# Introduction to computational models of language acquisition


Lisa Pearl

University of California, Irvine

Computation of Language Laboratory
UC Irvine

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


August 3, 2017:

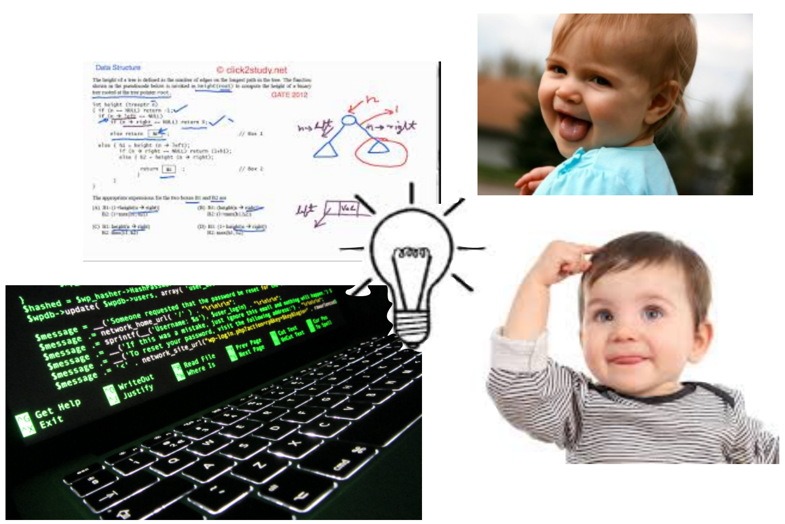
Norwegian Summer Institute on Language & Mind

University of Oslo

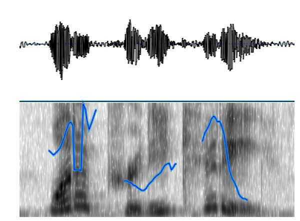

# Today's Plan:
# Computational models of language acquisition

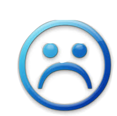
I. Why

II. How



III. What we can learn





Noun

✓KI tty

*Who does… is pretty?*

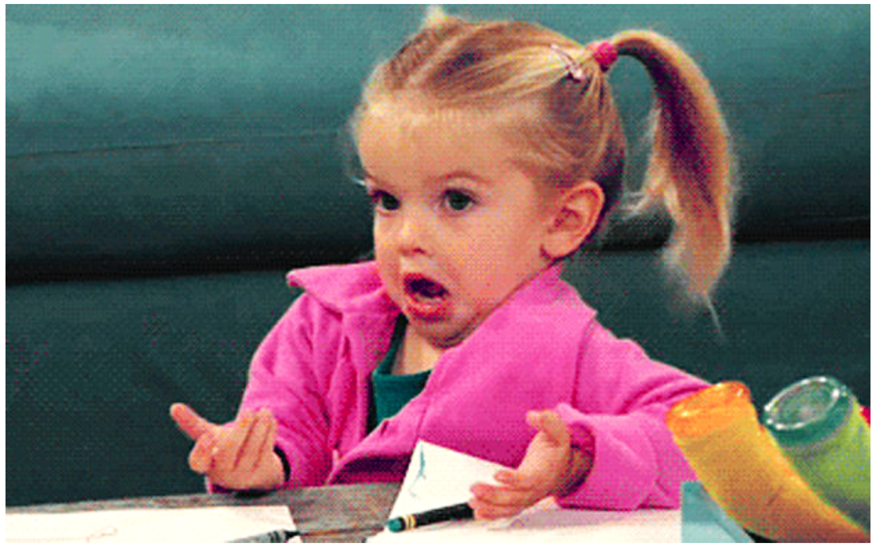
*another one*

*Every kitty didn't …*

# Why do you want to model language acquisition?

# Why do you want to model language acquisition?

*What does it mean to model something?*



# What does it mean to model something?

It's a scientific technique, like running an experiment. So saying "I want to model $thing" is just like saying "I want to run an experiment about $thing." Basically, it's a fine plan, but the important question is *why* you're doing it. That is, what question are you trying to answer?

Once you know **what question you're trying to answer**, you can design the right test of it — whether that's an experiment or a model or something else entirely.

# So what questions should we be using models for?

"…these questions tend to concern the process of acquisition that yields adult knowledge – that is, *how exactly acquisition proceeds, using particular learning strategies*." - Pearl 2017

## The importance of theory

"…an informative model of acquisition is the **embodiment of a specific theory about acquisition**."
- Pearl 2017

# The importance of theory

"…you need to first have a theory about how acquisition works. Then, the model can be used to

(1) make all the components of that acquisition theory explicit,

(2) evaluate whether it actually works, and

(3) determine precisely what makes it work (or not work)."

- Pearl 2017

# Making the components explicit

"It often turns out that the acquisition theories that seem explicit to humans don't actually specify all the details necessary to implement the strategies these theories describe."

- Pearl 2017

**Example: Learning linguistic parameter values from triggers in the input**

**Specific example:**

The trigger for *wh*-movement is seeing a *wh*-word in a position different from where it's understood (e.g., *what* in the question *What did the penguin do* $\_\__{what}$?)

# Making the components explicit

The trigger for *wh*-movement is seeing a *wh*-word in a position different from where it's understood (e.g., *what* in the question *What did the penguin do* $\_\_\__{what}$?)

**What do children need to know or be able to do in order to recognize the appropriate *wh*-movement trigger in their input?**

- *Know*: a certain word is one of these special *wh*-words
- *Do*: reliable segmentation of words in the utterance in order to recognize a *wh*-word not appearing where it's understood
- *Do*: remember the fronted *wh*-word in the utterance reliably enough to update the internal parameter value
- *Know*: ignore utterances where the *wh*-word doesn't move (e.g., echo questions like *The penguin did what?!*)

# Making the components explicit

The trigger for ***wh*-movement** is seeing a *wh*-word in a position different from where it's understood (e.g., *what* in the question *What did the penguin do* $\_\__{what}$?)

**Now, what about the *wh-in-situ* option (for languages like Mandarin Chinese and Japanese)?**

- Does this have a trigger too? What is it?
- If not, is *wh-in-situ* the default option that gets overridden by the presence of *wh*-movement triggers? If so, how many does it take?
- If there are no defaults but *wh-in-situ* also has no trigger, does the child use indirect negative evidence to decide her language is *wh-in-situ*? How much indirect negative evidence does it take?

# The importance of theory

"…you need to first have a theory about how acquisition works. Then, the model can be used to

(1) make all the components of that acquisition theory explicit,

(2) evaluate whether it actually works, and

(3) determine precisely what makes it work (or not work)."

- Pearl 2017

# Evaluating whether the theory works and determining what makes it work

"Once an acquisition theory is specified enough to implement in a computational model, we can then evaluate it by comparing the predictions it generates against the empirical data available from children."

- Pearl 2017

**Two basic outcomes:**

- the model predictions match children's data

- the model predictions don't match children's data

# Evaluating whether the theory works and determining what makes it work

The model predictions match children's data

This is an existence proof that the acquisition theory, as implemented in the model, is a way acquisition *could* proceed.

Note: Doesn't rule out alternative acquisition theories

**Two basic outcomes:**

- the model predictions don't match children's data

# Evaluating whether the theory works and determining what makes it work

The model predictions match children's data

This is an existence proof that the acquisition theory, as implemented in the model, is a way acquisition *could* proceed.

The model predictions don't match children's data

This is then evidence against that acquisition theory, *as implemented by the model*.

Remember: A model often specifies components of a theory that the original theory didn't. So, if this particular theory implementation doesn't work, maybe it's a problem with those components, and not the theory more broadly.

# Evaluating whether the theory works and determining what makes it work

The model predictions match children's data

This is an existence proof that the acquisition theory, as implemented in the model, is a way acquisition *could* proceed.

The model predictions don't match children's data

This is then evidence against that acquisition theory, *as implemented by the model*.

If you have an implemented model (whether it succeeds or fails), you can look inside it to determine what exactly makes it work or not work. This is something that's much more difficult to do with children's minds.

# Evaluating whether the theory works and determining what makes it work

*What did the penguin do* $\_\__{what}$?

Suppose we have a successful model of the acquisition of *wh*-movement from triggers.

# Evaluating whether the theory works and determining what makes it work

*What did the penguin do $\_\_\__{what}$?*

We can see if it's important for English children to ignore *wh*-echo questions where there's no *wh*-movement, or how necessary a Mandarin Chinese default *wh-in-situ* value is.

*The ~~penguin~~ did what?!*

*default = don't move*

# Evaluating whether the theory works and determining what makes it work

*What did the penguin do $\_\_\__{what}$?*

*The ~~penguin~~ did what?!*

*default = don't move*

This is **useful**!

If the model's predictions don't match children's behavior without these, we can say they're necessary components of the learning strategy this theory describes and we can explain *why* (e.g., they filter the input or help the child navigate the hypothesis space).

# Modeling as a useful tool

Modeling can be used as a tool for both developing and refining acquisition theories.

Notably, an acquisition theory actually includes two types of theories:

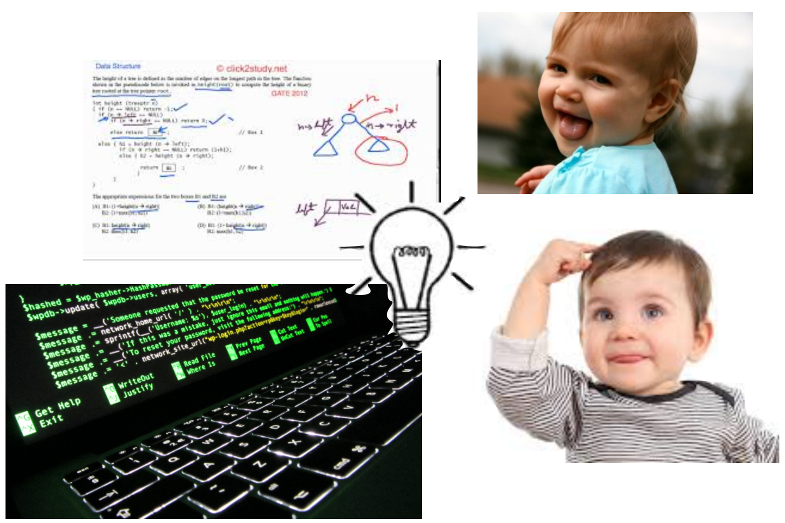
- theories of the learning process
- theories of the representations to be learned

An informative model incorporates both.

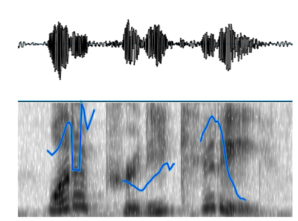
# Today's Plan:
# Computational models of language acquisition

I. Why

II. How



III. What we can learn



Noun

KI tty

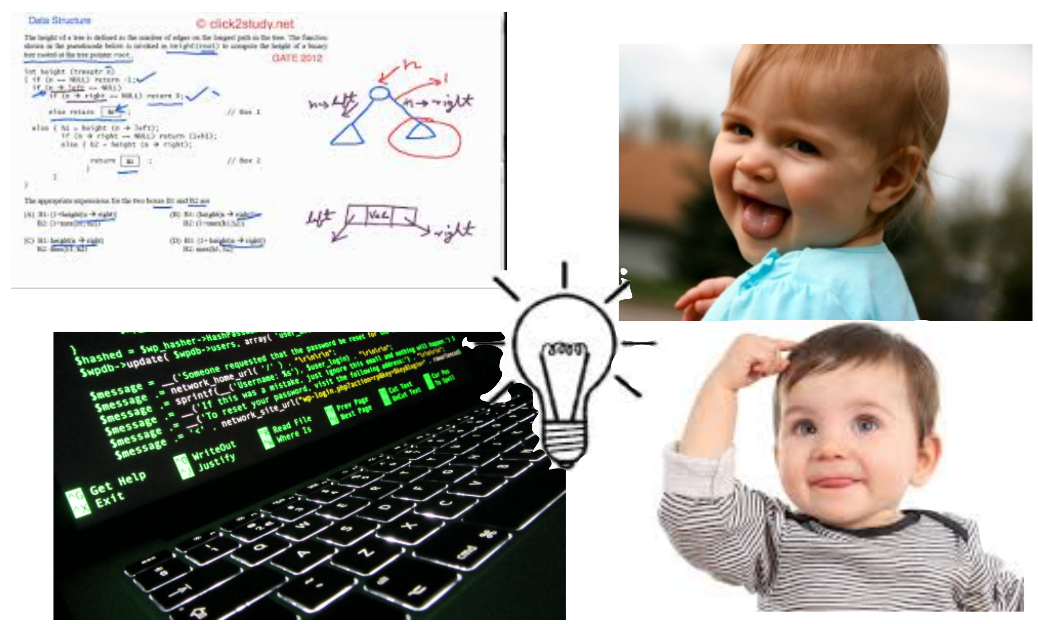
*Who does… is pretty?*

*another one*

*Every kitty didn't …*

# Today's Plan:
# Computational models of language acquisition

## II. How



# Language acquisition = An information processing task

# Language acquisition = An information processing task

Given the available input,

*Look at that kitty!*
*There's another one.*

Input

*Where did he hide?*
*What happened?*

## Language acquisition = An information processing task

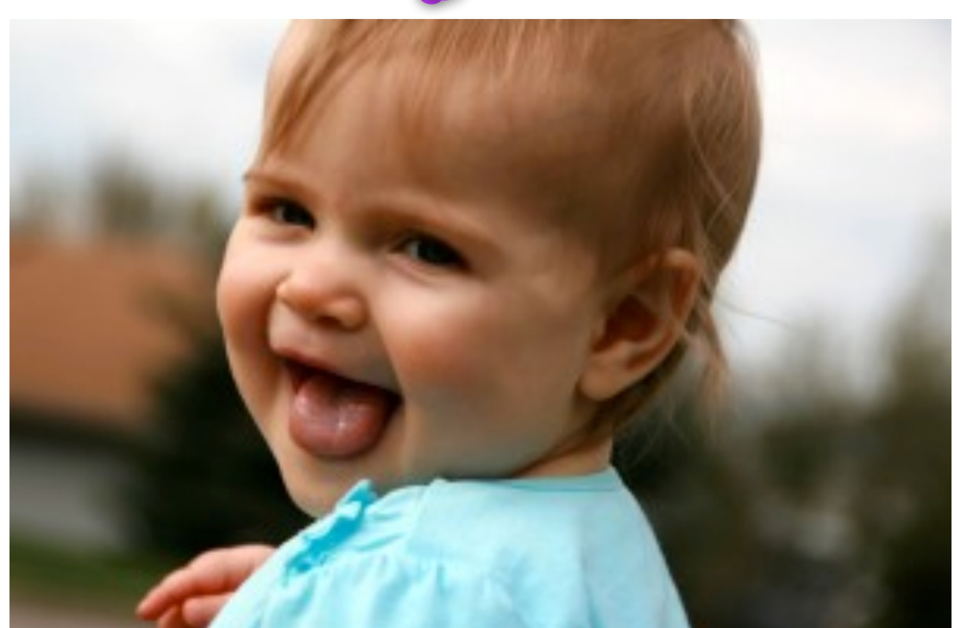
Given the available input, information processing done by human minds

processing & generalization

*Look at that kitty!*
*There's another one.*

Input

*Where did he hide?*
*What happened?*



# Language acquisition = An information processing task

Given the available input, information processing done by human minds to build a system of linguistic knowledge

words & morphemes

metrical phonology

syntactic categories

syntax

semantics

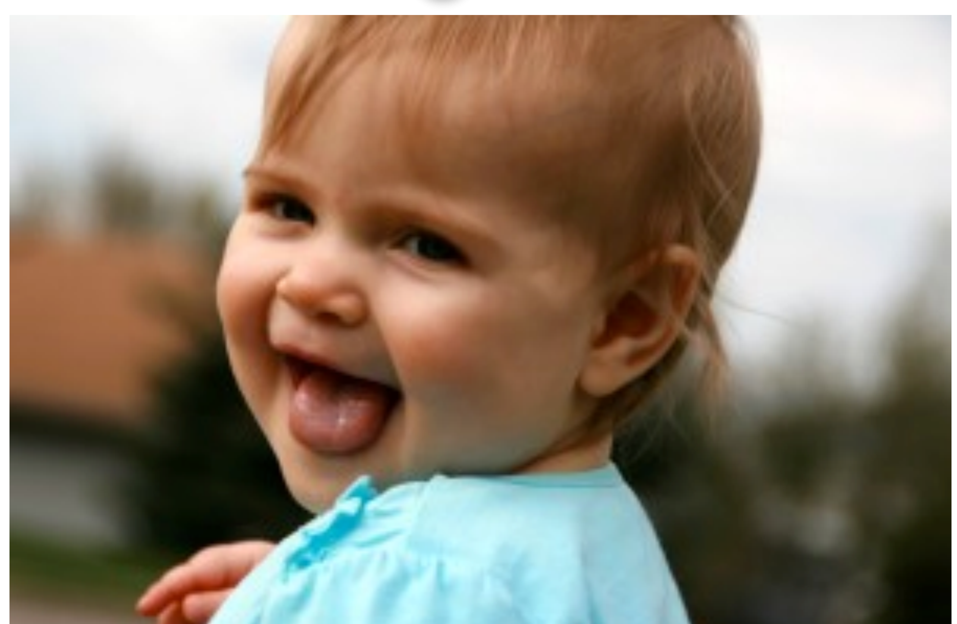
pragmatics

processing & generalization

*Look at that kitty!*
*There's another one.*

Input

*Where did he hide?*
*What happened?*



# Language acquisition = An information processing task

Given the available input, information processing done by human minds to build a system of linguistic knowledge whose output we observe

words & morphemes

metrical phonology

syntactic categories

syntax

semantics

pragmatics

processing & generalization

*Look at that kitty!*
*There's another one.*

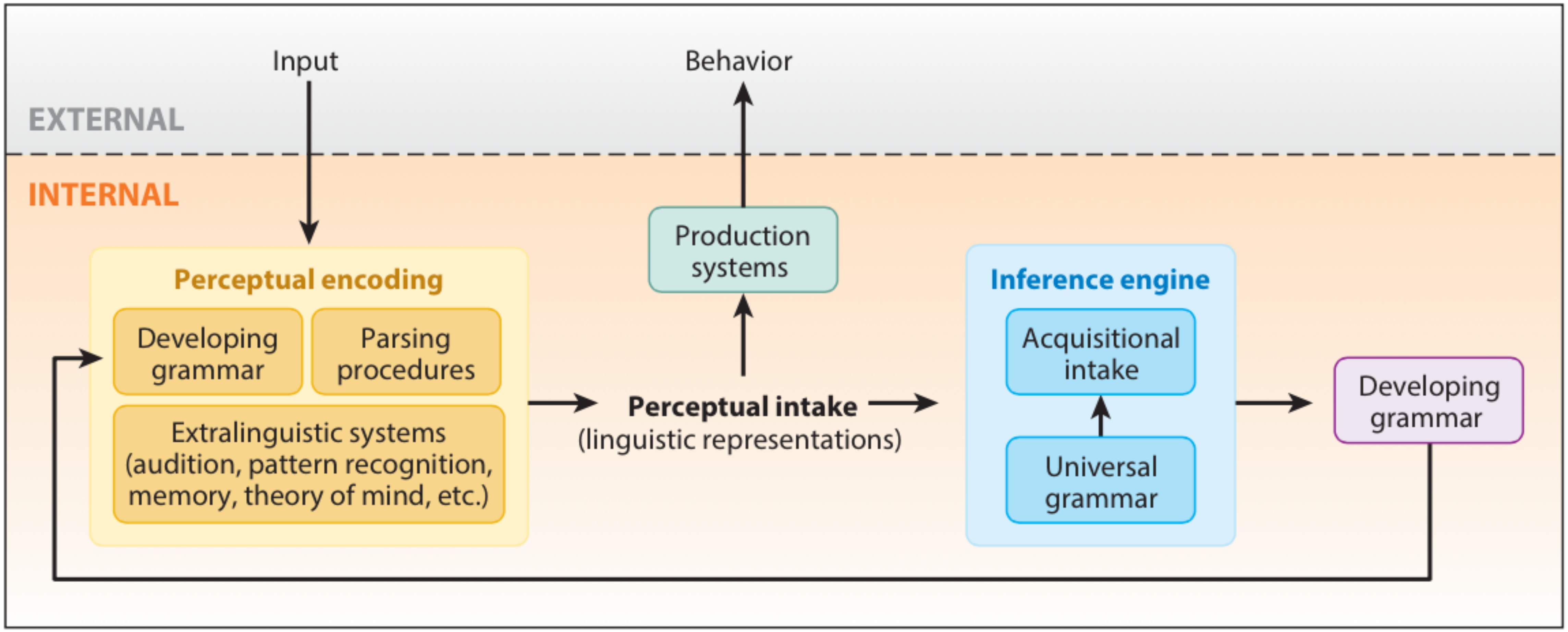
Input

*Where did he hide?*
*What happened?*

*Where's the kitty?*

*That one's really cute.*

# Language acquisition = An information processing task



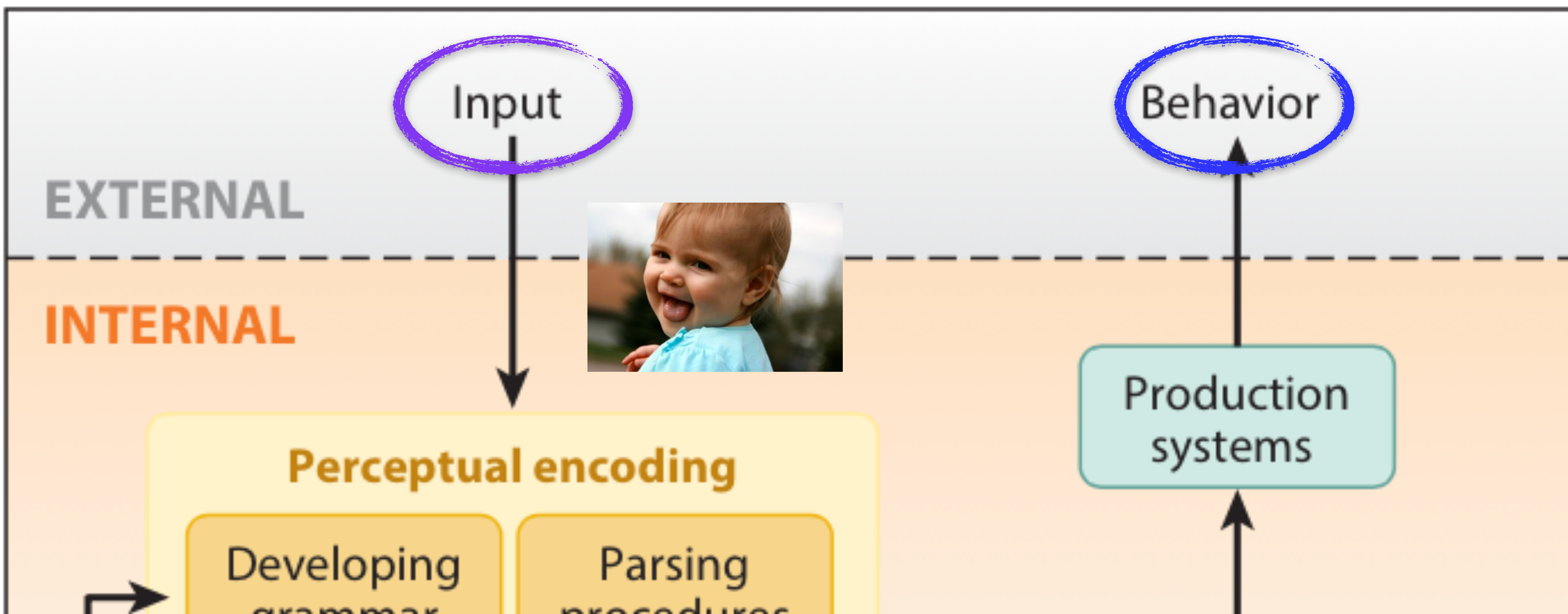

Lidz & Gagliardi 2015

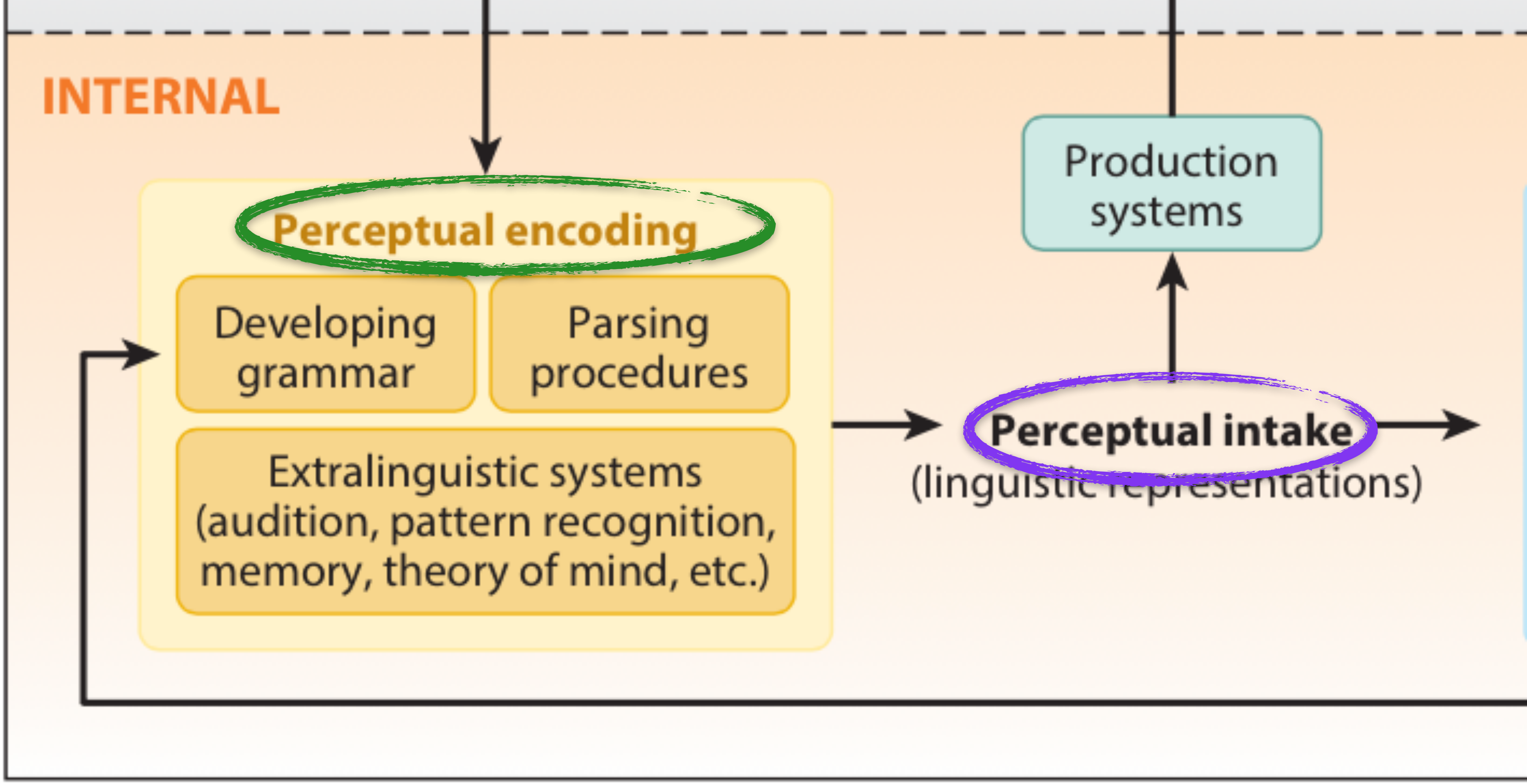
**A framework that makes components of the acquisition task more explicit.**

**A framework that makes components of the acquisition task more explicit.**

Distinguishes between things external to the child that we can observe (input signal, child's behavior) vs. things internal to the child (everything else).

**Corpus methods** **Experimental methods**

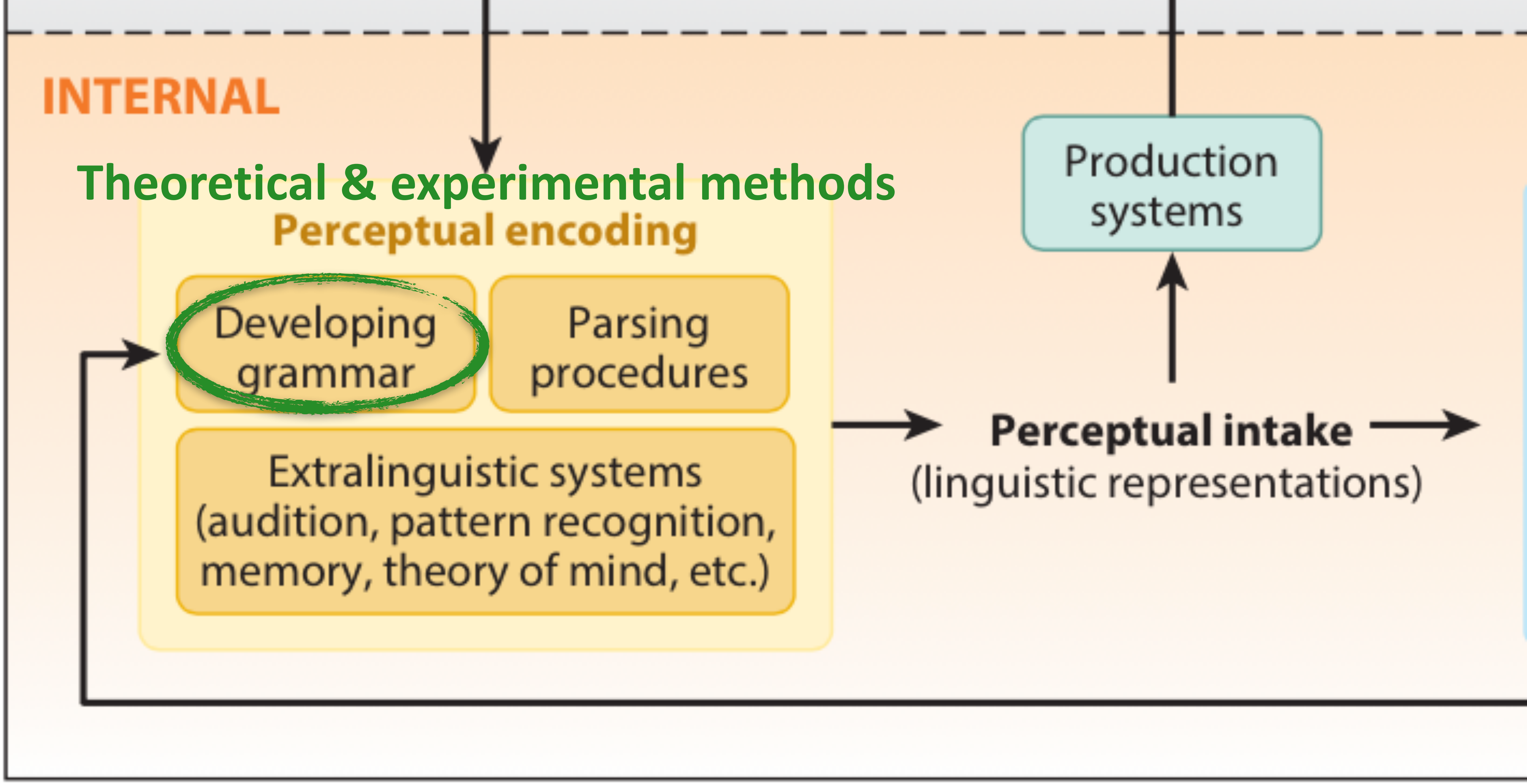




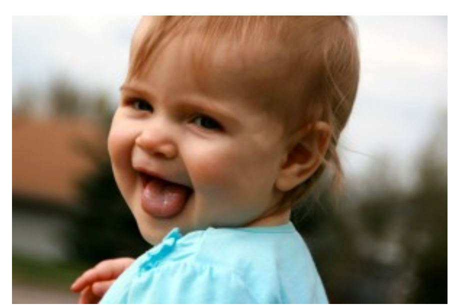

Lidz & Gagliardi 2015

**Perceptual encoding:**

Turning the input signal into an internal linguistic representation = perceptual intake.

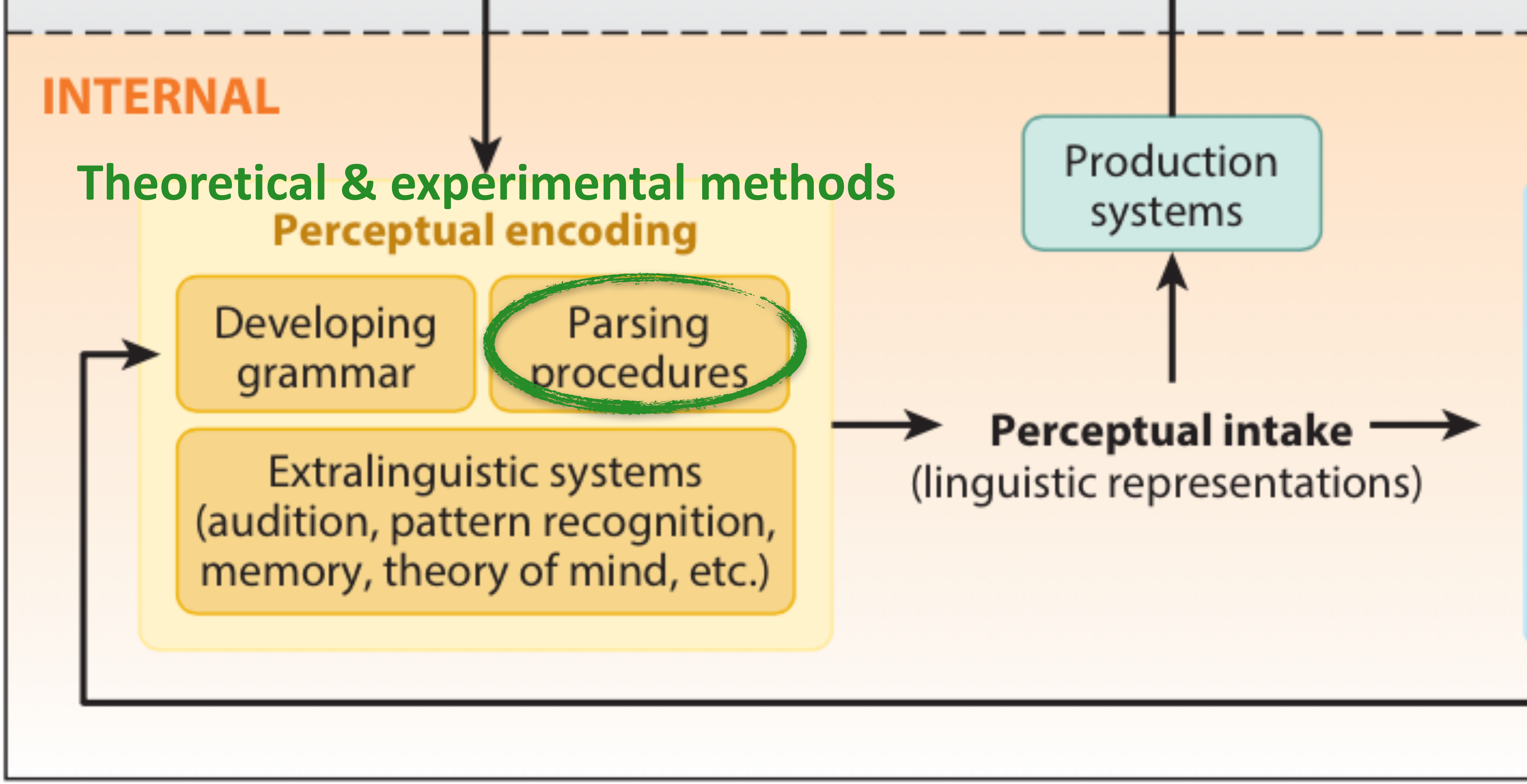



Lidz & Gagliardi 2015



**Perceptual encoding**:

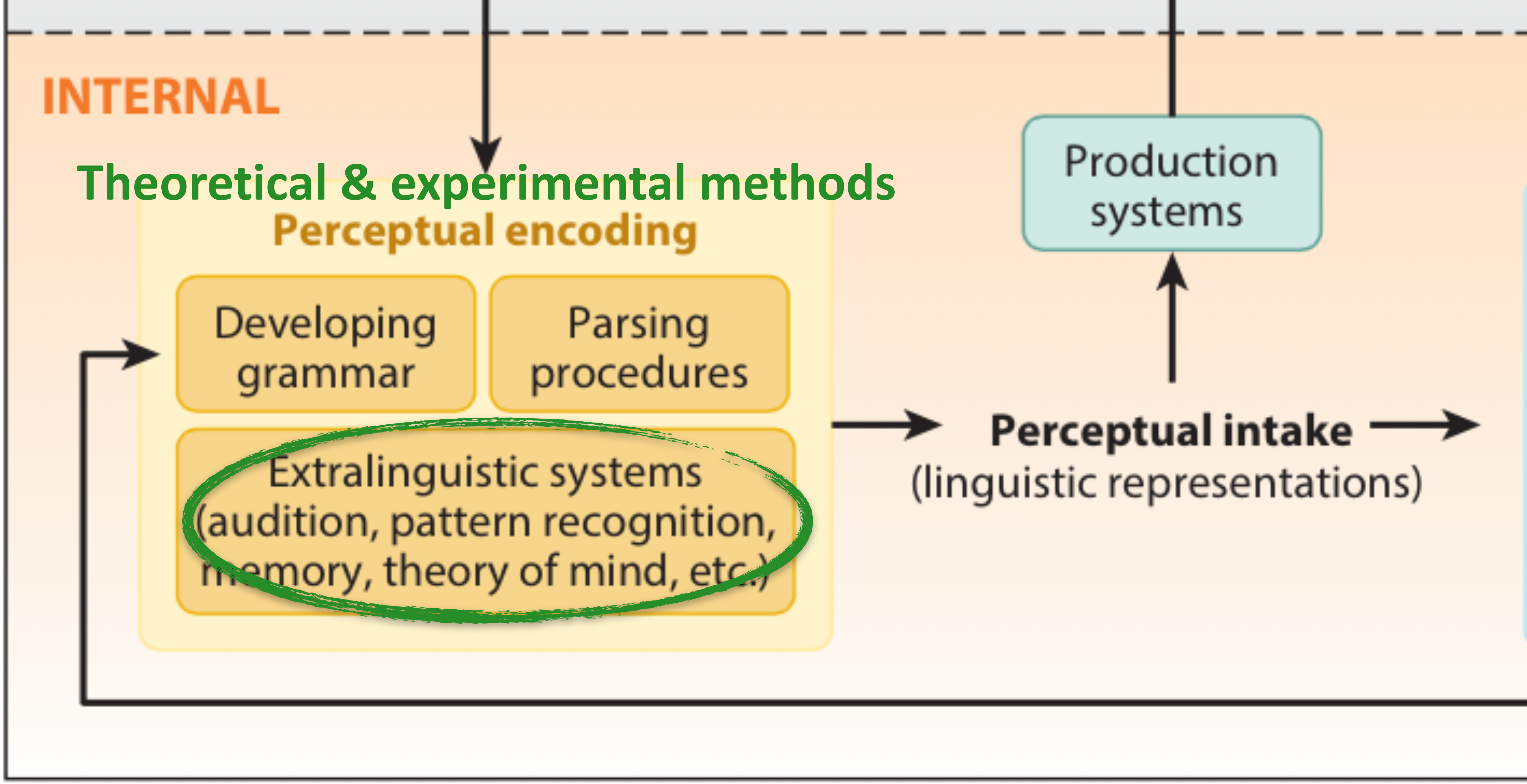
Involves current grammar



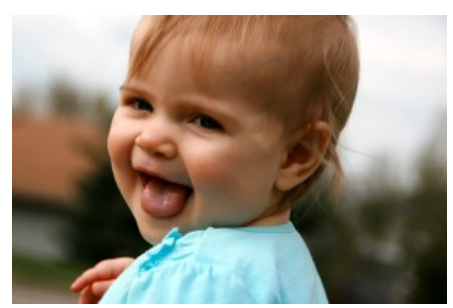

Lidz & Gagliardi 2015

**Perceptual encoding**:

Involves current grammar being deployed in real time to parse the input

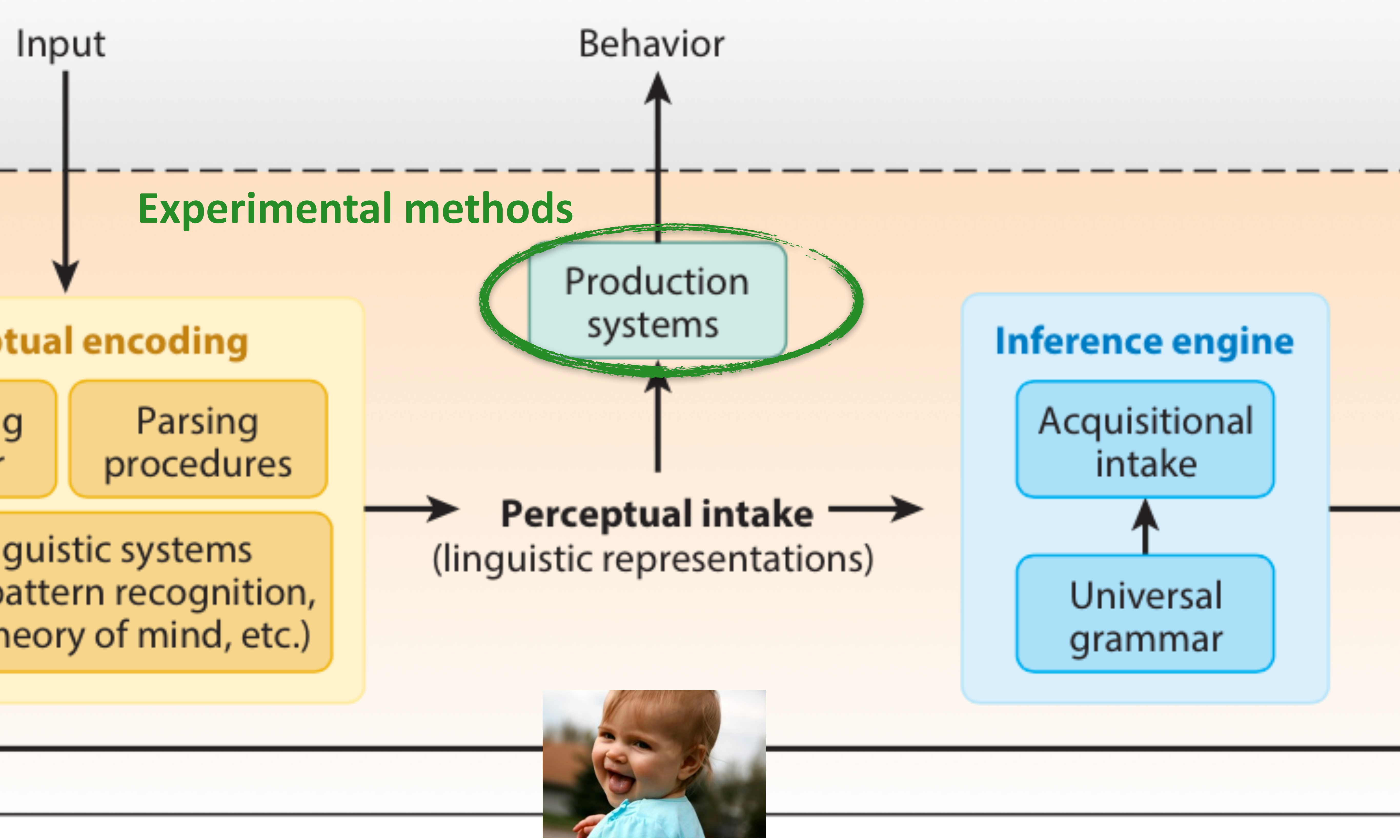



Lidz & Gagliardi 2015



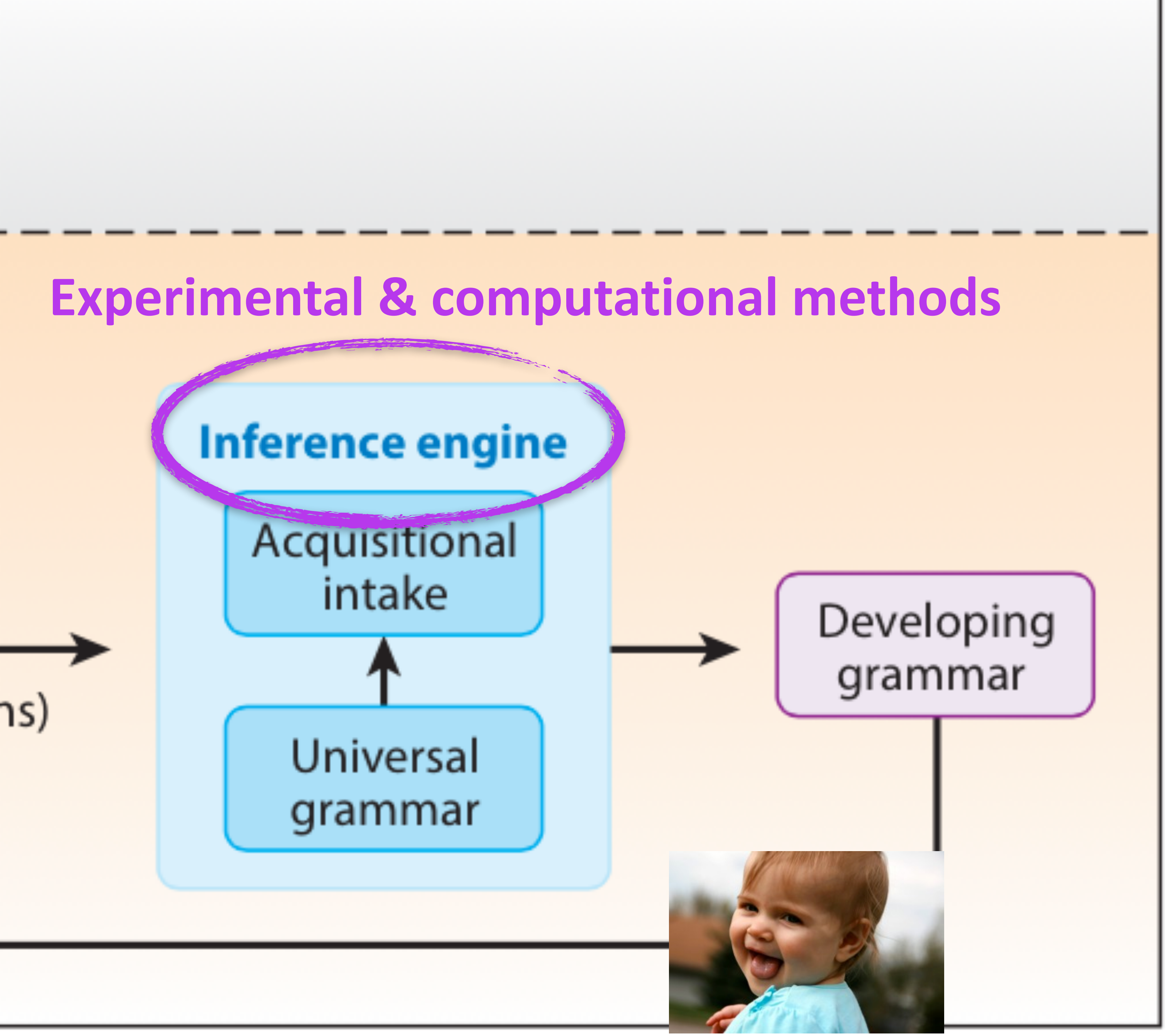
**Perceptual encoding**:

Involves current grammar being deployed in real time to parse the input often drawing on extralinguistic systems



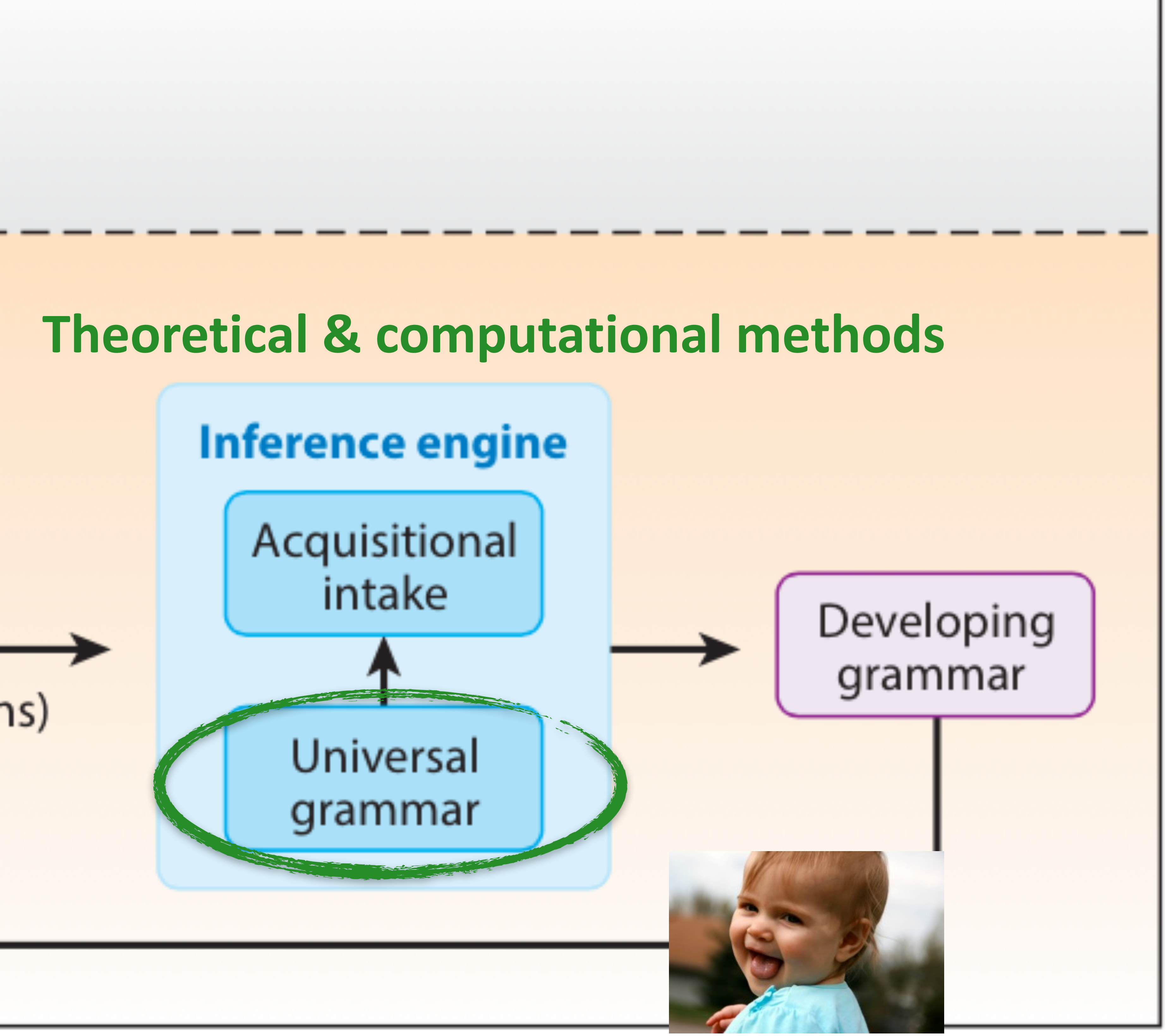

**Generating observable behavior**

Involves current linguistic representations being used by production systems.



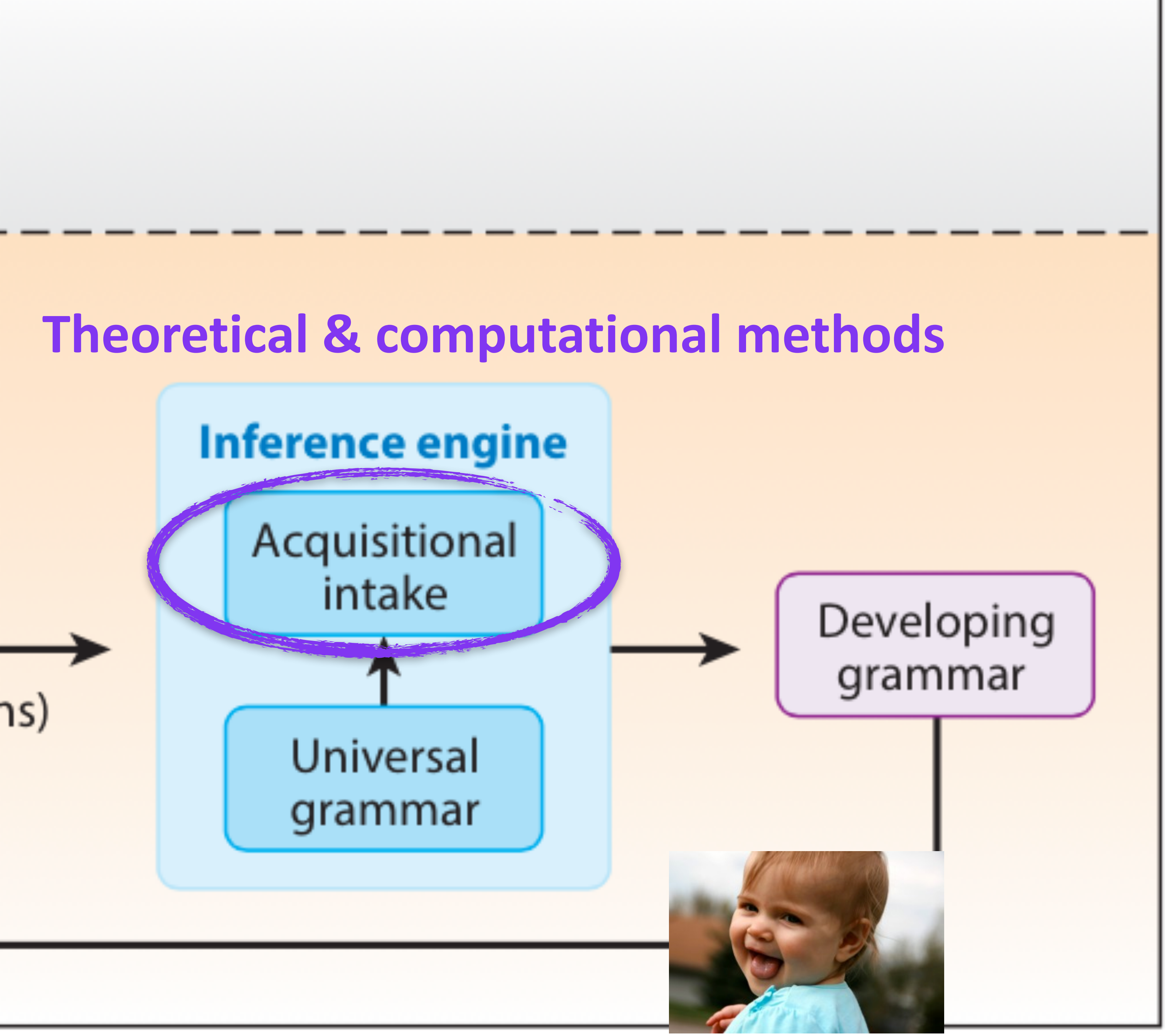

**Doing inference**

Generalization happens



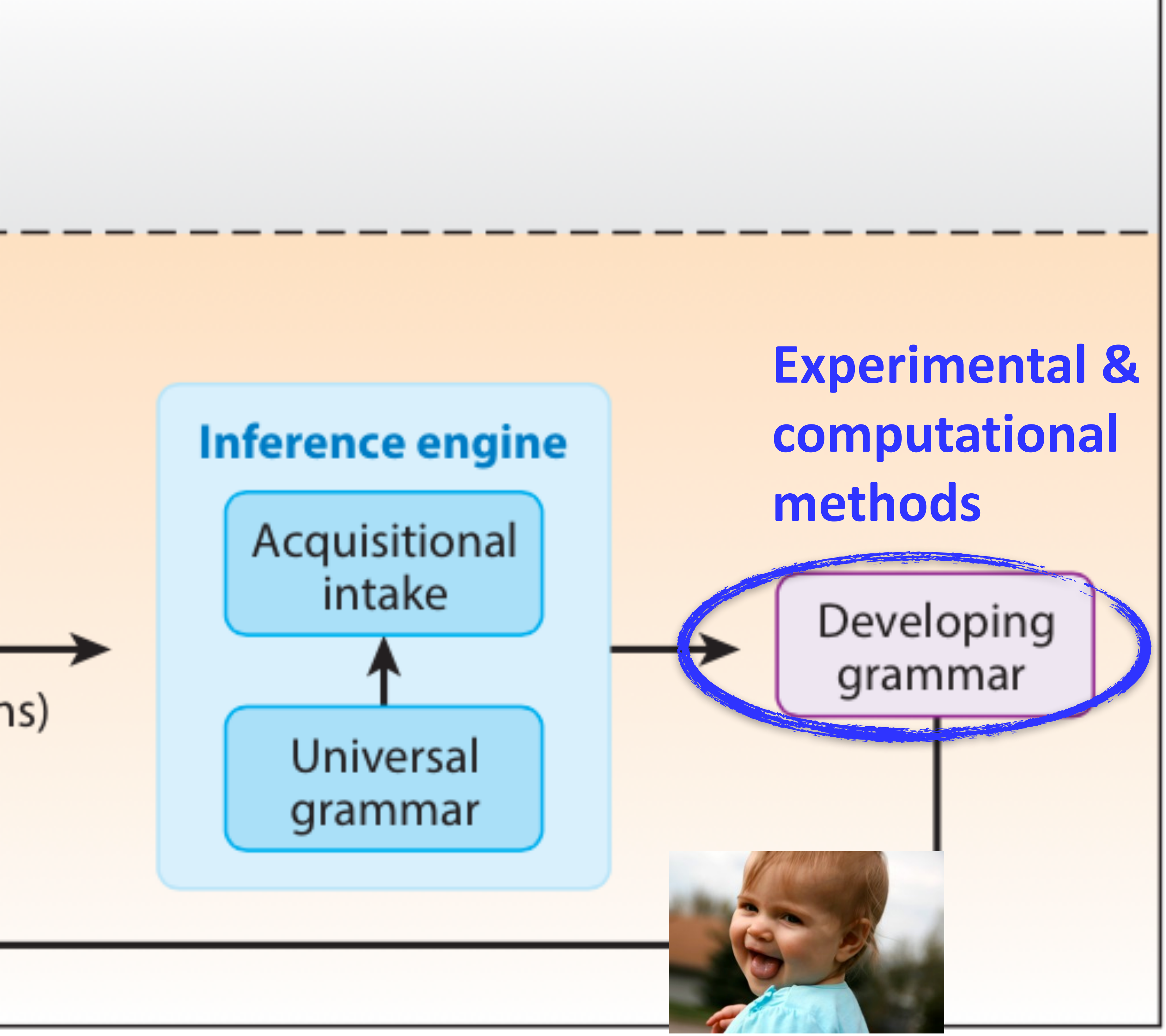

**Doing inference**

Generalization happens by using existing learning biases, (some of which may be innate and language-specific)



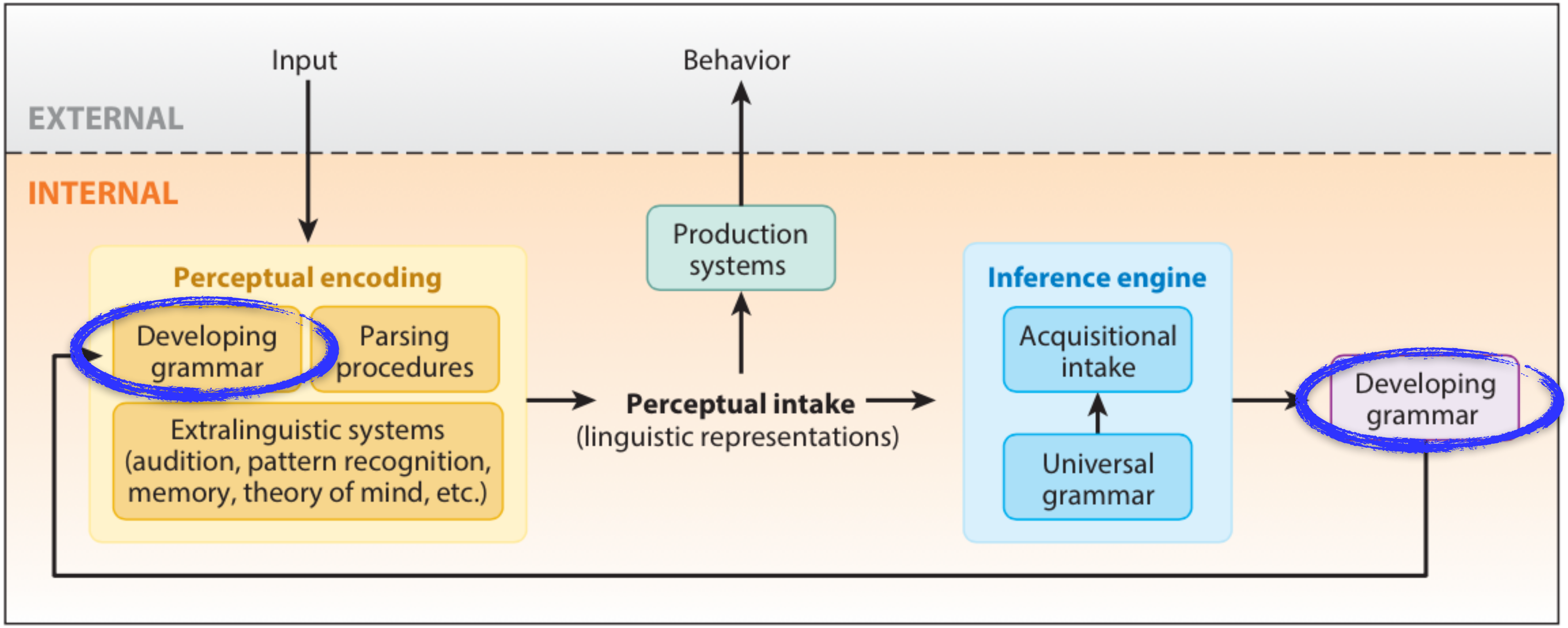

## Doing inference

Generalization happens by using existing learning biases, (some of which may be innate and language-specific) operating over the acquisitional intake — what's perceived as relevant for acquisition



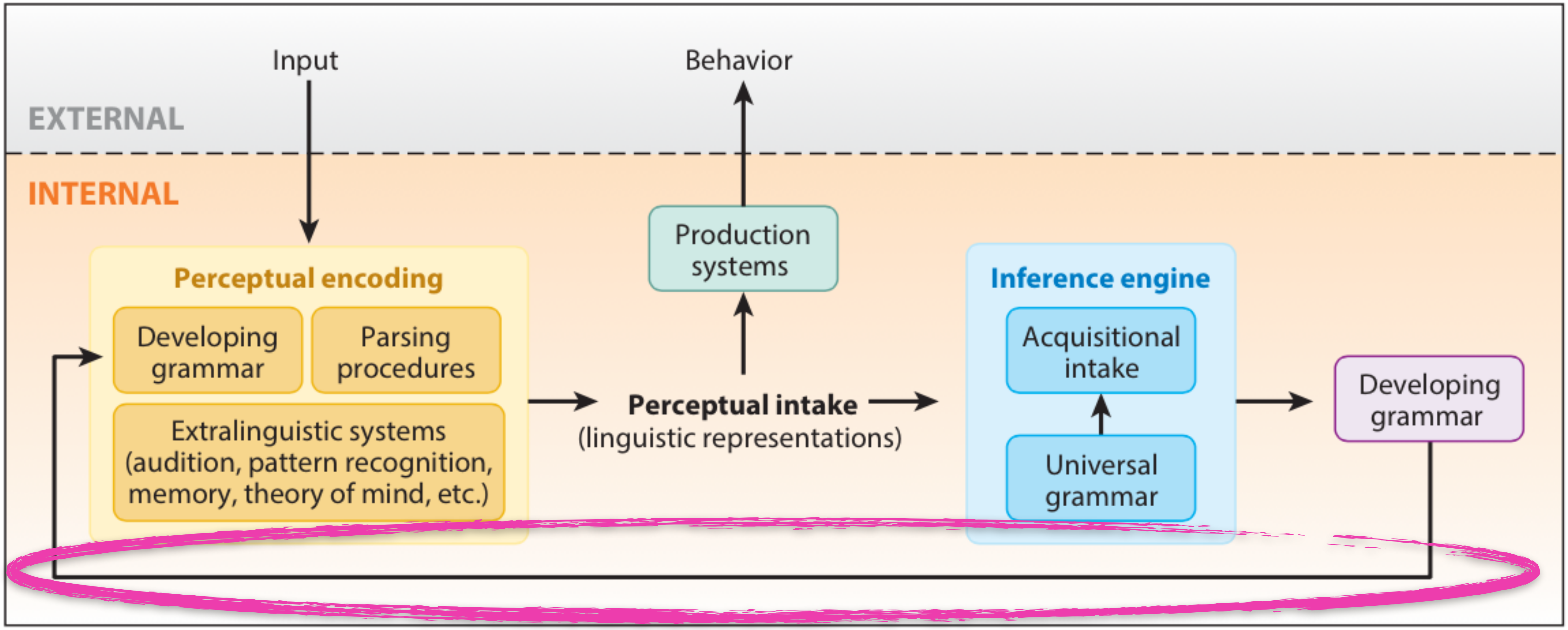

## Doing inference

Generalization happens by using existing learning biases, (some of which may be innate and language-specific) operating over the acquisitional intake — what's perceived as relevant for acquisition to produce the most up-to-date hypotheses about linguistic knowledge

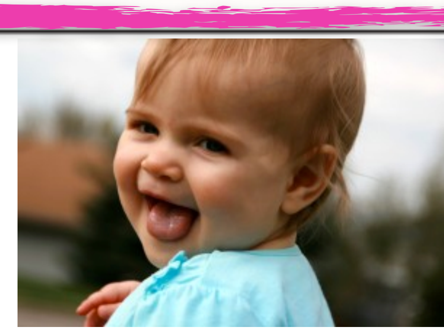



Lidz & Gagliardi 2015

**The current linguistic hypotheses are**
**used in subsequent perceptual encoding**



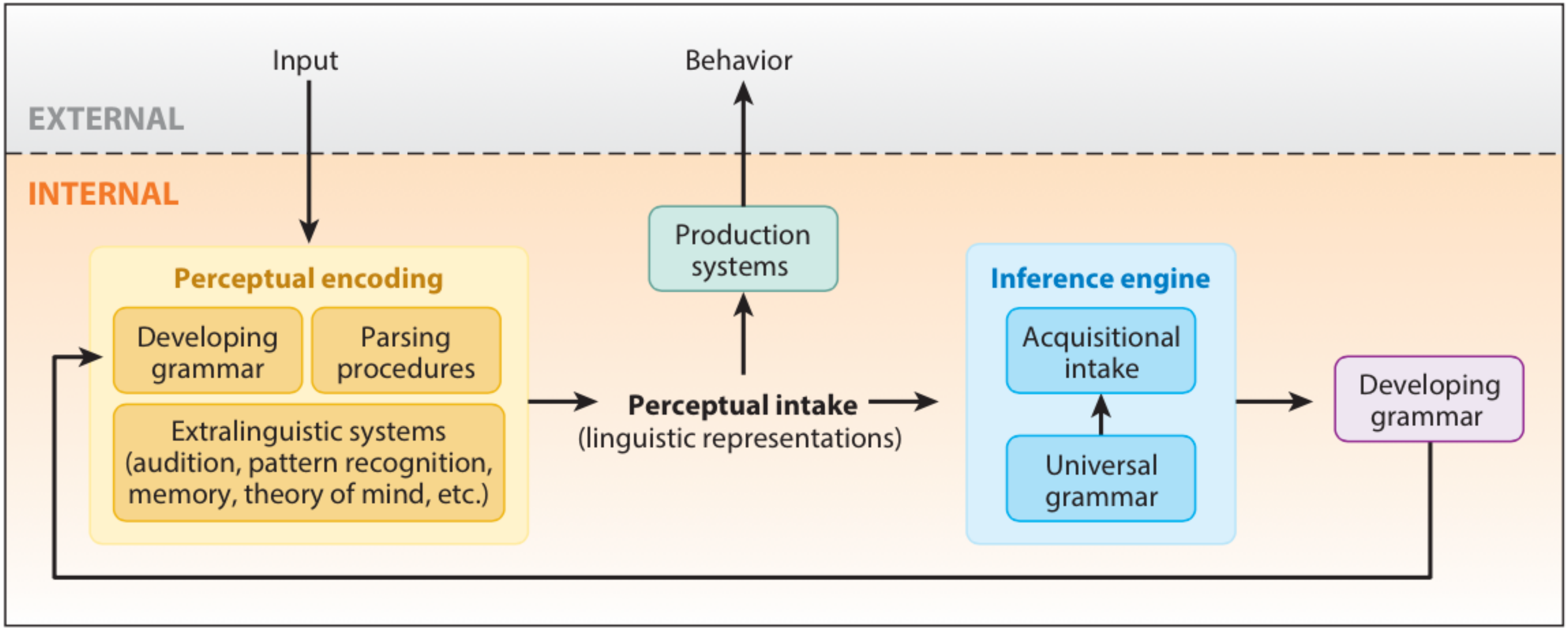

Lidz & Gagliardi 2015



Experimental methods

This whole process **happens over and over again** throughout the learning period

# This is language acquisition



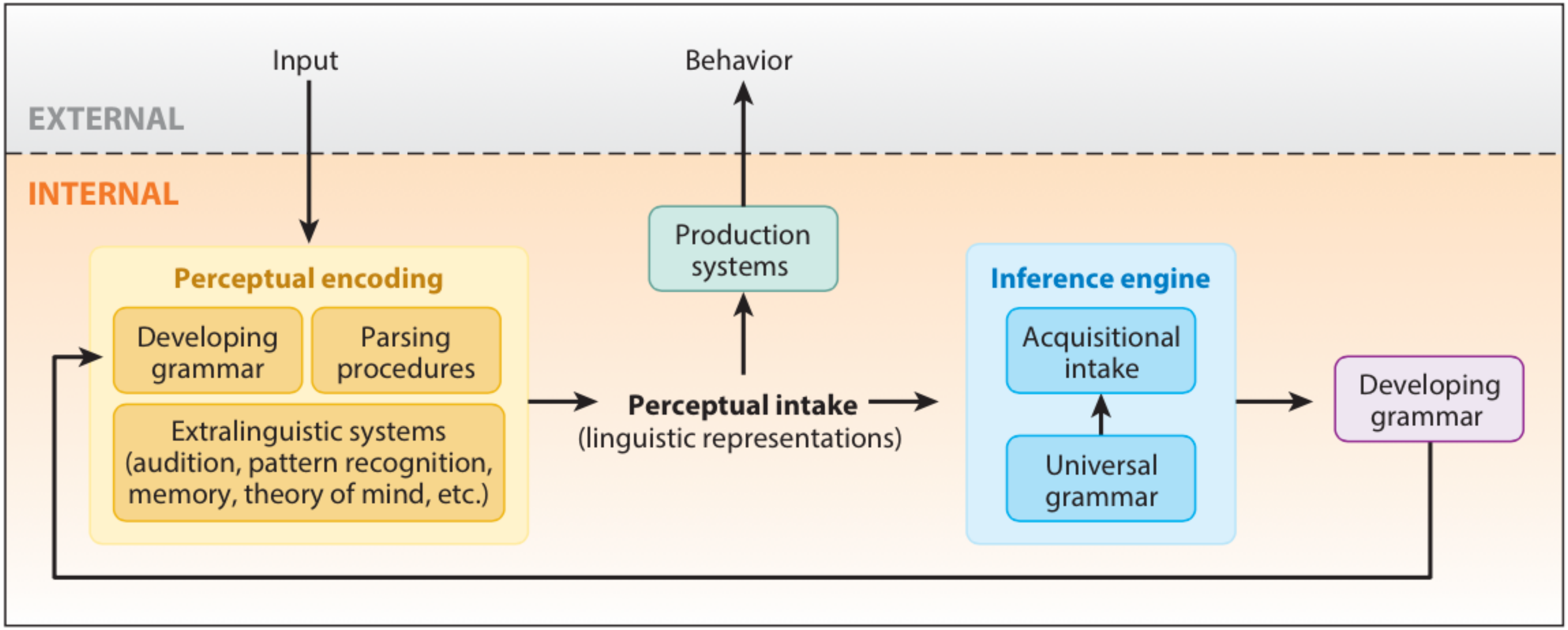

Lidz & Gagliardi 2015

Corpus Experimental

Theoretical Computational

An informative computational model of language acquisition captures these important pieces in an empirically-grounded way.

# This is language acquisition

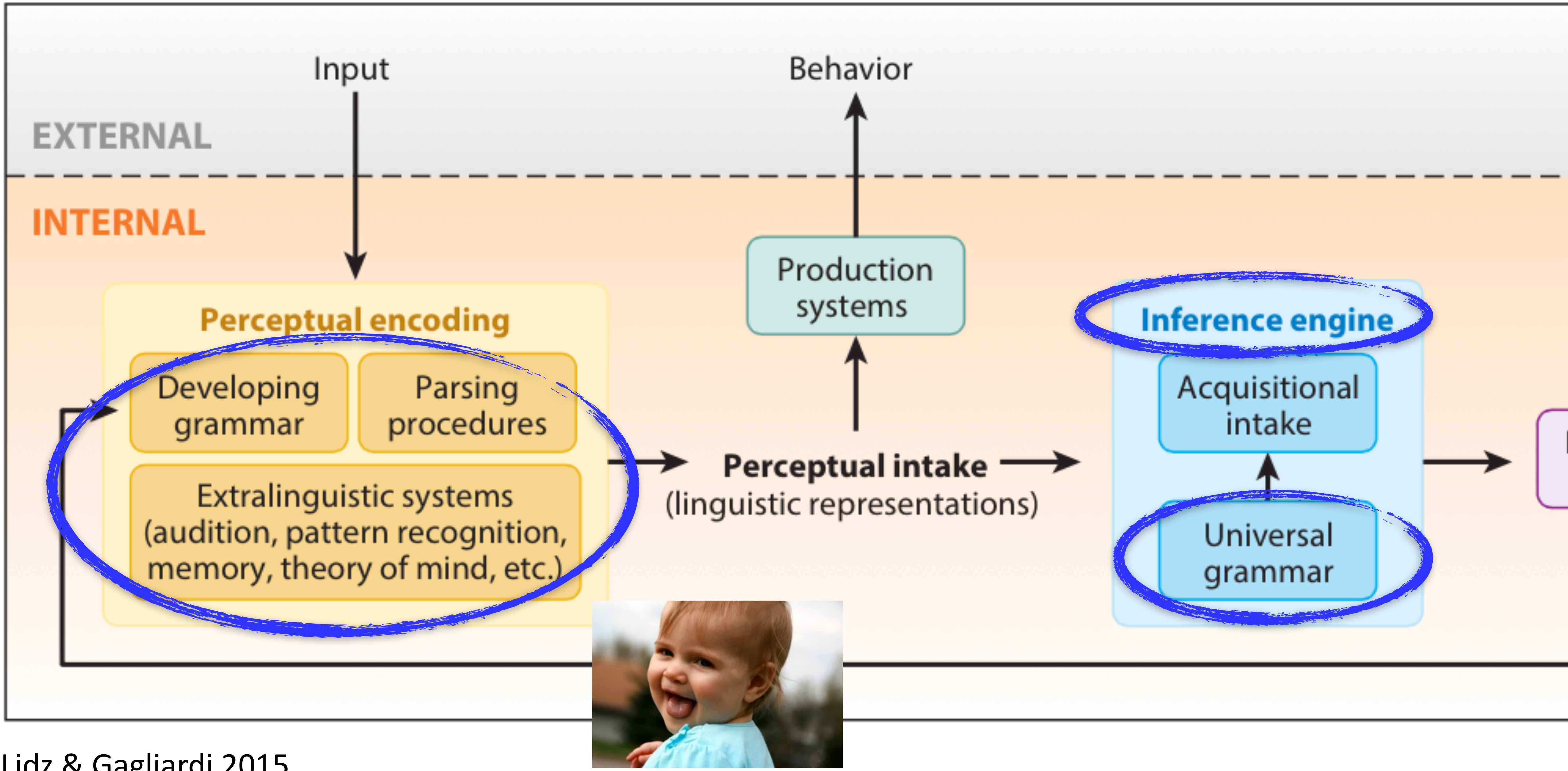



Lidz & Gagliardi 2015

Informative computational models = informative about the learning strategies children use

# Learning strategies children use

A successful learning strategy is an existence proof that linguistic knowledge is attainable using the knowledge, learning biases, and capabilities comprising that strategy.

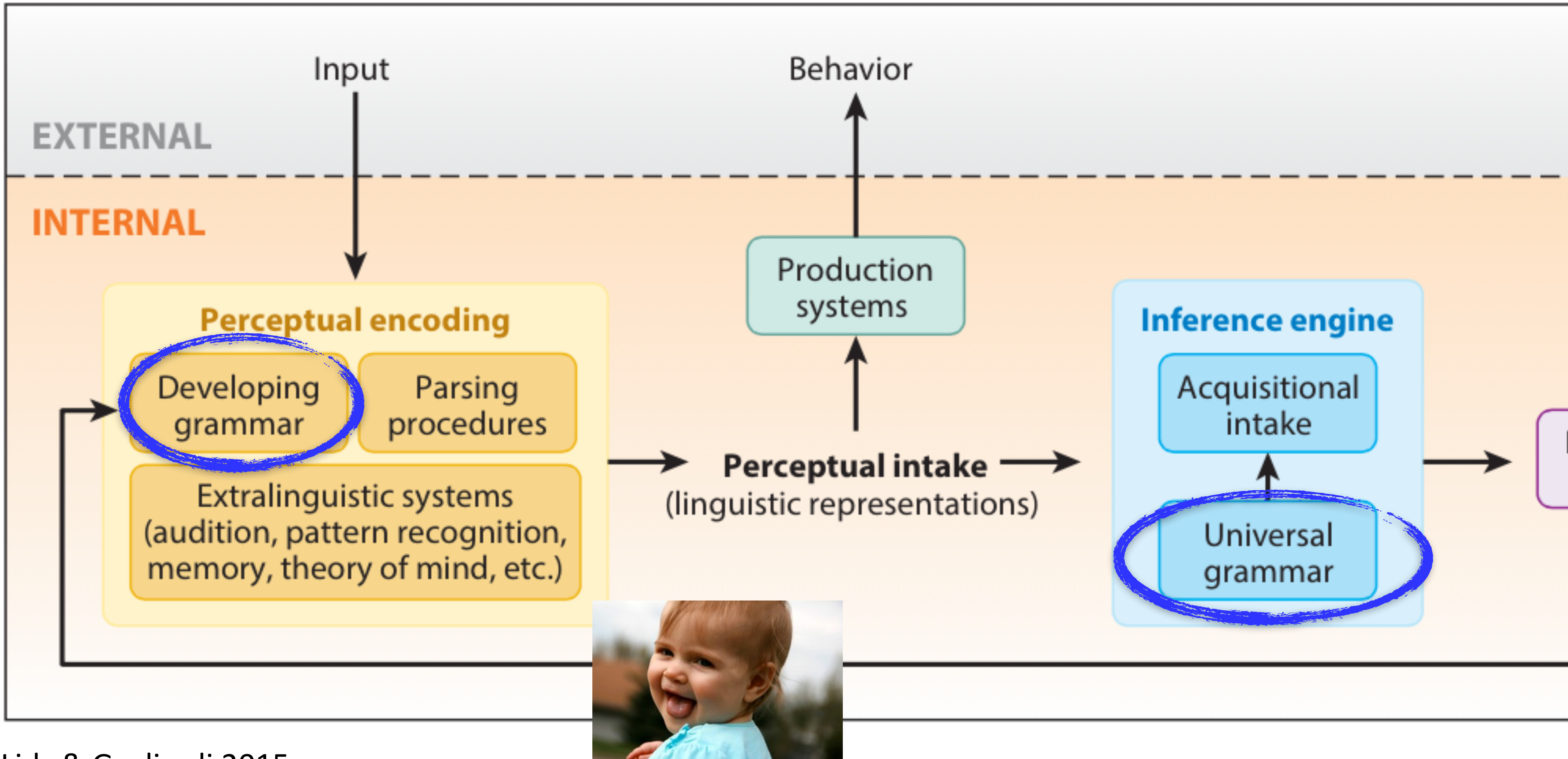



Lidz & Gagliardi 2015

# Learning strategies children use

Important learning strategy components include

- knowledge (= theories of **representation**)

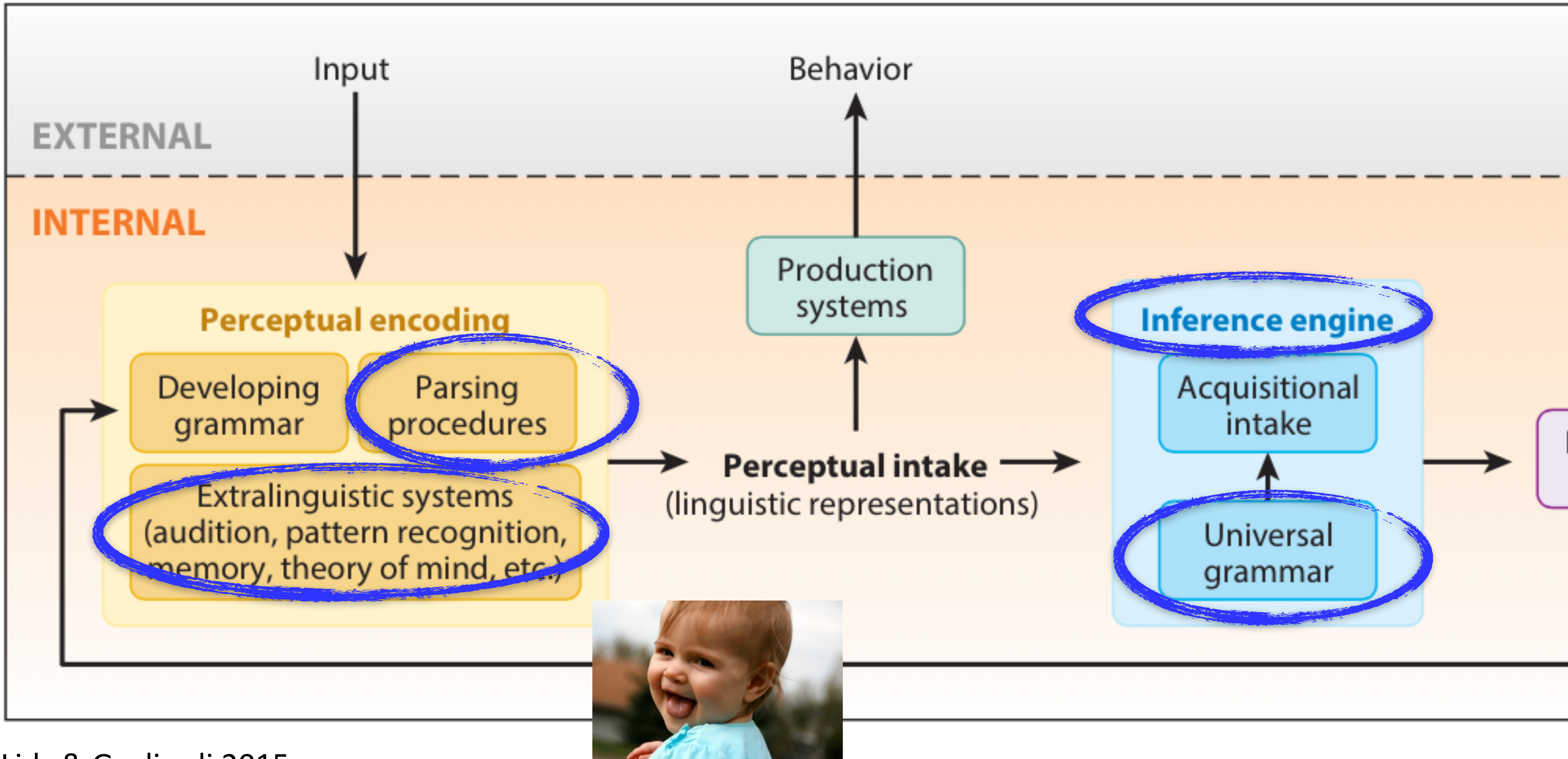



Lidz & Gagliardi 2015

# Learning strategies children use

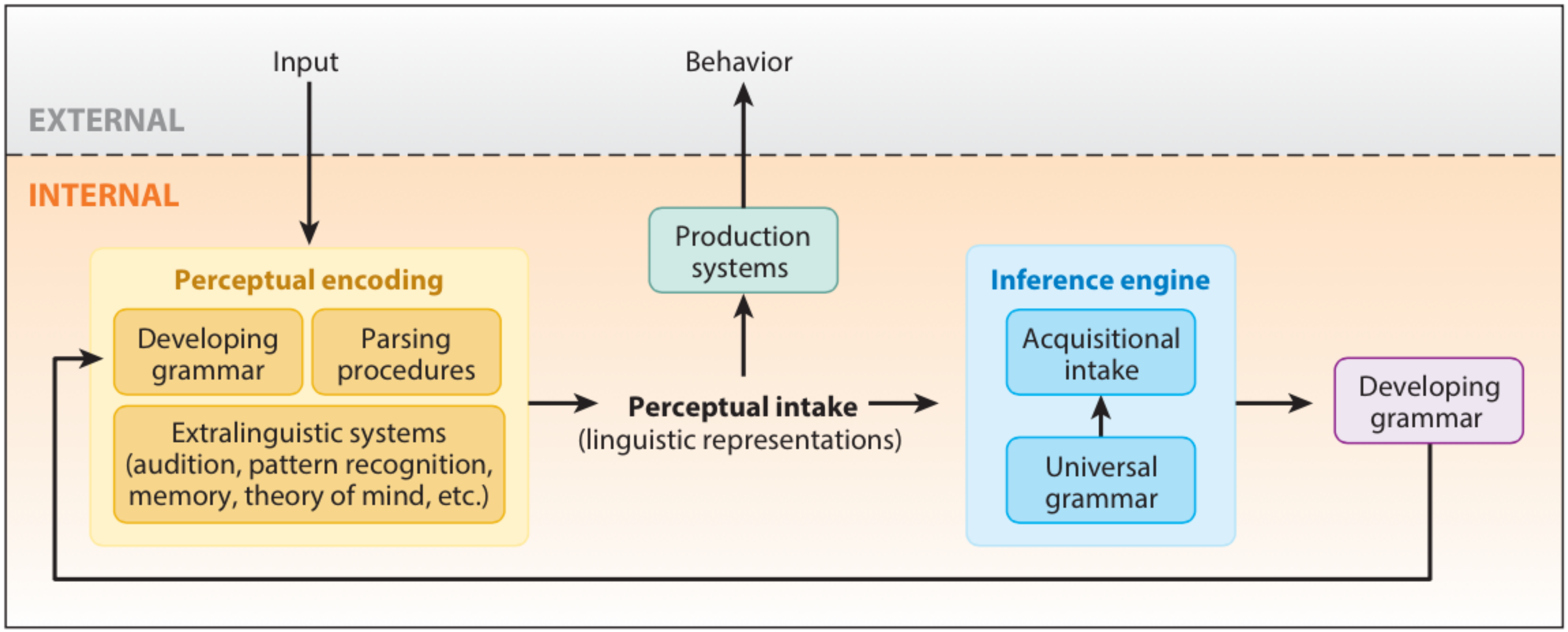
Important learning strategy components include

- theories of **representation**
- biases & capabilities that must exist for that knowledge to be successfully deployed during acquisition (= theories of the **learning process**).




Lidz & Gagliardi 2015

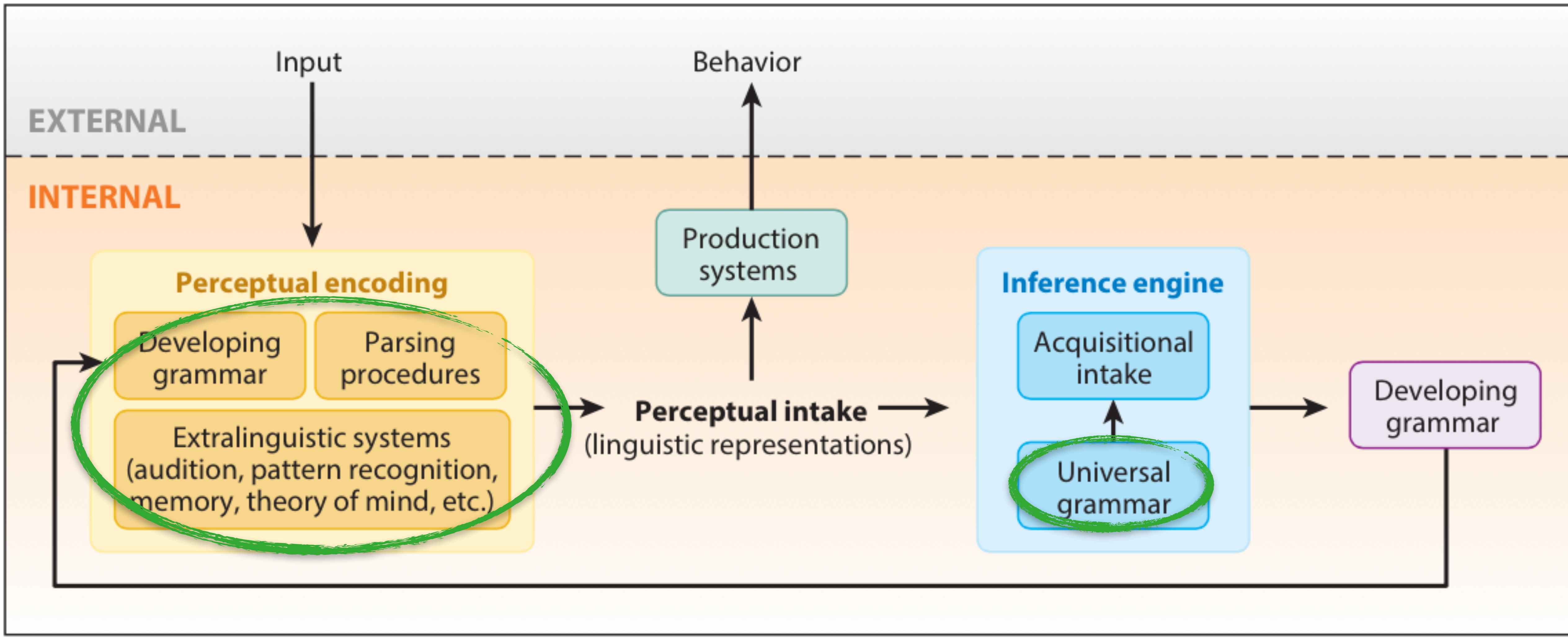
When building a specific model, it can be helpful to think about these different acquisition pieces in five main parts



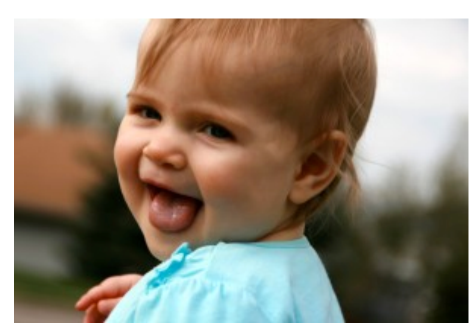

Lidz & Gagliardi 2015

# Initial state

What does the child start with? What knowledge, abilities, and learning biases does the child already have?




Lidz & Gagliardi 2015



# Initial state

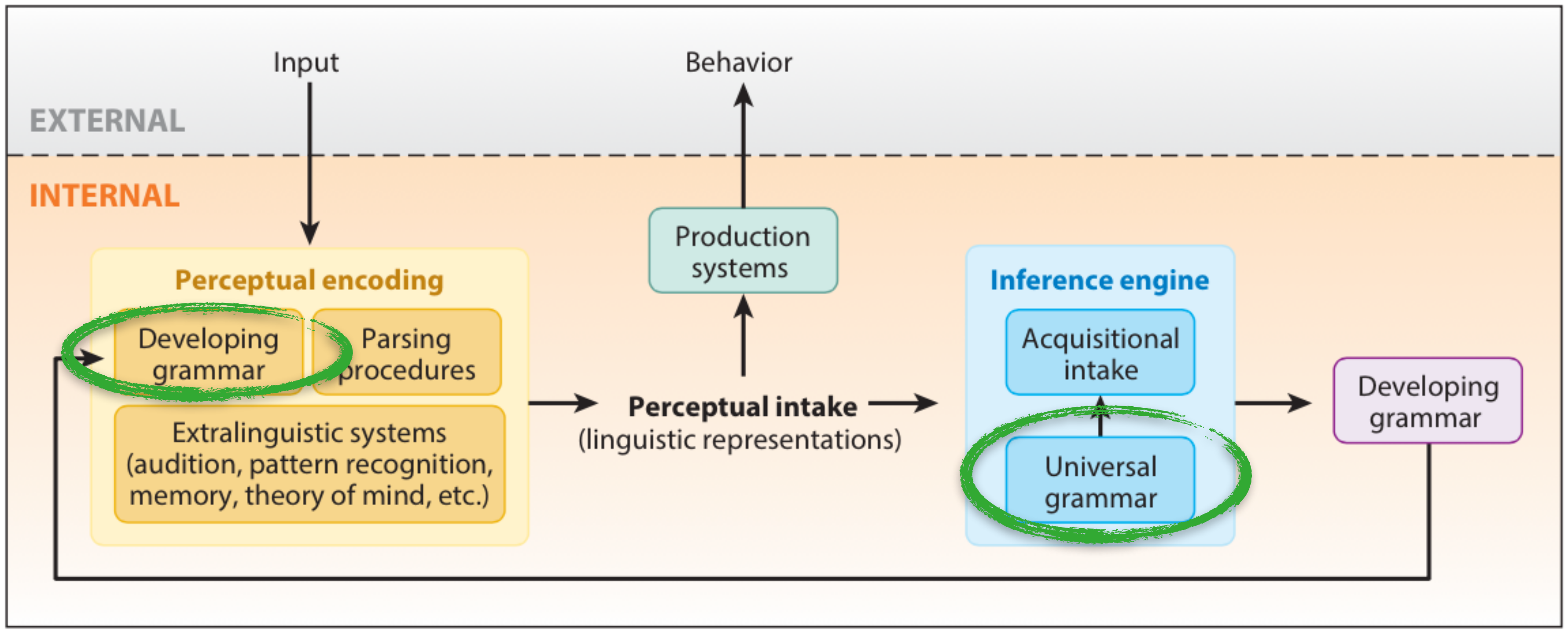
What does the child start with? What **knowledge**, abilities, and learning biases does the child already have?

Example knowledge:

syntactic categories exist and can be identified

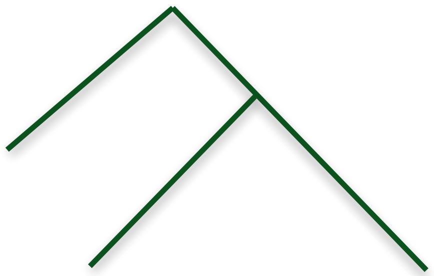
phrase structure exists and can be identified

participant roles can be identified




Lidz & Gagliardi 2015

N, V, Adj, P, …



Agent, Patient, Goal, …

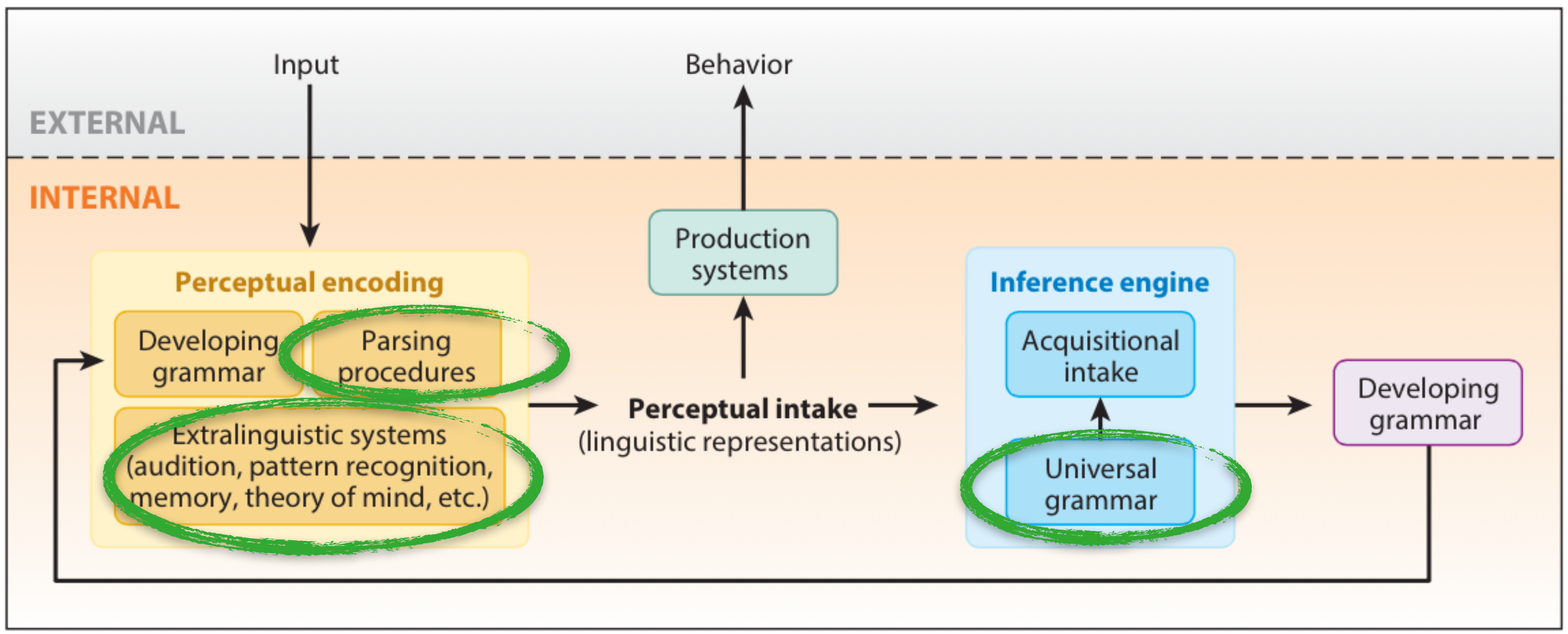
# Initial state

What does the child start with? What knowledge, **abilities**, and learning **biases** does the child already have?

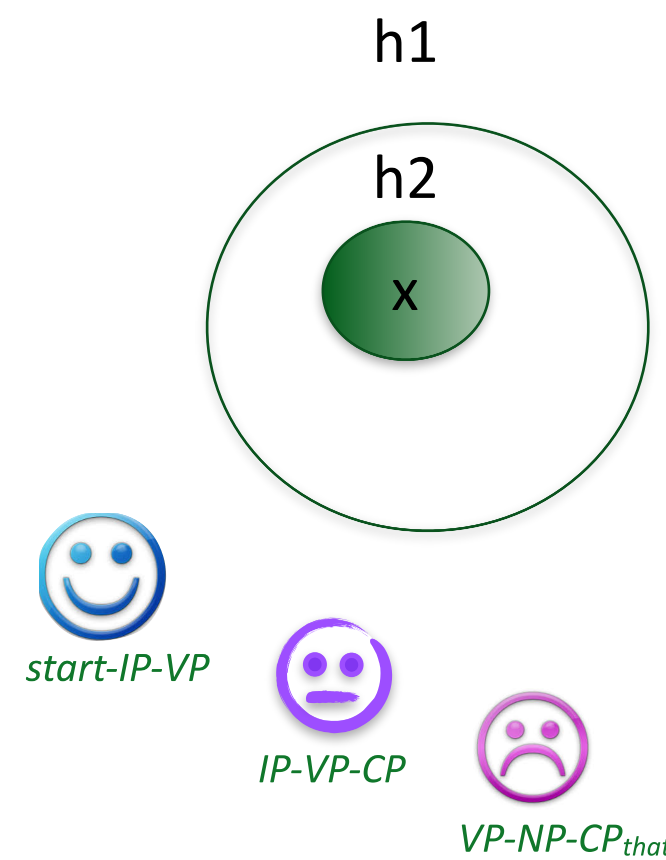
Example abilities & biases:

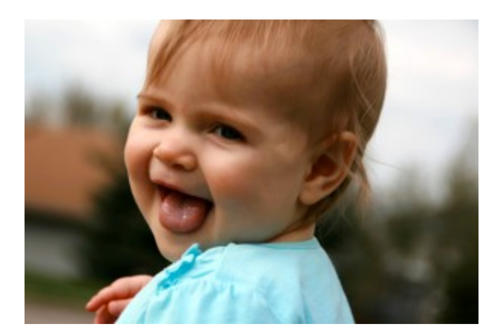
frequency information can be tracked

distributional information can be leveraged



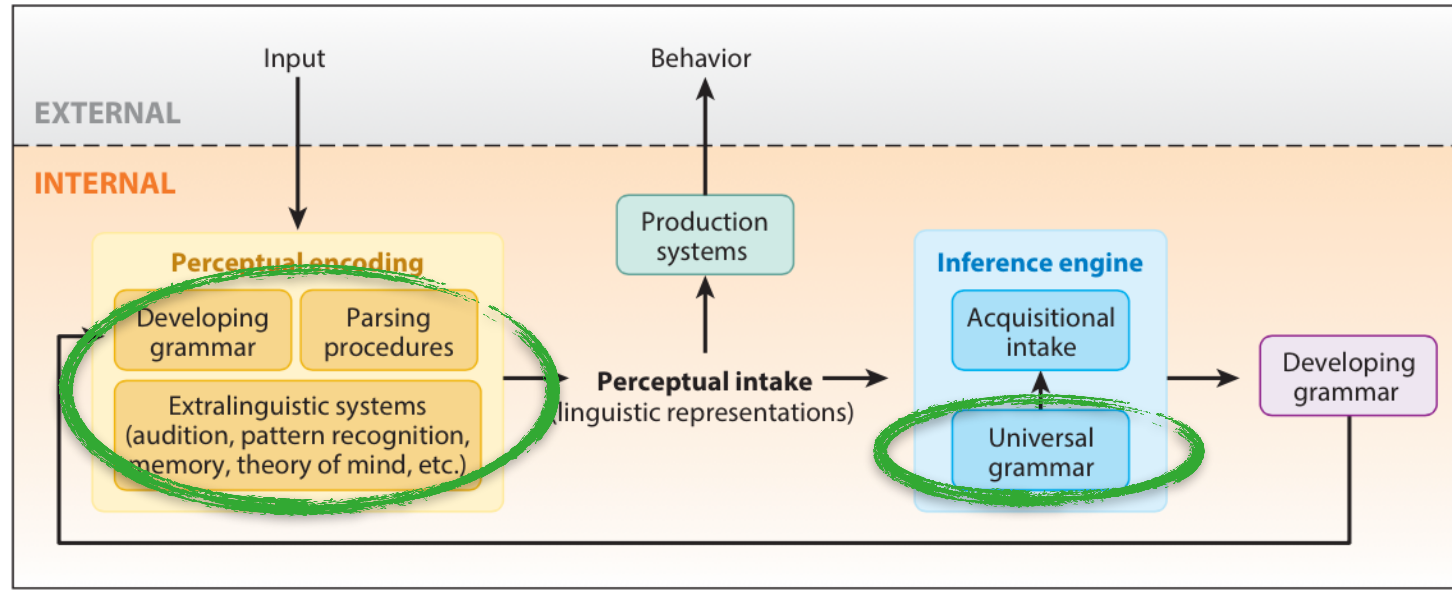

Lidz & Gagliardi 2015



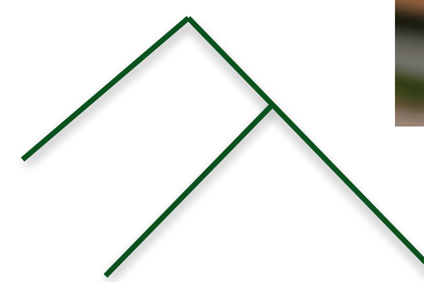
# Initial state

What does the child start with? What knowledge, **abilities**, and learning **biases** does the child already have?



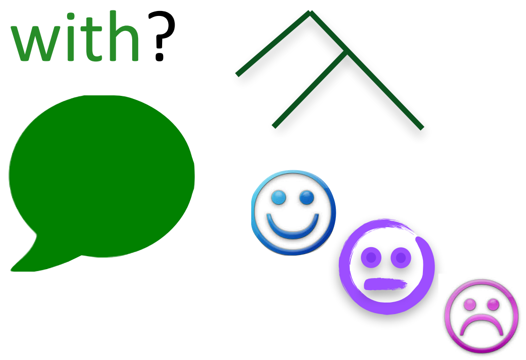

Lidz & Gagliardi

Example initial state: A strategy that depends on the frequency of certain syntactic structures would need the child to know about that syntactic structure via the developing grammar and/or Universal Grammar, recognize it in the input via the developing language processing abilities, and be able to track the frequency of that structure.



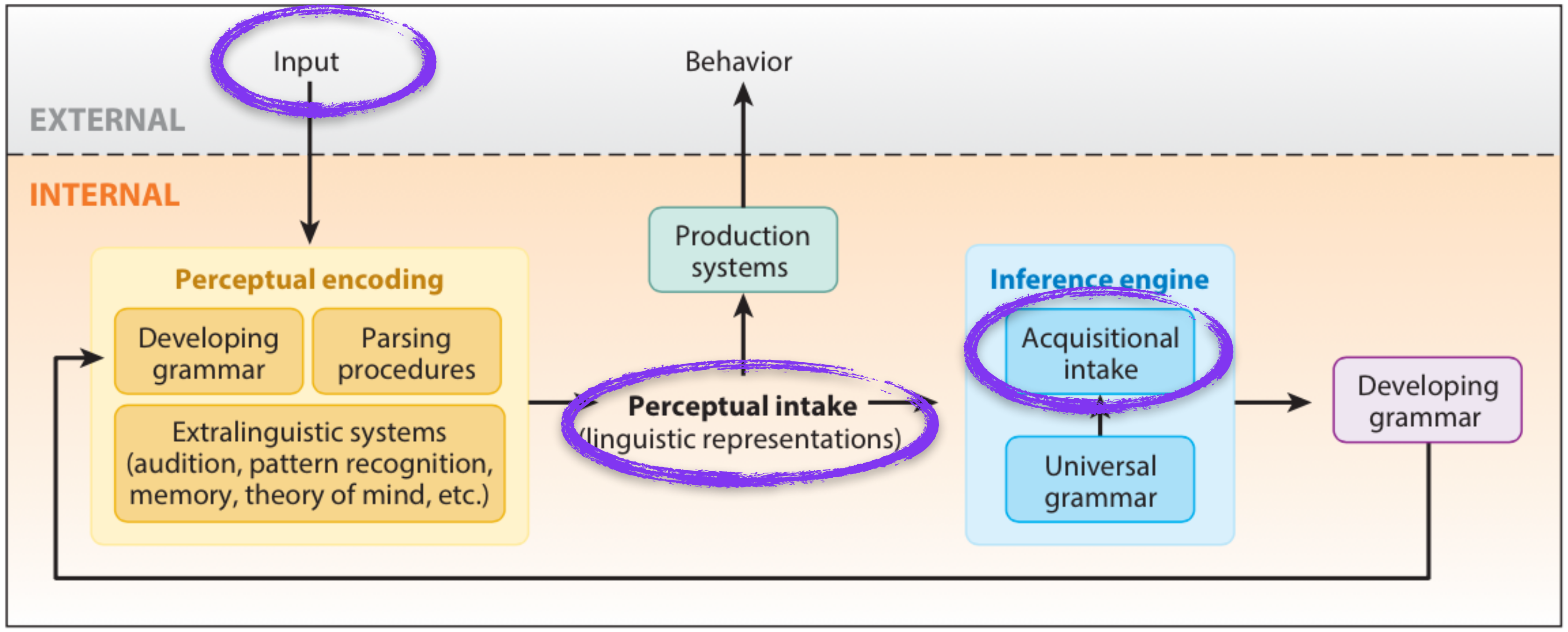
## Initial state

What knowledge, abilities, and learning biases does the child start with?



## Data intake

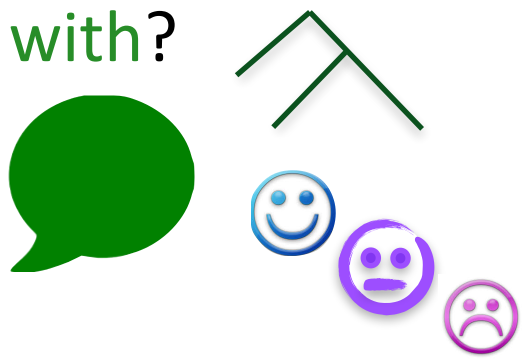
How does the modeled child perceive the input (=perceptual intake)? What part of the perceived data is used for acquisition (=acquisitional intake)?




Lidz & Gagliardi 2015

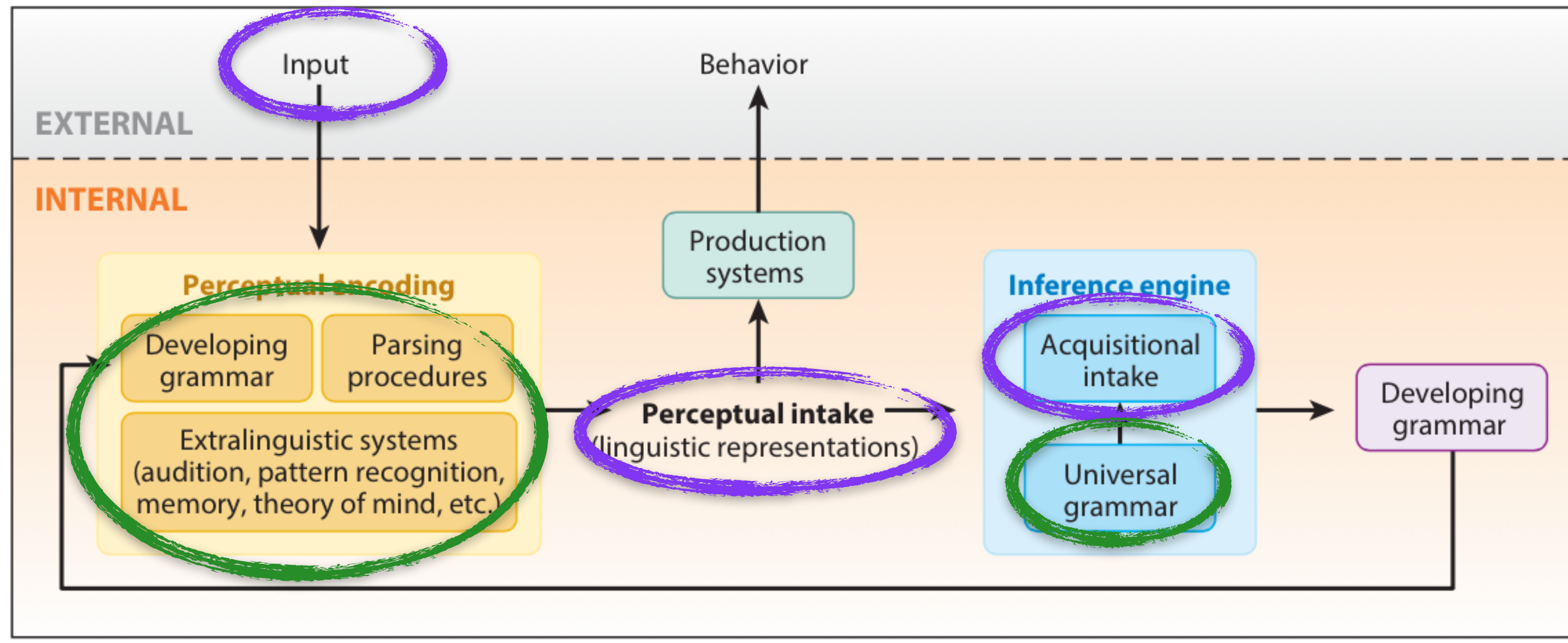
## Initial state

What knowledge, abilities, and learning biases does the child start with?



# Data intake

How does the modeled child perceive the input (=perceptual intake)? What part of the perceived data is used for acquisition (=acquisitional intake)?

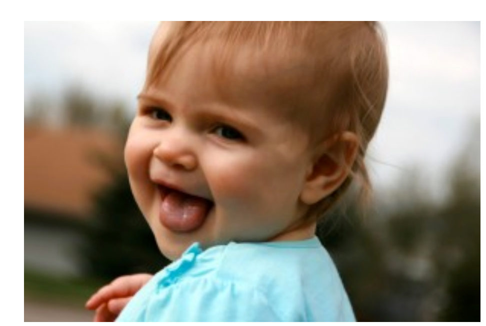



Lidz & Gagliardi 2015

ex: all *wh*-utterances for learning about *wh*-dependencies

ex: all pronoun data when learning about anaphoric *one*

ex: syntactic and conceptual data for learning syntactic knowledge that links with conceptual knowledge



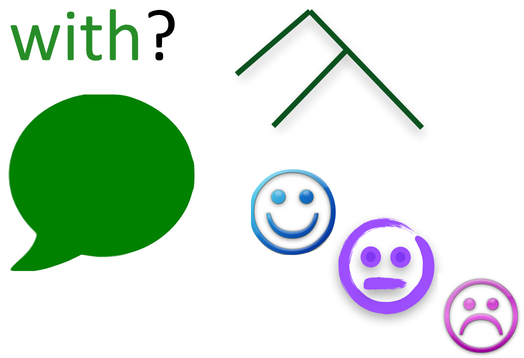
[defined by knowledge & biases/capabilities in the initial state]

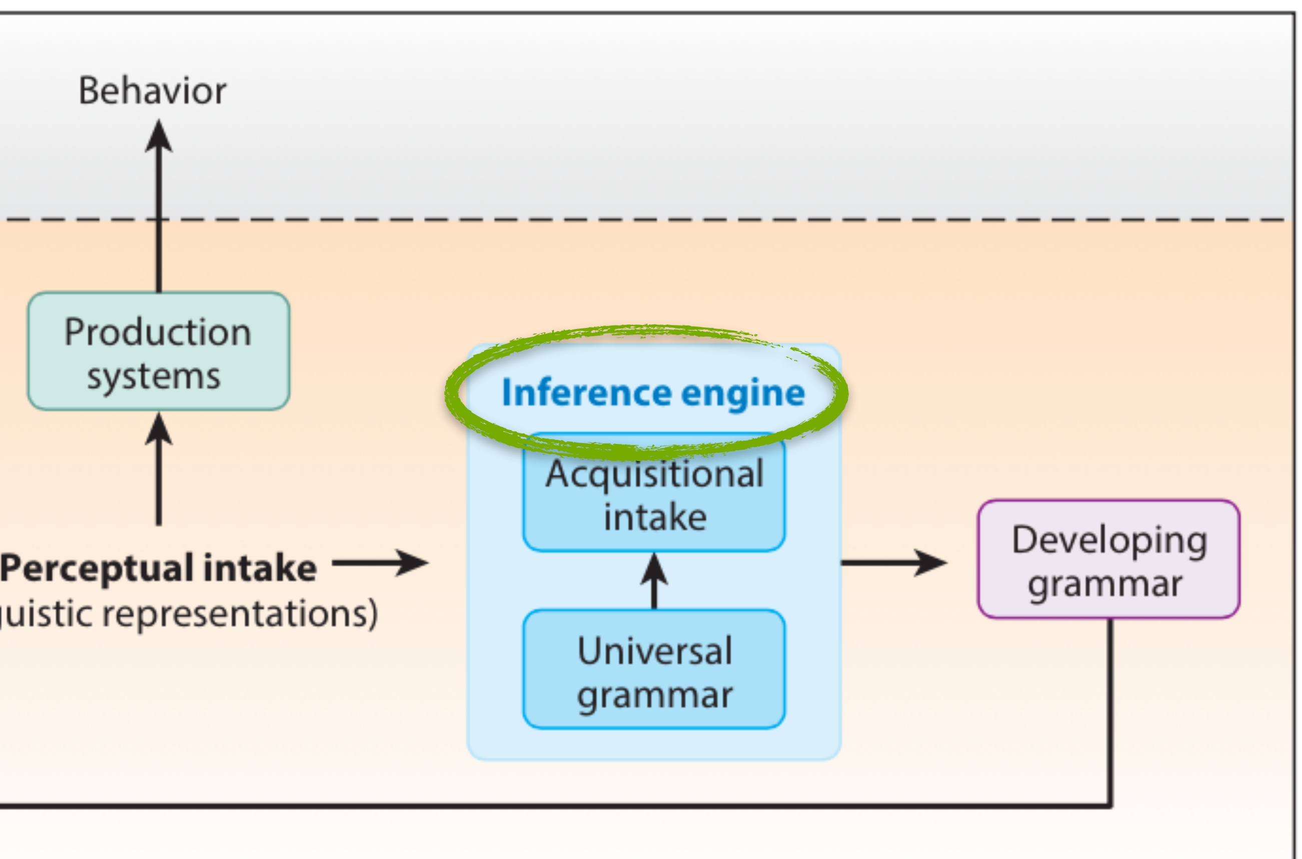
## Data intake

What is the acquisitional intake?

## Initial state

What knowledge, abilities, and learning biases does the child start with?





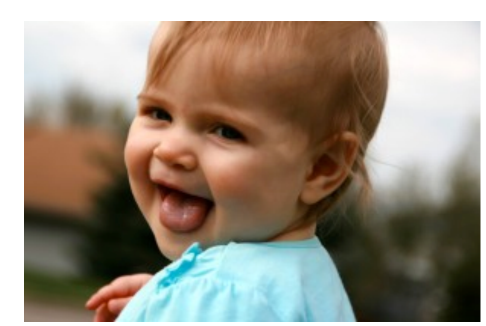

## Inference

How are updates made to the modeled child's internal representations?

ex: probabilistic integration of available information (e.g., Bayesian inference)

ex: sequential hypothesis testing

[defined by knowledge & biases/ capabilities in the initial state]

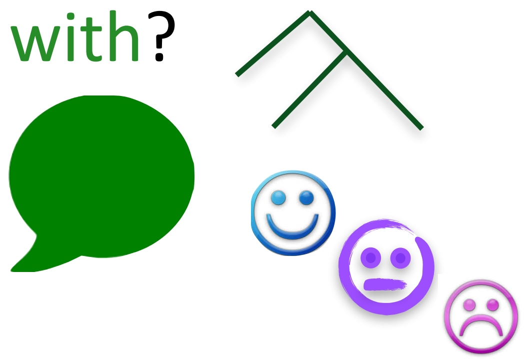


## Data intake

What is the acquisitional intake?

## Initial state

What knowledge, abilities, and learning biases does the child start with?

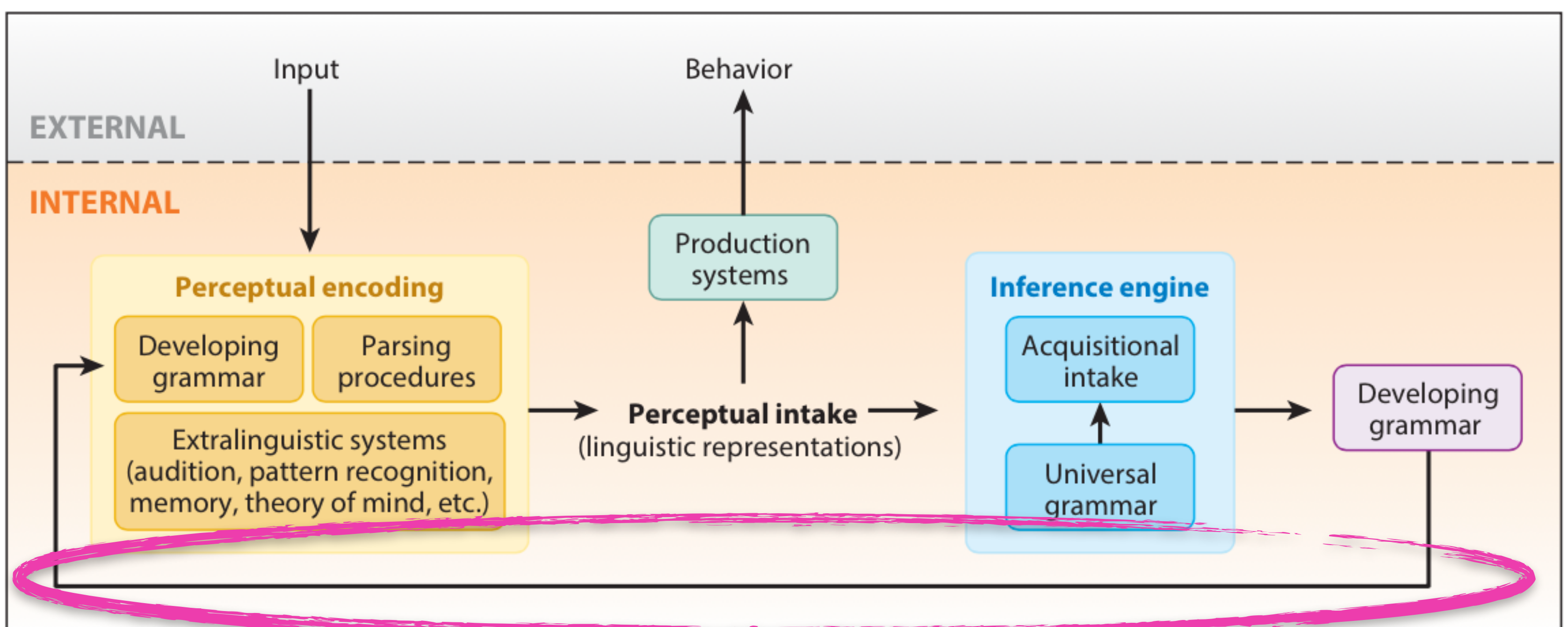


## Inference

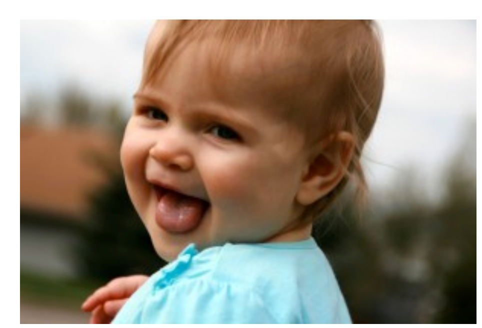
How are updates made?

## Learning period

How long does the child have to learn?




Lidz & Gagliardi 2015

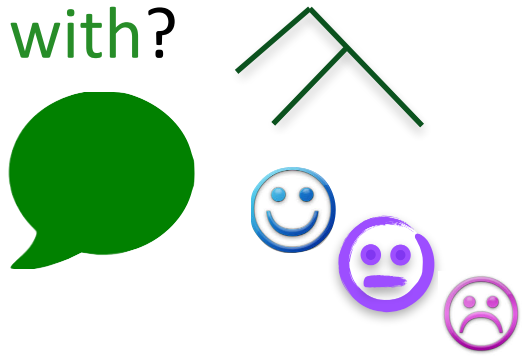


## Data intake

What is the acquisitional intake?

## Initial state

What knowledge, abilities, and learning biases does the child start with?



## Inference

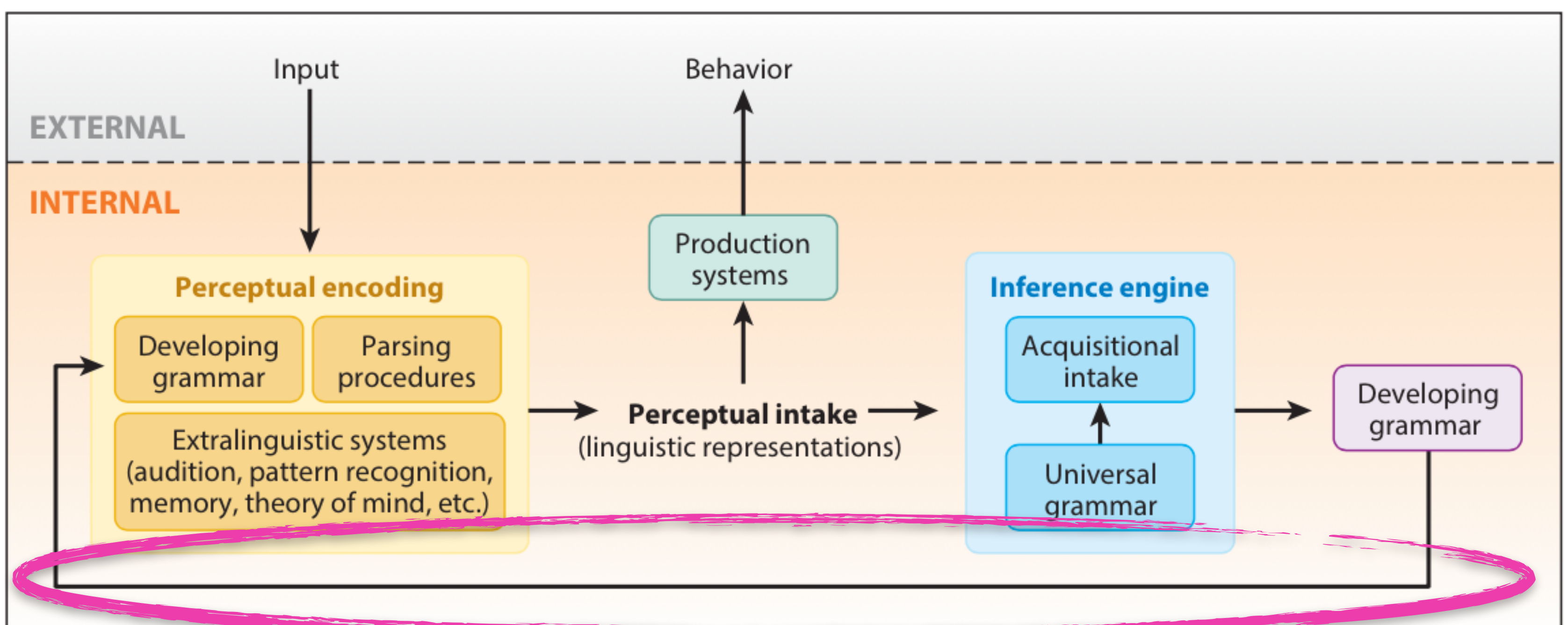
How are updates made?

## Learning period

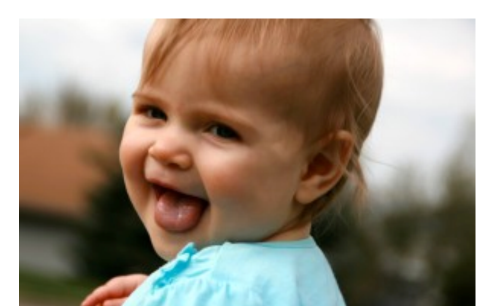
How long does the child have to learn?

ex: 3 years, ~1,000,000 data points

ex: 4 months, ~36,500 data points




Lidz & Gagliardi 2015

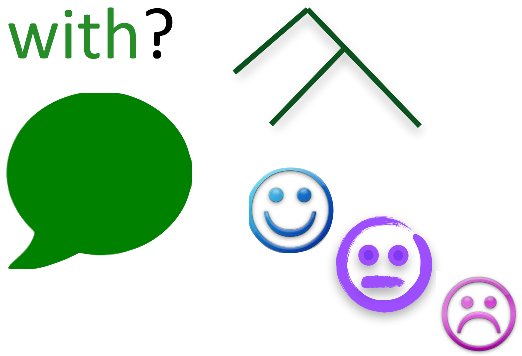


## Data intake

What is the acquisitional intake?

## Initial state

What knowledge, abilities, and learning biases does the child start with?



## Inference

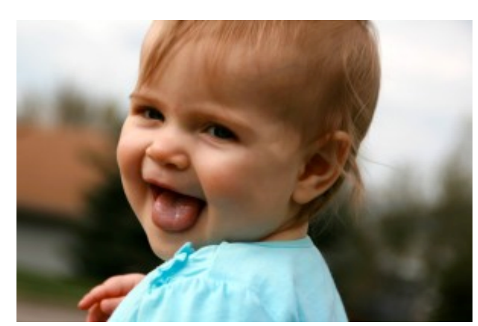
How are updates made?

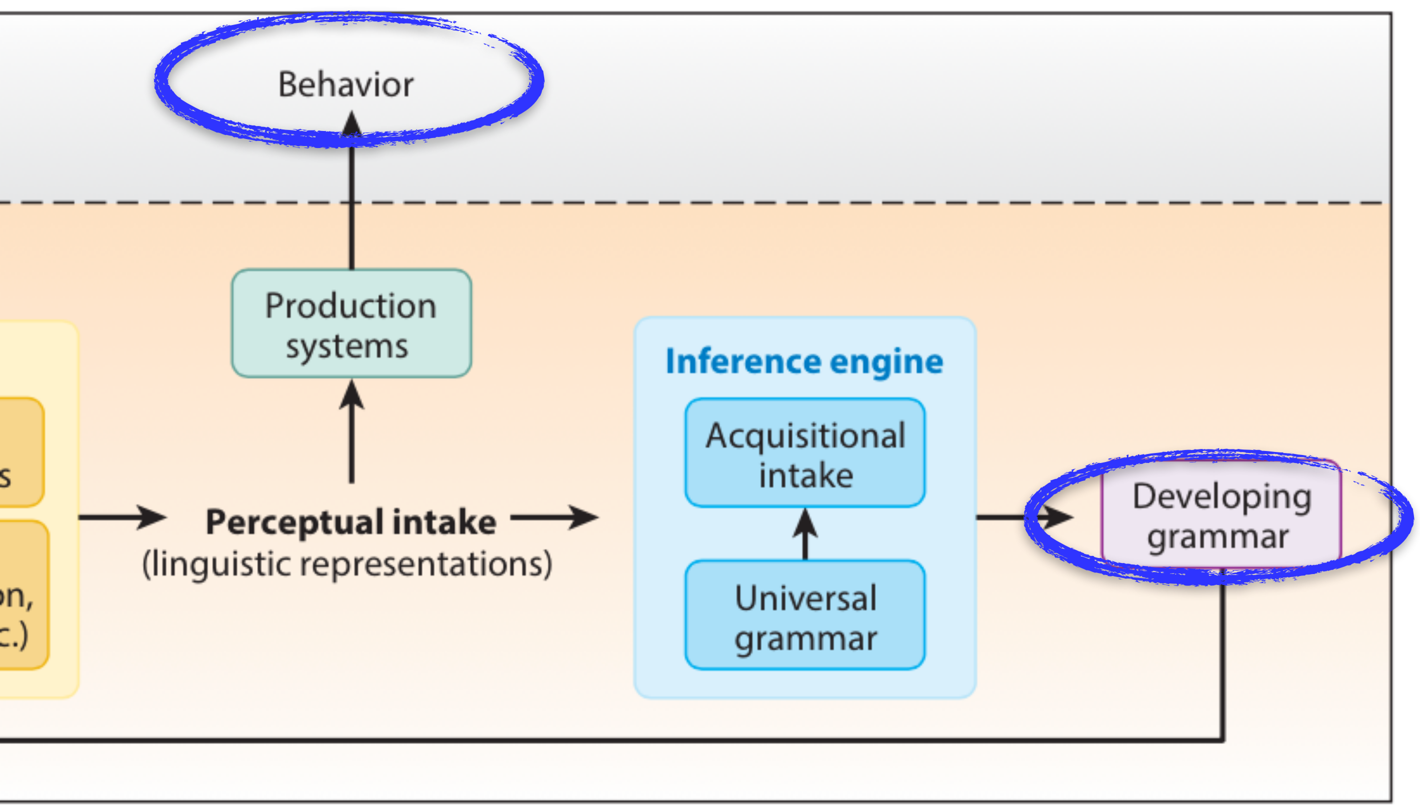
## Learning period

How long does the child have to learn?

## Target state

What does successful acquisition look like? What knowledge is the child trying to attain (often assessed in terms of observable behavior)?



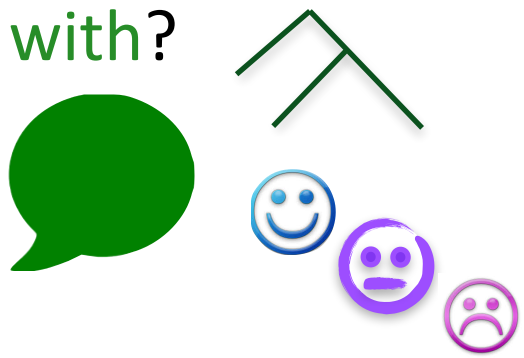


# Data intake

What is the acquisitional intake?

## Initial state

What knowledge, abilities, and learning biases does the child start with?

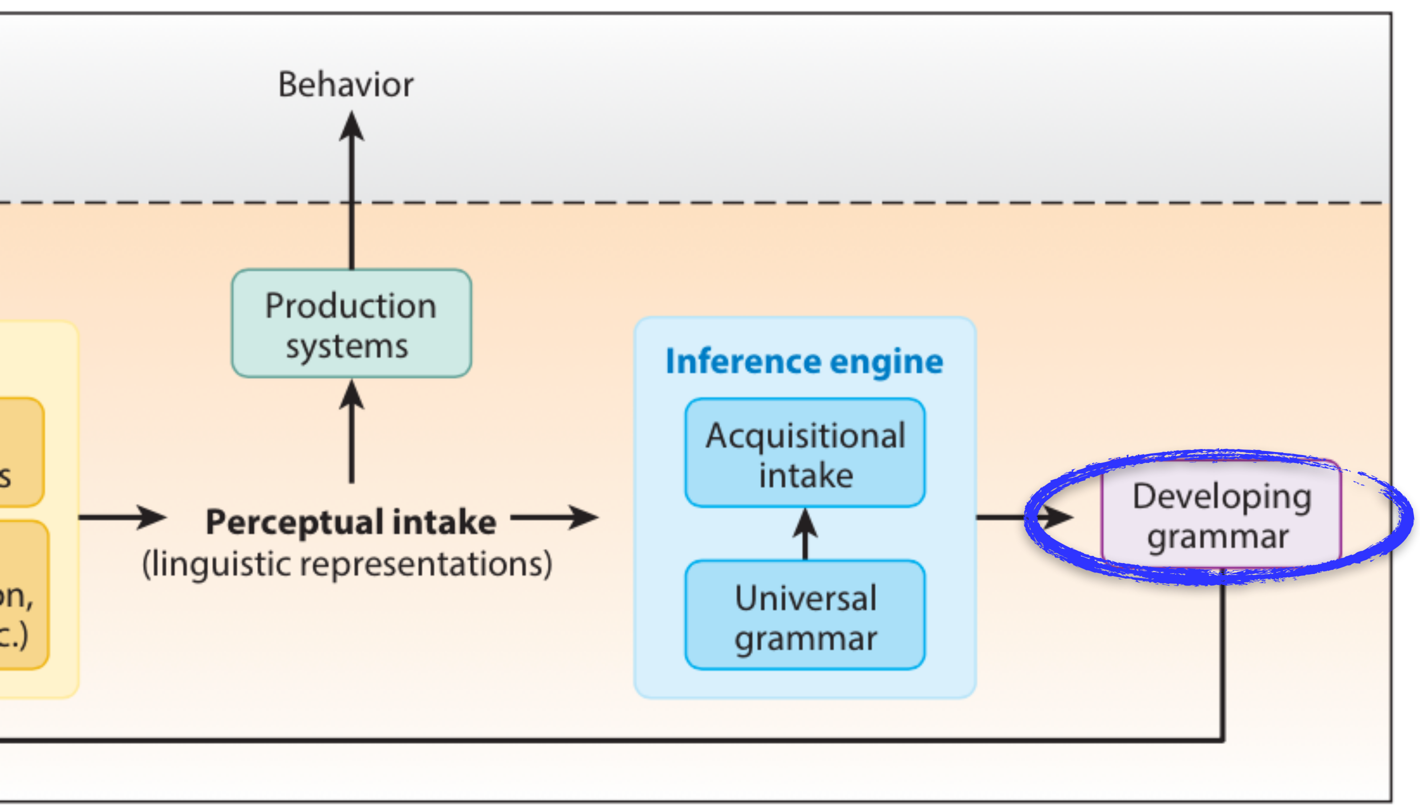


## Inference

How are updates made?

## Learning period

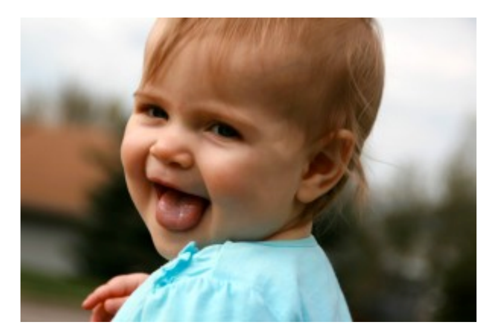
How long does the child have to learn?




## Target state

What does successful acquisition look like?

**knowledge**



ex: *Where did Jack think the necklace from __ was too expensive? 

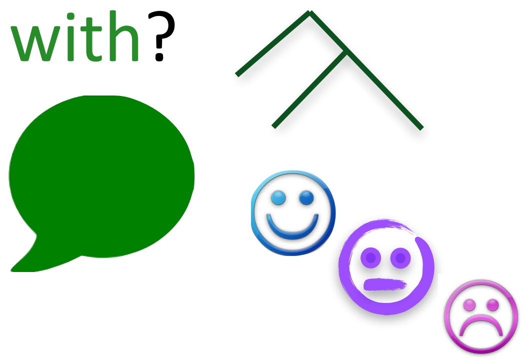
ex: Where did Jack buy a necklace from __ for Lily for her birthday? 

# Data intake

What is the acquisitional intake?

## Initial state

What knowledge, abilities, and learning biases does the child start with?

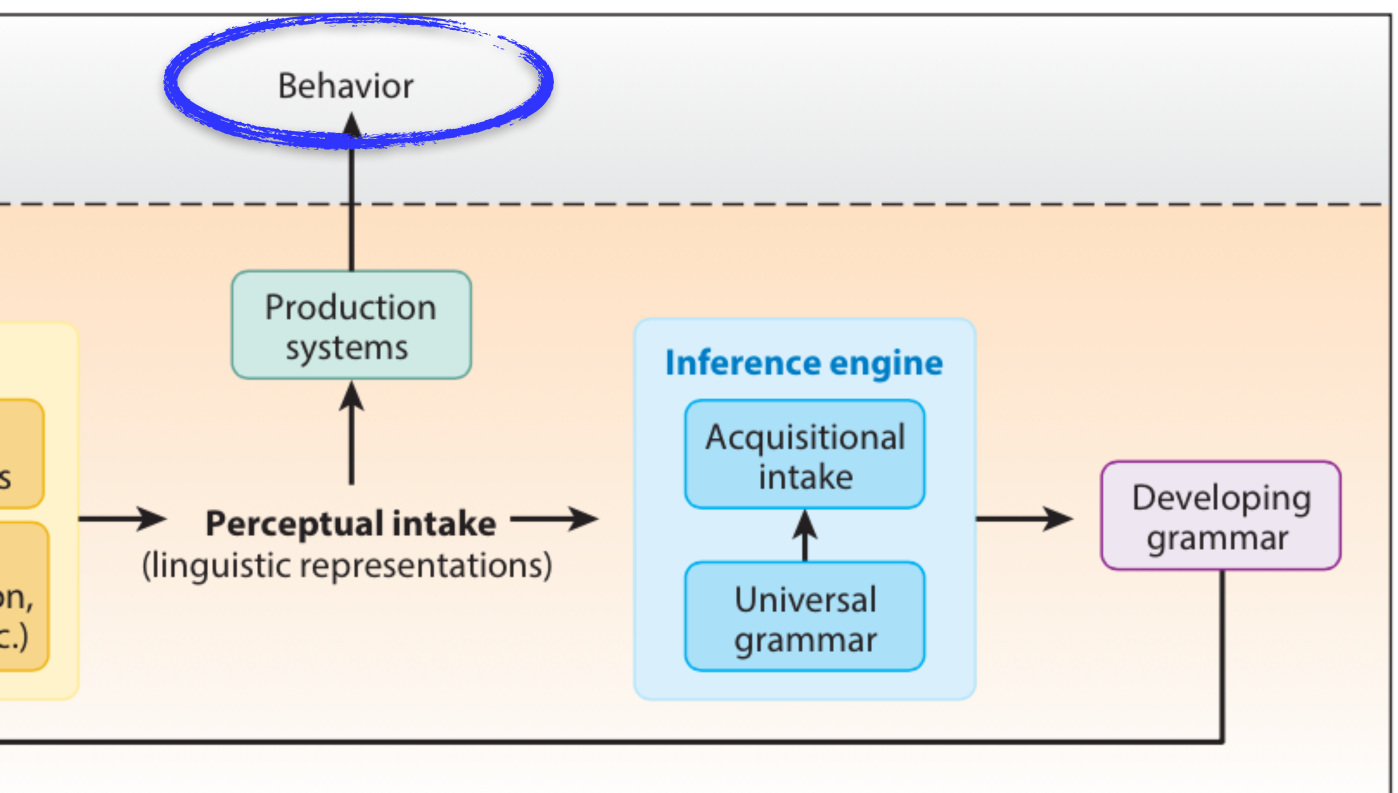


## Inference

How are updates made?

## Learning period

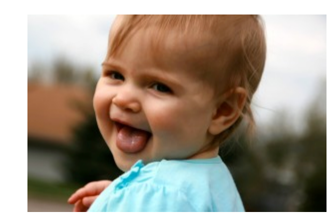
How long does the child have to learn?

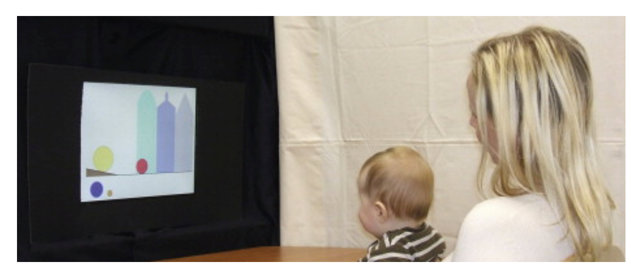



## Target state

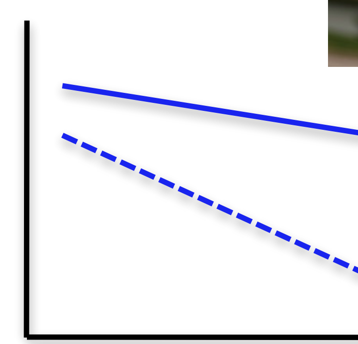
What does successful acquisition look like?

**behavior**





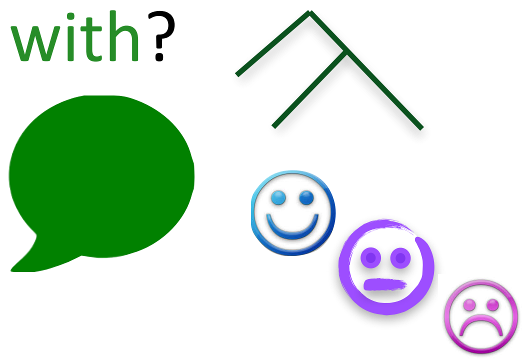
looking time preferences



z-score rating

## Initial state

What knowledge, abilities, and learning biases does the child start with?



## Data intake

What is the acquisitional intake?

## Inference

How are updates made?

## Target state

What does successful acquisition look like?

**knowledge**

**behavior**

## Learning period

How long does the child have to learn?

**Defining each of these pieces for a model (as relevant) can help streamline the modeling process and make sure we're building an informative model.**

# Building an informative model about…

Which learning strategies could children be using?

(Phillips & Pearl in press, Pearl 2017, Bar-Sever & Pearl 2016, Phillips & Pearl 2015a, 2015b, 2014a, 2014b, 2012; Pearl 2014, Pearl et al. 2011, Pearl et al. 2010)

# Building an informative model about...

Which learning strategies could children be using?

Which learning biases are necessary?

(Pearl & Sprouse in prep., Pearl, Ho, & Detrano in press, 2014; Pearl & Mis 2016, Pearl & Sprouse 2015, 2013a, 2013b, Pearl & Mis 2011, Pearl & Lidz 2009, Pearl 2008, Pearl & Weinberg 2007)

# Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

Which knowledge representations are learnable — and which aren't?

(Pearl, Ho, & Detrano in press, 2014; Pearl 2017, Pearl 2011, Pearl 2009)

## Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

Which knowledge representations are learnable — and which aren't?

When do children learn different aspects of the linguistic system?

(Bates, Pearl, & Braunwald in prep., Nguyen & Pearl in press, Caponigro, Pearl et al. 2012, Caponigro, Pearl et al. 2011)

# Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

Which knowledge representations are learnable — and which aren't?

When do children learn different aspects of the linguistic system?

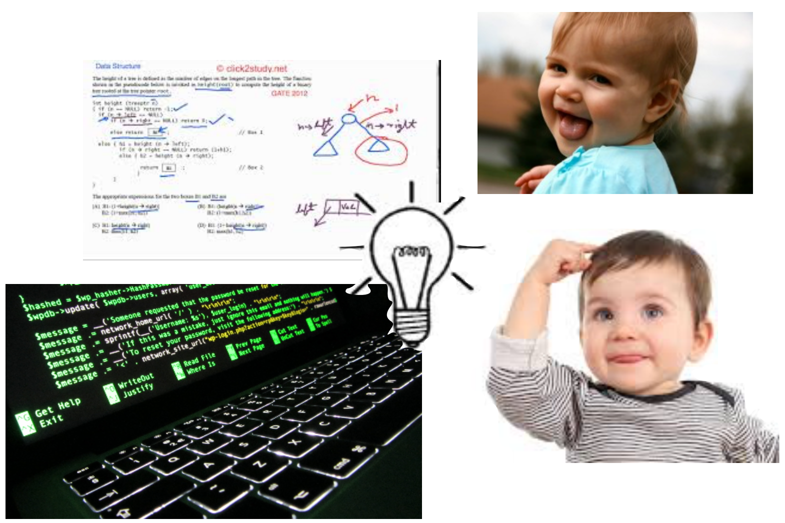
What factors affect children's observable behavior?

(Savinelli, Scontras, & Pearl in prep., Nguyen & Pearl in press, Savinelli, Scontras, & Pearl 2017)

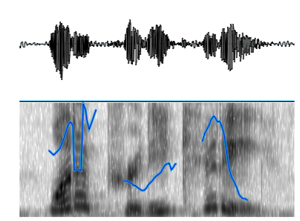
# Today's Plan:
# Computational models of language acquisition

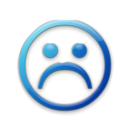
I. Why

II. How



III. What we can learn





Noun

✓KI tty

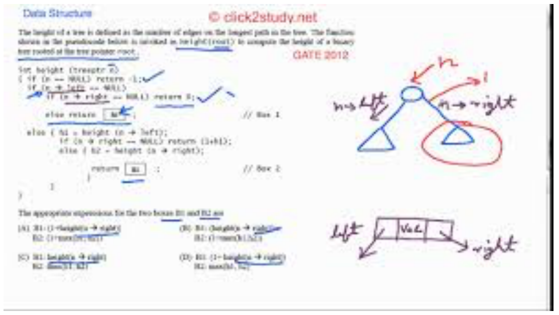
*Who does… is pretty?*

*another one*

*Every kitty didn't …*

# Today's Plan:
# Computational models of language acquisition

## II. How

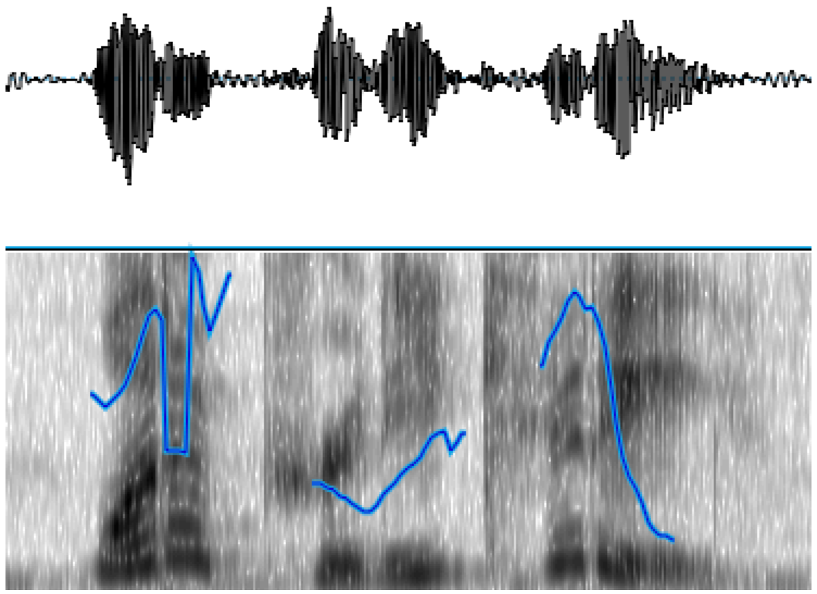



A concrete example

# How do we model language acquisition?

A concrete example with speech segmentation

# How do we model language acquisition?
## An example with speech segmentation



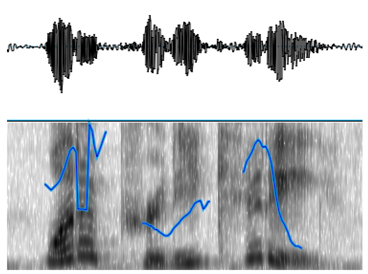
= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

# How do we model language acquisition?

## An example with speech segmentation



what a pretty kitty!

**(1) Decide what kind of learner the model represents**

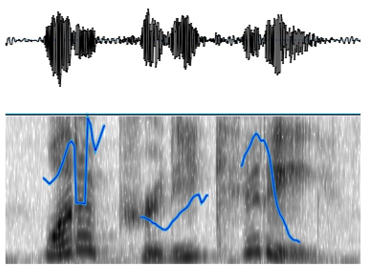
This depends on what task you're modeling

For the first stages of speech segmentation:

Typically developing 6- to 8-month-old child learning first language

# How do we model language acquisition?

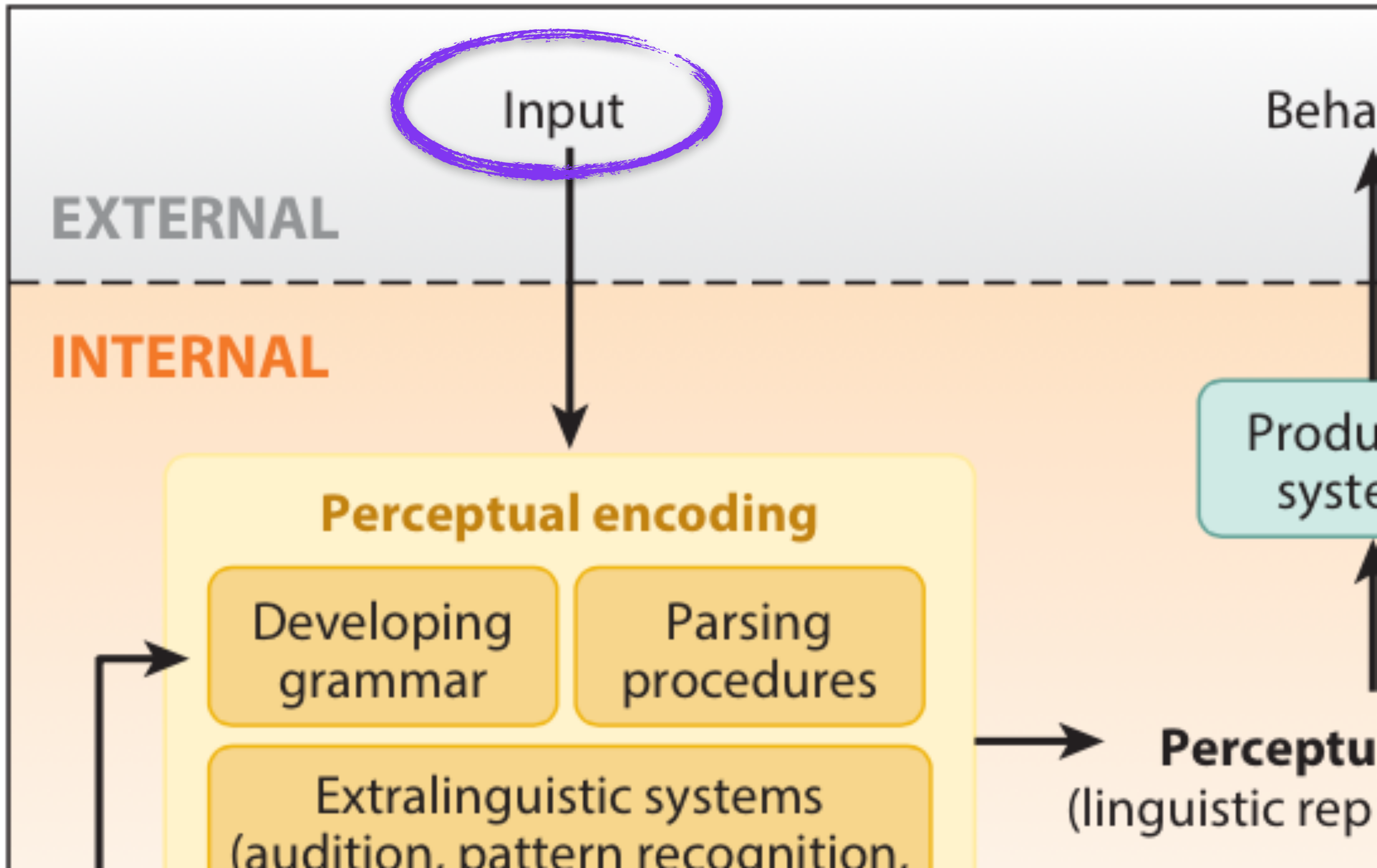
## An example with speech segmentation



what a pretty kitty!

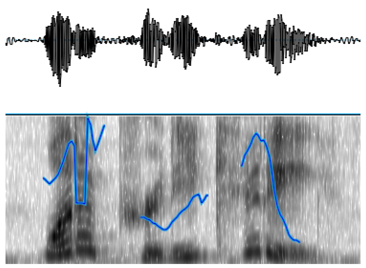
### (2) Decide what data the child learns from (input)

This depends on your acquisition theory and the empirical data available



# How do we model language acquisition?

## An example with speech segmentation



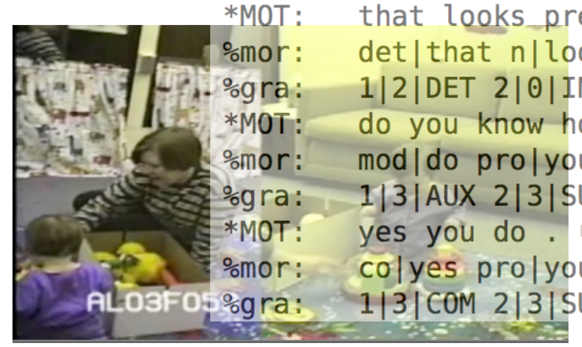
what a pretty kitty!

## (2) Decide what data the child learns from (input)

Example empirical data: CHILDES database
[http://childes.talkbank.org](http://childes.talkbank.org)

**CHILDES** Child Language Data Exchange System

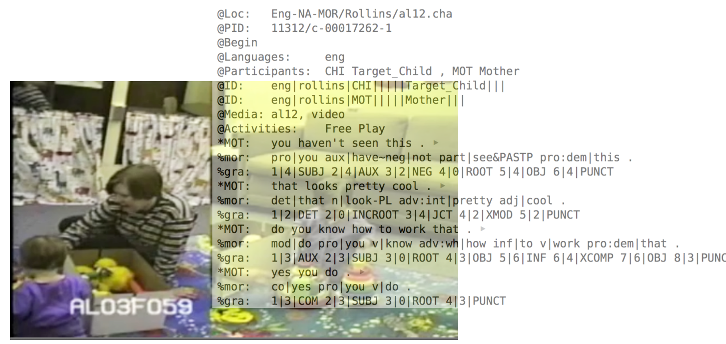
Video/audio recordings of speech samples, along with transcriptions and some structural annotations.

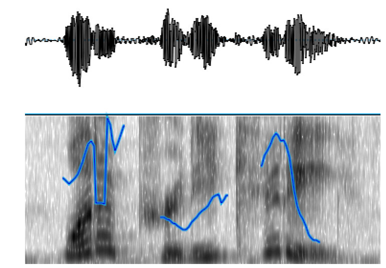



```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI|||||Target_Child|||
@ID:    eng|rollins|MOT|||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PU
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

# How do we model language acquisition?

## An example with speech segmentation



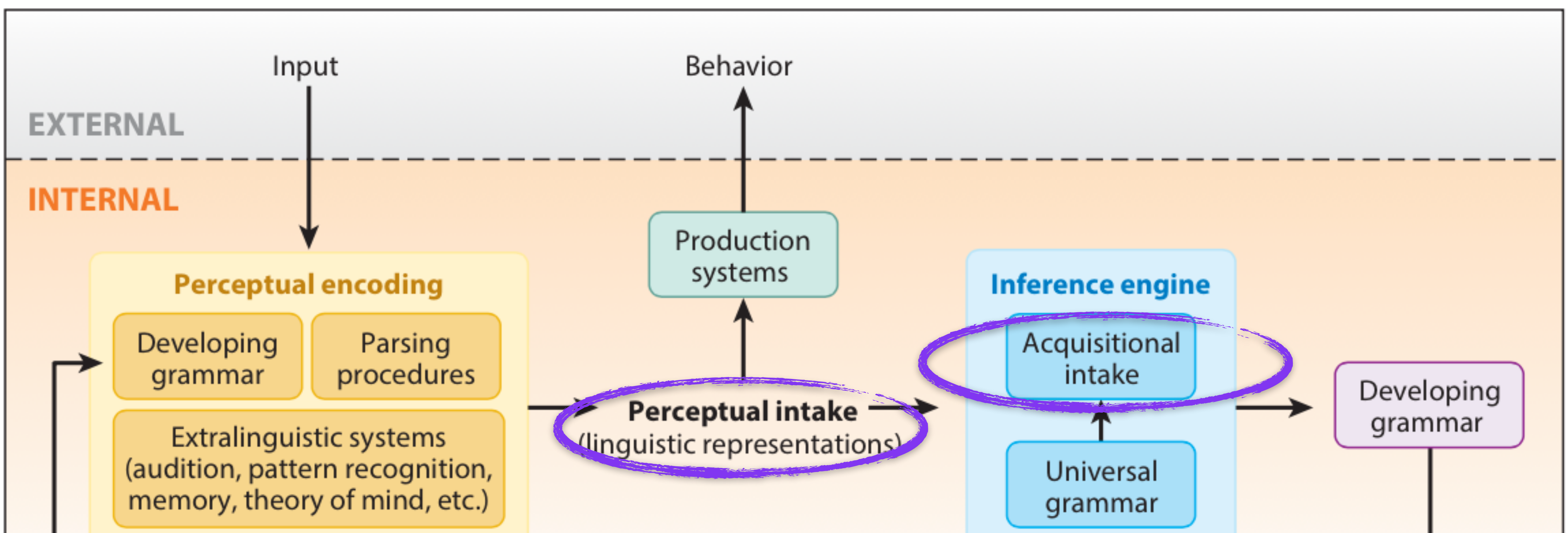



what a pretty kitty!

**(3) Decide how the child perceives the data,**

**and which data are relevant (intake)**

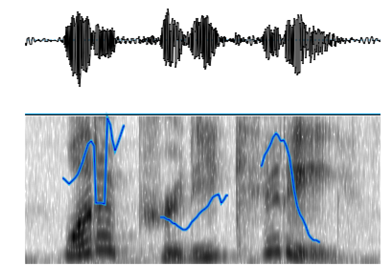
This depends on your acquisition theory



# How do we model language acquisition?

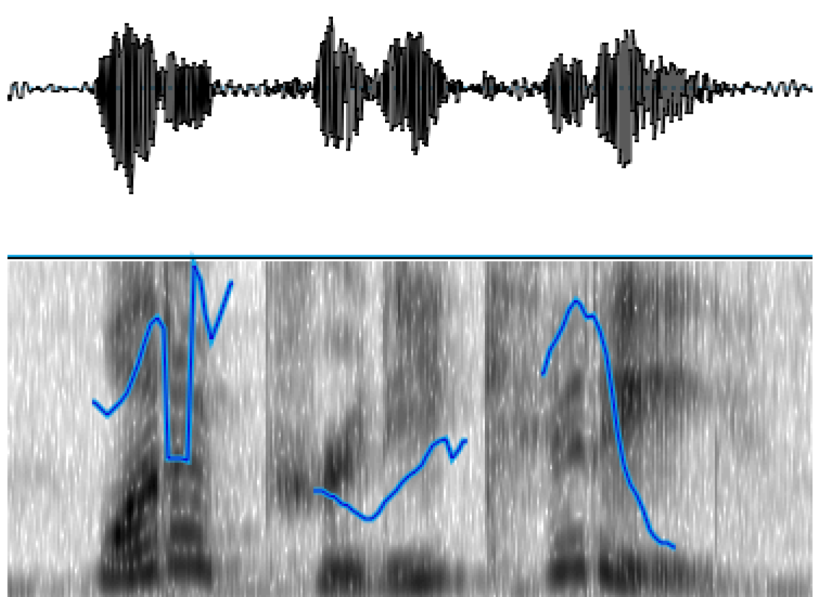
## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT|||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```



what a pretty kitty!

**(3) Decide how the child perceives the data, and which data are relevant (intake)**



*syllables with stress*

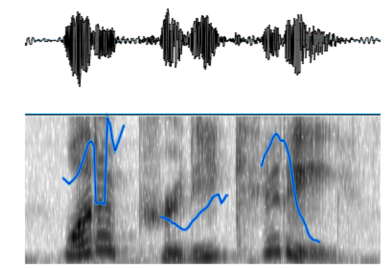
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

# How do we model language acquisition?

## An example with speech segmentation

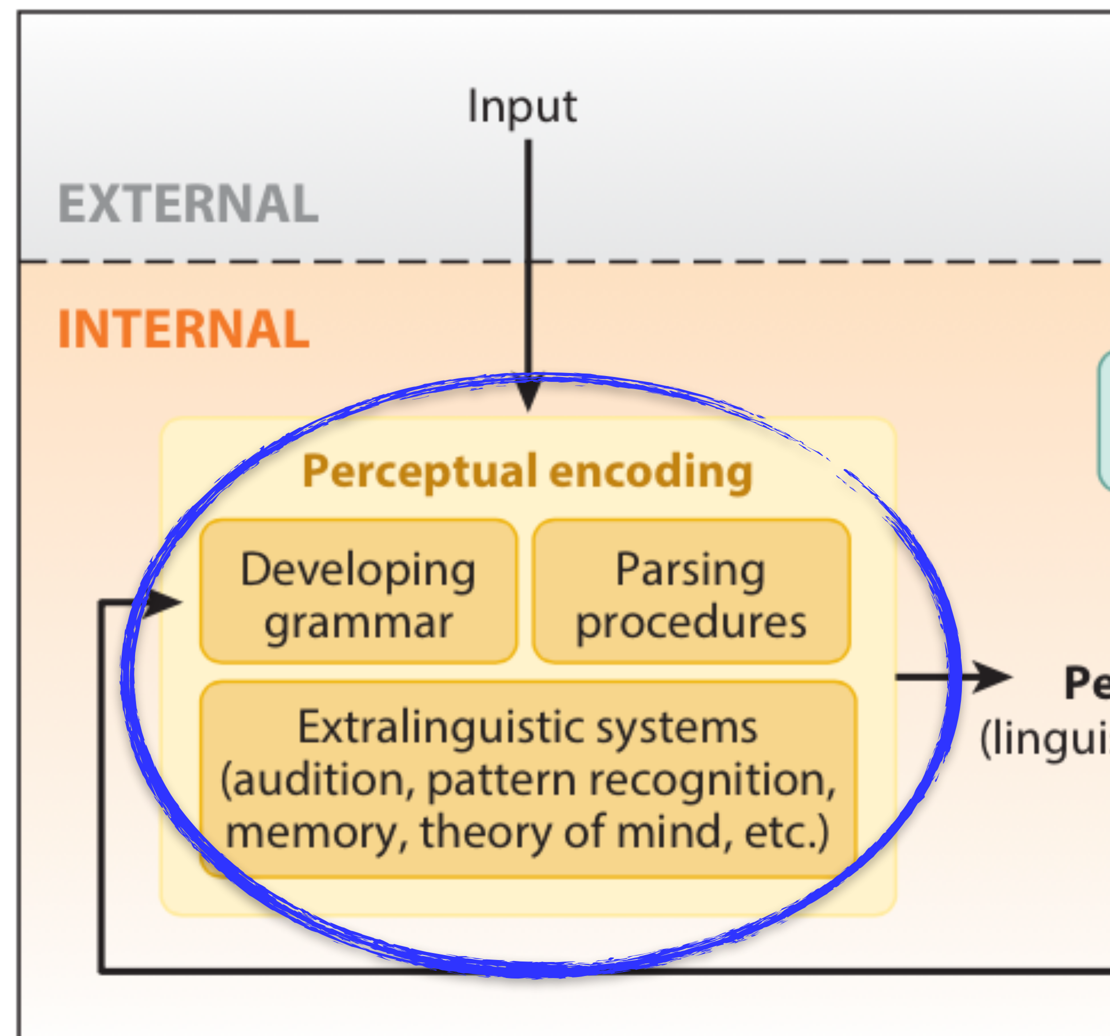
```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

Many models will try to make **cognitively plausible** assumptions about how the child is representing and processing input data

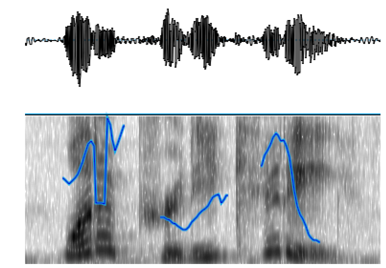


# How do we model language acquisition?

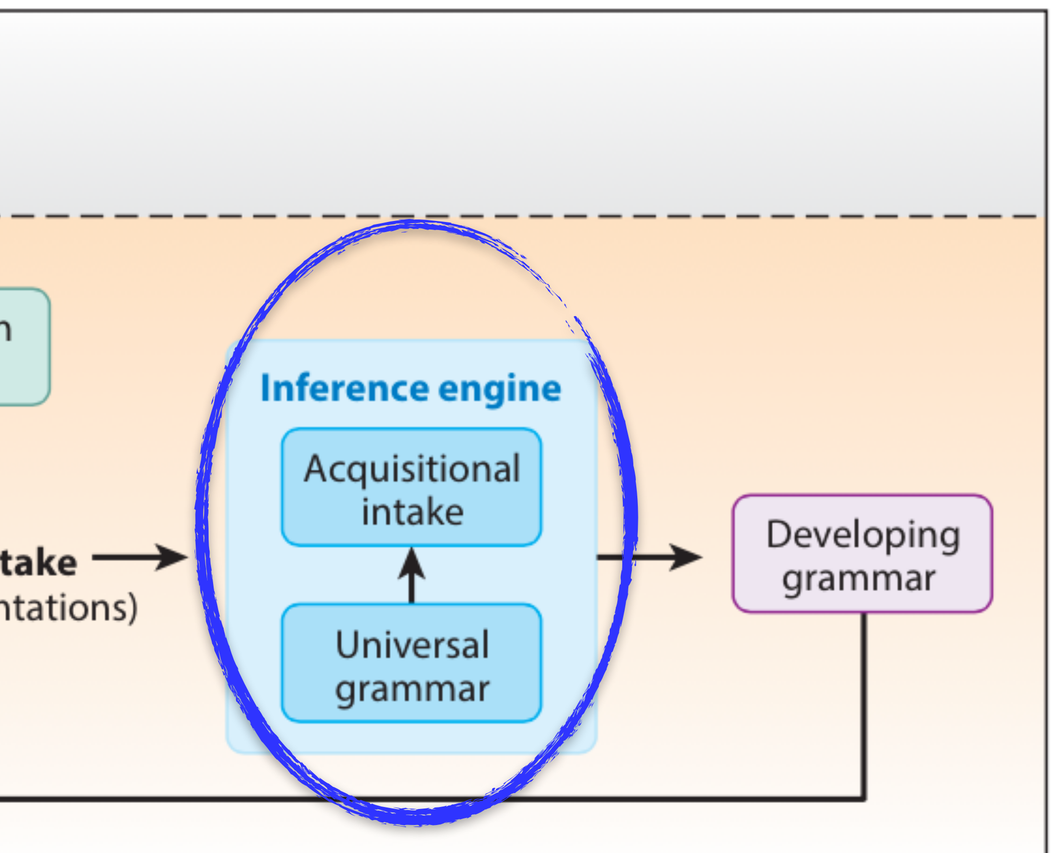
## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!




**(4) Decide what hypotheses the child has and what information is being tracked in the input**

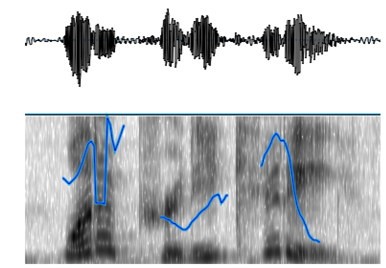
This depends on your acquisition theory

# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

**(4) Decide what hypotheses the child has and what information is being tracked in the input**

Example hypotheses: what the words are

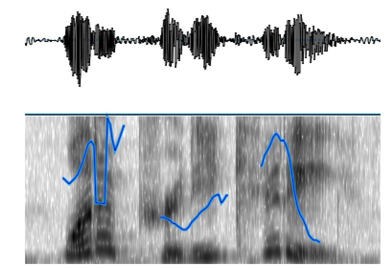
| wˈʌɾə | wˈʌ | wˈʌɾə |
|---|---|---|
| pɹˈɪɾi | ɾə | pɹˈɪɾikˈɪɾi |
| kˈɪɾi | pɹˈɪɾikˈɪɾi | |

# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

| wˈʌɾə<br>pɹˈɪɾi<br>kˈɪɾi | wˈʌ<br>ɾə<br>pɹˈɪɾikˈɪɾi | wˈʌɾə<br>pɹˈɪɾikˈɪɾi |
|---|---|---|

**(4) Decide what hypotheses the child has and what information is being tracked in the input**

Example information:

transitional probability between syllables,

stress on syllables

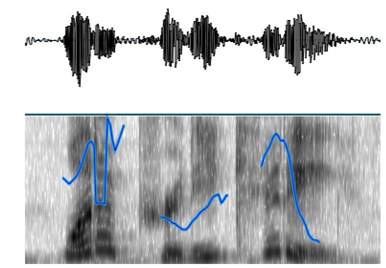
wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:	Eng-NA-MOR/Rollins/al12.cha
@PID:	11312/c-00017262-1
@Begin
@Languages:	eng
@Participants:	CHI Target_Child , MOT Mother
@ID:	eng|rollins|CHI||||Target_Child|||
@ID:	eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:	Free Play
*MOT:	you haven't seen this .
%mor:	pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:	1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:	that looks pretty cool .
%mor:	det|that n|look-PL adv:int|pretty adj|cool .
%gra:	1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:	do you know how to work that .
%mor:	mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:	1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:	yes you do .
%mor:	co|yes pro|you v|do .
%gra:	1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

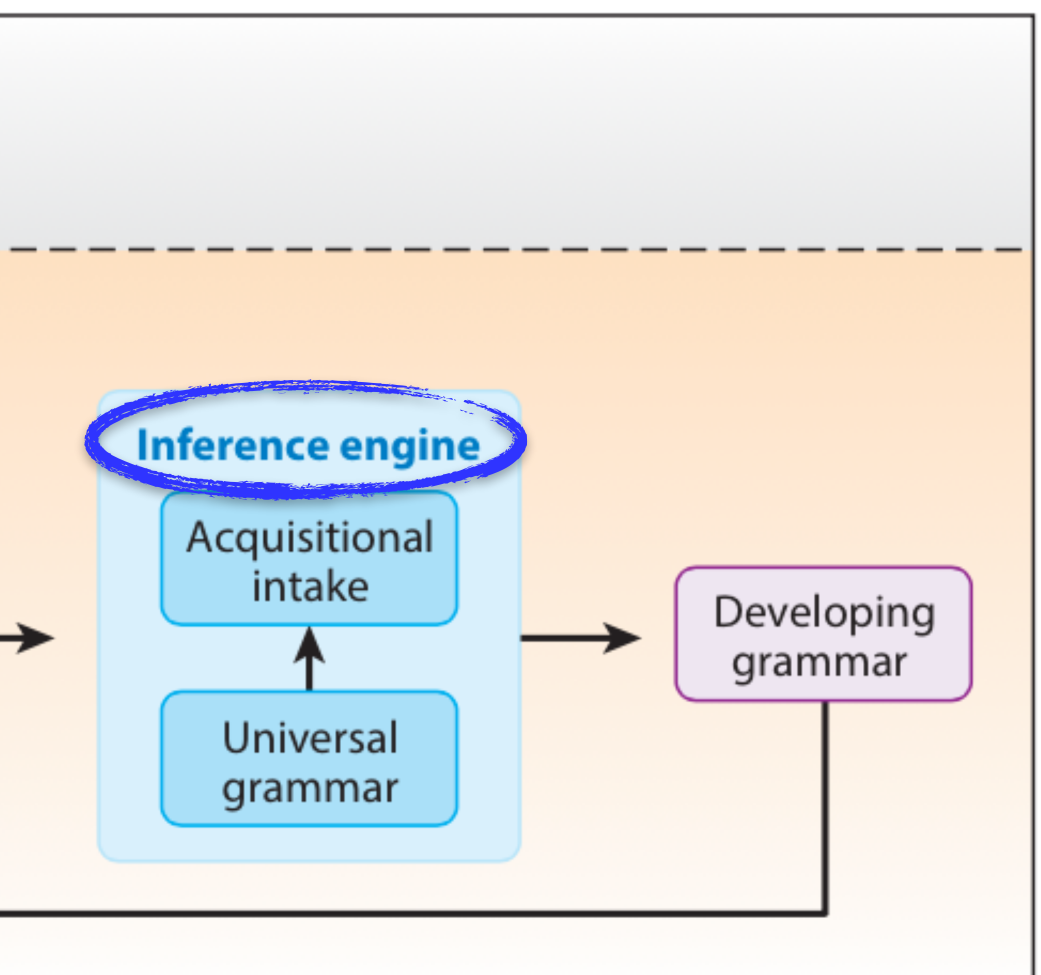
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

| wˈʌɾə<br>pɹˈɪɾi<br>kˈɪɾi | wˈʌ<br>ɾə<br>pɹˈɪɾikˈɪɾi | wˈʌɾə<br>pɹˈɪɾikˈɪɾi |
|---|---|---|



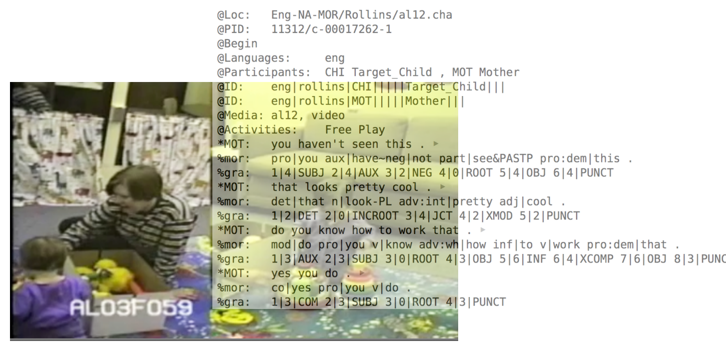

**(5) Decide how belief in different hypotheses is updated**

This depends on your acquisition theory

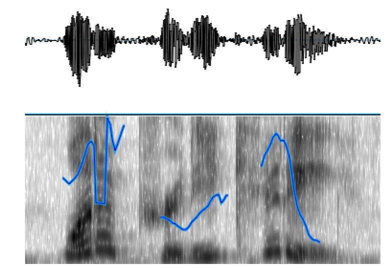
Example: based on transitional probability between syllables

# How do we model language acquisition?

## An example with speech segmentation




= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



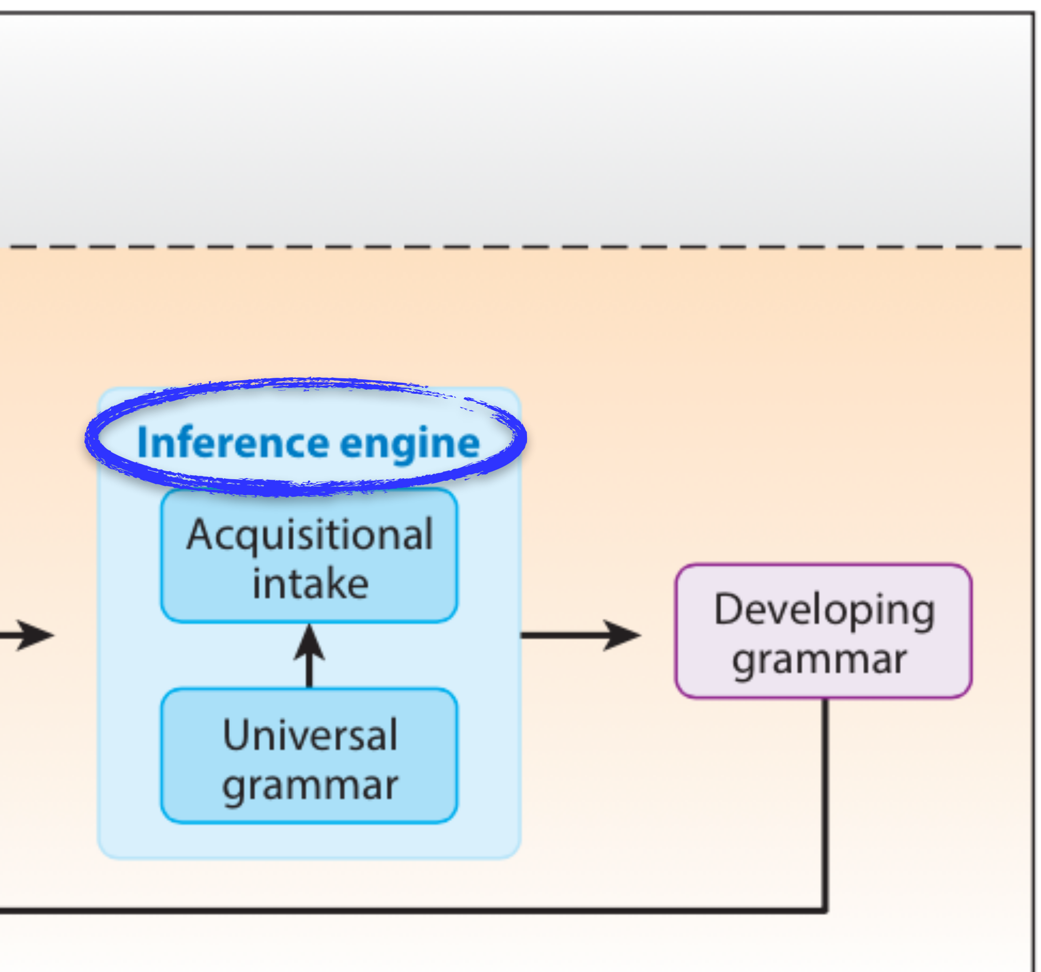
what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi




**(5) Decide how belief in different hypotheses is updated**

This depends on your acquisition theory

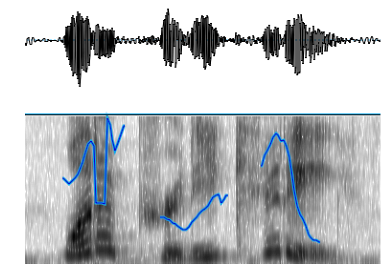
Example: based on transitional probability between syllables

# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



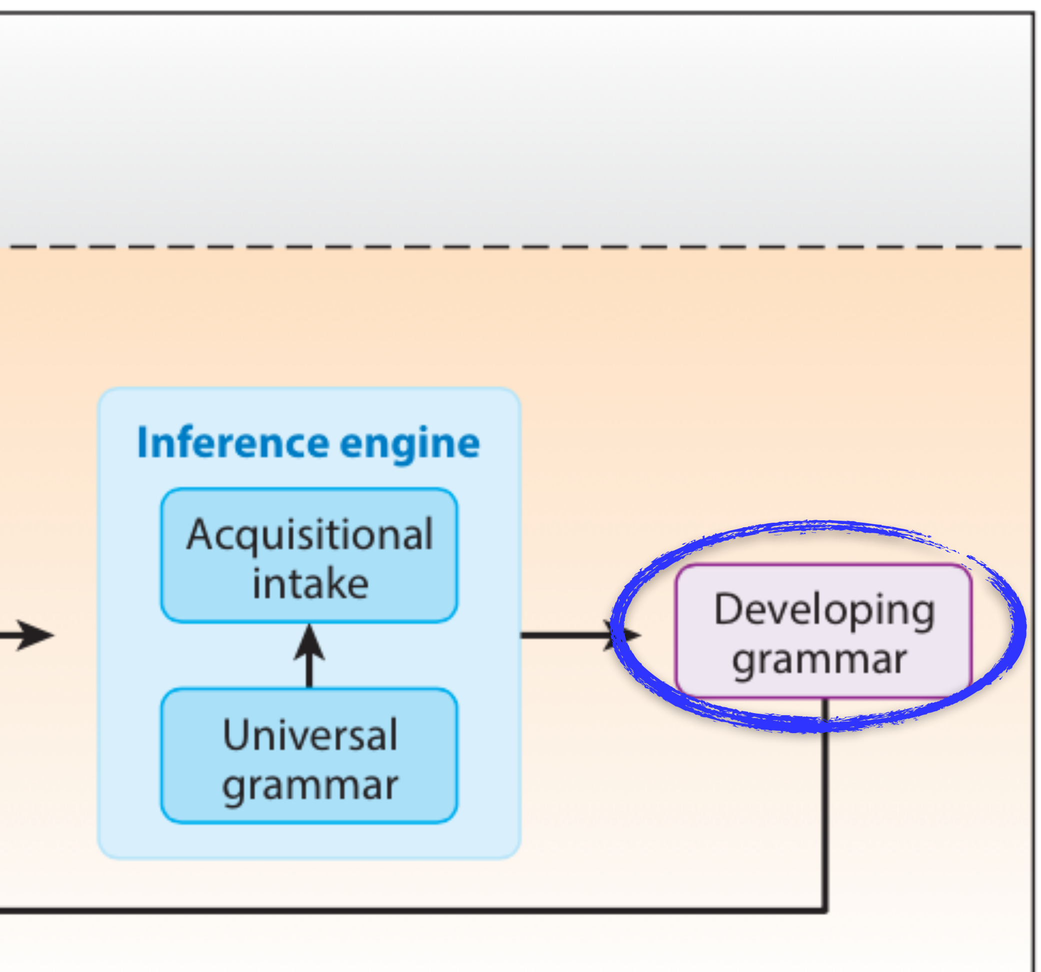
what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi




**(6) Decide what the measure of success is**

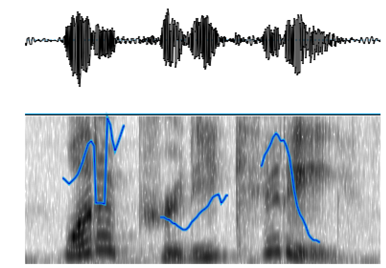
This can be based on your theory…

# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

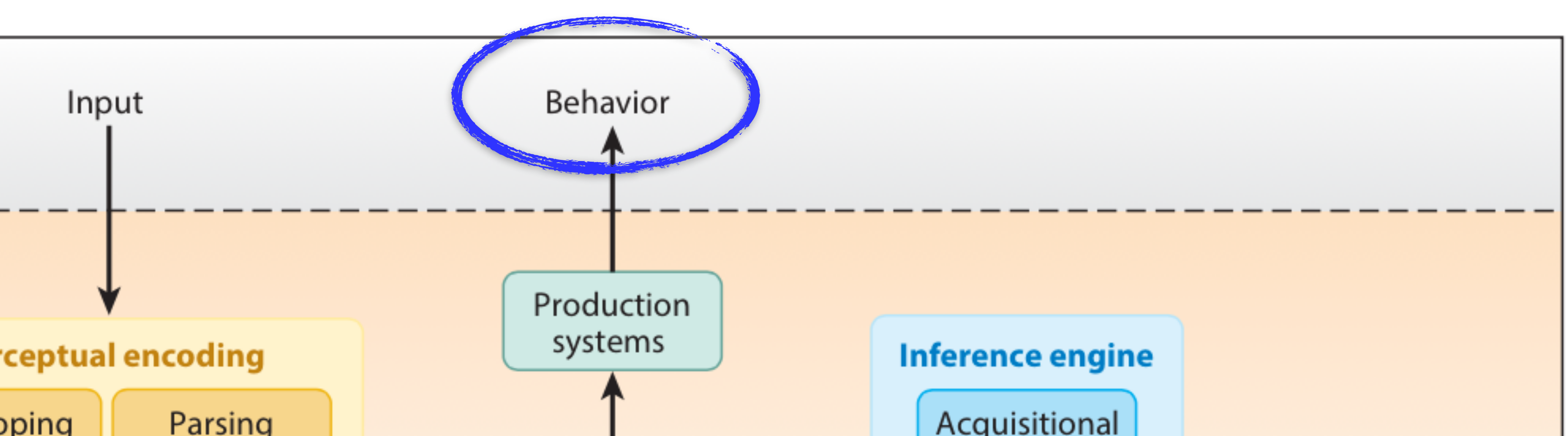
wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi

**(6) Decide what the measure of success is**

This can be based on your theory
or empirical data about behavior

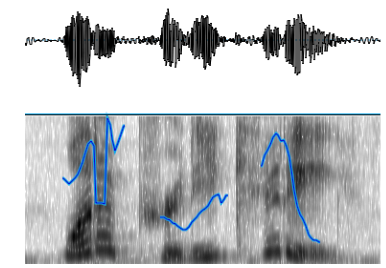


# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this . ▸
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool . ▸
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that . ▸
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do . ▸
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi → 

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi

### (6) Decide what the measure of success is

Example developing knowledge

Proto-lexicon of word forms

| | |
|---|---|
| wˈʌɾ | *what* |
| ə | *a* |
| pɹˈɪɾi | *pretty* |
| kˈɪɾi | *kitty* |

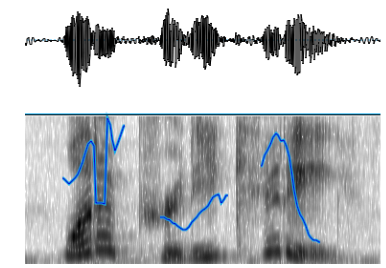
This can be based on your theory or empirical data about behavior

# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:    eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

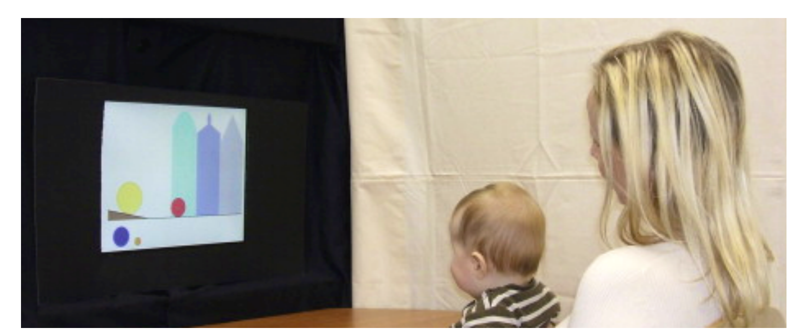
wˈʌɾə
pɹˈɪɾikˈɪɾi

| | |
|---|---|
| wˈʌɾ | *what* |
| ə | *a* |
| pɹˈɪɾi | *pretty* |
| kˈɪɾi | *kitty* |

**(6) Decide what the measure of success is**

This can be based on your theory
or empirical data about behavior

Example behavior indicating developed knowledge: Recognizing useful units (such as words) in a fluent speech stream, as indicated by looking time behavior

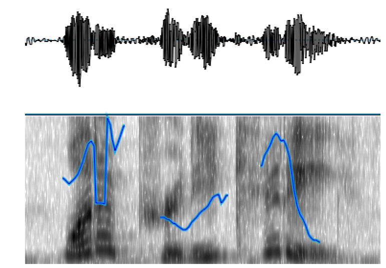


# How do we model language acquisition?

## An example with speech segmentation


```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   pro|you aux|have~neg|not part|see&PASTP pro:dem|this .
%gra:   1|4|SUBJ 2|4|AUX 3|2|NEG 4|0|ROOT 5|4|OBJ 6|4|PUNCT
*MOT:   that looks pretty cool .
%mor:   det|that n|look-PL adv:int|pretty adj|cool .
%gra:   1|2|DET 2|0|INCROOT 3|4|JCT 4|2|XMOD 5|2|PUNCT
*MOT:   do you know how to work that .
%mor:   mod|do pro|you v|know adv:wh|how inf|to v|work pro:dem|that .
%gra:   1|3|AUX 2|3|SUBJ 3|0|ROOT 4|3|OBJ 5|6|INF 6|4|XCOMP 7|6|OBJ 8|3|PUNCT
*MOT:   yes you do .
%mor:   co|yes pro|you v|do .
%gra:   1|3|COM 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
```

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi → 

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

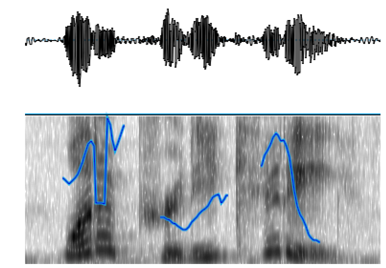
wˈʌɾə
pɹˈɪɾikˈɪɾi

**This is the heart of the model**

| | |
|---|---|
| wˈʌɾ | *what* |
| ə | *a* |
| pɹˈɪɾi | *pretty* |
| kˈɪɾi | *kitty* |

# How do we model language acquisition?

## An example with speech segmentation



= w'ʌ rə pɹ'ɪ ri k'ɪ ri


what a pretty kitty!

w'ʌrə

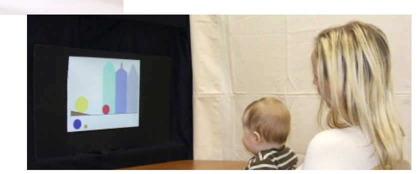
pɹ'ɪrik'ɪri

w'ʌ rə pɹ'ɪ ri k'ɪ ri

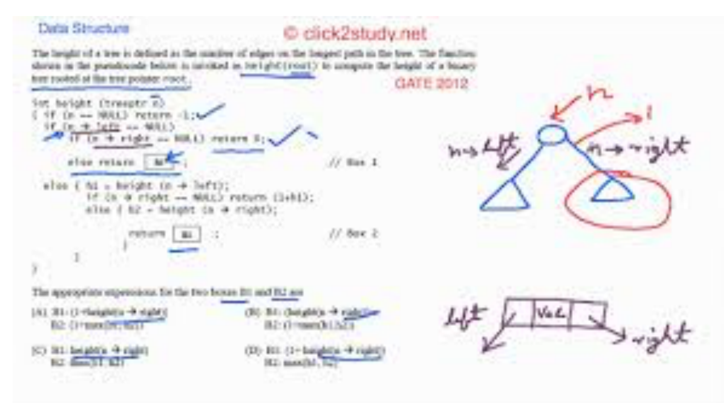
pɹ'ɪri

k'ɪri

what
a
pretty
kitty

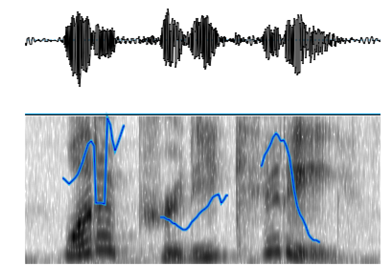


**(7) Implement the model in a programming language of choice**




# How do we model language acquisition?

## An example with speech segmentation



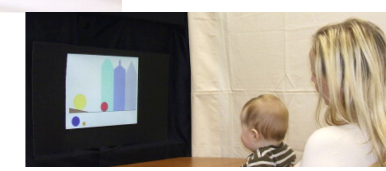
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi


wˈʌɾə

pɹˈɪɾikˈɪɾi



**(8) See how well the model did w.r.t. the measure of success**

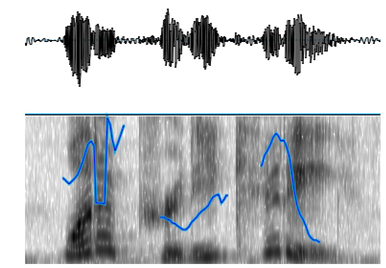
Example developing knowledge

Proto-lexicon of word forms

| | | |
|---|---|---|
| | wˈʌɾ | *what* |
| | ə | *a* |
| **???** | pɹˈɪɾi | *pretty* |
| | kˈɪɾi | *kitty* |

# How do we model language acquisition?

## An example with speech segmentation



= w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi

what a pretty kitty!

w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi


pɹ'ɪɾi pretty

k'ɪɾi kitty

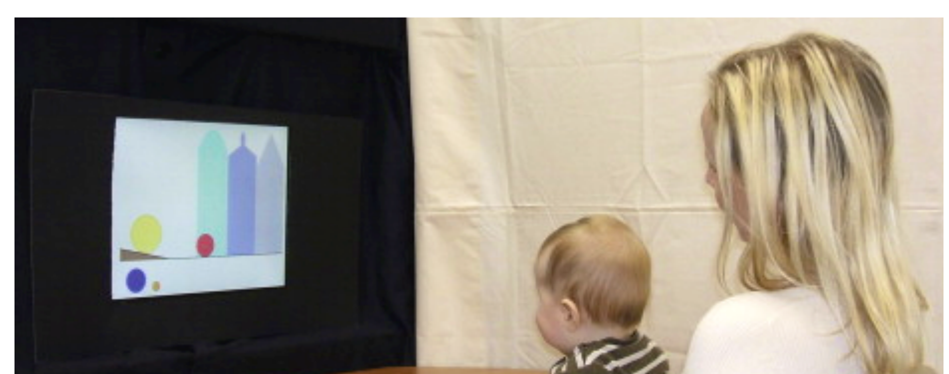
w'ʌɾə

pɹ'ɪɾik'ɪɾi

**(8) See how well the model did w.r.t. the measure of success**

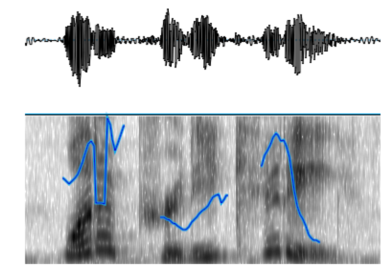
Recognizing useful units (such as words) in a fluent speech stream, as indicated by looking time behavior

**???**



# How do we model language acquisition?

## An example with speech segmentation

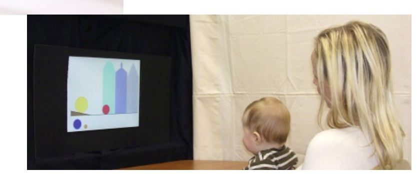


= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi


wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

???

**(8) See how well the model did w.r.t. the measure of success**

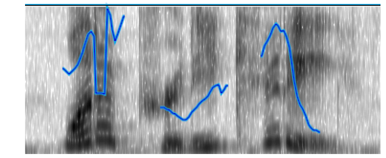


From this, we can determine how well the model did — and more importantly, how well the strategy implemented concretely in the model did.

# How do we model language acquisition?

## An example with speech segmentation

= w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi



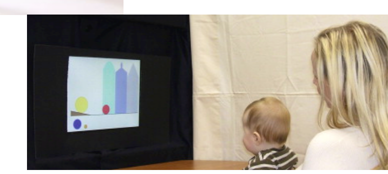
what a pretty kitty!


w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi

???

w'ʌɾə

pɹ'ɪɾik'ɪɾi



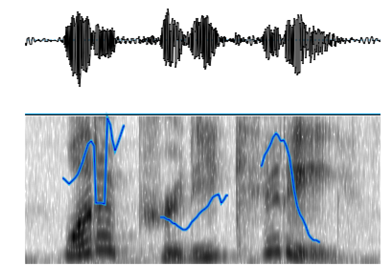
**(9) Interpret the results for other people who aren't you so they know why they should care**

"The modeled child has the same developing knowledge as we think 8-month-olds do. This strategy can be what they're using!"

| | |
|---|---|
| w'ʌɾ | *what* |
| ə | *a* |
| pɹ'ɪɾi | *pretty* |
| k'ɪɾi | *kitty* |

# How do we model language acquisition?

## An example with speech segmentation



= w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi

what a pretty kitty!

w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi

???

w'ʌɾə

pɹ'ɪɾi

k'ɪɾi

w'ʌ

ɾə

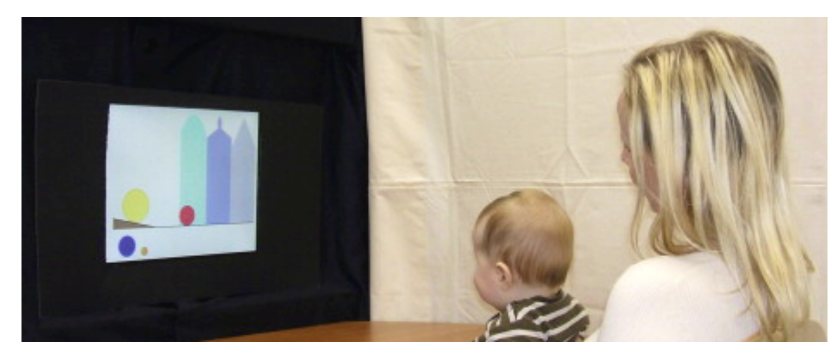
pɹ'ɪɾik'ɪɾi

w'ʌɾə

pɹ'ɪɾik'ɪɾi

| | |
|---|---|
| w'ʌɾ | what |
| ə | a |
| pɹ'ɪɾi | pretty |
| k'ɪɾi | kitty |

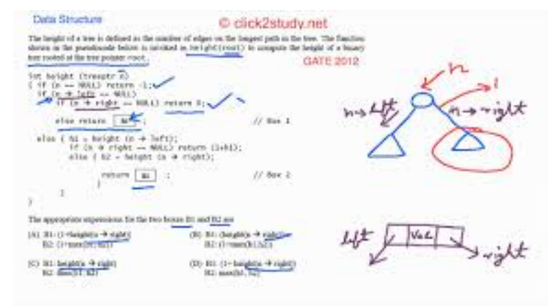
**(9) Interpret the results for other people who aren't you so they know why they should care**

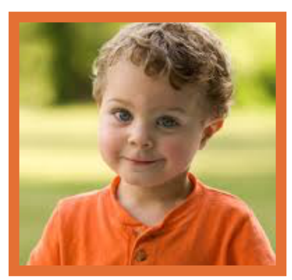


"The modeled child can reproduce the behavior we see in 8-month-olds. This strategy could be what they're using to generate that behavior!"

# Today's Plan:
# Computational models of language acquisition

## II. How



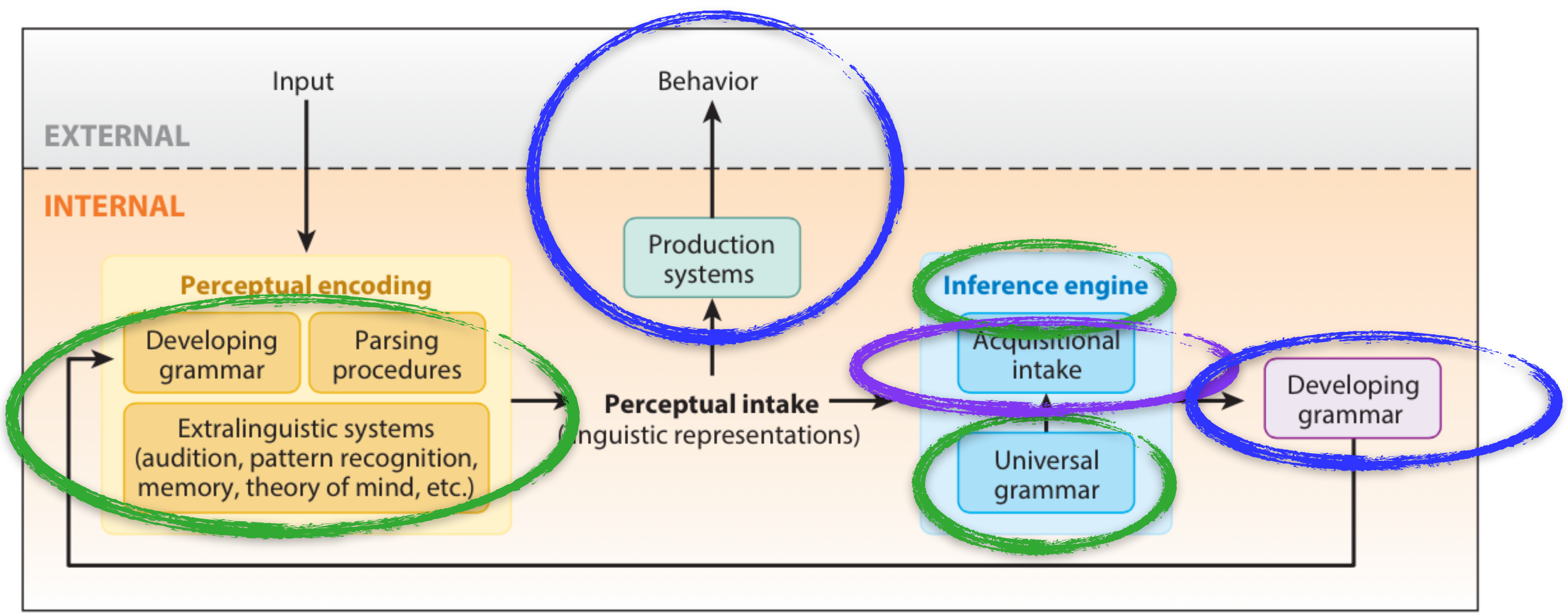



Levels of explanation

# How do we model language acquisition?

**What level of model do you want to build?**

A very basic question:
Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state?




Lidz & Gagliardi 2015

Computational-level (Marr 1982)

**Is this the right conceptualization of the acquisition task? Do we have the right goal in mind?**

# How do we model language acquisition?

**What level of model do you want to build?**

**Computational-level**
A very basic question:
Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state?

Helpful for determining **if this implementation of the acquisition task is the right one**.

Are these useful learning assumptions for children to have? Are these useful linguistic representations?

# How do we model language acquisition?

**What level of model do you want to build?**

**Computational-level**

A very basic question:

Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state?

This is typically implemented as an ideal learner model, which isn't concerned with the cognitive limitations and incremental learning restrictions children have.

(That is, useful for children is different from useable by children in real life.)

# How do we model language acquisition?

**What level of model do you want to build?**

**Computational-level**
A very basic question:
Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state?

Practical note:
Doing a computational-level analysis is often a really good idea to make sure we've got the right conceptualization of the acquisition task (see Pearl 2011 for the trouble you can get into when you don't do this first).

# How do we model language acquisition?

**What level of model do you want to build?**

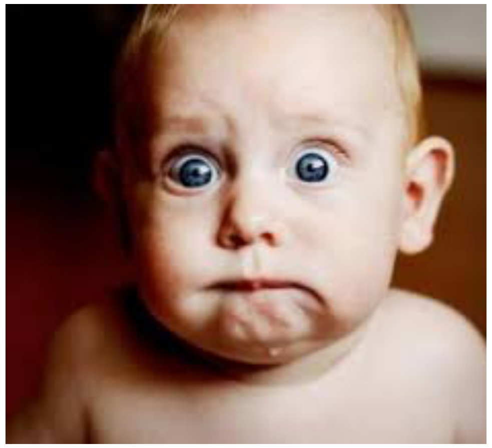
**Computational-level**

A very basic question:

Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state?

*(What happened in a nutshell in Pearl 2011)*

Why do none of these learning strategies work?



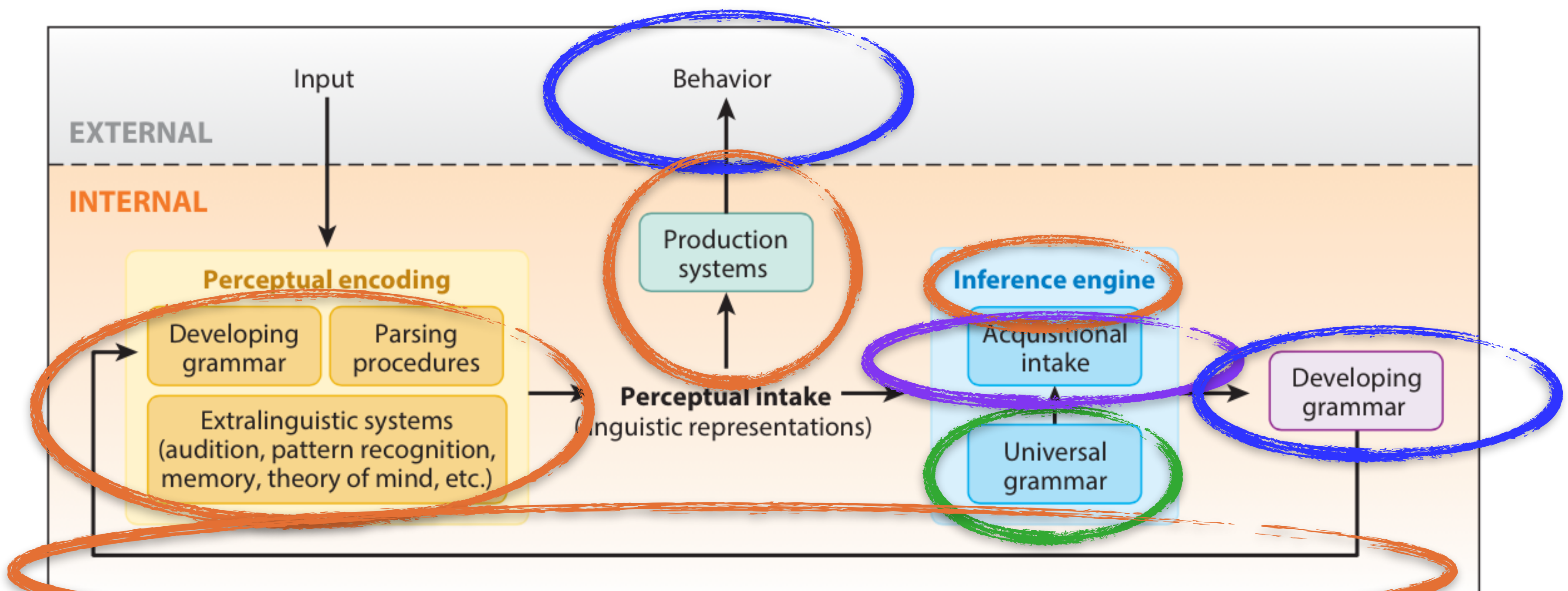
**Because they're solving the wrong acquisition task…oops.**

# How do we model language acquisition?

**What level of model do you want to build?**

Computational-level

Another basic question:
Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state in the amount of time children typically get to do it, given the incremental nature of learning and children's cognitive constraints?




Lidz & Gagliardi 2015

# How do we model language acquisition?

**What level of model do you want to build?**

**Computational-level**

Another basic question:

Is it possible for the child with a specific initial state to use the acquisitional intake to achieve the target state in the amount of time children typically get to do it, given the incremental nature of learning and children's cognitive constraints?

**Algorithmic-level** (Marr 1982)

**Is it possible for children to use this strategy? That is, once we know it's useful for children, it's important to make sure it's also useable by children.**

# How do we model language acquisition?

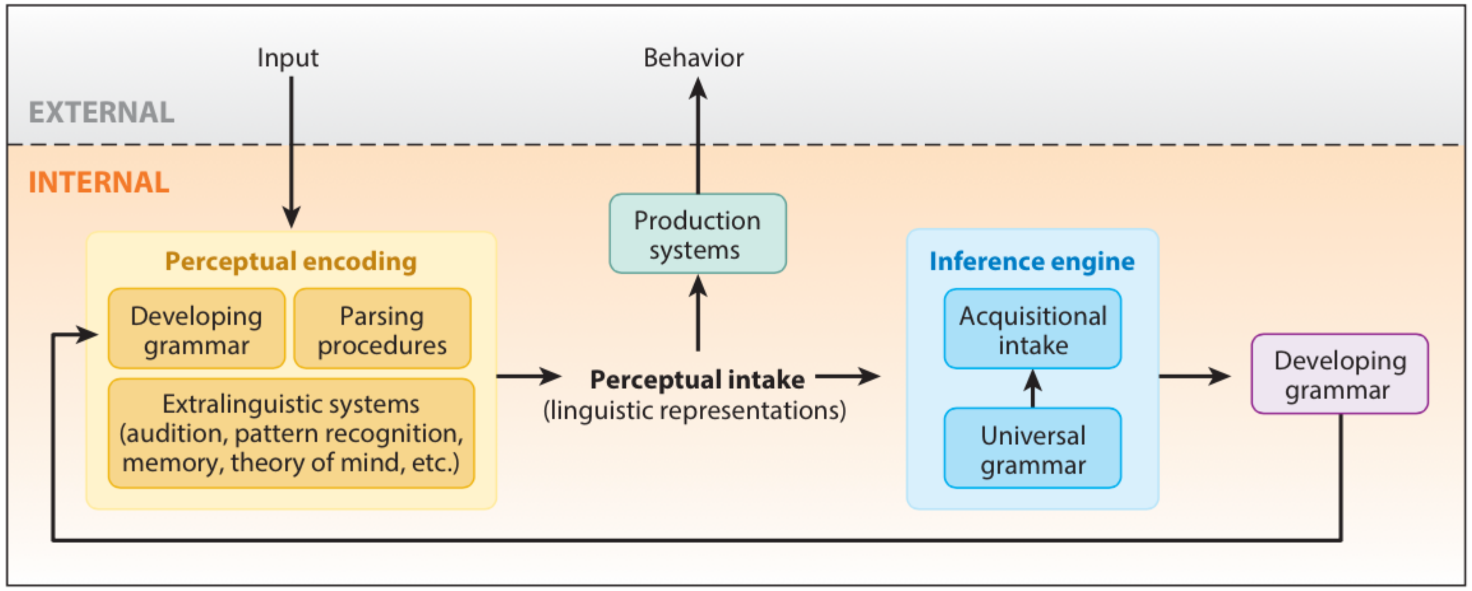
What level of model do you want to build?

Computational-level

Algorithmic-level

Another important (not so basic) question: If we have an algorithm that seems useable by children to usefully solve an acquisition task, how is it implemented in the brain?

## Implementational-level




Lidz & Gagliardi 2015

# How do we model language acquisition?

**What level of model do you want to build?**

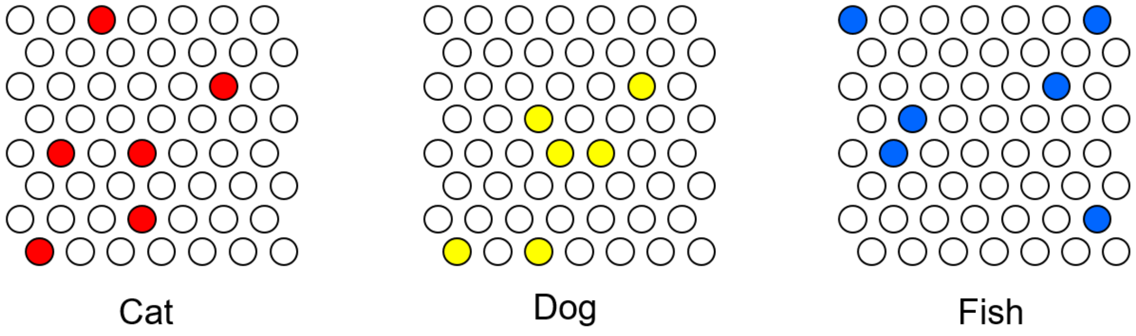
**Computational-level**

Another important (not so basic) question: If we have an algorithm that seems useable by children to usefully solve an acquisition task, how is it implemented in the brain?

**Algorithmic-level**

**Implementational-level**

*This isn't easy to model yet.*
*Advances in natural language processing: ways to encode complex information into distributed representations like what we think the brain uses.*




(Rashkin et al. 2016, Levy & Goldberg 2014, Iyyer et al 2014)

# How do we model language acquisition?

What level of model do you want to build?

The types I generally work with

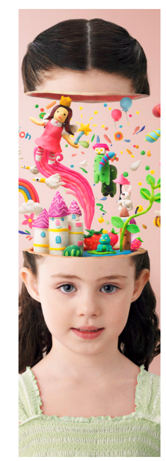
Computational-level

Algorithmic-level

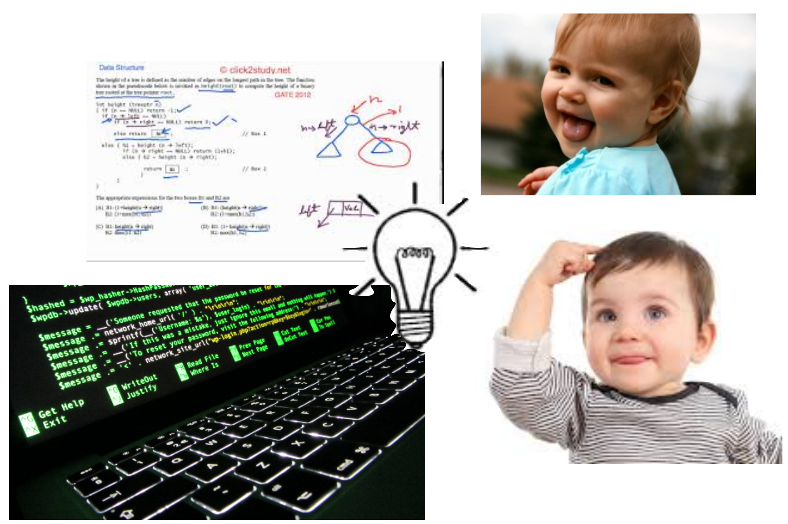
Implementational-level

# Today's Plan:
# Computational models of language acquisition

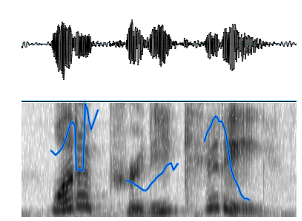
I. Why



II. How



III. What we can learn



Noun

✓KI tty

*Who does… is pretty?*

*another one*

*Every kitty didn't …*

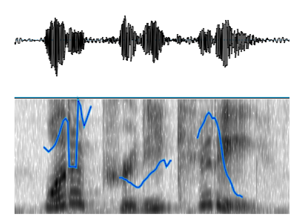
# Today's Plan:
# Computational models of language acquisition

## III. What we can learn

metrical phonology

syntactic categorization

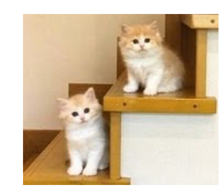
speech segmentation



Noun

syntax

✓KI tty



*Who does… is pretty?*

*another one*

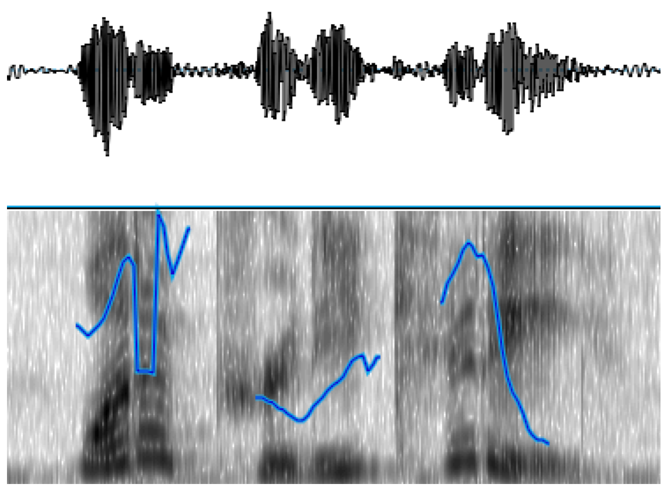
*Every kitty didn't …*

pragmatics

syntax, semantics

# What we can learn

speech segmentation



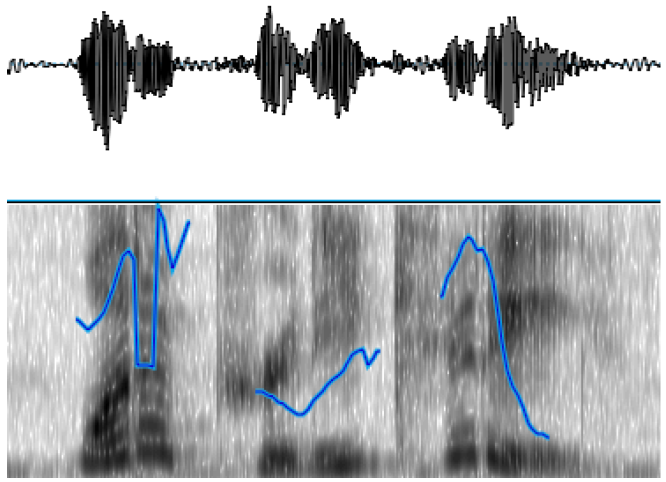
= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

# What we can learn

speech segmentation

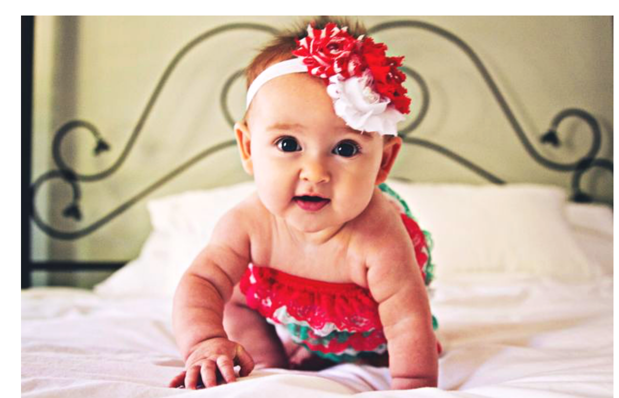


= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

**Investigating a Bayesian inference strategy for the very early stages of speech segmentation occurring around six months**

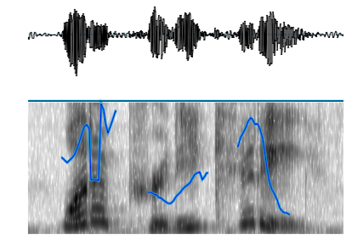


Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

$$P(s|u) \propto P(s)P(u|s)$$

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

**Strategy: Identify a proto-lexicon of words that best generates the observable fluent speech utterances**

Mathematically encoded preferences:

wʌɾə
pɹɪɾi
kɪɾi

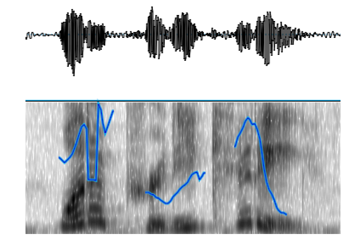
wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

**Strategy: Identify a proto-lexicon of words that best generates the observable fluent speech utterances**

Mathematically encoded preferences:

(1) Prefer shorter words

wʌɾə
pɹɪɾi
kɪɾi

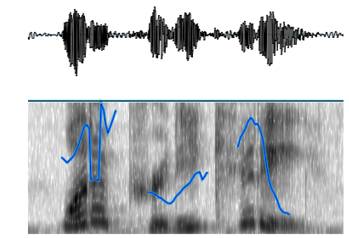
wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

**Bayesian inference**

$$P(s|u) \propto P(s)P(u|s)$$

**Strategy: Identify a proto-lexicon of words that best generates the observable fluent speech utterances**

Mathematically encoded preferences:

(1) Prefer shorter words

(2) Prefer lexicons with fewer words

wʌɾə
pɹɪɾi
kɪɾi

wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

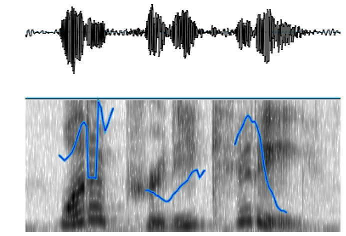
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

**Strategy: Identify a proto-lexicon of words that best generates the observable fluent speech utterances**

Mathematically encoded preferences:

(1) Prefer shorter words

(2) Prefer lexicons with fewer words

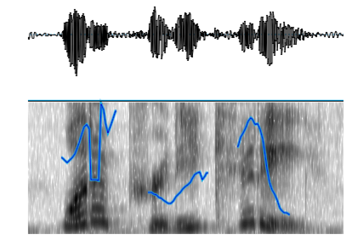
| | | |
|---|---|---|
| wʌɾə | wʌ | wʌɾə |
| pɹɪɾi | ɾə | pɹɪɾikɪɾi |
| kɪɾi | pɹɪɾikɪɾi | |

**Find the best segmentation**

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

**Strategy: Identify a proto-lexicon of words that best generates the observable fluent speech utterances**

Mathematically encoded preferences:

(1) Prefer shorter words

(2) Prefer lexicons with fewer words

wʌɾə
pɹɪɾi
kɪɾi

wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

**Find the best segmentation that balances these proto-lexicon preferences**

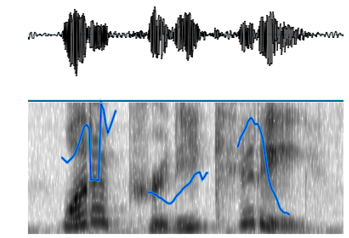
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

**Strategy: Identify a proto-lexicon of words that best generates the observable fluent speech utterances**

Mathematically encoded preferences:

(1) Prefer shorter words

(2) Prefer lexicons with fewer words

wʌɾə
pɹɪɾi
kɪɾi

wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

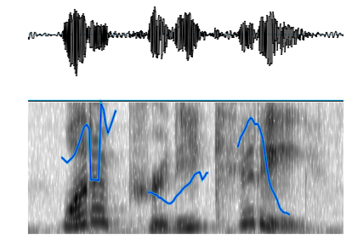
**Find the best segmentation that balances these proto-lexicon preferences**

**and can generate the observable fluent speech utterances**

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

**Is it useful?**

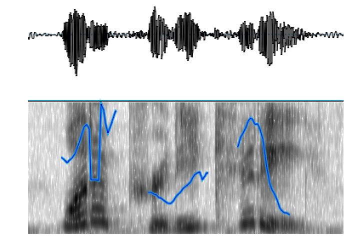
Computational-level modeled learners using this strategy segment fairly well, given realistic English child-directed speech data.

The inferred proto-lexicons, while not perfect, are very useful for subsequent stages of language acquisition.

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

✓ Is it **useful**?

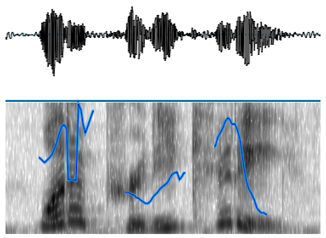
✓ Is it **useable**?

Algorithmic-level modeled learners with cognitive constraints on their inference and memory can still use this strategy and segment English quite well.

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

✓ Is it **useful**?

✓ Is it **useable**?

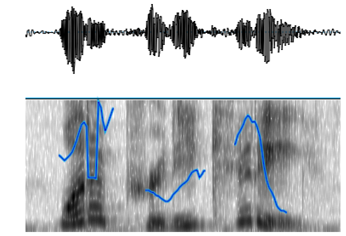
✓ Does it work for **different languages**?

It segments well for languages with different morphology and syllable properties: Spanish, Italian, German, Hungarian, Japanese, Farsi

Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

## Bayesian inference

$$P(s|u) \propto P(s)P(u|s)$$

Is it **useful**?

Is it **useable**?

Does it work for **different languages**?

**This kind of Bayesian inference seems to be a good proposal for a very early speech segmentation strategy.**

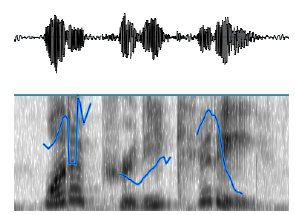
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

# What we can learn

metrical phonology

syntactic categorization

speech segmentation



Noun

syntax

KI tty

*Who does … is pretty?*

*another one*

*Every kitty didn't …*

pragmatics

syntax, semantics

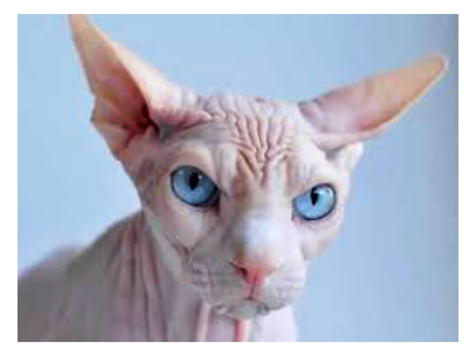
# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Look — there's another one!"



# What we can learn

**syntax, semantics**

*another one*

“Oh look — a pretty kitty!”

“Look — there’s another one!”

Interpretation: another pretty kitty

**same**

**syntactic category**

**???**

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Look — there's another one!"

Interpretation: another

same
syntactic category
???

bigger than a plain **Noun**

**Noun**
|
pretty **kitty**

# What we can learn

syntax, **semantics**

*another one*

"Oh look — a pretty kitty!"

"Look — there's another one!"

Interpretation: another ~~the pretty kitty~~

same
syntactic category
???

smaller than a full **Noun Phrase**

**Noun Phrase**

*the*

Noun

pretty kitty

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Look — there's another one!"

Interpretation: another

same
syntactic category
???

In-between category **Noun'**
that includes strings with nouns
and modifiers+nouns

Noun Phrase

*the* **Noun'**

**Noun'**

Noun

**pretty kitty**

# What we can learn

syntax, **semantics**

*another one*

"Oh look — a pretty kitty!"

"Look — there's another one!"

Interpretation: another

same
syntactic category

Noun Phrase

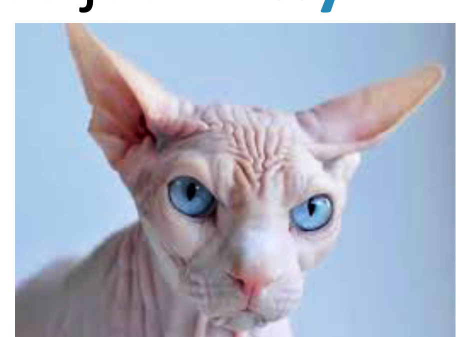
*the* Noun'

**Noun'**

Noun

pretty **kitty**

This is why we can also interpret one as just **kitty.**



# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Do you see another one?"

Lidz, Waxman, & Freedman 2003:
18-month-old interpretations

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Do you see  another one?"

**pretty kitty**

**Noun'**

Lidz, Waxman, & Freedman 2003:
18-month-old interpretations

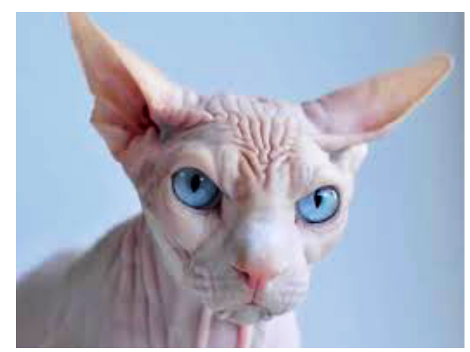
# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

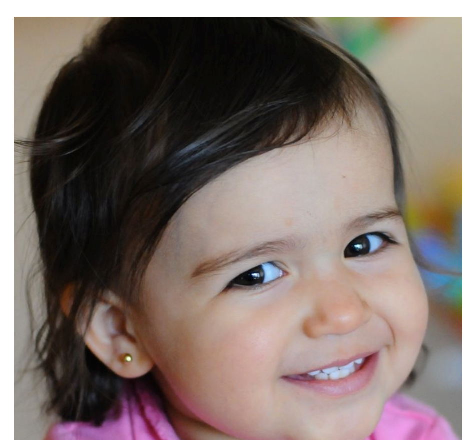
"Do you see another kitty?"



another one

**pretty kitty**

**Noun'**



Lidz, Waxman, & Freedman 2003:
18-month-old interpretations

# What we can learn

**syntax, semantics**

*another one*

"Oh look — a pretty kitty!"

"Do you see another kitty?"

another one

**pretty kitty**

**Noun'**

Lidz, Waxman, & Freedman 2003:
18-month-old interpretations

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Do you see another pretty kitty?"

another one

**pretty kitty**

**Noun'**

Lidz, Waxman, & Freedman 2003:
18-month-old interpretations

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

"Do you see another pretty kitty?"

another one

**pretty kitty**

**Noun'**

Lidz, Waxman, & Freedman 2003:
18-month-old interpretations

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

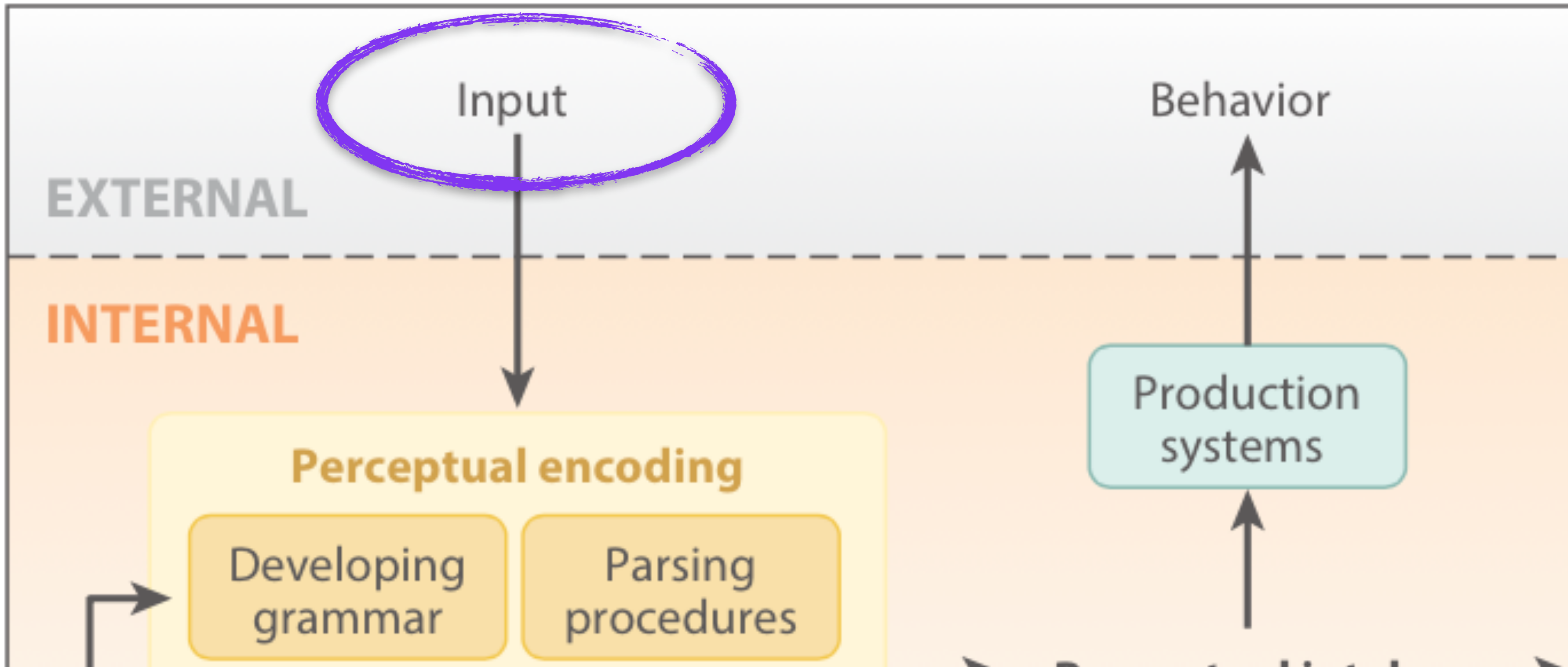
Noun'

pretty kitty

"Do you see another one ?"

Several learning strategies implemented with **algorithmic-level** modeled learners, given realistic samples of English child-directed speech.

Pearl & Mis 2016



# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

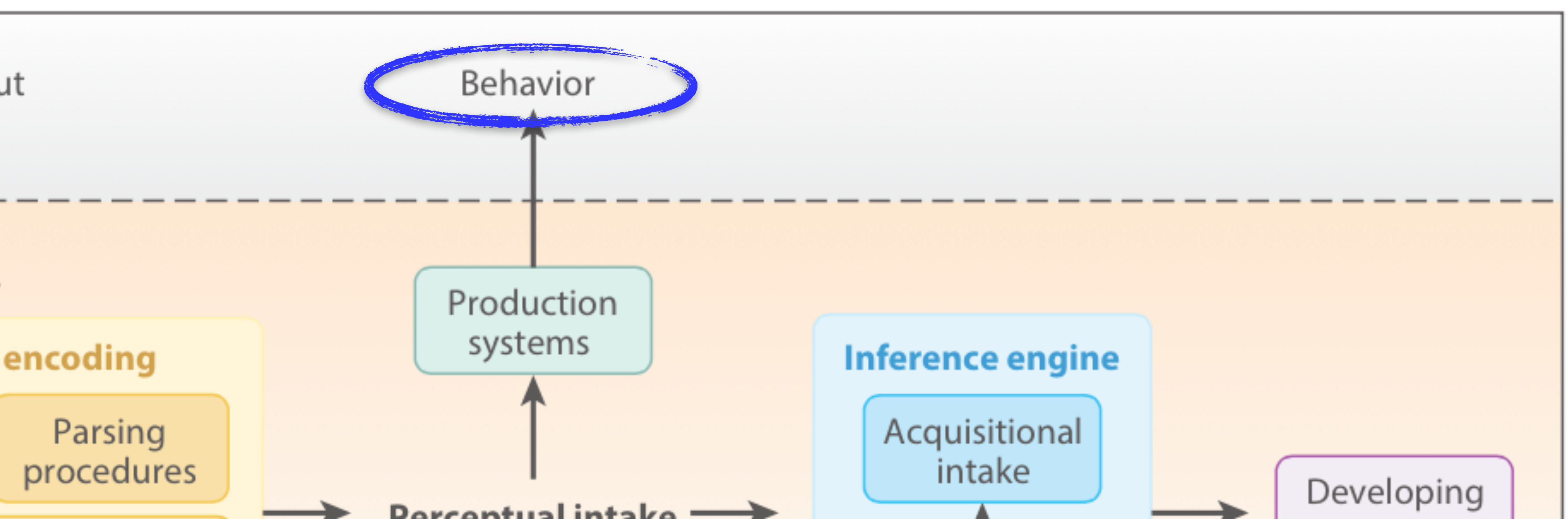
Noun'

pretty kitty

"Do you see another one ?"

**Algorithmic-level**

Evaluated on whether they matched 18-month-old looking preferences.




Pearl & Mis 2016

# What we can learn

syntax, semantics

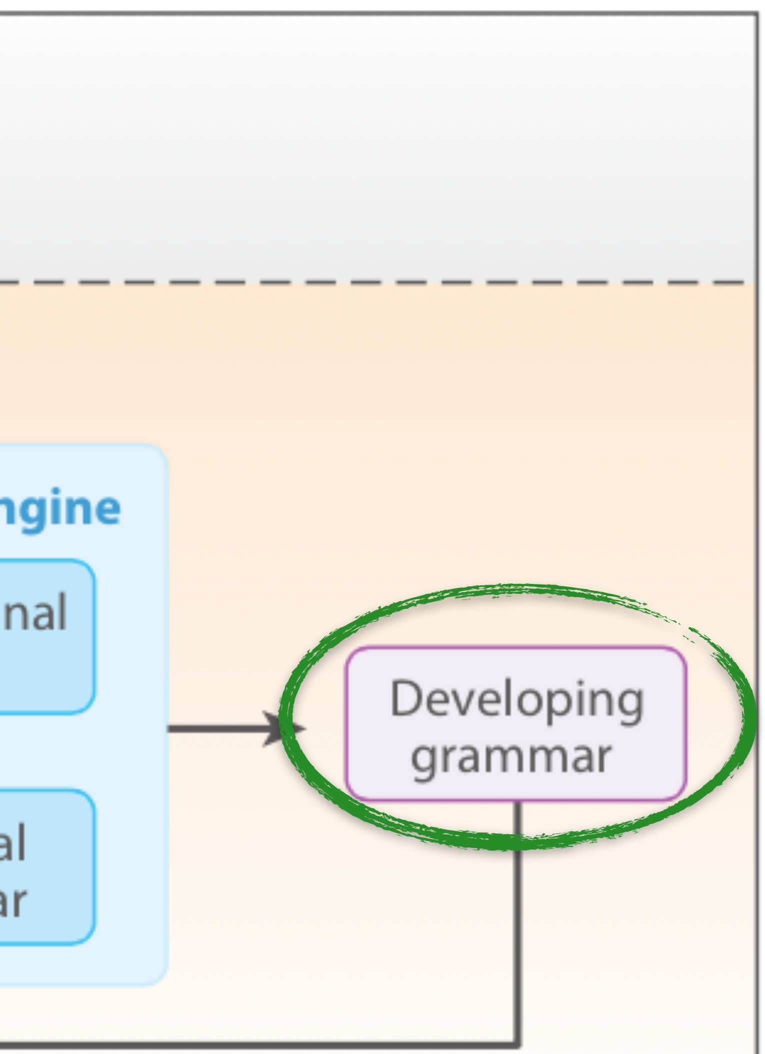
*another one*

"Oh look — a pretty kitty!"

Noun'

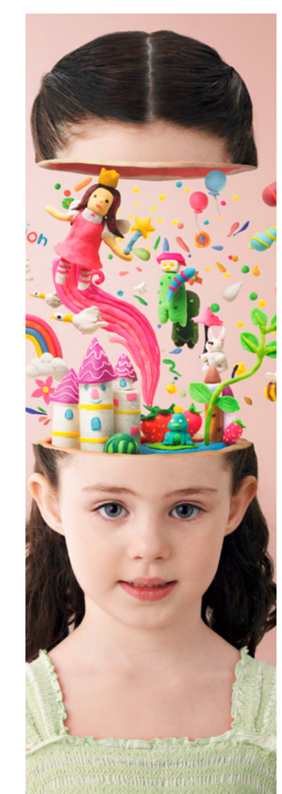
pretty kitty

"Do you see another one ?"




**Algorithmic-level**

Two strategies were successful at generating the 18-month-old behavior. We can then look inside the modeled learner and see what the underlying representations were.



Pearl & Mis 2016

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

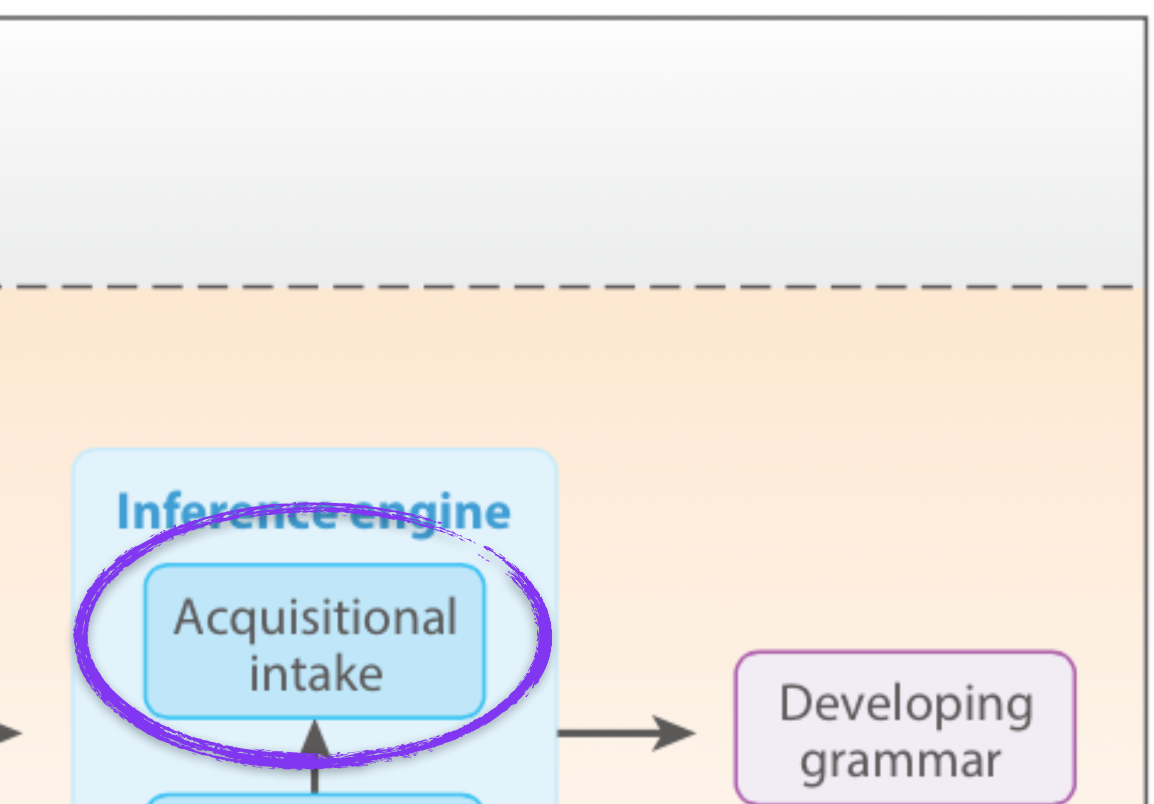
**Algorithmic-level**

Noun'

pretty kitty

"Do you see another one ?"

Strategy 1: Ignore some of the available one data in the input




Pearl & Mis 2016

# What we can learn

syntax, semantics

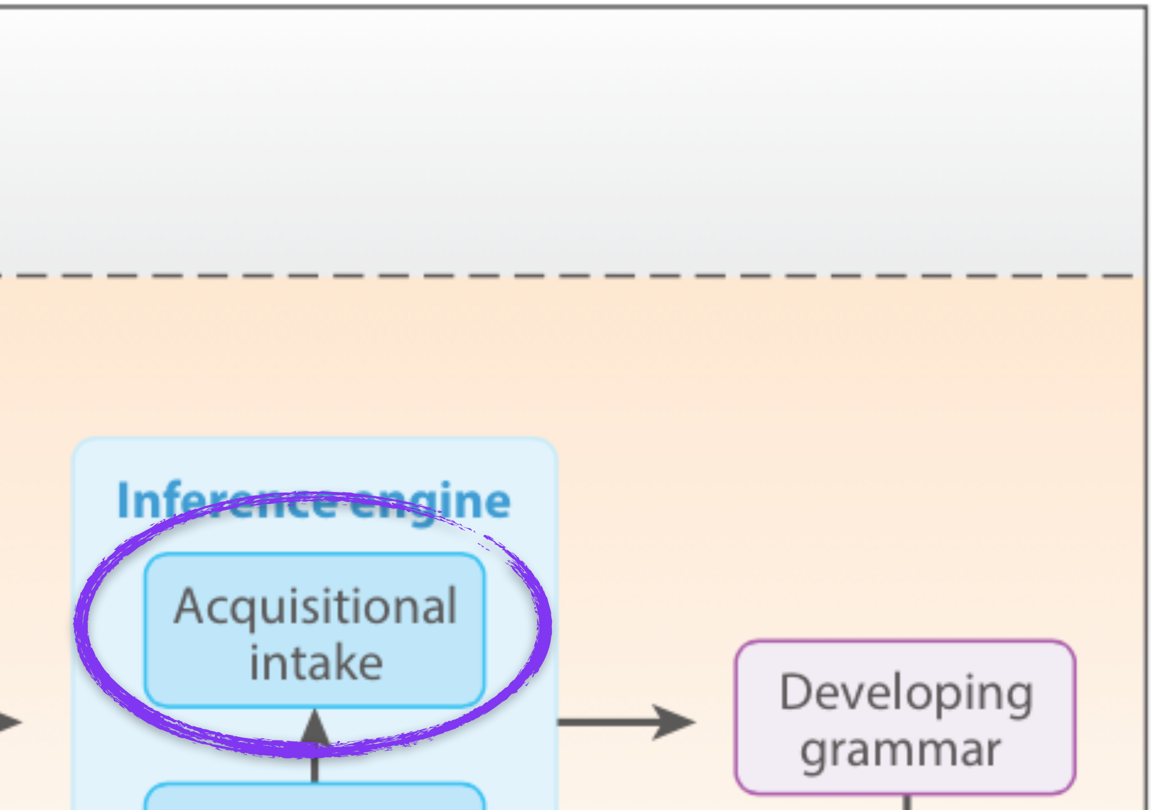
*another one*

"Oh look — a pretty kitty!"

"Do you see another one ?"

Algorithmic-level

Strategy 1: Ignore some of the available one data in the input




Adult representations

✓ Noun'
pretty kitty

But…required additional situational context to be present to succeed.

Less robust

Pearl & Mis 2016

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

Noun'

pretty kitty

"Do you see another one ?"

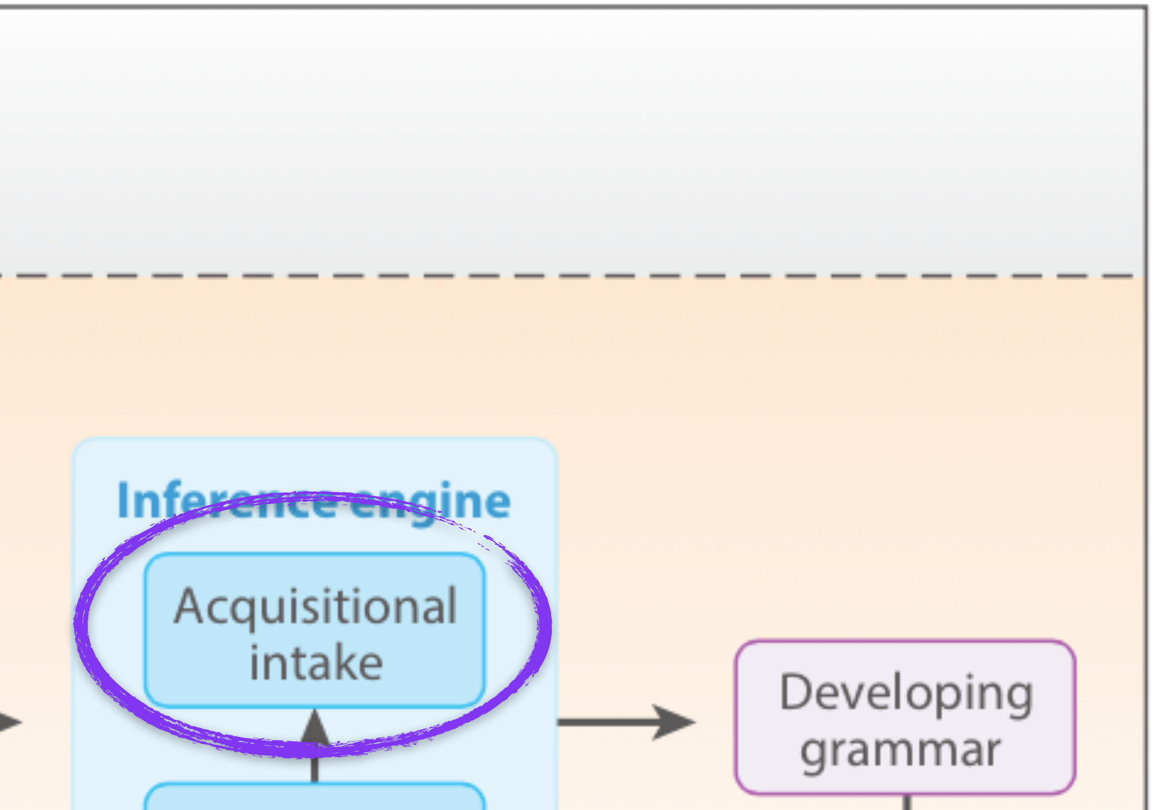
Algorithmic-level

✓

Strategy 1: Ignore

Less robust

Strategy 2: Include other pronoun data besides one data in the intake




Pearl & Mis 2016

# What we can learn

**syntax, semantics**

*another one*

"Oh look — a pretty kitty!"

"Do you see another one ?"

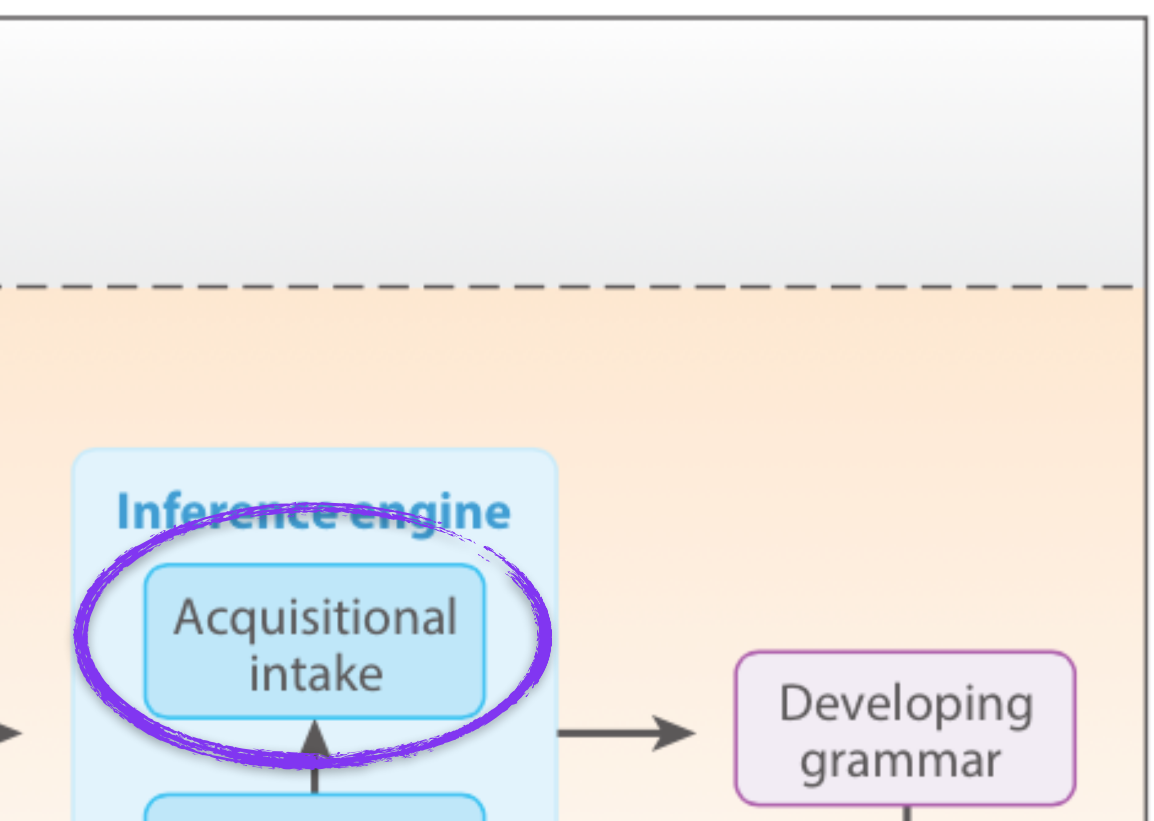
**Algorithmic-level**

✓

Strategy 1: Ignore

**Less robust**

Strategy 2: Include other pronoun data besides one data in the intake




Immature representations

**Noun'** only in certain linguistic contexts

**pretty kitty** ✓

✗ otherwise **Noun**

But…does this for pretty much any situational context.

**More robust**

Pearl & Mis 2016

# What we can learn

syntax, semantics

*another one*

"Oh look — a pretty kitty!"

Noun'

pretty kitty

"Do you see another one ?"

Algorithmic-level

Strategy 1: Ignore ✓

Less robust

Strategy 2: Include other ✓✗

More robust

**By modeling, we have two concrete proposals for how children learn the knowledge they do by 18 months.**

This also motivates future experimental work to distinguish these two possibilities.

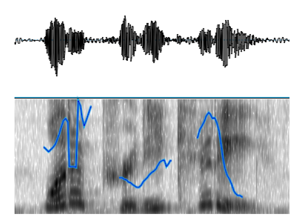
Pearl & Mis 2016

# What we can learn

metrical phonology

syntactic categorization

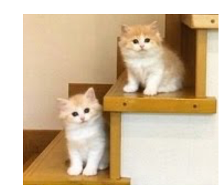
speech segmentation



Noun

syntax

KI tty



*Who does … is pretty?* *another one*

*Every kitty didn't …*

pragmatics

syntax, semantics

# What we can learn

syntax

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

*Lily thinks this kitty is pretty.*

**What's going on here?**

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

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

# What we can learn

syntax

*Who does*

*Lily think the kitty for is pretty?*

**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 ___ is pretty?*

This dependency is not allowed in English.

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

# What we can learn

syntax

*Who does*

*Lily think the kitty for is pretty?*

**What's going on here?** syntactic island

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

*Jack is somewhat tricksy.*

*He claimed he bought something.*

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

# What we can learn

syntax

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

**What's going on here?** syntactic island

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

*What did Jack make the claim that he bought___ ?*

*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 we can learn

syntax

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

**What's going on here?** syntactic island

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

*What did Jack make the claim that he bought___ ?*

*What did Elizabeth wonder whether Jack bought ___?*

*What did Elizabeth worry if Jack bought ___ ?*

*Jack is somewhat tricksy.*

*He claimed he bought something.*

*Elizabeth worried it was something dangerous.*

# What we can learn

syntax

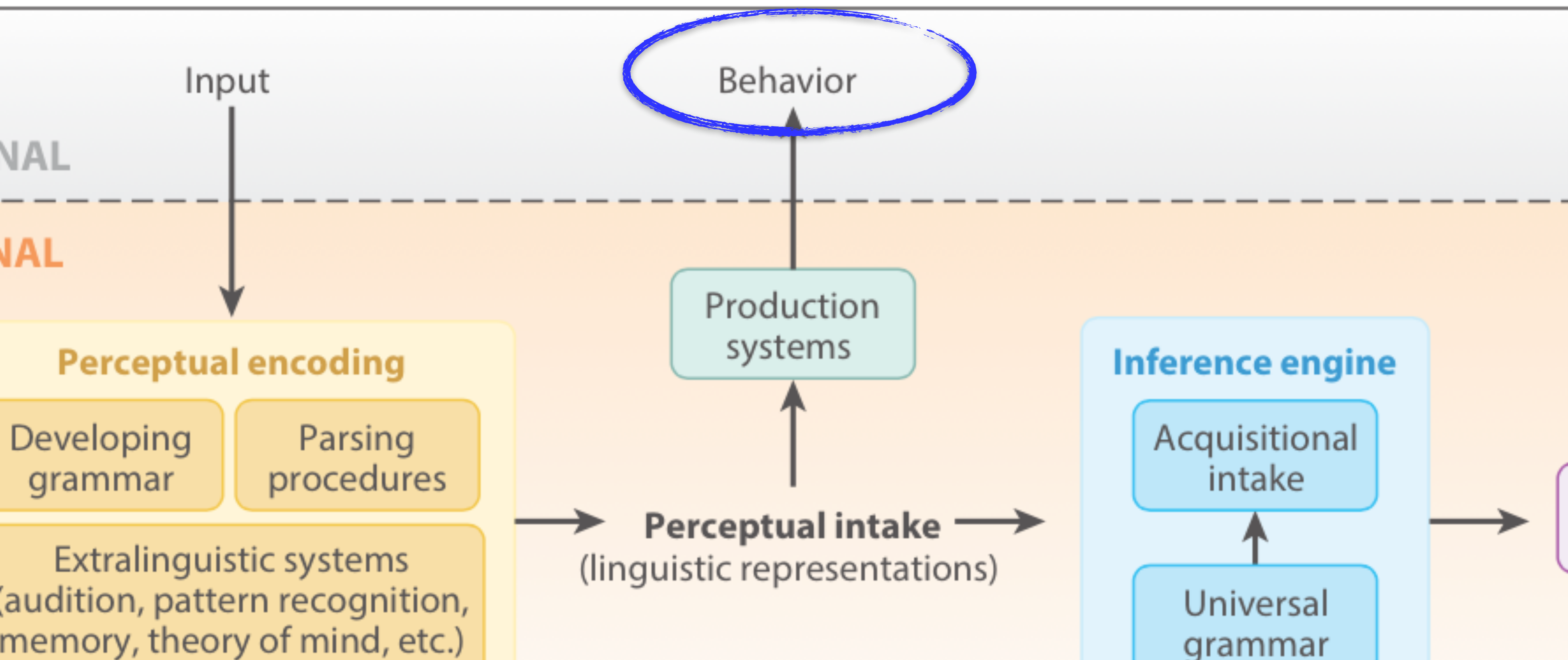
syntactic island

*Who does*

*Lily think the kitty for is pretty?*

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

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).



# What we can learn

syntax

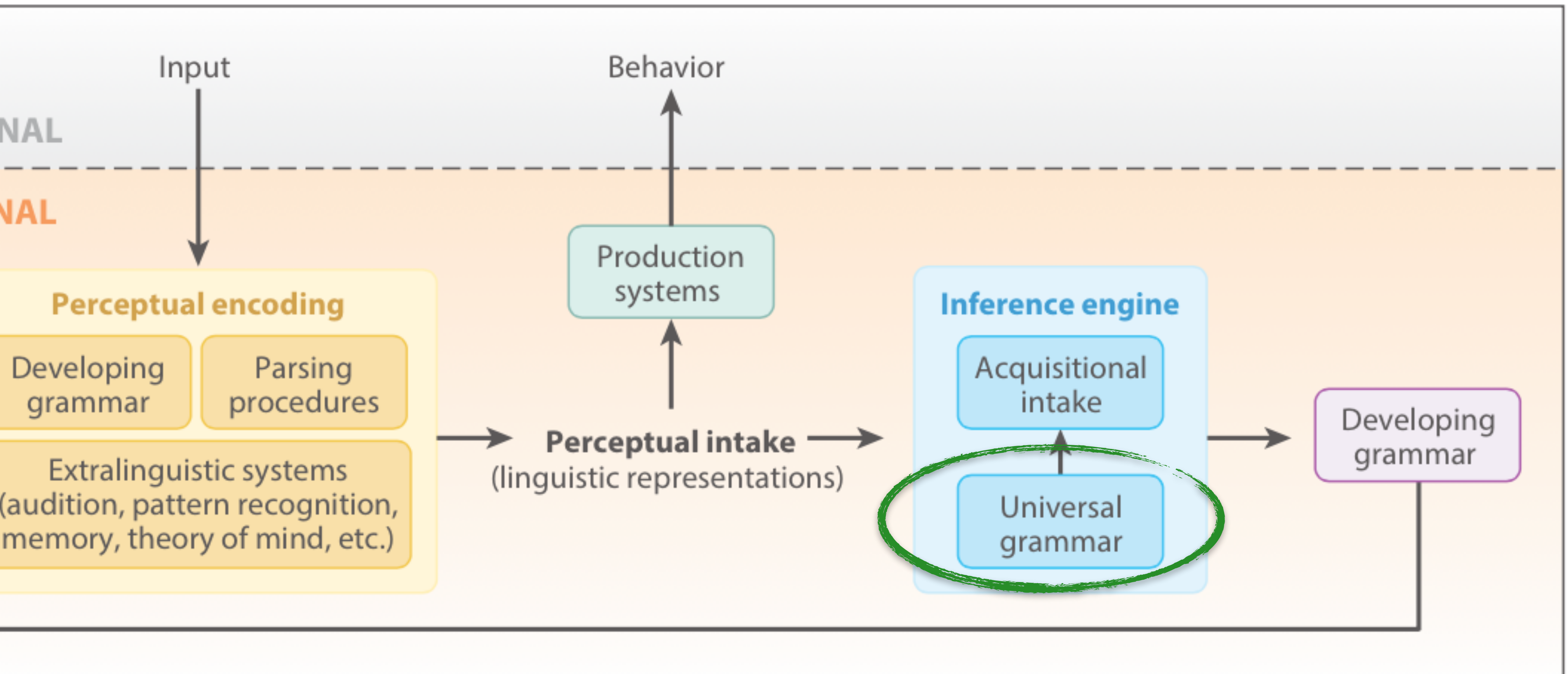
syntactic island

*Who does* ☹

*Lily think the kitty for is pretty?*

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

Previous learning theories suggested children need syntactic-island-specific innate knowledge.



# What we can learn

syntax

*Who does*

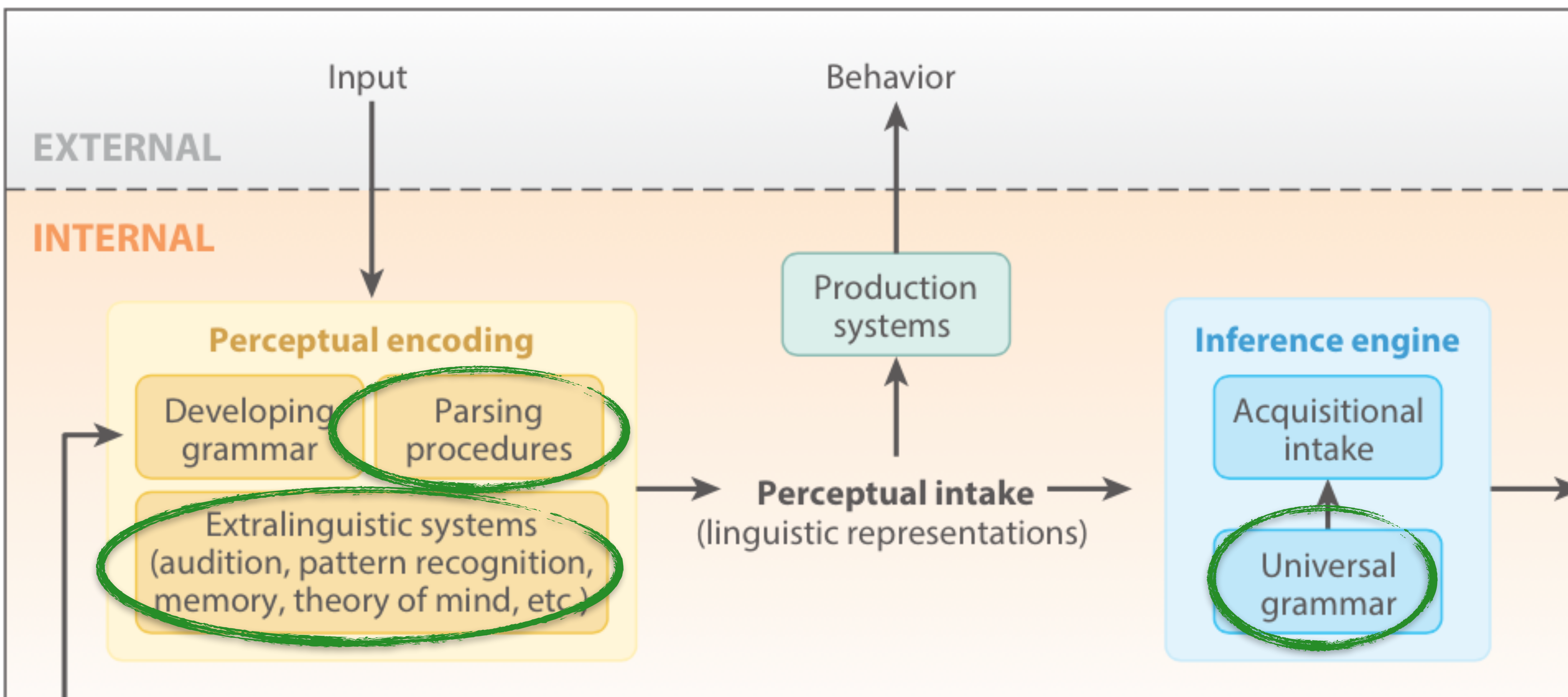
*Lily think the kitty for is pretty?*

syntactic island

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

An alternative learning strategy suggests children need less-specific linguistic prior knowledge along with probabilistic learning.

Pearl & Sprouse (2013a, 2013b, 2015)



# What we can learn

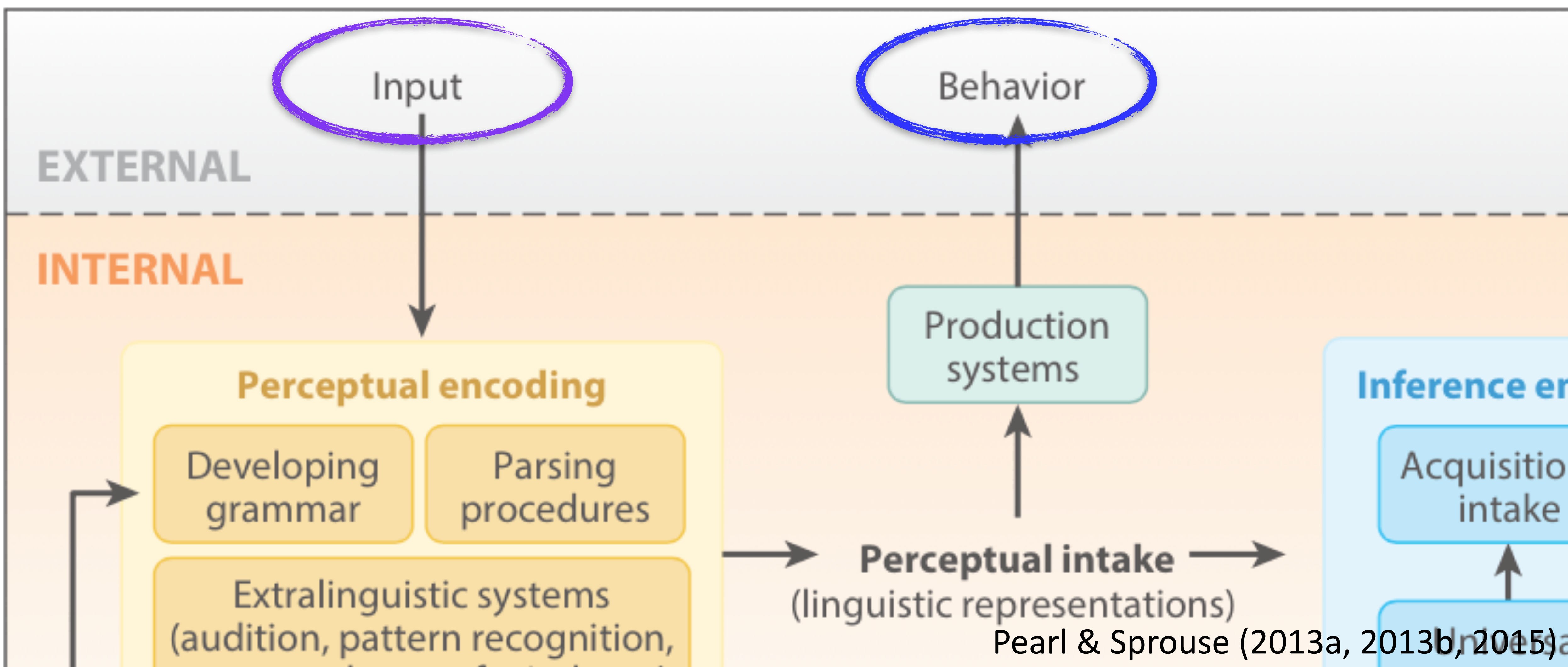
syntax

syntactic island

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

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

This alternative strategy was implemented in an algorithmic-level learning model that learned from realistic samples of child-directed speech. The modeled learner was able to reproduce the pattern of adult judgments.




Pearl & Sprouse (2013a, 2013b, 2015)

# What we can learn

syntax

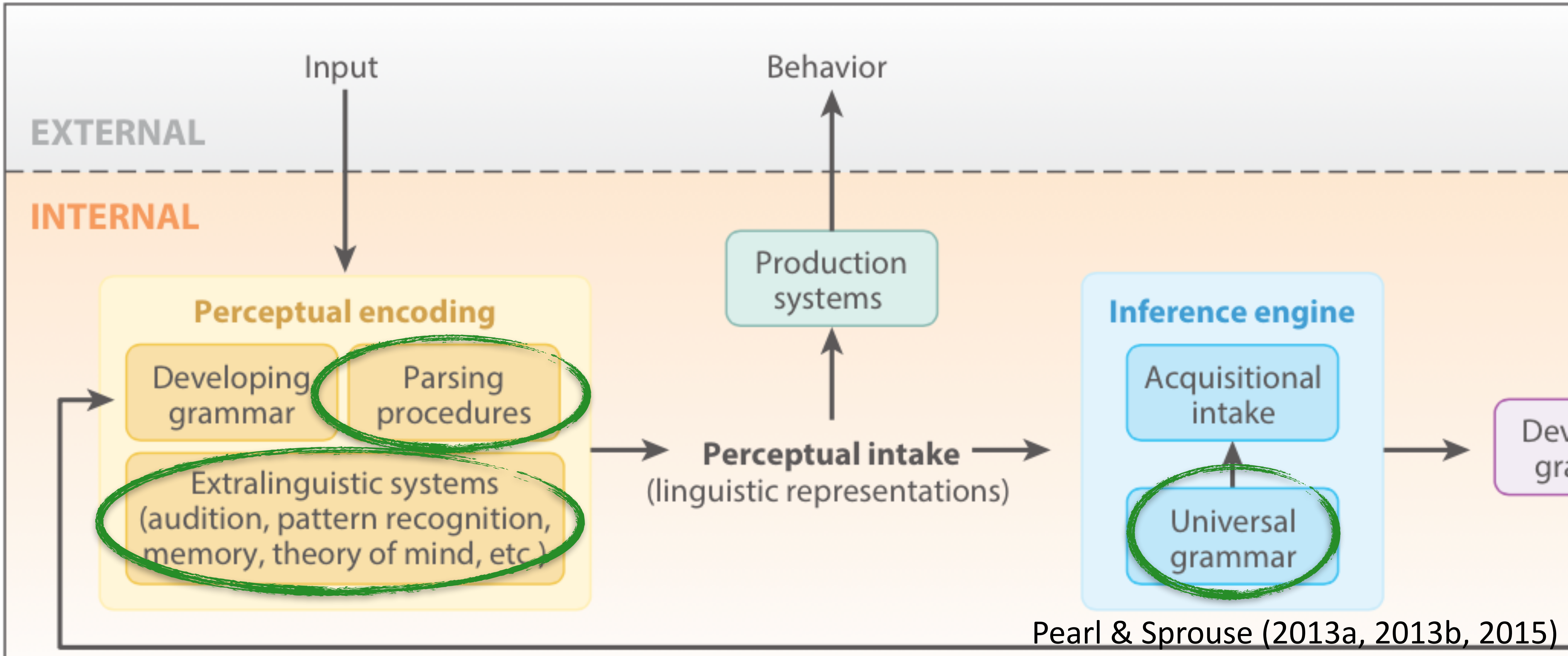
Who does 

Lily think the kitty for is pretty?

syntactic island

Who does Lily think the kitty for ___ is pretty?

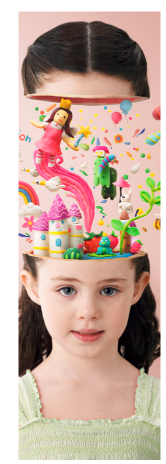
Upshot: Children can learn these sophisticated restrictions without relying as much on very specific linguistic knowledge that's necessarily innate.



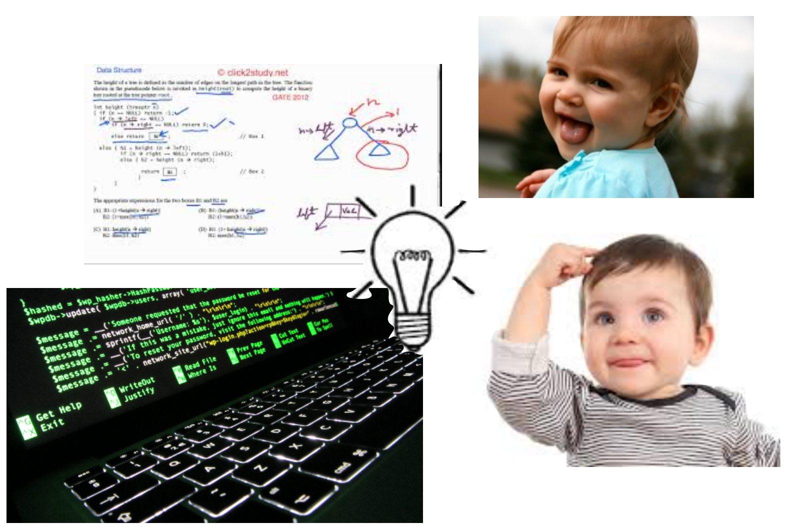

Pearl & Sprouse (2013a, 2013b, 2015)

# Today's Plan:
# Computational models of language acquisition

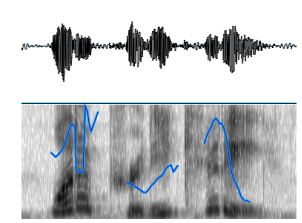
I. Why



II. How



III. What we can learn



Noun

✓KI tty

Who does… is pretty?

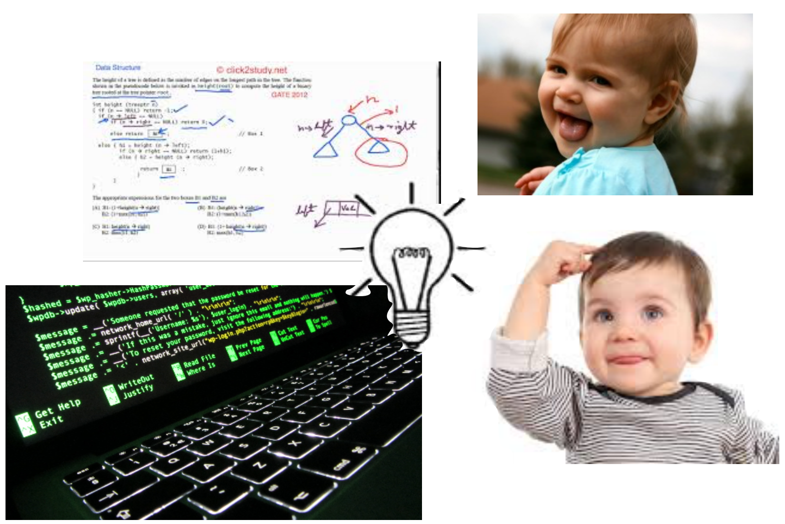
another one

Every kitty didn't …

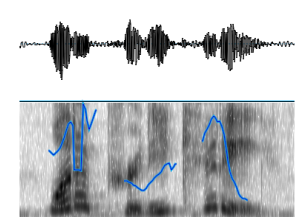
# Today's Plan:
# Computational models of language acquisition

**I. Why: Because language acquisition is pretty amazing and we want to understand how it works**

II. How



III. What we can learn



*Who does… is pretty?*

**Noun**

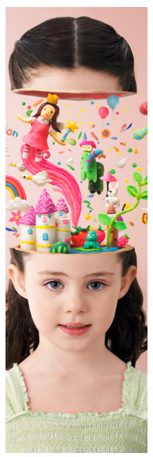
✓KI tty

*another one*

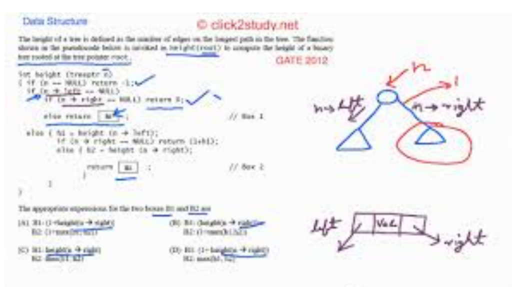
*Every kitty didn't …*

# Today's Plan:
# Computational models of language acquisition

I. Why: Because language acquisition is pretty amazing
and we want to understand how it works

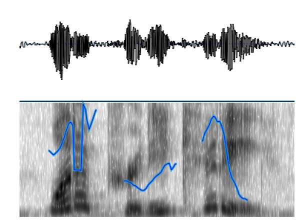



## II. How: By building informative computational models






III. What we can learn



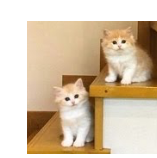

Noun

✓KI tty

*Who does… is pretty?*

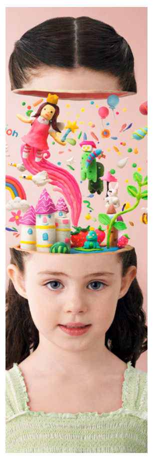
*another one*

*Every kitty didn't …*

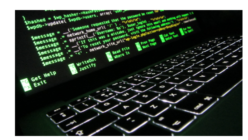
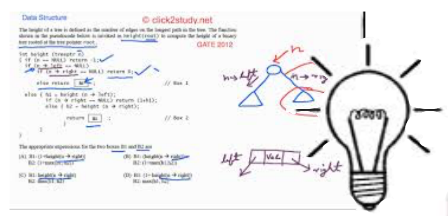




# Today's Plan:
# Computational models of language acquisition

I. Why : Because language acquisition is pretty amazing and we want to understand how it works

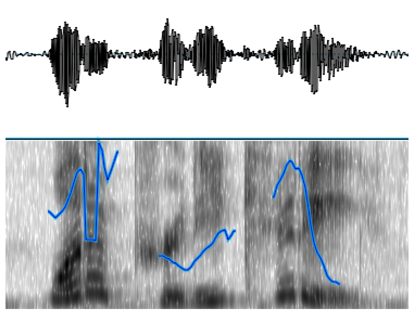





II. How : By building informative computational models

## III. What we can learn : A lot about a lot



speech segmentation

KI tty

metrical phonology

*Who does... is pretty?*

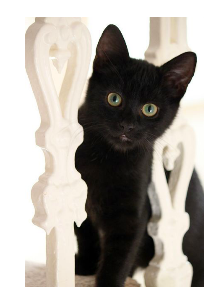
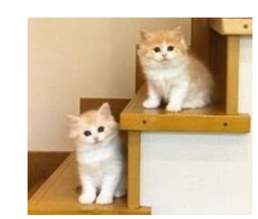
syntax

Noun

syntactic categorization

*another one*

syntax, semantics

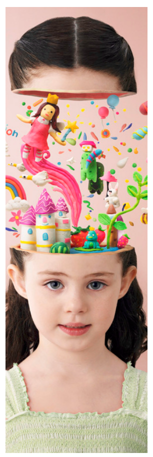



pragmatics

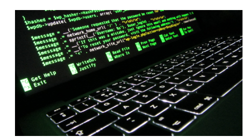
*Every kitty didn't ...*

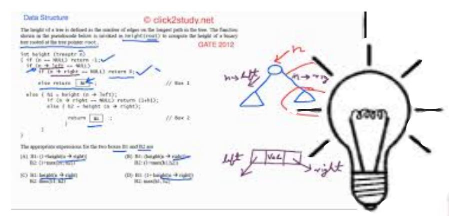
# Today's Plan:
# Computational models of language acquisition

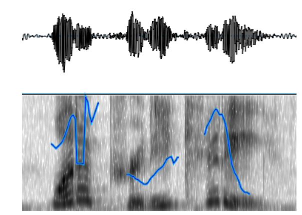
I. Why : Because language acquisition is pretty amazing
and we want to understand how it works



II. How : By building informative
computational models



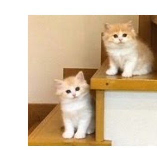


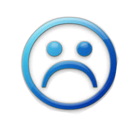


Noun

III. What we can learn : A lot about a lot

✓KI tty



  *another one*

*Who does… is pretty?*

*Every kitty didn't …*

**This is a great tool - so let's use it to understand how linguistic representations develop!**

# Thank you!

Computation of Language Laboratory

UC Irvine

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

This work was supported in part by NSF grants BCS-0843896 and BCS-1347028.

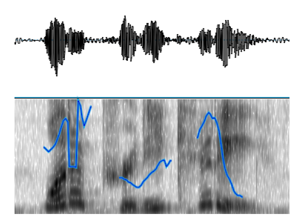
# Extra material

# What we can learn

metrical phonology

syntactic categorization

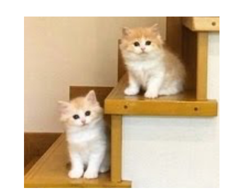
speech segmentation



Noun

syntax

✓KI tty



*Who does… is pretty?*

*another one*

*Every kitty didn't …*

pragmatics

syntax, semantics

# What we can learn

metrical phonology

✓ KI tty

✗ ki TTY

# What we can learn

metrical phonology

✓ a DO ra ble

✗ A do RA ble

✗ a DO ra BLE

✓ KI tty

✗ ki TTY

# What we can learn

metrical phonology

✓ a DO ra ble

✗ A do RA ble

✗ a DO ra BLE

✓ KI tty

✗ ki TTY

Our underlying knowledge representation of the metrical phonology system allows us to generate these metrical stress preferences.

# What we can learn

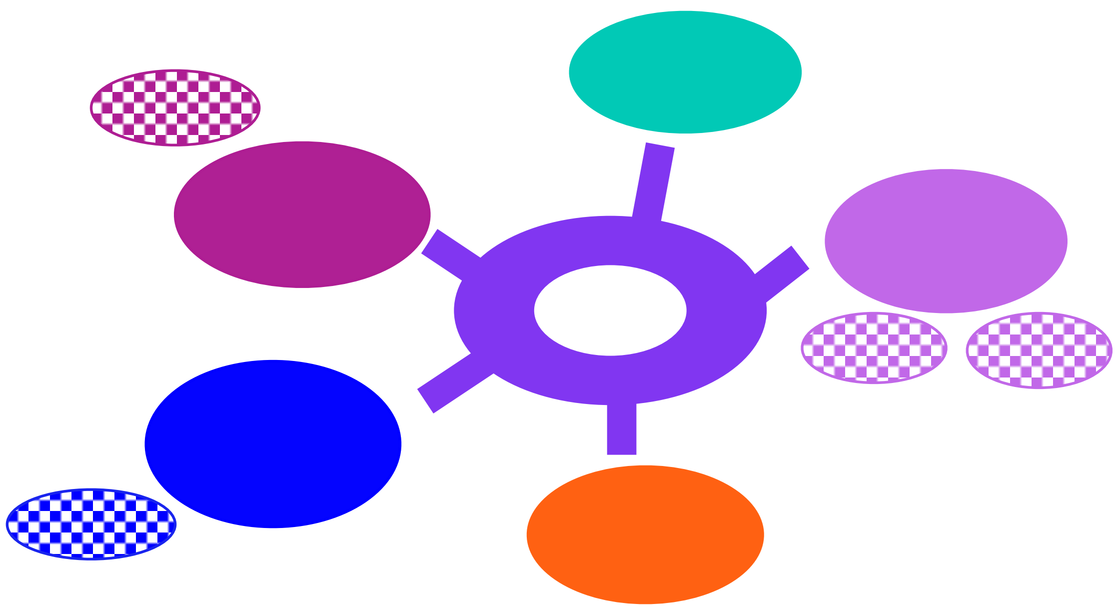
metrical phonology

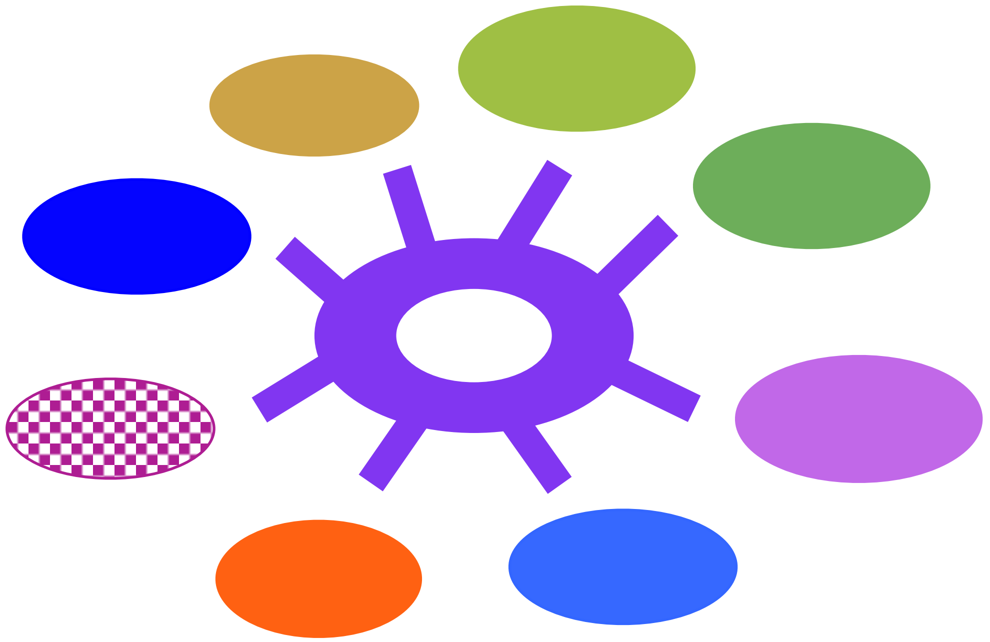
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

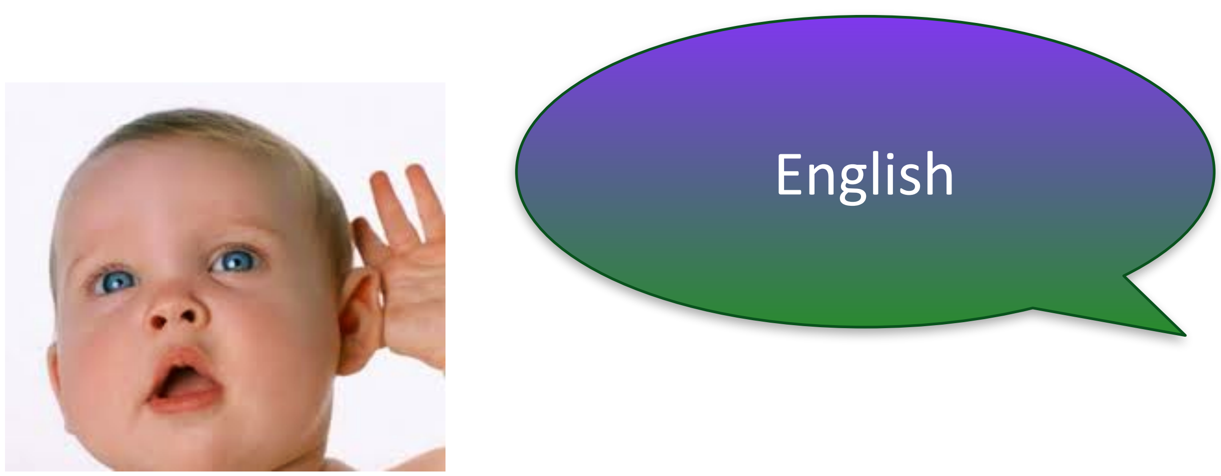
✓ KI tty
✗ ki TTY

knowledge representation options

parameters whose values must be set







# What we can learn

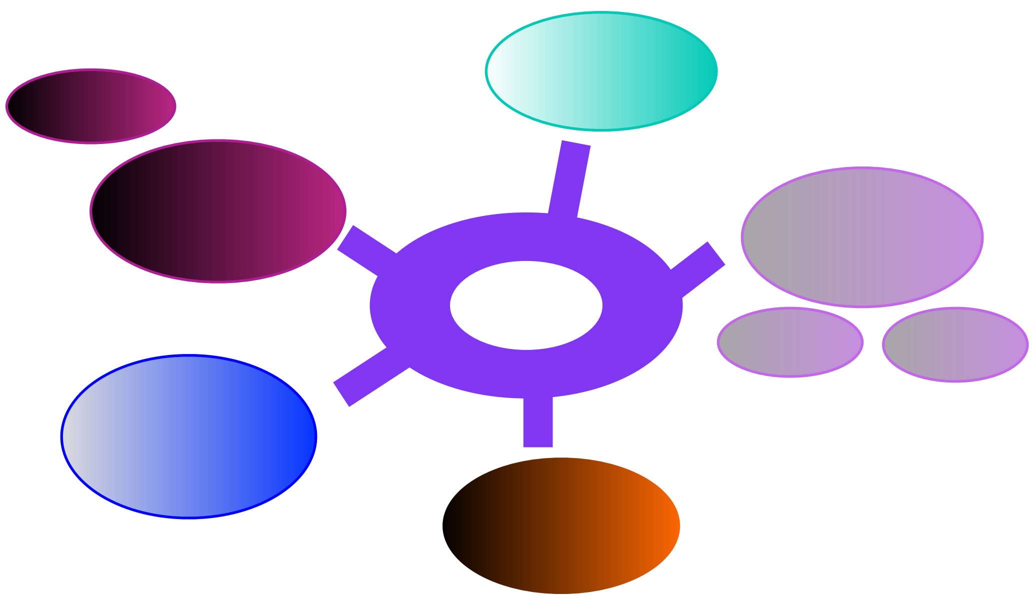
metrical phonology

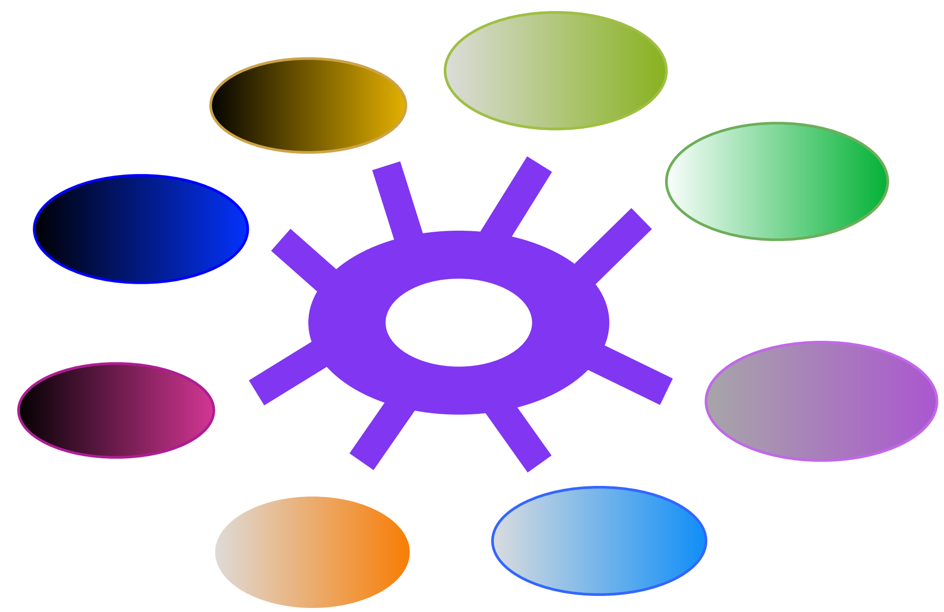
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

✓ KI tty
✗ ki TTY

## knowledge representation options

### parameters whose values must be set





English

# What we can learn

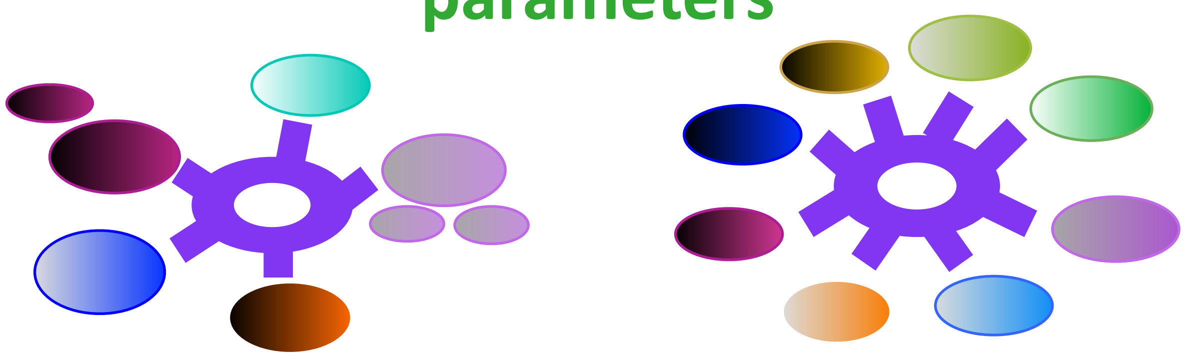
metrical phonology

✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

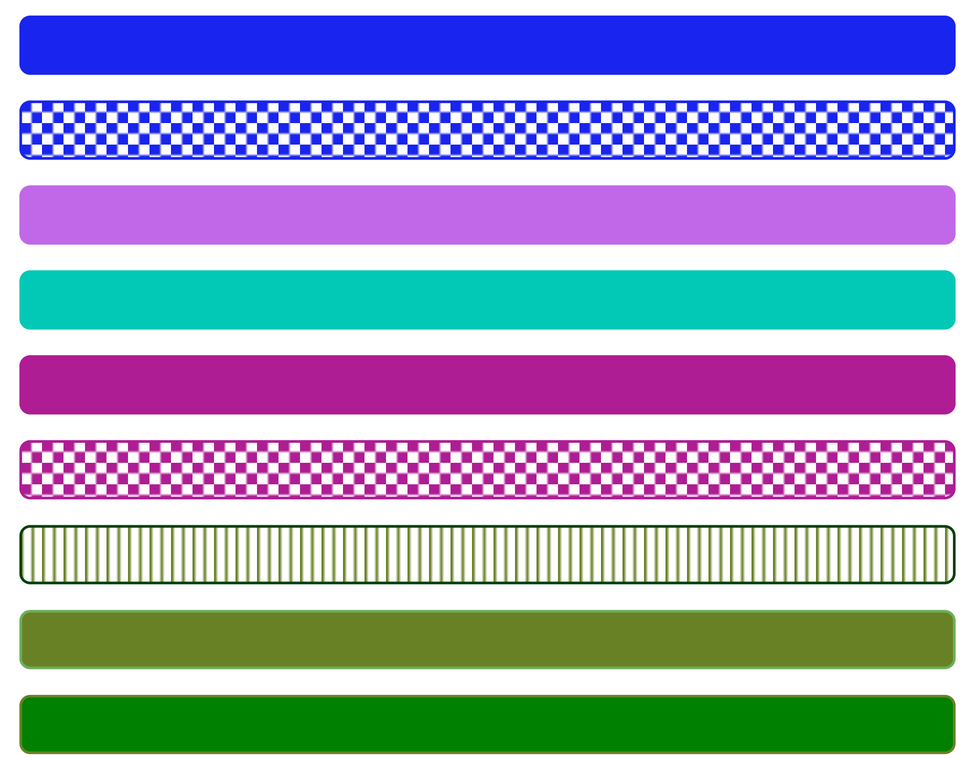
✓ KI tty
✗ ki TTY

## knowledge representation options

### parameters



### Violable constraints that must be ranked



English

# What we can learn

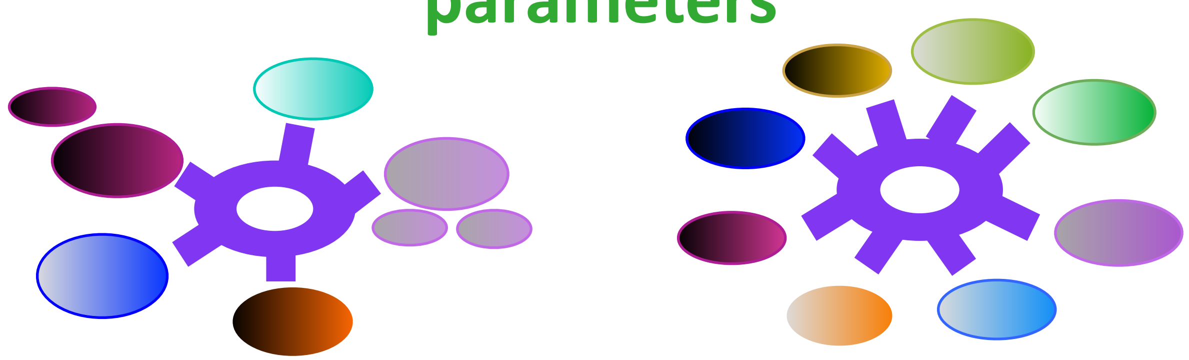
metrical phonology

✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

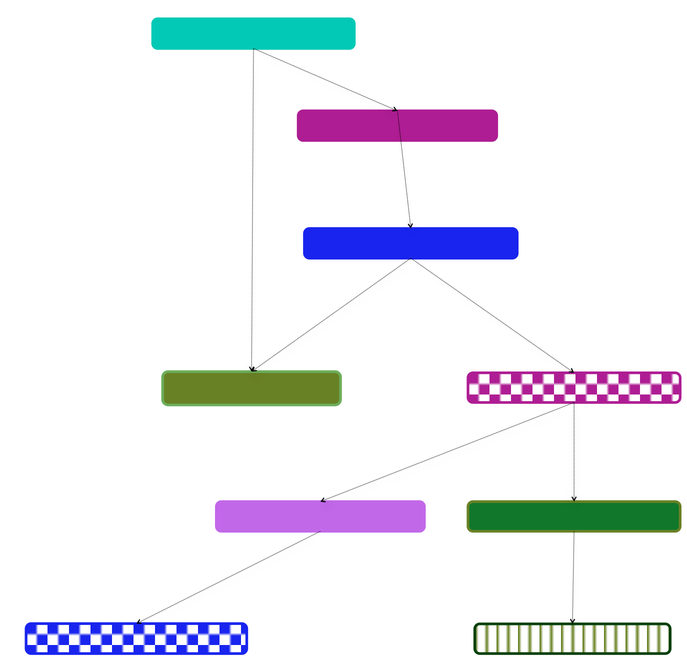
✓ KI tty
✗ ki TTY

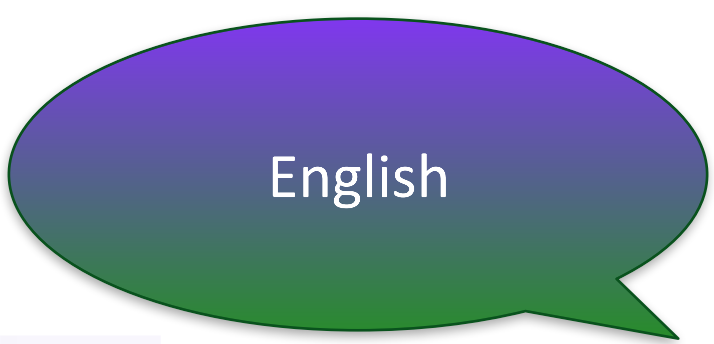
## knowledge representation options

### parameters



### Violable constraints that must be ranked






# What we can learn

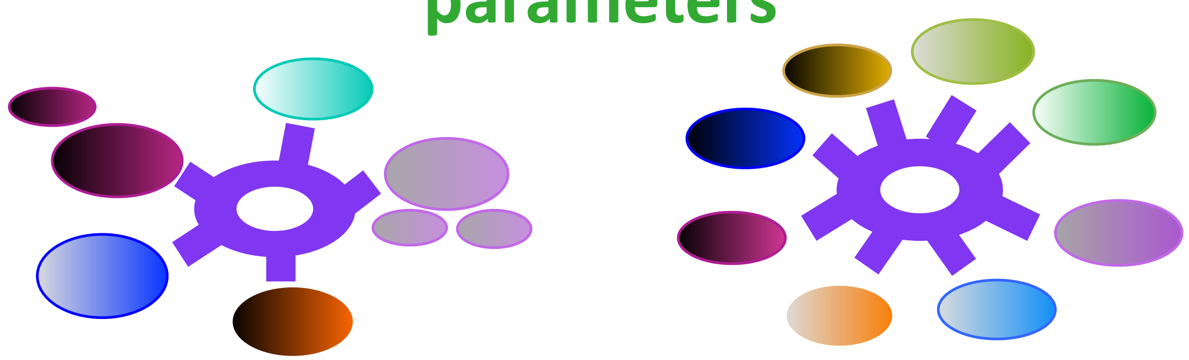
metrical phonology

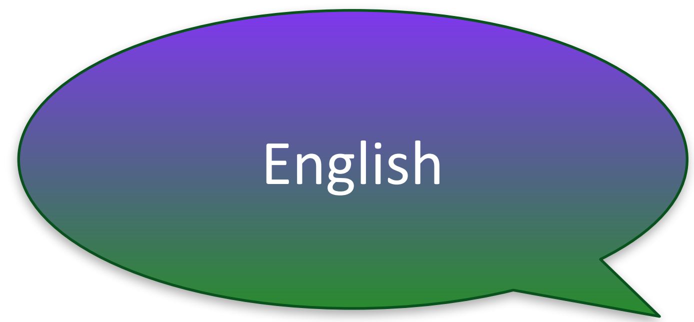
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

✓ KI tty
✗ ki TTY

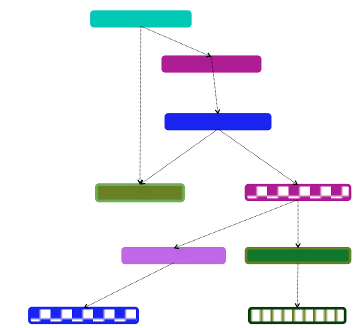
**knowledge representation options**

**parameters**






**constraints**



**These representations have some similarities, but aren’t obviously using identical variables.**

How do we choose among these representations and their English versions?

# What we can learn

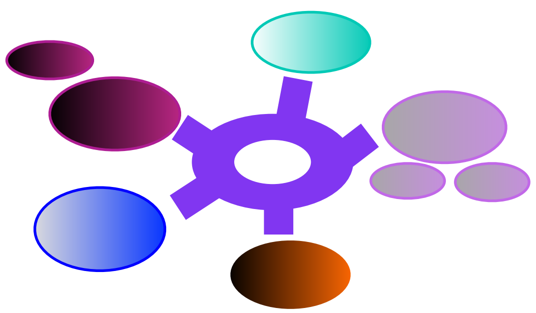
metrical phonology

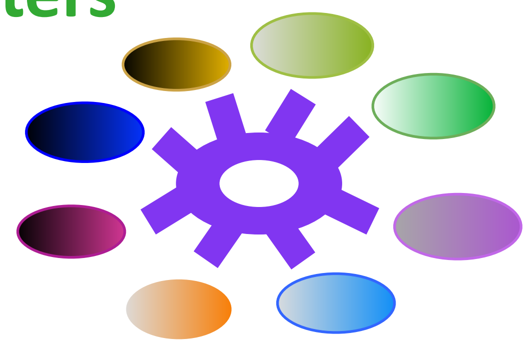
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

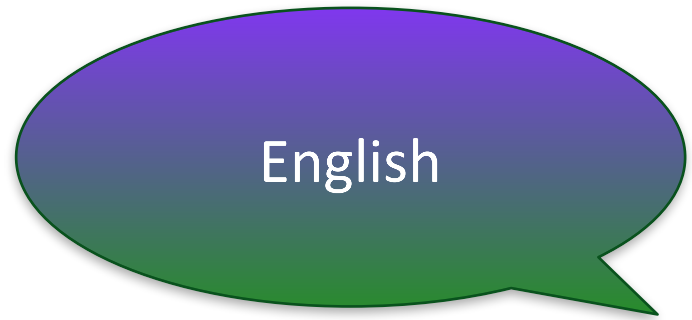
✓ KI tty
✗ ki TTY

## knowledge representation options

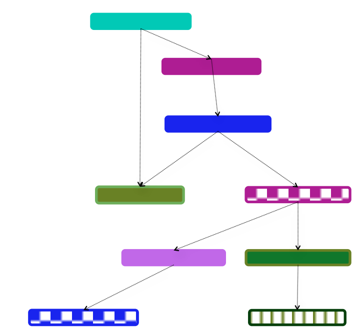
parameters








constraints



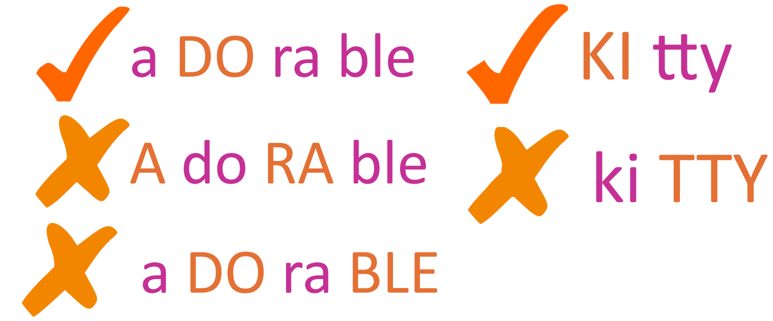
## How do we choose among these representations and their English versions?

Answer: Let's see how learnable they are from the English data children typically encounter!

Pearl et al. 2014, Pearl 2017, Pearl et. al in press

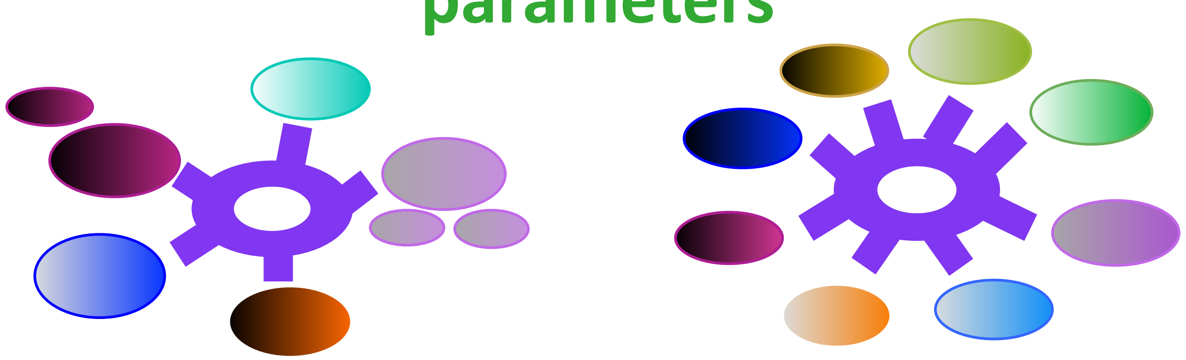
# What we can learn

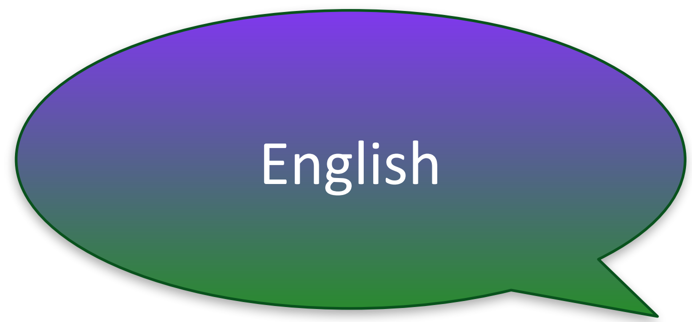
metrical phonology



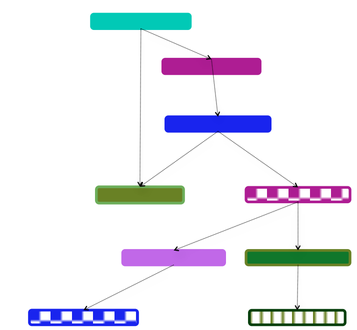

## knowledge representation options

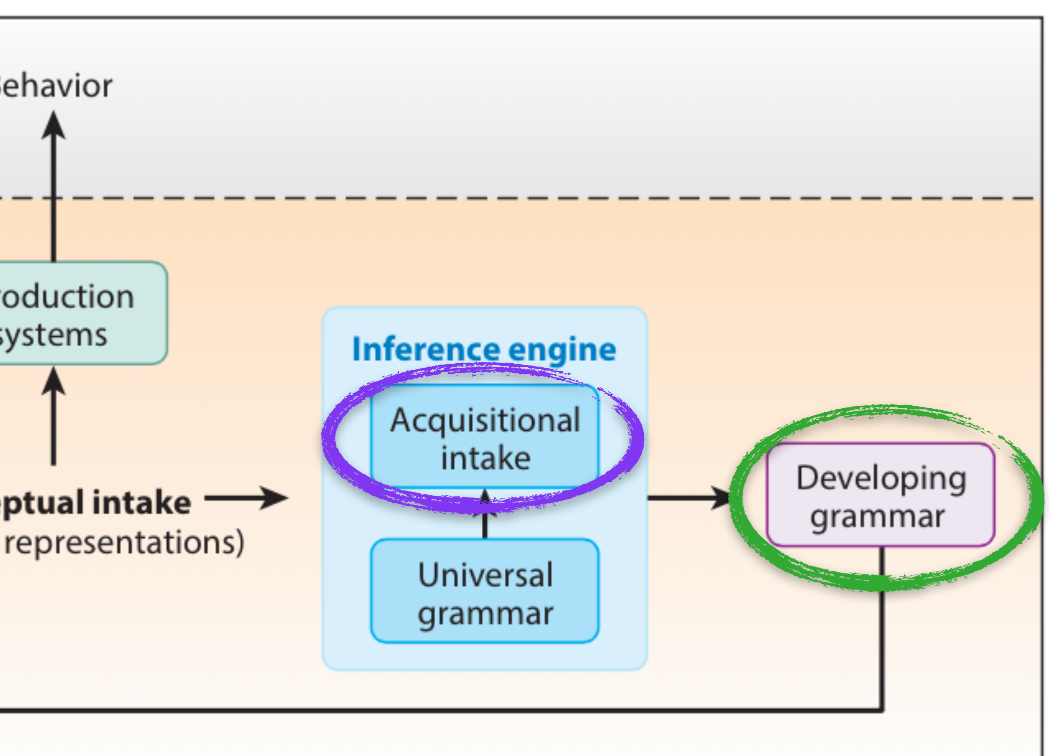
parameters






constraints






## how learnable they are

Computational-level analysis

Modeled learners given realistic samples of English child-directed speech can identify parameter combinations or constraint rankings that are very good at accounting for the input especially if children use a data filter.

Pearl et al. 2014, Pearl 2017 , Pearl et. al in press

# What we can learn

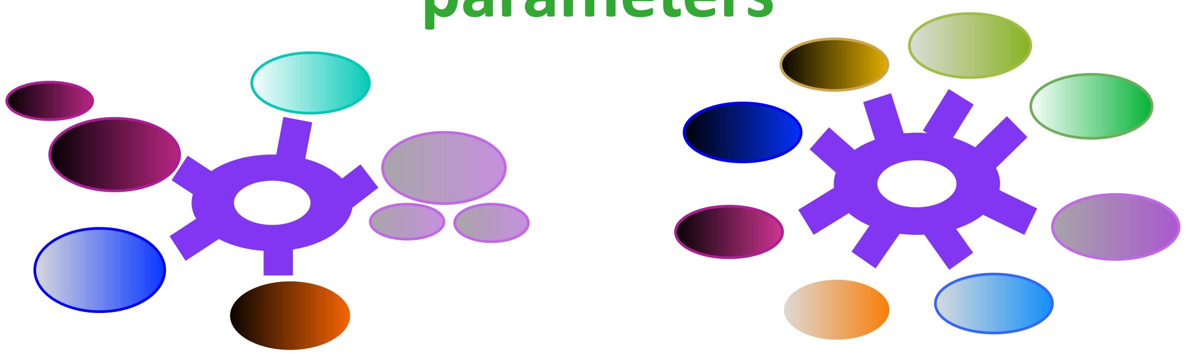
metrical phonology

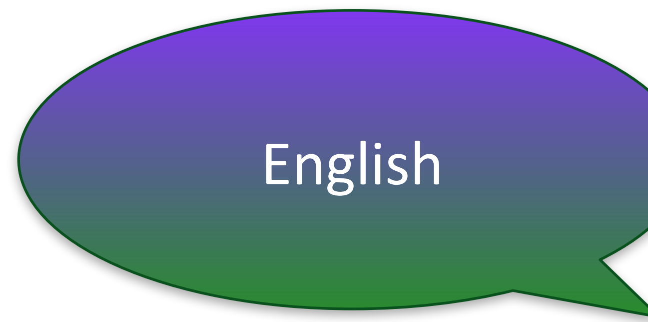
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

✓ KI tty
✗ ki TTY

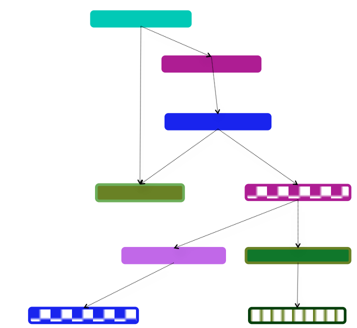
## knowledge representation options

parameters





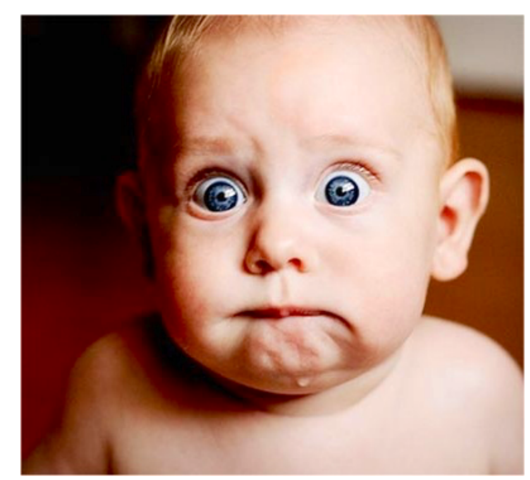

constraints



how learnable they are

Computational-level analysis

But the best options for English data aren't the ones currently proposed for English.



Pearl et al. 2014, Pearl 2017 , Pearl et. al in press

# What we can learn

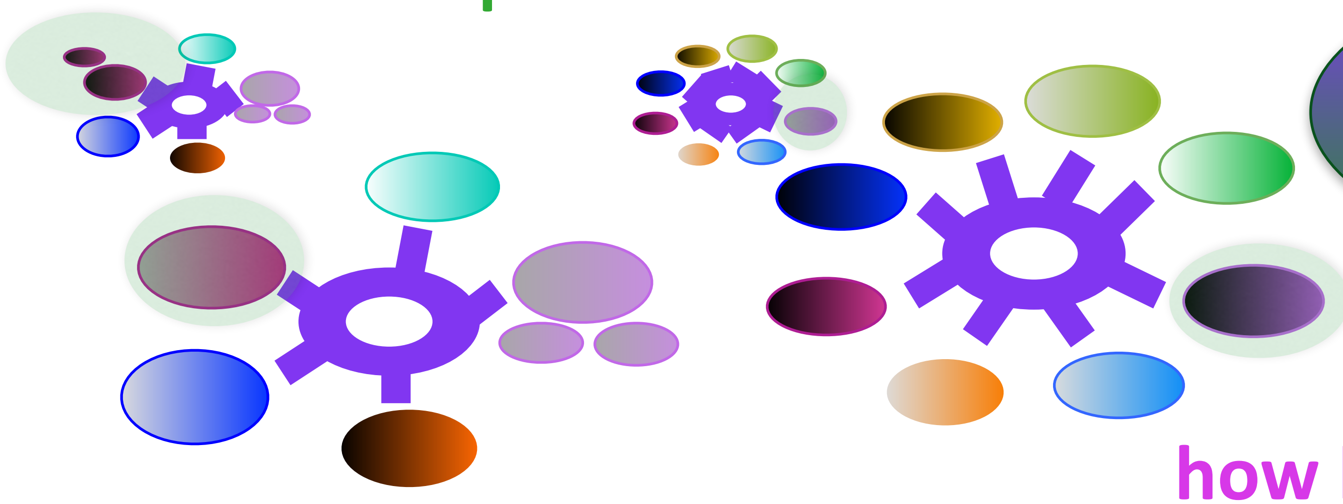
metrical phonology

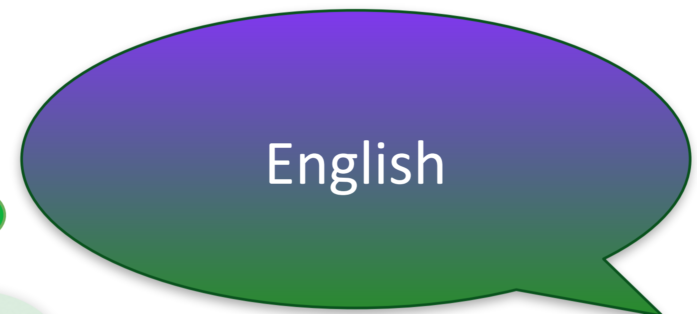
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

✓ KI tty
✗ ki TTY

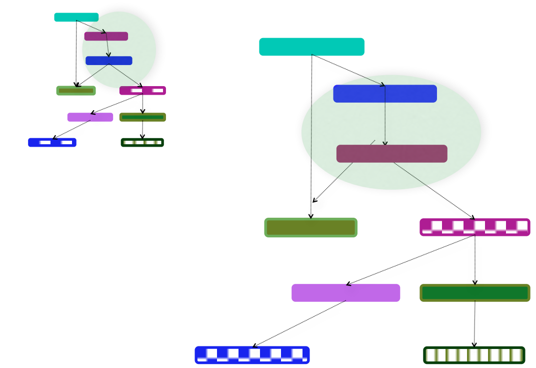
## knowledge representation options

**parameters**






**constraints**



**how learnable they are**

Computational-level analysis

Other options (differing very slightly) are much more easily learnable.

Pearl et al. 2014, Pearl 2017 , Pearl et. al in press

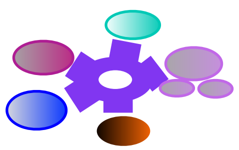
# What we can learn

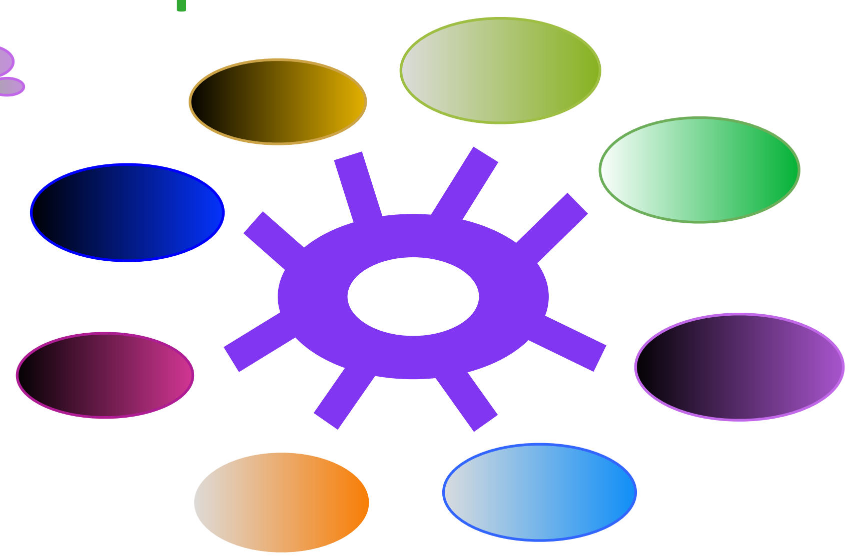
metrical phonology

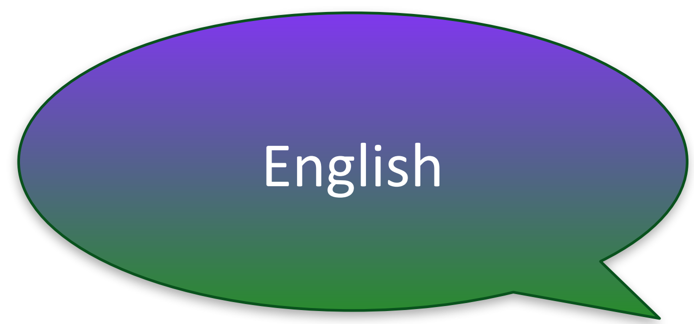
- ✓ a DO ra ble
- ✗ A do RA ble
- ✗ a DO ra BLE
- ✓ KI tty
- ✗ ki TTY

knowledge representation options

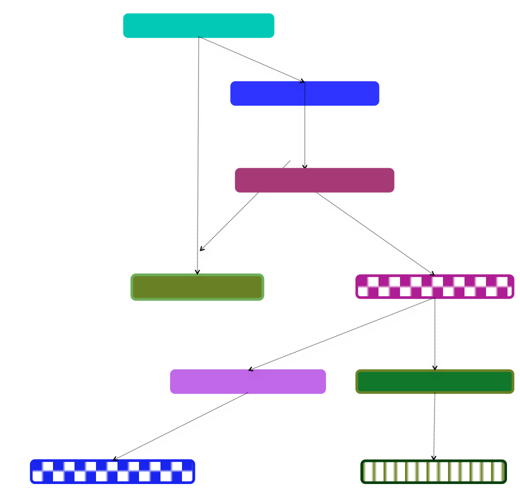
parameters





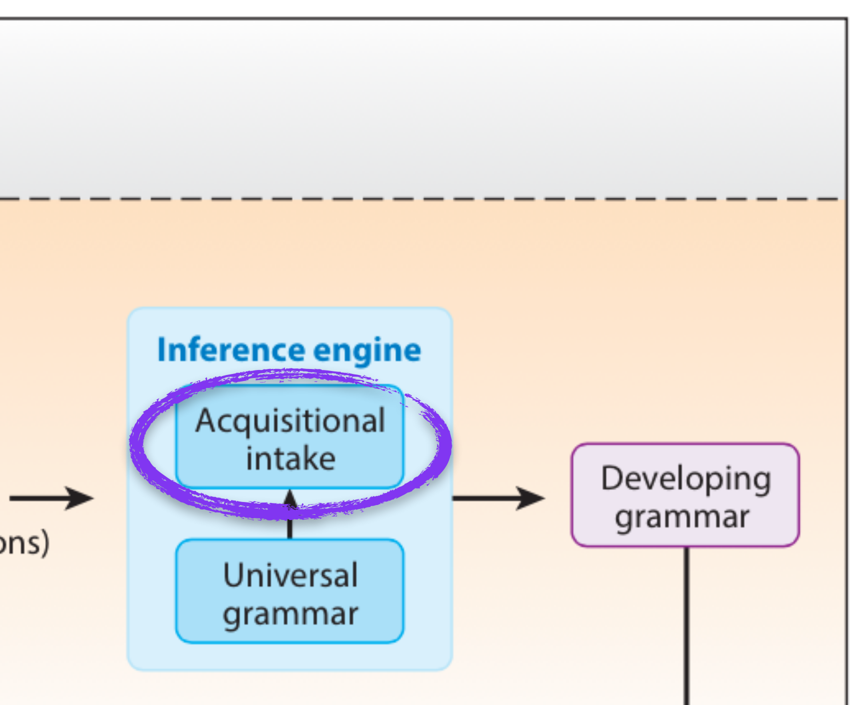



constraints



how learnable they are

Computational-level analysis




And two do particularly well when a data filter is in place.

Pearl et al. 2014, Pearl 2017 , Pearl et. al in press

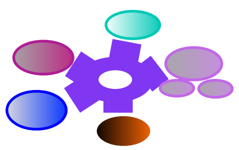
# What we can learn

metrical phonology

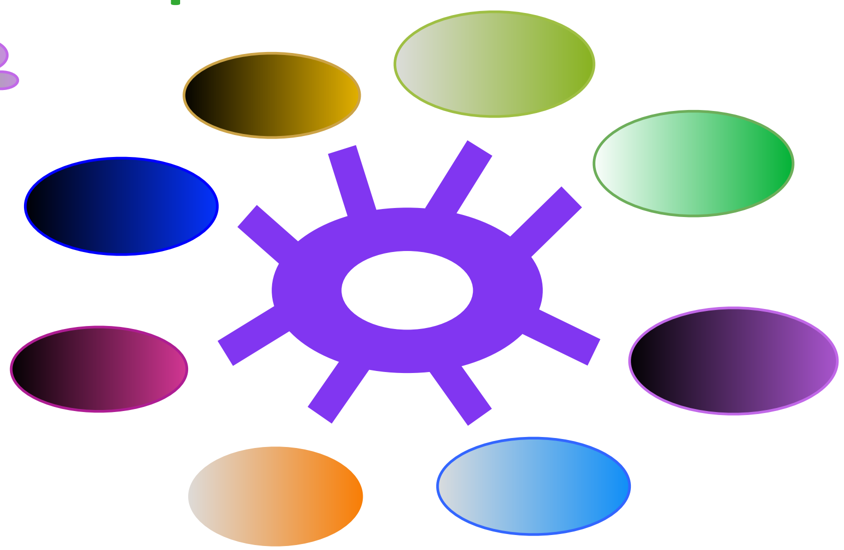
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

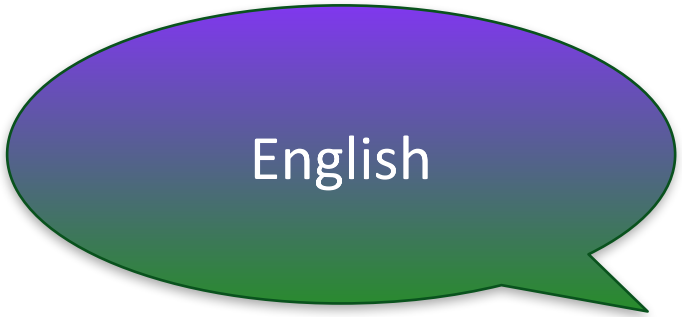
✓ KI tty
✗ ki TTY

## knowledge representation options

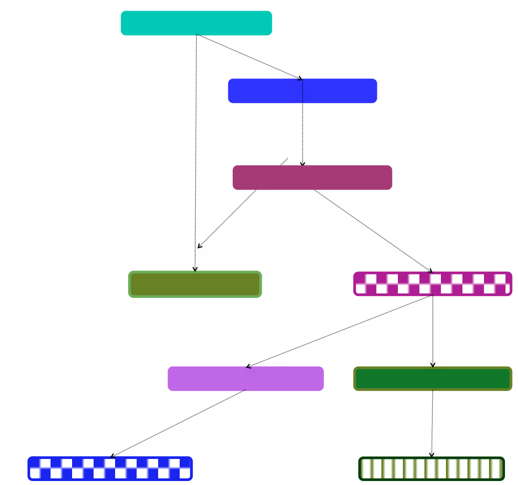


**parameters**






**constraints**



**By modeling acquisition, we provide support for these two theories of English representation.**

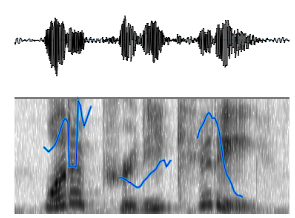
Pearl et al. 2014, Pearl 2017 , Pearl et. al in press

# What we can learn

metrical phonology

syntactic categorization

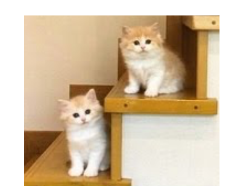
speech segmentation



Noun

syntax

✓KI tty



*Who does… is pretty?*

*another one*

*Every kitty didn't …*

pragmatics

syntax, semantics

# What we can learn

syntactic categorization

idea

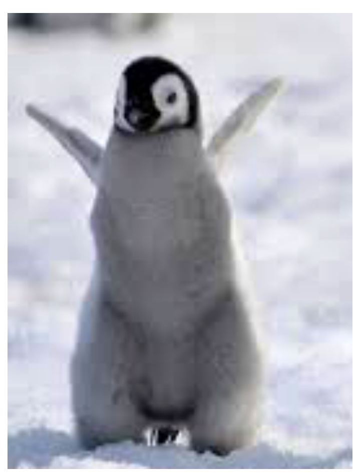
glitter

unicorn

penguin

**Noun**

owl



kitty

# What we can learn

syntactic categorization

idea

glitter

unicorn

penguin

**Noun**

owl

kitty

Nouns behave similarly:

They can combine with certain types of words to make larger units (like Noun Phrases).

# What we can learn

syntactic categorization

Determiner + Noun ("the kitty")

[NP → Det + N]

idea

glitter

unicorn

**Noun**

penguin

owl

Nouns behave similarly:

They can combine with certain types of words to make larger units (like Noun Phrases).

# What we can learn

syntactic categorization

Determiner + Noun (“the ”)

[NP → Det + N]

dax

idea

glitter

unicorn

**Noun**

penguin

owl

kitty

Rule with category Noun = new phrases with words of category Noun

**This is very handy for generating new expressions we haven’t heard before.**

# What we can learn

syntactic categorization

Determiner + Noun ("the dax ")

[NP → Det + N]

idea

glitter

unicorn

**Noun**

penguin

owl

kitty

Rule with category Noun = new phrases with words of category Noun

**This is very handy for generating new expressions we haven't heard before.**

# What we can learn

syntactic categorization

idea
glitter
unicorn
**Noun**
owl
penguin
kitty

We have many categories in human language.

Some are open-class — it's easy to add new words to them.

# What we can learn

syntactic categorization

idea
glitter
unicorn
**Noun**
owl
penguin
kitty

We have many categories in human language.

Some are open-class — it's easy to add new words to them.

[VP → Negation + V]

It's not daxing
- it's dancing!

surprise
stand
**Verb**
find
dance
adore

# What we can learn

syntactic categorization

idea
glitter unicorn
**Noun**
penguin owl
kitty

We have many categories in human language.

Some are open-class — it's easy to add new words to them.

dax
surprise stand
**Verb**
find dance
adore

# What we can learn

syntactic categorization

idea
glitter unicorn
**Noun** owl
penguin
kitty
surprise stand
find **Verb** dance
adore

We have many categories in human language.

Some are closed-class — the words in them are fixed.

[VP → Negation + V]

It's not daxing
- it's dancing!

didn't not
**Negation**
wouldn't can't
won't

# What we can learn

syntactic categorization

didn't
not
**Negation**
wouldn't
can't
won't

surprise
stand
find
**Verb**
dance
adore

idea
glitter
unicorn
**Noun**
owl
penguin
kitty

We have many categories in human language.

Some are closed-class — the words in them are fixed.

[VP → Auxiliary + V]
It would sing
if it could sing

might
would
could
**Auxiliary**
will
can
should

# What we can learn

syntactic categorization

didn't
not
Negation
wouldn't
can't
won't

might
would
could
Auxiliary
will
should
can

surprise
stand
find
Verb
dance
adore

idea
glitter
unicorn
Noun
owl
penguin
kitty

There's significant debate on when these categories develop.

# What we can learn

syntactic categorization

**Auxiliary**: might, would, could, will, should, can

**Negation**: didn't, not, wouldn't, can't, won't

**Verb**: surprise, stand, find, adore

**Noun**: idea, glitter, unicorn, owl, penguin, kitty

There's significant debate on when these categories develop.

Easy to observe: When children know individual words.


dance

# What we can learn

syntactic categorization

didn't
not
**Negation**
wouldn't
can't
won't

might
would
could
**Auxiliary**
will
should
can

idea
glitter
unicorn
**Noun**
owl
penguin
kitty

There's significant debate on when these categories develop.

Harder to observe: When children have recognized these words belong to categories.


surprise
adore
stand
find **Verb** ???
dance

# What we can learn

syntactic categorization

didn't
not
**Negation**
wouldn't
can't
won't

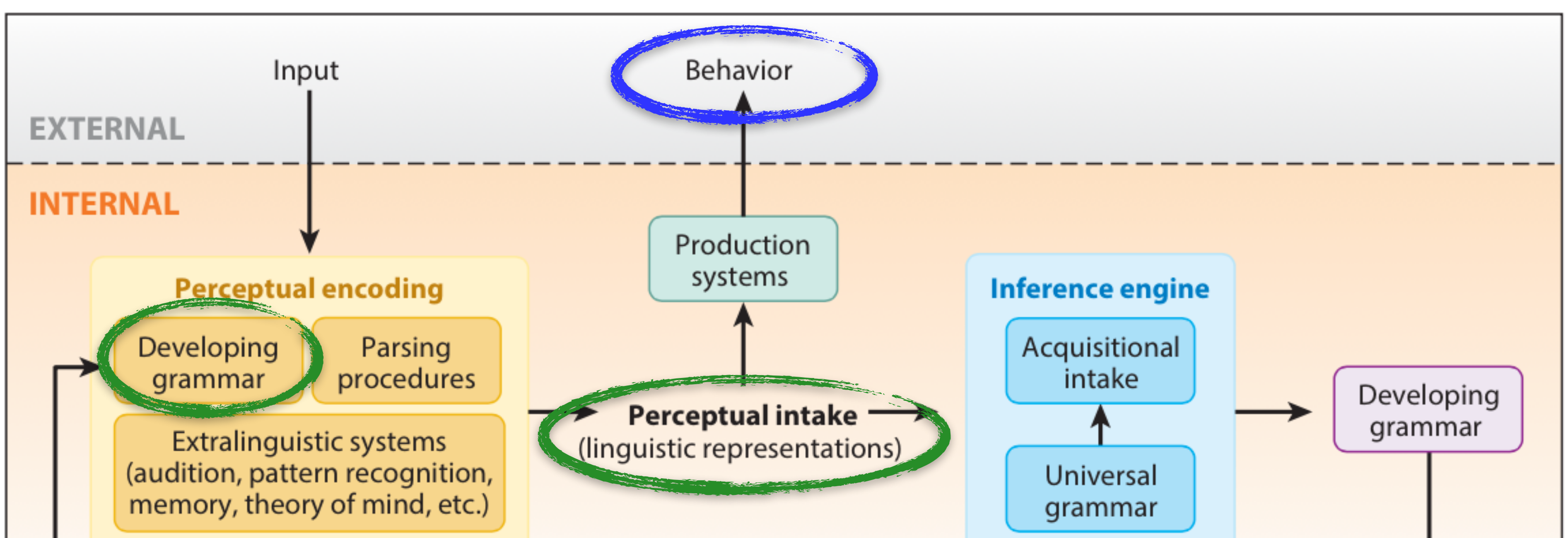
might
would
could
will
**Auxiliary**
should
can

surprise
stand
find
**Verb**
dance
adore

idea
glitter
unicorn
**Noun**
owl
penguin
kitty

What we can do: Computational-level analysis of children's productions, using formal metrics that describe how children generate their utterances given their underlying representations

Bates, Pearl & Braunwald, in prep.



# What we can learn

syntactic categorization

didn't
not
**Negation**
wouldn't
can't
won't

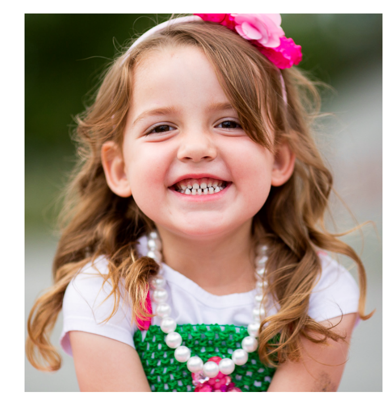
might
would
could
will
**Auxiliary**
can
should

surprise
stand
find
**Verb**
dance
adore

idea
glitter
unicorn
**Noun**
owl
penguin
kitty

Computational-level

Analyzing the utterances produced by a single American English child between the ages of 20 and 24 months



Bates, Pearl & Braunwald, in prep.

# What we can learn

syntactic categorization

✓ **Auxiliary**: might, would, could, will, should, can

✓ **Negation**: didn't, not, wouldn't, can't, won't

~~**Verb**~~: surprise, stand, find, dance, adore

~~**Noun**~~: idea, glitter, unicorn, penguin, owl, kitty

Computational-level

Analyzing the utterances produced by a single American English child between the ages of 20 and 24 months

Utterances compatible with having adult-like closed-class categories, but not adult-like open-class categories.

Bates, Pearl & Braunwald, in prep.

# What we can learn

syntactic categorization

might would could will **Auxiliary** should can

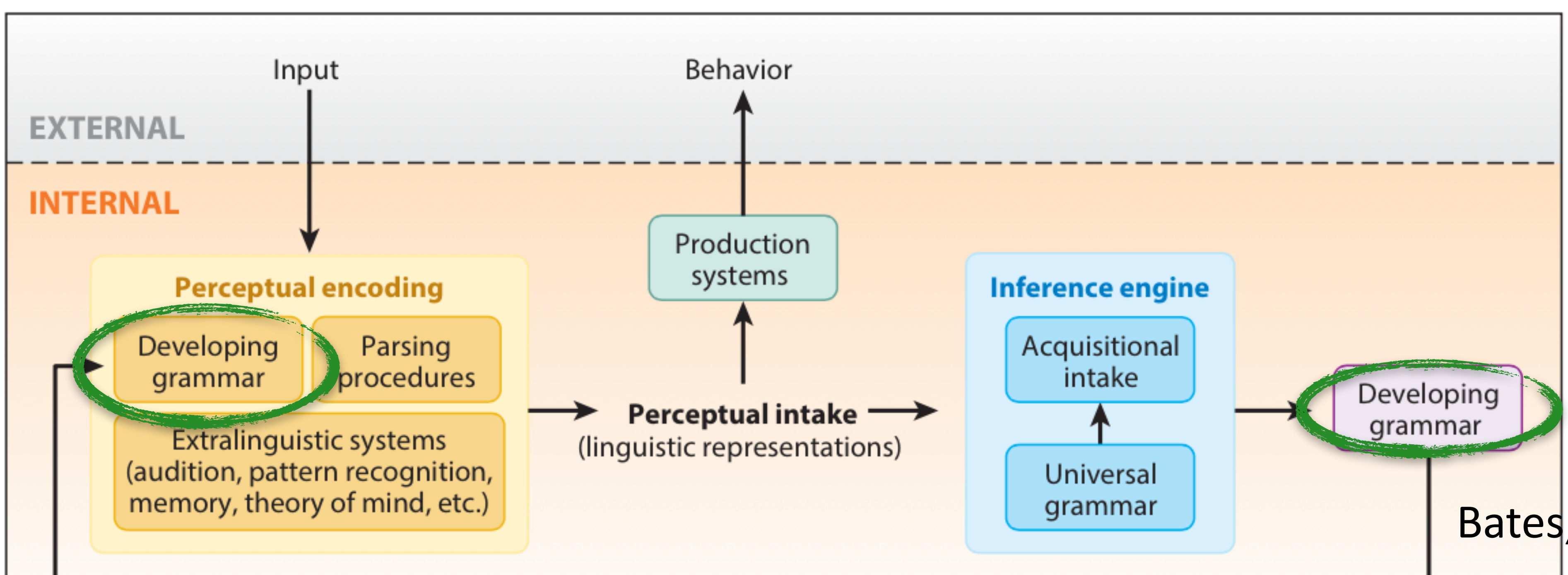
didn't not **Negation** wouldn't can't won't

surprise stand find **Verb** dance adore

idea glitter unicorn **Noun** owl penguin kitty

## Computational-level

This suggests that closed-class categories may develop into an adult-like state earlier than open-class categories and much earlier than previously thought.



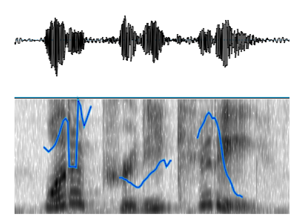

Bates, Pearl & Braunwald, in prep.

# What we can learn

metrical phonology

syntactic categorization

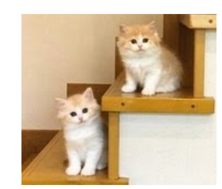
speech segmentation



Noun

syntax


KI tty


*Who does … is pretty?*

*another one*

*Every kitty didn't …*

pragmatics

syntax, semantics

# What we can learn

pragmatics

"Every kitty didn't sit on the stairs"

X No kitties sat on the stairs.

✓ Not all kitties sat on the stairs.

**Why are two interpretations available?**

**Quantifier scope**

# What we can learn

pragmatics

## Quantifier scope

"**Every** kitty did**n't** sit on the stairs"

$\forall$ $\neg$

✗ No kitties sat on the stairs.

✓ Not all kitties sat on the stairs.

# What we can learn

**pragmatics**

## Quantifier scope

" **Every** kitty did**n't** sit on the stairs"

$\forall \qquad \neg$

**surface** $\forall$ kitties ***k*** $\neg$ ***k*** sat on the stairs

*"For all kitties k, it's not true that k sat on the stairs"*

✗ No kitties sat on the stairs.

✓ Not all kitties sat on the stairs.

# What we can learn

**pragmatics**

## Quantifier scope

" **Every** kitty did**n't** sit on the stairs"

$\forall$ $\neg$

**surface** $\forall$ kitties ***k*** $\neg$ ***k*** sat on the stairs

*"For all kitties k, it's not true that k sat on the stairs"*

✗ No kitties sat on the stairs.

**inverse** $\neg$ $\forall$ kitties ***k*, *k*** sat on the stairs

*"It's not true that for all kitties k, k sat on the stairs"*

✓ Not all kitties sat on the stairs.

# What we can learn

pragmatics

## Quantifier scope

✓ “**Every** kitty did**n’t** sit on the stairs”

$\forall \quad \neg$

✗

**surface** $\forall \neg$ No kitties sat on the stairs.

✓

**inverse** $\neg \forall$ Not all kitties sat on the stairs.

**Adults**

# What we can learn

pragmatics

## Quantifier scope

X "**Every** kitty did**n't** sit on the stairs"

$\forall \quad \neg$

X

**surface** $\forall \neg$ No kitties sat on the stairs.

??

**inverse** $\neg \forall$ Not all kitties sat on the stairs.

**5-year-olds**

**But why?**

# What we can learn

pragmatics

## Quantifier scope

X " **Every** kitty did**n't** sit on the stairs"

$\forall$ $\neg$

??

inverse $\neg$ $\forall$ Not all kitties sat on the stairs.

**5-year-olds**

One idea: **grammatical processing** problem

# What we can learn

**pragmatics**

## Quantifier scope

X "**Every** kitty did**n't** sit on the stairs"

$\forall \quad \neg$

??

**inverse** $\neg \quad \forall$ Not all kitties sat on the stairs.

## 5-year-olds

One idea: **grammatical processing** problem

The **inverse scope** is harder to get from the surface string.

# What we can learn

pragmatics

## Quantifier scope

“**Every** kitty did**n’t** sit on the stairs”

$\forall$ $\neg$

??

**inverse** $\neg$ $\forall$ Not all kitties sat on the stairs.

**5-year-olds**

One idea: **grammatical processing** problem

Another idea: **pragmatic context** management problem.

# What we can learn

pragmatics

## Quantifier scope

✗ "**Every** kitty did**n't** sit on the stairs"

$\forall$ $\neg$

??

**inverse** $\neg$ $\forall$ Not all kitties sat on the stairs.

*Did none of the kitties sit on the stairs?*

*Do kitties like stairs?* **QUD** *How many kitties sat on the stairs?*

**5-year-olds**

One idea: **grammatical processing** problem

Another idea: **pragmatic context** management problem.

Children thought the topic of conversation (the implicit **Q**uestion **U**nder **D**iscussion) was something else and this utterance doesn't answer that QUD very well.

# What we can learn

pragmatics

## Quantifier scope

"**Every** kitty did**n't** sit on the stairs"

$\forall \quad \neg$

??

**inverse** $\neg \; \forall$ Not all kitties sat on the stairs.

*Kitties don't like stairs*

**expectations about the world**

*Kitties love stairs.*

*Kitties don't care about stairs.*

**5-year-olds**

One idea: **grammatical processing** problem

Another idea: **pragmatic context** management problem.

**QUD**

Children's prior **expectations about the world** make this utterance less informative.

# What we can learn

pragmatics

## Quantifier scope

“**Every** kitty did**n’t** sit on the stairs”

$\forall \qquad \neg$

??

**inverse** $\neg \quad \forall$ Not all kitties sat on the stairs.

**QUD**

**grammatical processing**

**expectations about the world**

**5-year-olds**

It’s hard to manipulate only one of these factors in experimental research investigating children’s responses.

# What we can learn

pragmatics

## Quantifier scope

"**Every** kitty did**n't** sit on the stairs"

$\forall \quad \neg$

??

**inverse** $\neg \forall$ Not all kitties sat on the stairs.

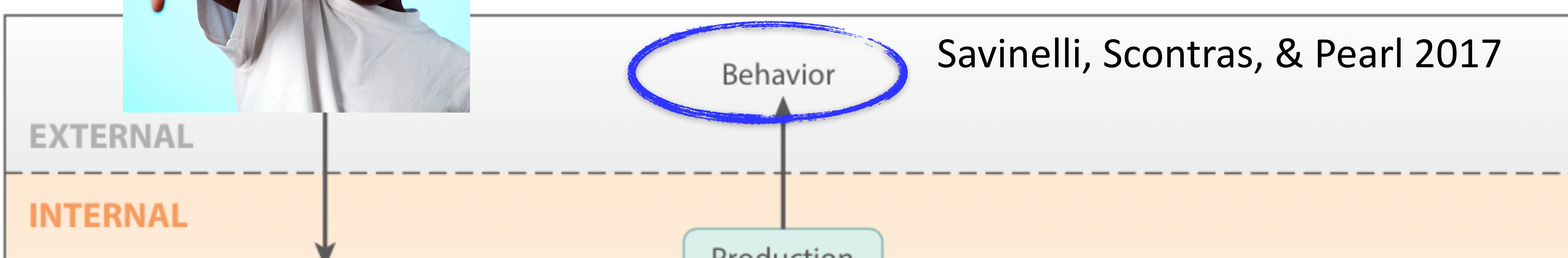
**QUD**

**grammatical processing**

**expectations about the world**

**5-year-olds**

Using a computational-level model that formalizes the separate contribution of each factor, we can determine which ones have the largest impact on children's observed behavior.




Savinelli, Scontras, & Pearl 2017

# What we can learn

pragmatics

## Quantifier scope

✗ "**Every** kitty did**n't** sit on the stairs"

$\forall$ $\neg$

??

**inverse** $\neg\ \forall$ Not all kitties sat on the stairs.

**QUD**

**grammatical processing**

**expectations about the world**

**5-year-olds**

The pragmatic factors seem to be the driving force behind children's behavior. This suggests that 5-year-olds are still developing their ability to manage the pragmatic context of a conversation as well as adults do.

Savinelli, Scontras, & Pearl 2017