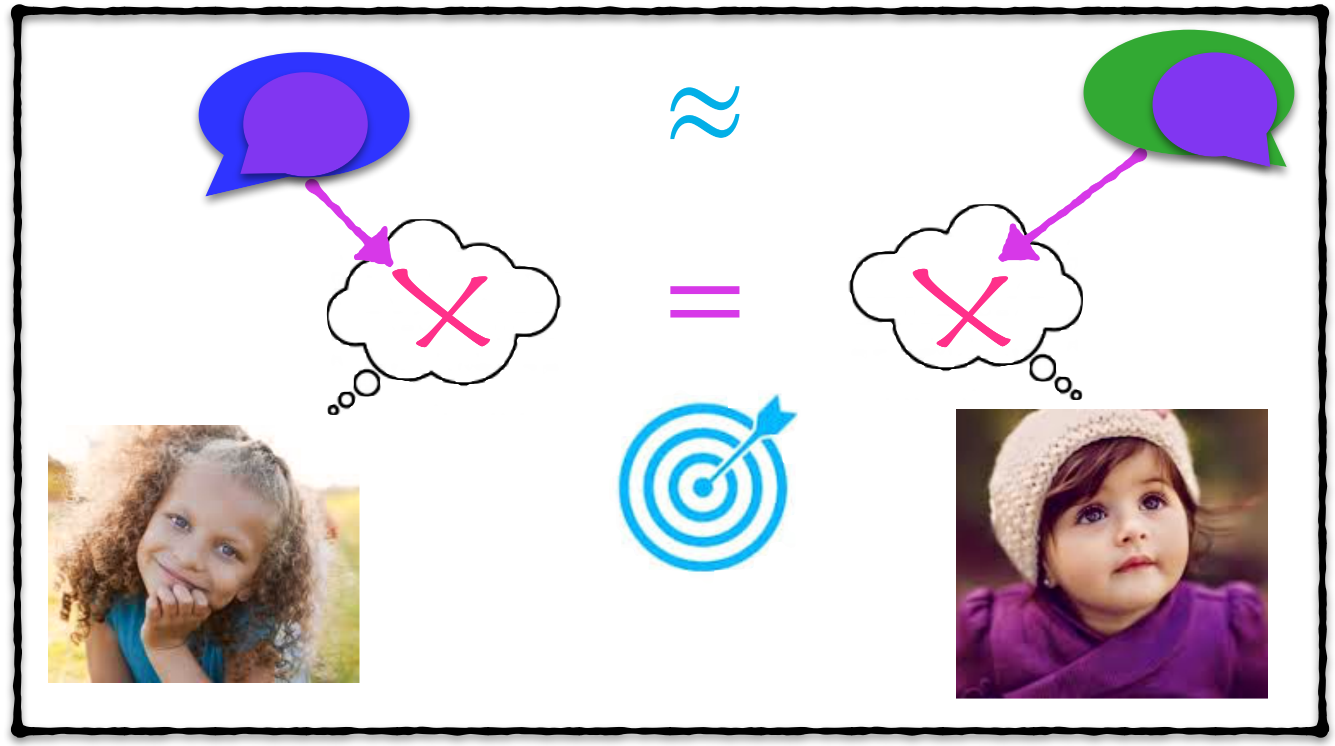
What do you think **that** $\_\__{what}$ happens?

Bates & Pearl 2019, submitted

Who does

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



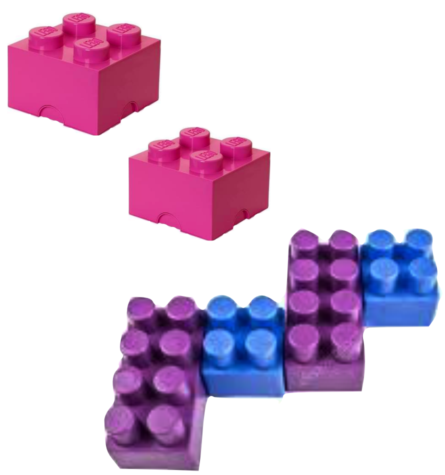


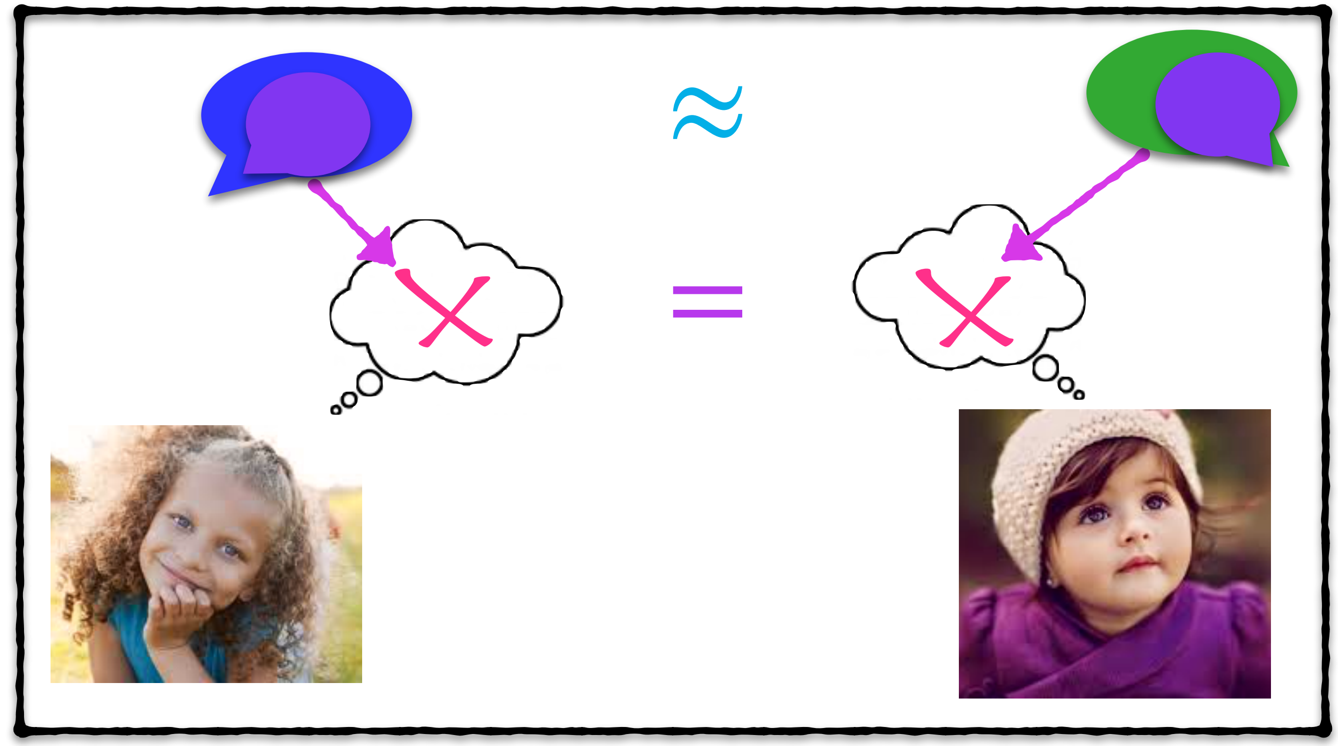
No meaningful input differences predicted

So now what?

Bates & Pearl 2019, submitted

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





No meaningful input differences predicted

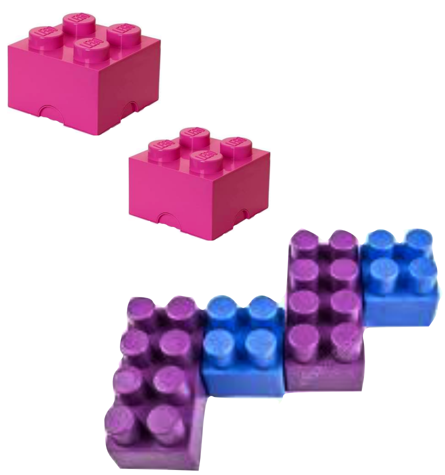
We should measure the learning outcomes in children across SES to see if in fact there are any learning outcome differences.

We already know there's no difference for Complex NP islands, despite measurable input differences.

Bates & Pearl 2019, submitted

Who does

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



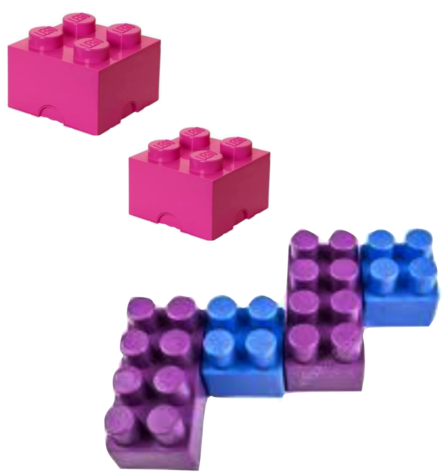
No meaningful input differences predicted

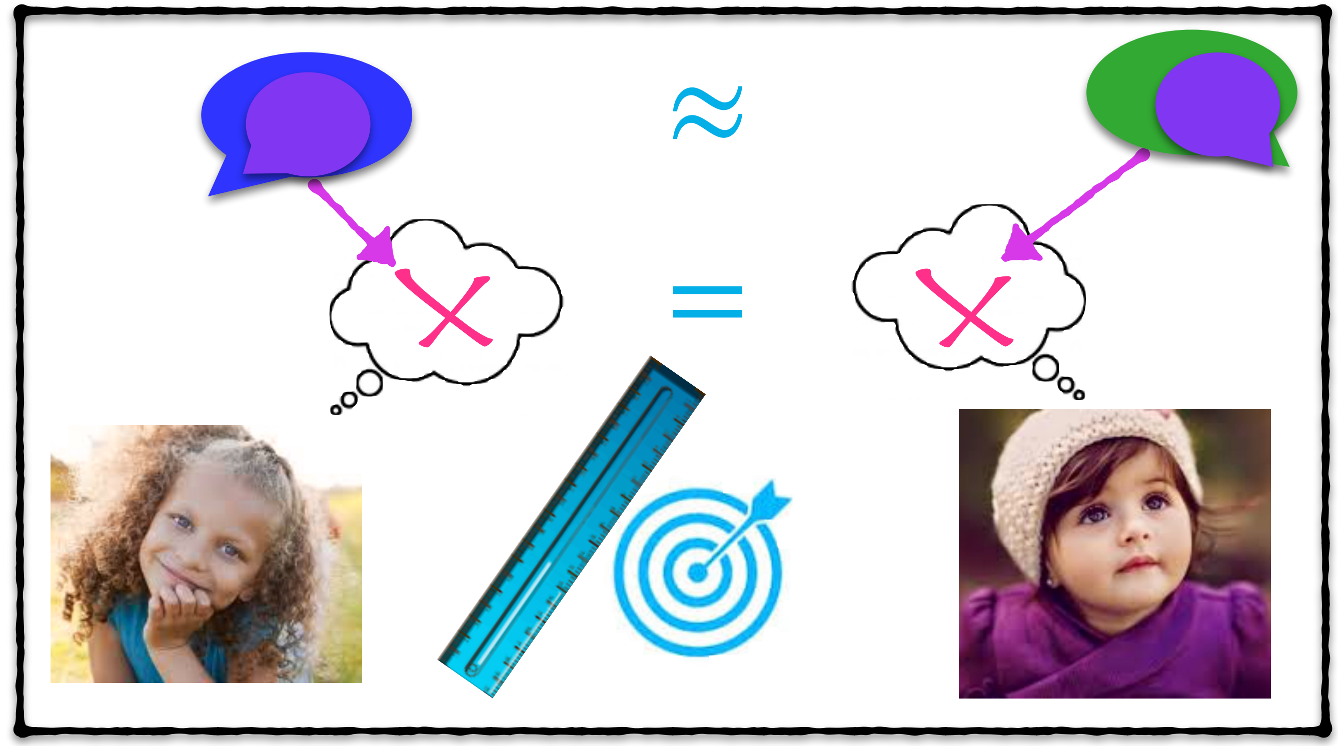
One caveat: If there are in fact differences, it could be due to other factors besides input differences.

Example factor: Language processing ability is known to differ across SES, with low-SES children sometimes slower compared to their high-SES counterparts (Fernald et al. 2013, Weisleder & Fernald 2013). If low-SES children are less able to harness the information in their input (even if it's there), they might be delayed in acquiring syntactic island knowledge.

Bates & Pearl 2019, submitted

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





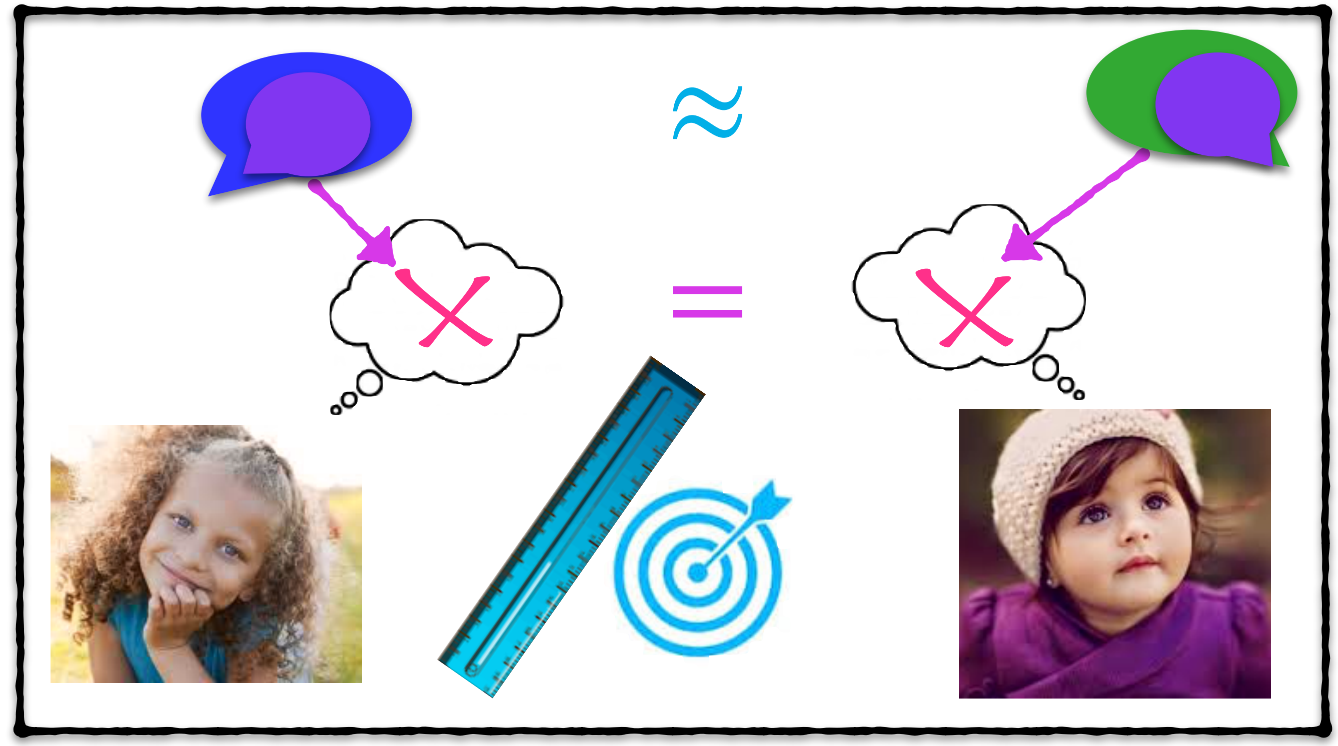
No meaningful input differences predicted

But, if there aren't outcome differences (perhaps after any language processing ability differences have resolved), then this supports syntactic island input quality being the same across SES (and also that the learning theory is plausible).

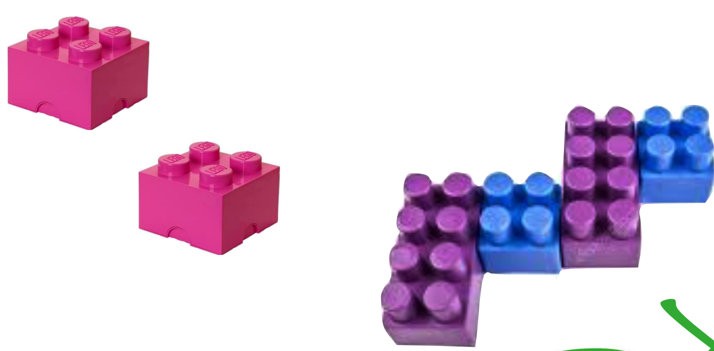
Bates & Pearl 2019, submitted

Who does

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

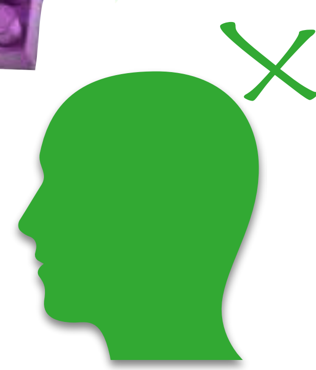


Building block origins



Low-SES
child-directed

What do you think **that** __$_{what}$ happens?

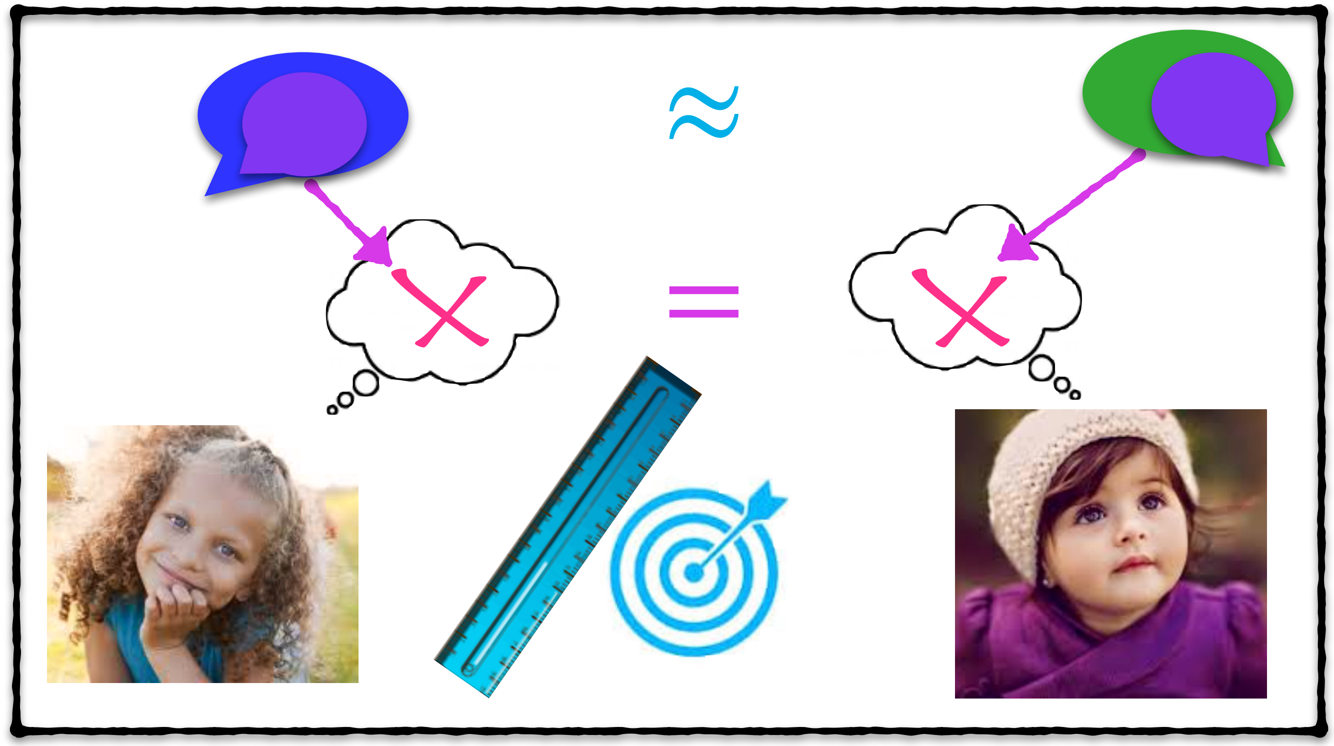


Remember that key building blocks involving CP$_{that}$ are predicted to come from a particular *wh*-dependency in low-SES child-directed speech that's ungrammatical in the high-SES dialect.

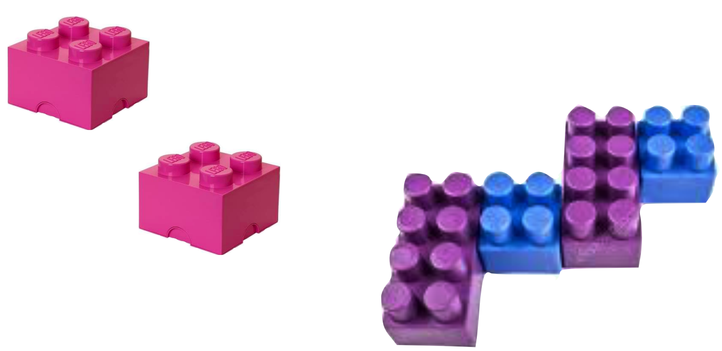
Bates & Pearl 2019, submitted

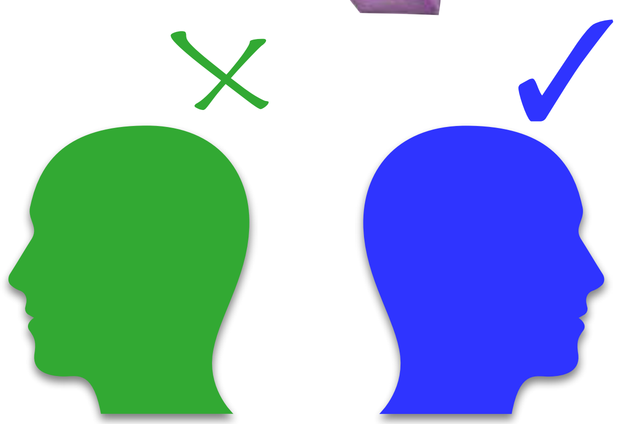
Who does

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



Building block origins

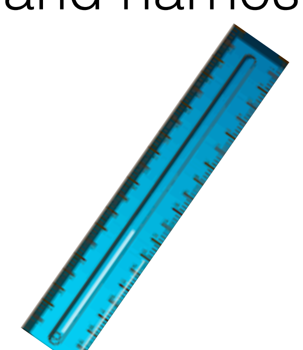




Low-SES
child-directed

What do you think **that** $\_\_{}_{what}$ happens?

This means low-SES adults are predicted to view this *wh*-dependency as grammatical if we expect low-SES children to hear it and harness those crucial CP_{that} building blocks from it.

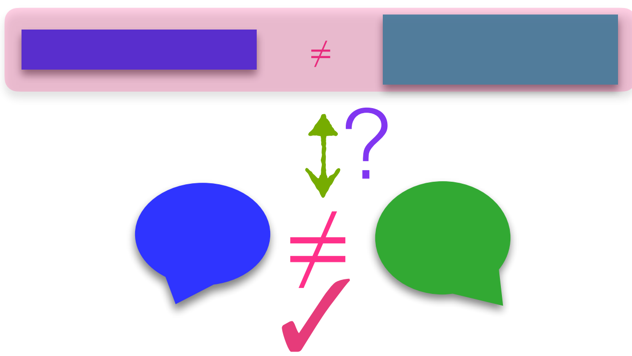


We can test this.

Bates & Pearl 2019, submitted

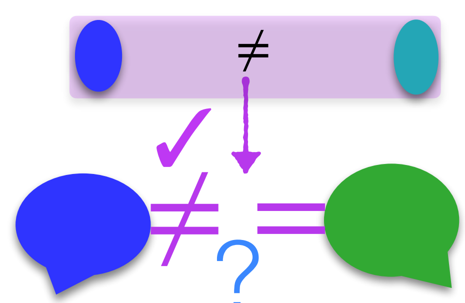
The big picture

One (standard) way



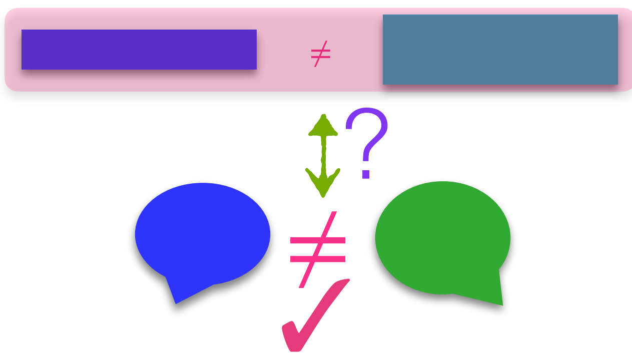
Developmental computational modeling complements existing techniques for identifying if input differences may be developmentally meaningful.

A new (complementary) way



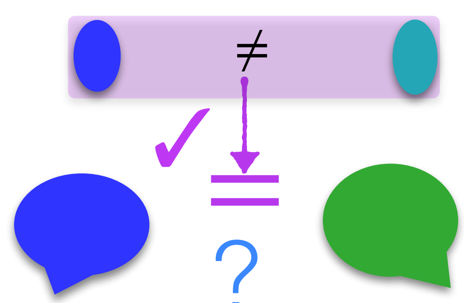
The big picture

One (standard) way



Developmental computational modeling complements existing techniques for identifying if input differences may be developmentally meaningful.

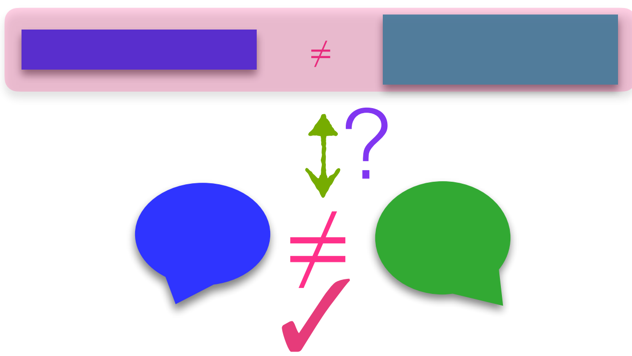
A new (complementary) way



We demonstrated this approach for syntactic island knowledge, and predicted no meaningful input differences across SES for knowledge by age four.

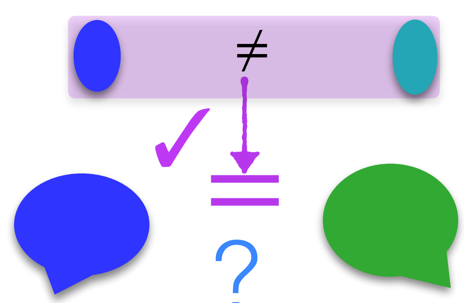
The big picture

One (standard) way



Developmental computational modeling complements existing techniques for identifying if input differences may be developmentally meaningful.

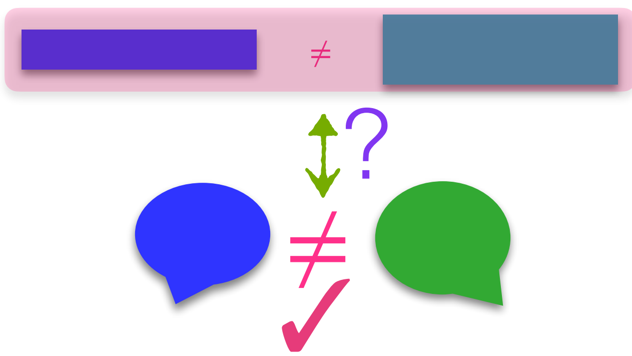
A new (complementary) way



This means we predict that input-based interventions wouldn't be impactful if there actually are any differences in the acquisition of these syntactic islands across SES.

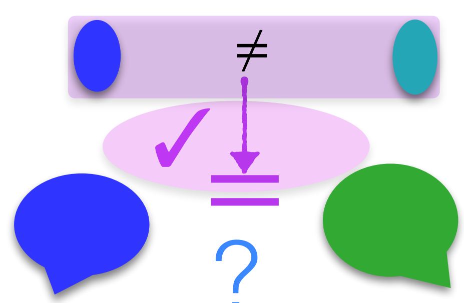
The big picture

One (standard) way



Something useful: This technique can provide a causal explanation (using the learning theory implemented in the model) for how input differences could affect learning outcomes.

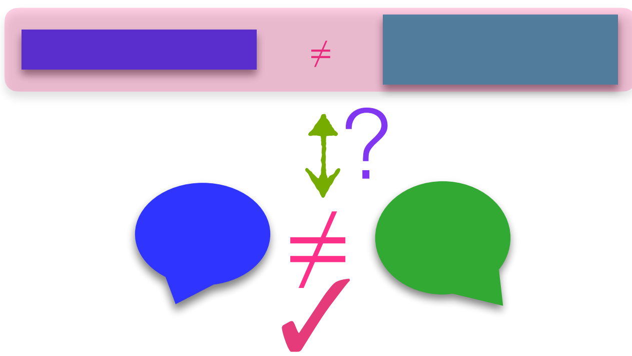
A new (complementary) way



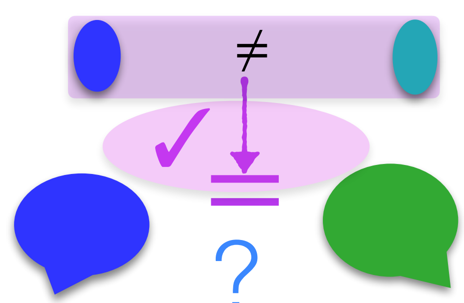
The big picture

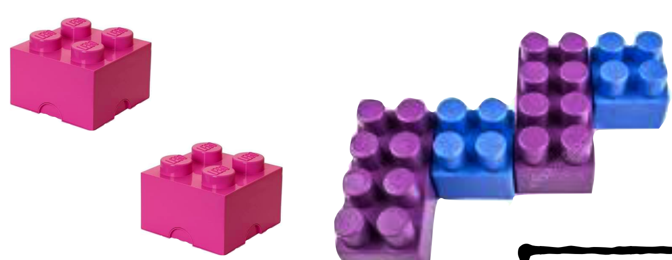
Something useful: This technique can provide a causal explanation (using the learning theory implemented in the model) for how input differences could affect learning outcomes.

One (standard) way



A new (complementary) way



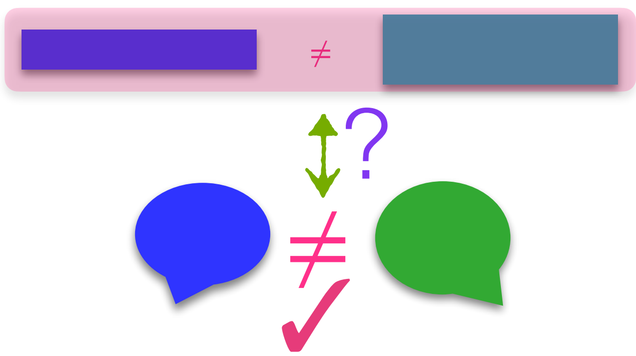


For syntactic islands, the building blocks needed for this knowledge don't seem to differ enough across SES to matter by age four.

Who does

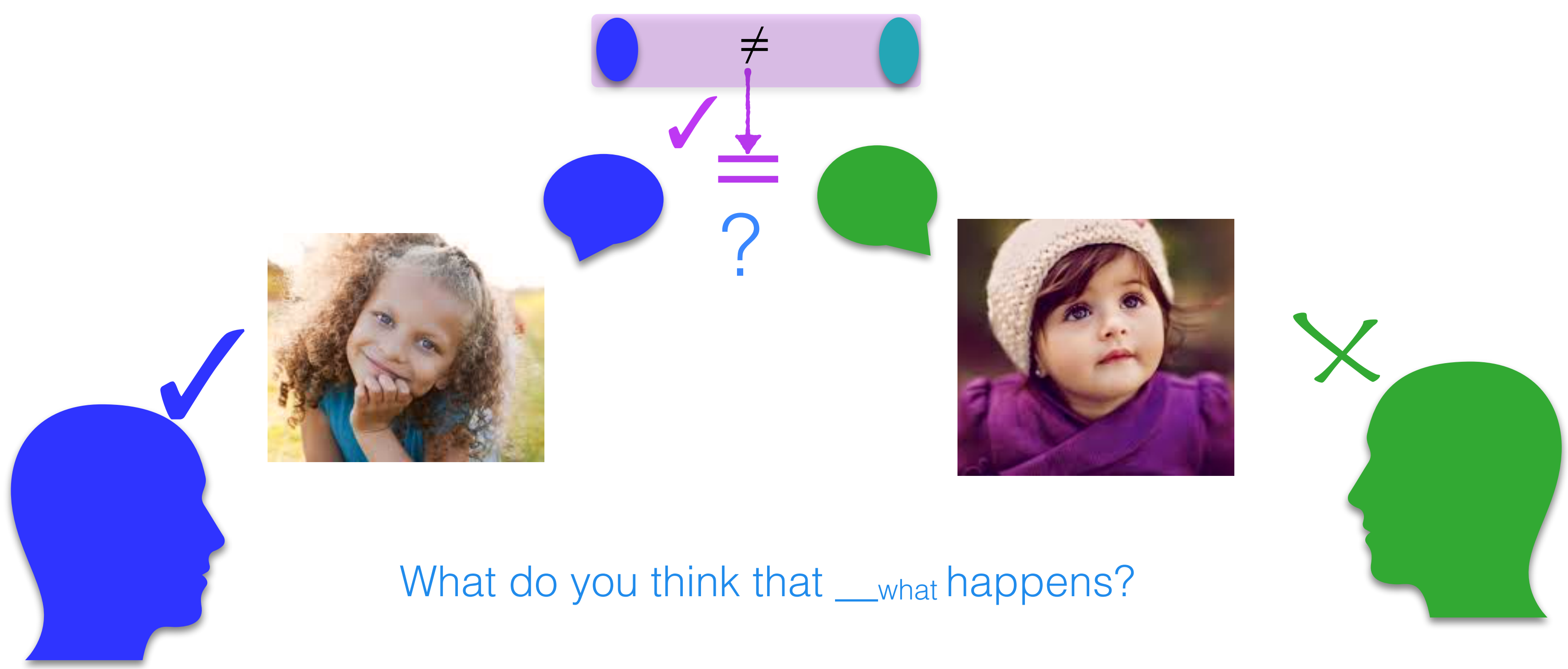
The big picture

One (standard) way



Something else useful: This technique can make predictions about differences we might expect in both child outcomes and eventual adult knowledge.

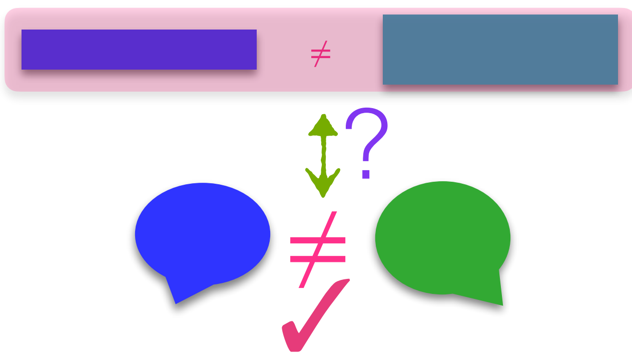
A new (complementary) way



What do you think that ___$_{\text{what}}$ happens?

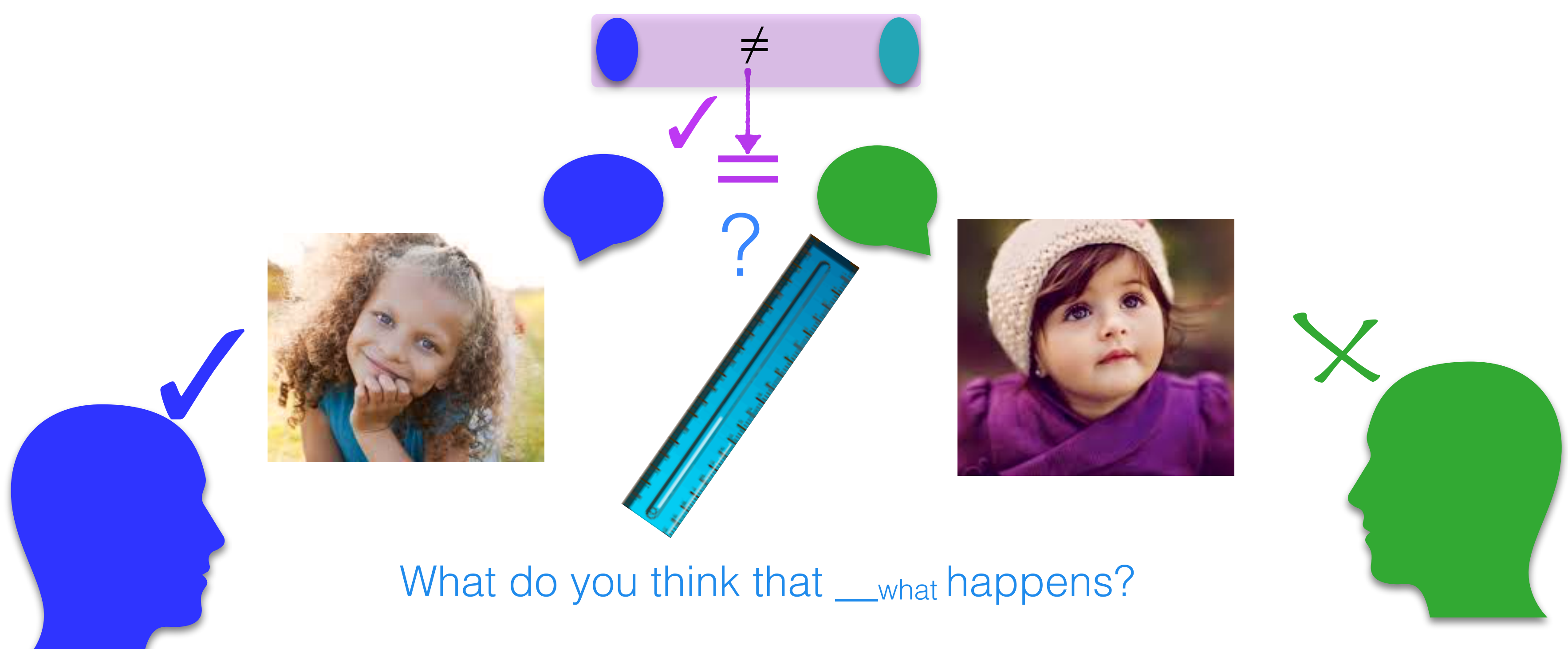
The big picture

One (standard) way



Something important: Any predicted differences still need to be measured to see if they're actually present. But at least we have a better idea what to look for.

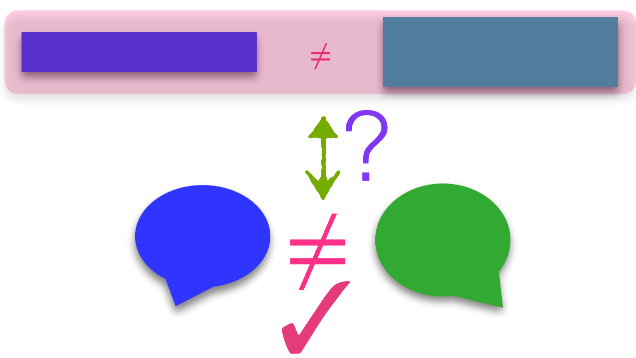
A new (complementary) way



What do you think that __what happens?

The big picture

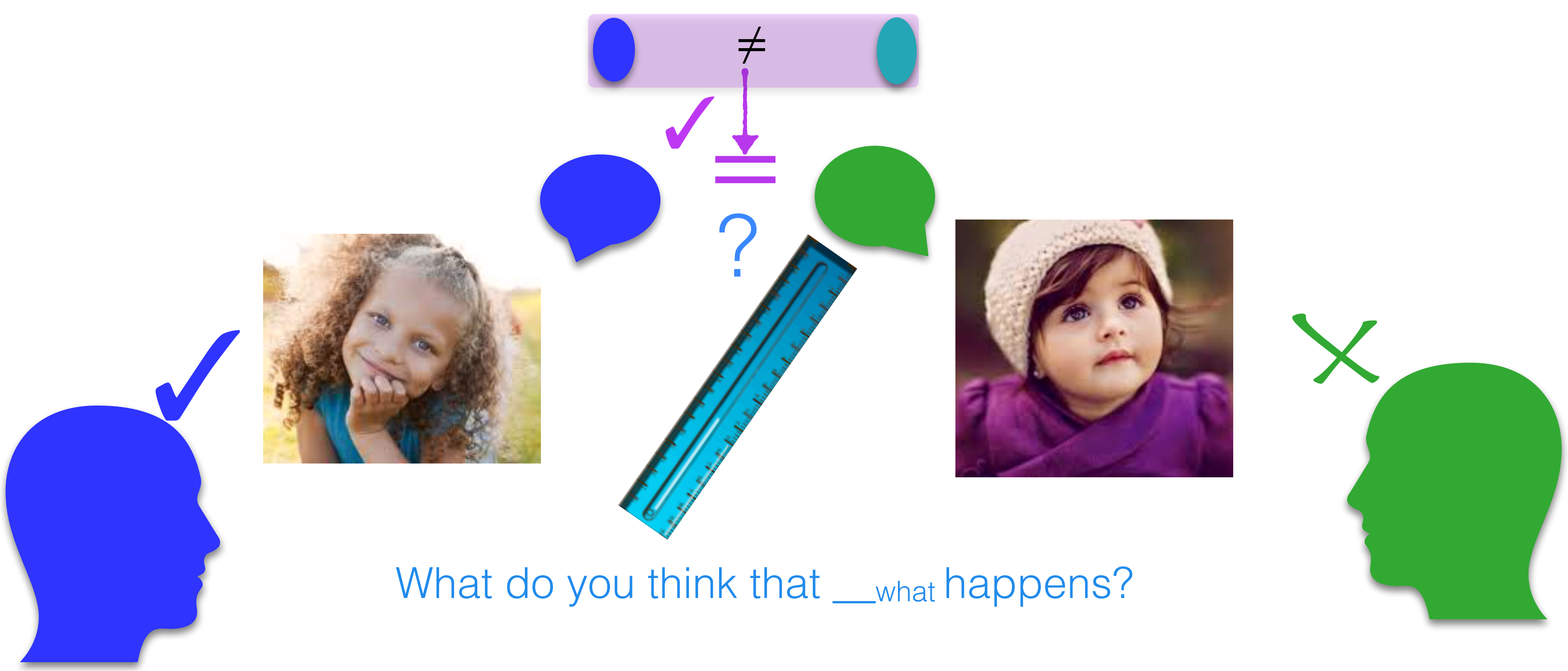
One (standard) way

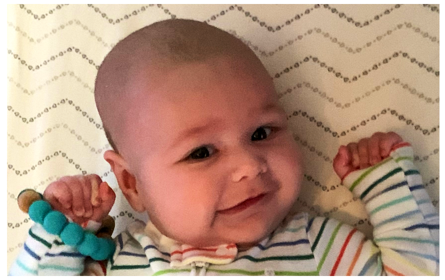


Bonus: Modeling is often faster (and cheaper to do) than behavioral work. So it can be very useful as a first-pass input-quality assessment.

Extra bonus: Possible to do in pandemic times.

A new (complementary) way

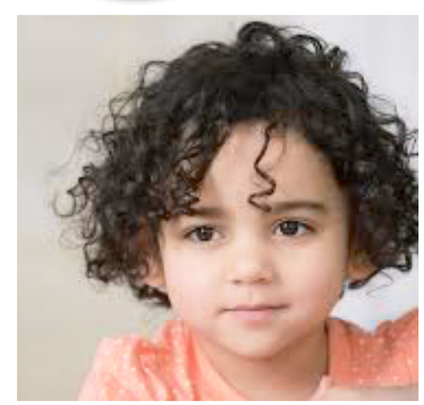
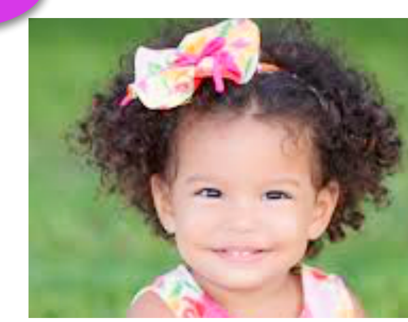



So let's use developmental computational modeling when we want to identify and understand if input variation could be developmentally meaningful!





Thank you!

Alandi Bates

BUCLD 2018

UCSD Linguistics 2020

ForMA Group 2020

UMD Linguistics 2020

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