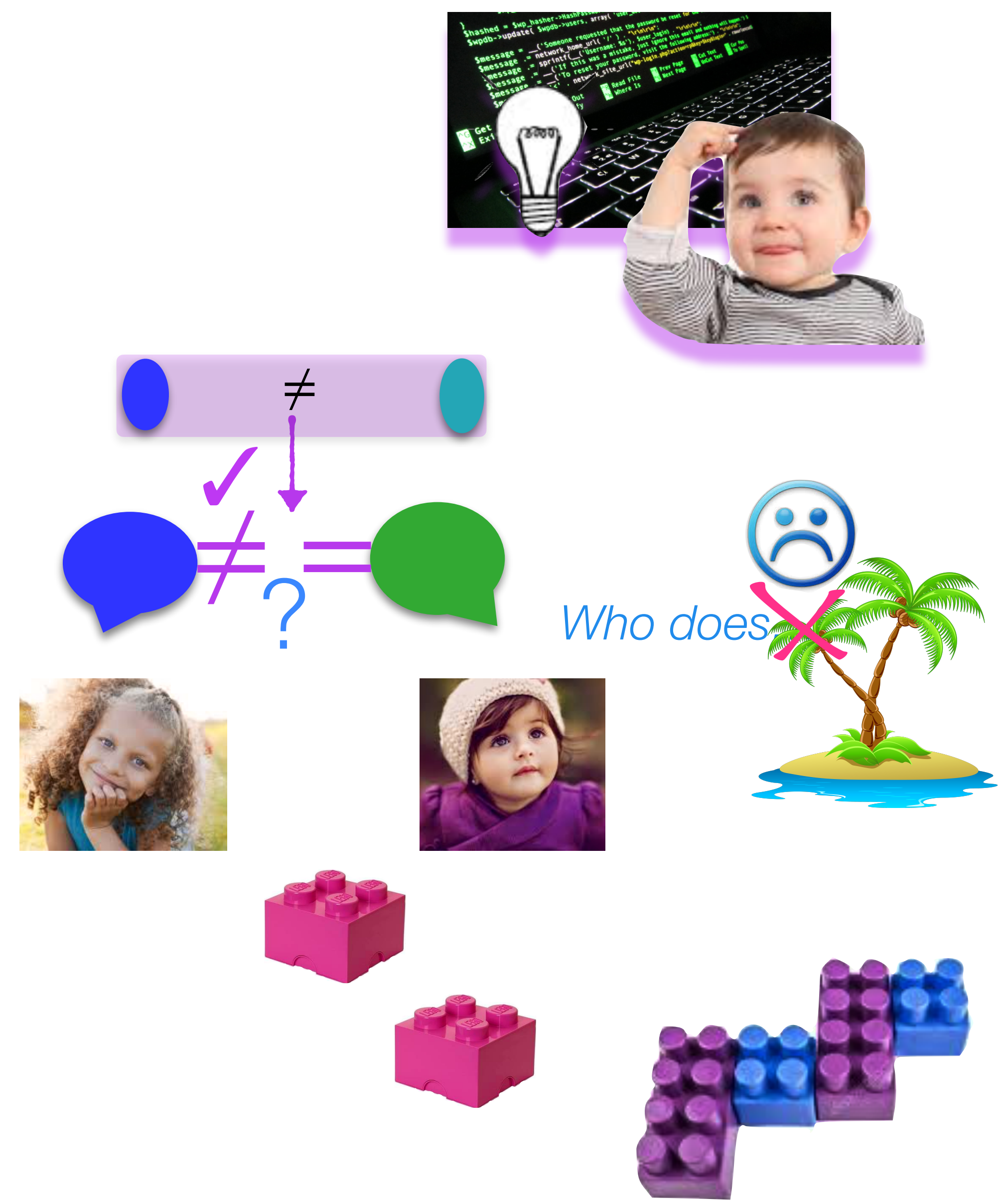
A new way to find meaningful variation in children's input across socio-economic status

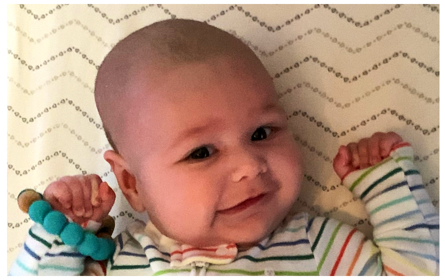
Lisa Pearl
University of California, Irvine

Computation of Language Laboratory
UC Irvine

Lisa S. Pearl
Professor
Department of Language Science
Department of Cognitive Sciences
SSPB 2219, SBSG 2314
University of California, Irvine
lpearl@uci.edu

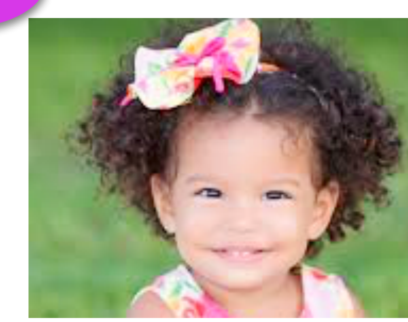
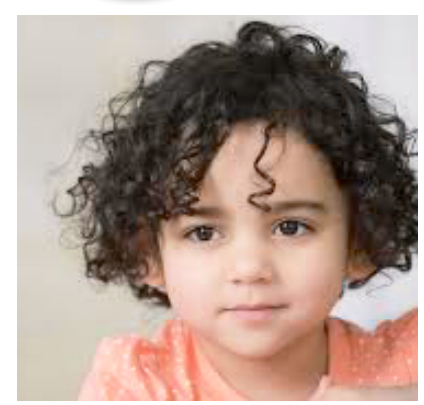
May 18, 2020
Linguistics Colloquium
University of California, San Diego

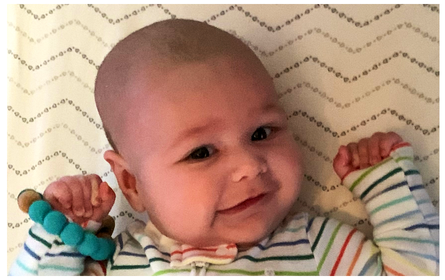



There's lots of variation in children's input



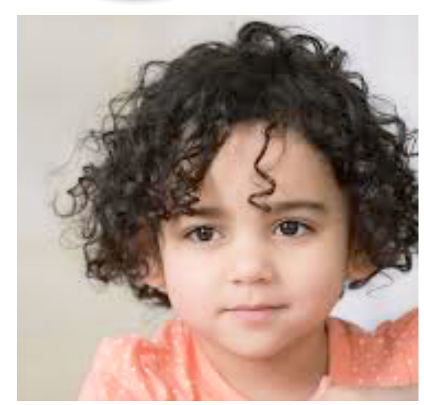
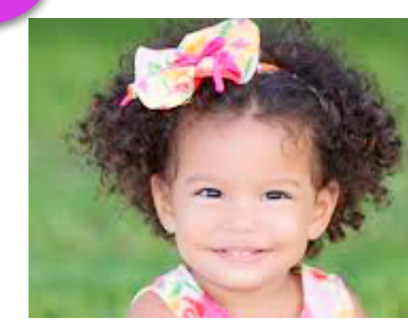



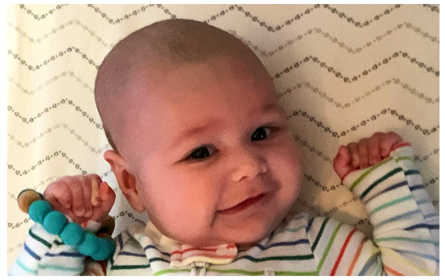




Meaningful variation impacts language development



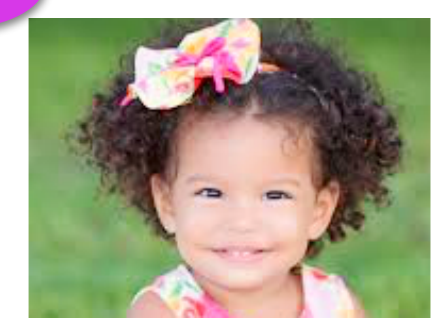


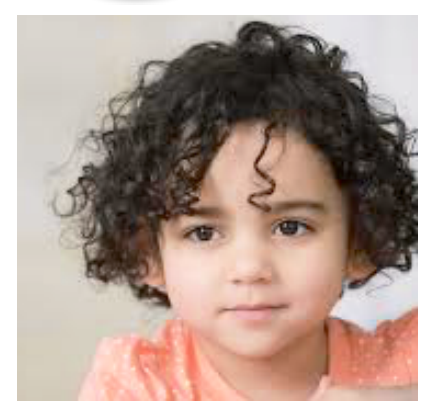




Meaningful input deficits can lead to language delays








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).

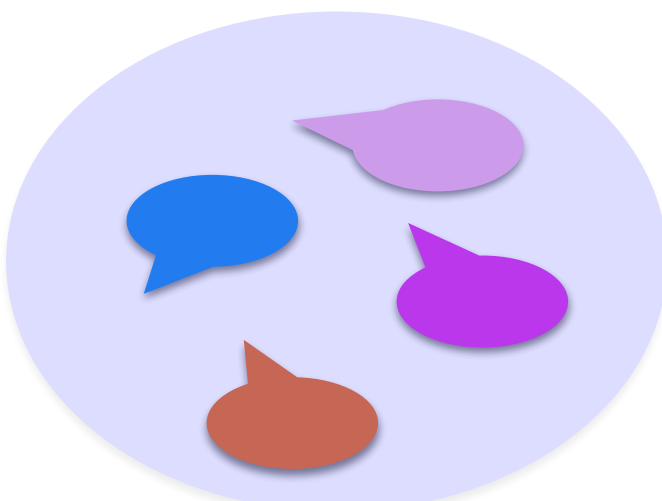
Quality can be measured by different aspects of the input, like diversity of vocabulary …

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

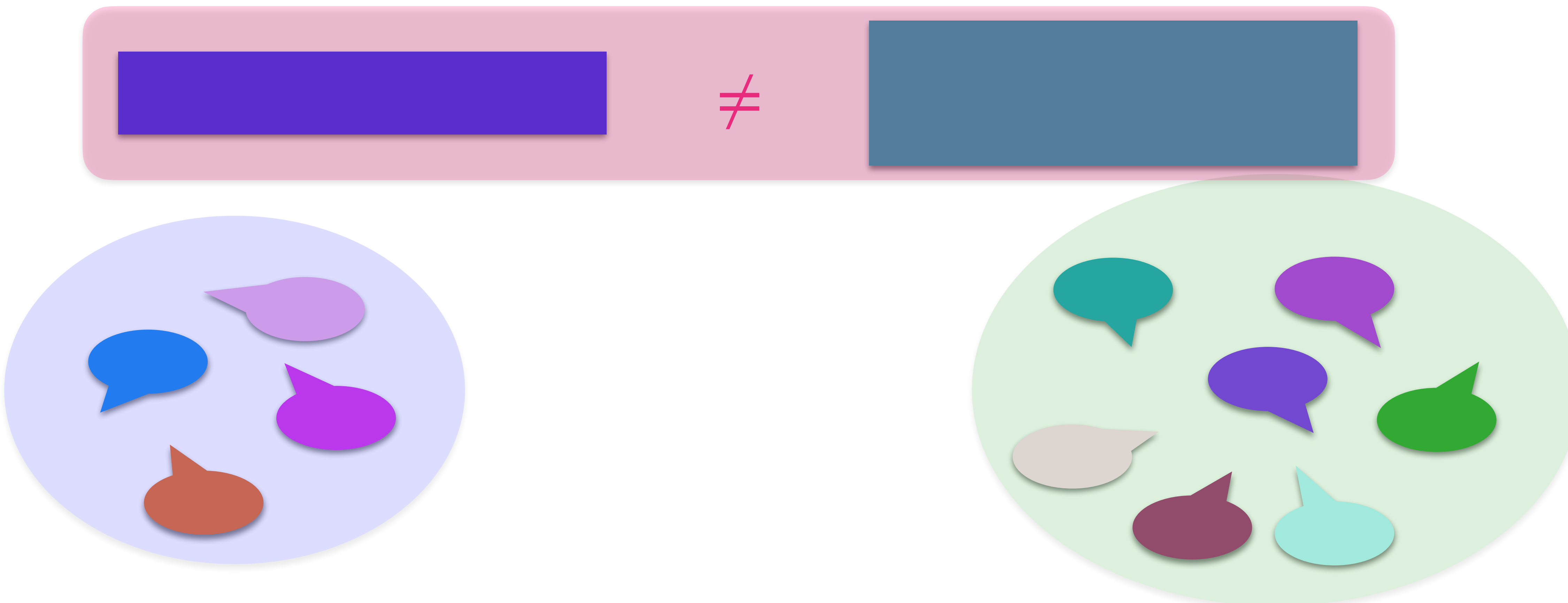
Because we're going tomorrow

We'll see the kitty on Friday

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

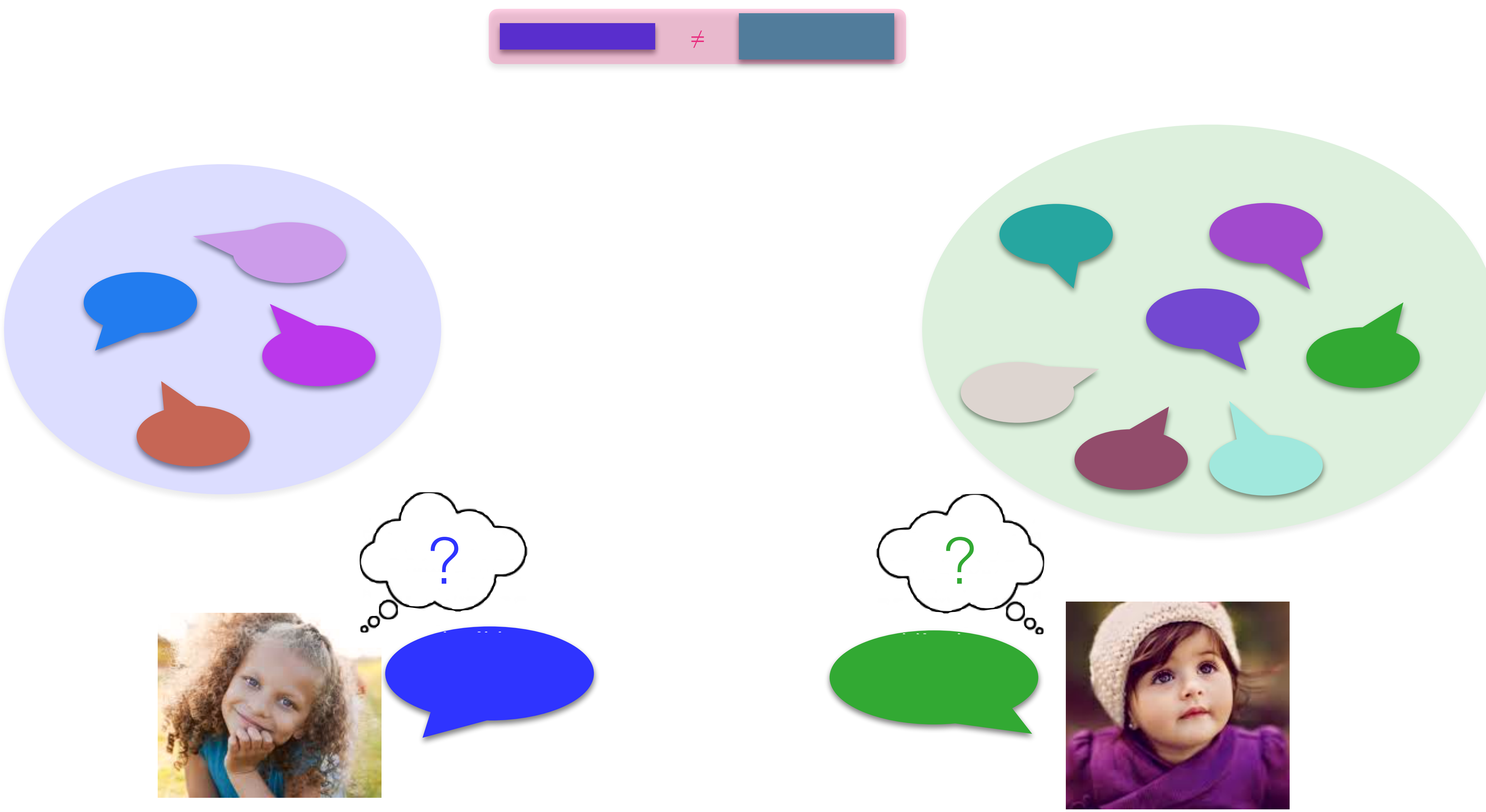
One (standard) way:

- Notice that there's a difference



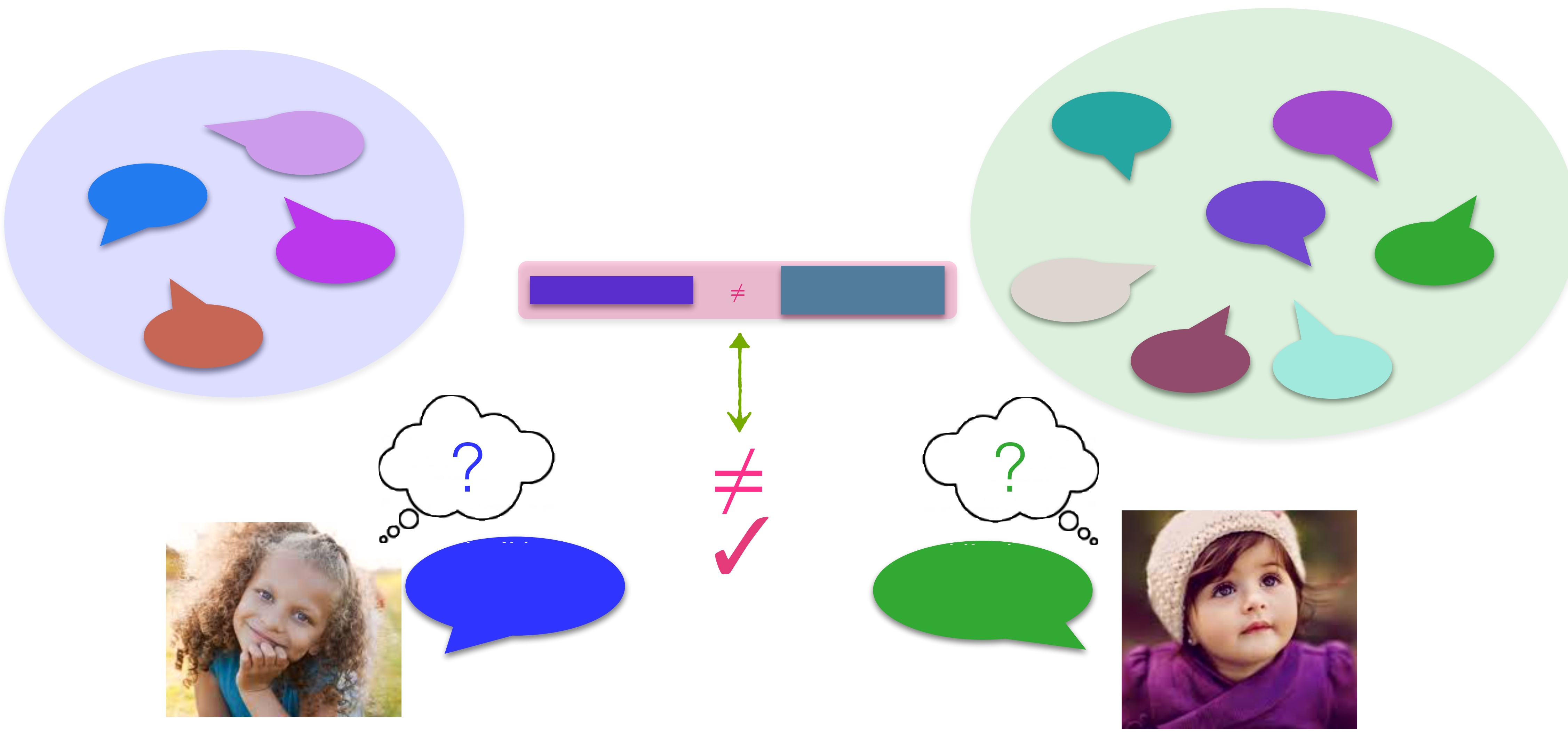
One (standard) way:

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



One (standard) way:

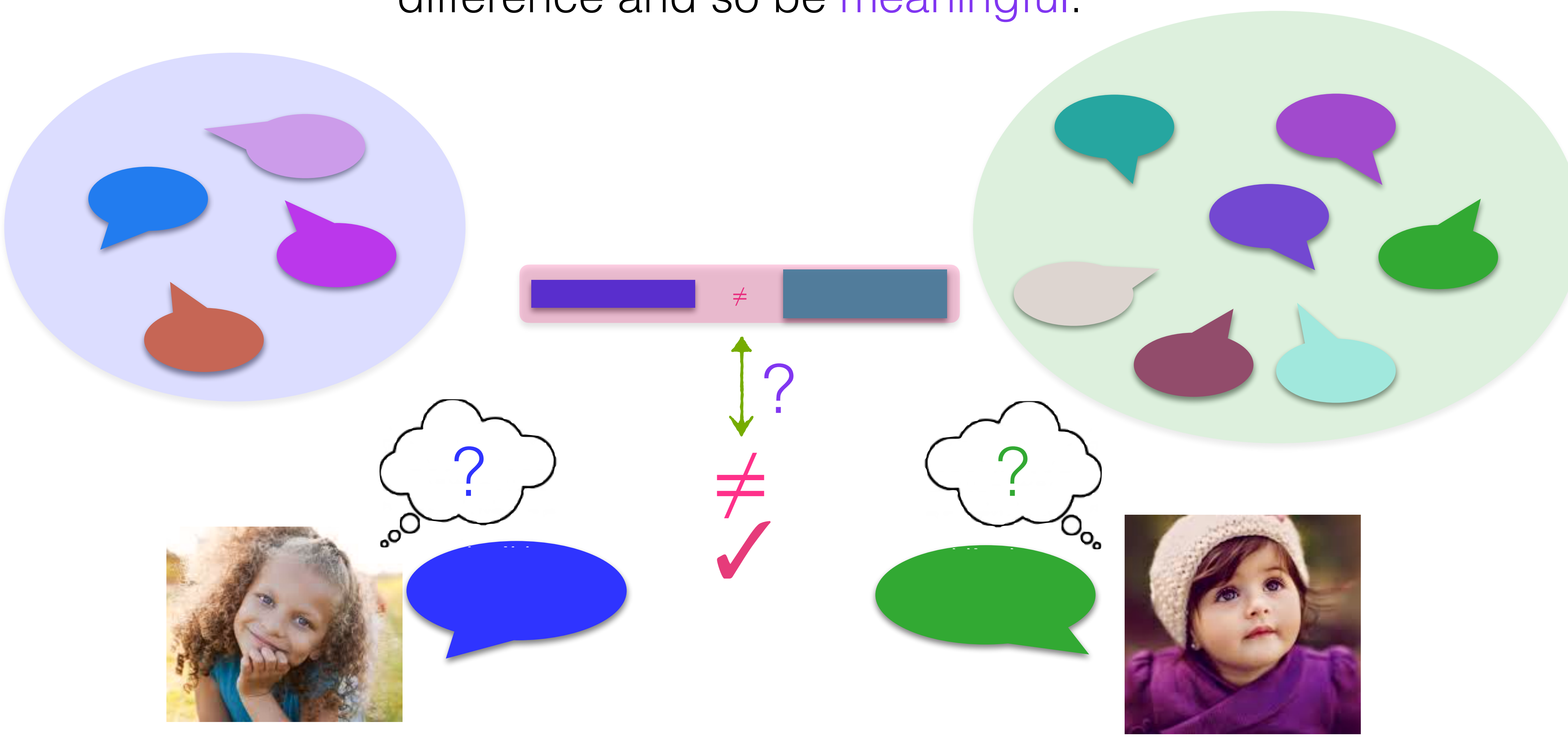
- Notice that there's a difference
- Measure language acquisition outcomes
- See if that input difference correlates with any outcome differences



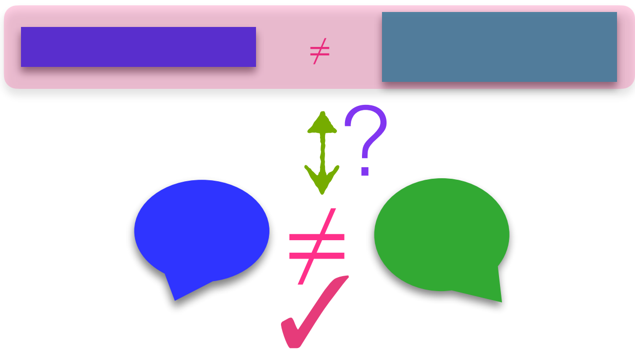
One (standard) way:

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

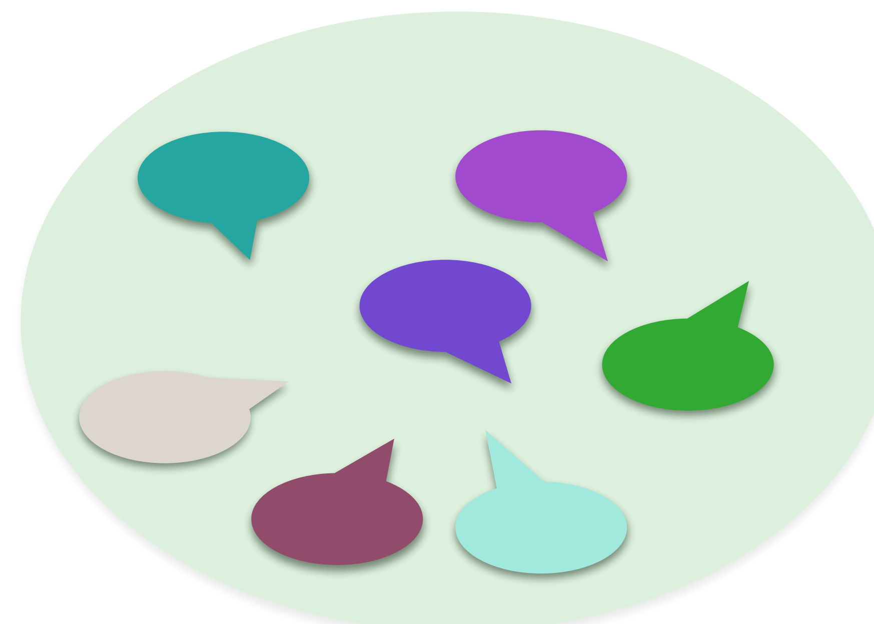
If so, then the input difference *might* cause the outcome difference and so be meaningful.



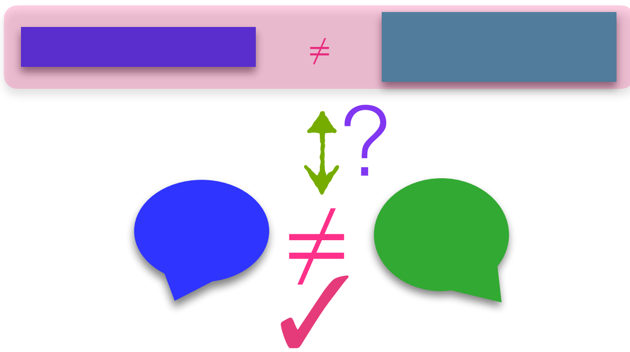
One (standard) way



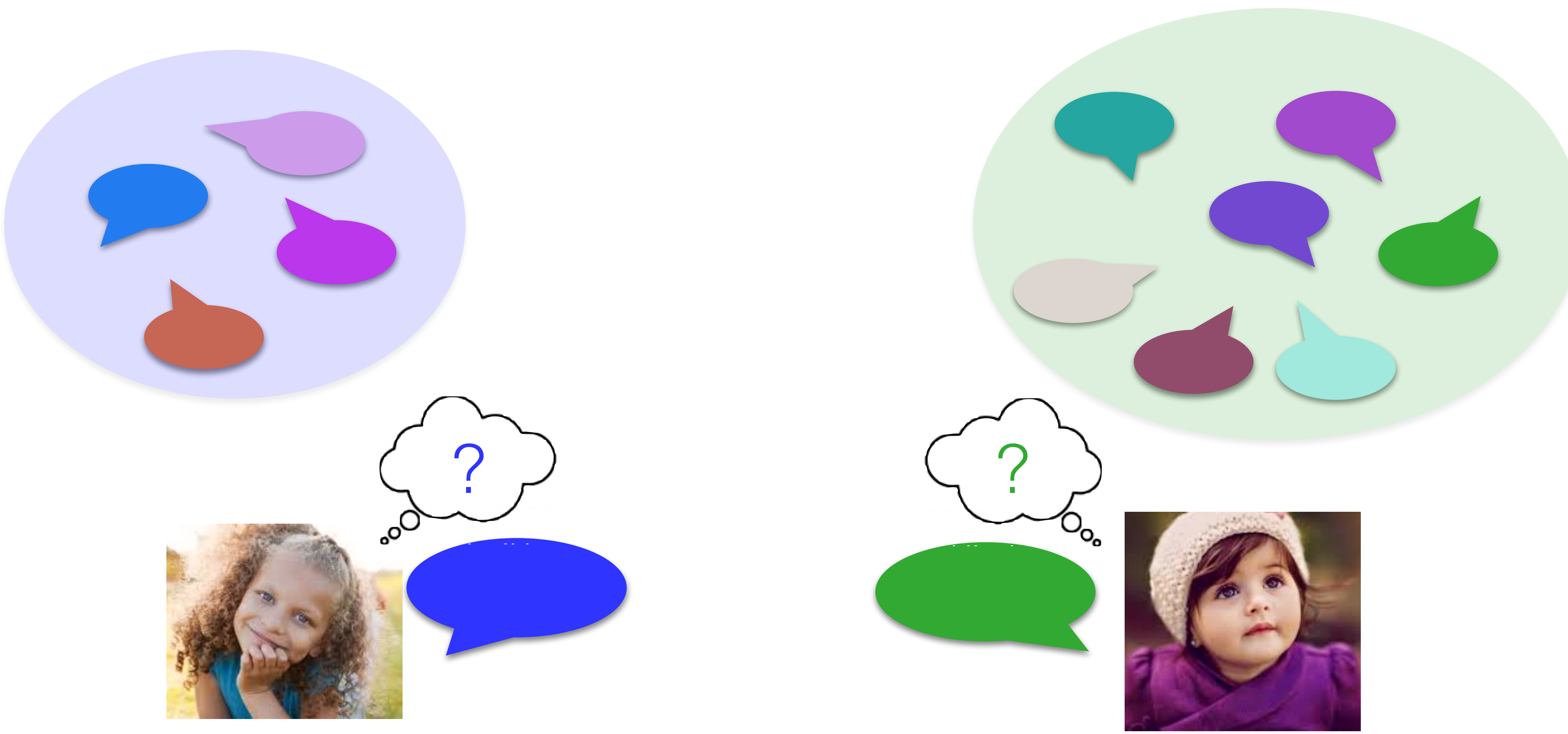
A new (complementary) way uses developmental computational modeling.



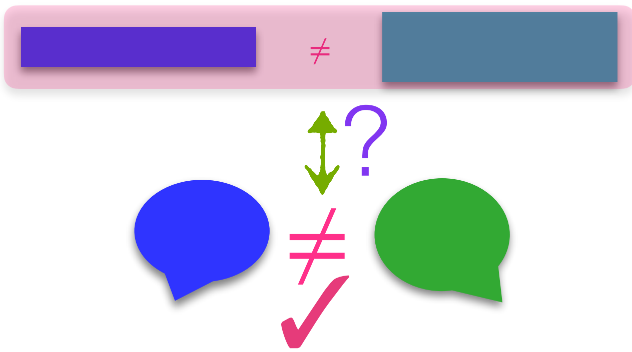
One (standard) way



A developmental computational model implements a specific learning theory …

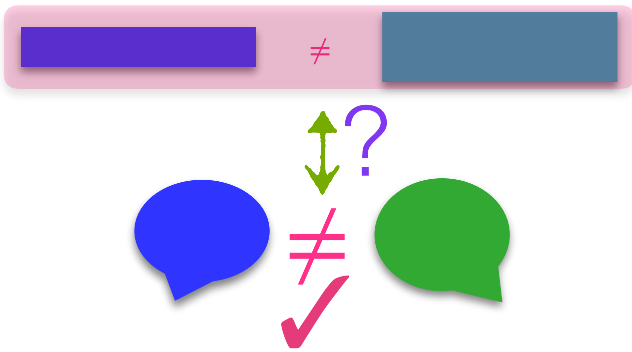


One (standard) way

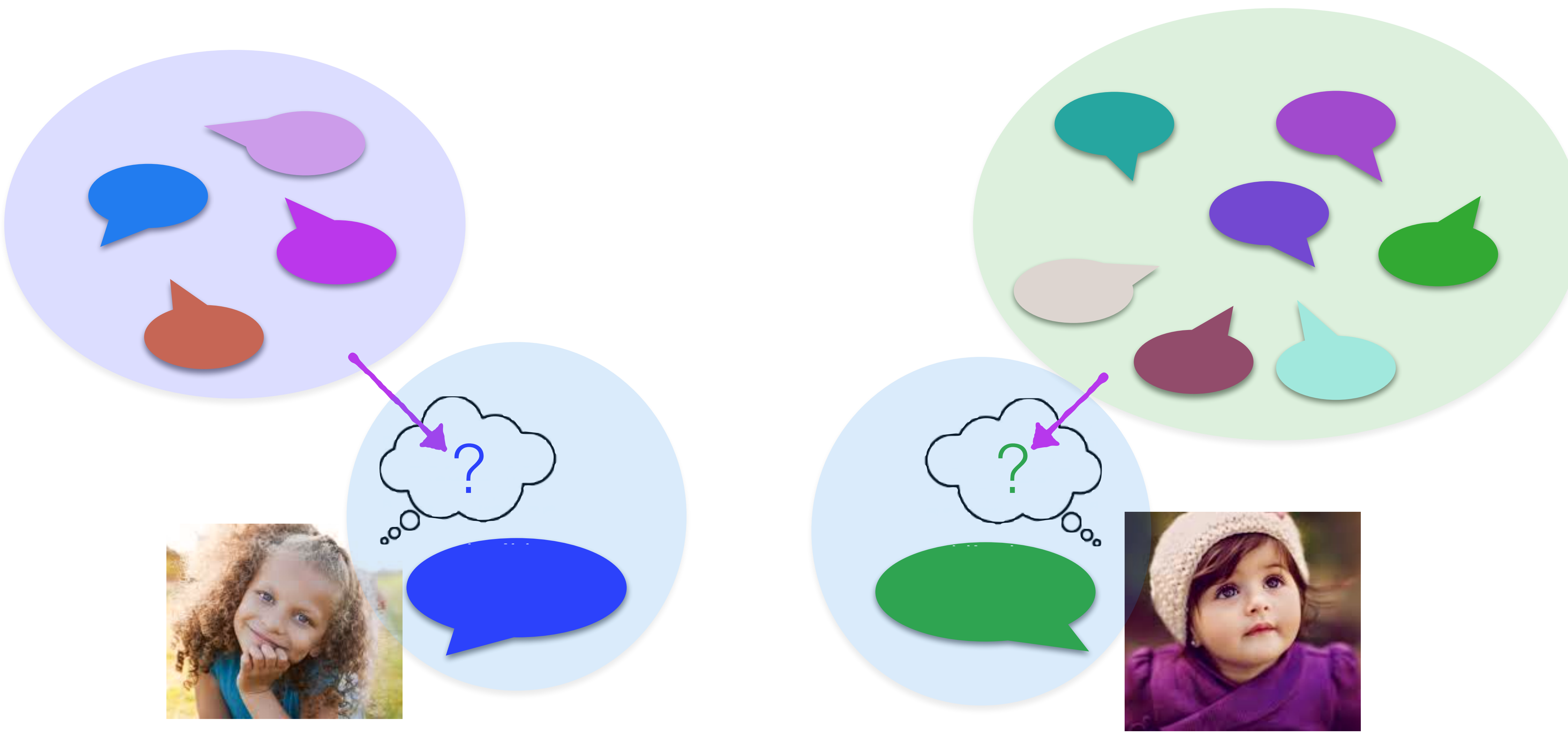


A developmental computational model implements a specific learning theory about how children use their input …

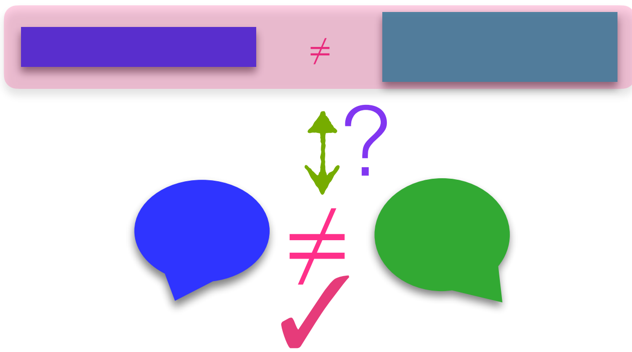
One (standard) way



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

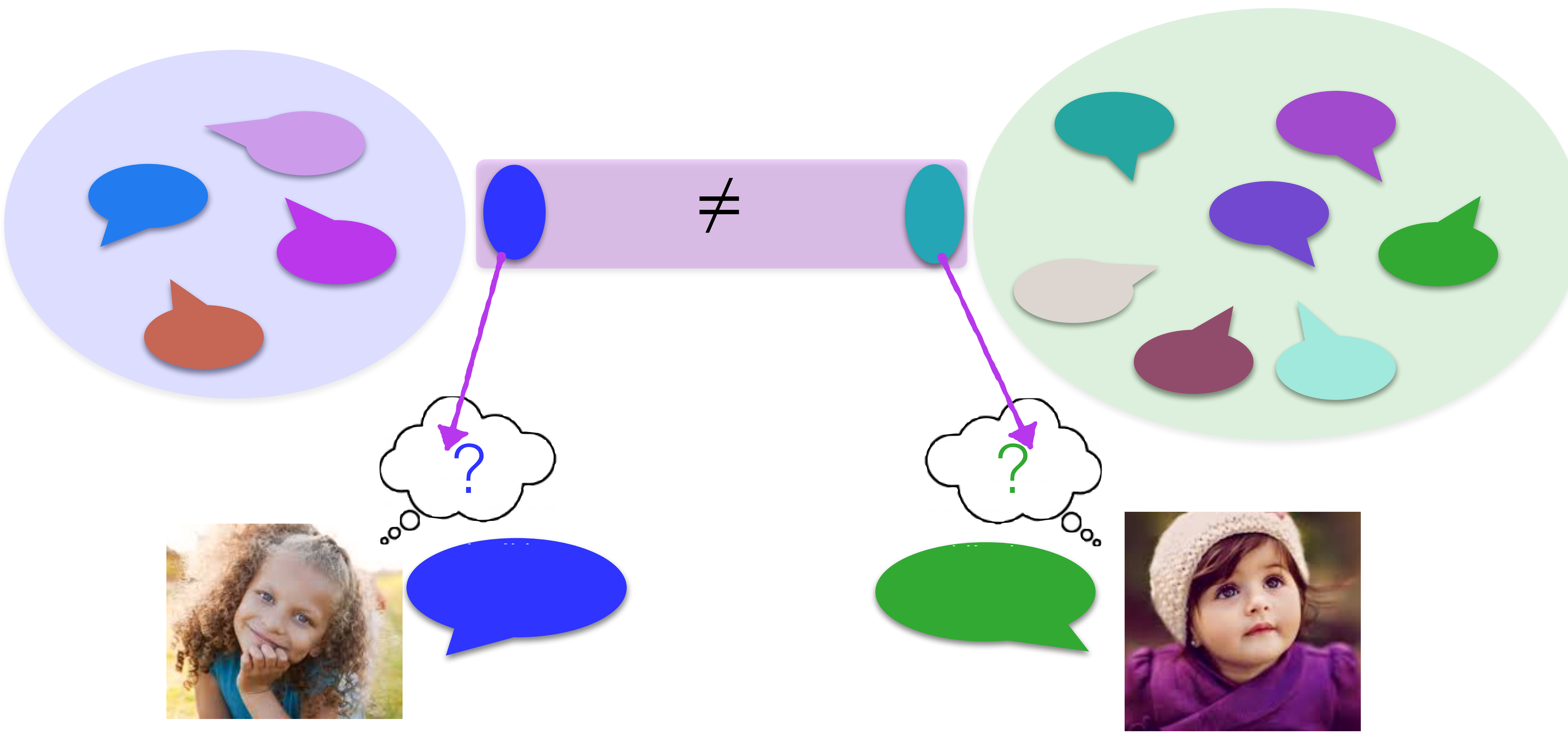


One (standard) way

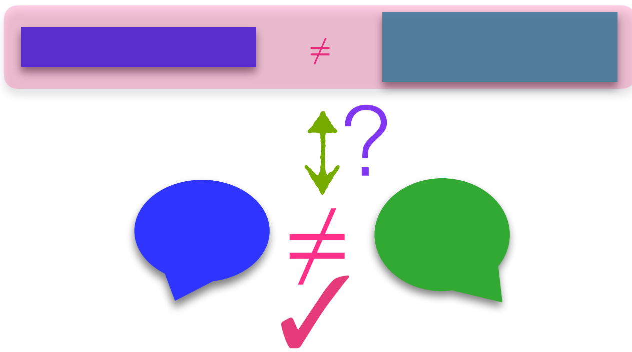


A developmental computational model implements a specific learning theory about how children use their input to acquire the knowledge to generate their output.

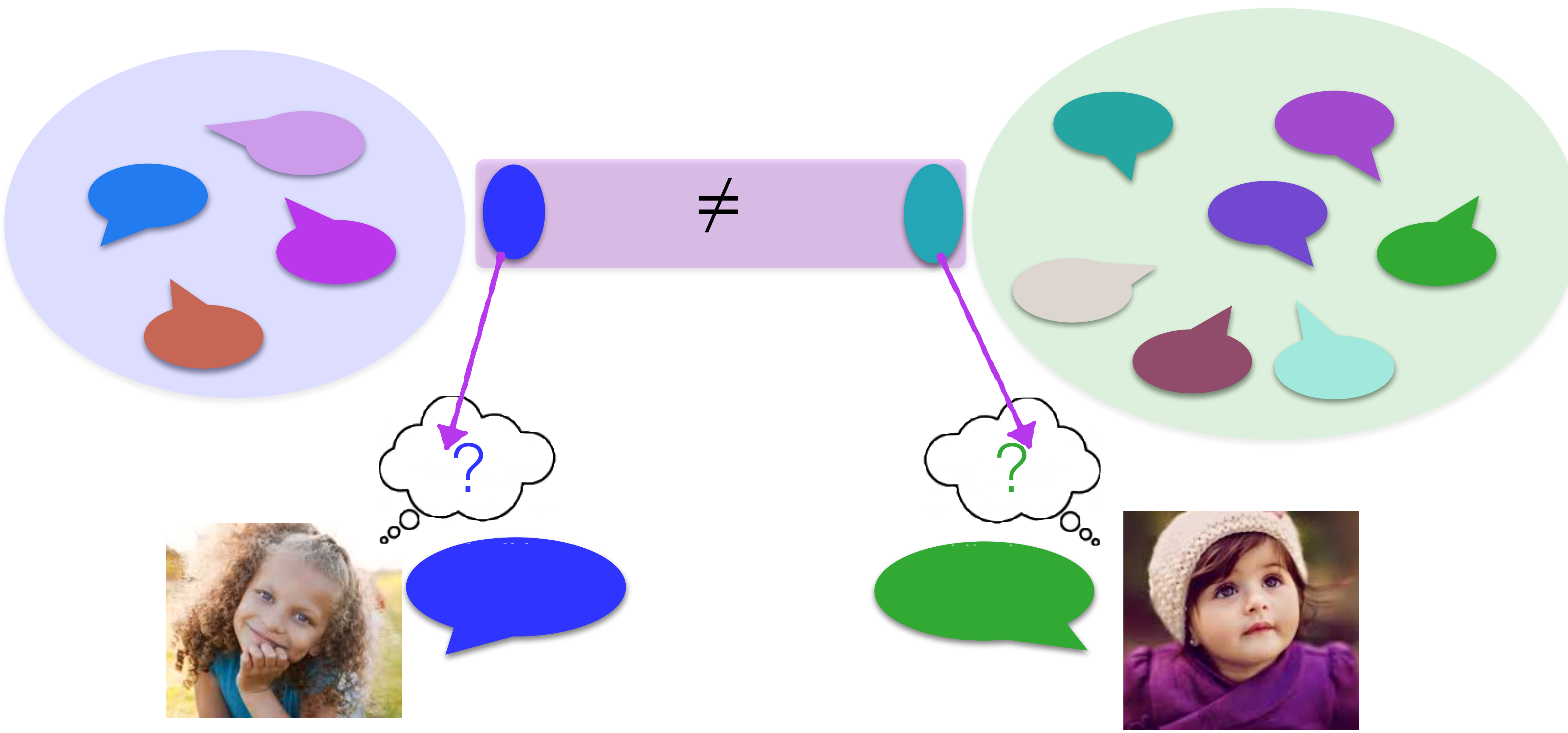
Important: the learning theory implemented by the model specifies what aspect of the input matters.



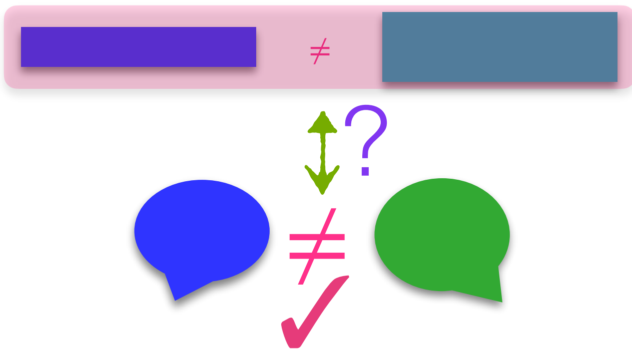
One (standard) way



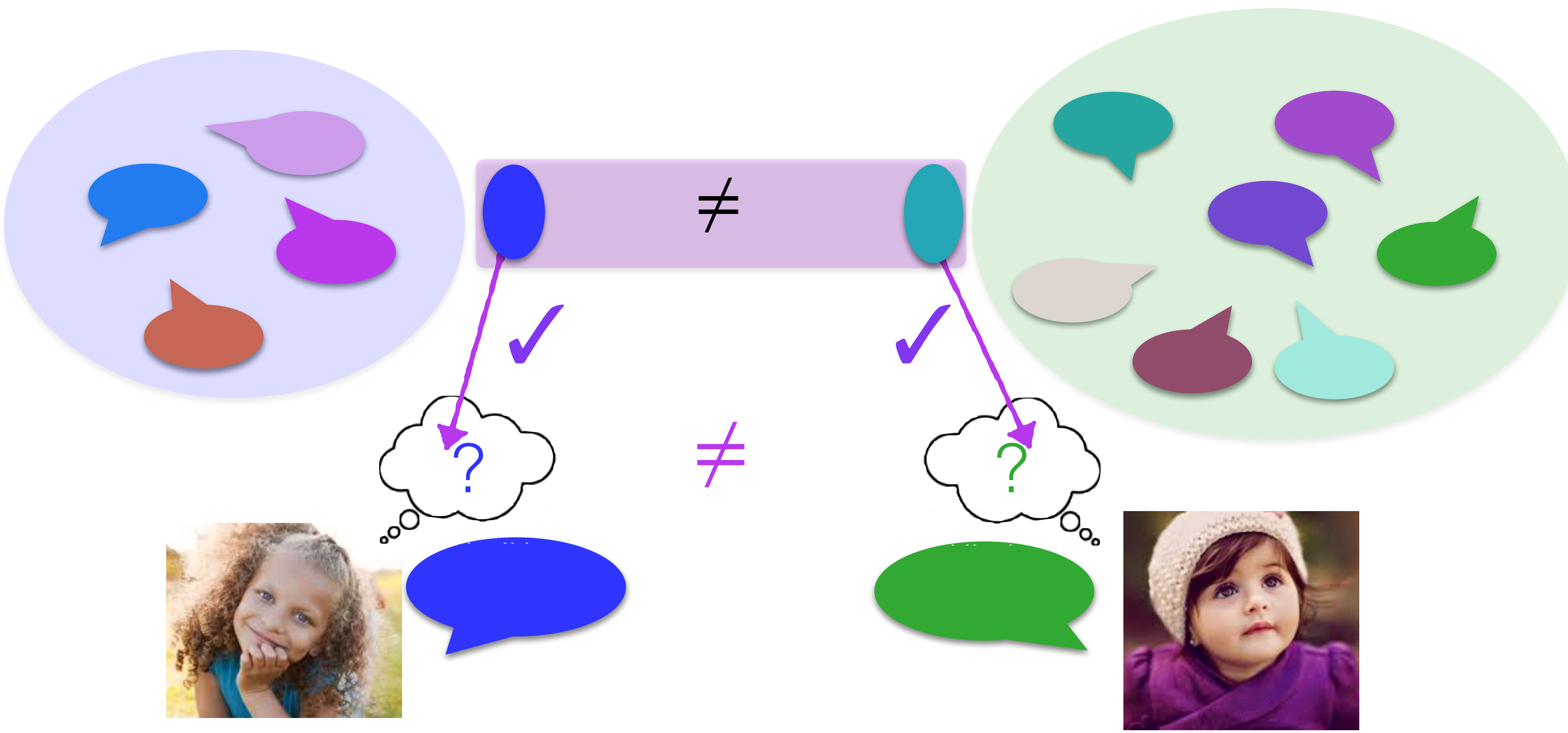
So, a developmental computational model can predict the language outcome on the basis of the input.



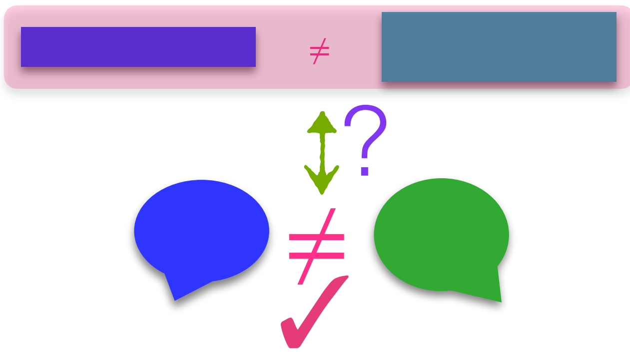
One (standard) way



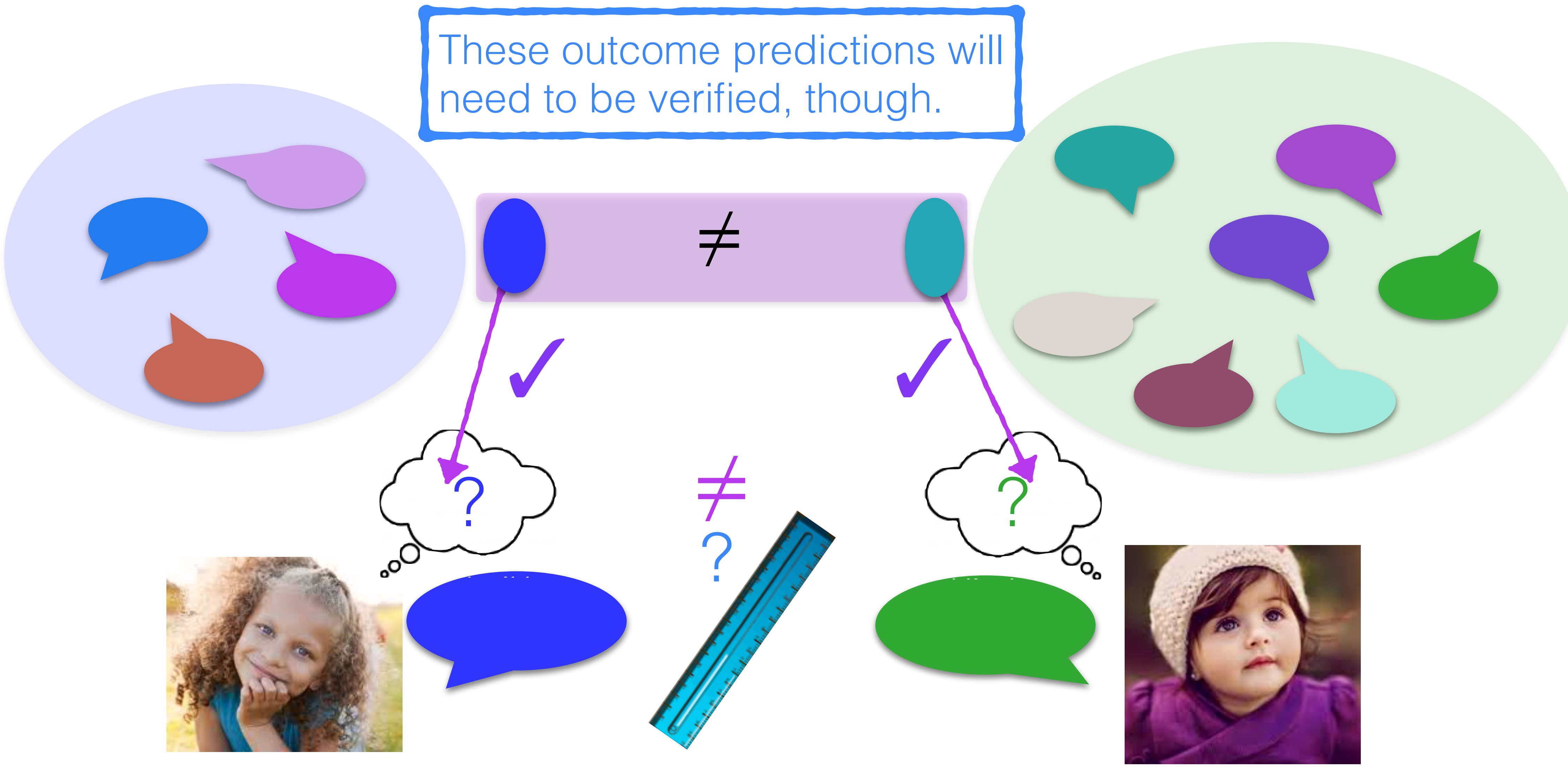
If the predicted outcomes differ, then it's because the input difference caused that outcome difference. So, the input difference is predicted to be meaningful.



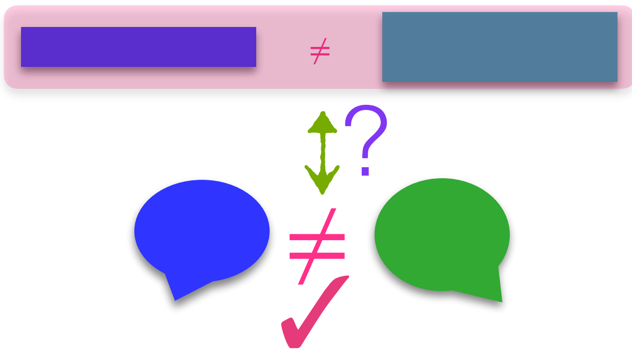
One (standard) way



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

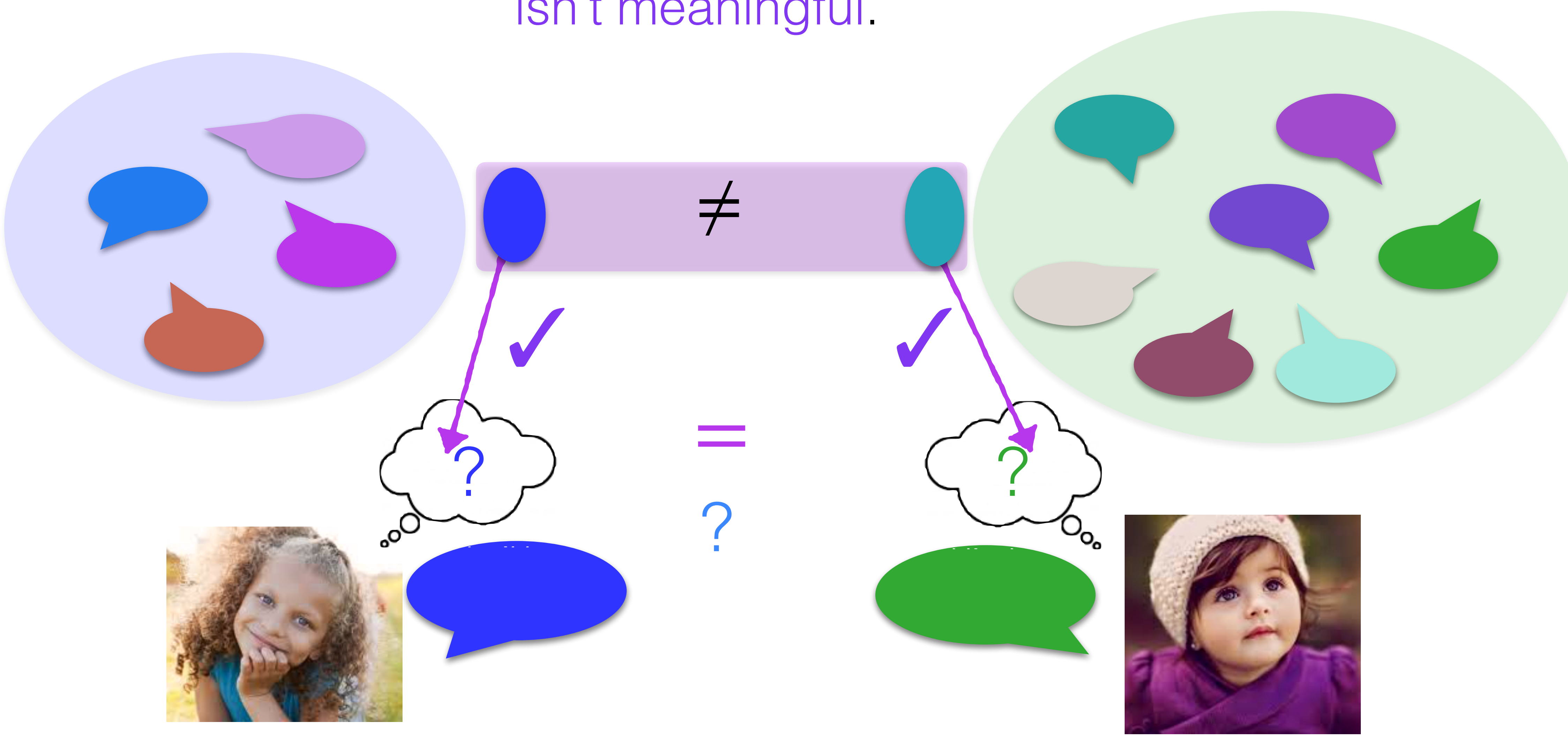


One (standard) way



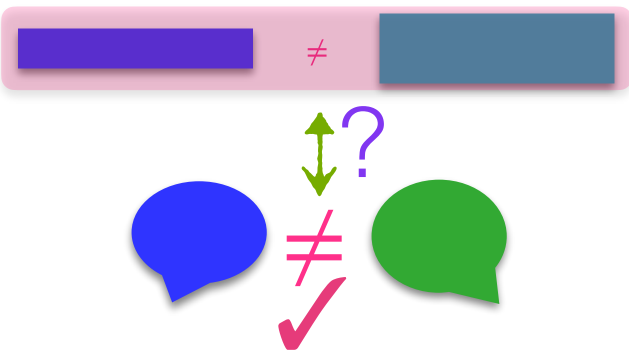
Bonus: Because the learning theory in the model is causal, we can predict if the input should cause similar outcomes, too.

In that case, the input difference isn’t meaningful.

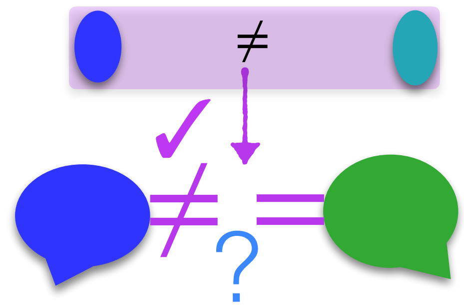


Detecting meaningful input differences

One (standard) way

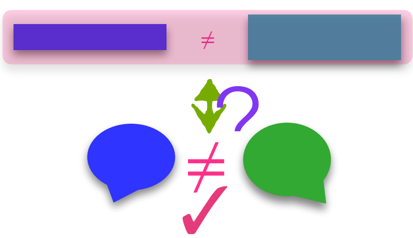


A new (complementary) way



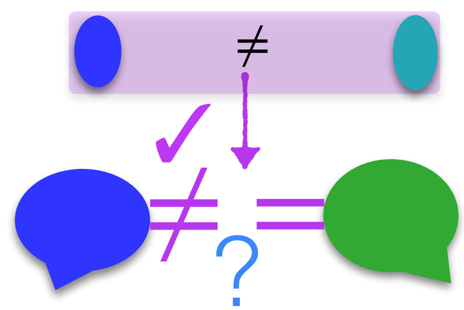
Detecting meaningful input differences

One (standard) way



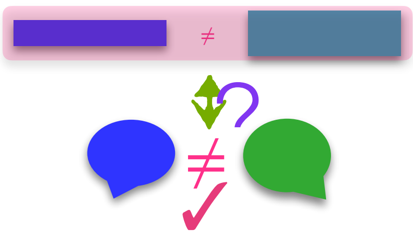
Today's focus

A new (complementary) way



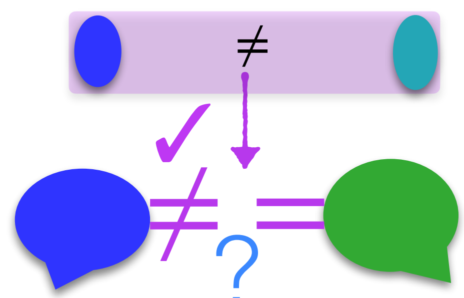
Detecting meaningful input differences

One (standard) way

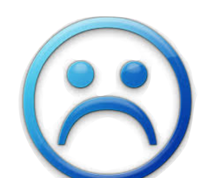


Today's focus

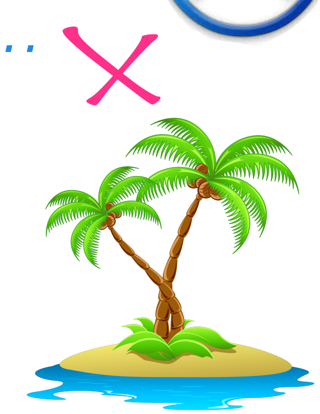
A new (complementary) way



Case study:
Syntactic island acquisition

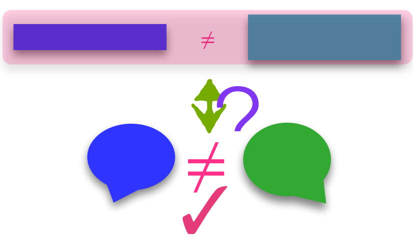


Who does…



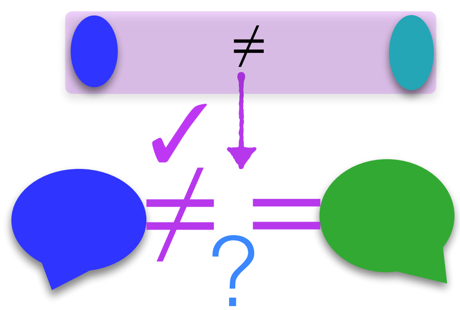
Detecting meaningful input differences

One (standard) way



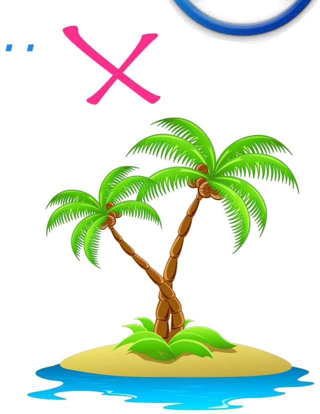
Today's focus

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 meaningful input differences 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? 

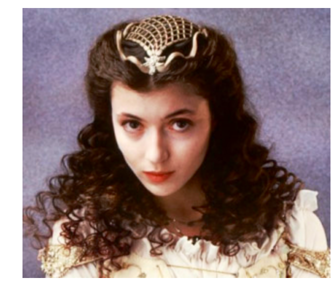
What does Lily think is pretty, and who does she think it's for? 

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 ___$_{who}$ 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.

Elizabeth

What did Elizabeth think __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

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

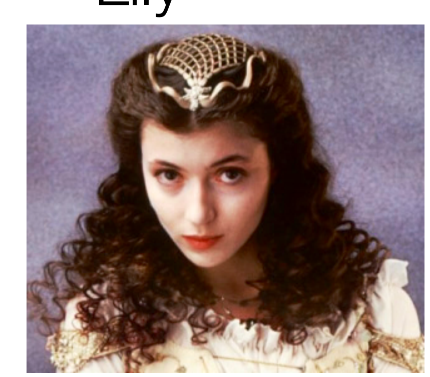
Jack

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

What did Elizabeth think Jack said Lily saw __what?

✔

Lily



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 an observable behavior signaling that acquisition of syntactic island knowledge has occurred.

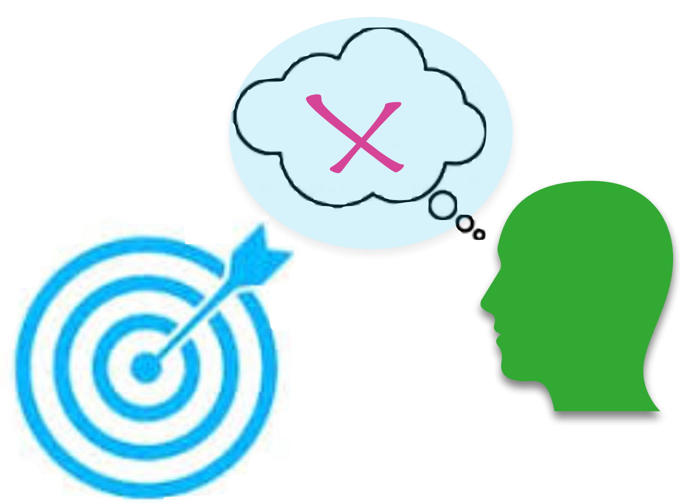
Syntactic island constraints

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.



Who does

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**)

Subject island stimuli

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

Sprouse et al. 2012

Syntactic 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

Syntactic island constraints

High-SES adult judgments

= behavioral target outcome

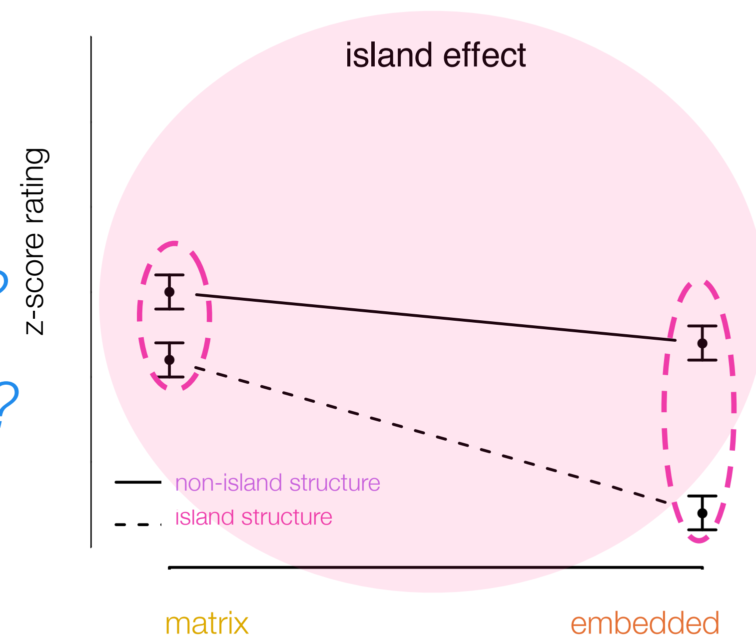
Adult knowledge as measured by acceptability judgment behavior

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

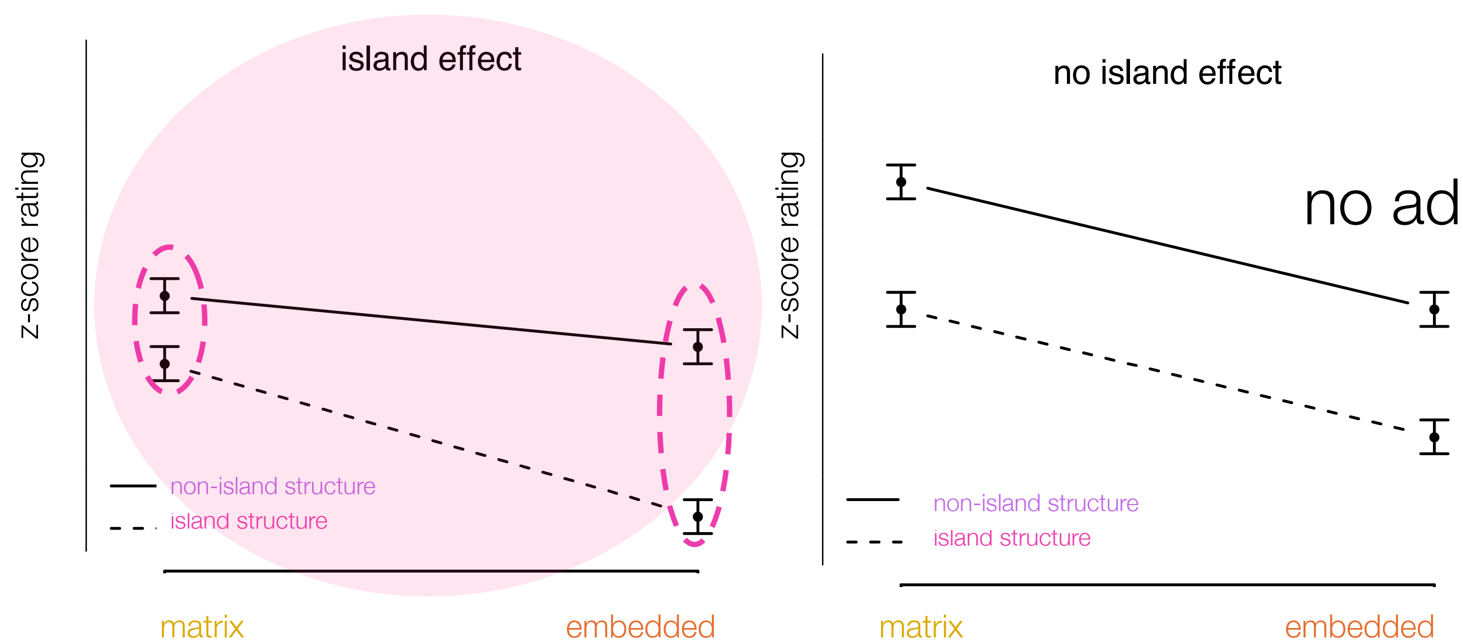
Adult knowledge as measured by acceptability judgment behavior

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

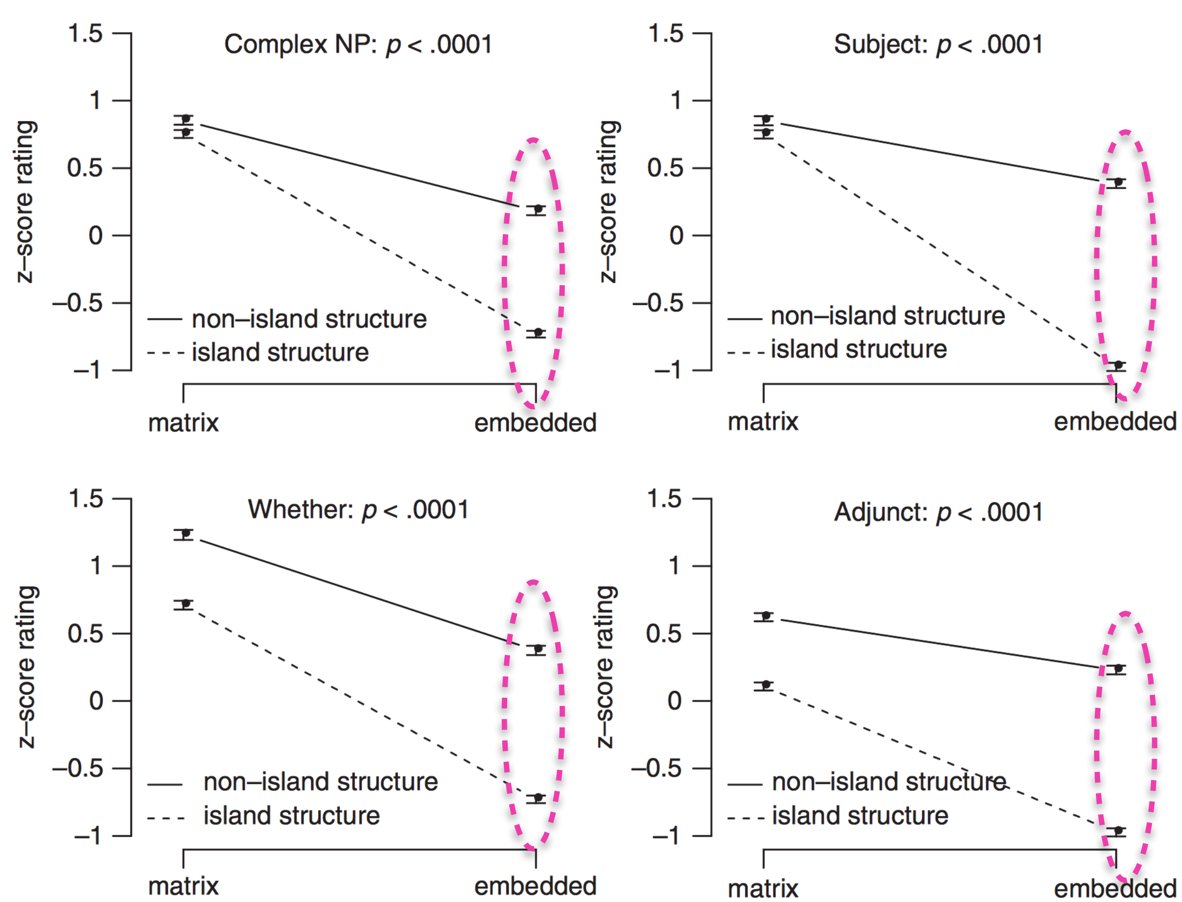
length of dependency
(**matrix vs. embedded**)

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

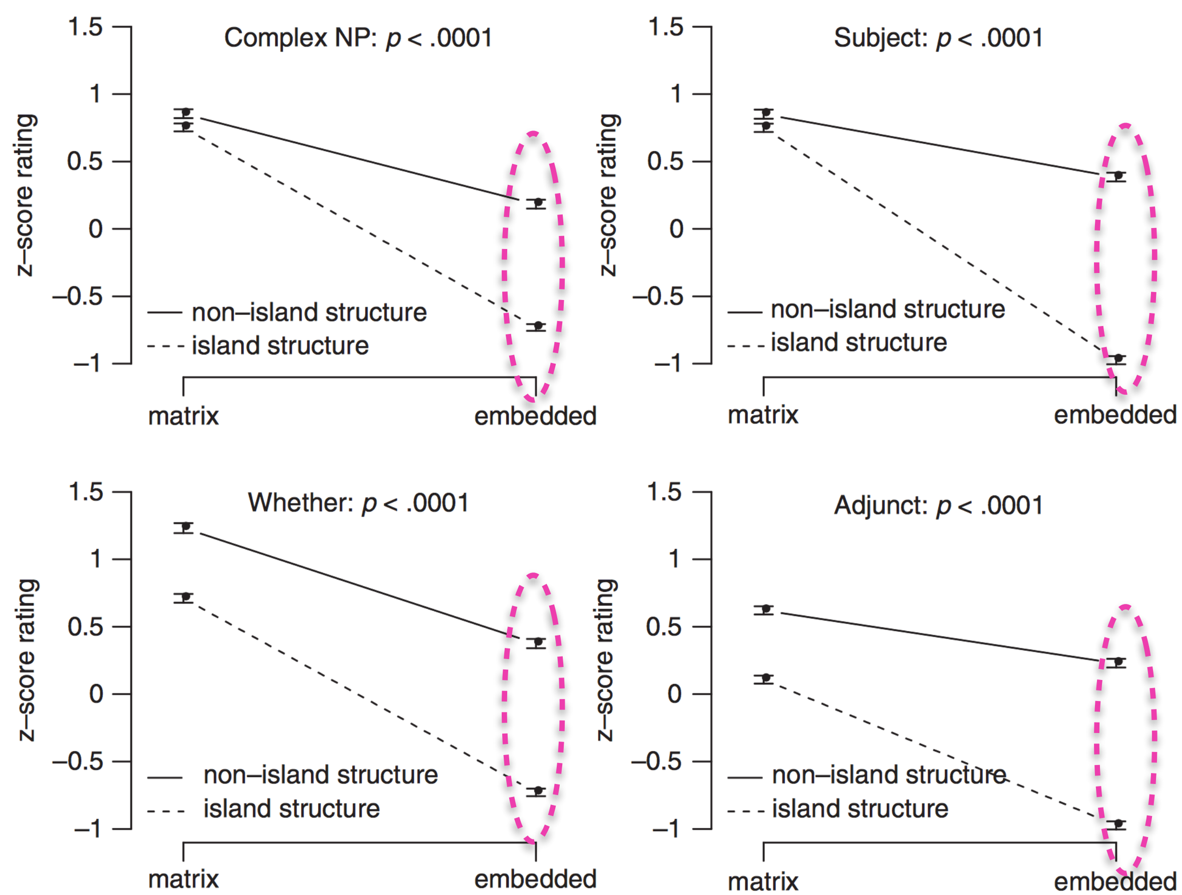
= behavioral target outcome

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

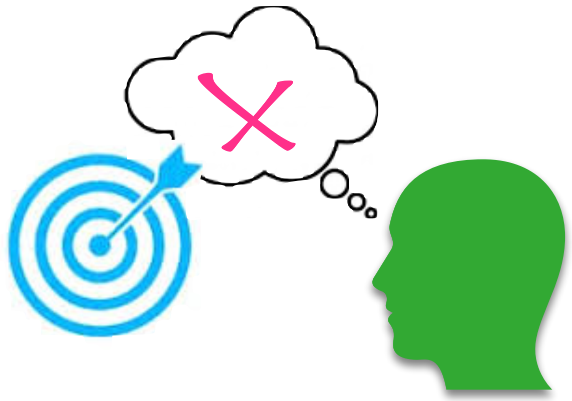


superadditivity for all four island types

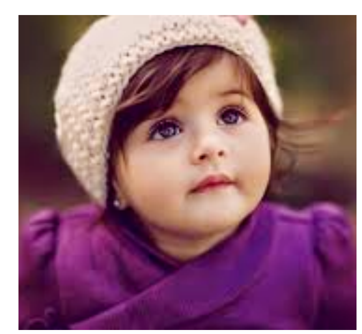
= knowledge that dependencies can't cross these island structures.

Sprouse et al. 2012

Who does



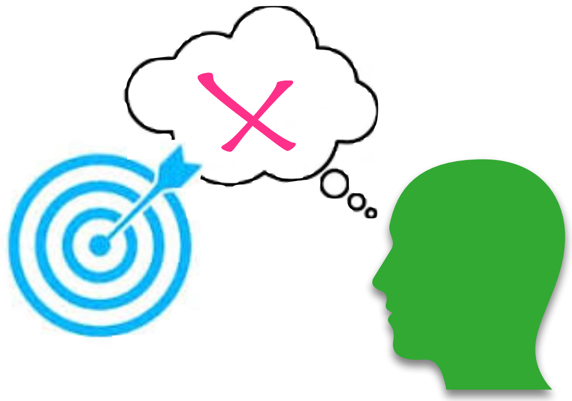
Okay, so what's the relevant input for learning this target knowledge?



Who does 

That depends on how we think children learn it.

Who does



That depends on how we think children learn it.

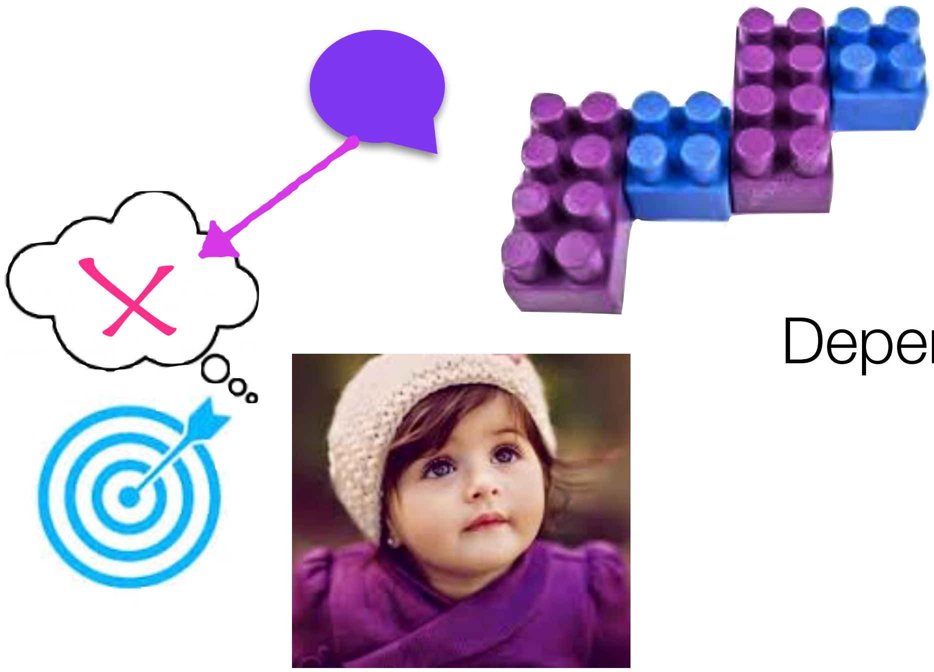
Pearl & Sprouse 2013 intuition:

- 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 syntactic islands.

Who does

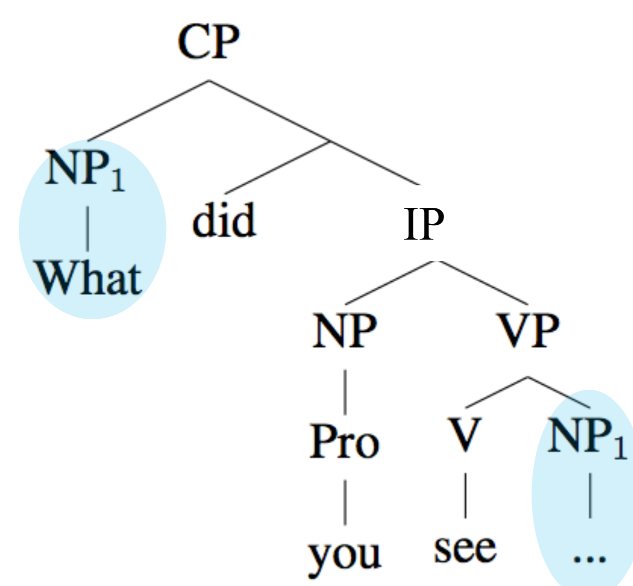
A concrete learning strategy (Pearl & Sprouse 2013): View *wh*-dependencies in terms of their building blocks and track those building blocks in the input.

A strategy for learning syntactic islands



Dependencies represented as a sequence of container nodes

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

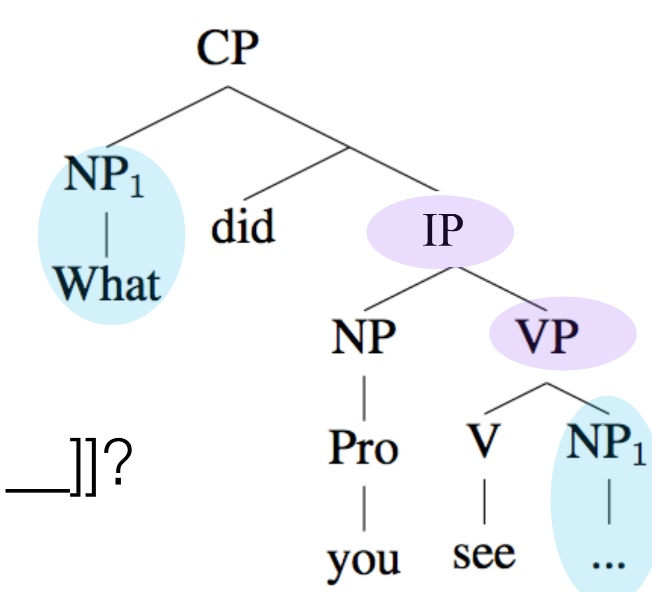


Pearl & Sprouse 2013

A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

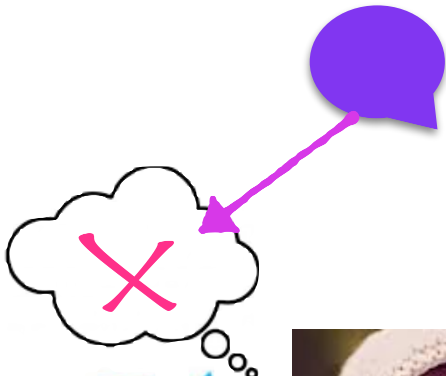
What phrases contain the gap
(but not the *wh*-word)?



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

Pearl & Sprouse 2013

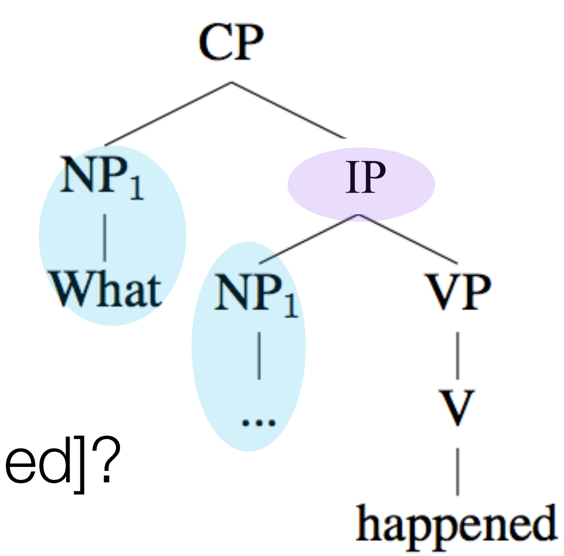
A strategy for learning syntactic islands



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*



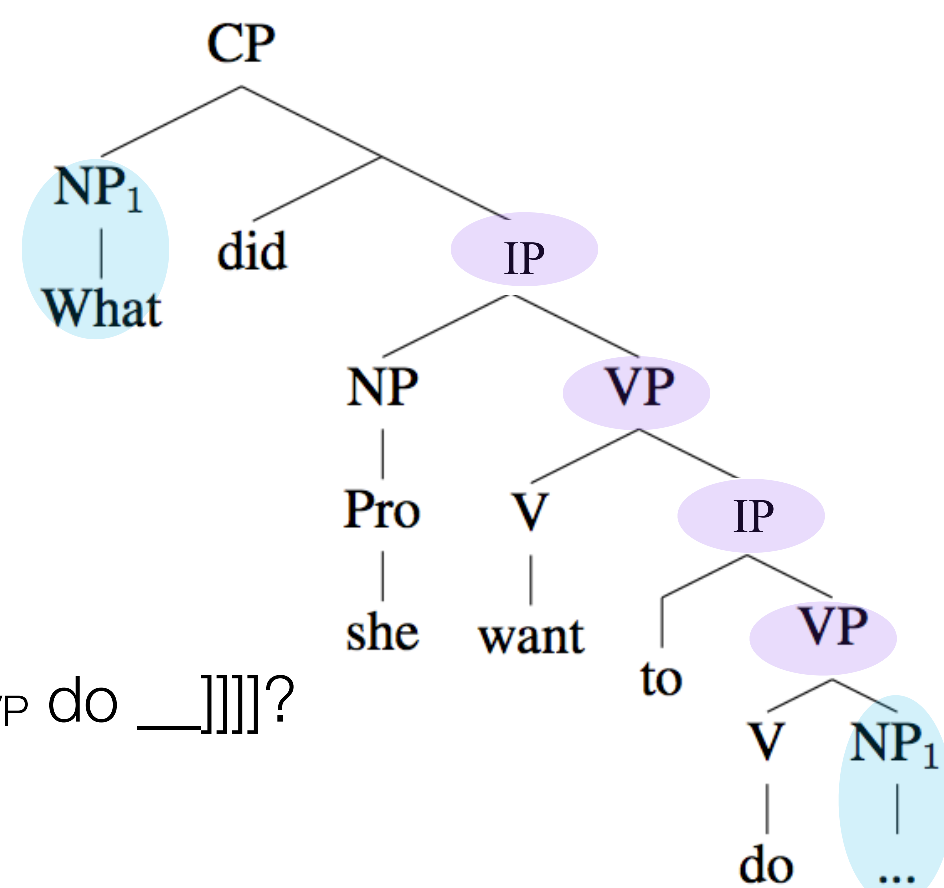
What __ happened?
= What [$_{IP}$ __ happened]?
= *start*-IP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

Dependencies represented as a sequence of container nodes

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



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

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

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

Pearl & Sprouse 2013

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*

Ungrammatical dependencies have low probability segments

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

start-IP-VP-CP-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*~~ [$_{CP}$ [$_{IP}$ [$_{NP}$ *the kitty* [$_{PP}$ *for* __]] *was pretty* ?]]]]

start-IP-VP-CP-IP-NP-PP-*end*

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

Pearl & Sprouse 2013

Who does

A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-VP-CP-IP-NP-PP-*end*

start-IP-*end*

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

syntactic trigrams

The building blocks: trigrams of container nodes

Pearl & Sprouse 2013

A strategy for learning syntactic islands

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

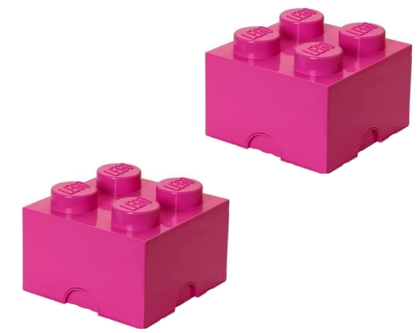
A strategy for learning syntactic islands

Who does

start-IP-VP-*end*

start-IP-VP-CP-IP-NP-PP-*end*

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



syntactic trigrams

The building blocks: trigrams of container nodes

start-IP-VP

IP-VP-IP

VP-IP-VP

IP-VP-*end*

start-IP-*end*

start-IP-*end*

Pearl & Sprouse 2013

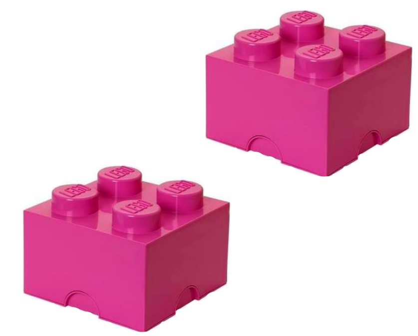
A strategy for learning syntactic islands

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-IP-NP-PP-*end*

start-IP-VP

IP-VP-CP

VP-CP-IP

CP-IP-NP

IP-NP-PP

NP-PP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-VP-CP-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

IP-VP-PP

VP-PP-*end*

VP-CP-IP

NP-PP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-VP-CP-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

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-CP-IP

VP-IP-VP

NP-PP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-VP-CP-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-IP-NP

IP-NP-PP

IP-VP-CP

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-CP-IP

VP-IP-VP

NP-PP-*end*

Pearl & Sprouse 2013

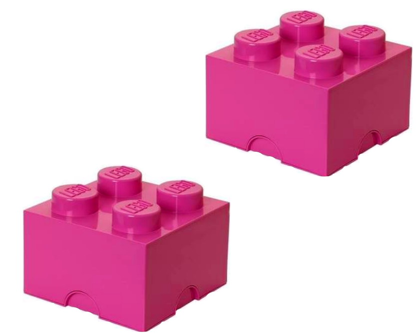
A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-VP-CP-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*

CP-IP-NP

IP-NP-PP

IP-VP-CP

IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

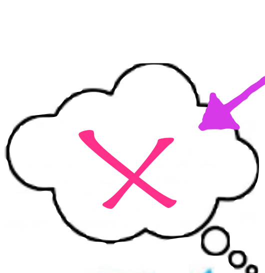
NP-PP-*end*

VP-CP-IP

Pearl & Sprouse 2013

Who does

A strategy for learning syntactic islands



start-IP-*end*

start-IP-VP-*end*

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

start-IP-VP-CP-IP-NP-PP-*end*

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

start-IP-VP

start-IP-*end*

IP-VP-*end*

IP-VP-CP

IP-VP-IP

IP-VP-PP

VP-PP-*end*

VP-CP-IP

VP-IP-VP

NP-PP-*end*

CP-IP-NP

IP-NP-PP

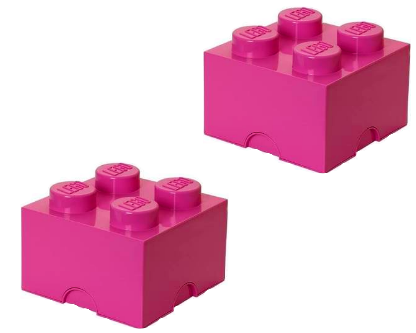
Pearl & Sprouse 2013

A strategy for learning syntactic islands

start-IP-VP-CP-IP-NP-PP-*end*

start-IP-*end*

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



start-IP-VP-*end*

start-IP-VP

IP-VP-*end*

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

start-IP-*end*

CP-IP-NP

IP-VP-CP

IP-NP-PP

IP-VP-PP

VP-PP-*end*

IP-VP-IP

VP-IP-VP

NP-PP-*end*

VP-CP-IP

Pearl & Sprouse 2013

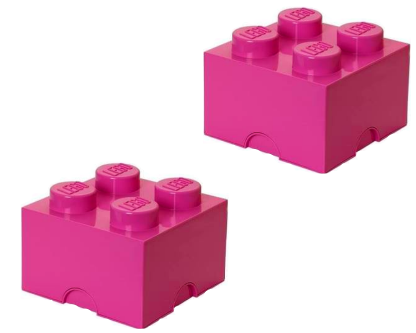
Who does

A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-*end*

start-IP-VP-CP-IP-NP-PP-*end*



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

start-IP-VP

IP-VP-IP

VP-IP-VP

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

IP-VP-*end*

start-IP-*end*

CP-IP-NP

IP-NP-PP

IP-VP-CP

IP-VP-PP

VP-PP-*end*

NP-PP-*end*

VP-CP-IP

Pearl & Sprouse 2013

A strategy for learning syntactic islands

Who does

start-IP-VP-*end*

start-IP-*end*

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

start-IP-VP-CP-IP-NP-PP-*end*

start-IP-VP

IP-VP-CP

VP-CP-IP

CP-IP-NP

IP-NP-PP

NP-PP-*end*

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

start-IP-*end*

IP-VP-IP

IP-VP-PP

VP-IP-VP

VP-PP-*end*

Pearl & Sprouse 2013

A strategy for learning syntactic islands

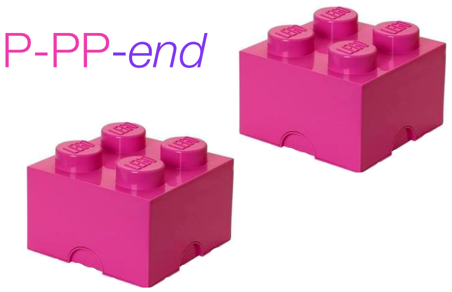
Who does

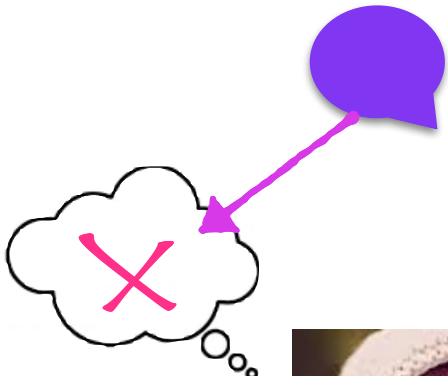
start-IP-VP-*end*

start-IP-VP-CP-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 judged acceptability.

Pearl & Sprouse 2013

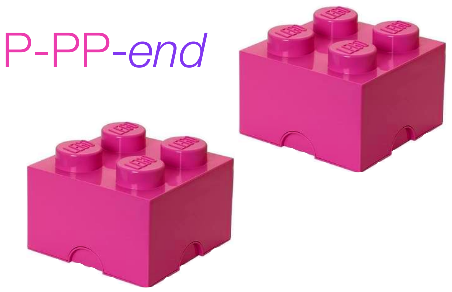
A strategy for learning syntactic islands

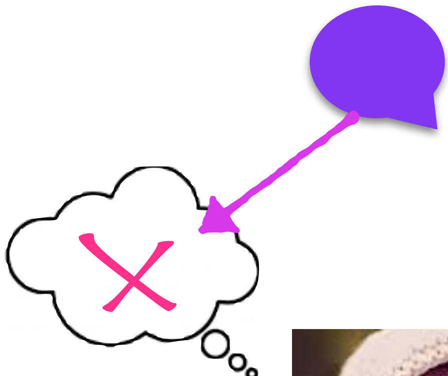
start-IP-VP-*end*

start-IP-VP-CP-IP-NP-PP-*end*

start-IP-*end*

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





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

Lower probability dependencies are dispreferred, compared to higher probability dependencies.

Pearl & Sprouse 2013

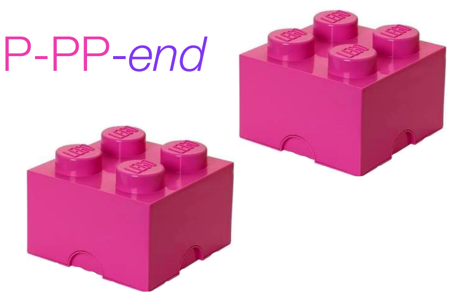
A strategy for learning syntactic islands

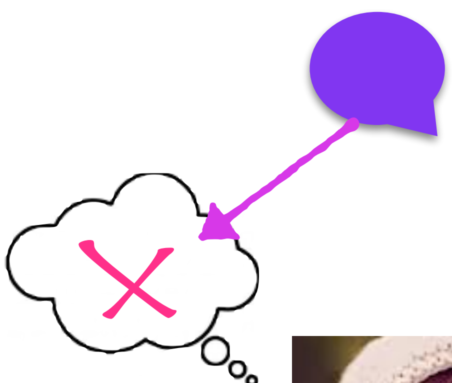
start-IP-VP-*end*

start-IP-VP-CP-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

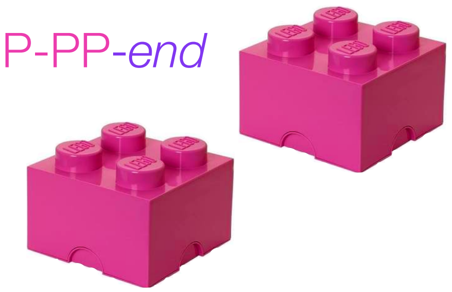
A strategy for learning syntactic islands

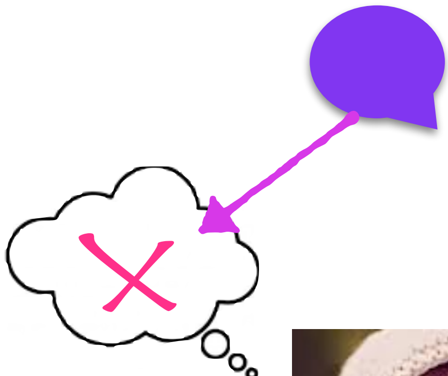
start-IP-VP-*end*

start-IP-VP-CP-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

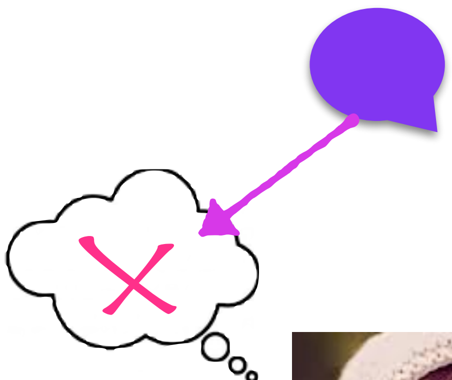
A strategy for learning syntactic islands

start-IP-VP-*end*

start-IP-VP-CP-IP-NP-PP-*end*

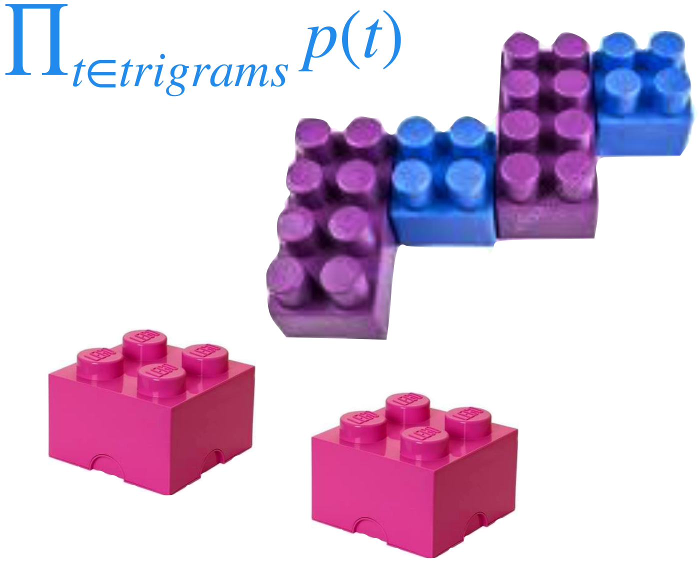
start-IP-*end*

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

Who does

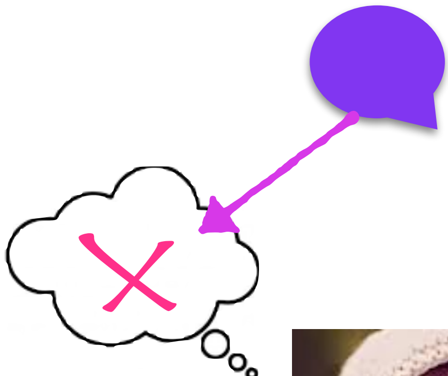
A strategy for learning syntactic islands

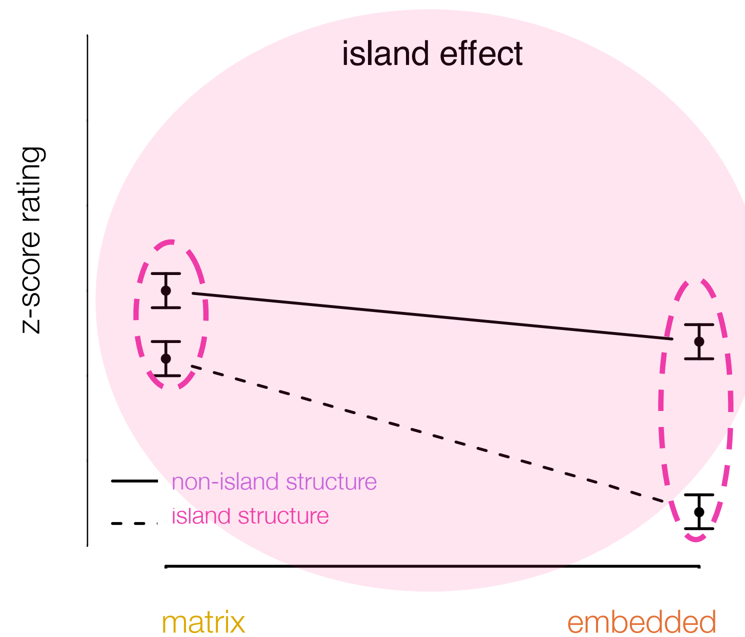
start-IP-VP-*end*

start-IP-VP-CP-IP-NP-PP-*end*

start-IP-*end*

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



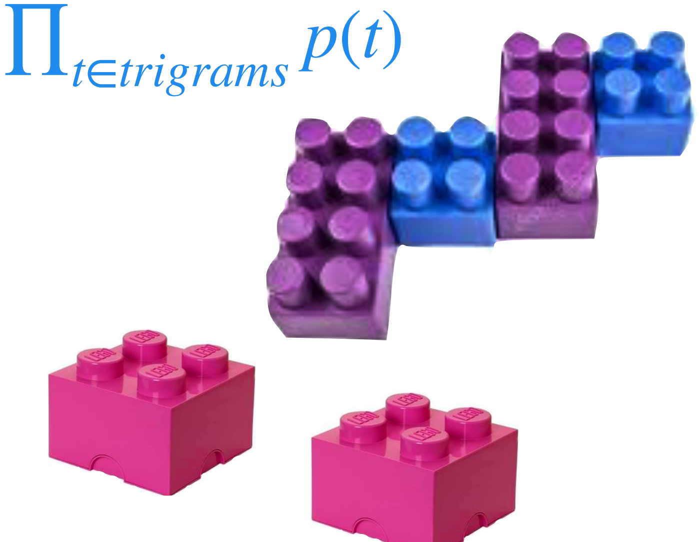


These probabilities can then be plotted to see if superadditivity is present.

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

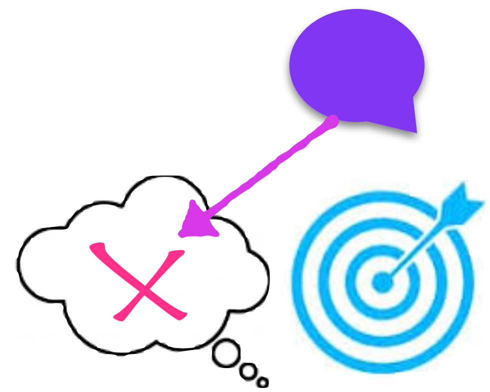
A strategy for learning syntactic islands

start-IP-VP-*end*

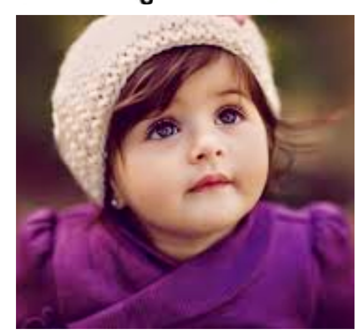
start-IP-VP-CP-IP-NP-PP-*end*

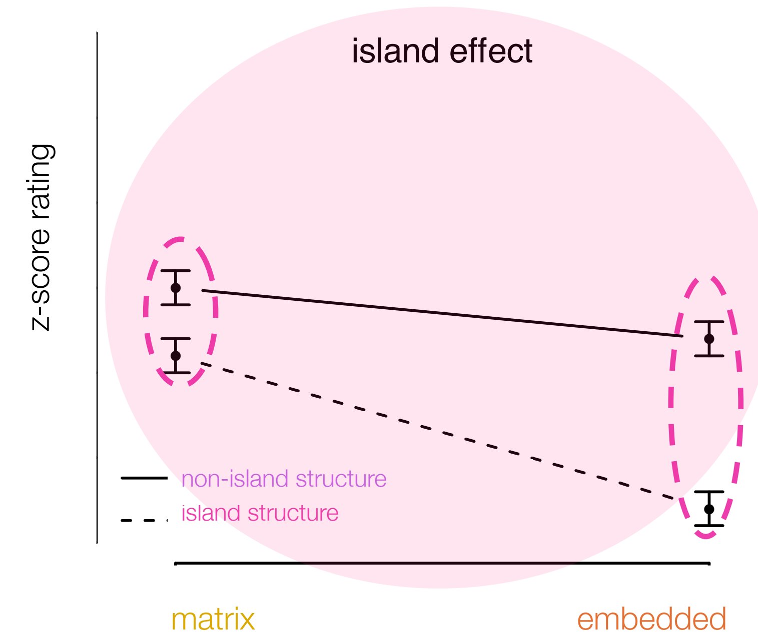
start-IP-*end*

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



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

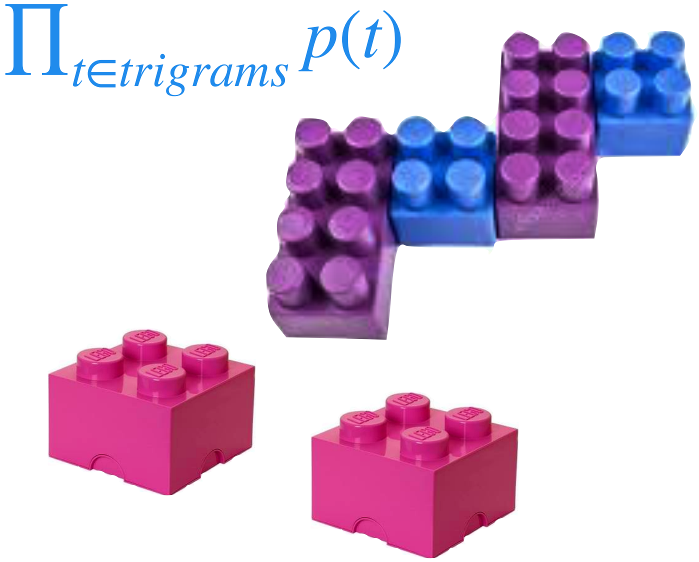




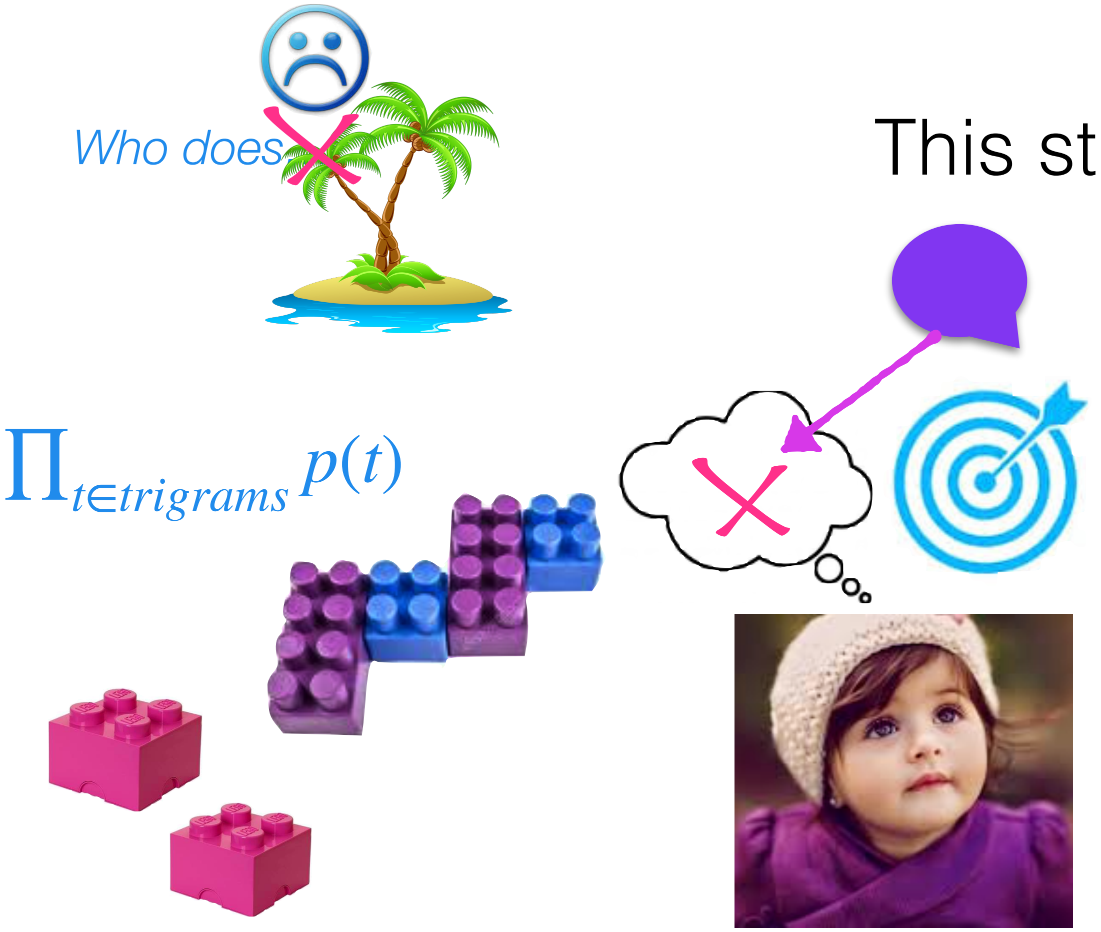
Complex NP island stimuli

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

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

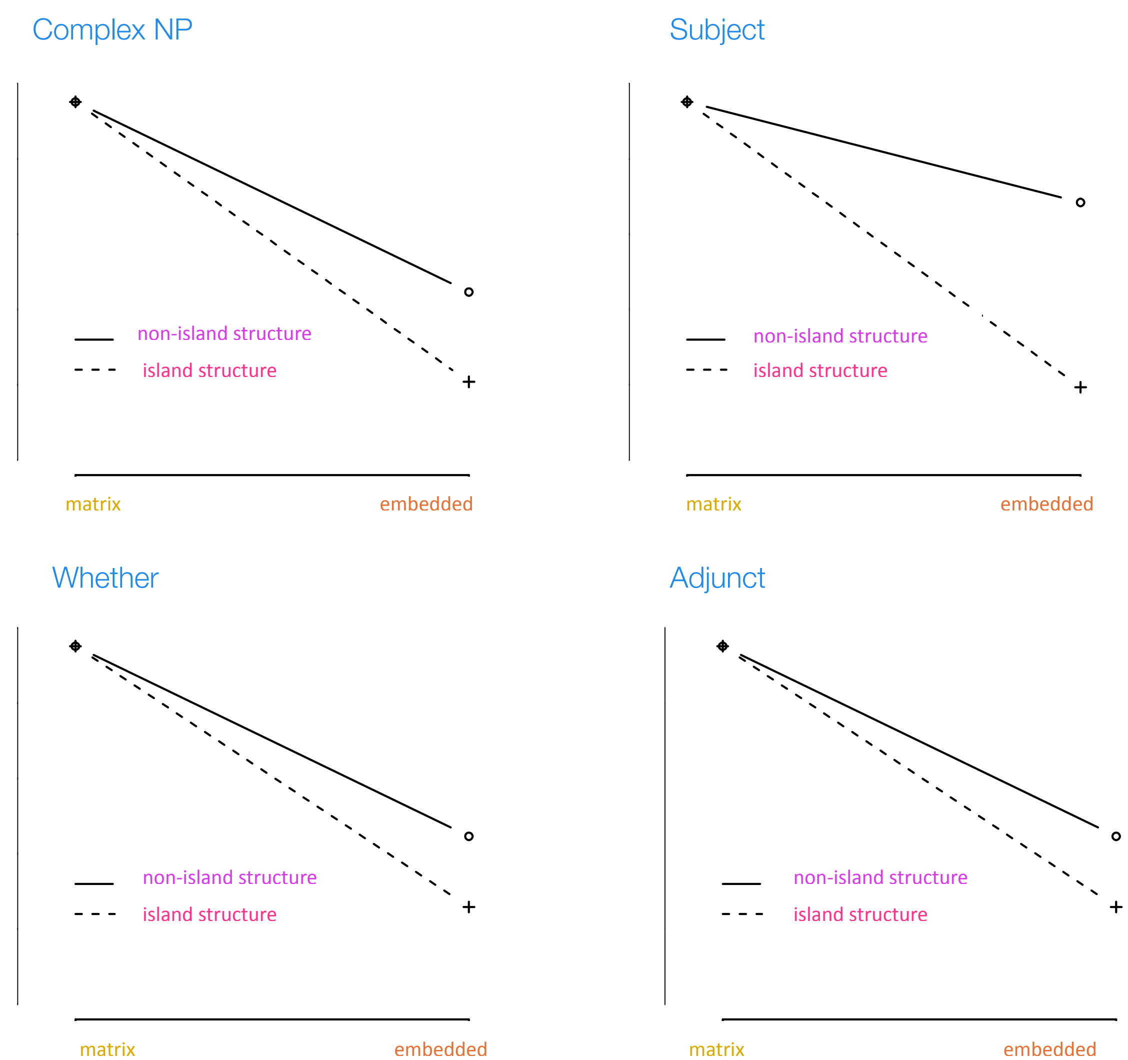


Pearl & Sprouse 2013



This strategy works for high-SES children's input

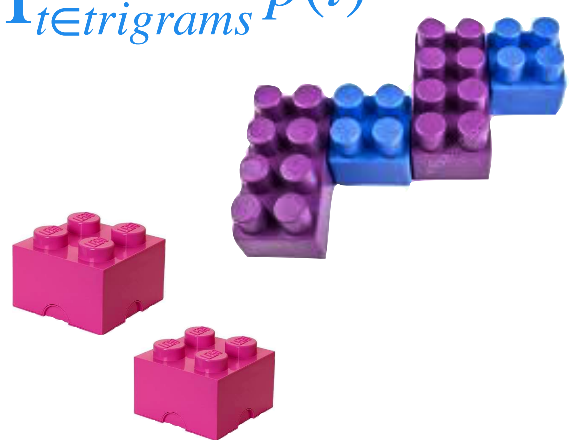
Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data.

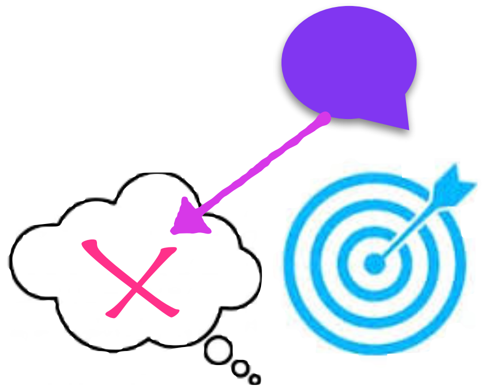


Pearl & Sprouse 2013

This strategy works for high-SES children's input

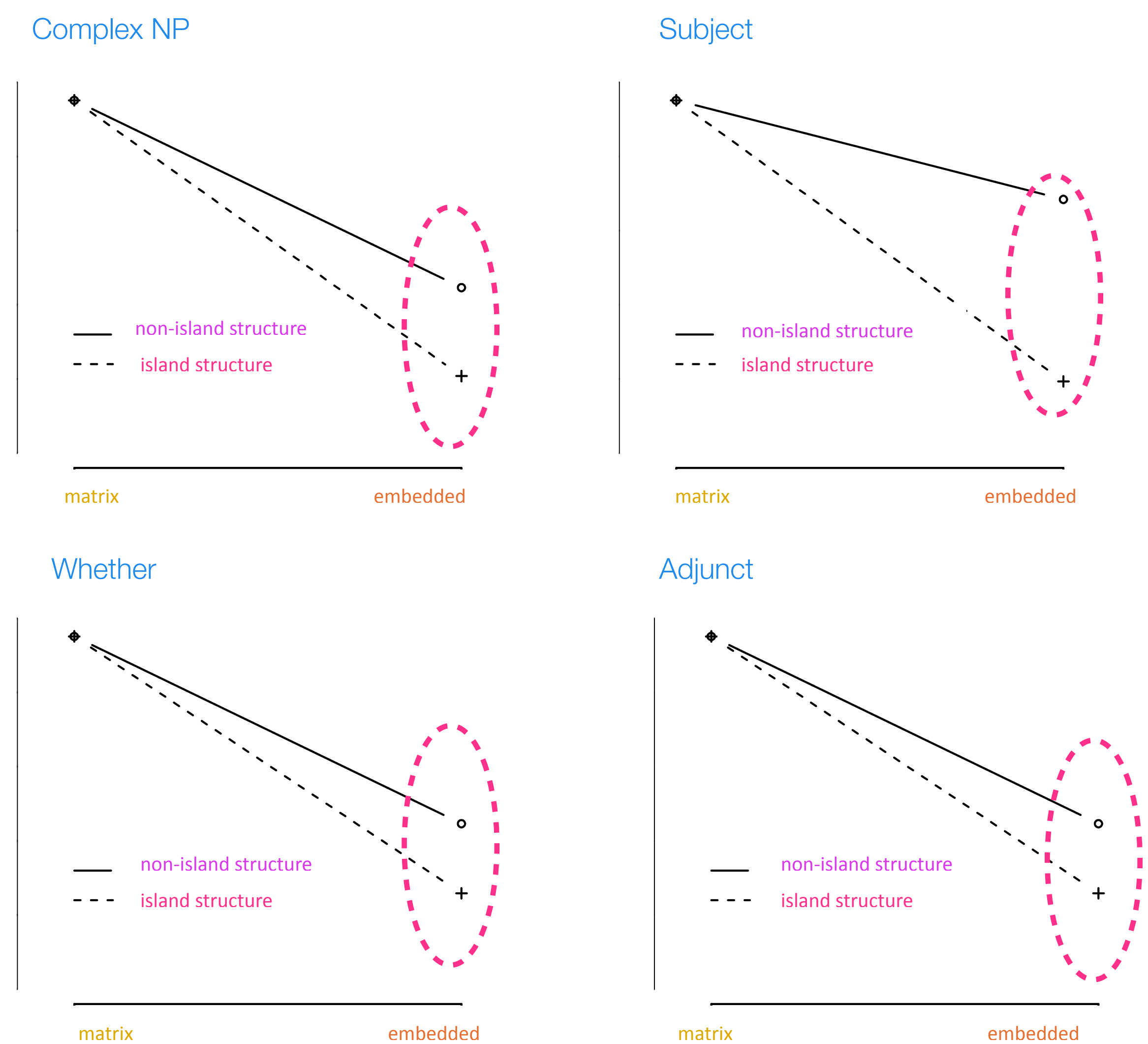
$\prod_{t \in trigrams} p(t)$





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.

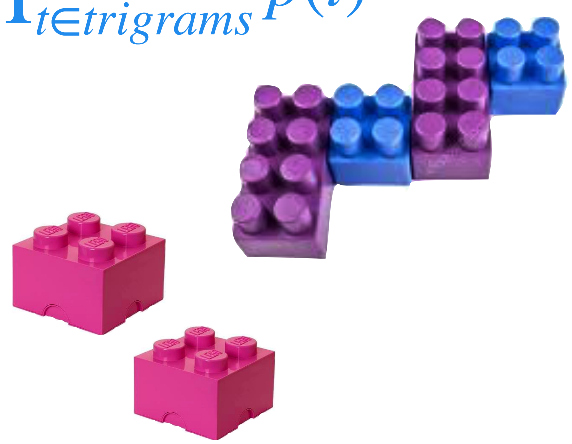
Superadditivity for all four islands.

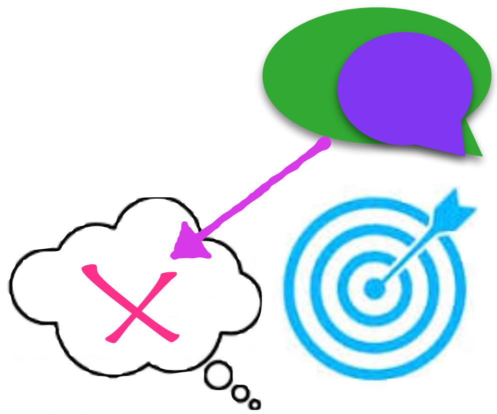


Pearl & Sprouse 2013

This strategy works for high-SES children's input

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

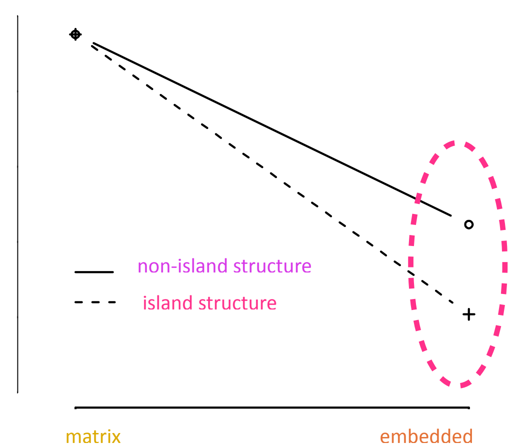




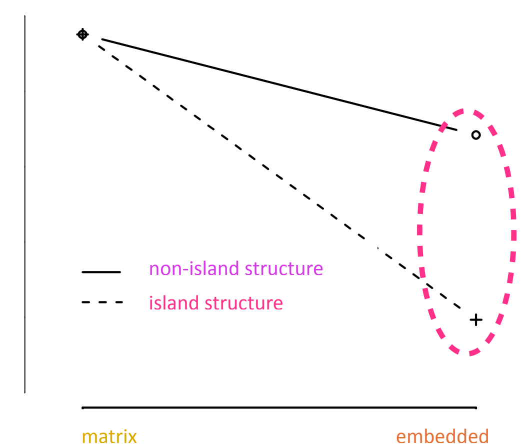
Judgments from a modeled child learning from the same amount of data as high-SES children seem to, with those data having the same composition as high-SES child-directed speech data.

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.

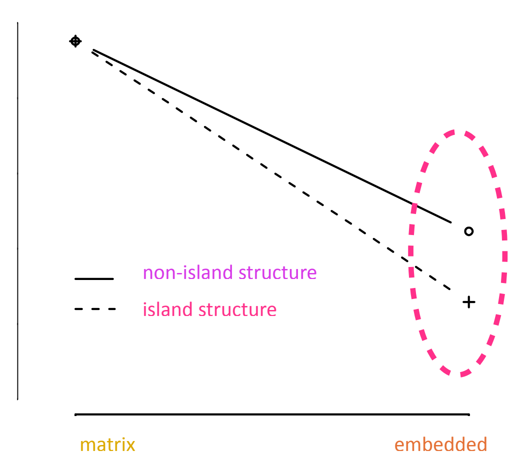
Complex NP



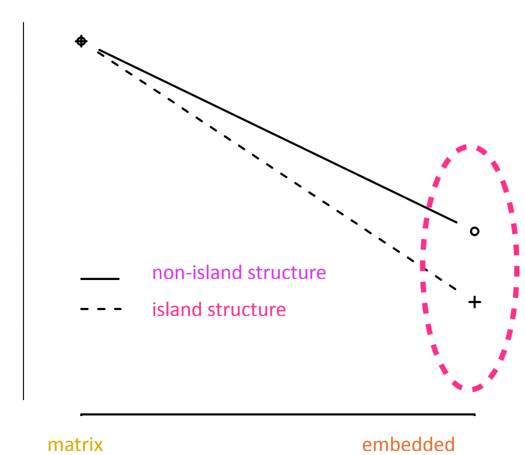
Subject



Whether

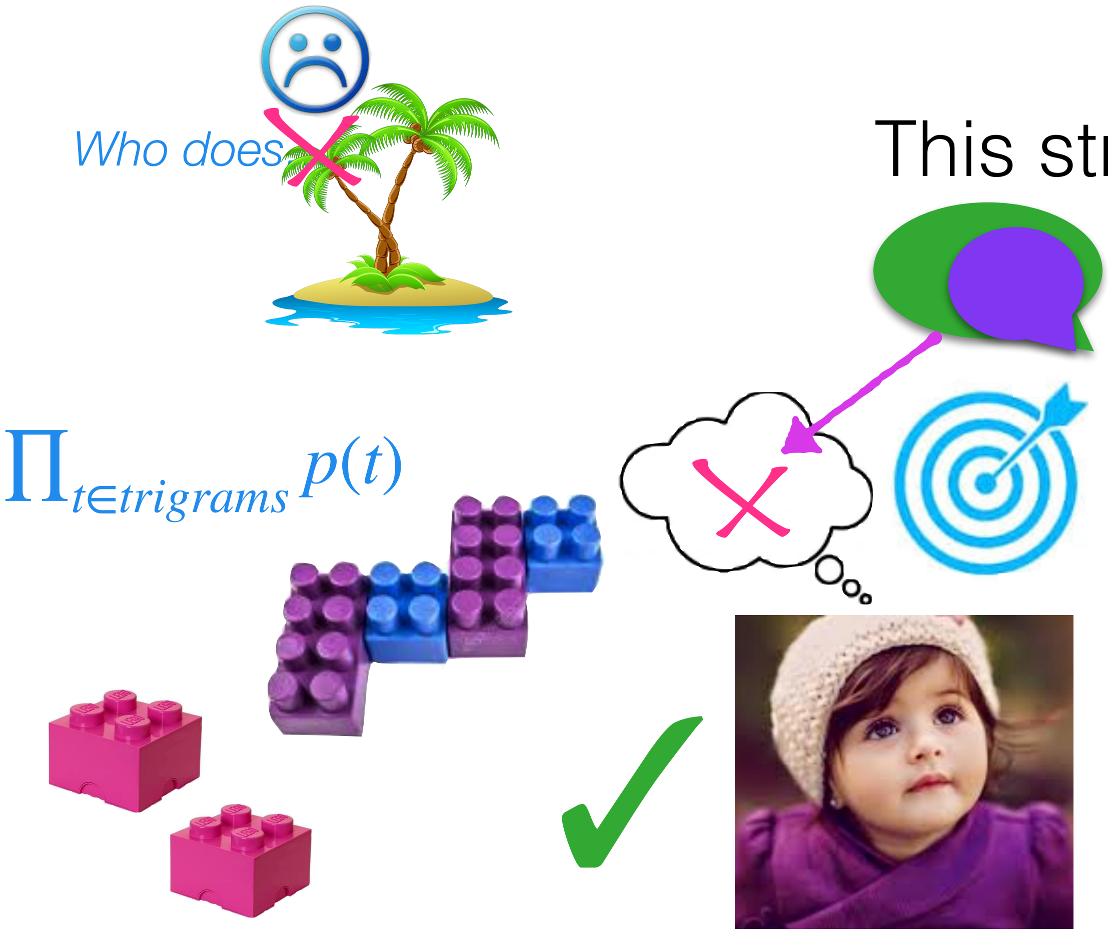


Adjunct



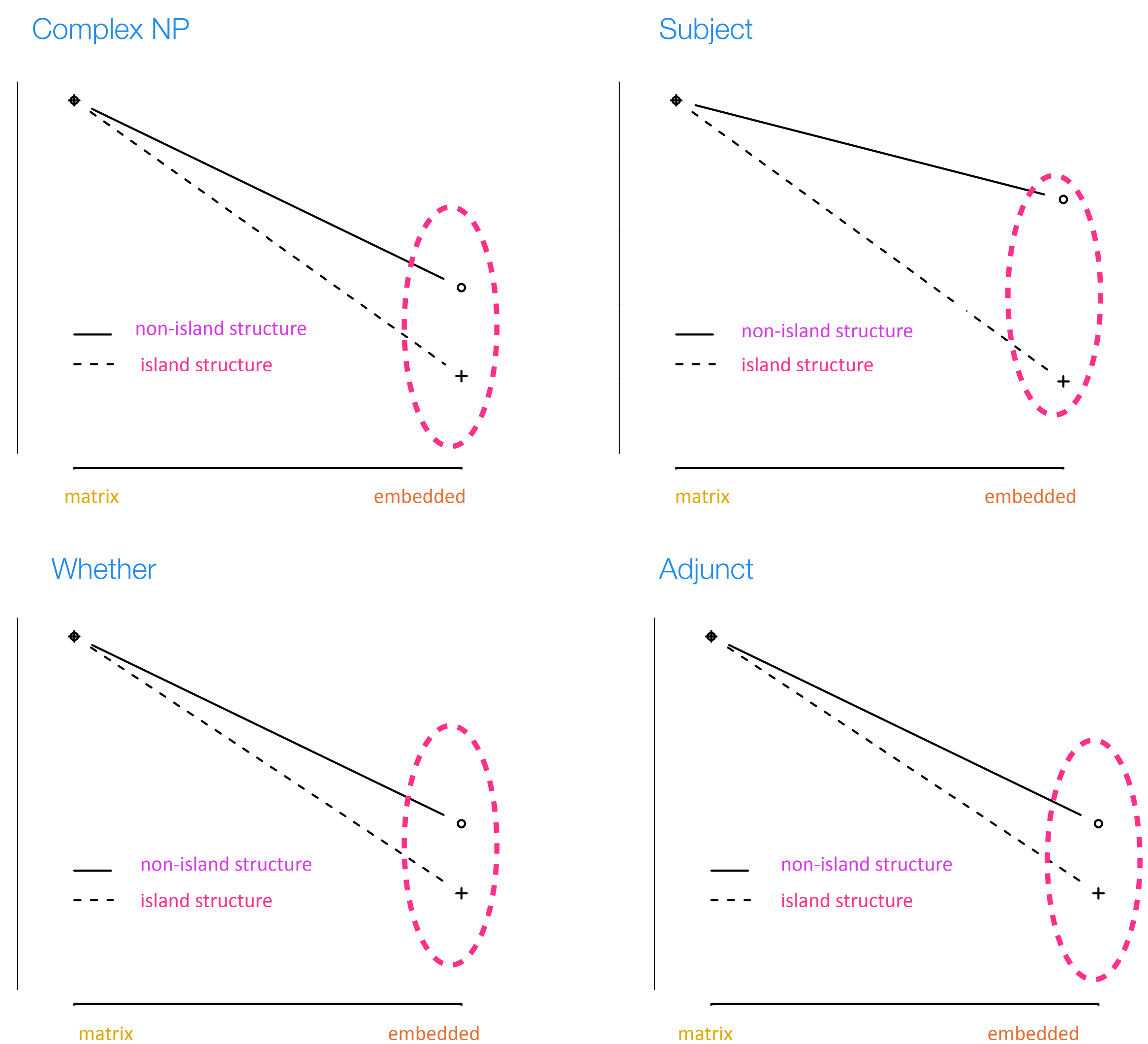
Pearl & Sprouse 2013

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.

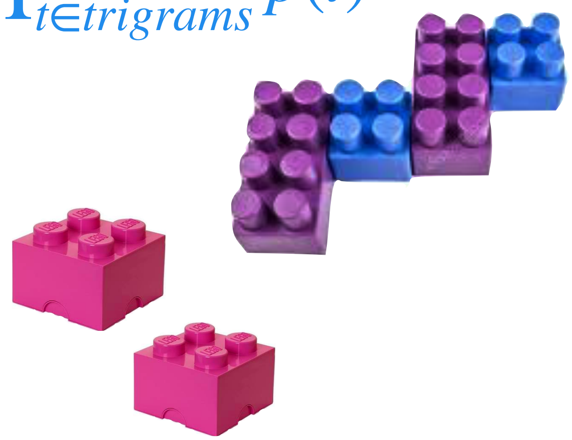
That input part is the *wh*-dependencies, and their building blocks (the syntactic trigrams).



Pearl & Sprouse 2013

Who does

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

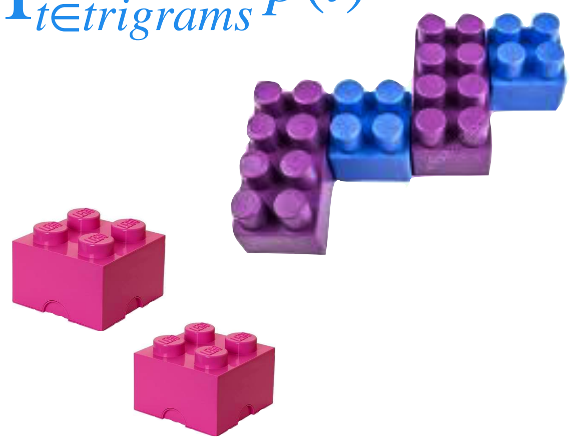


Are there meaningful differences across SES in this part of the input (the *wh*-dependencies and syntactic trigrams)?

Bates & Pearl 2019, in prep.

Who does

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

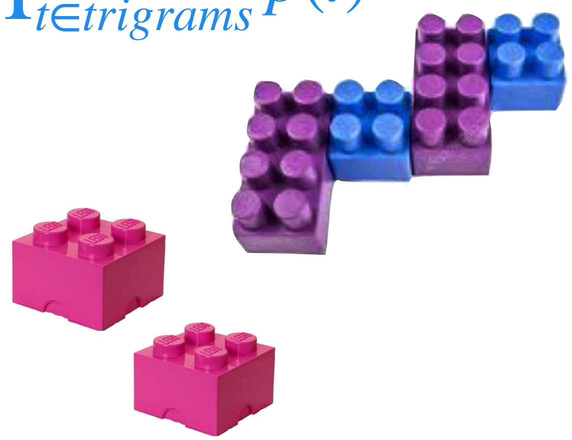


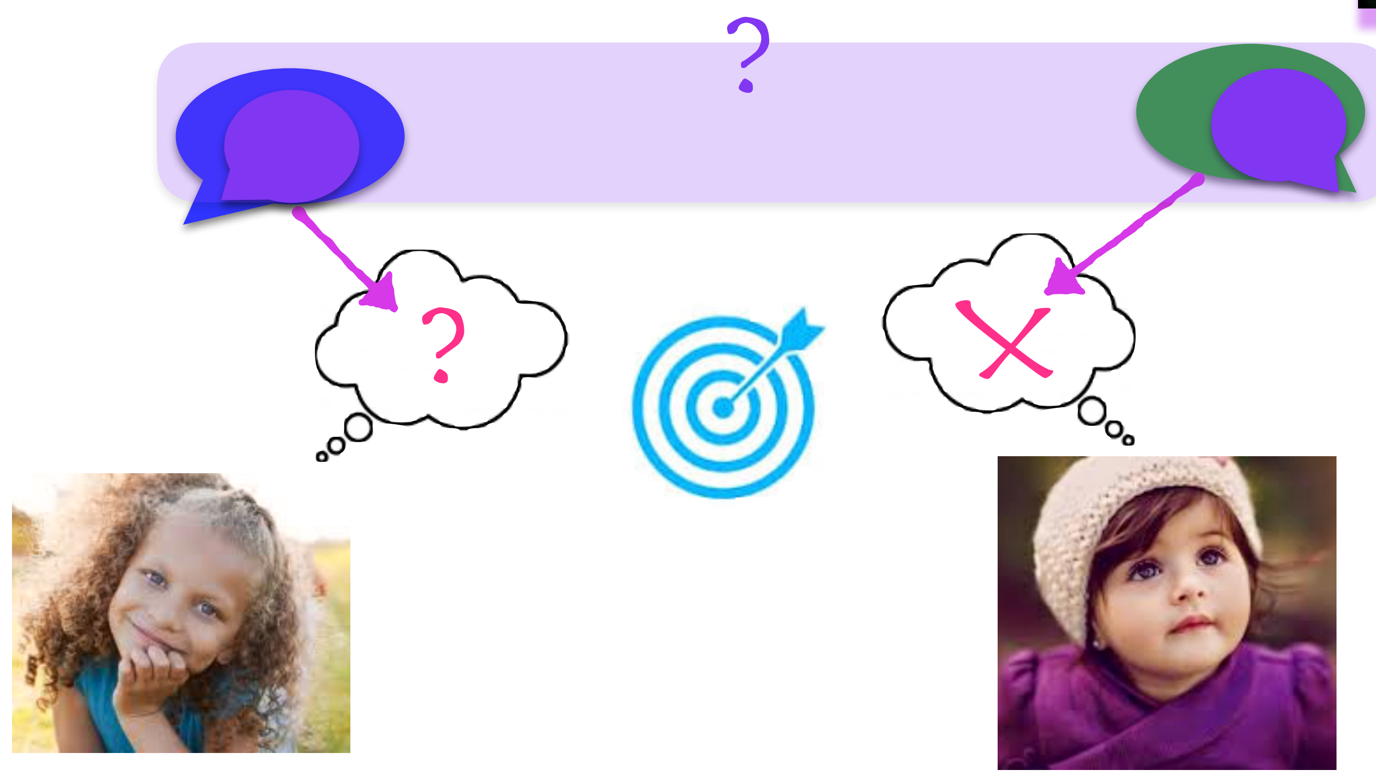
Are there meaningful differences across SES in this part of the input (the *wh*-dependencies and syntactic trigrams)?

Let's use developmental modeling to find out.

Bates & Pearl 2019, in prep.

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





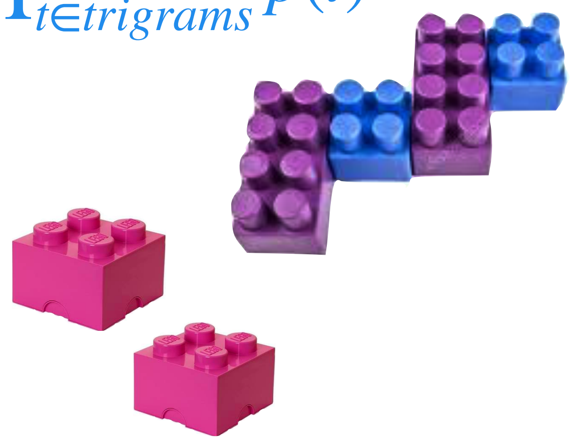
But first…how different does this input look across SES?

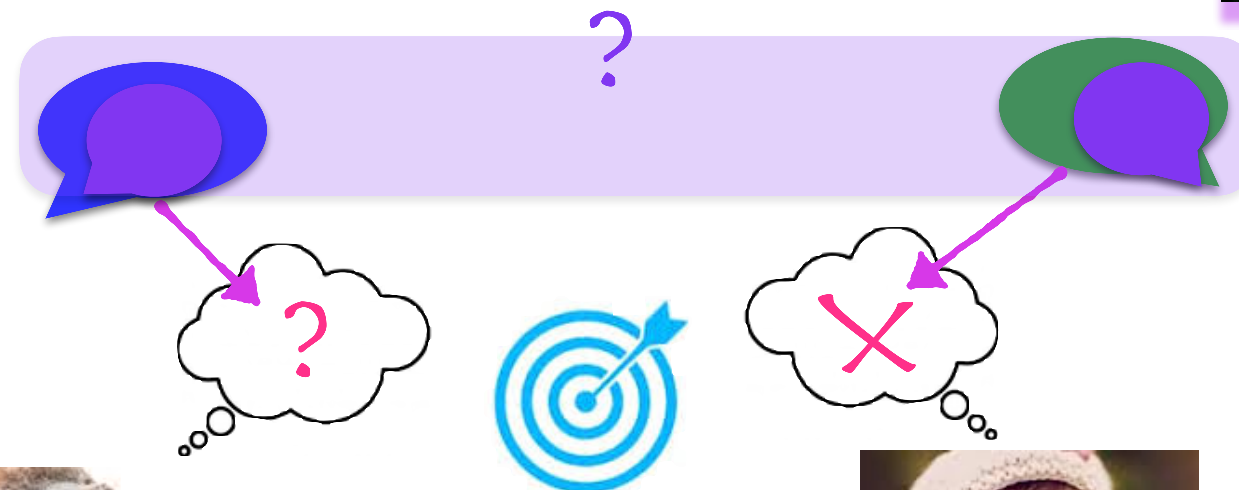
Let’s look at the distribution of the relevant parts:
the *wh*-dependencies and the syntactic trigrams.

Bates & Pearl 2019, in prep.

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

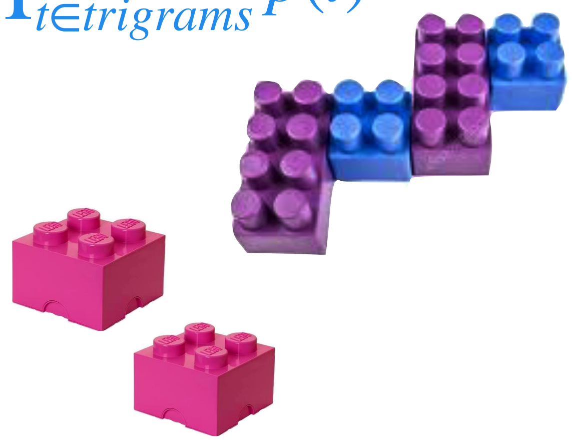
$\neq$

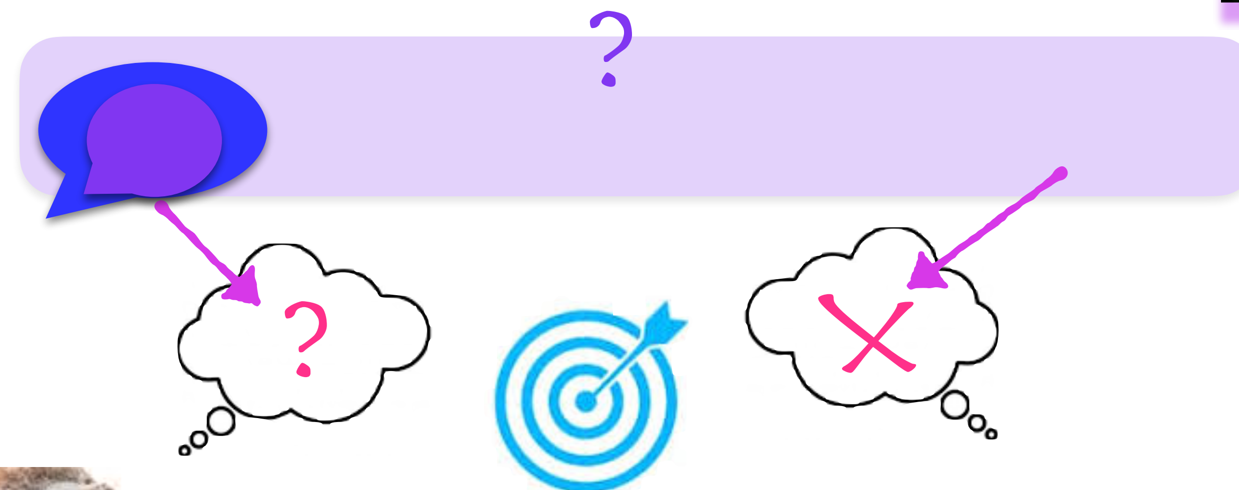
Bates & Pearl 2019, in prep.

Measurable input differences

Who does

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





$=$ $\neq$

$0 \leq JSDiv \leq 1$

The input samples

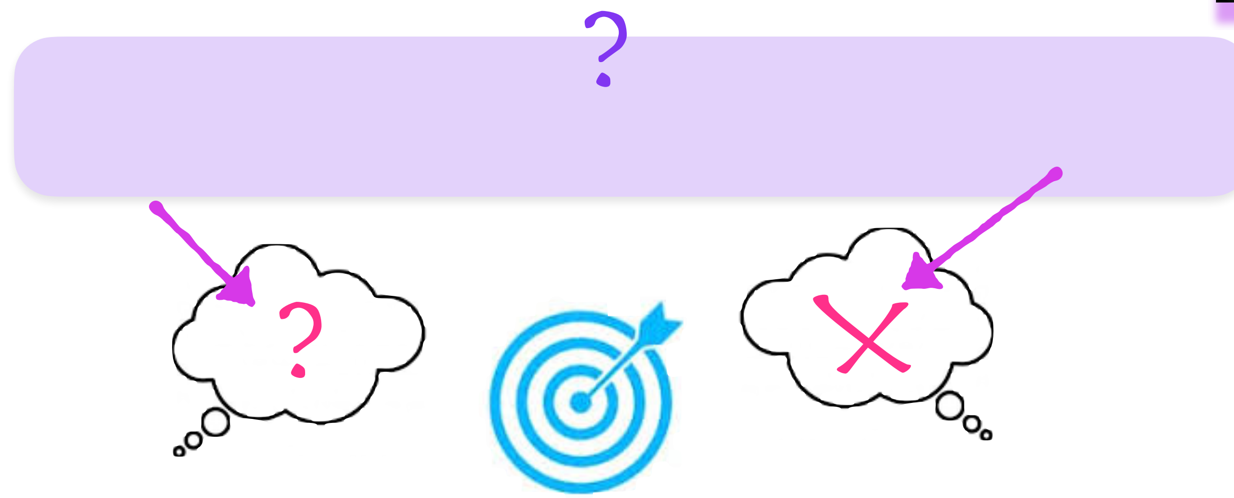
High-SES
child-directed

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.

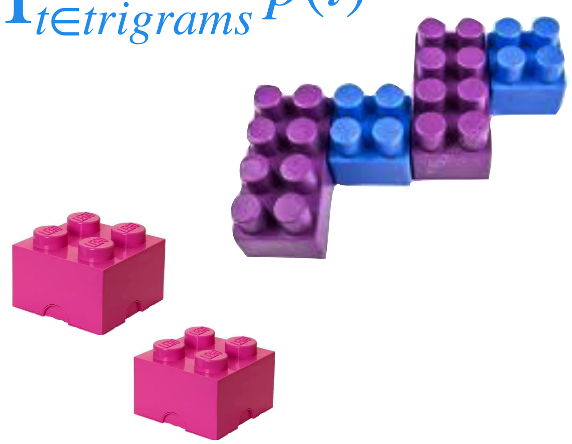
Bates & Pearl 2019, in prep.

Measurable input differences

Who does



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

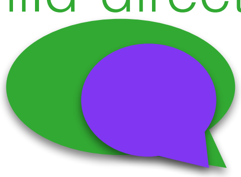


$=$ $\neq$

$$0 \leq JSDiv \leq 1$$

$\boxed{\text{The input samples}}$

High-SES
child-directed



21K *wh*-dependencies

Low-SES
child-directed

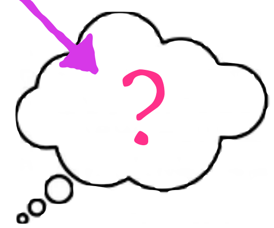
31.8K utterances (3.9K *wh*-dependencies) from a subpart of the HSLLD corpus (Dickinson & Tabors 2001) in the CHILDES Treebank (Pearl & Sprouse 2013) of speech directed at 78 low-SES children between the ages of 3 and 5.

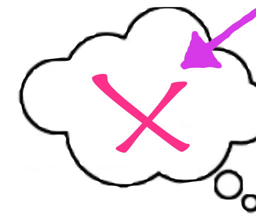
Bates & Pearl 2019, in prep.

Measurable input differences

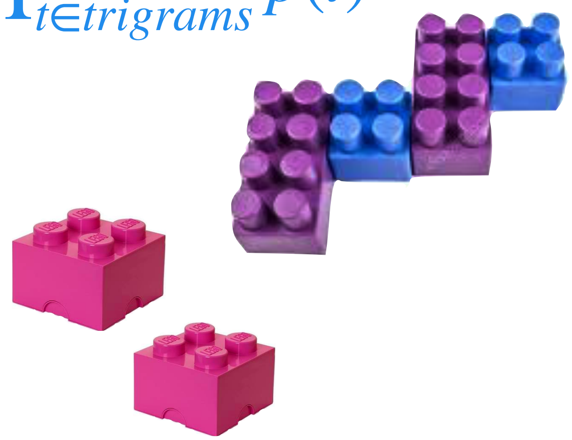
Who does

?





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

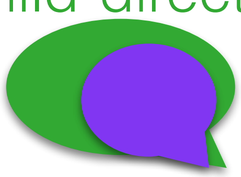


$=$ $\neq$

$0 \leq JSDiv \leq 1$

The input samples

High-SES
child-directed



21K *wh*-dependencies

Low-SES
child-directed

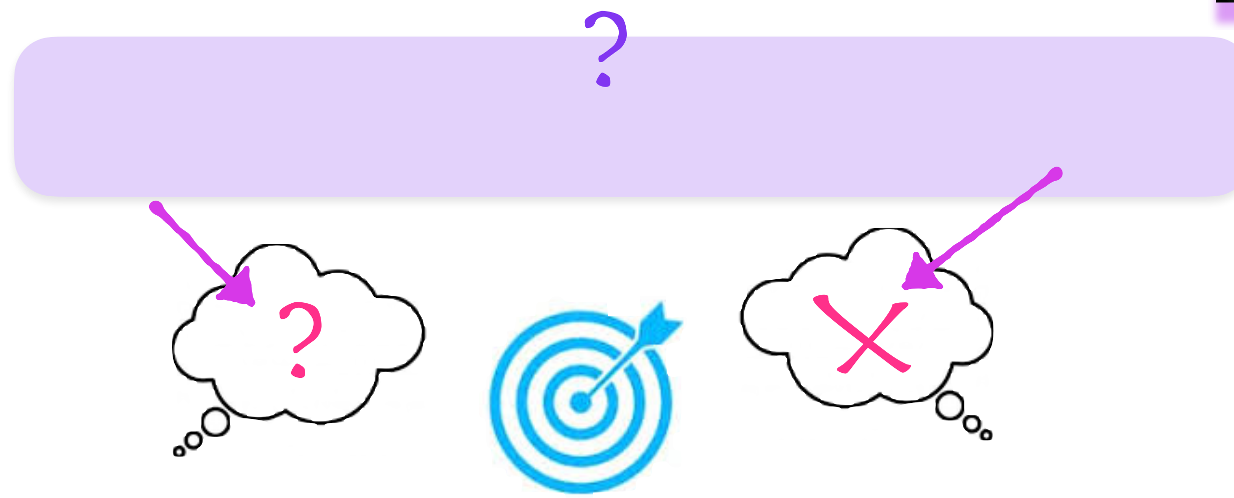
3.9K *wh*-dependencies

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).

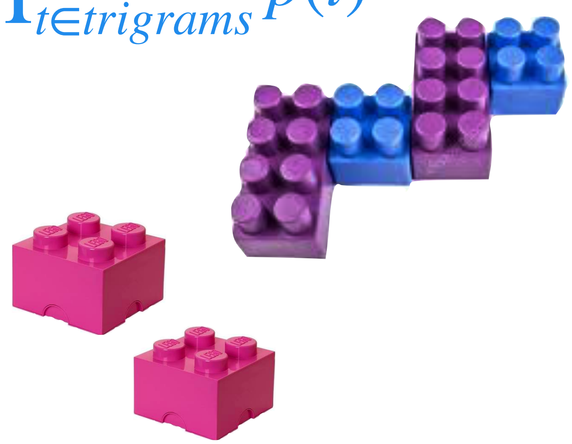
Bates & Pearl 2019, in prep.

Measurable input differences

Who does



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



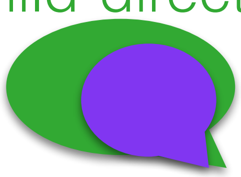
$=$ $\neq$

$0 \leq JSDiv \leq 1$

The input samples

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

High-SES child-directed

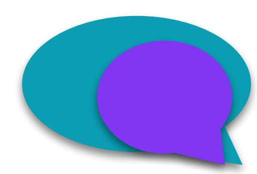


21K *wh*-dependencies

Low-SES child-directed

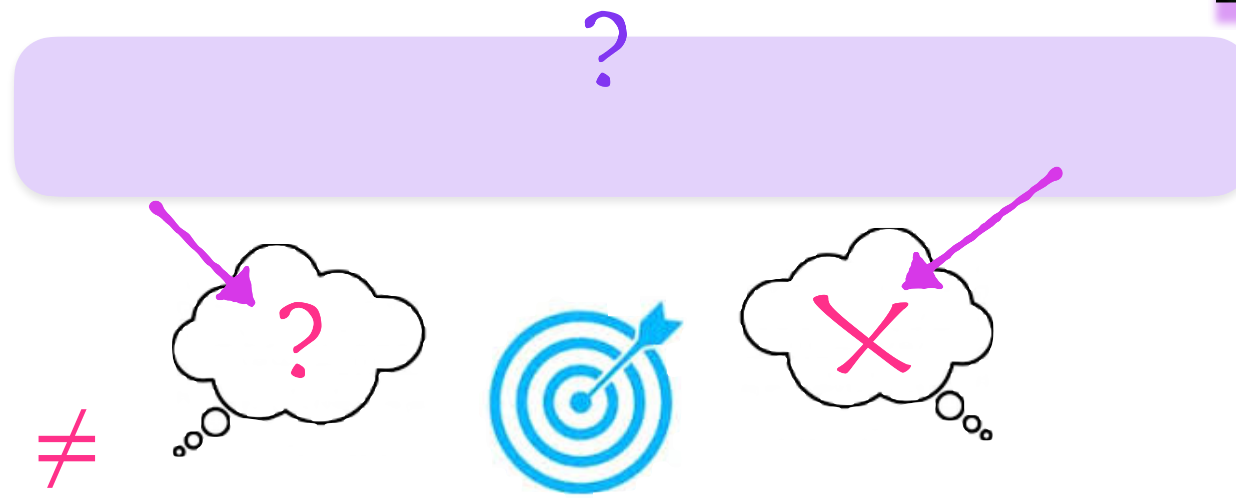
3.9K *wh*-dependencies

High-SES adult-directed

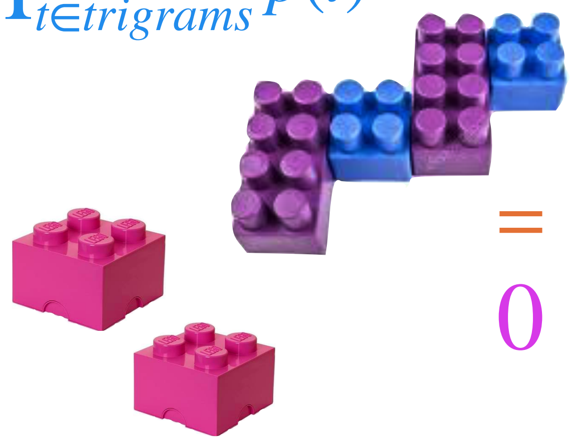


Bates & Pearl 2019, in prep.

Measurable input differences



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



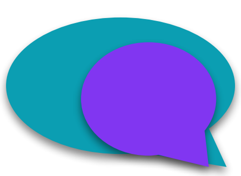
$=$ $\neq$

$$0 \leq JSDiv \leq 1$$

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?

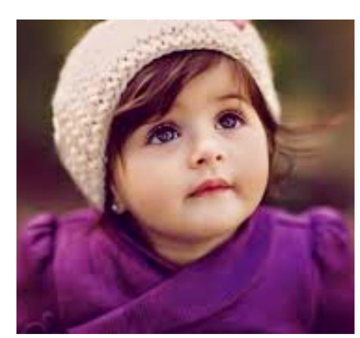
High-SES adult-directed



8.5K *wh*-dependencies

High-SES child-directed

21K *wh*-dependencies



Low-SES child-directed

3.9K *wh*-dependencies

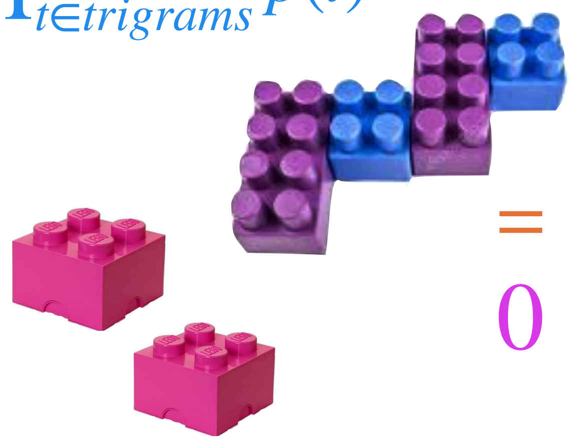
Bates & Pearl 2019, in prep.

Who does 

Measurable input differences

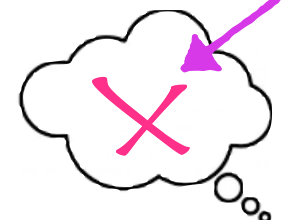
?

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



=

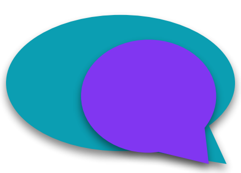
$\neq$



$0 \leq JSDiv \leq 1$

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

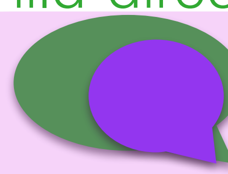


8.5K *wh*-dependencies

High-SES child-directed

SES differences

Low-SES child-directed



21K *wh*-dependencies

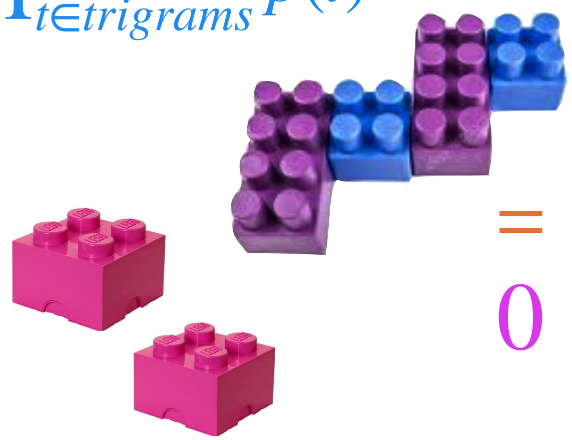
3.9K *wh*-dependencies

Bates & Pearl 2019, in prep.

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 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

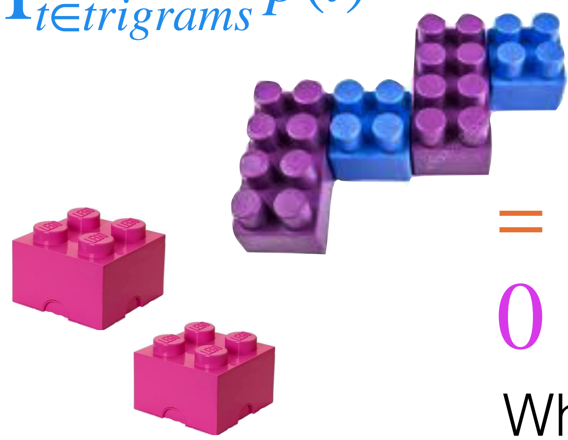
Bates & Pearl 2019, in prep.

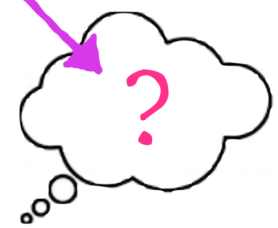
Measurable input differences

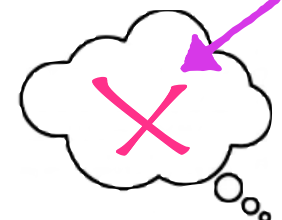
Who does

?

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





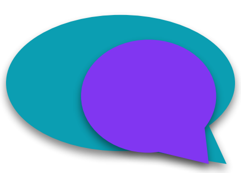


$=$ $\neq$

$0 \leq JSDiv \leq 1$

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

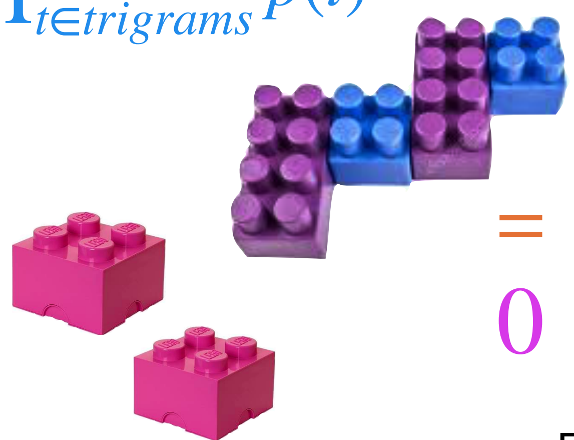
Bates & Pearl 2019, in prep.

Who does

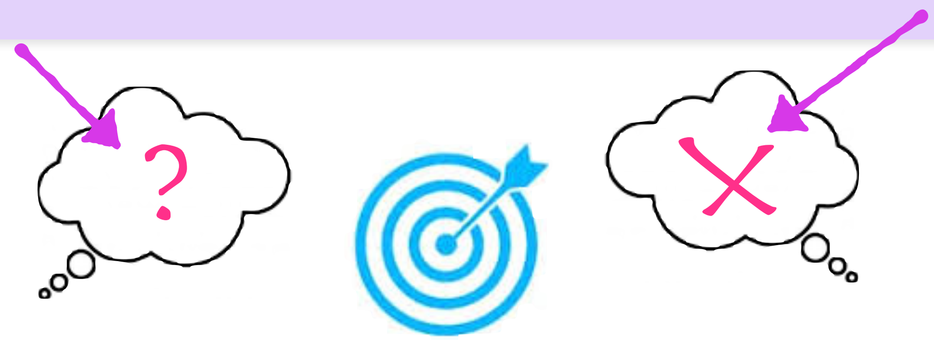
Measurable input differences

?

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



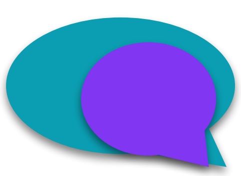
$=$ $\neq$



$$0 \leq JSDiv \leq 1$$

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 High-SES child-directed .00445 Low-SES child-directed



8.5K *wh*-dependencies

21K *wh*-dependencies

3.9K *wh*-dependencies

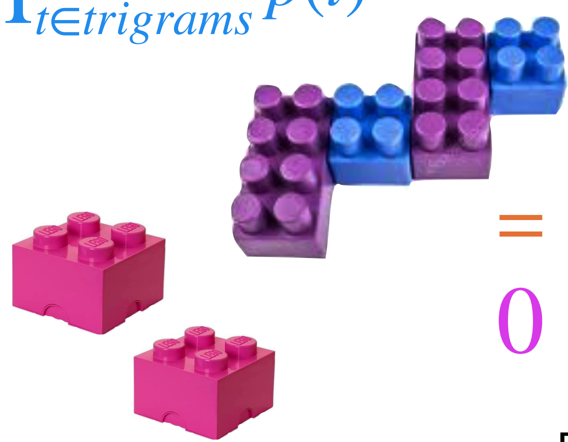
Bates & Pearl 2019, in prep.

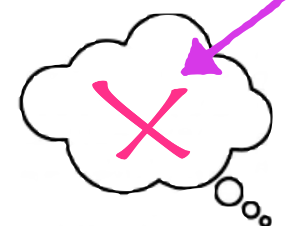
Measurable input differences

Who does

?

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



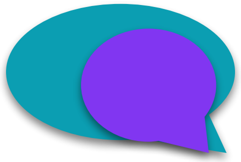


$=$ $\neq$

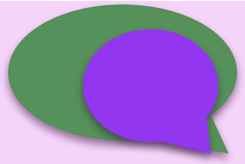
$$0 \leq JSDiv \leq 1$$

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

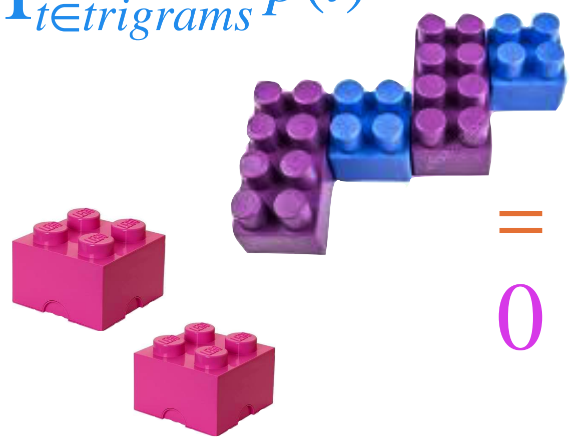
Bates & Pearl 2019, in prep.

Who does

Measurable input differences

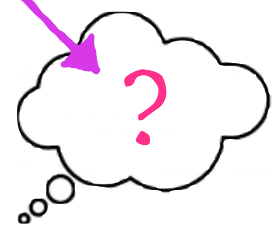
?

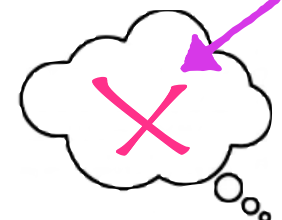
$\prod_{t \in trigrams} p(t)$



$=$

$\neq$





$0 \leq JSDiv \leq 1$

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

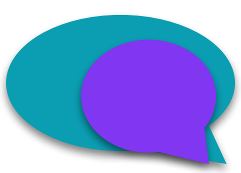
High-SES adult-directed

directed at who differences

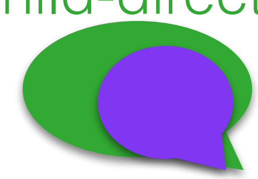
High-SES child-directed

SES differences

Low-SES child-directed



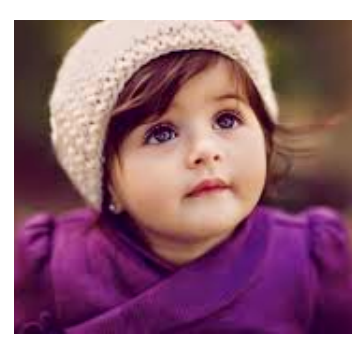
8.5K *wh*-dependencies



21K *wh*-dependencies

$\approx$

3.9K *wh*-dependencies

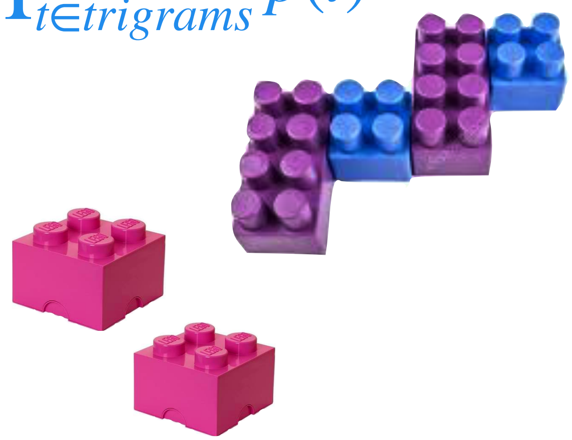


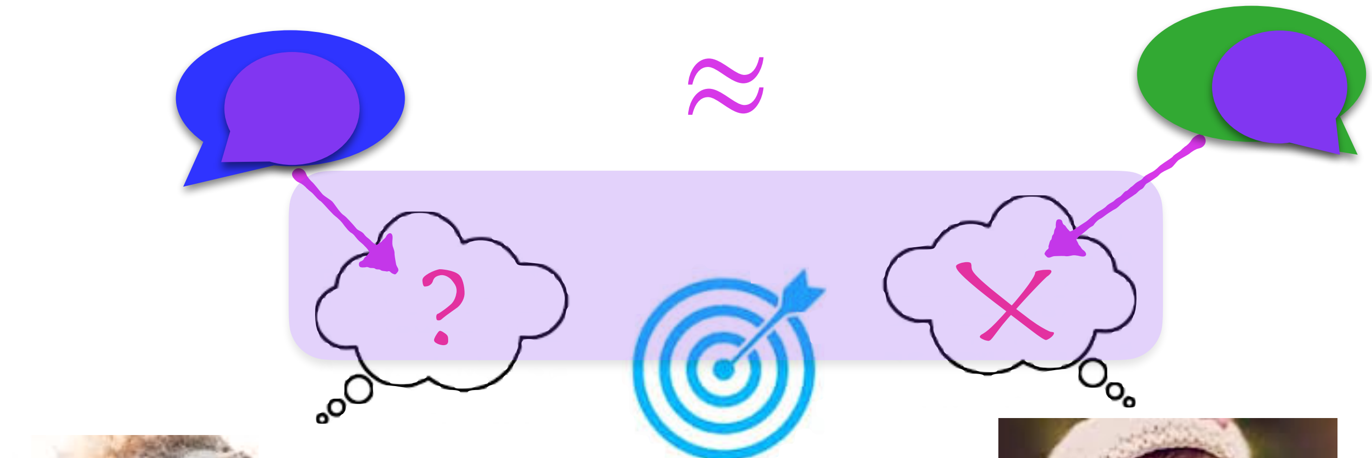
Bates & Pearl 2019, in prep.

Meaningful input differences

Who does

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



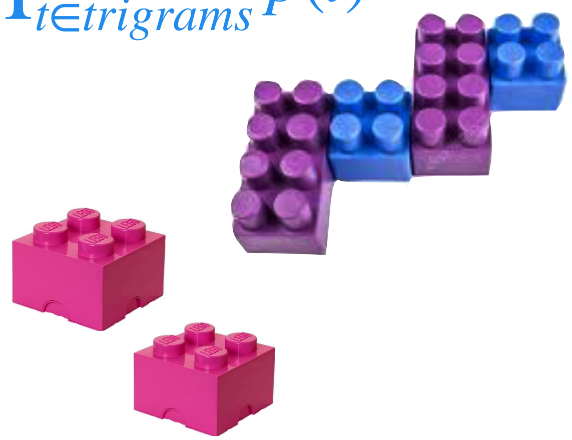


But does this part of the input act differently? That is, are any differences (even if they're small) meaningful?

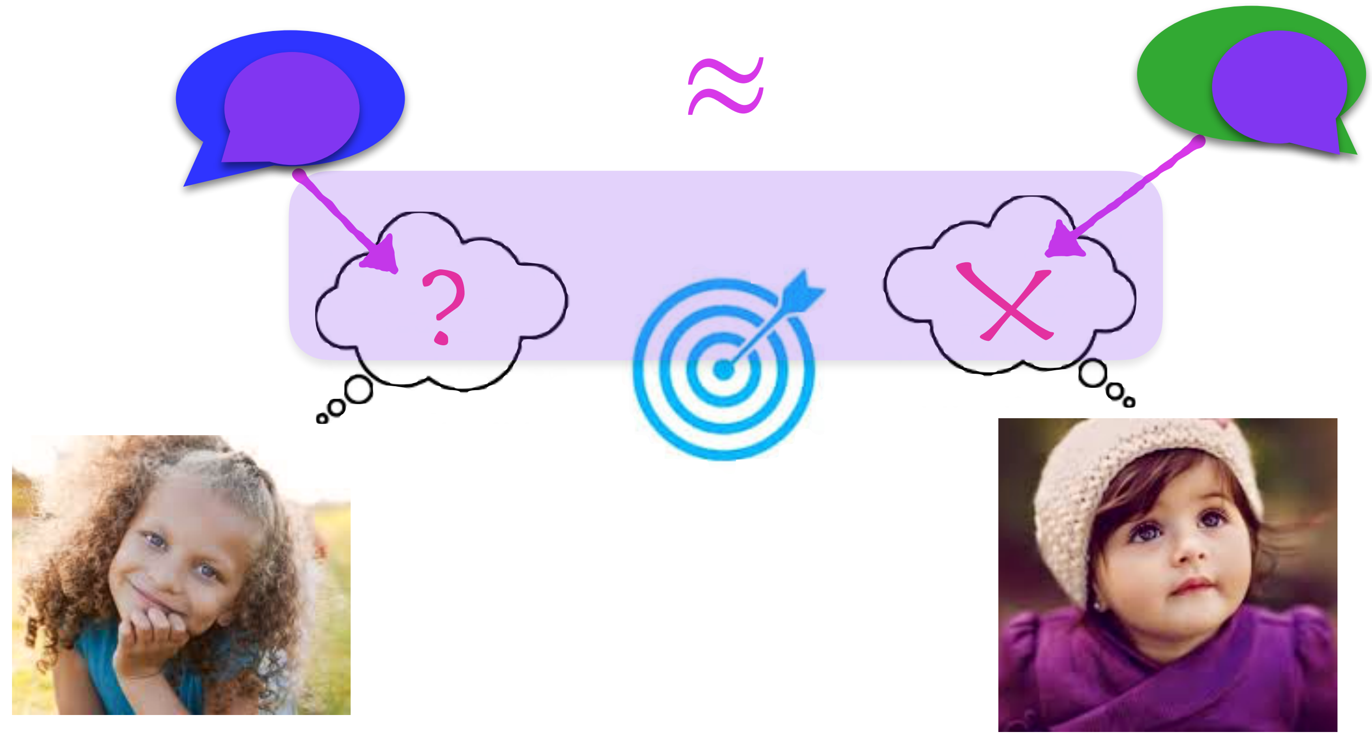
They might be — small differences in the input distribution might snowball into learning outcome differences.

Bates & Pearl 2019, in prep.

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



Meaningful input differences



But does this part of the input act differently? That is, are any differences (even if they're small) meaningful?

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

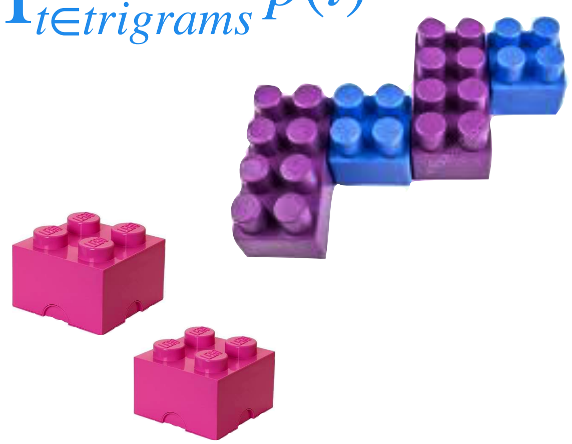
wh-dependencies

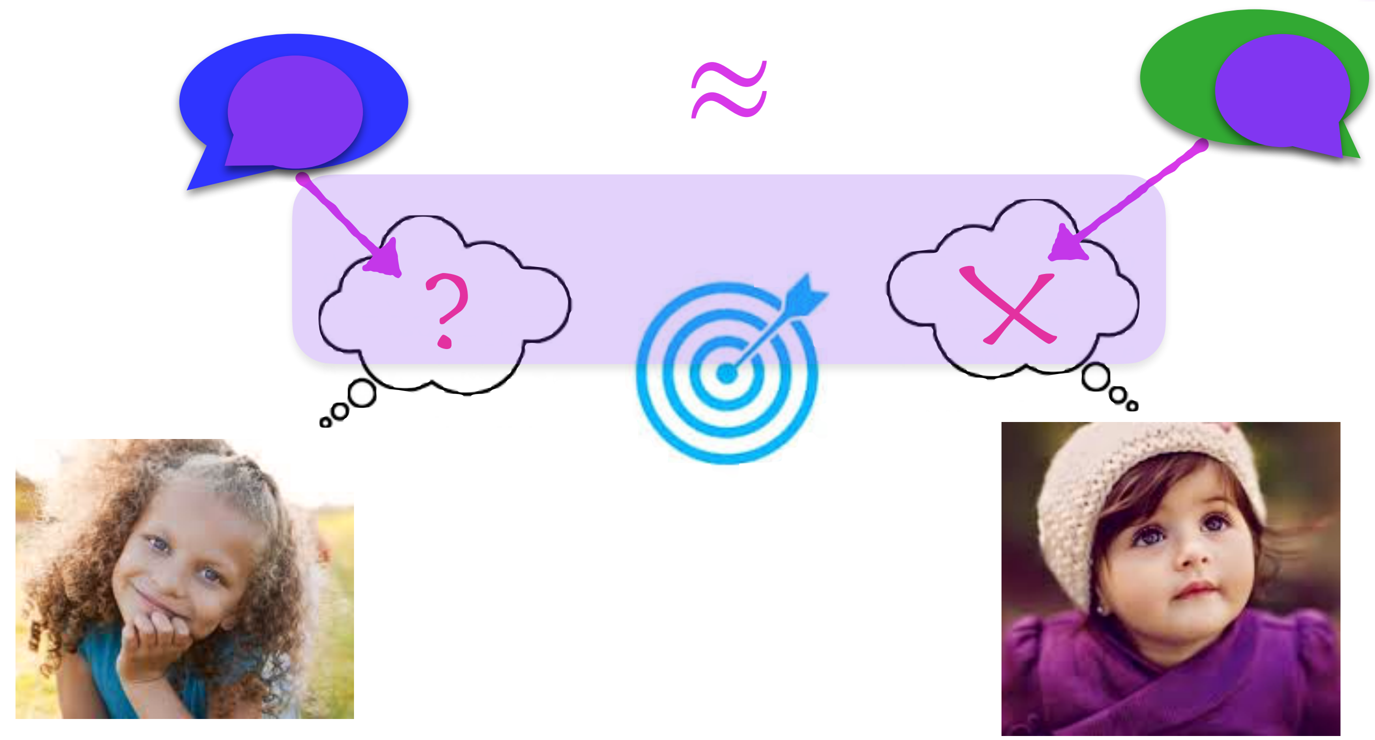
76.7%	*start*-IP-VP-*end* What did Lily read $\_\__{what}$?	75.5%
10.3%	*start*-IP-*end* What $\_\__{what}$ happened?	12.8%

Bates & Pearl 2019, in prep.

Meaningful input differences

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





But does this part of the input act differently? That is, are any differences (even if they're small) meaningful?

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

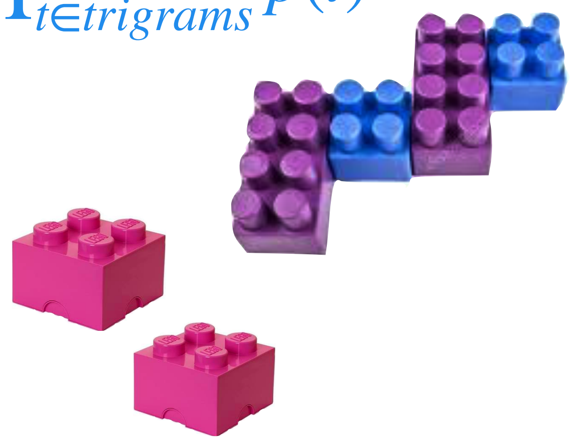
syntactic trigrams

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

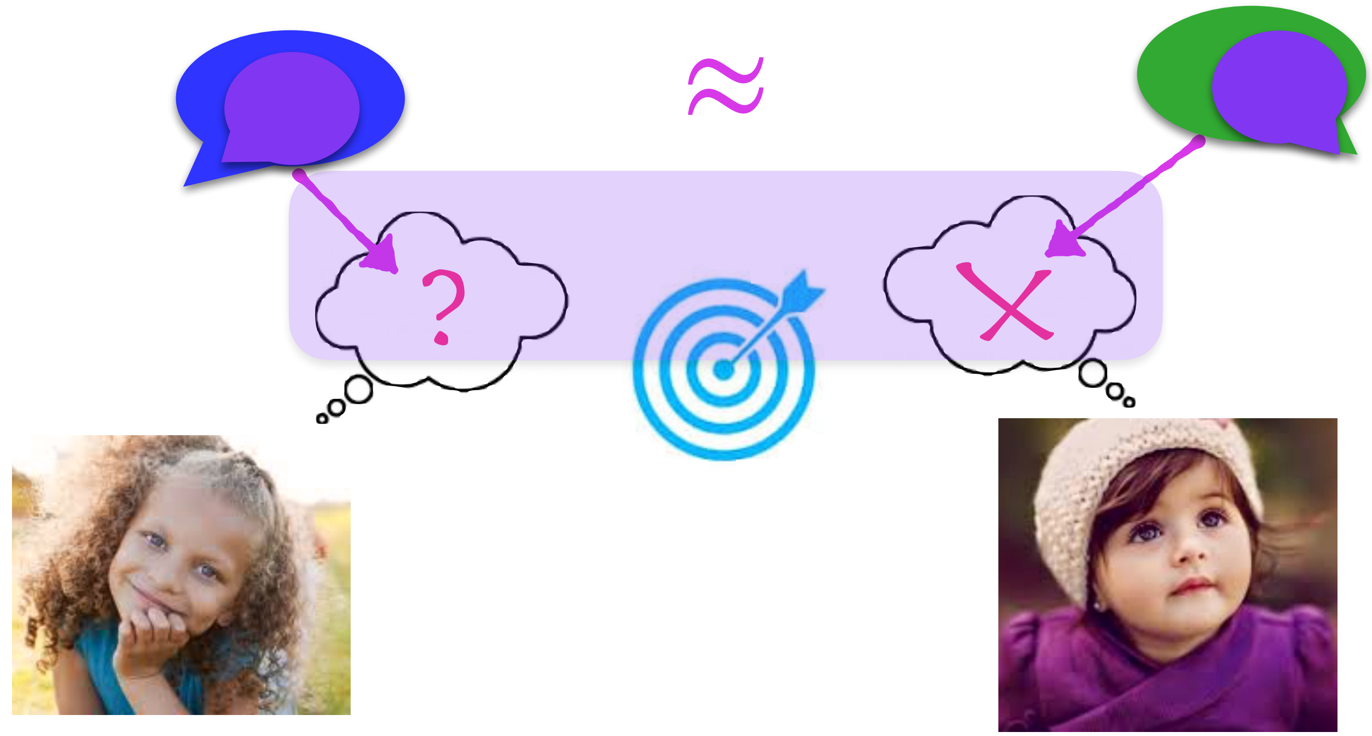
Bates & Pearl 2019, in prep.

Who does

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



Meaningful input differences



Let's use developmental computational modeling to find out.

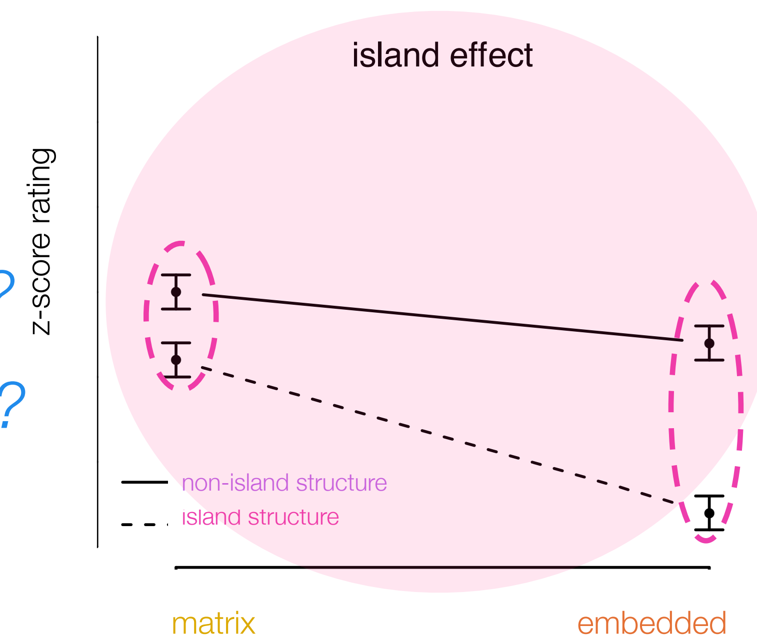
Bates & Pearl 2019, in prep.

Meaningful input differences

Who does

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

Judgments from a modeled child learning from the same amount of data as low-SES children seem to, with those data having the same composition as low-SES child-directed speech data.



Who [non-island]?

Who [island]?

Looking for superadditivity as the sign of syntactic islands knowledge

Who __who?

Who [CP… __who]?

Bates & Pearl 2019, in prep.

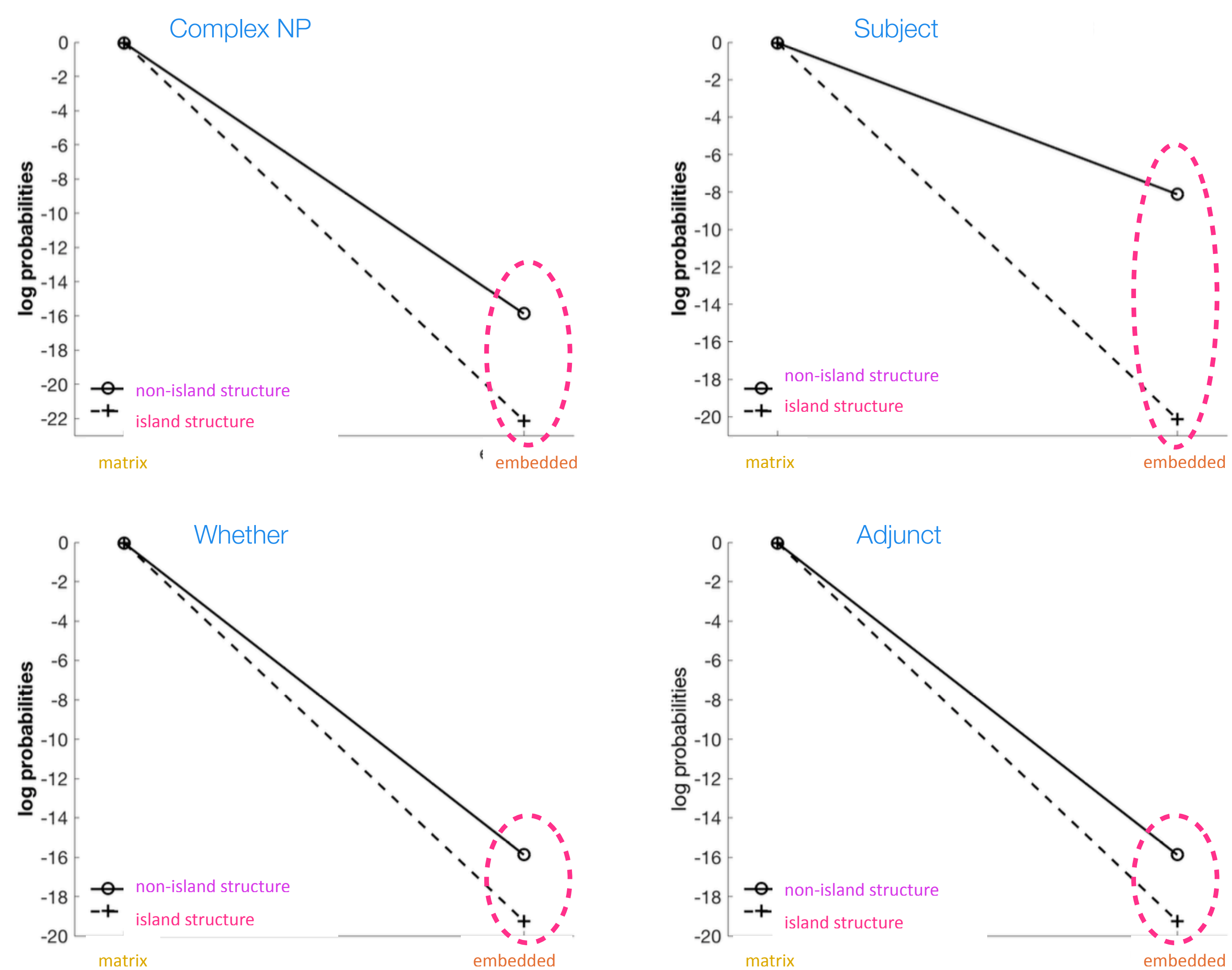
Meaningful input differences

Who does

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

Judgments from a modeled child learning from the same amount of data as low-SES children seem to, with those data having the same composition as low-SES child-directed speech data.

Superadditivity for all four islands!



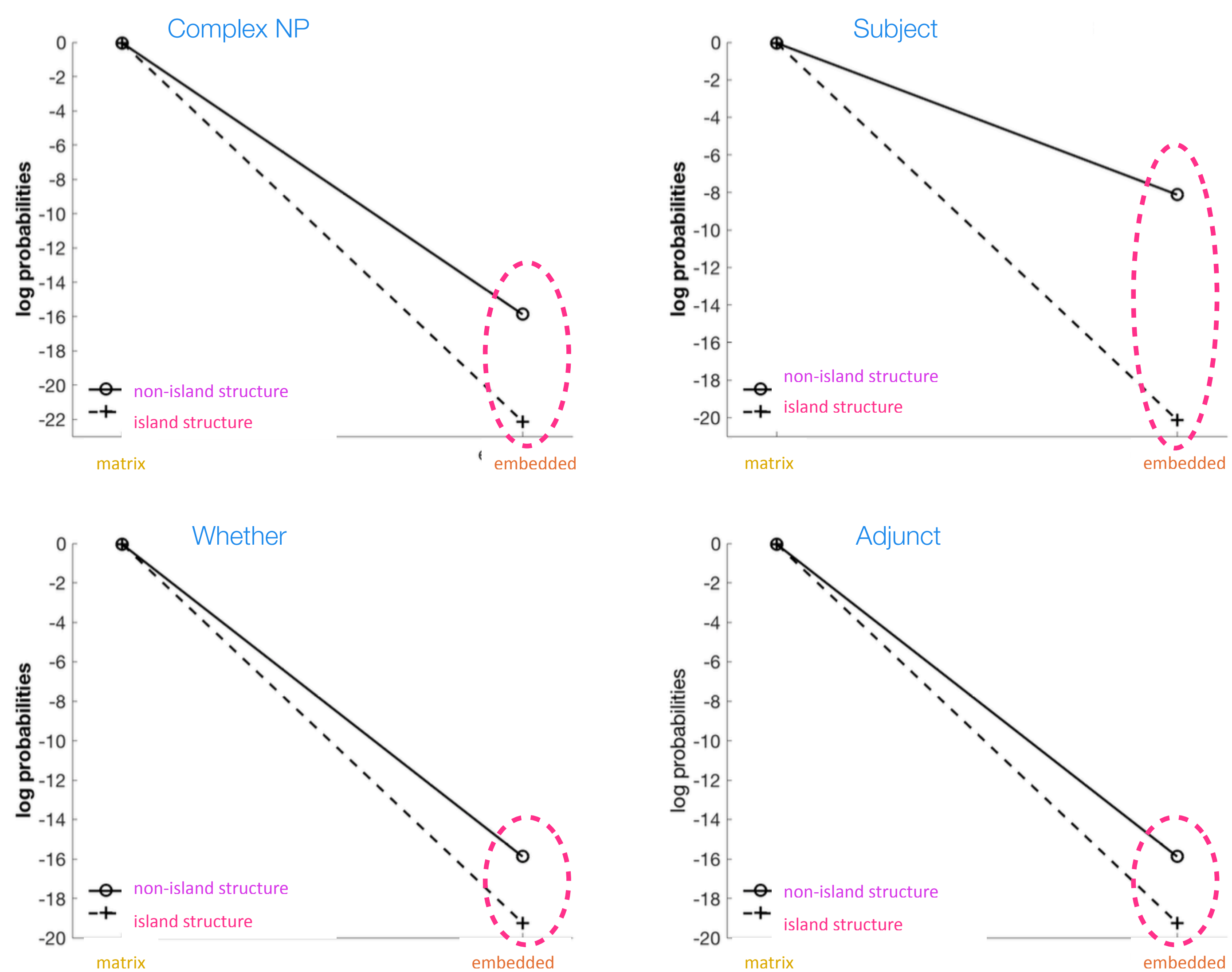
Bates & Pearl 2019, in prep.

Meaningful input differences

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

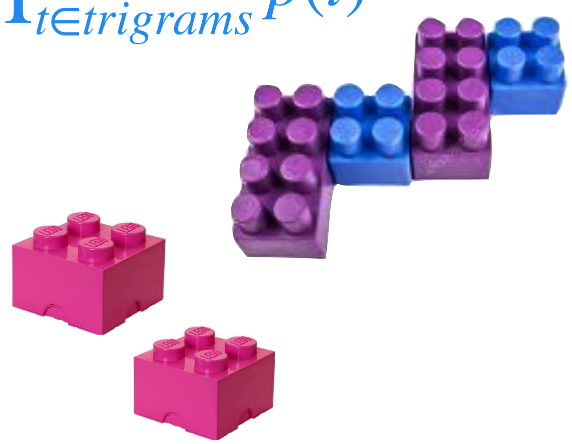
Judgments from a modeled child learning from the same amount of data as low-SES children seem to, with those data having the same composition as low-SES child-directed speech data.

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

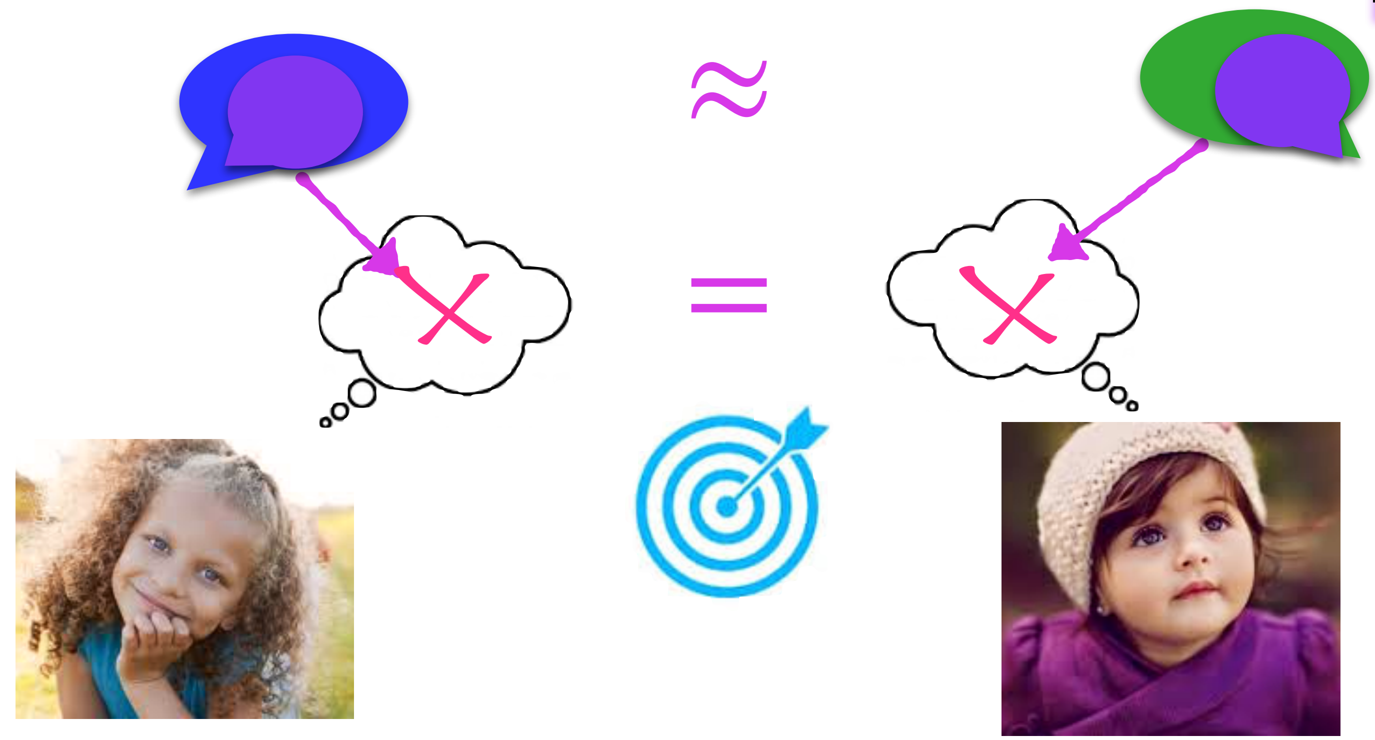


Bates & Pearl 2019, in prep.

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



No meaningful input differences predicted

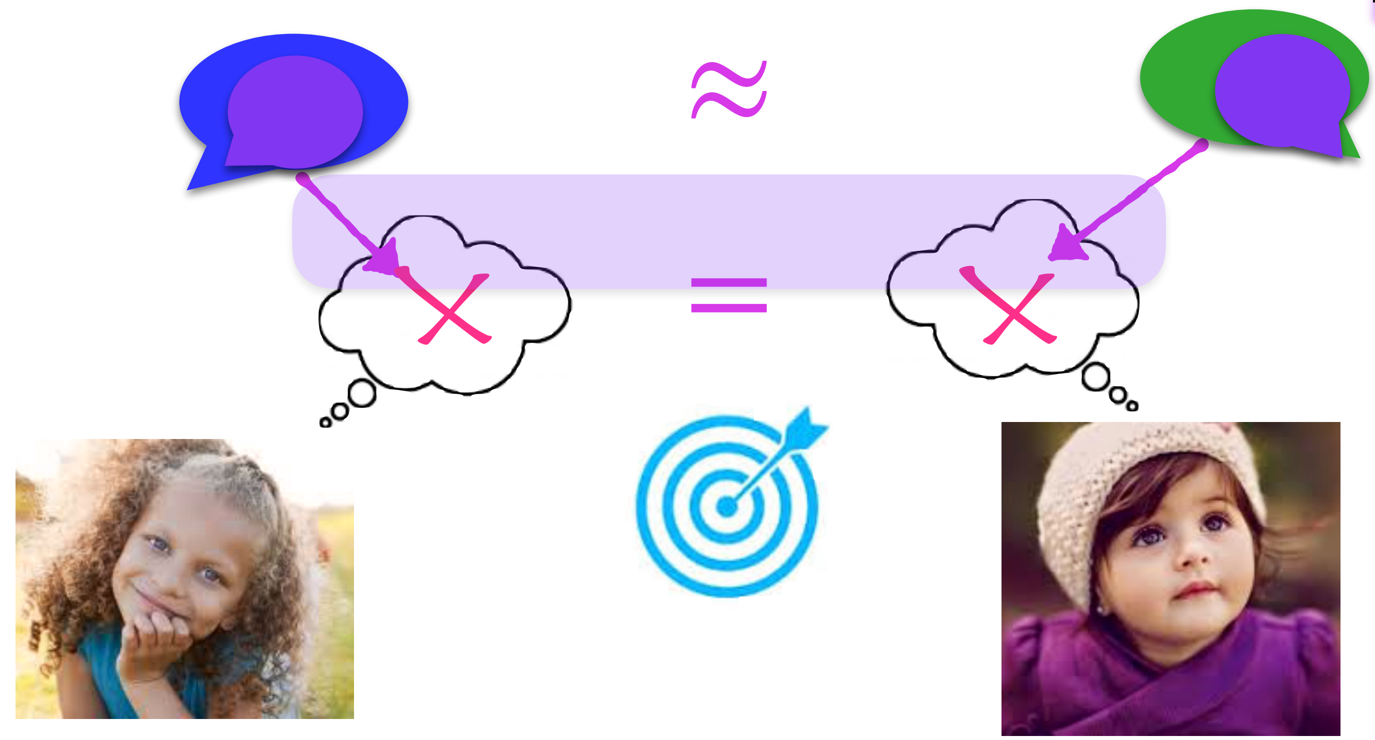


So, our developmental computational model predicts no meaningful input differences across SES when it comes to learning this syntactic island knowledge from this part of the input.

Bates & Pearl 2019, in prep.

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

No meaningful input differences predicted

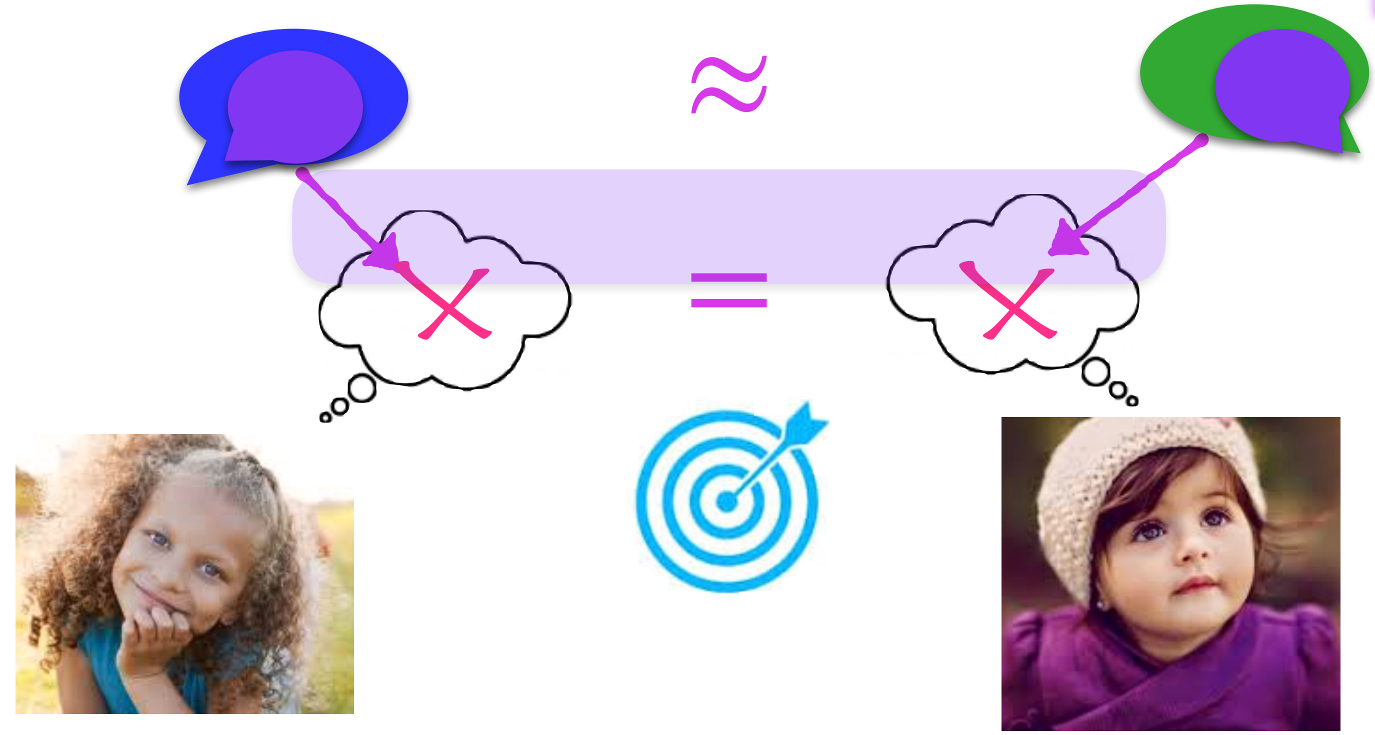


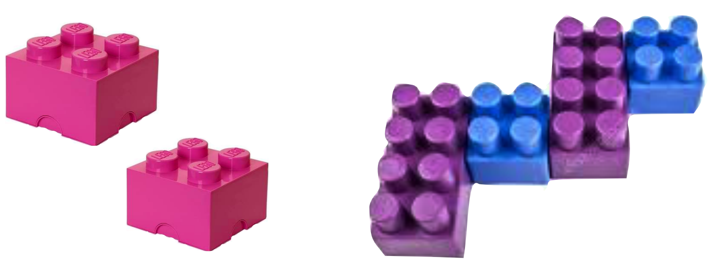
Useful: Because we know how the input is predicted to cause the knowledge to develop, we know which building blocks are particularly important.

Bates & Pearl 2019, in prep.

No meaningful input differences predicted

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





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

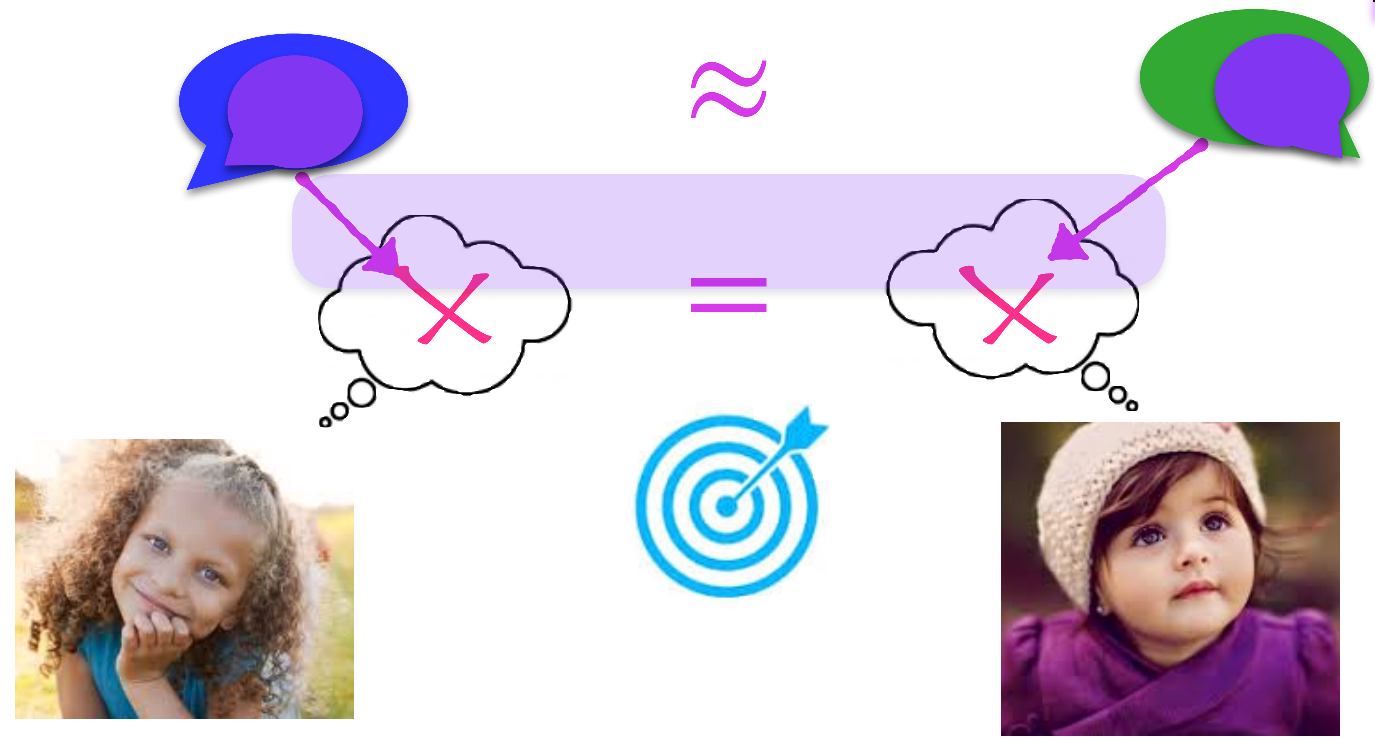
whether *What does the teacher wonder [whether Lily forgot __]? embedded | island

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

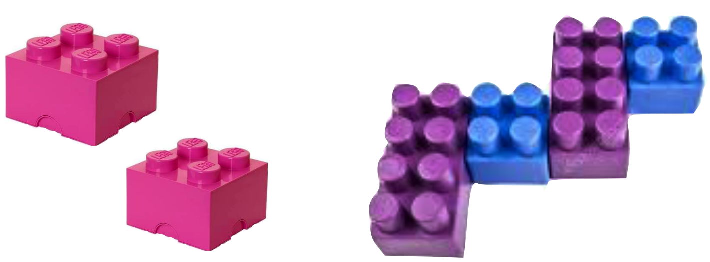
Bates & Pearl 2019, in prep.

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

No meaningful input differences predicted



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.

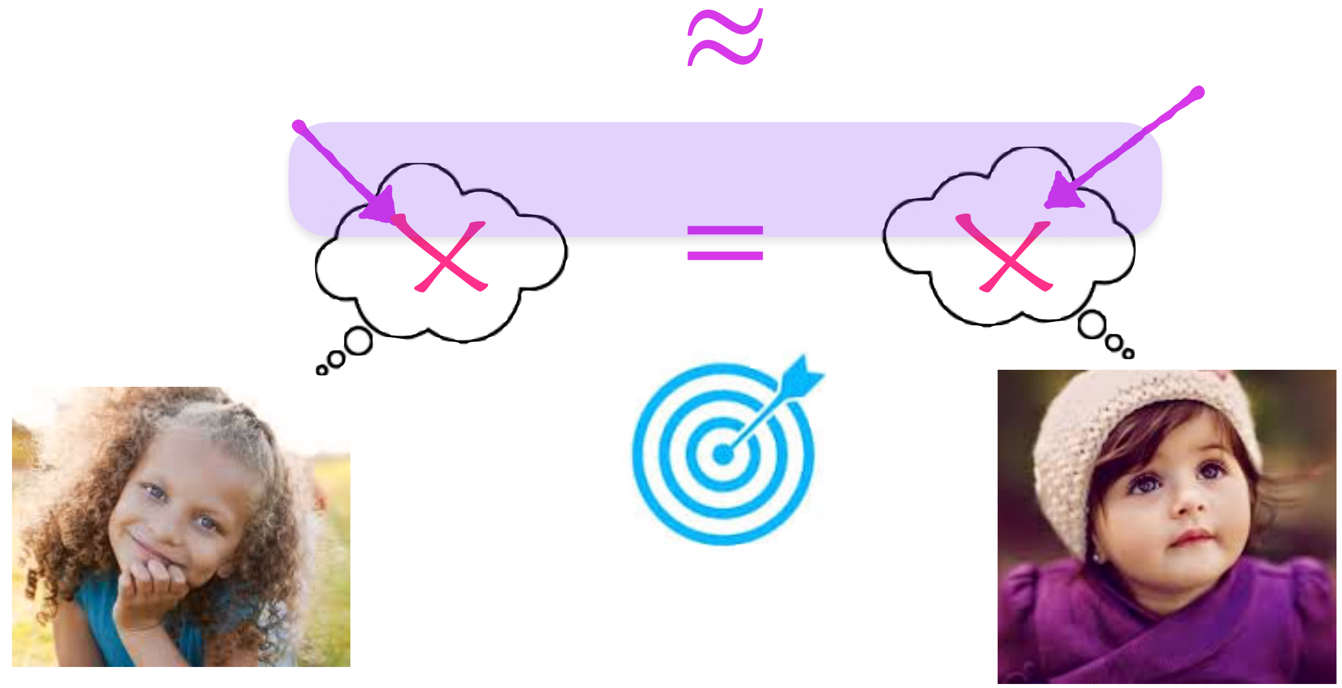


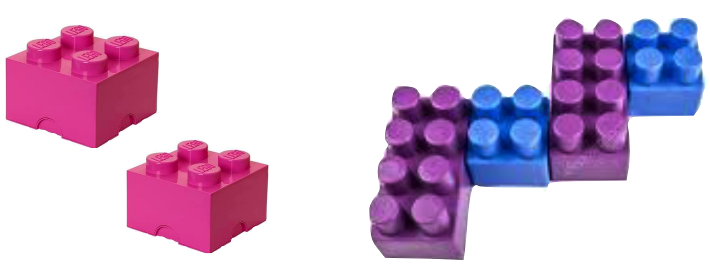
	start-IP-VP-CP_{that}- IP-VP-*end*	embedded | non-island
whether	* *start*-IP-VP-$CP_{whether}$-IP-VP-*end*	embedded | island
adjunct	* *start*-IP-VP-CP_{if}- IP-VP-*end*	embedded | island

Bates & Pearl 2019, in prep.

No meaningful input differences predicted

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





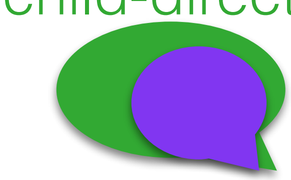
So, children need to encounter grammatical *wh*-dependencies that involve CP_{that}. These are actually pretty rare in child-directed speech.

Low-SES child-directed

2 instances of 3.9K (=.05%)

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

High-SES child-directed



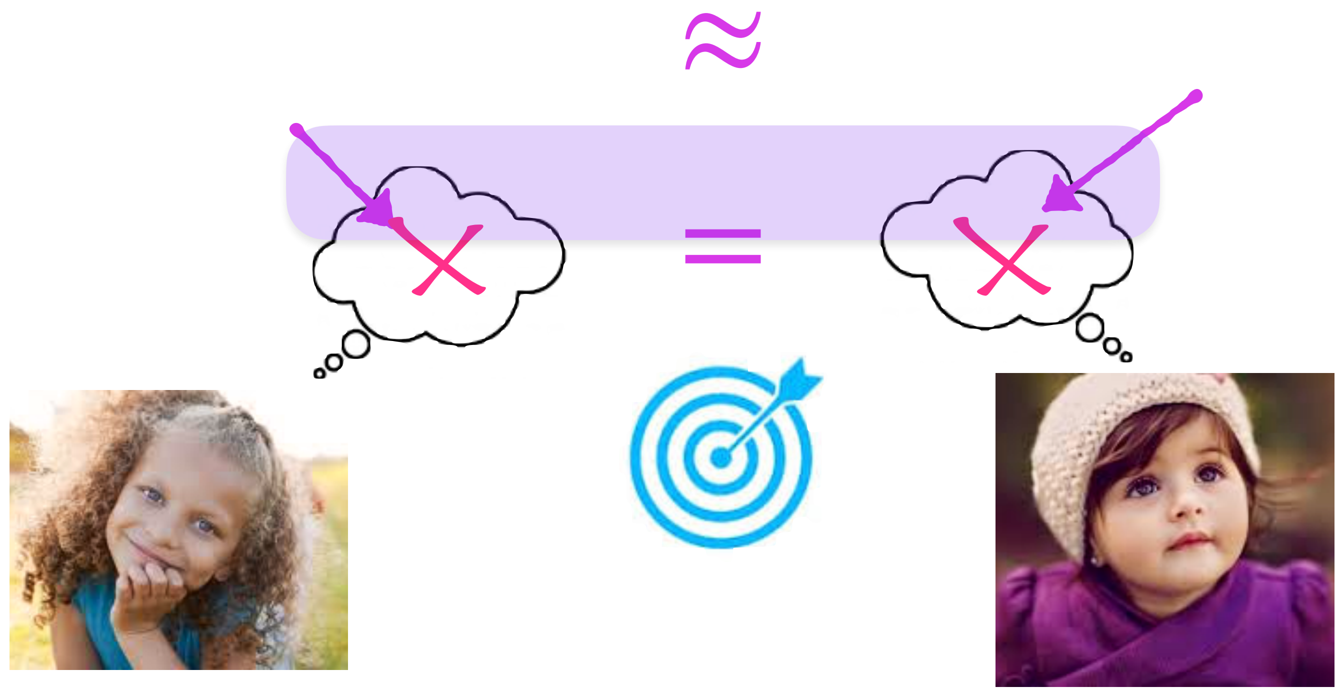
2 instances of 21K (<.01%)

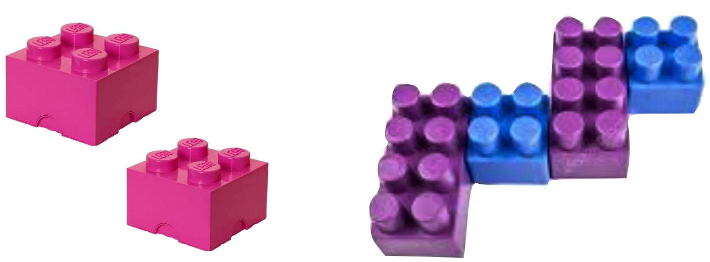
What do you think that Jack read $\_\__{\text{what}}$?

Bates & Pearl 2019, in prep.

No meaningful input differences predicted

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





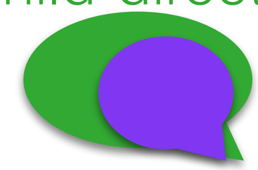
But with enough input (over several years), even these rare cases are predicted to support learning.

Low-SES child-directed

2 instances of 3.9K (=.05%)

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

High-SES child-directed



2 instances of 21K (<.01%)

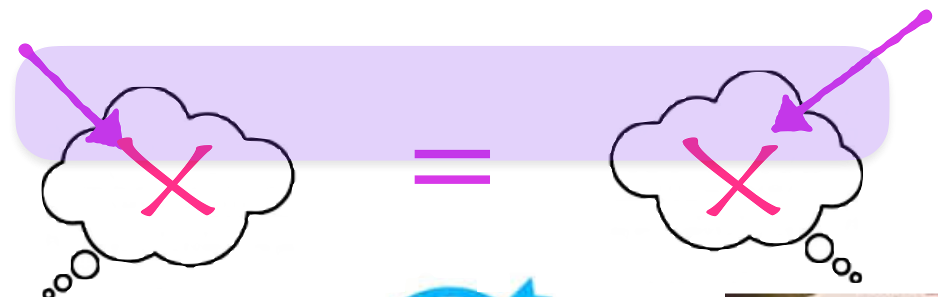
What do you think that Jack read $\_\__{what}$?

Bates & Pearl 2019, in prep.

No meaningful input differences predicted

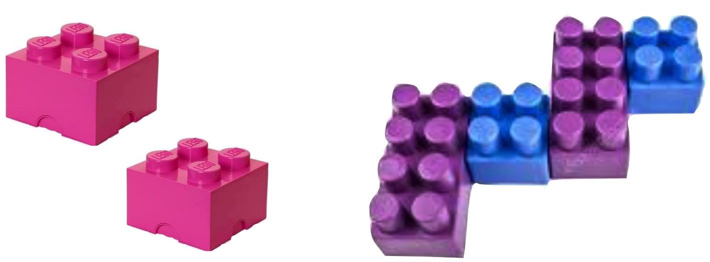
$\prod_{t \in trigrams} p(t)$

≈



=

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

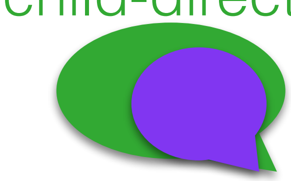


Low-SES child-directed

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

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

High-SES child-directed



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

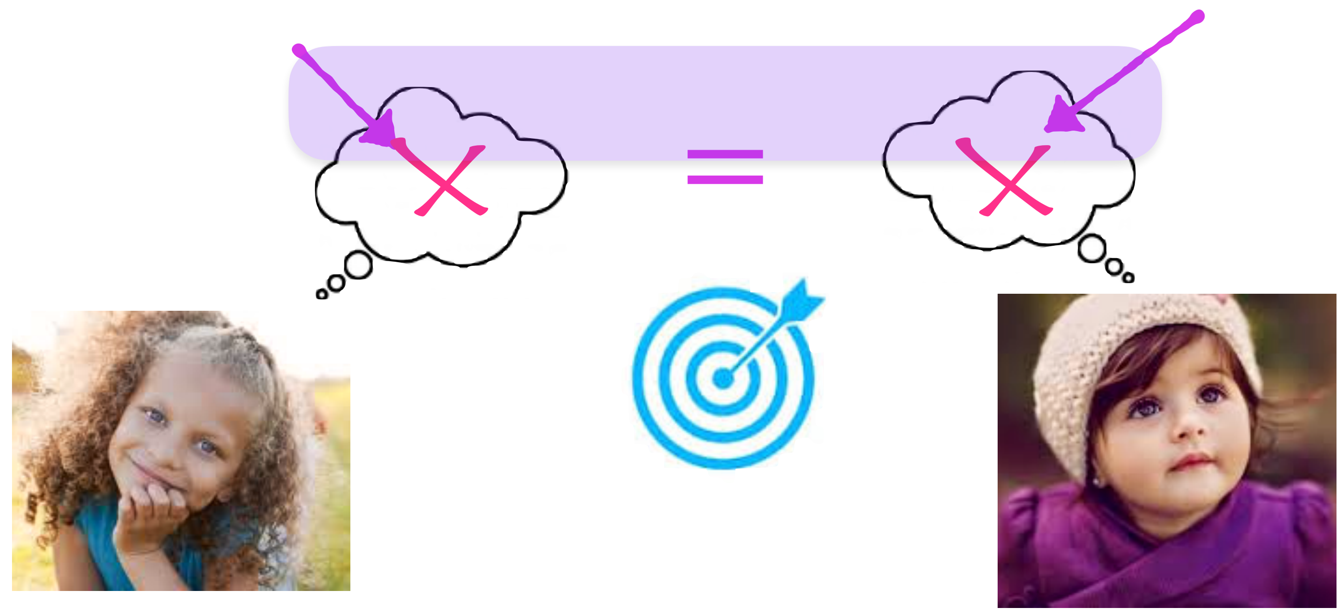
What do you think that Jack read $\_\__{what}$?

Bates & Pearl 2019, in prep.

Who does

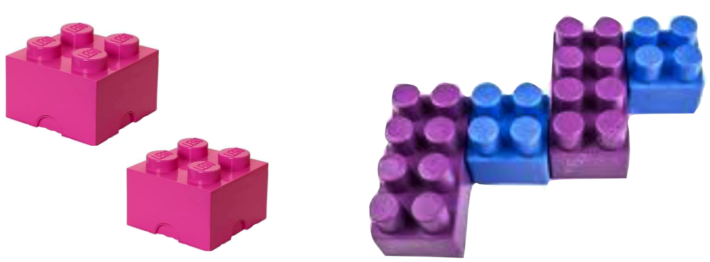
$\prod_{t \in trigrams} p(t)$

No meaningful input differences predicted



$\approx$

$=$



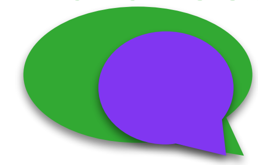
Interesting: The *wh*-dependency with this building block is typically judged to be ungrammatical in the high-SES dialect (a *that*-trace violation).

Low-SES child-directed

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

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

High-SES child-directed



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

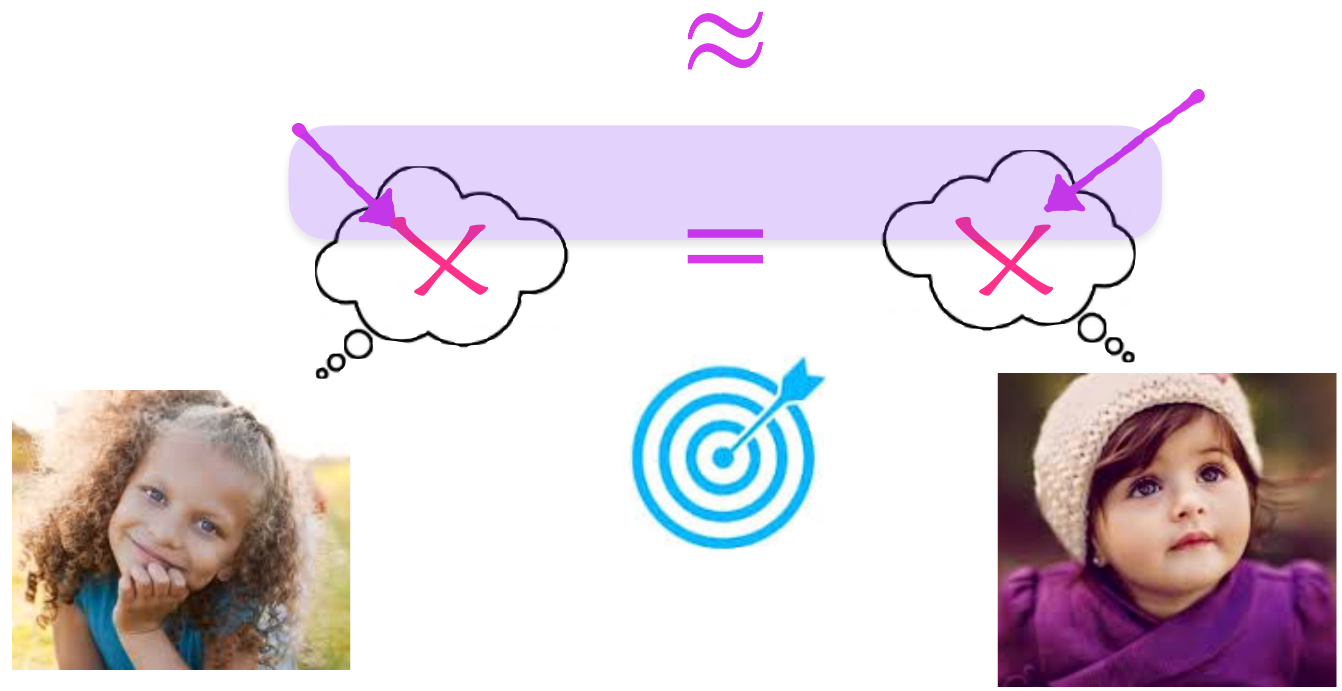
What do you think that Jack read $\_\__{what}$?

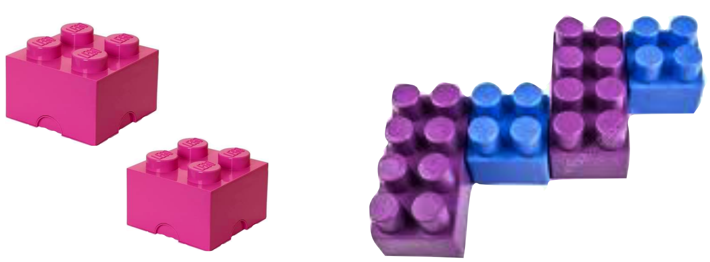
Bates & Pearl 2019, in prep.

Who does

No meaningful input differences predicted

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





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

Low-SES child-directed

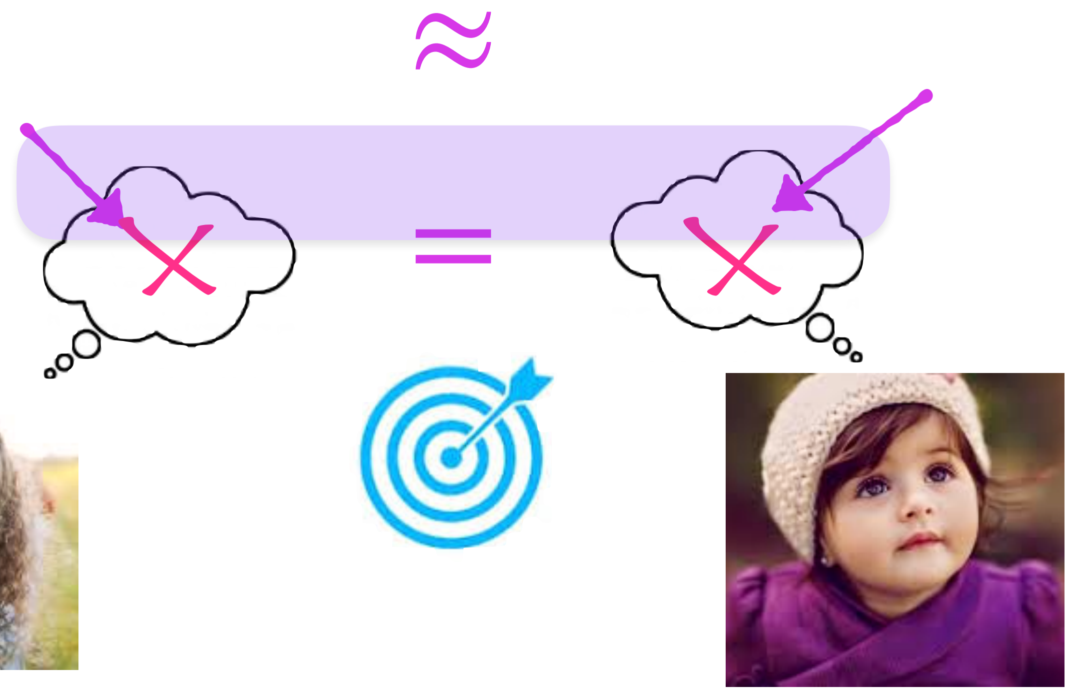
High-SES child-directed

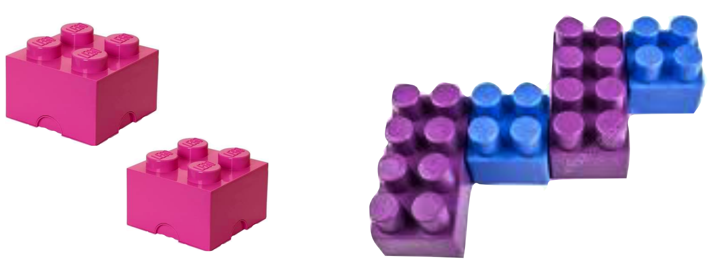
What do you think that $\_\__{what}$ happens?

Bates & Pearl 2019, in prep.

No meaningful input differences predicted

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





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

Low-SES child-directed

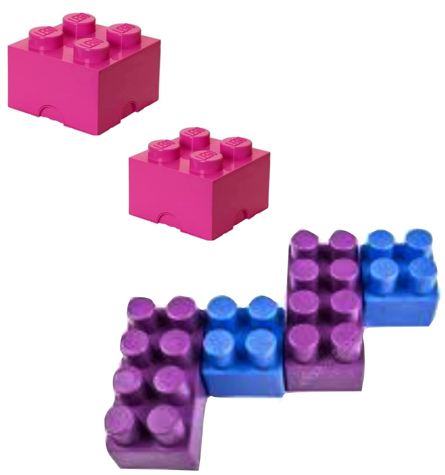
High-SES child-directed

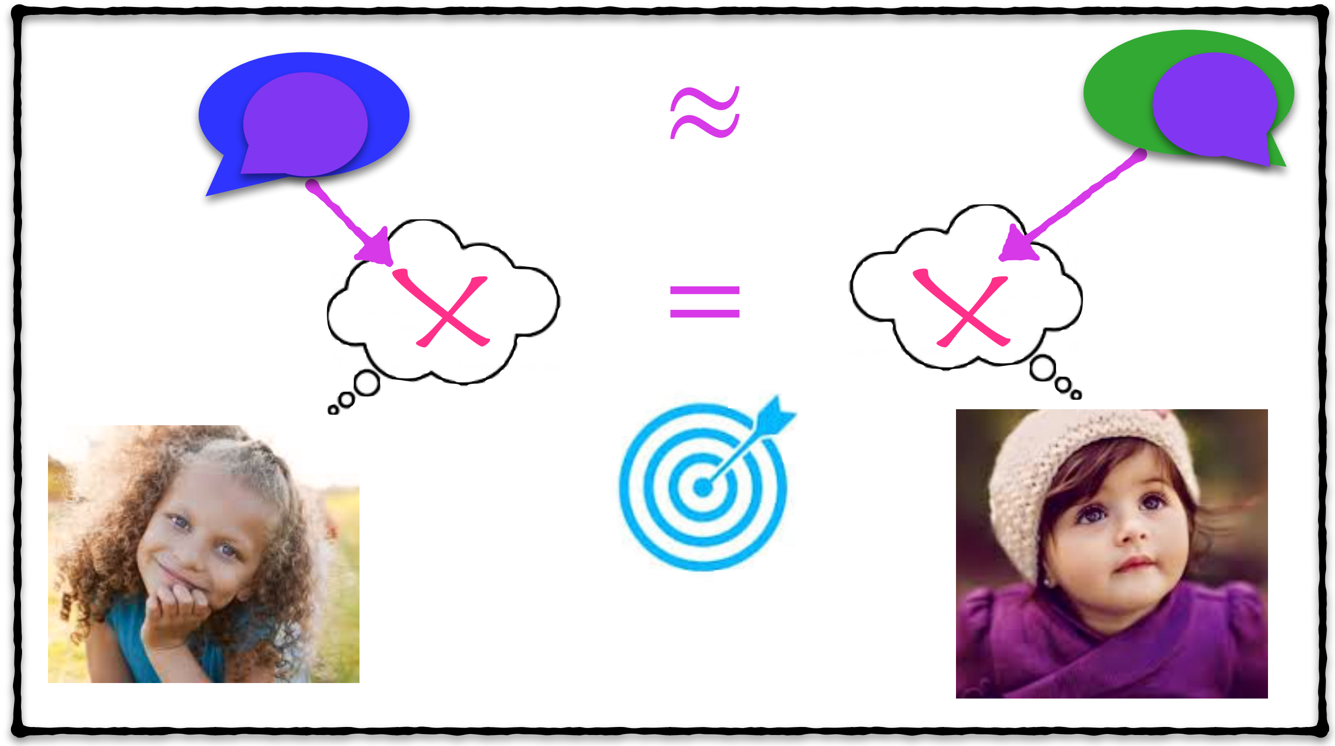
What do you think that $\_\__{what}$ happens?

Bates & Pearl 2019, in prep.

Who does

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





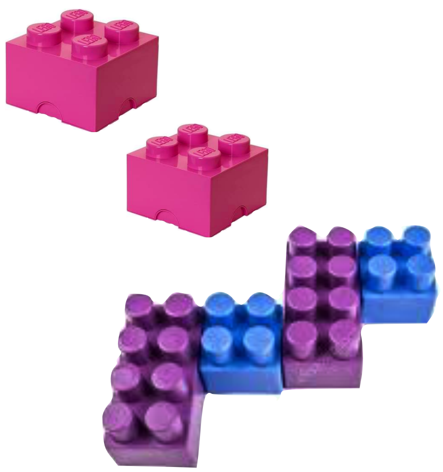
No meaningful input differences predicted

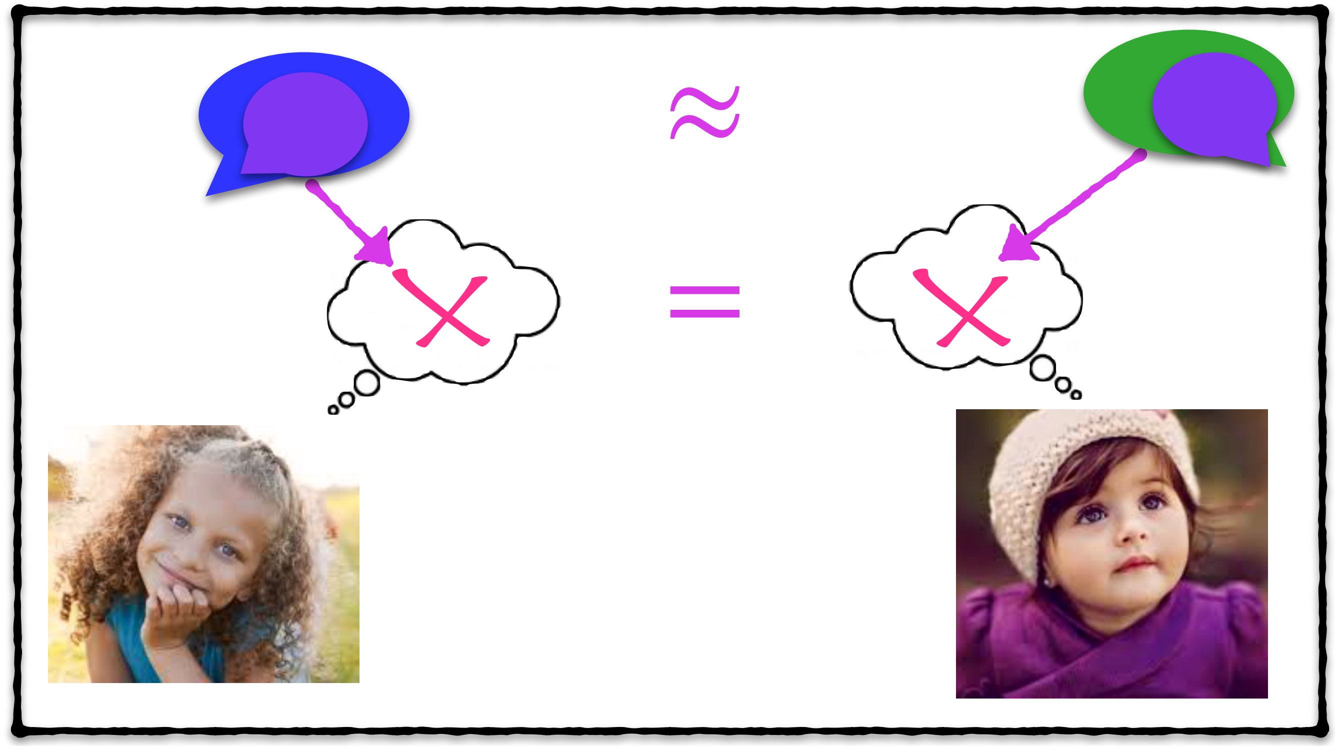
So now what?

Bates & Pearl 2019, in prep.

Who does

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



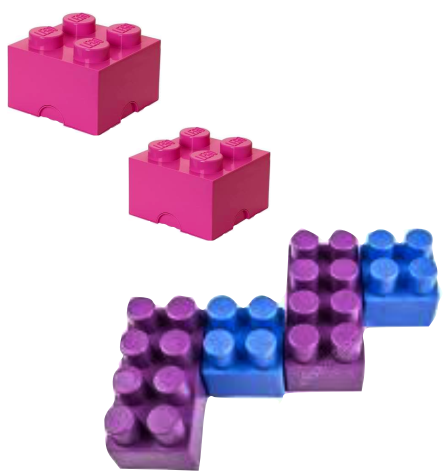


No meaningful input differences predicted

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

Bates & Pearl 2019, in prep.

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



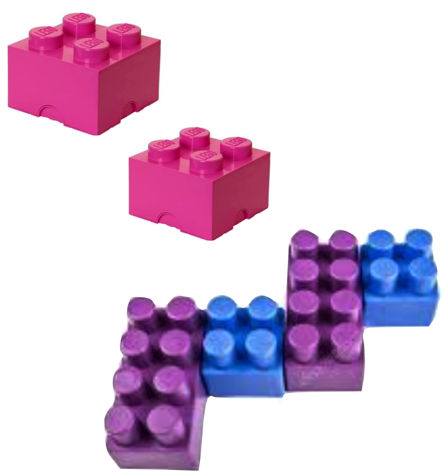
No meaningful input differences predicted

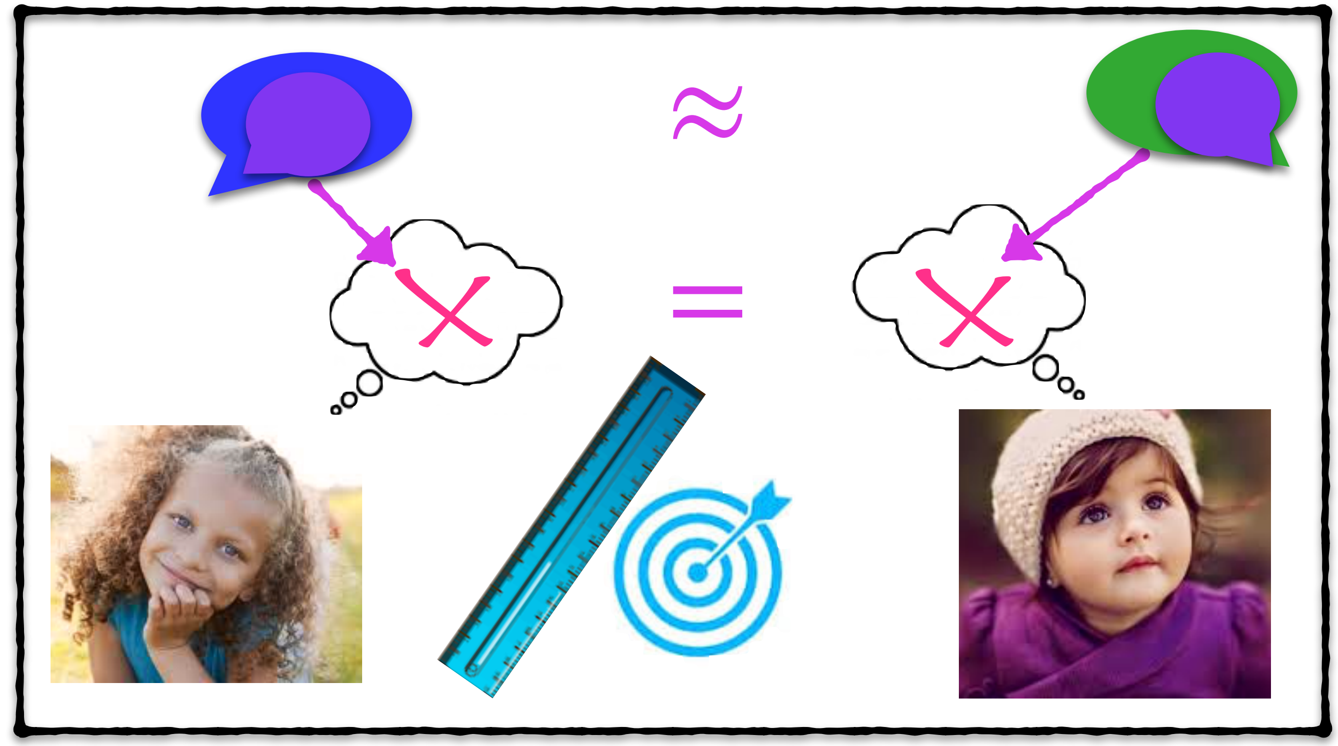
One caveat: If there are in fact differences, it could be due to other factors besides input differences.

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, in prep.

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



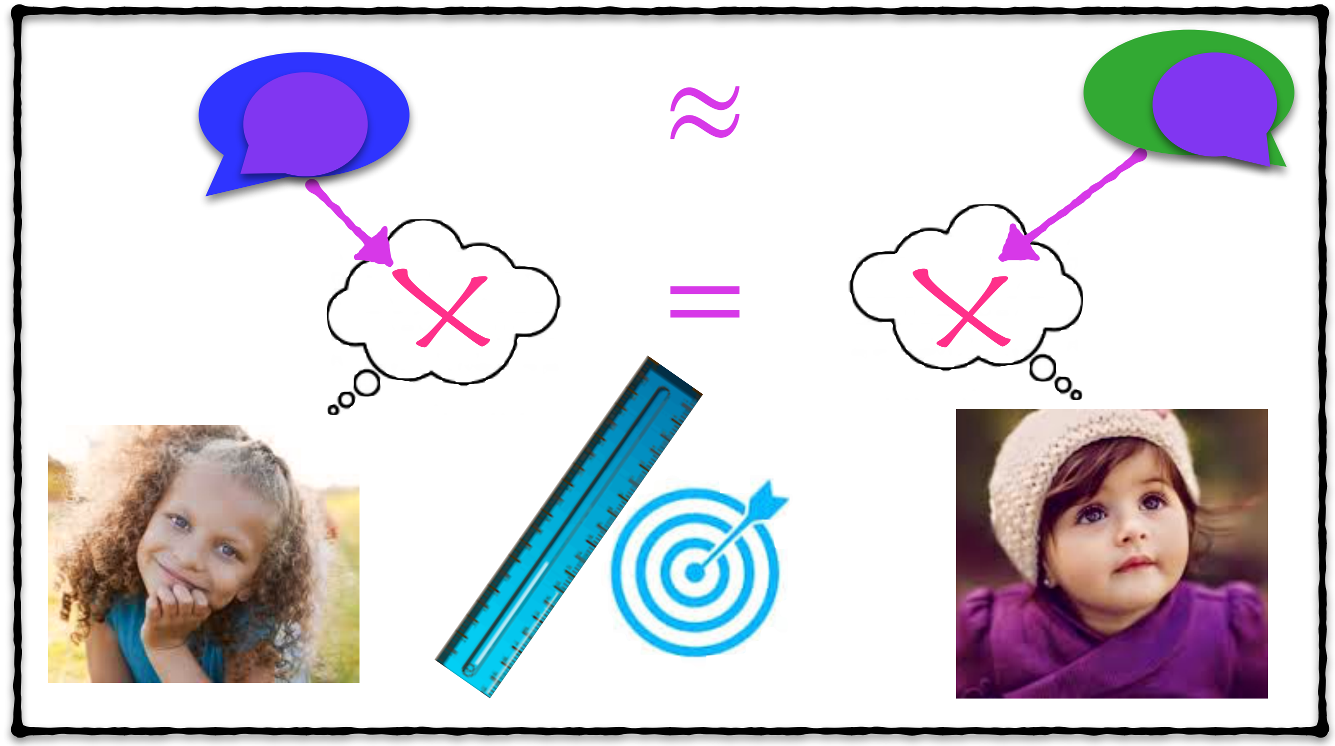


No meaningful input differences predicted

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.

Bates & Pearl 2019, in prep.

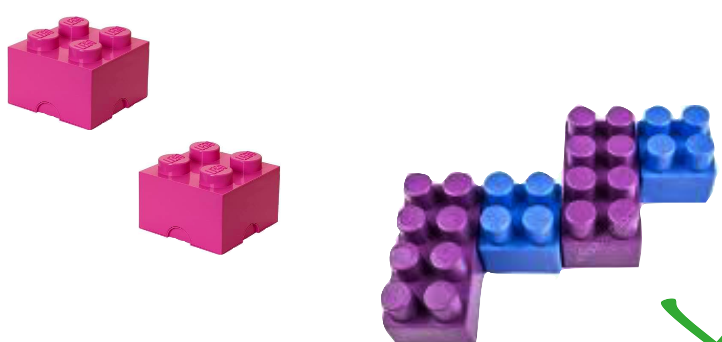
$\prod_{t \in trigrams} p(t)$

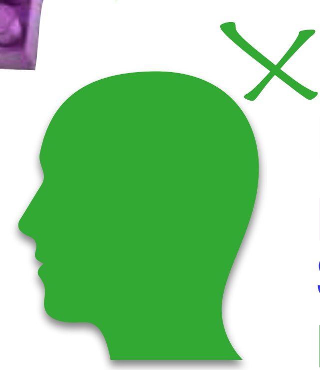


Building block origins

Low-SES
child-directed

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



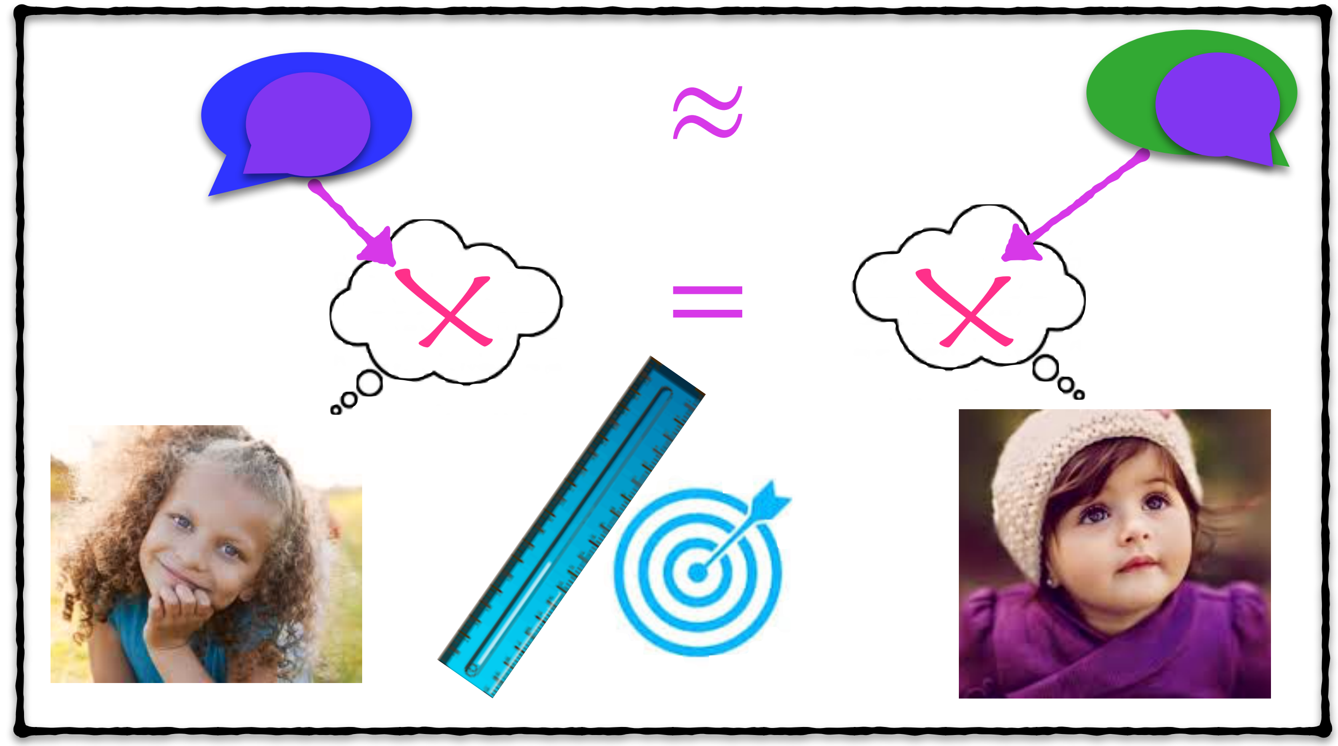


Remember that key building blocks involving CP_{that} are predicted to come from a particular *wh*-dependency in low-SES child-directed speech that's ungrammatical in the high-SES dialect.

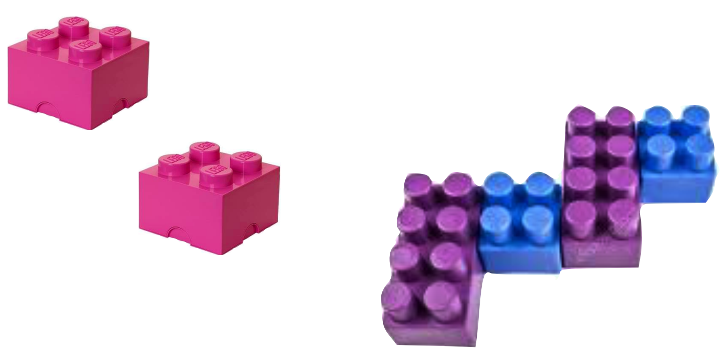
Bates & Pearl 2019, in prep.

Who does

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

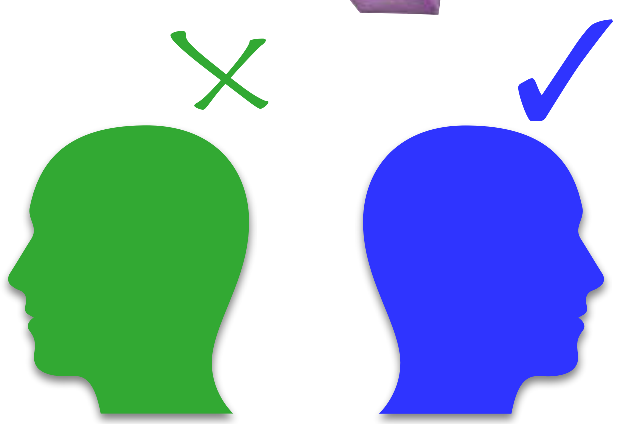


Building block origins

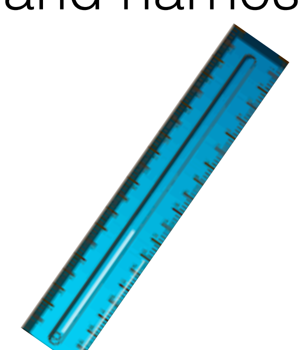


Low-SES
child-directed

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 $\mathrm{CP}_{\mathrm{that}}$ building blocks from it.

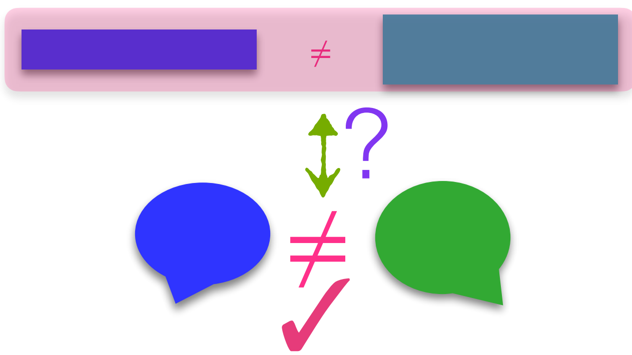


We can test this.

Bates & Pearl 2019, in prep.

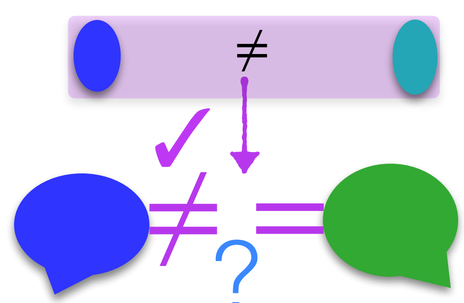
The big picture

One (standard) way



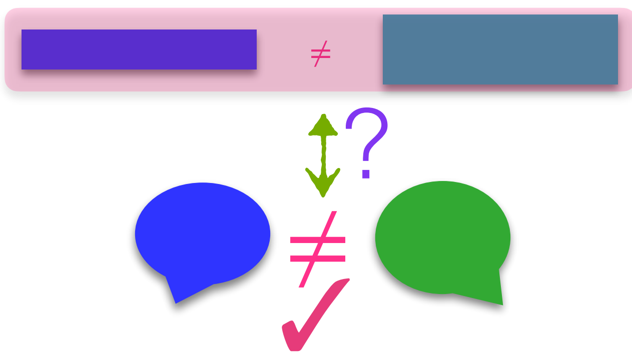
Developmental computational modeling complements existing techniques for assessing meaningful input differences.

A new (complementary) way



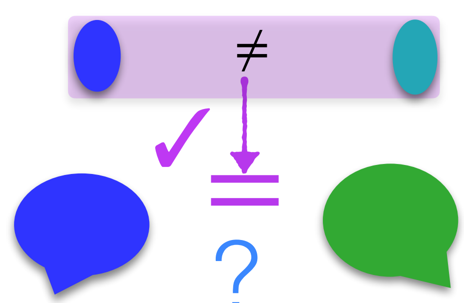
The big picture

One (standard) way



Developmental computational modeling complements existing techniques for assessing meaningful input differences.

A new (complementary) way

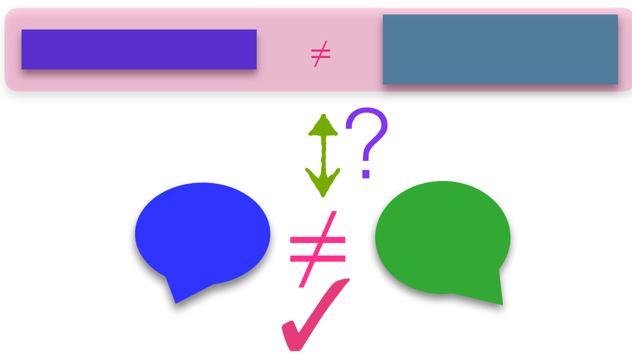


Who does

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

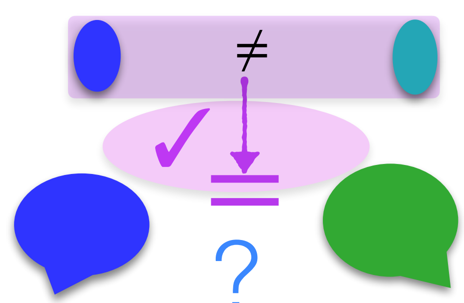
The big picture

One (standard) way



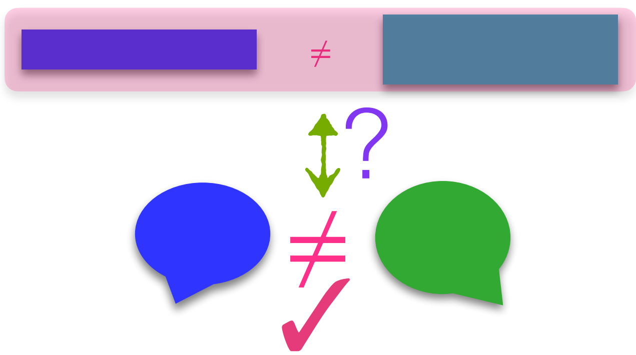
Something useful: This technique can provide a causal explanation for how input differences could affect learning outcomes.

A new (complementary) way



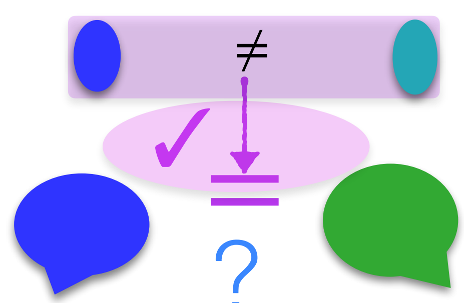
The big picture

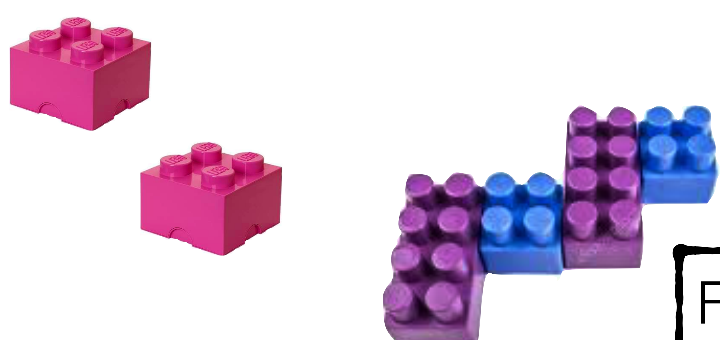
One (standard) way



Something useful: This technique can provide a causal explanation for how input differences could affect learning outcomes.

A new (complementary) way

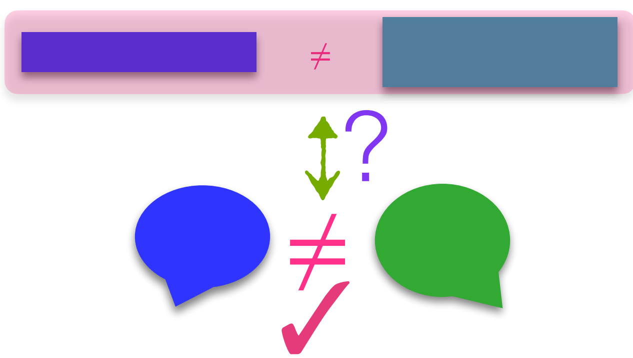




For syntactic islands, the building blocks needed for this knowledge don't seem to differ enough to matter.

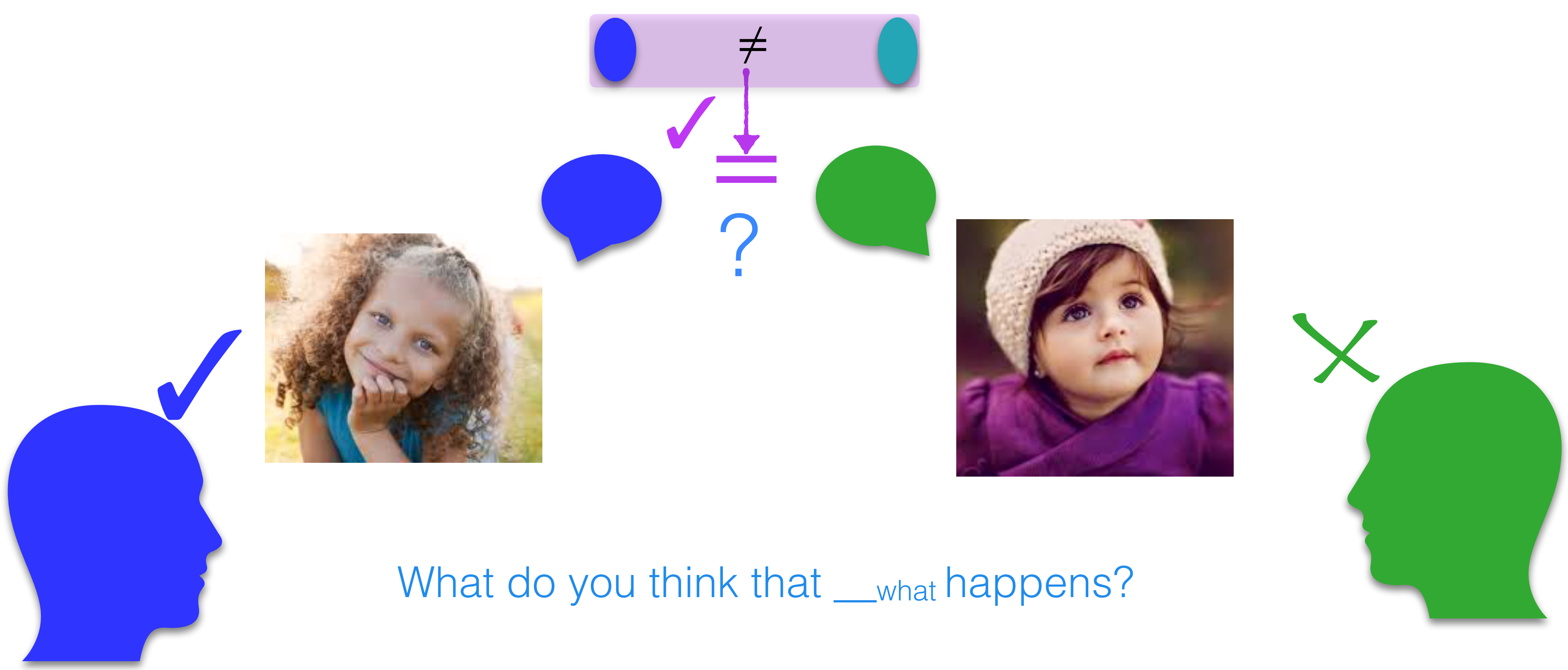
The big picture

One (standard) way



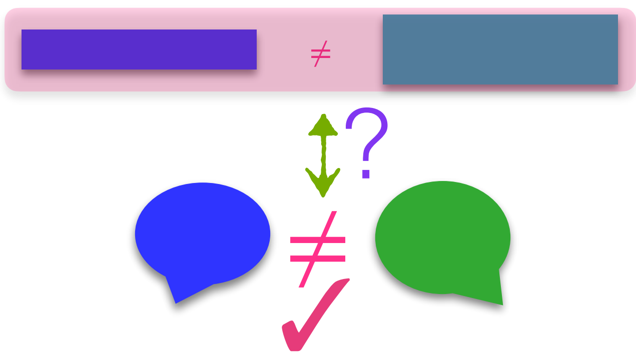
Something else useful: This technique can make predictions about differences we expect in both child outcomes and eventual adult knowledge.

A new (complementary) way



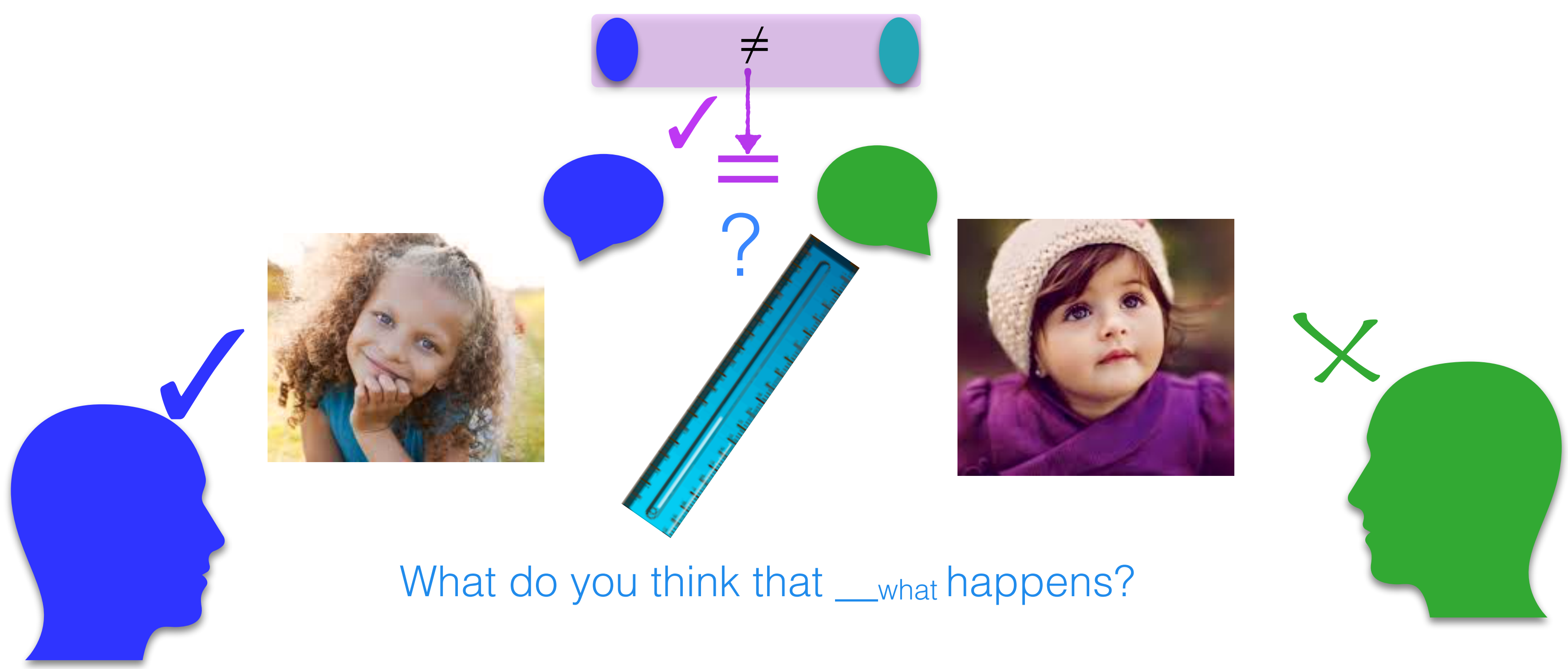
The big picture

One (standard) way



Something important: Any predicted differences still need to be measured. But at least we know what to look for.

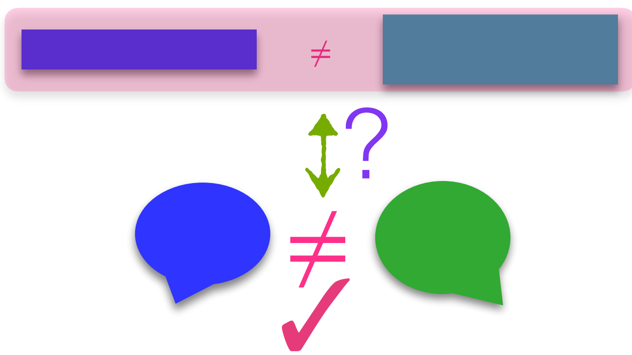
A new (complementary) way



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

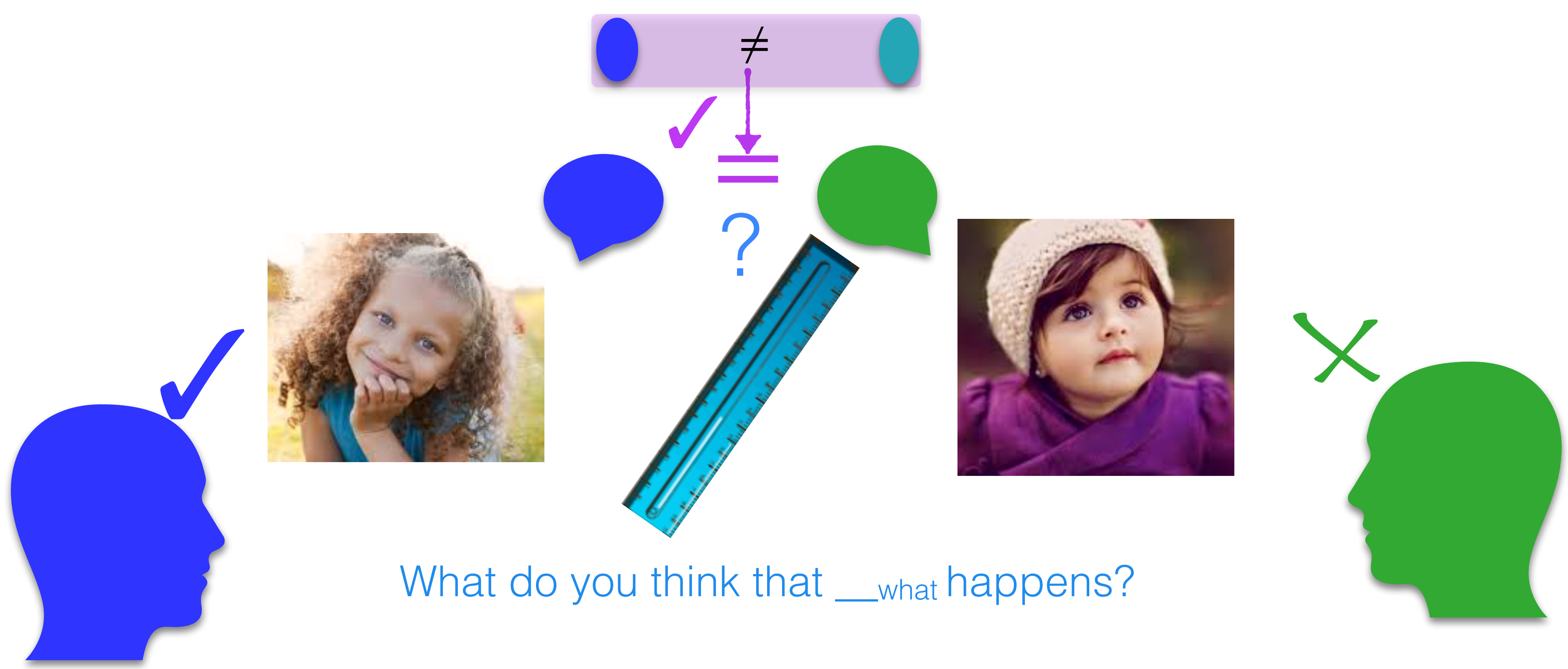
The big picture

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.

A new (complementary) way



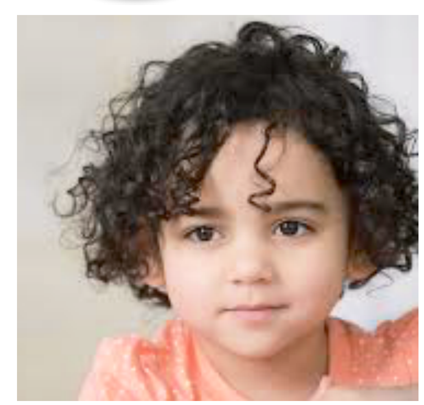




So let's use
developmental computational modeling
when we want to identify and understand
meaningful input variation!





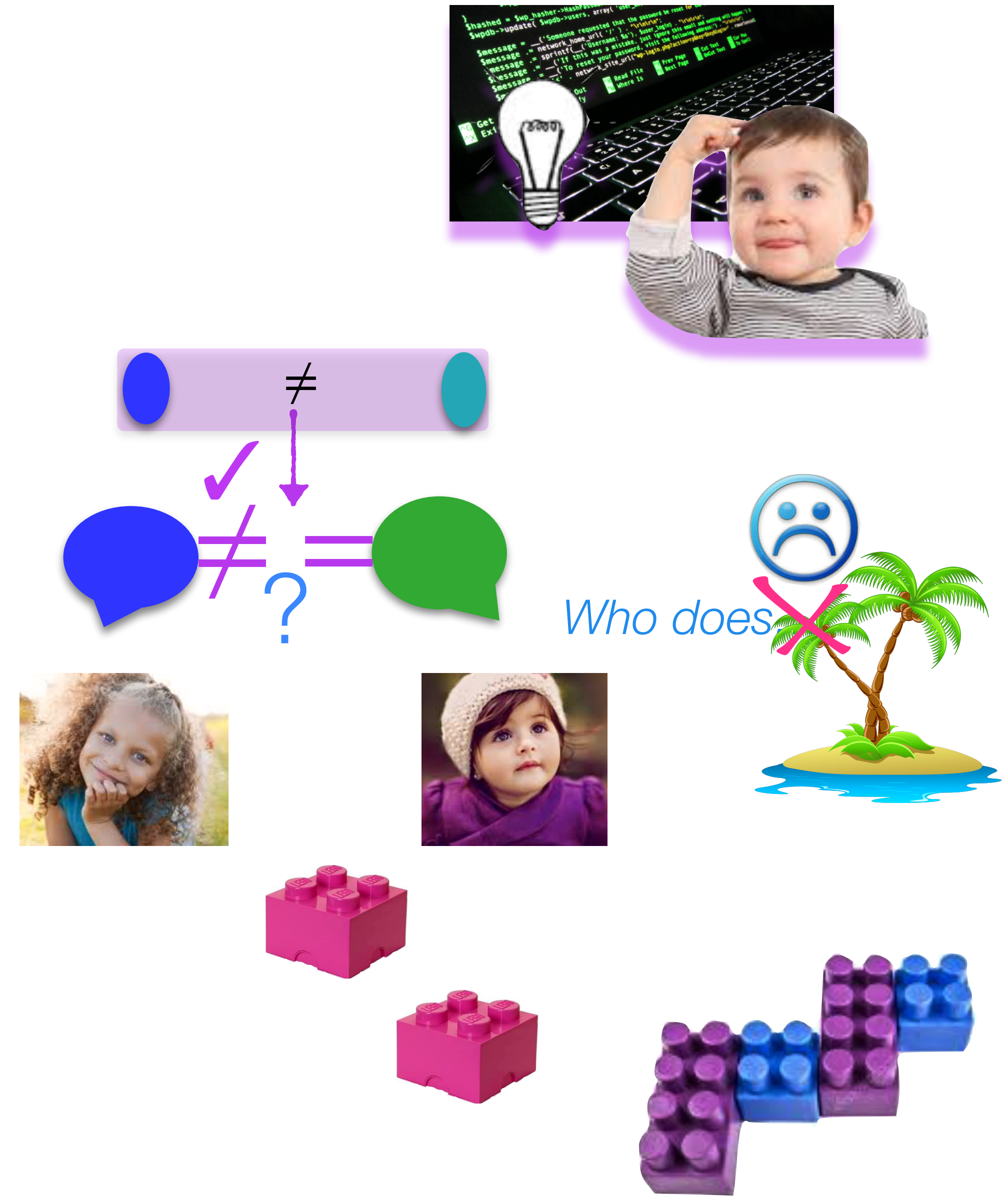

Thank you!

Alandi Bates

BUCLD 2018

UCI Institute for Mathematical Behavioral Sciences 2019

UCI QuantLang Collective



Computation of Language Laboratory

UC Irvine

Lisa S. Pearl
Professor
Department of Language Science
Department of Cognitive Sciences
SSPB 2219, SBSG 2314
University of California, Irvine
lpearl@uci.edu