Computational models for language acquisition: Why, how, and what we can learn

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

June 3, 2017:

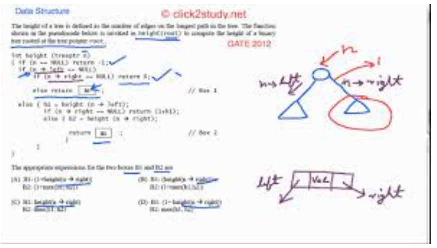
Cognitive Science 1st Annual Workshop

Simon Fraser University

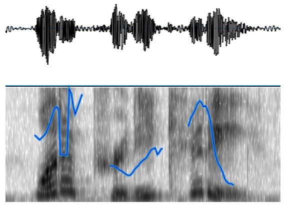
Today's Plan:
Computational models of language acquisition

I. Why


II. How



III. What we can learn



Noun

KI tty


Who does… is pretty?

another one

Every kitty didn't …

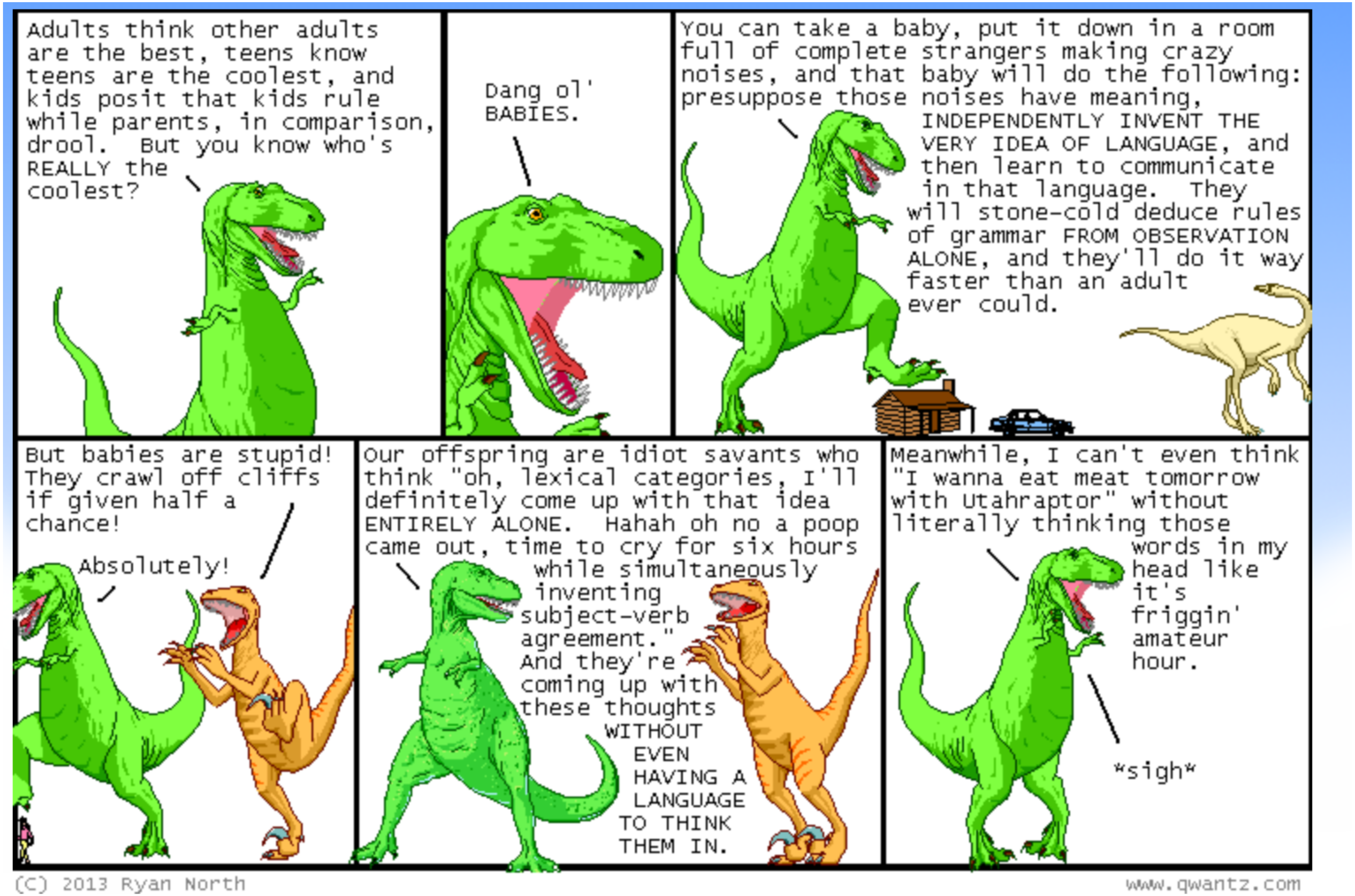
Today's Plan:
Computational models of language acquisition

I. Why

Why language acquisition?

Babies are amazing at learning language

Babies are amazing at learning language



http://www.qwantz.com/index.php?comic=2479

Babies are amazing at learning language

Wait…what exactly do you know when you know a language?

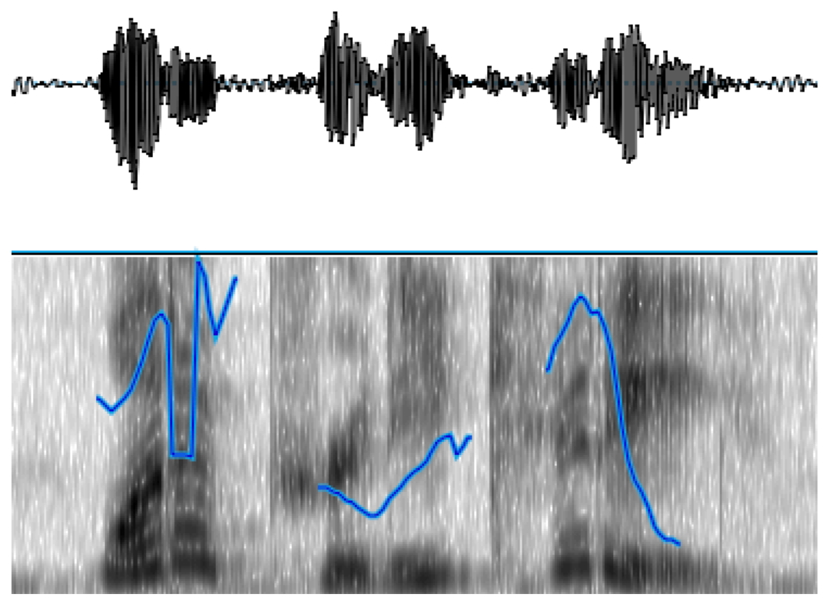
Wait…what exactly do you know when you know a language?

A lot!

Wait…what exactly do you know when you know a language?

A lot!

You know how to identify words in fluent speech (speech segmentation)



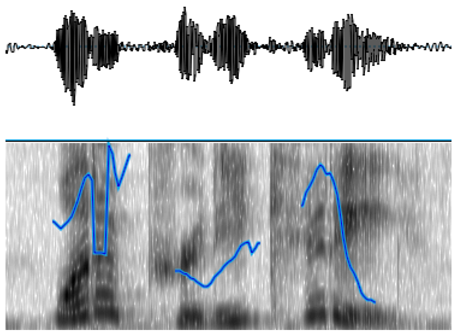
= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

Wait…what exactly do you know when you know a language?

A lot!



what a pretty kitty!

speech segmentation

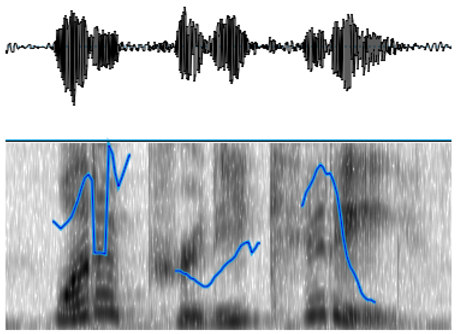
You know how to pronounce words (metrical phonology)

✓ KI tty

✗ ki TTY

Wait…what exactly do you know when you know a language?

A lot!



what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

metrical phonology

You know that certain words behave like other words (syntactic categorization)

Noun

what a pretty ___!

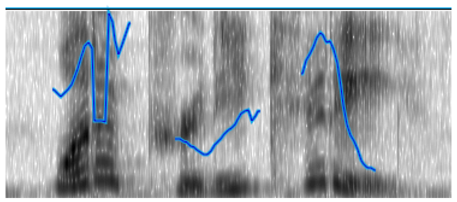
owl

penguin

kitty

Wait…what exactly do you know when you know a language?

A lot!



what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

metrical phonology

Noun

penguin owl

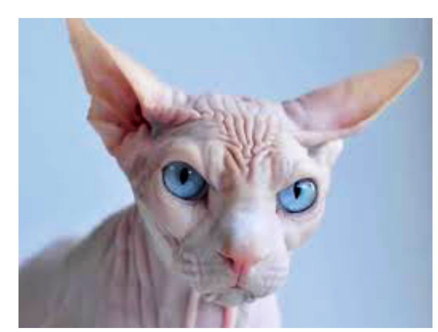
kitty

syntactic categorization

You know how to interpret words in context (syntax, semantics)

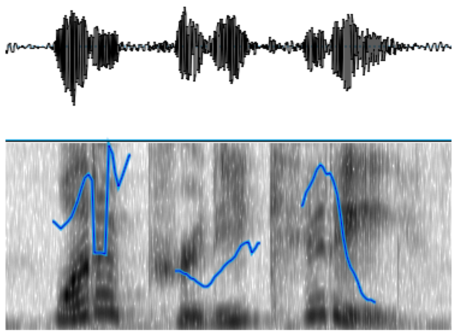
"Oh look — a pretty kitty!"

"Look — there's another one!"



Wait…what exactly do you know when you know a language?

A lot!



what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

metrical phonology

Noun

penguin owl

kitty

syntactic categorization

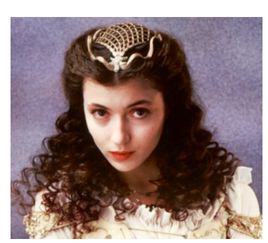
“Oh look — a pretty kitty!”

“Look — there’s another one!”

syntax, semantics

You know how to put words together to ask questions (syntax)

This kitty was bought as a present for someone.



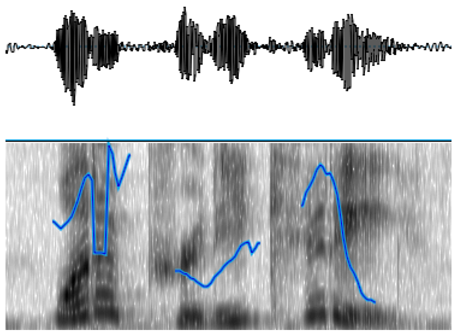
Lily thinks this kitty is pretty.

☹

Who does Lily think the kitty for is pretty?

Wait…what exactly do you know when you know a language?

A lot!



what a pretty kitty!

speech segmentation

KI tty

ki TTY

metrical phonology

Noun

penguin

owl

kitty

syntactic categorization

Who does Lily think the kitty for is pretty? ☹

syntax

"Oh look — a pretty kitty!"
"Look — there's another one!"

syntax, semantics

You know how to identify the right interpretation in context (pragmatics)

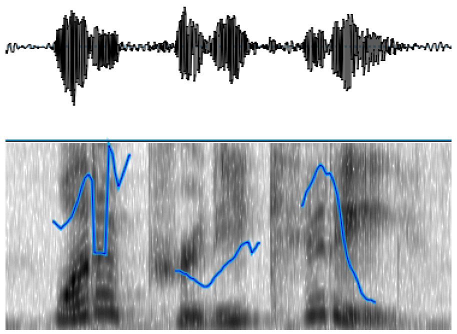
"Every kitty didn't sit on the stairs"

No kitties sat on the stairs.

Not all kitties sat on the stairs.

Wait…what exactly do you know when you know a language?

A lot!



what a pretty kitty!

speech segmentation

KI tty

ki TTY

metrical phonology

Noun

penguin owl

kitty

syntactic categorization

Who does Lily think the kitty for is pretty?

syntax

“Oh look — a pretty kitty!”

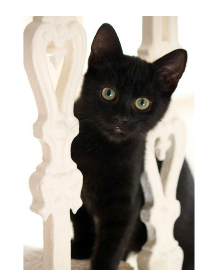
“Look — there’s another one!”

syntax, semantics

“Every kitty didn’t sit on the stairs”

Not all kitties sat on the stairs.

pragmatics



Wait…what exactly do you know when you know a language?

A lot!

metrical phonology

speech segmentation

syntactic categorization

pragmatics

syntax

syntax, semantics

So how exactly do children learn all this?

So how exactly do children learn all this?

We know they do it relatively quickly.

speech segmentation

metrical phonology

syntactic categorization

syntax

syntax, semantics

pragmatics

Much of the linguistic system is already known by **age 4**.

So how exactly do children learn all this?

They also don't seem to get a lot of explicit instruction. And when they do, they don't really pay attention to things that don't impact meaning.

(From Martin Braine)

Child: Want other one spoon, Daddy.
Father: You mean, you want the other spoon.
Child: Yes, I want other one spoon, please Daddy.
Father: Can you say "the other spoon"?
Child: Other…one…spoon.
Father: Say "other".
Child: Other.
Father: "Spoon."
Child: Spoon.
Father: "Other spoon."
Child: Other…spoon. **Now give me other one spoon**?

So how exactly do children learn all this?

They also don't seem to get a lot of explicit instruction. And when they do, they don't really pay attention to things that don't impact meaning.

What they're doing: **Extracting patterns** and **making generalizations** from the surrounding data mostly just by hearing examples of what's allowed in the language.

So how exactly do children learn all this?

What they're doing: **Extracting patterns** and **making generalizations** from the surrounding data mostly just by hearing examples of what's allowed in the language.

What's so hard about that?

So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

???

"birdie"

"What a pretty birdie!"

So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

???

"birdie"

"Look - a birdie!"

So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

???

"birdie"

"Look at that birdie!"

So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

How to generalize beyond the input?

???

"birdie"

So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

One hypothesis

+blue

"birdie"

So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

Another hypothesis

+on branch

"birdie"

So how exactly do children learn all this?

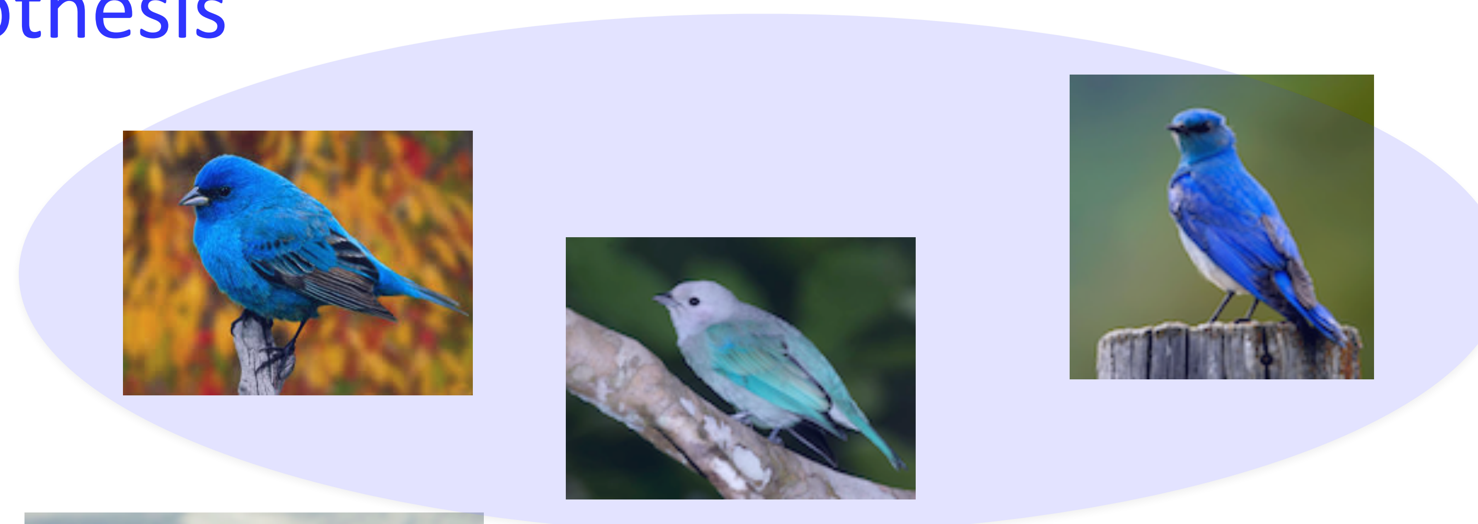
What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

The right hypothesis

+bird

"birdie"



So how exactly do children learn all this?

What's so hard about that?

There are often many ways to generalize beyond the input, and most of them aren't right.

speech segmentation

metrical phonology

syntactic categorization

syntax

syntax, semantics

pragmatics

These kind of **induction problems** are everywhere in cognitive development, including language acquisition.

Language acquisition = Solving a lot of induction problems.

Language acquisition = Solving a lot of induction problems.

We can also think about this as an information processing task.

Language acquisition = Solving a lot of induction problems.

We can also think about this as 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 = Solving a lot of induction problems.

We can also think about this as 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 = Solving a lot of induction problems.

We can also think about this as an information processing task.

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

processing & generalization

words & morphemes

metrical phonology

syntactic categories

syntax

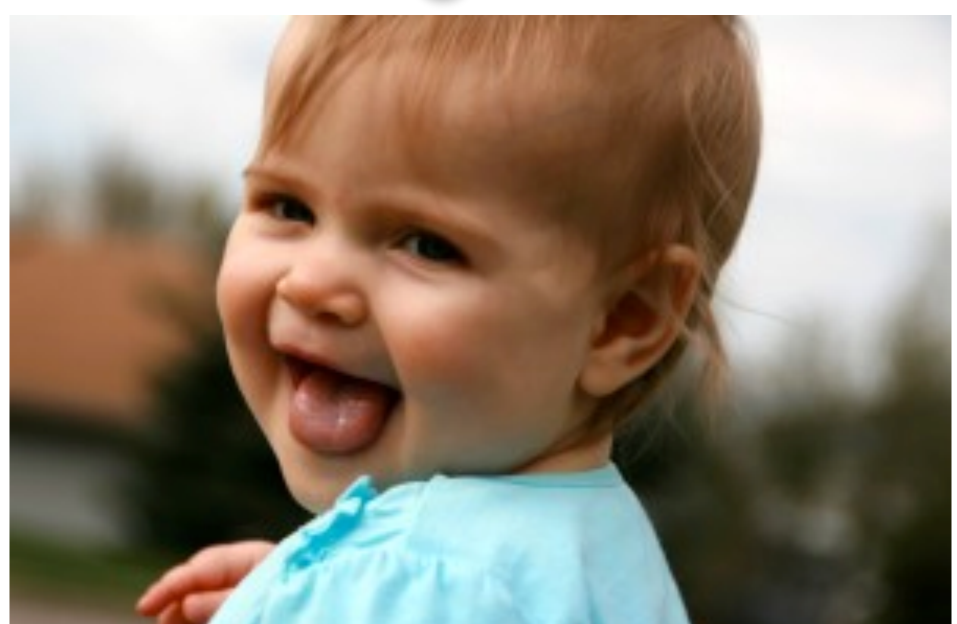
semantics

pragmatics

Look at that kitty!
There's another one.

Input

Where did he hide?
What happened?



Language acquisition = Solving a lot of induction problems.

We can also think about this as 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

Look at that kitty!
There's another one.

Input

Where did he hide?
What happened?

processing & generalization

words & morphemes

metrical phonology

syntactic categories

syntax

semantics

pragmatics

Where's the kitty?

That one's really cute.

Language acquisition = Solving a lot of induction problems.

We can also think about this as an information processing task.

To understand how children solve the acquisition task, we need theories of representation and theories of development.

processing & generalization

words & morphemes

metrical phonology

syntactic categories

syntax

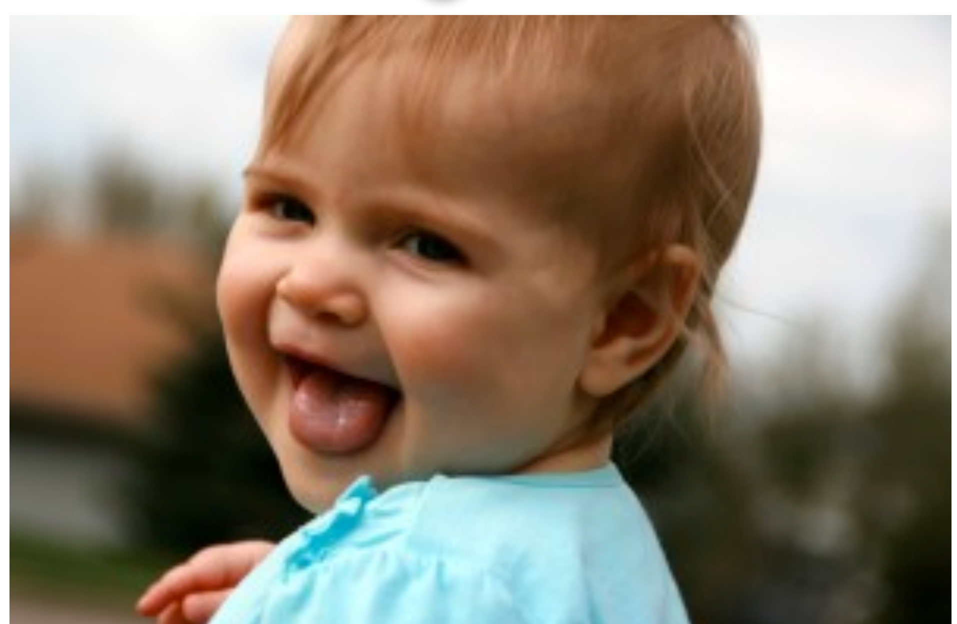
semantics

pragmatics

Look at that kitty!
There's another one.

Input

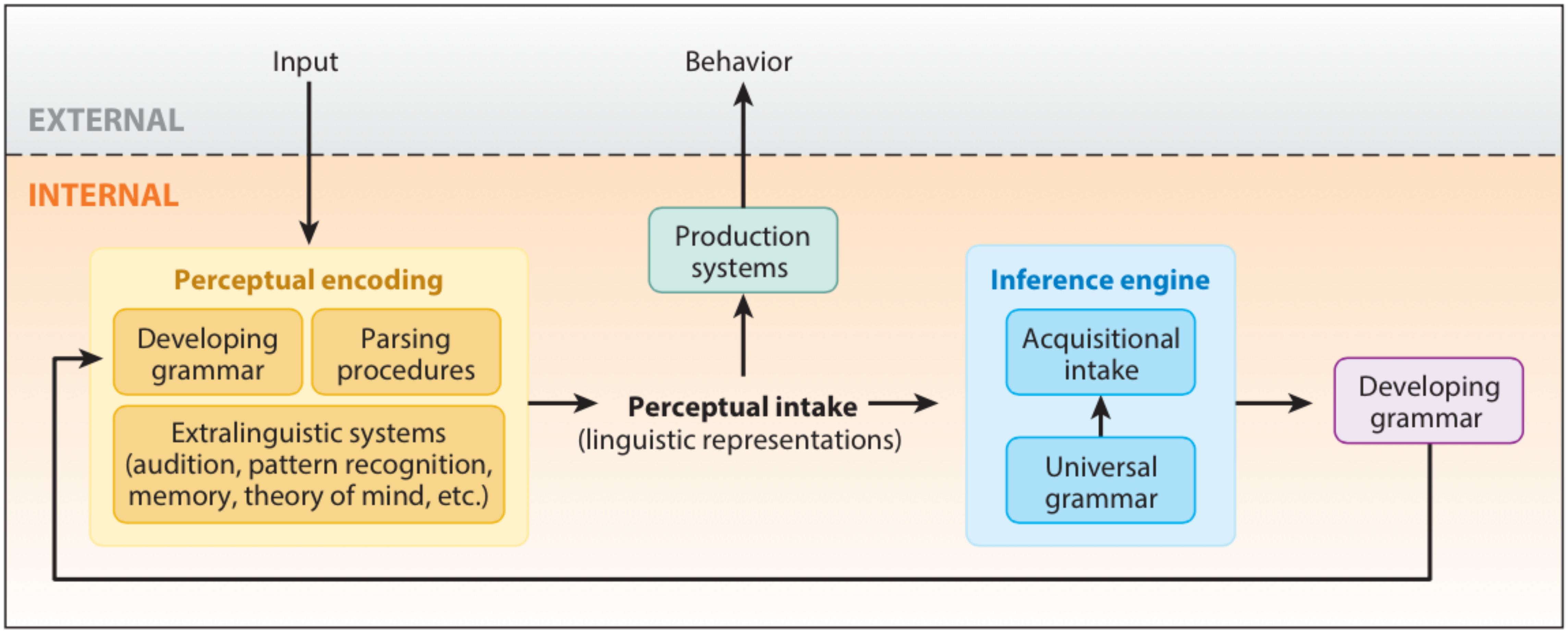
Where did he hide?
What happened?



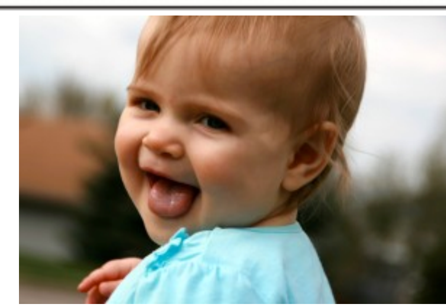
Where's the kitty?

That one's really cute.

Language acquisition = Solving a lot of induction problems.



Lidz & Gagliardi 2015



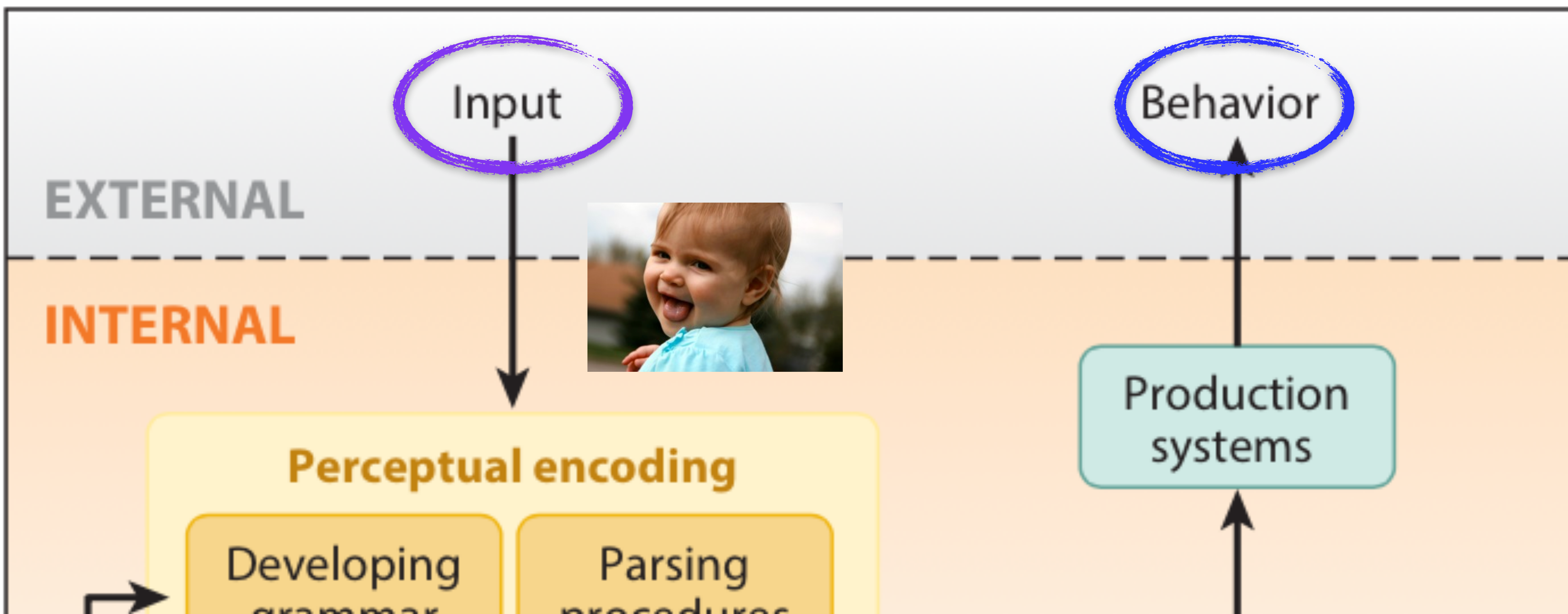
A framework that makes components of the acquisition task more explicit.

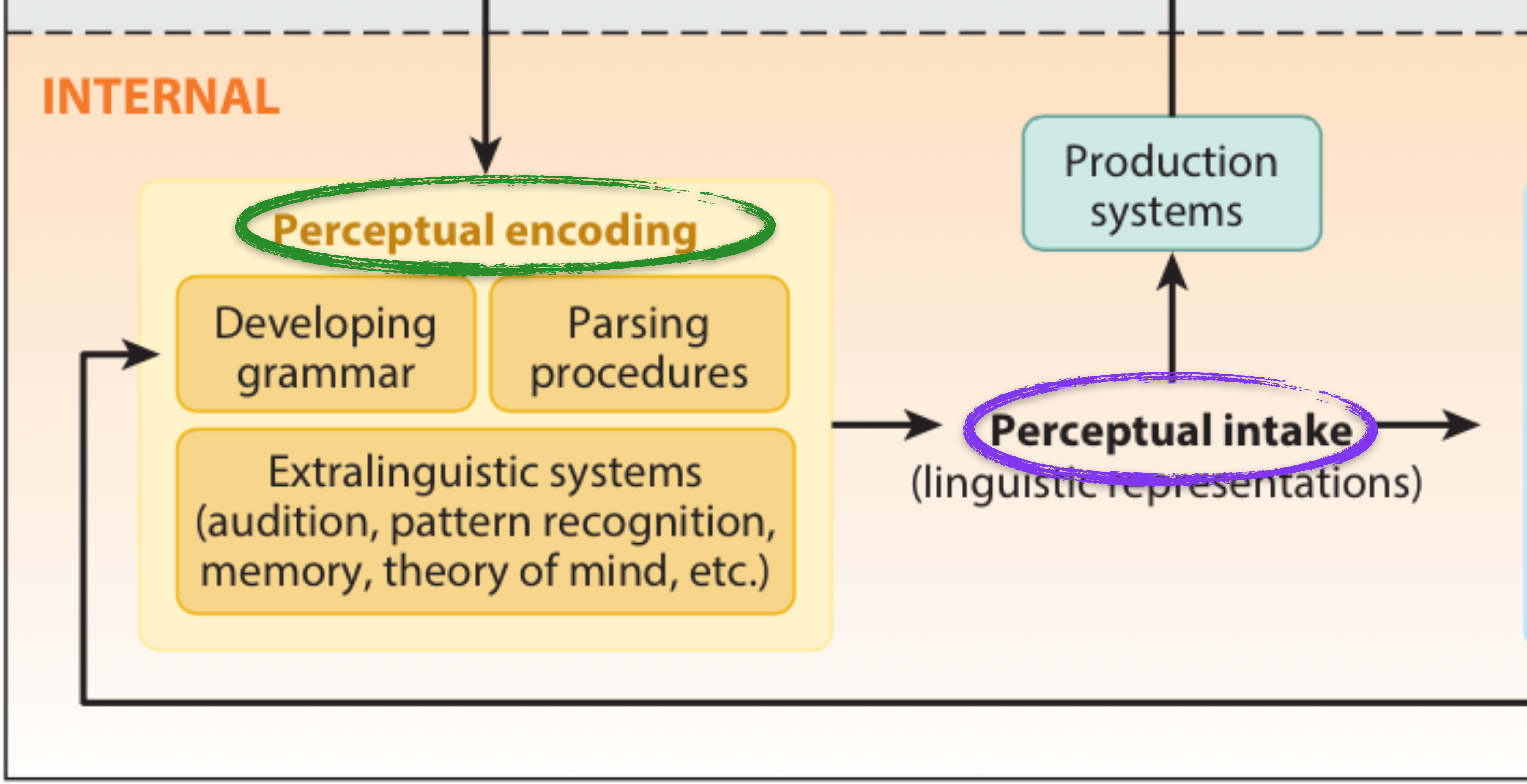
A framework that makes components of the acquisition task more explicit.

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

Experimental & Corpus methods

Experimental methods

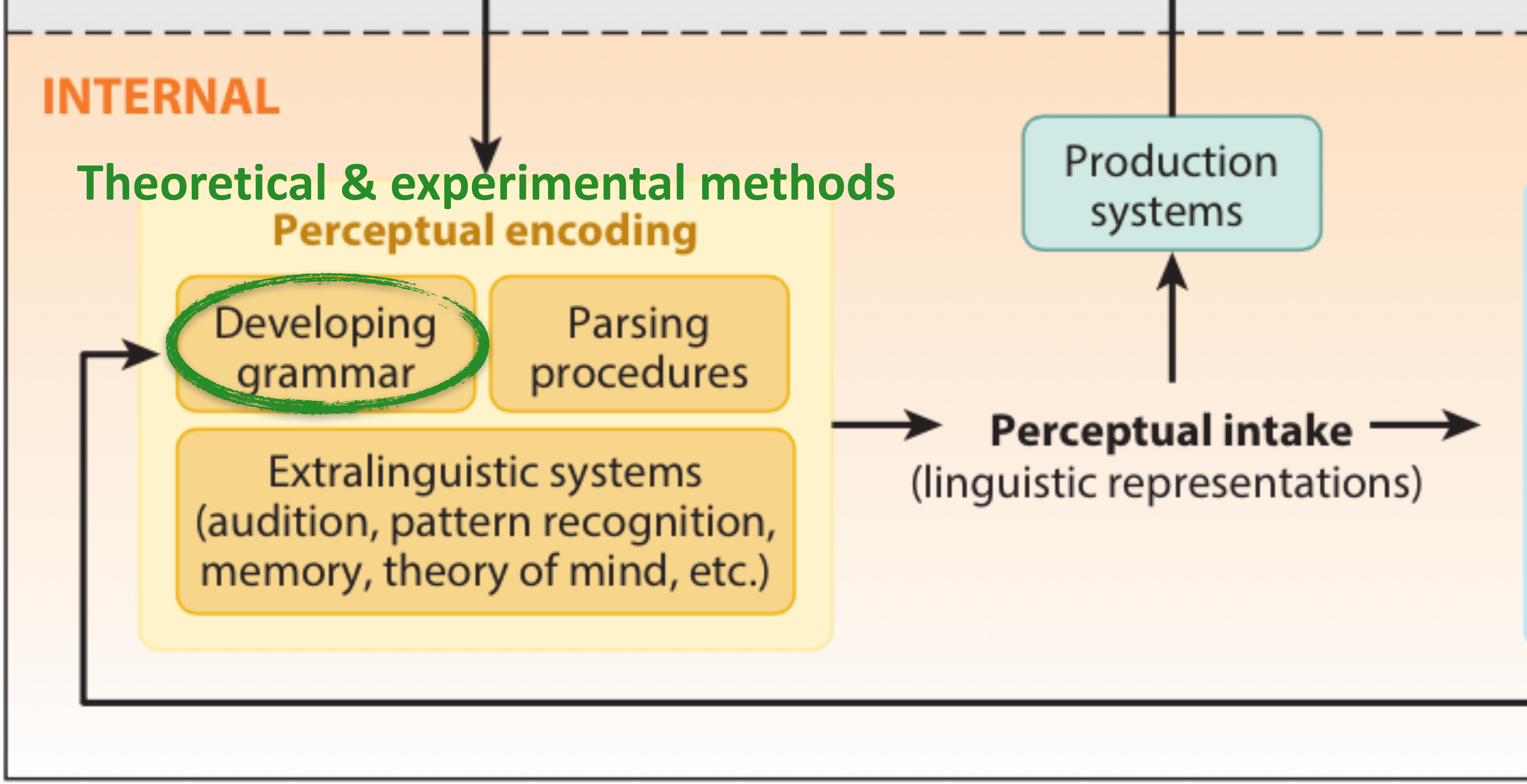




Lidz & Gagliardi 2015

Perceptual encoding:

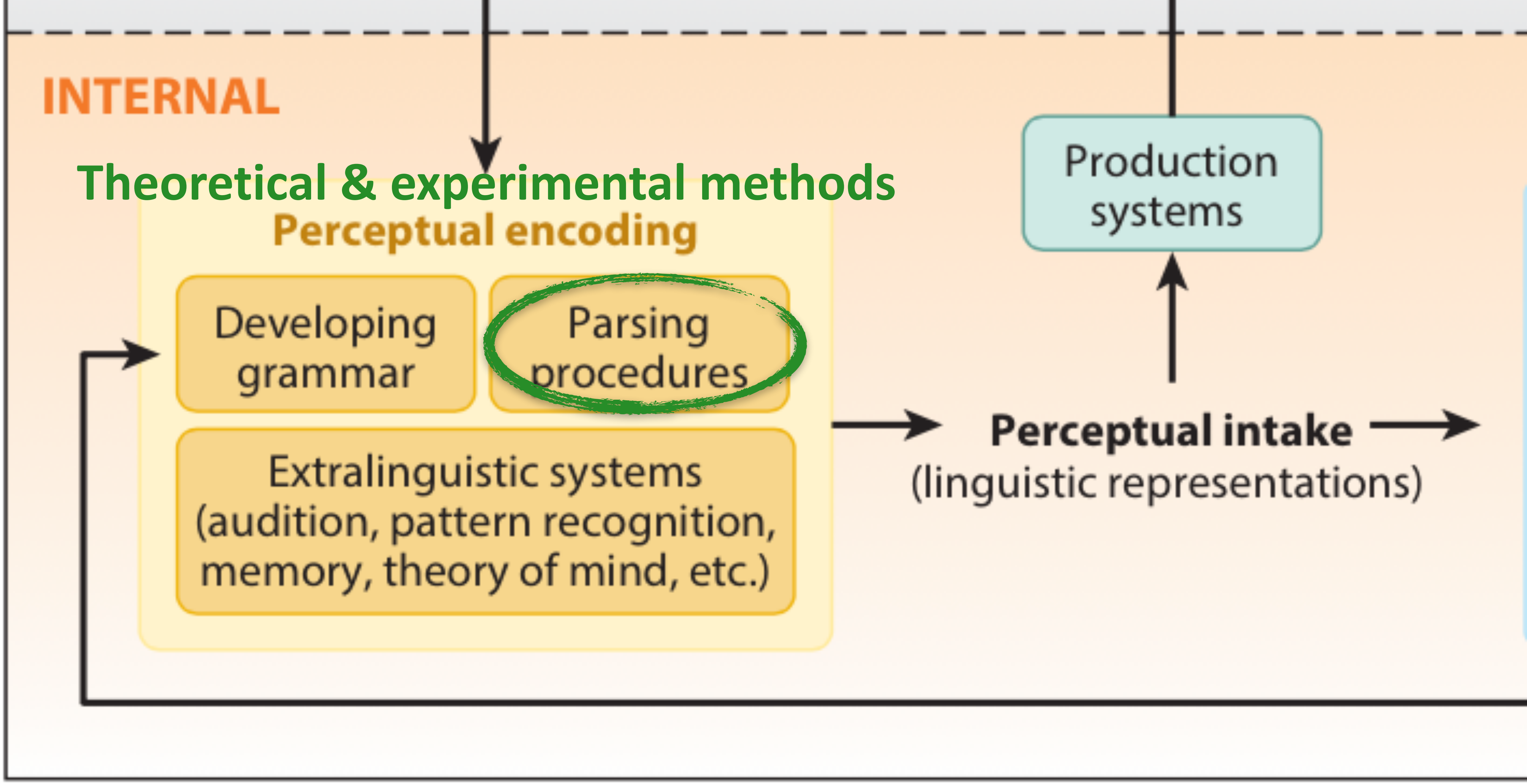
Turning the input signal into an internal linguistic representation = perceptual intake.



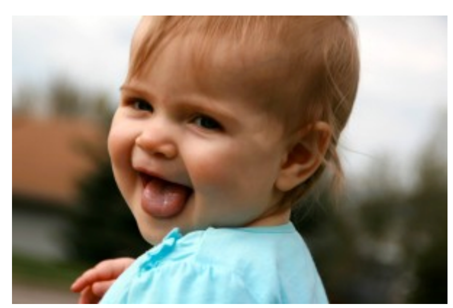
Lidz & Gagliardi 2015

Perceptual encoding:

Involves current grammar

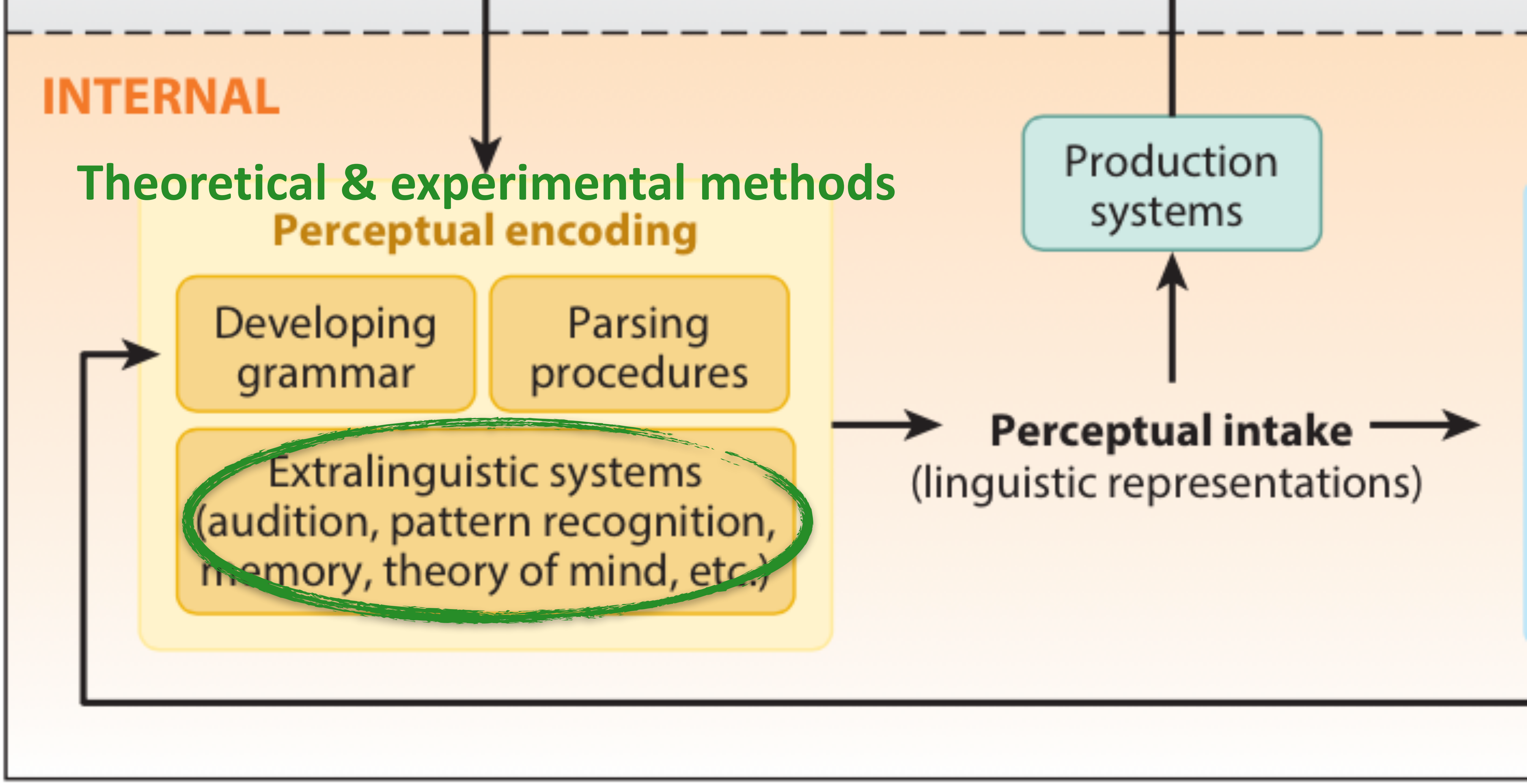


Lidz & Gagliardi 2015

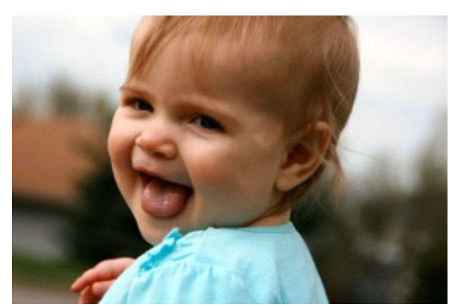


Perceptual encoding:

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

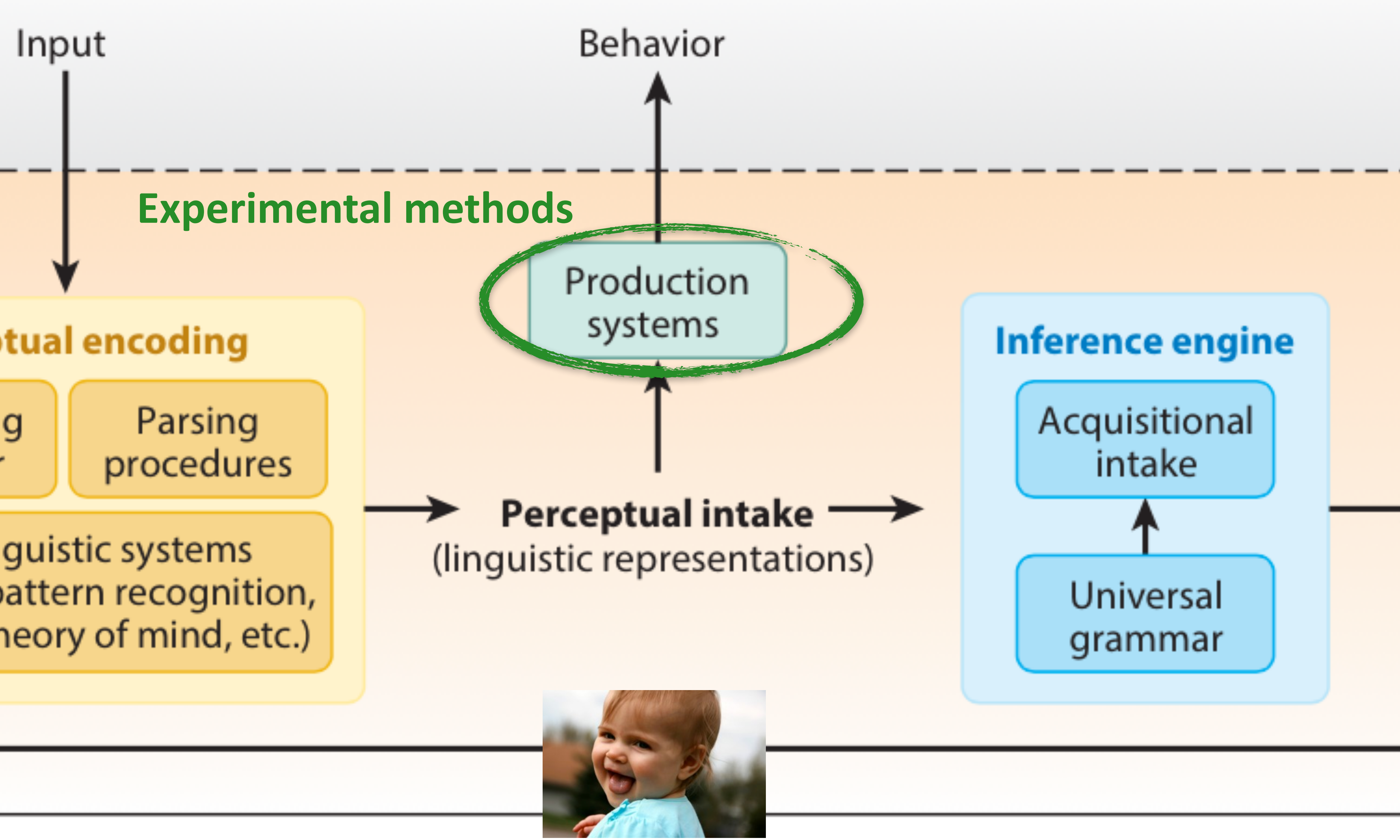


Lidz & Gagliardi 2015



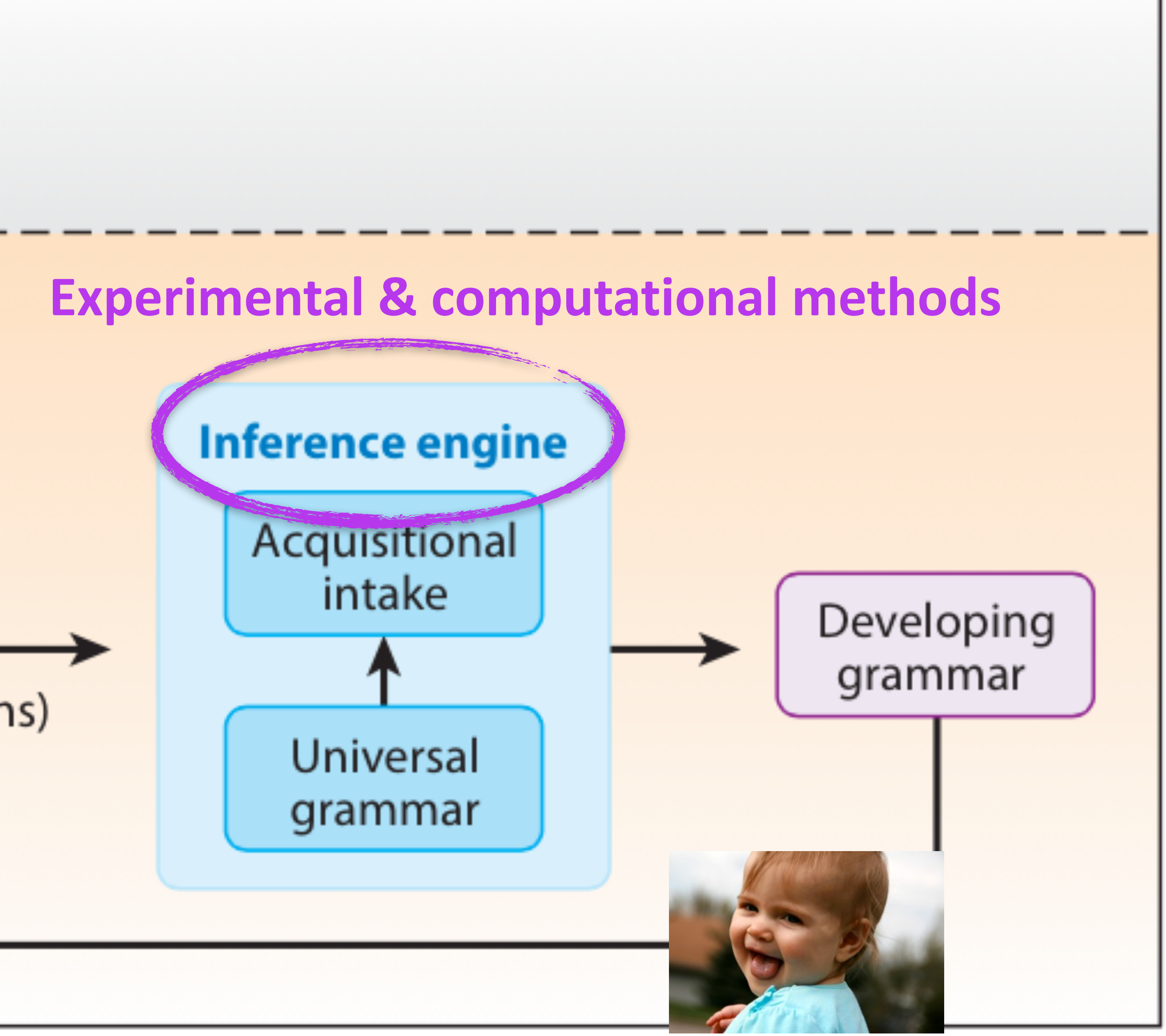
Perceptual encoding:

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



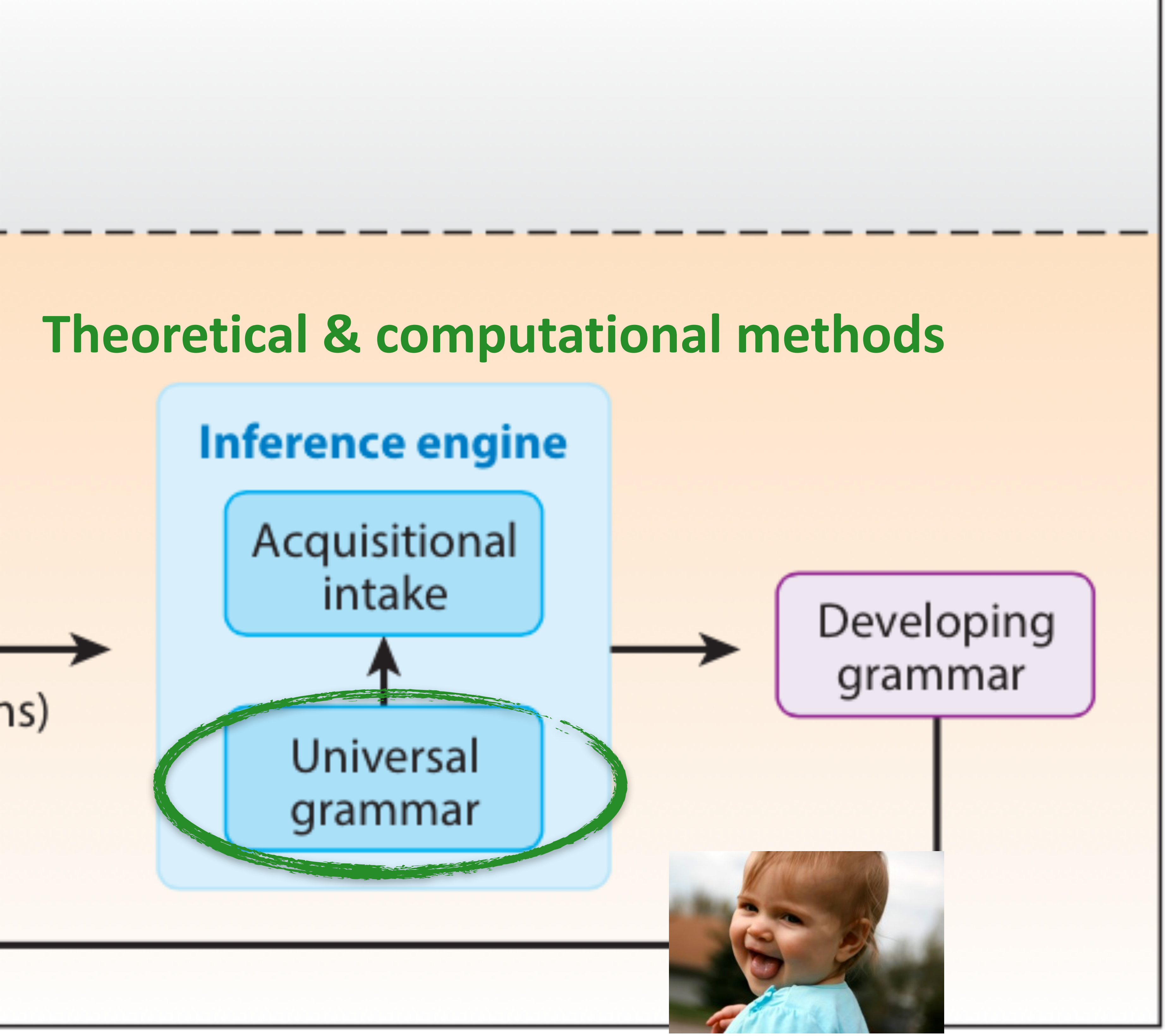
Generating observable behavior

Involves current linguistic representations being used by production systems.



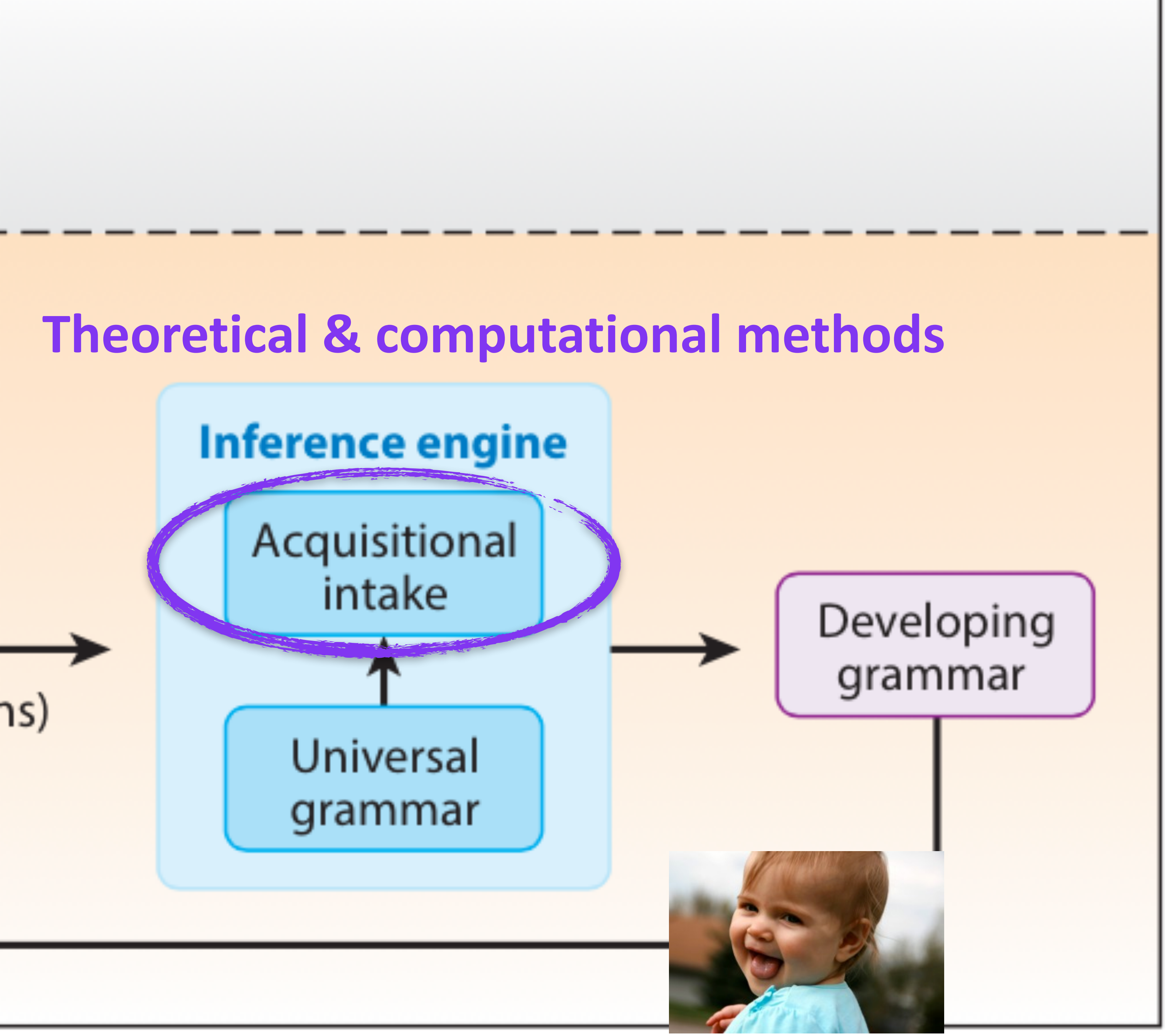
Doing inference

Generalization happens



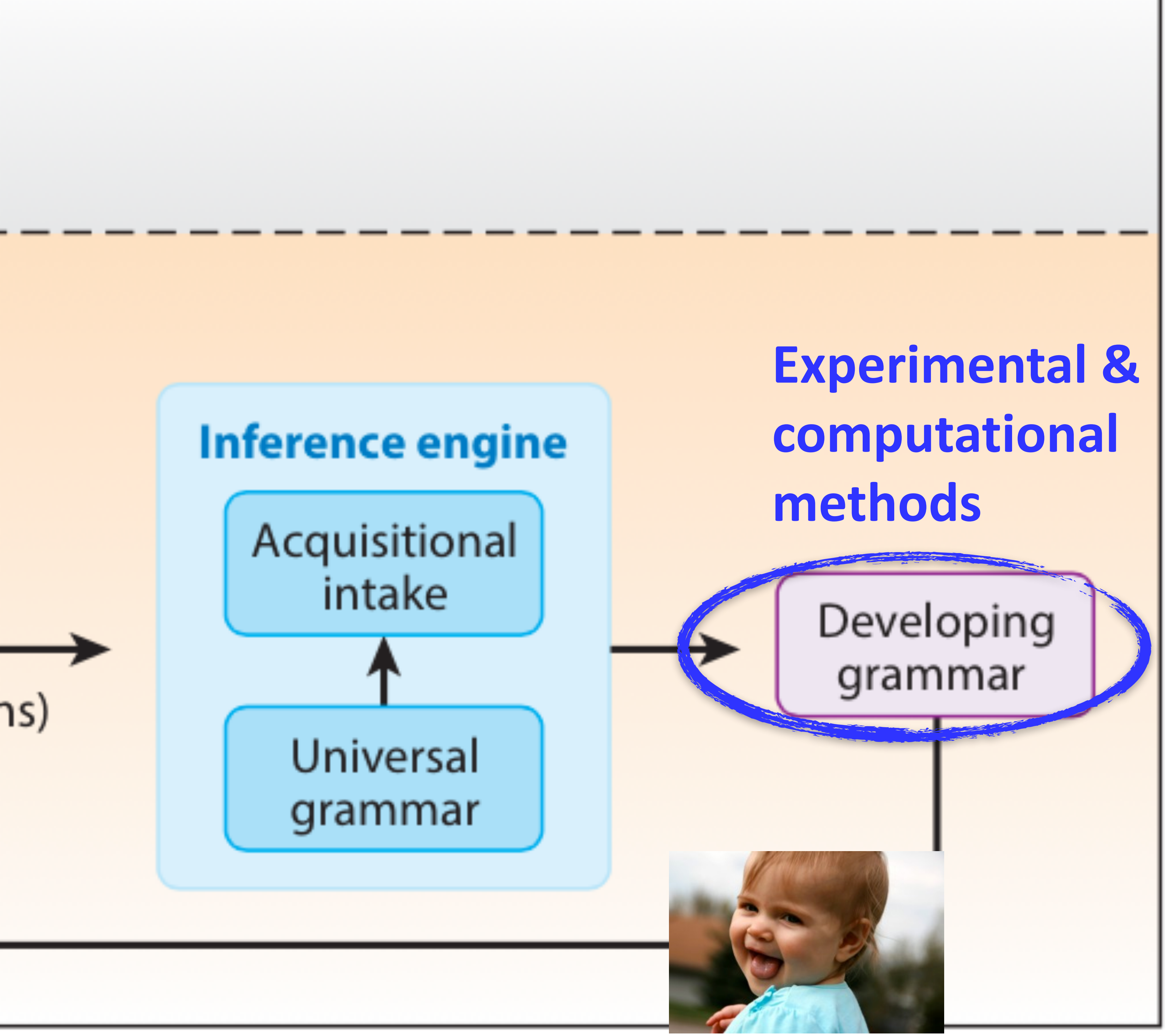
Doing inference

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



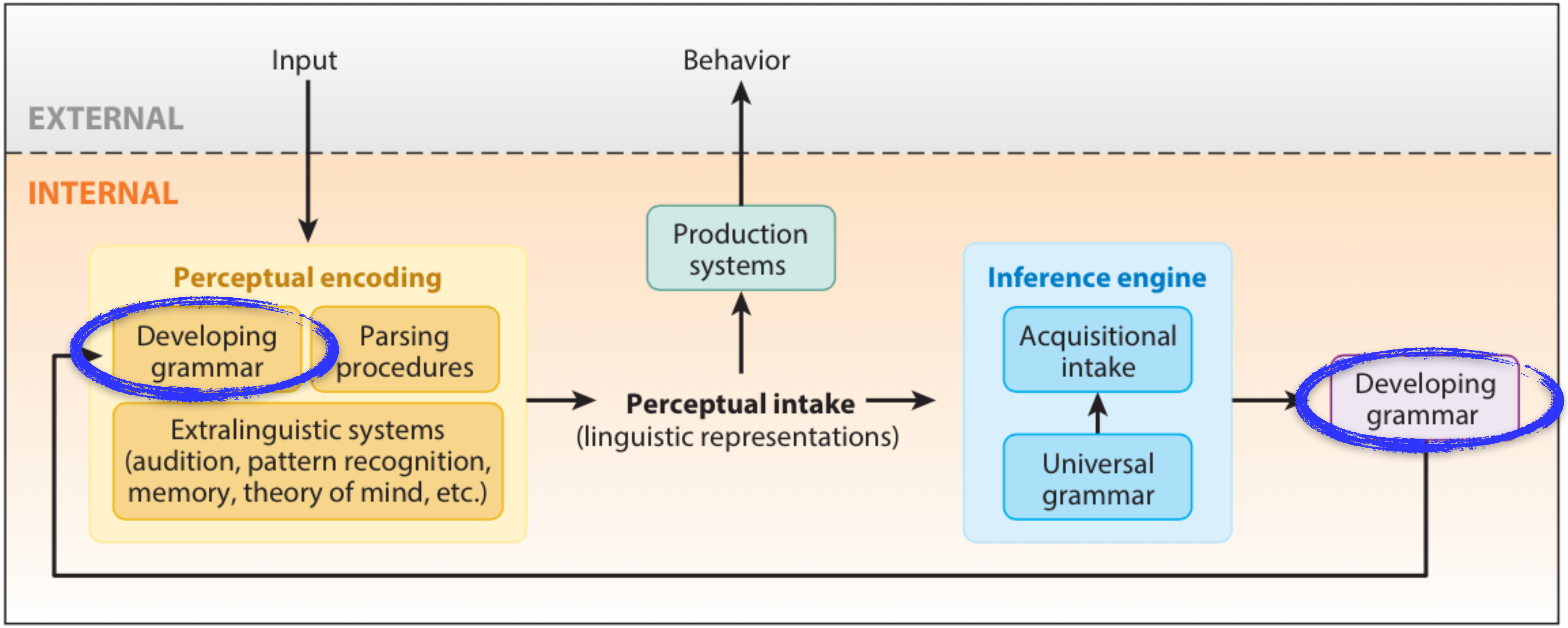
Doing inference

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



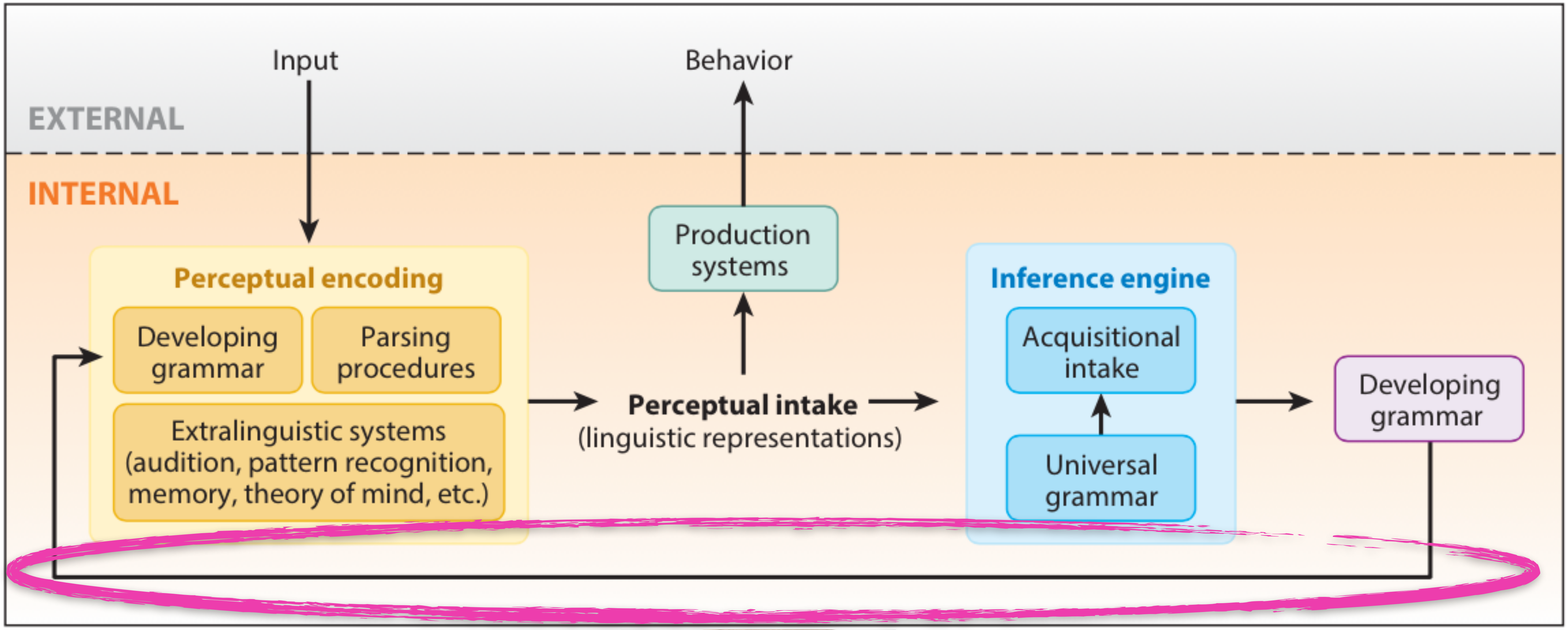
Doing inference

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

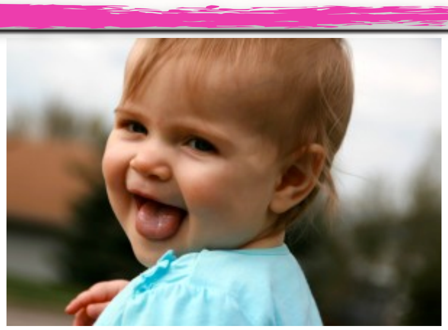


Lidz & Gagliardi 2015

The current linguistic hypotheses are
used in subsequent perceptual encoding



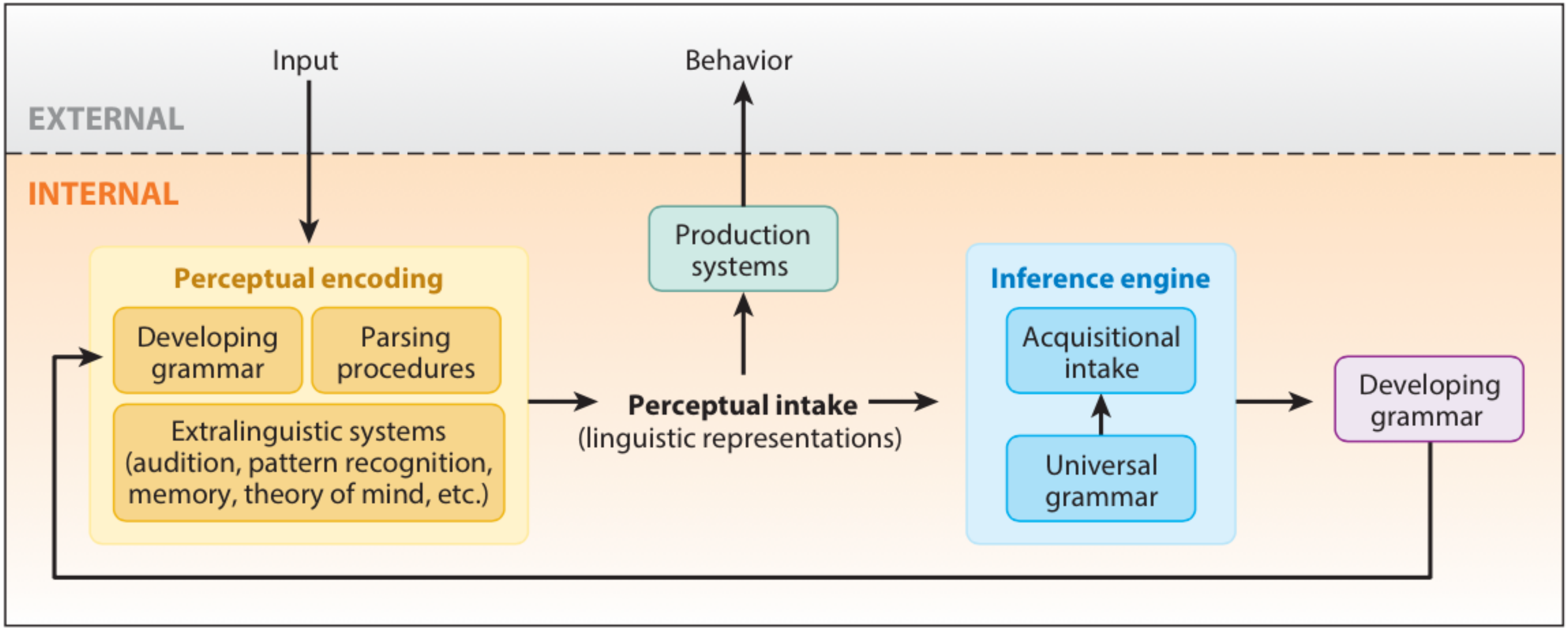
Lidz & Gagliardi 2015



Experimental methods

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

This is language acquisition



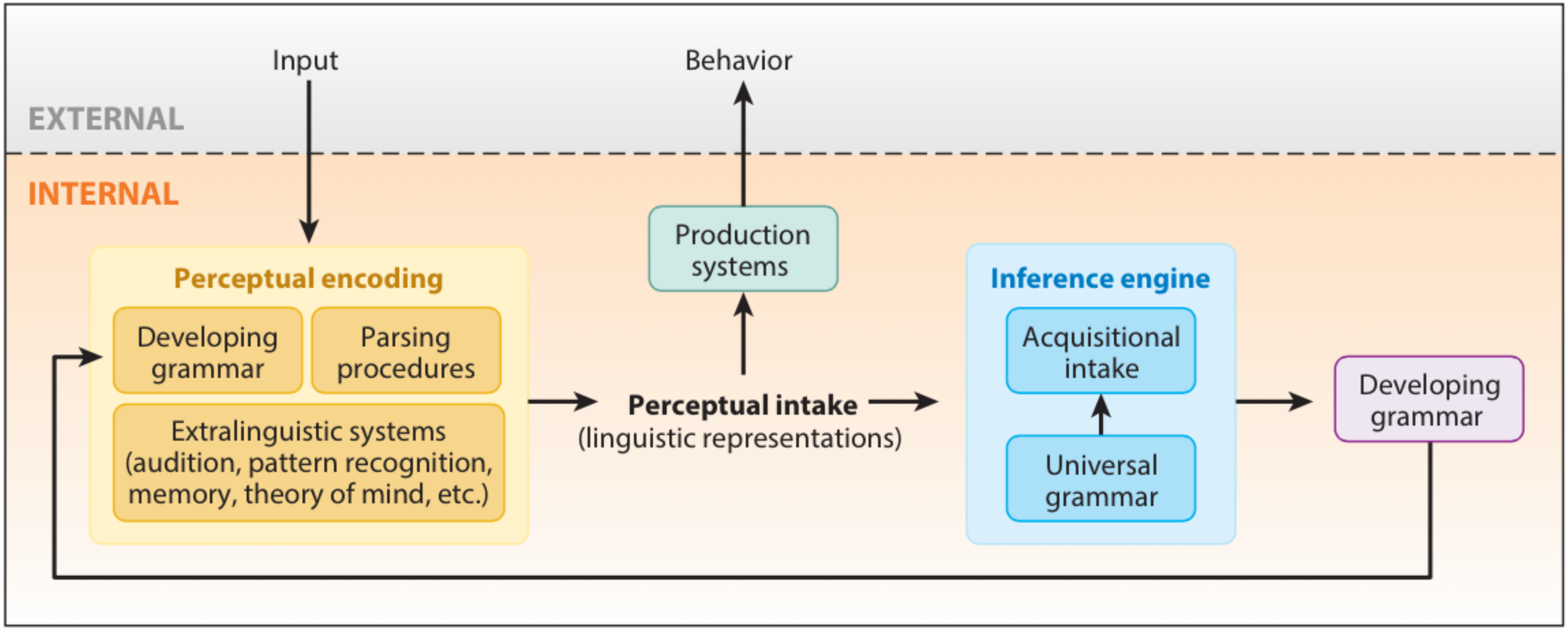
Lidz & Gagliardi 2015

Corpus Experimental

Theoretical Computational

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

This is language acquisition
…which involves solving induction problems

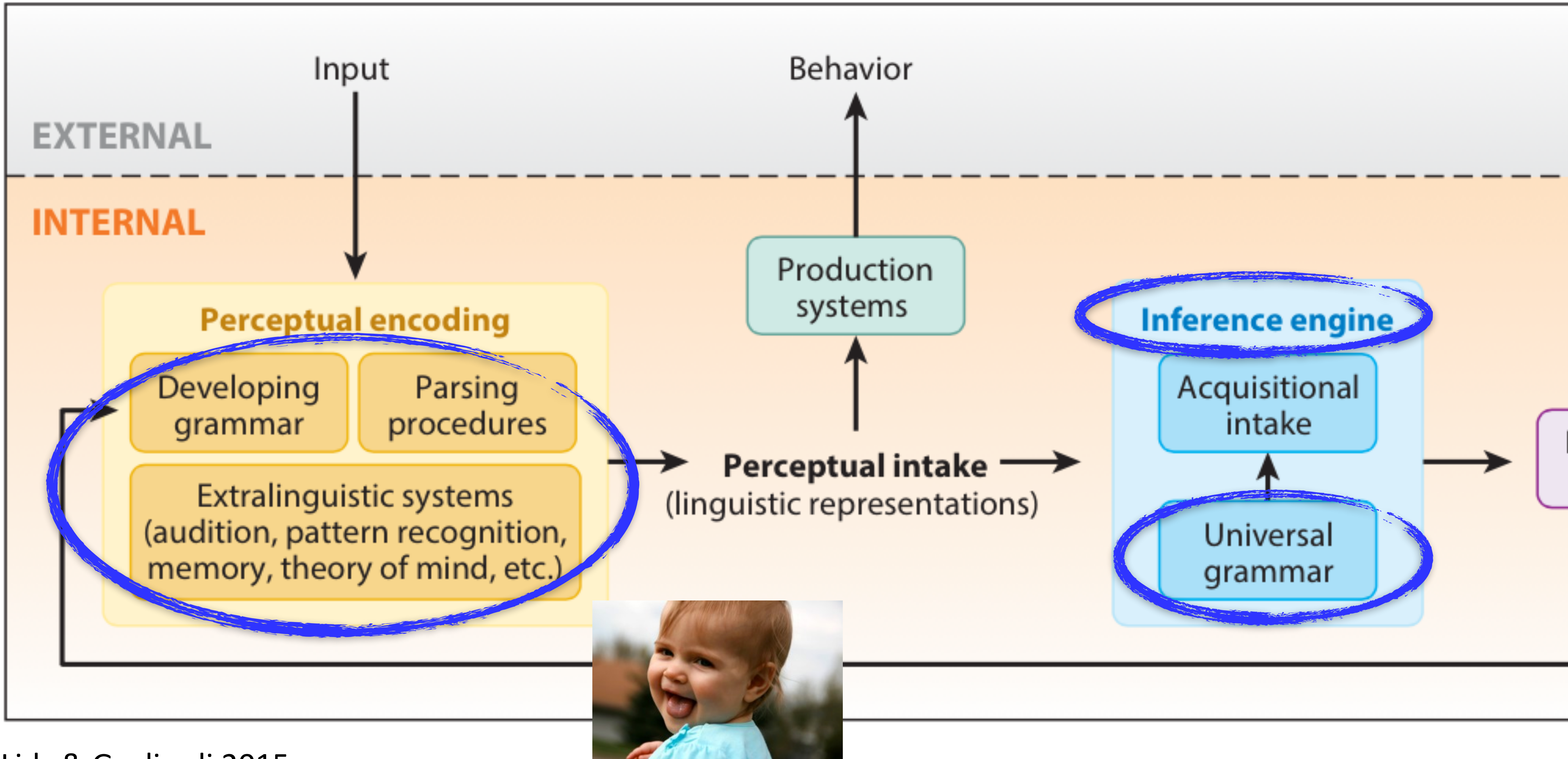


Lidz & Gagliardi 2015

Informative computational models = informative about
the learning strategies children use to solve **induction problems**

Learning strategies children use to solve **induction problems**

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

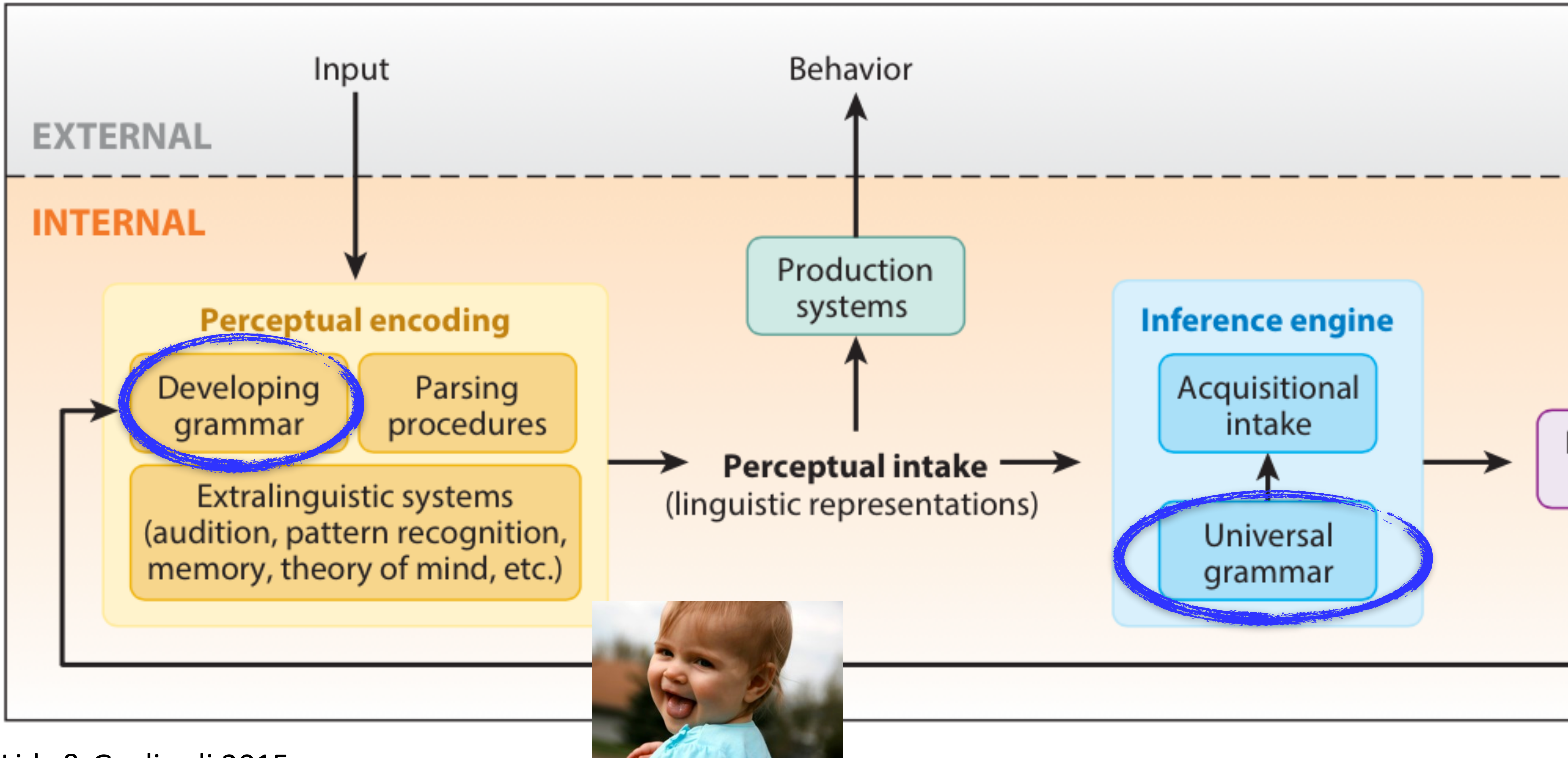


Lidz & Gagliardi 2015

Learning strategies children use to solve **induction problems**

Important learning strategy components include

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

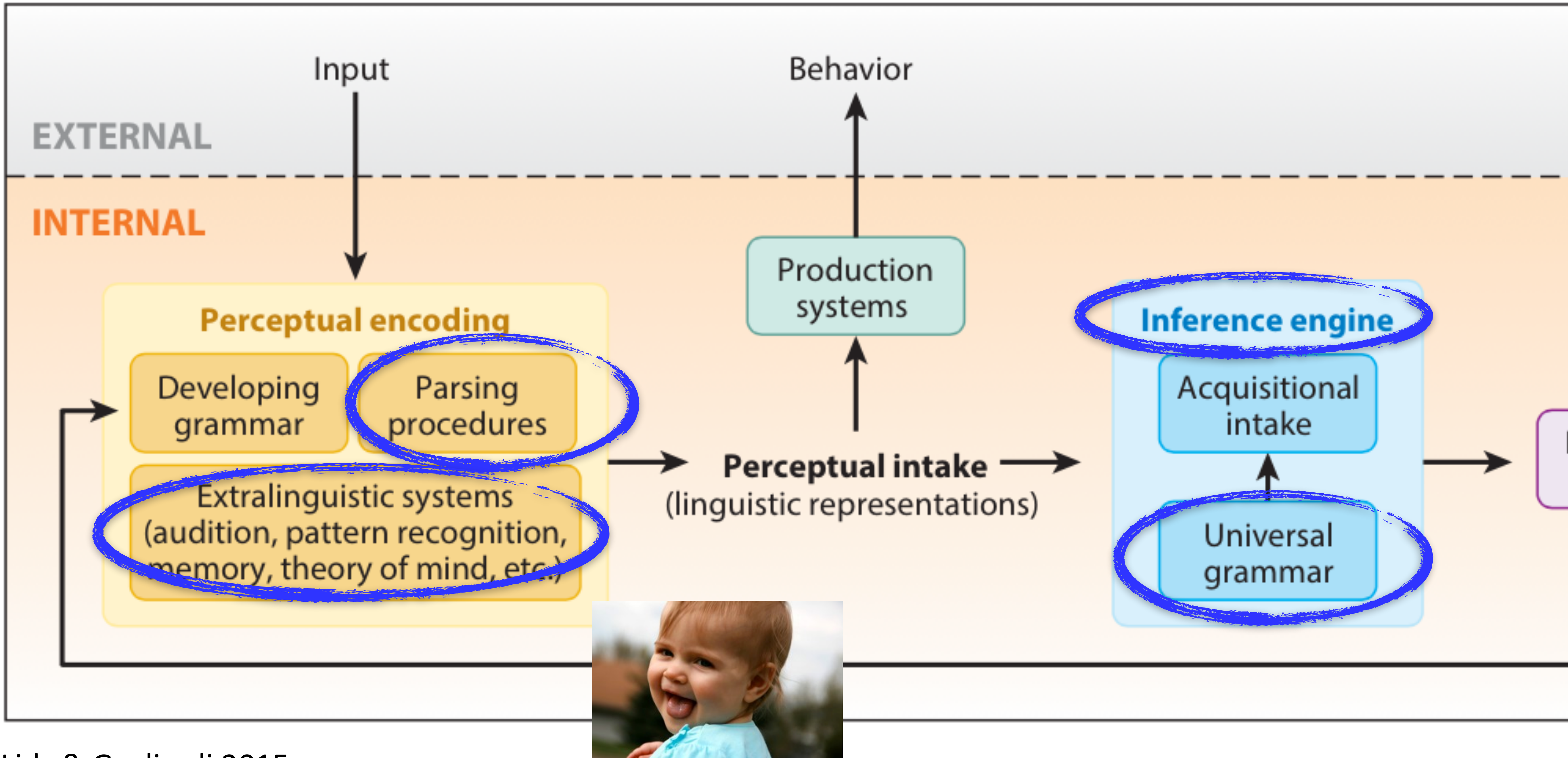


Lidz & Gagliardi 2015

Learning strategies children use to solve **induction problems**

Important learning strategy components include

- knowledge (= 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

Learning strategies children use to solve **induction problems**

And this is what we really want to know about!

And this is what we really want to know about!

Which learning strategies could children be using?

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

And this is what we really want to know about!

Which learning strategies could children be using?

Which learning biases are necessary?

(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)

And this is what we really want to know 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 in press, Pearl 2011, Pearl 2009)

And this is what we really want to know 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?

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

And this is what we really want to know 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?

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

And this is what we really want to know 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?

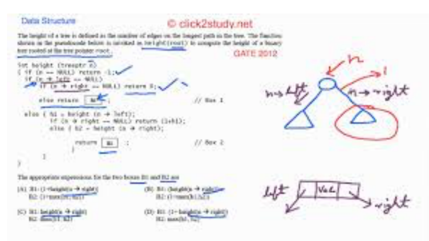
What factors affect children's observable behavior?

Why we do computational modeling: It can help us find out!

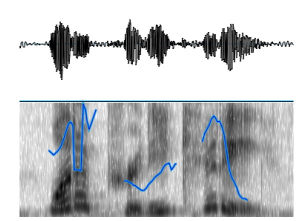
Today's Plan:
Computational models of language acquisition

I. Why

II. How



III. What we can learn



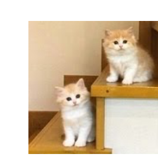
Noun

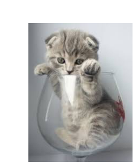
✓KI tty

Who does… is pretty?

another one

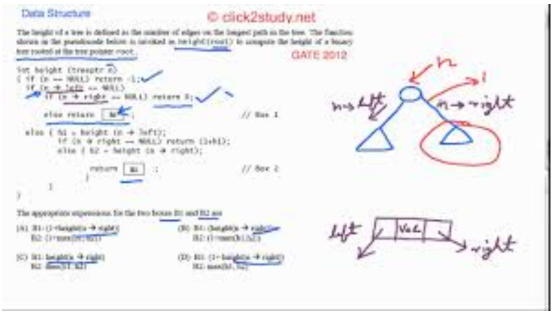
Every kitty didn't …





Today's Plan:
Computational models of language acquisition

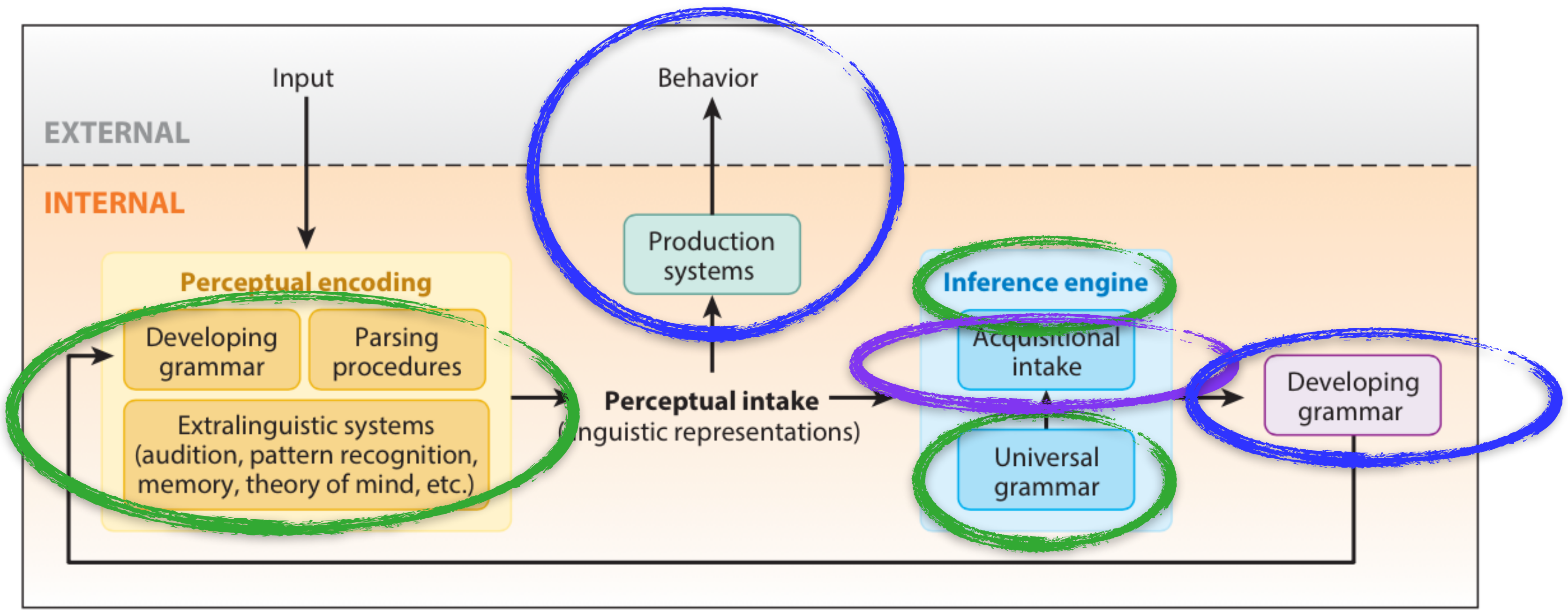
II. How



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?

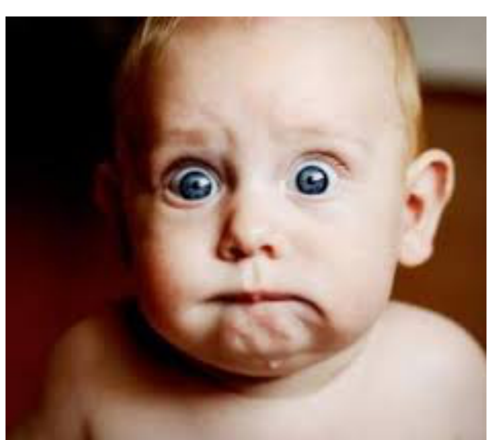
Computational-level

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?



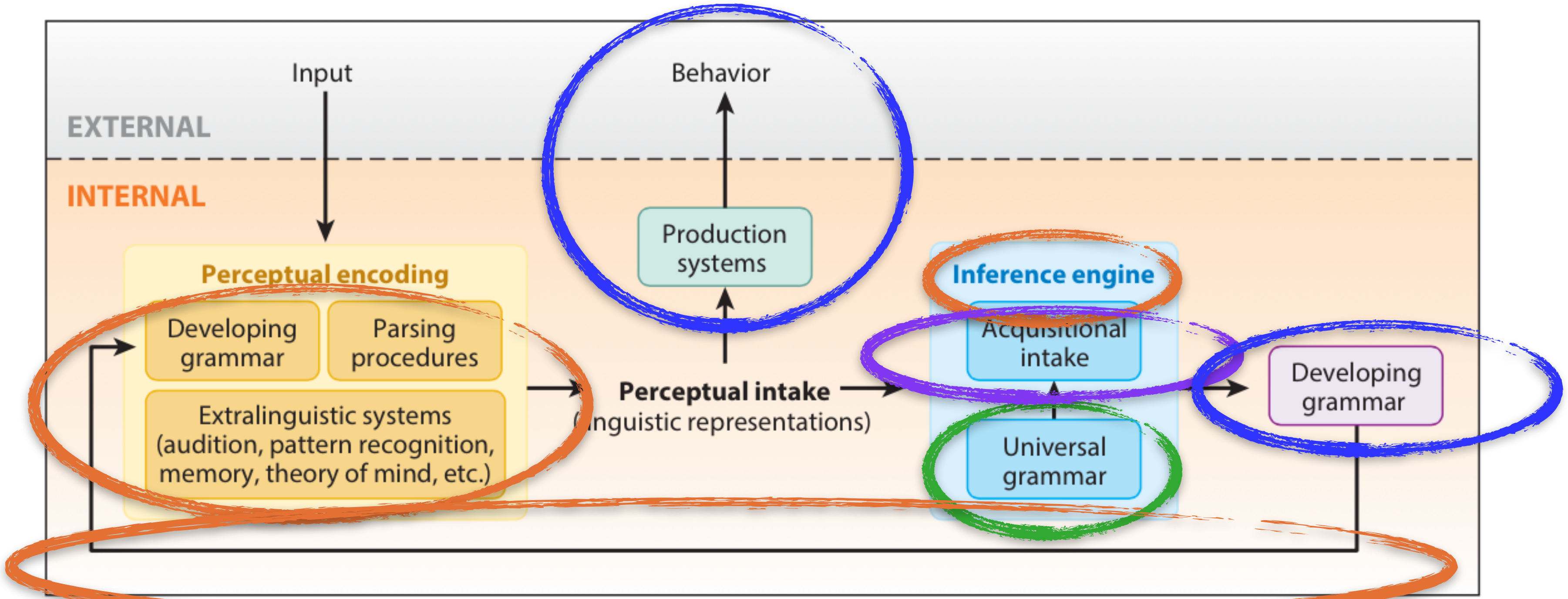
Because they're solving the wrong acquisition task…oops.

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?

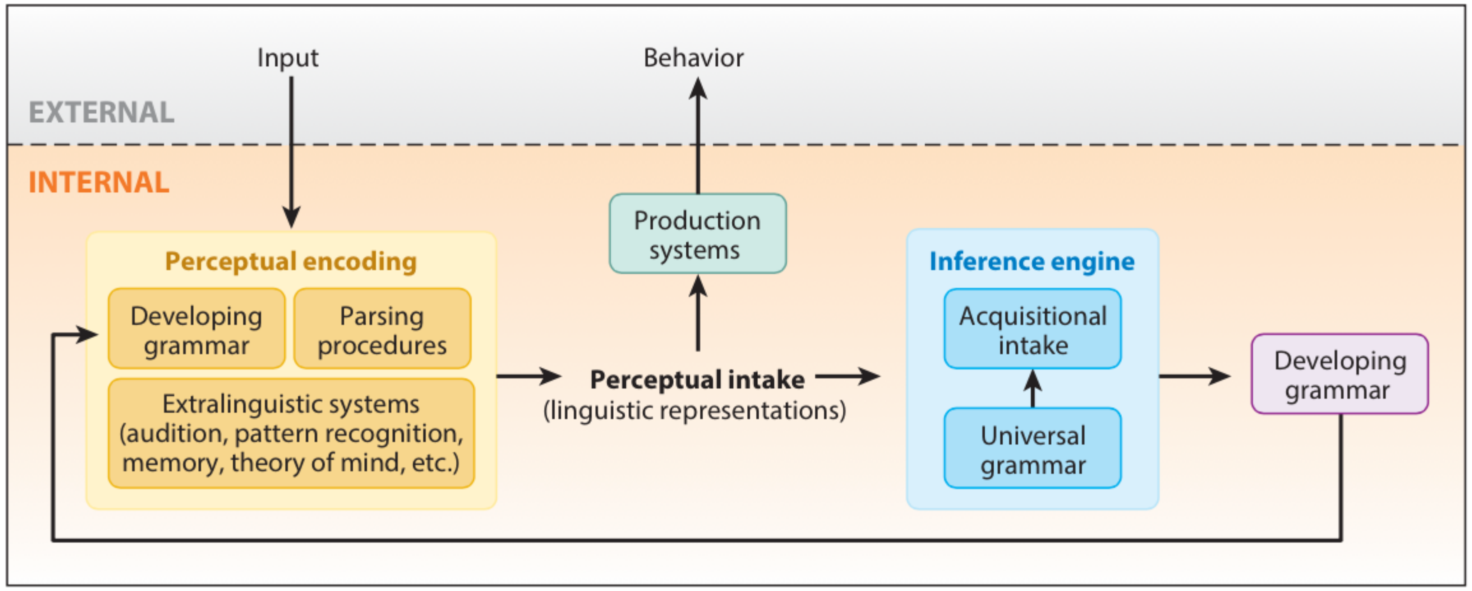
What level of model do you want to build?

Computational-level

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



Lidz & Gagliardi 2015

How do we model language acquisition?

What level of model do you want to build?

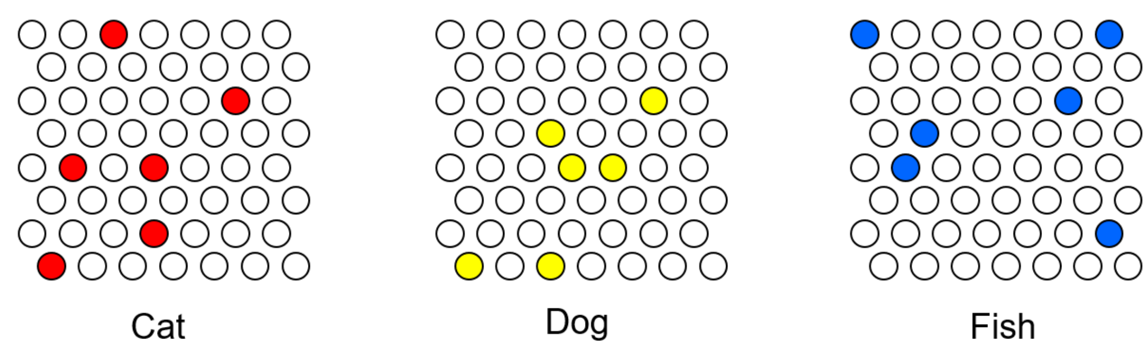
Computational-level

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.



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

How do we model language acquisition?

What level of model do you want to build?

The types I'll tell you about today

Computational-level

Algorithmic-level

Implementational-level

How do we model language acquisition?

Computational-level

So let's say you've figured out what level of model is appropriate to build. **Now what?**

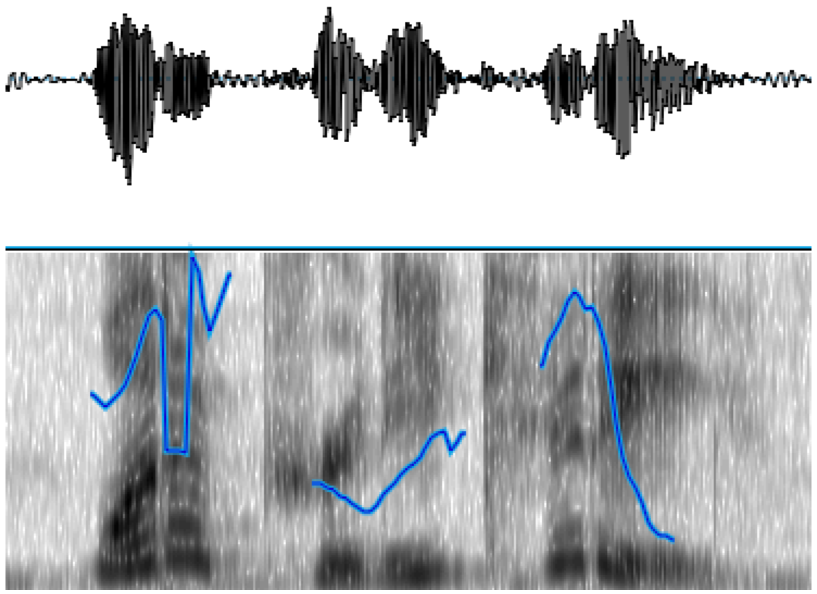
Time to actually build it!

Algorithmic-level

Let's look at an example with speech segmentation

How do we model language acquisition?

An example with speech segmentation



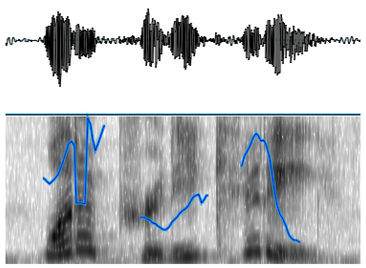
= wʌɾəpɹɪɾikɪɾi

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

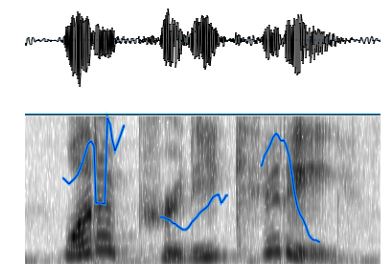
This depends on what task you're modeling

For the first stages of speech segmentation:

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

How do we model language acquisition?

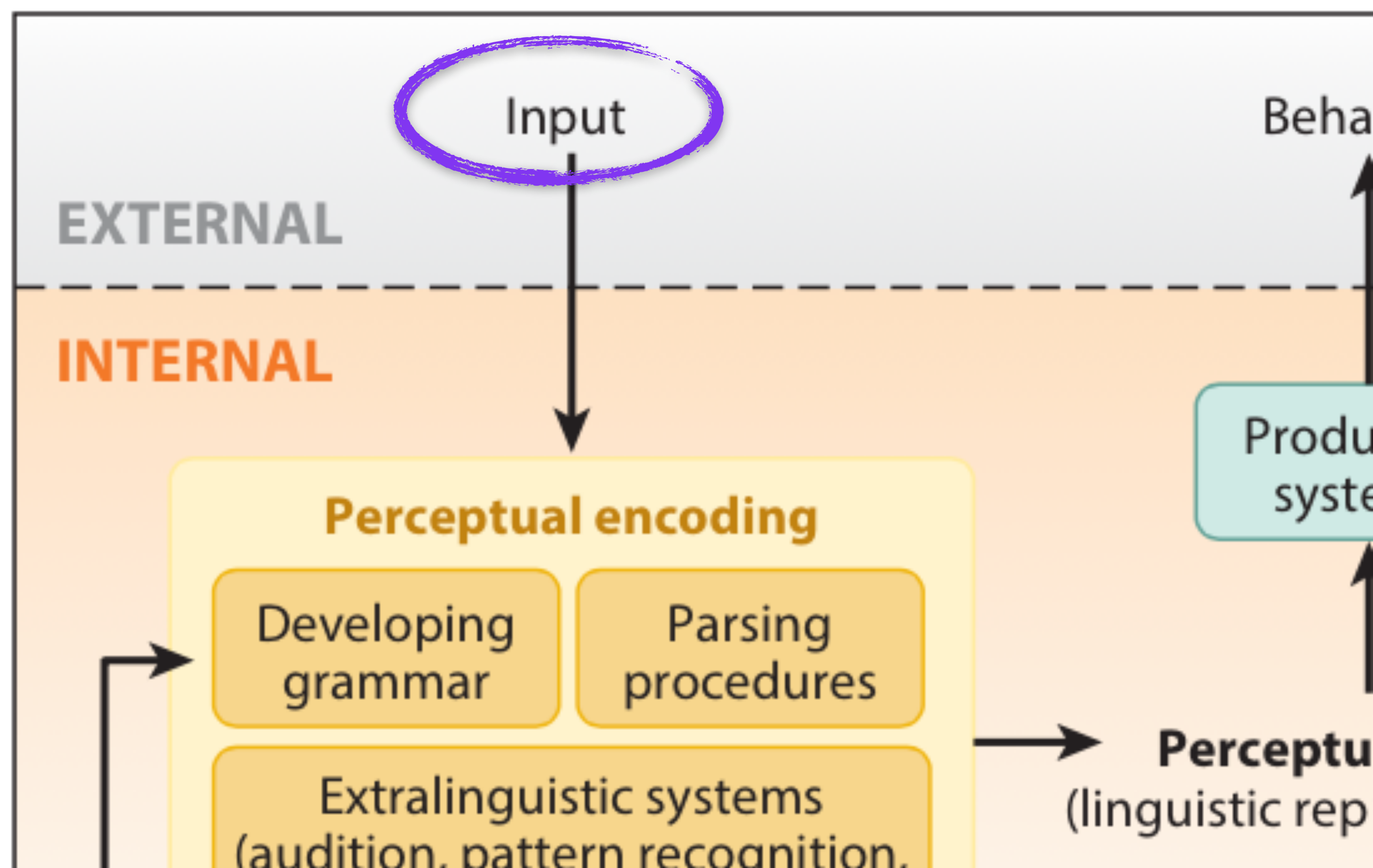
An example with speech segmentation



what a pretty kitty!

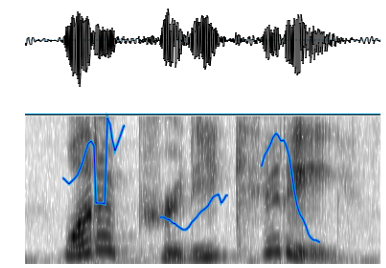
(2) Decide what data the child learns from (input)

This depends on your acquisition theory and the empirical data available



How do we model language acquisition?

An example with speech segmentation



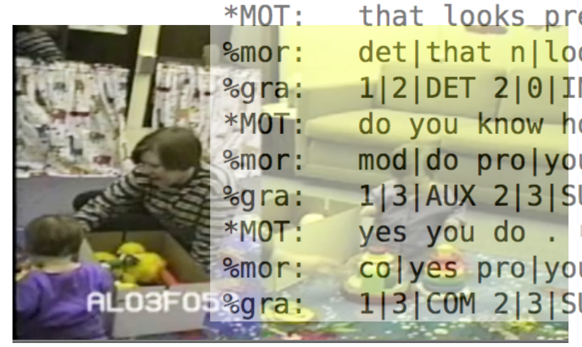
what a pretty kitty!

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

Example empirical data: CHILDES database
http://childes.talkbank.org

CHILDES Child Language Data Exchange System

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

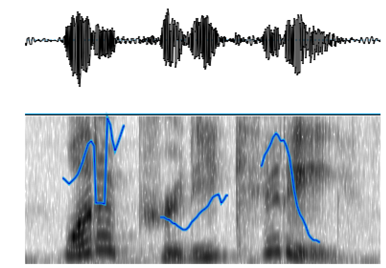



```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   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

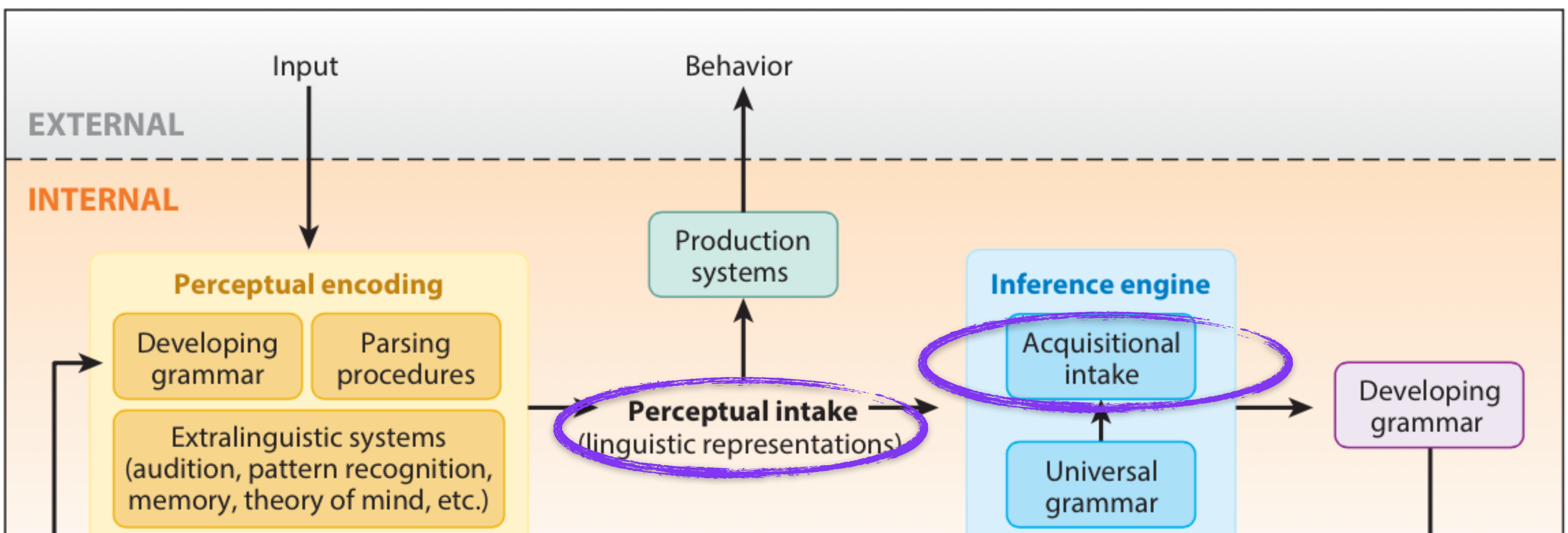

```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   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!

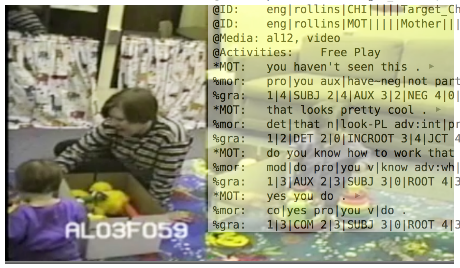
(3) Decide how the child perceives the data, and which data are relevant (intake)

This depends on your acquisition theory

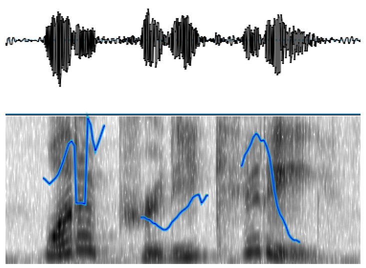


How do we model language acquisition?

An example with speech segmentation



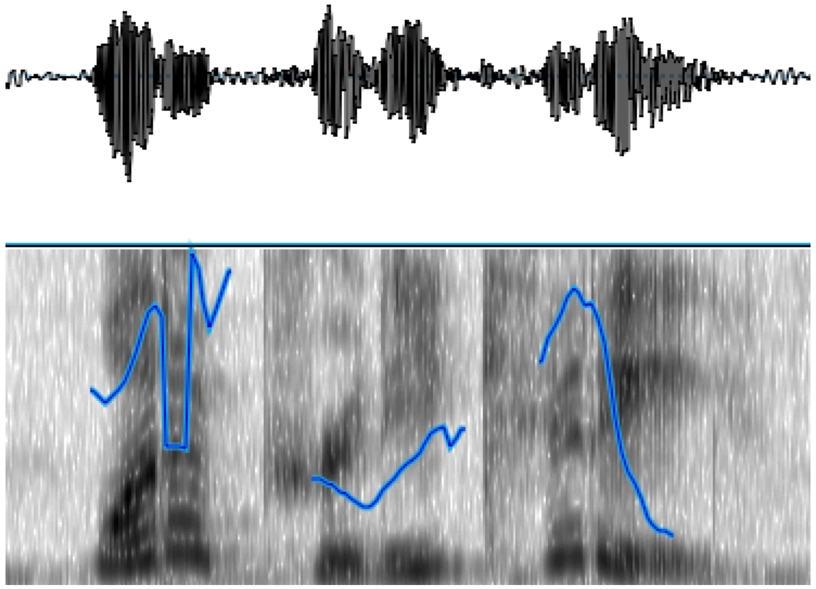

```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   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

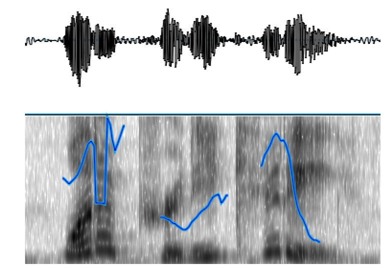
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

How do we model language acquisition?

An example with speech segmentation

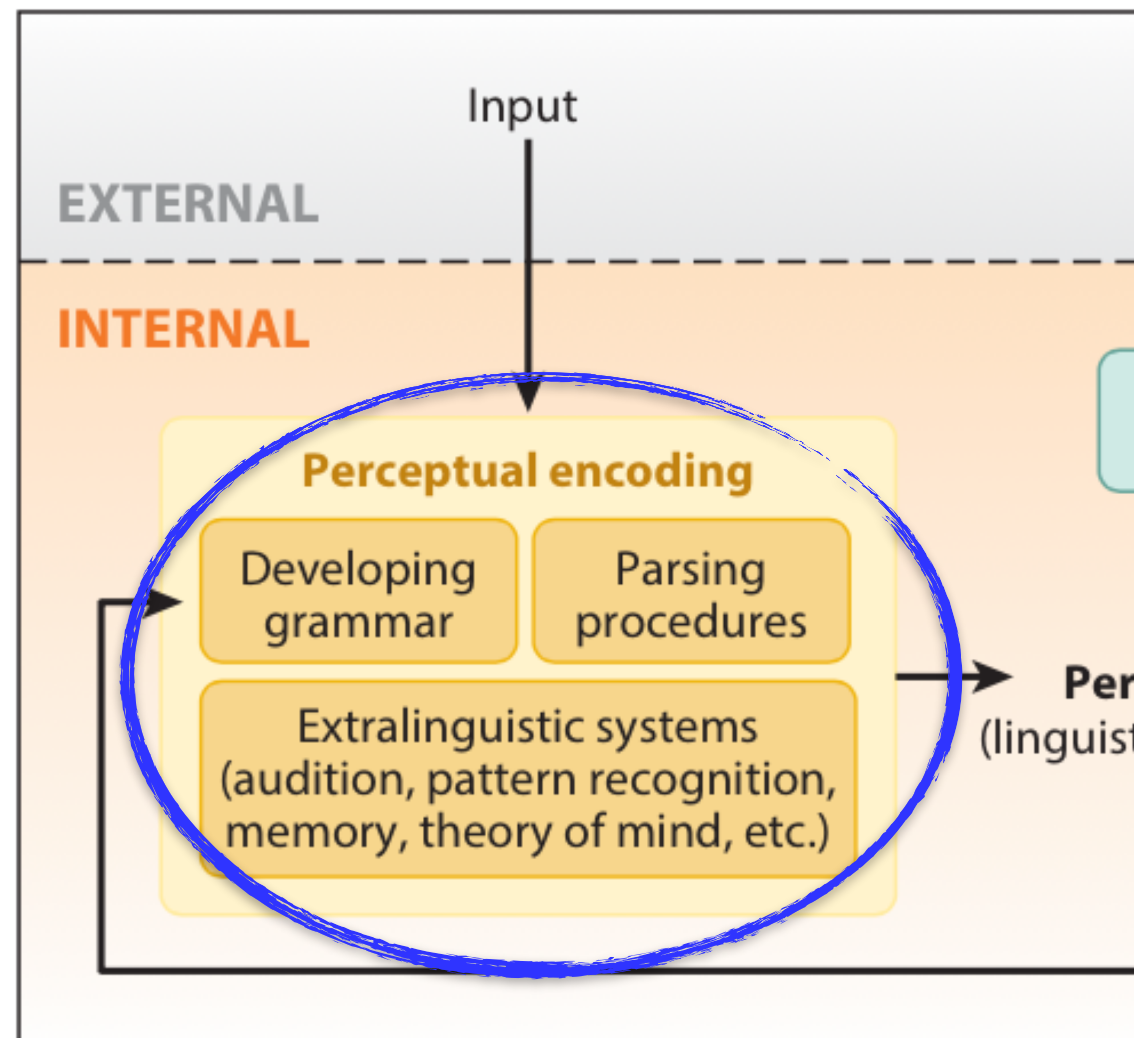

```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   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



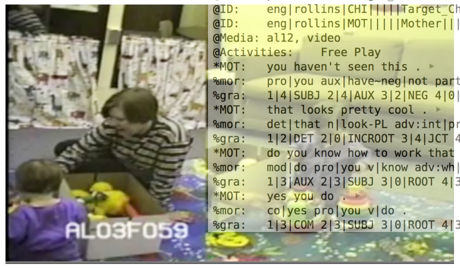
what a pretty kitty!

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



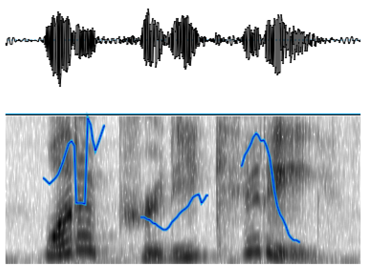
How do we model language acquisition?

An example with speech segmentation

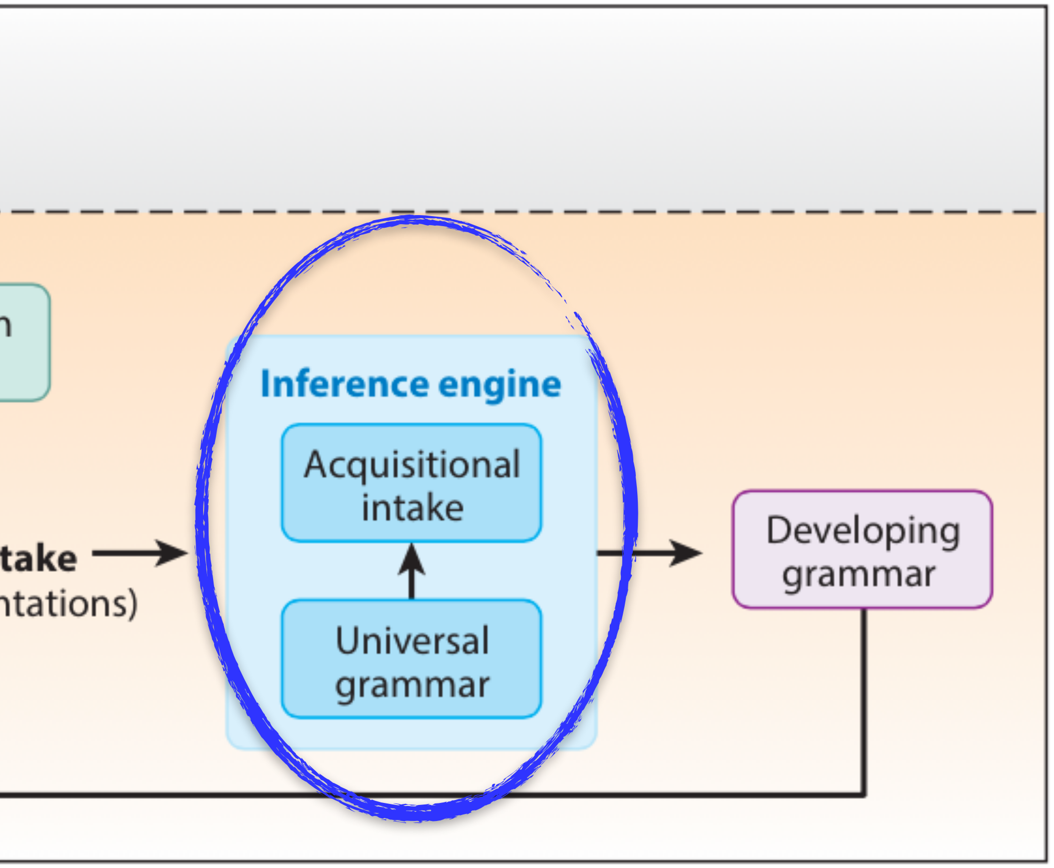



```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   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!



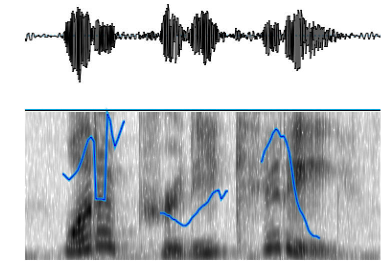
(4) Decide what hypotheses the child has and what information is being tracked in the input

This depends on your acquisition theory

How do we model language acquisition?

An example with speech segmentation

= 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

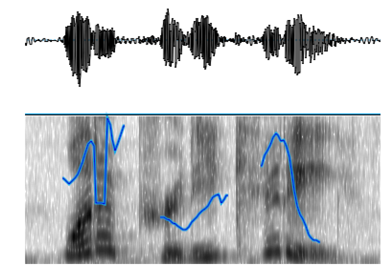
wˈʌɾə pɹˈɪɾi kˈɪɾi	wˈʌ ɾə pɹˈɪɾikˈɪɾi	wˈʌɾə pɹˈɪɾikˈɪɾi

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	wˈʌ	wˈʌɾə
pɹˈɪɾi	ɾə	pɹˈɪɾikˈɪɾi
kˈɪɾi	pɹˈɪɾikˈɪɾi	

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

Example information:

transitional probability between syllables,

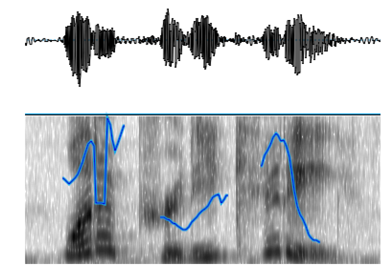
stress on syllables

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

How do we model language acquisition?

An example with speech segmentation

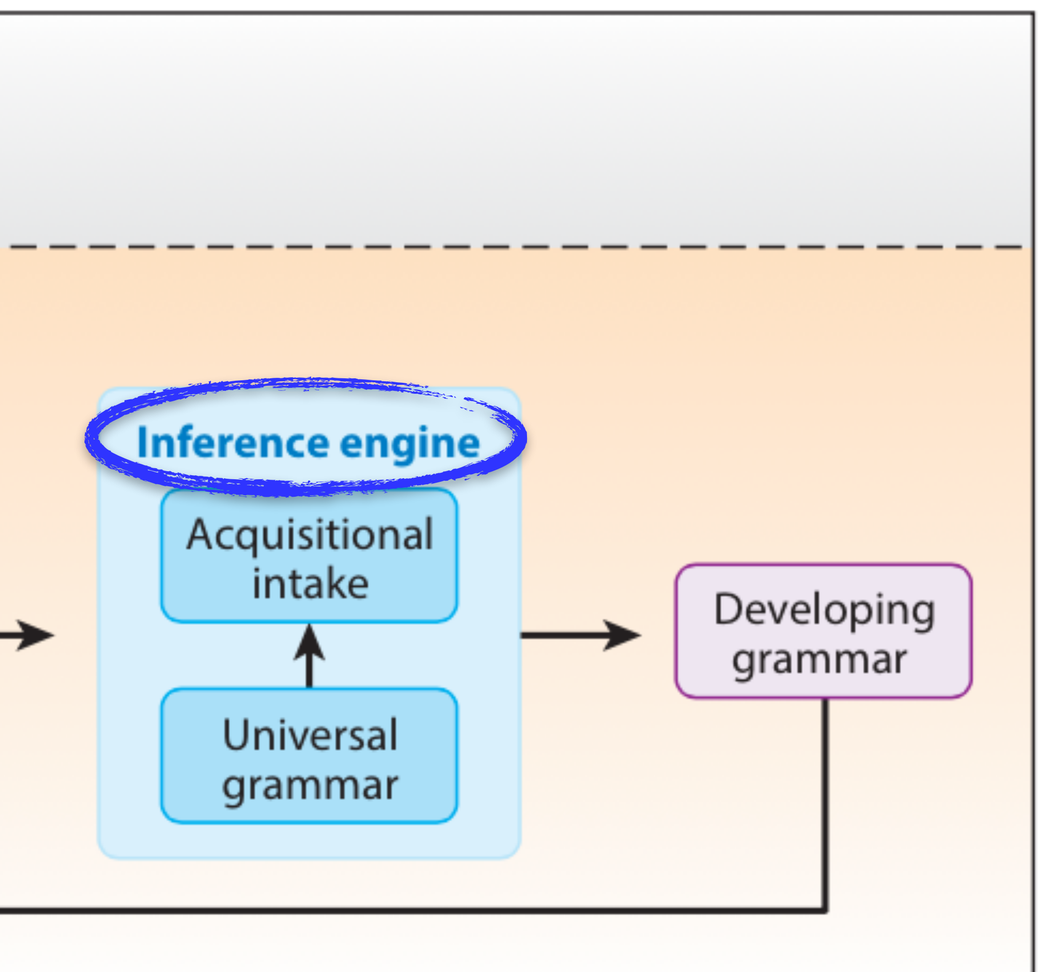
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə	wˈʌ	wˈʌɾə
pɹˈɪɾi	ɾə	pɹˈɪɾikˈɪɾi
kˈɪɾi	pɹˈɪɾikˈɪɾi	



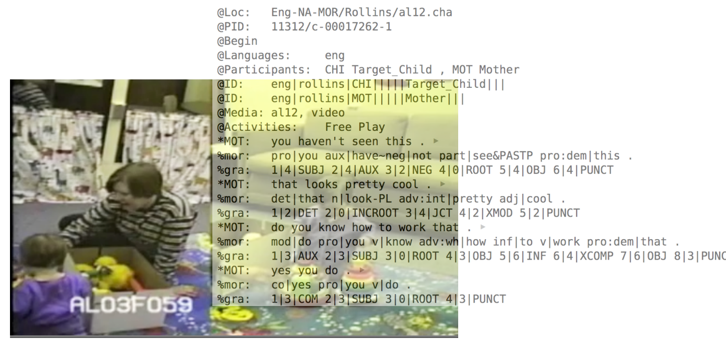
(5) Decide how belief in different hypotheses is updated

This depends on your acquisition theory

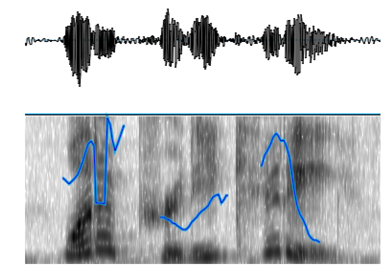
Example: based on transitional probability between syllables

How do we model language acquisition?

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



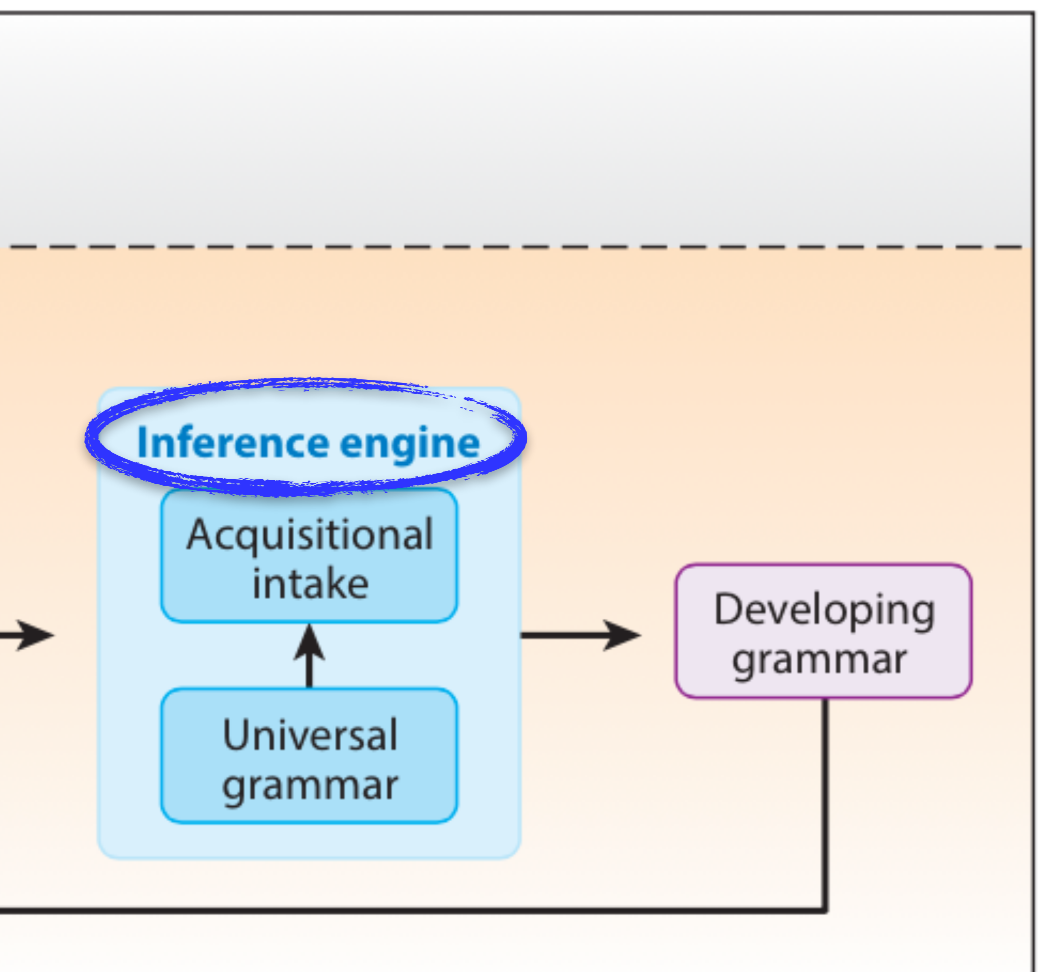
what a pretty kitty!

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

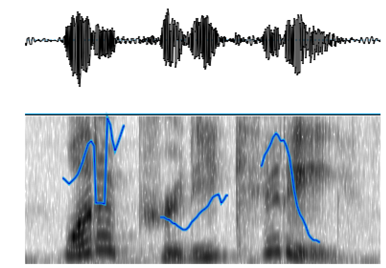
Example: based on transitional probability between syllables

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



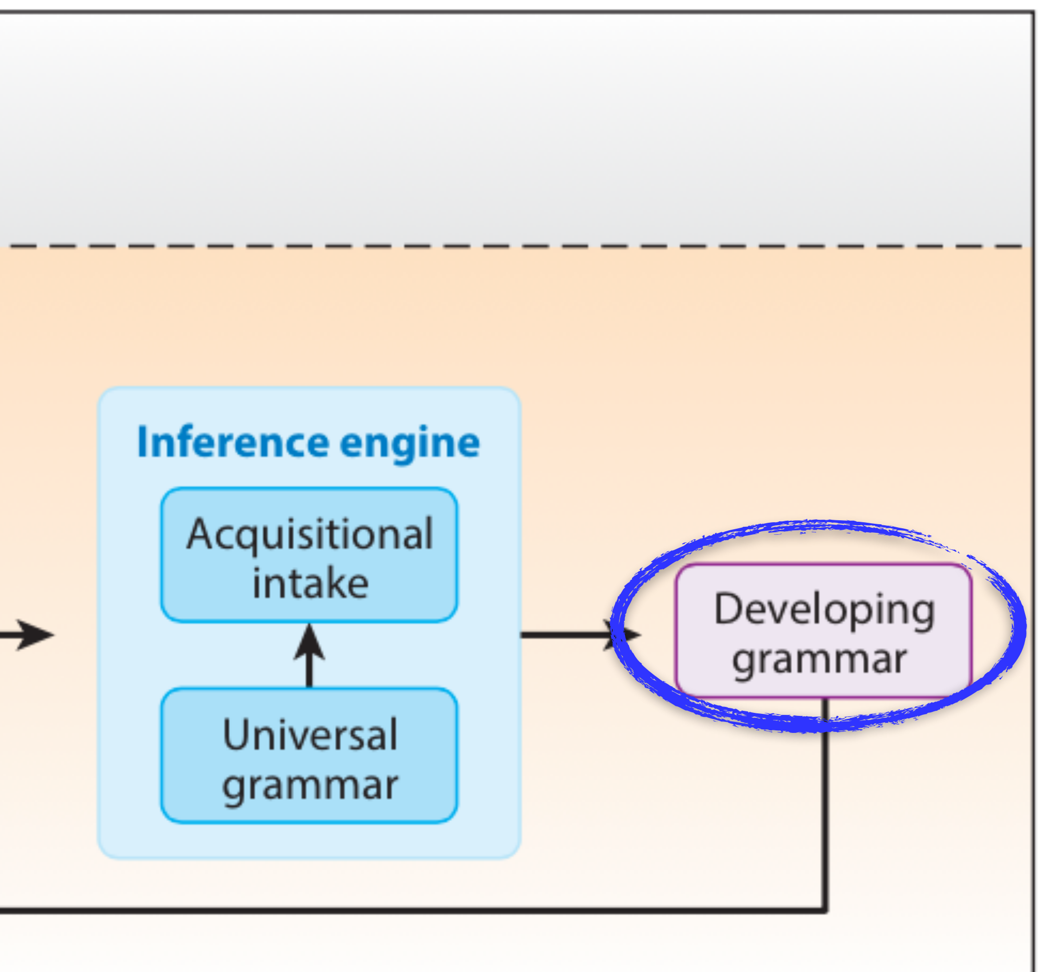
what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi



(6) Decide what the measure of success is

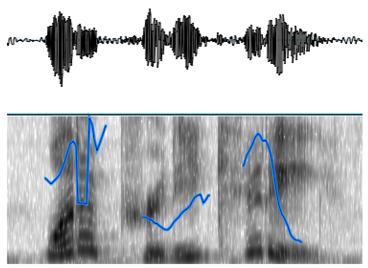
This can be based on your theory…

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

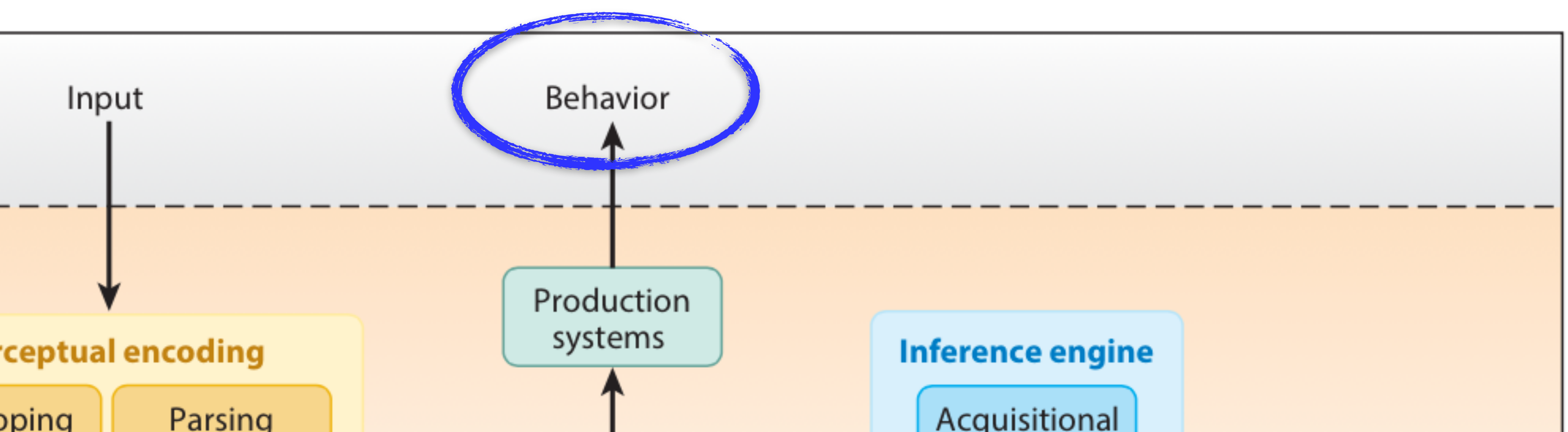
wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi

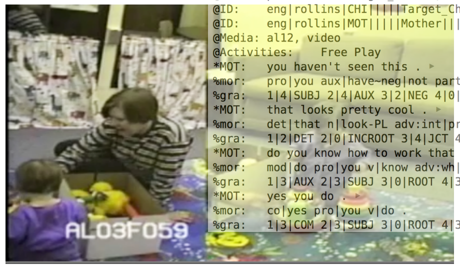
(6) Decide what the measure of success is

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



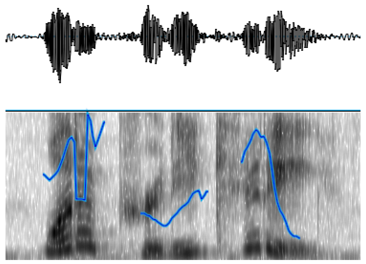
How do we model language acquisition?

An example with speech segmentation



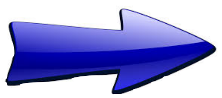

```
@Loc:   Eng-NA-MOR/Rollins/al12.cha
@PID:   11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:    eng|rollins|CHI||||Target_Child|||
@ID:    eng|rollins|MOT||||Mother|||
@Media: al12, video
@Activities:    Free Play
*MOT:   you haven't seen this .
%mor:   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

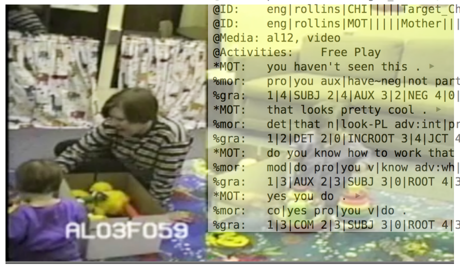
Proto-lexicon of word forms

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

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

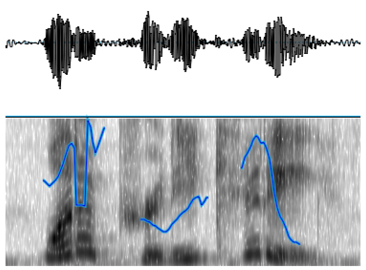
How do we model language acquisition?

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!

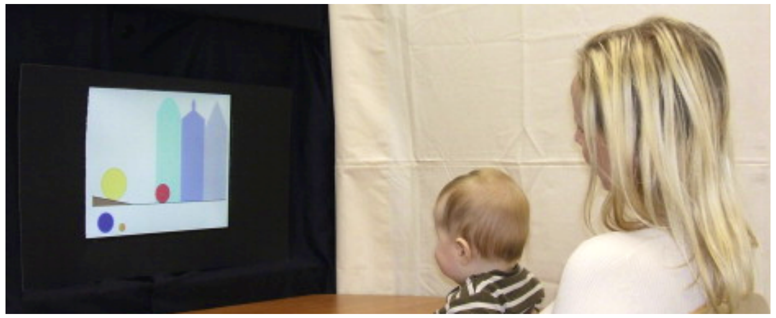
wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi → wˈʌɾə pɹˈɪɾi kˈɪɾi | wˈʌ ɾə pɹˈɪɾikˈɪɾi | wˈʌɾə pɹˈɪɾikˈɪɾi

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

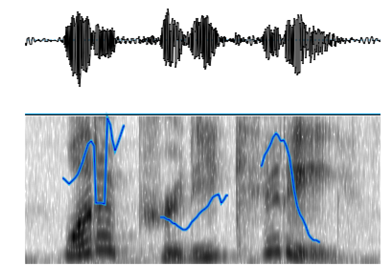


How do we model language acquisition?

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

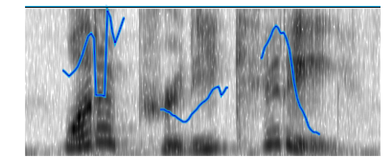
wˈʌɾə
pɹˈɪɾikˈɪɾi

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 ɾə pɹ'ɪ ɾi k'ɪ ɾi

what a pretty kitty!

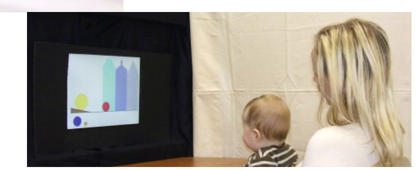
w'ʌ ɾə pɹ'ɪ ɾi k'ɪ ɾi

pɹ'ɪɾi pretty

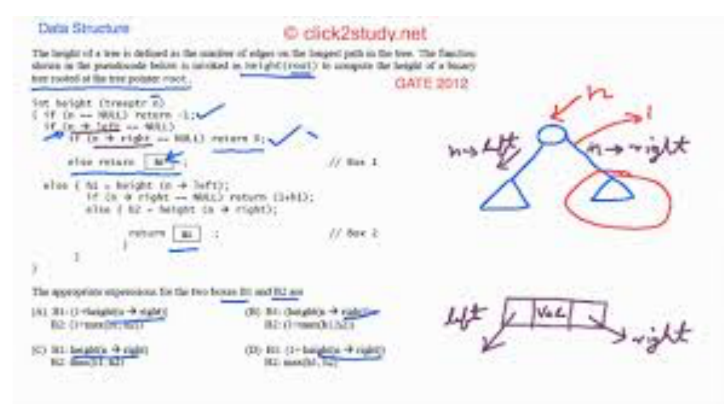
k'ɪɾi kitty

w'ʌɾə

pɹ'ɪɾik'ɪɾi

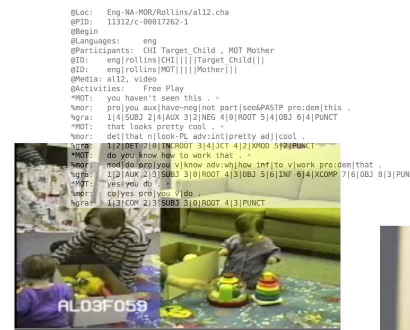


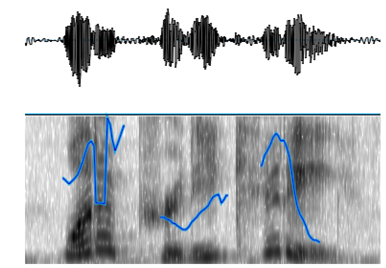
(7) Implement the model in a programming language of choice



How do we model language acquisition?

An example with speech segmentation





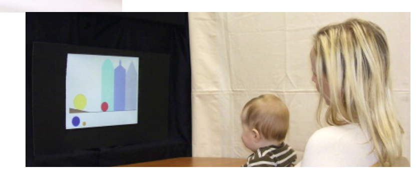
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə

pɹˈɪɾikˈɪɾi

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

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



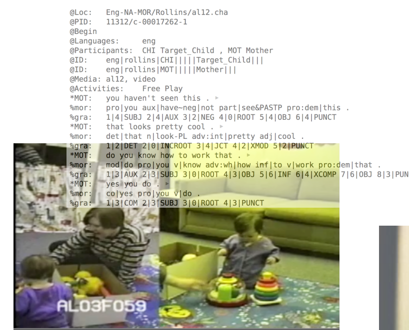
Example developing knowledge

Proto-lexicon of word forms

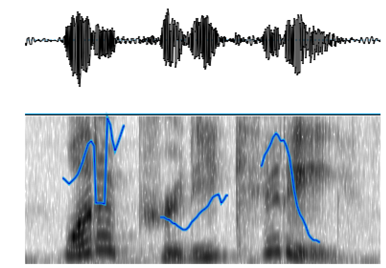
	wˈʌɾ	*what*
	ə	*a*
???	pɹˈɪɾi	*pretty*
	kˈɪɾi	*kitty*

How do we model language acquisition?

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

pɹˈɪɾi pretty

kˈɪɾi kitty

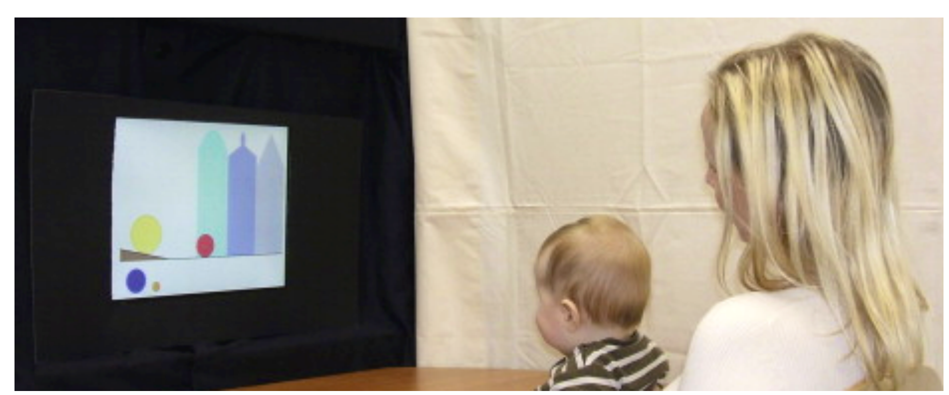
wˈʌɾə

pɹˈɪɾikˈɪɾi

(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

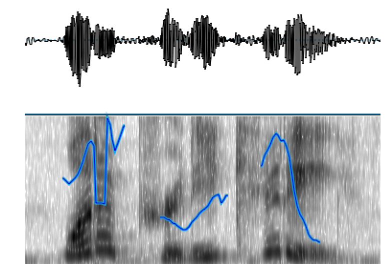
???



How do we model language acquisition?

An example with speech segmentation

= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə

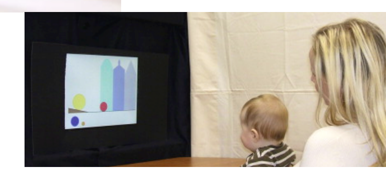
pɹˈɪɾikˈɪɾi

???

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

pɹˈɪɾi pretty

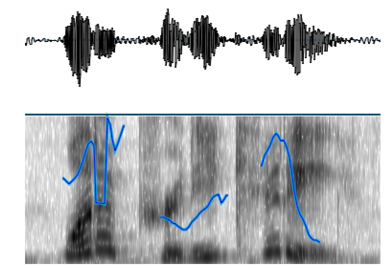
kˈɪɾi kitty



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



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

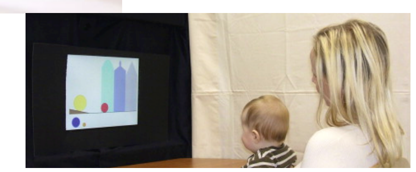
wˈʌɾə

pɹˈɪɾikˈɪɾi

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

???

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

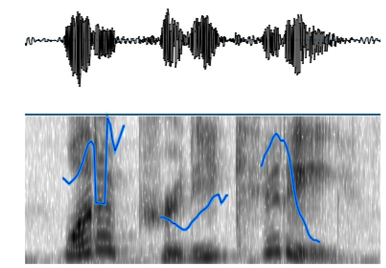


"The modeled child has the same developing knowledge as we think 8-month-olds do. This strategy can be what they're using!"

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

How do we model language acquisition?

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

what a pretty kitty!

???

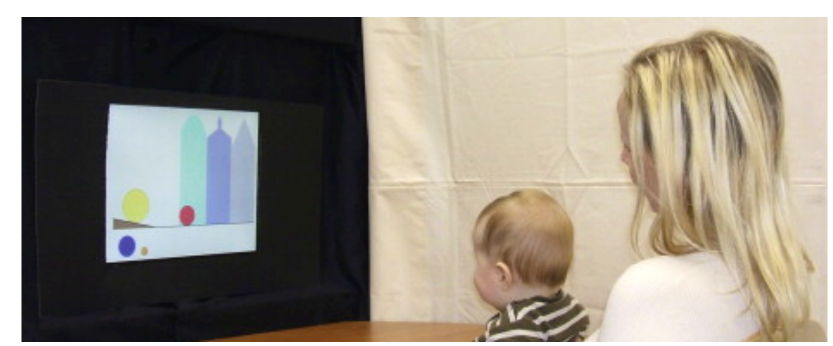
wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə

pɹˈɪɾikˈɪɾi

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

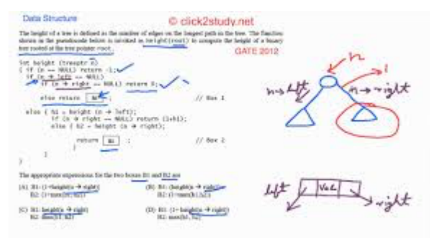
"The modeled child can reproduce the behavior we see in 8-month-olds. This strategy could be what they're using to generate that behavior!"



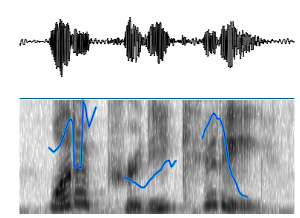
Today’s Plan:
Computational models of language acquisition

I. Why

II. How

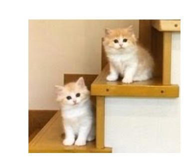


III. What we can learn



Noun

✓KI tty



Who does… is pretty?

another one

Every kitty didn’t …

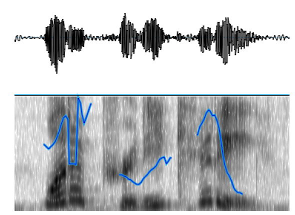
Today's Plan:
Computational models of language acquisition

III. What we can learn

metrical phonology

syntactic categorization

speech segmentation



Noun

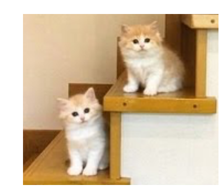
✓KI tty

syntax

Who does… is pretty?

another one

Every kitty didn't …

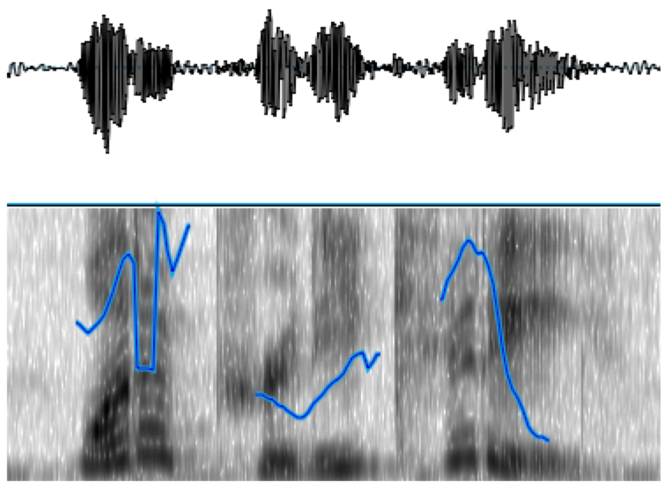


pragmatics

syntax, semantics

What we can learn

speech segmentation



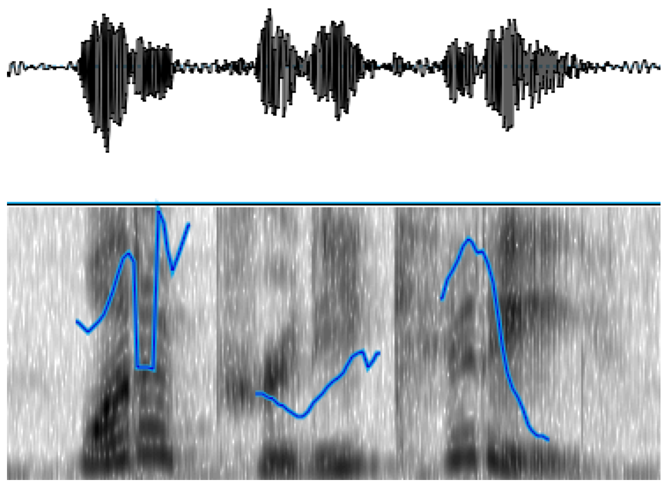
= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

What we can learn

speech segmentation

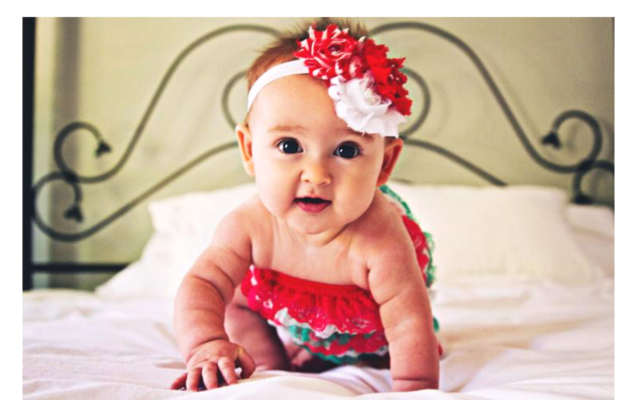


= wʌɾəpɹɪɾikɪɾi

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

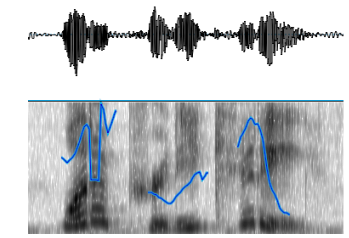


Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

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

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

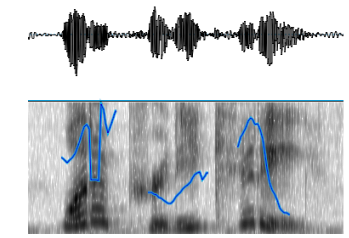
wʌ
ɾə
pɹɪɾikɪɾi

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

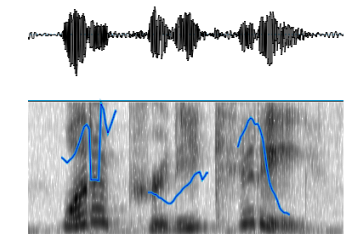
wʌ
ɾə
pɹɪɾikɪɾi

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

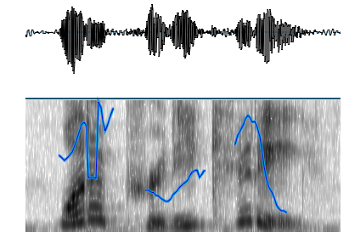
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

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
pɹɪɾi
kɪɾi

wʌ
ɾə
pɹɪɾikɪɾi

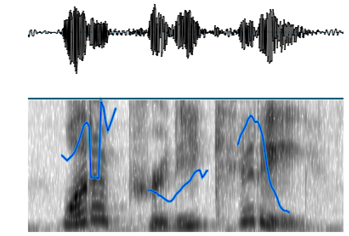
wʌɾə
pɹɪɾikɪɾi

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

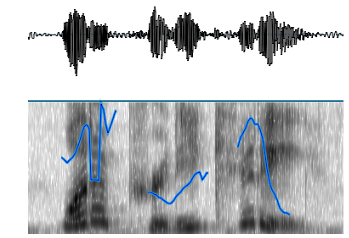
wʌɾə
pɹɪɾikɪɾi

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



= 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

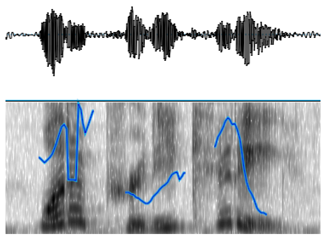
wʌɾə
pɹɪɾikɪɾi

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

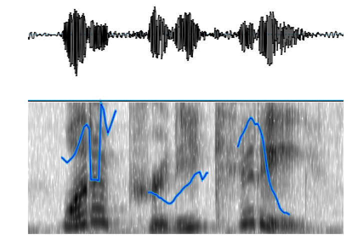
Computational-level modeled learners using this strategy segment fairly well, given realistic English child-directed speech data.

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?**

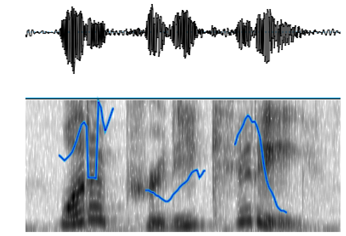
✓ **Is it useable?**

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?**

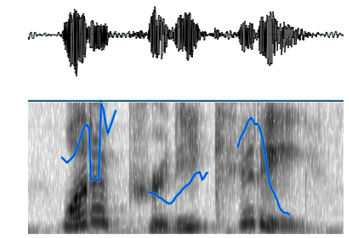
✓ **Does it work for different languages?**

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?

Bayesian inference seems to be a good proposal for a very early speech segmentation strategy.

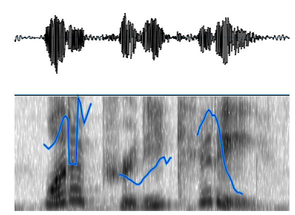
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

What we can learn

metrical phonology

syntactic categorization

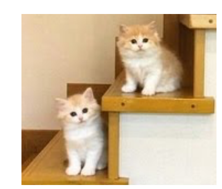
speech segmentation



Noun

syntax

✓KI tty



Who does… is pretty?

another one

Every kitty didn't …

pragmatics

syntax, semantics

What we can learn

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

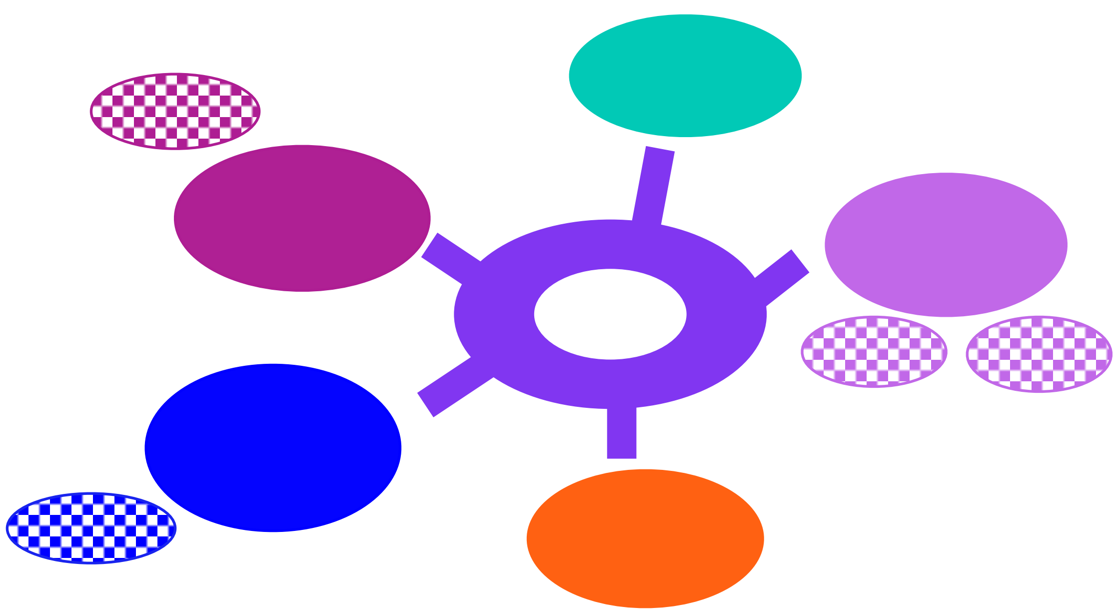
metrical phonology

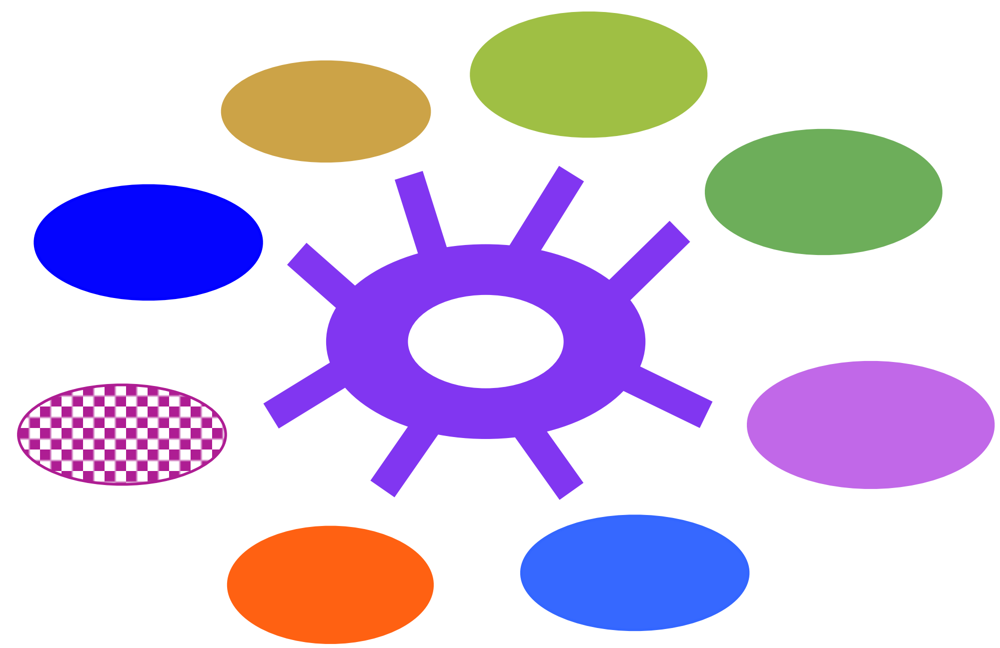
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

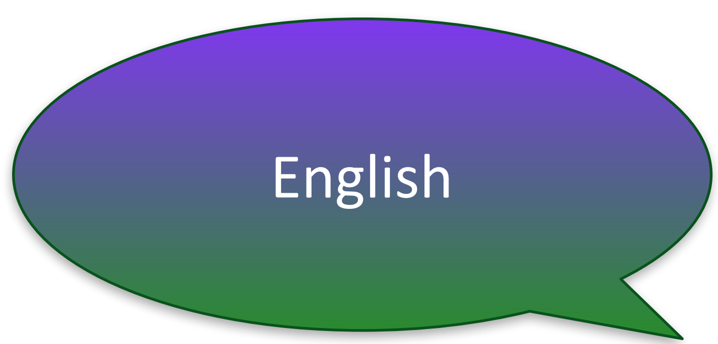
✓ KI tty
✗ ki TTY

knowledge representation options

parameters whose values must be set







What we can learn

metrical phonology

a DO ra ble

A do RA ble

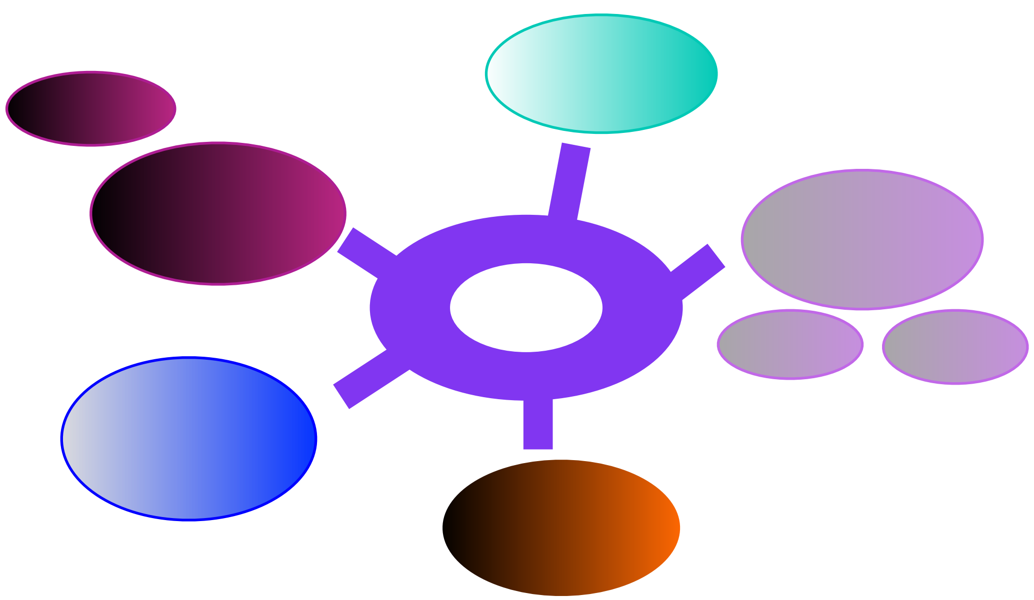
a DO ra BLE

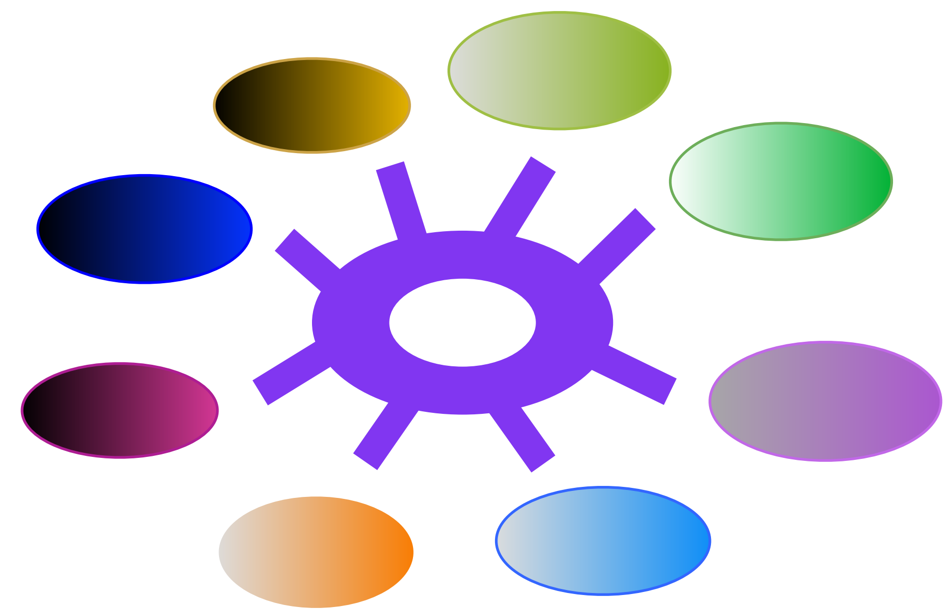
KI tty

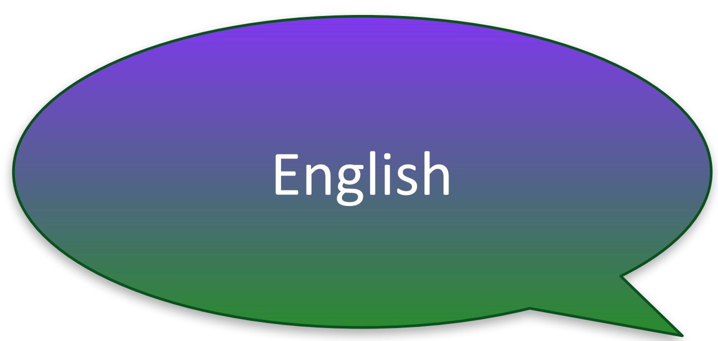
ki TTY

knowledge representation options

parameters whose values must be set







What we can learn

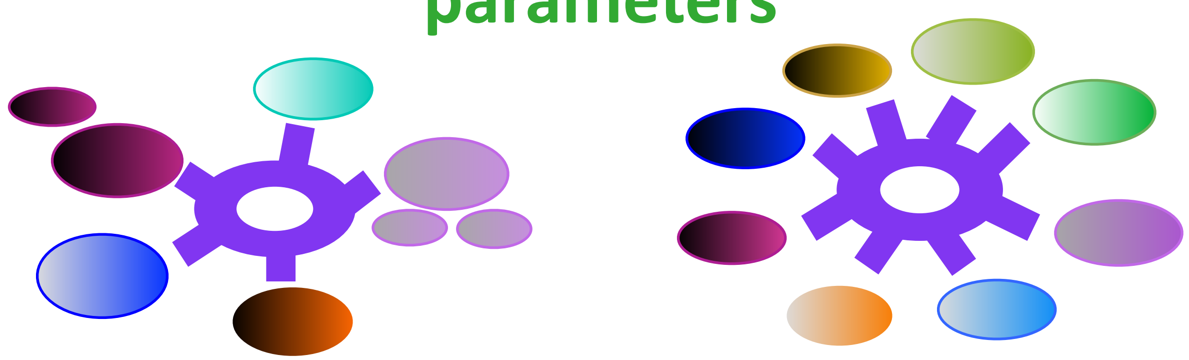
metrical phonology

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

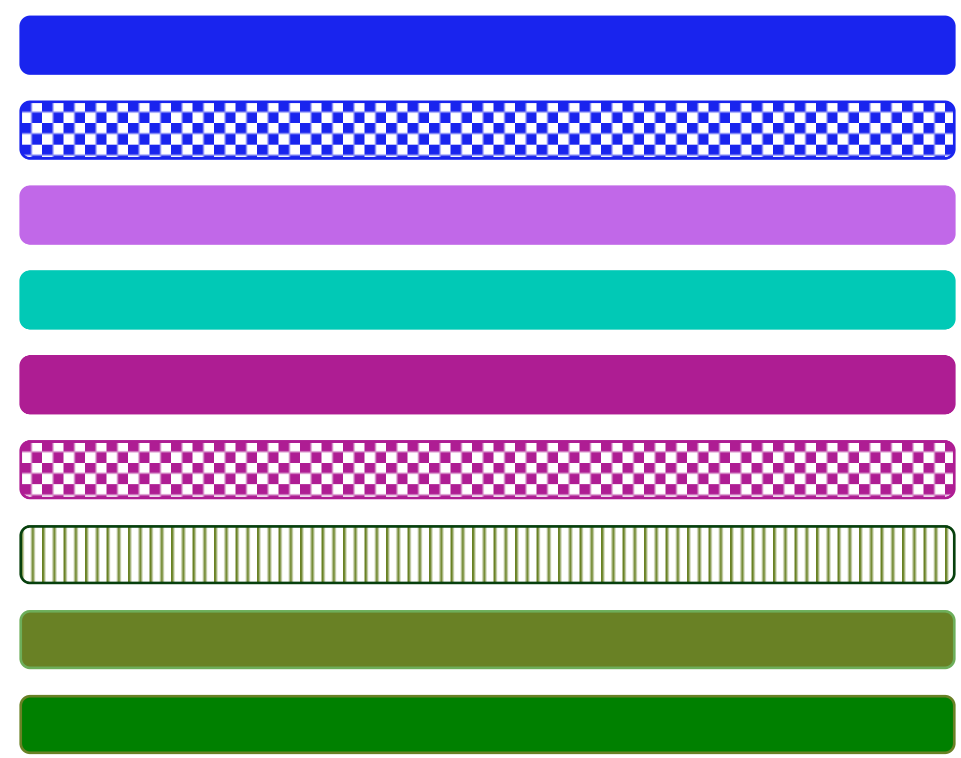
KI tty
ki TTY

knowledge representation options

parameters



Violable constraints that must be ranked



English

What we can learn

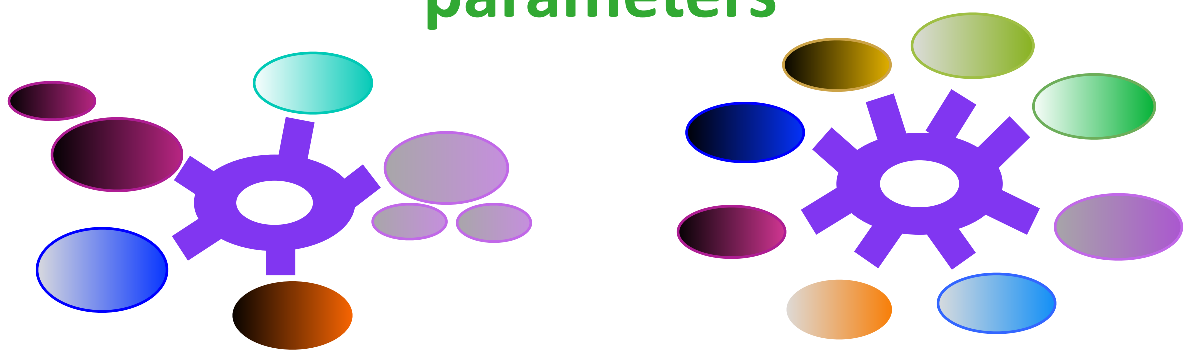
metrical phonology

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

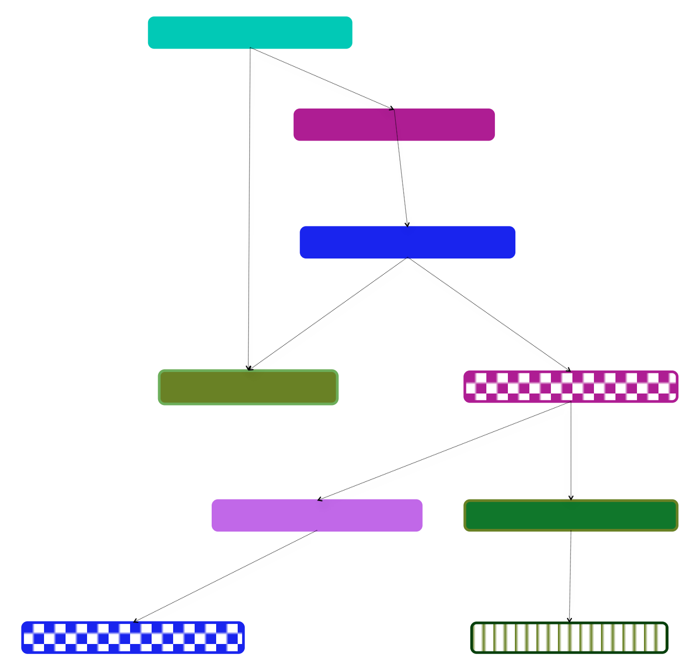
✓ KI tty
✗ ki TTY

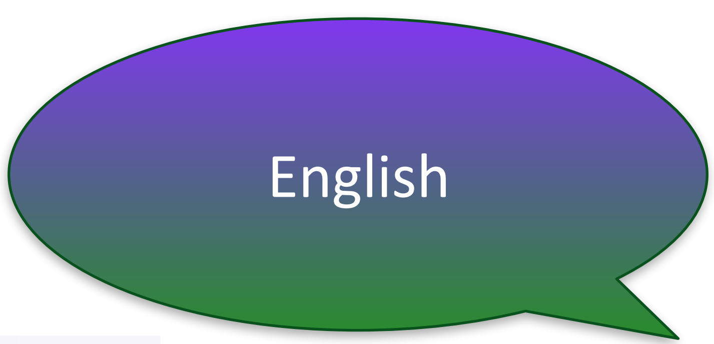
knowledge representation options

parameters



Violable constraints that must be ranked





What we can learn

metrical phonology

a DO ra ble

A do RA ble

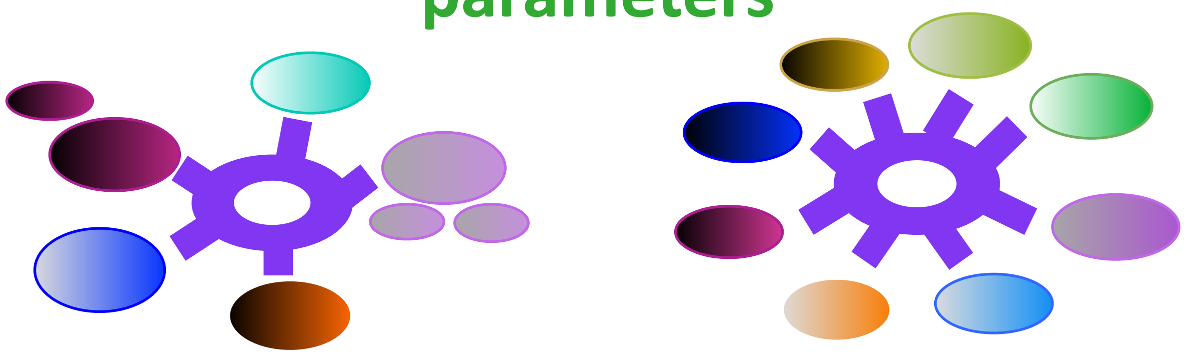
a DO ra BLE

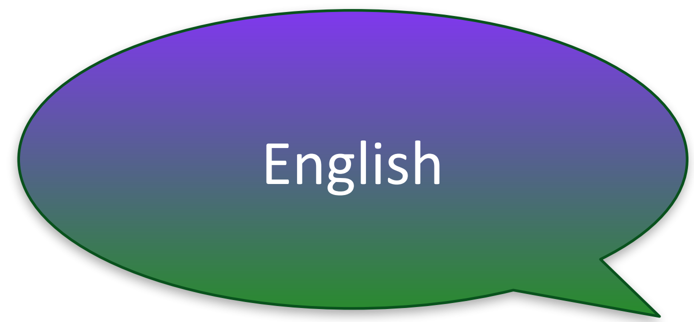
KI tty

ki TTY

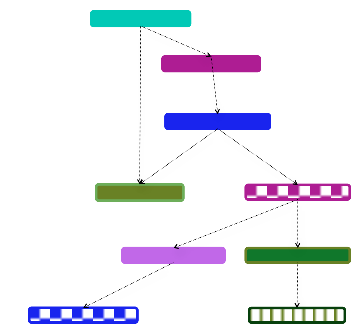
knowledge representation options

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

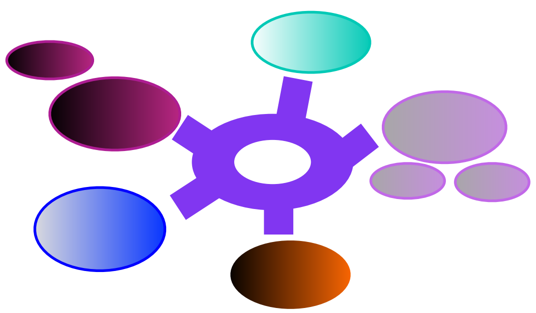
metrical phonology

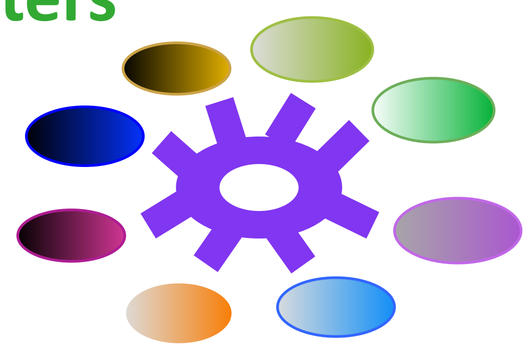
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

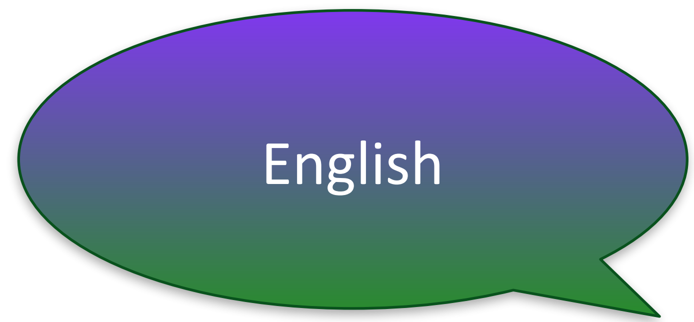
✓ KI tty
✗ ki TTY

knowledge representation options

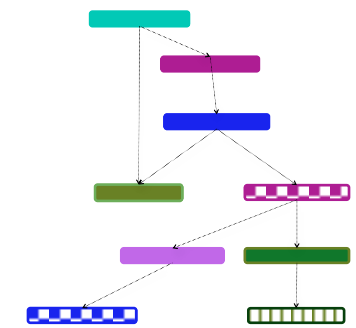
parameters







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

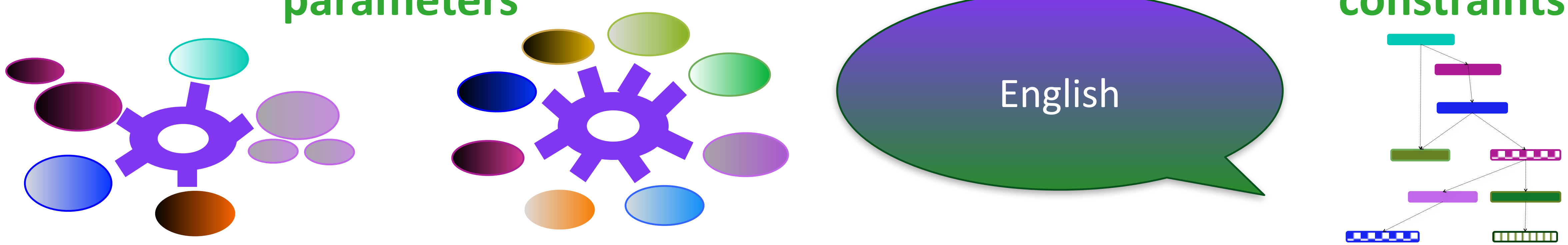
metrical phonology

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

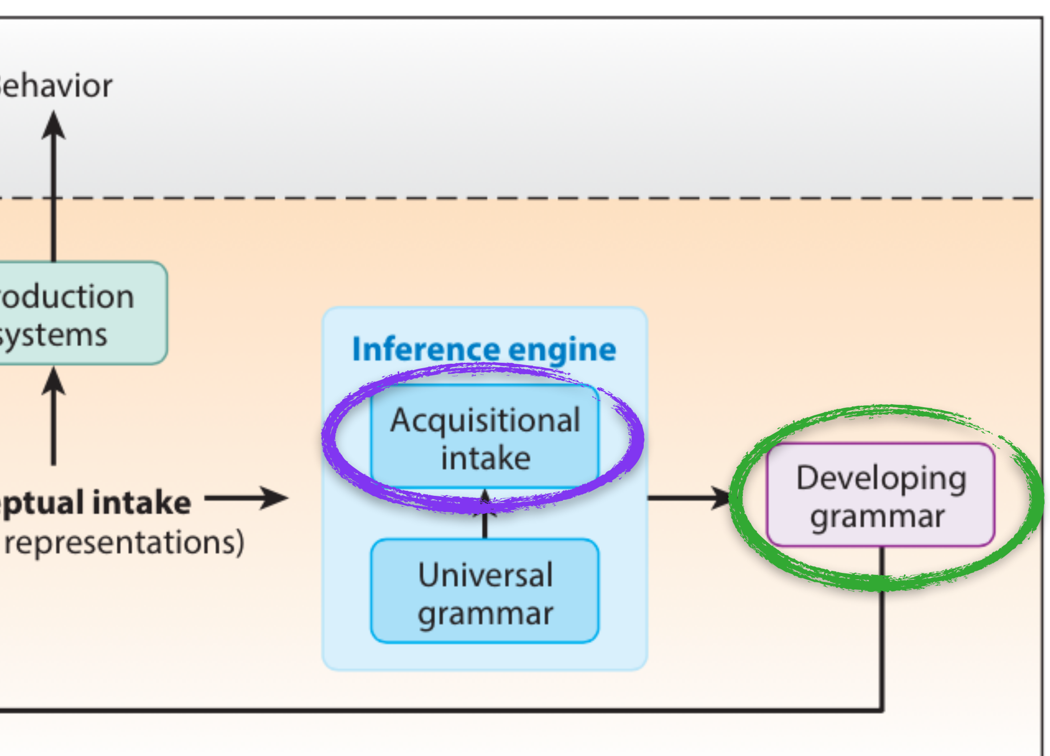
knowledge representation options

parameters

constraints



how learnable they are



Computational-level analysis

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

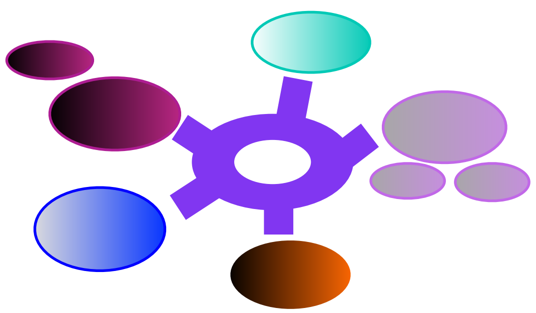
metrical phonology

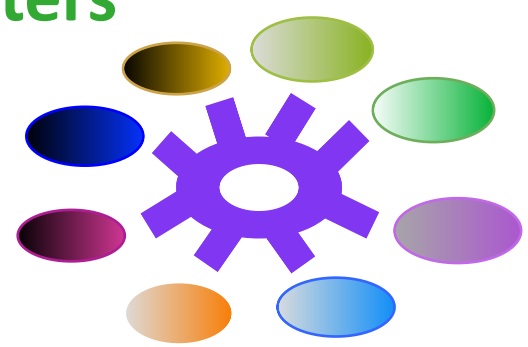
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

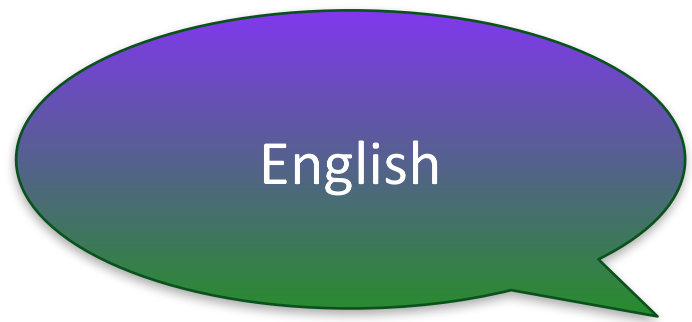
✓ KI tty
✗ ki TTY

knowledge representation options

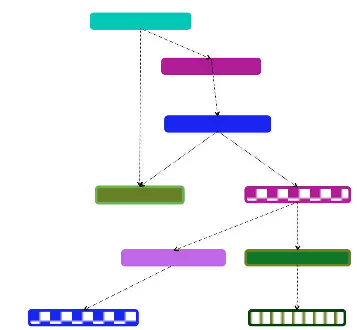
parameters







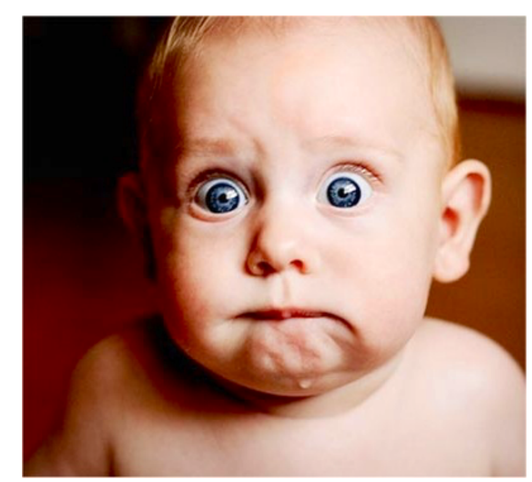
constraints



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

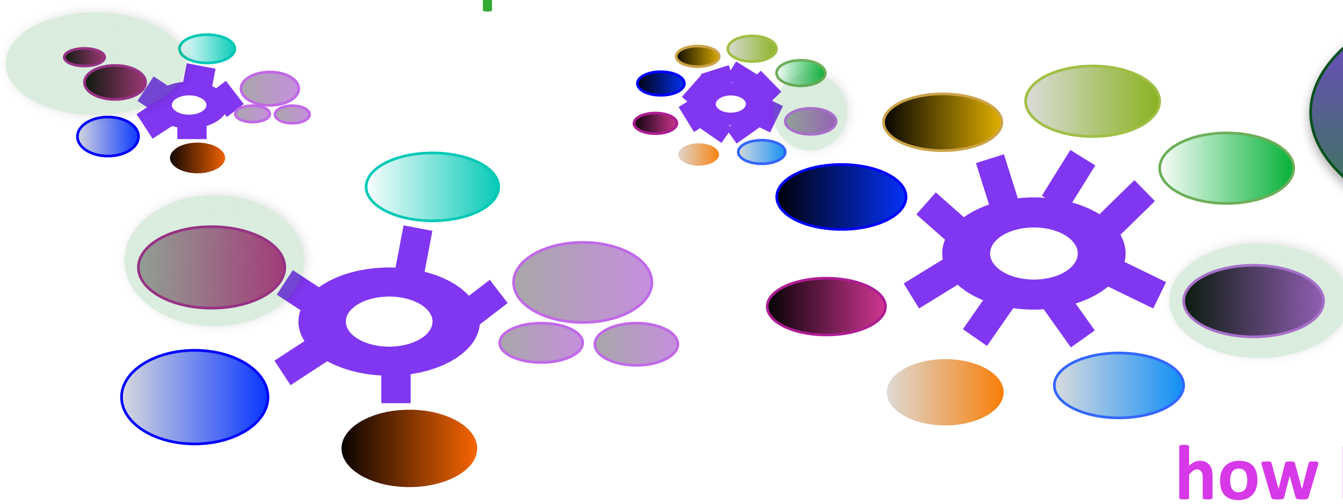
metrical phology

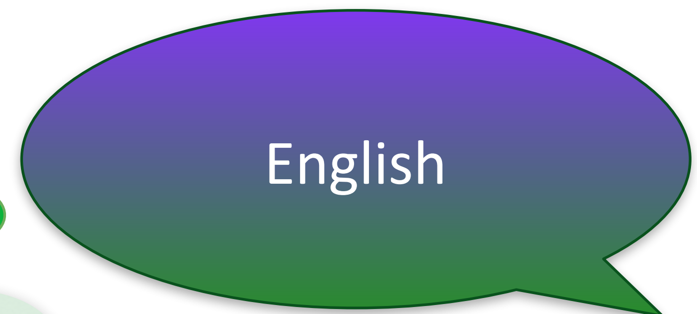
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

✓ KI tty
✗ ki TTY

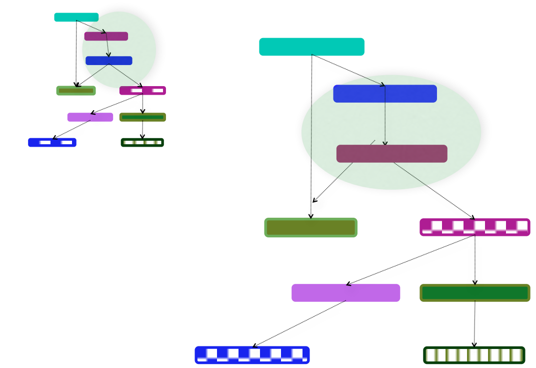
knowledge representation options

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

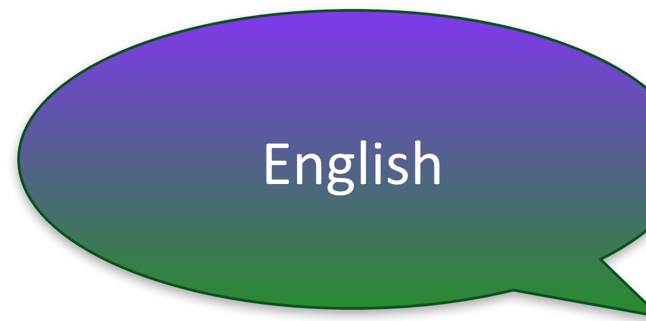
metrical phonology

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

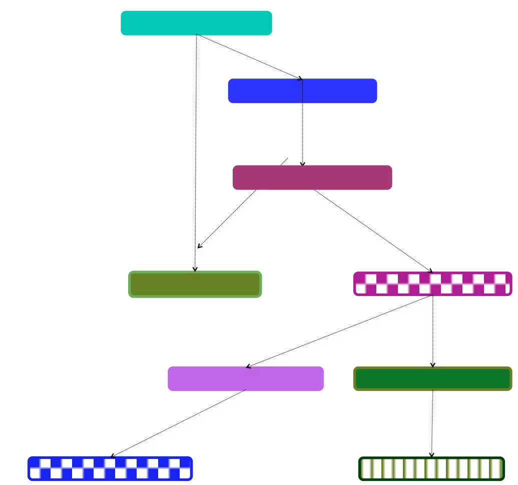
✓ KI tty
✗ ki TTY

knowledge representation options

parameters

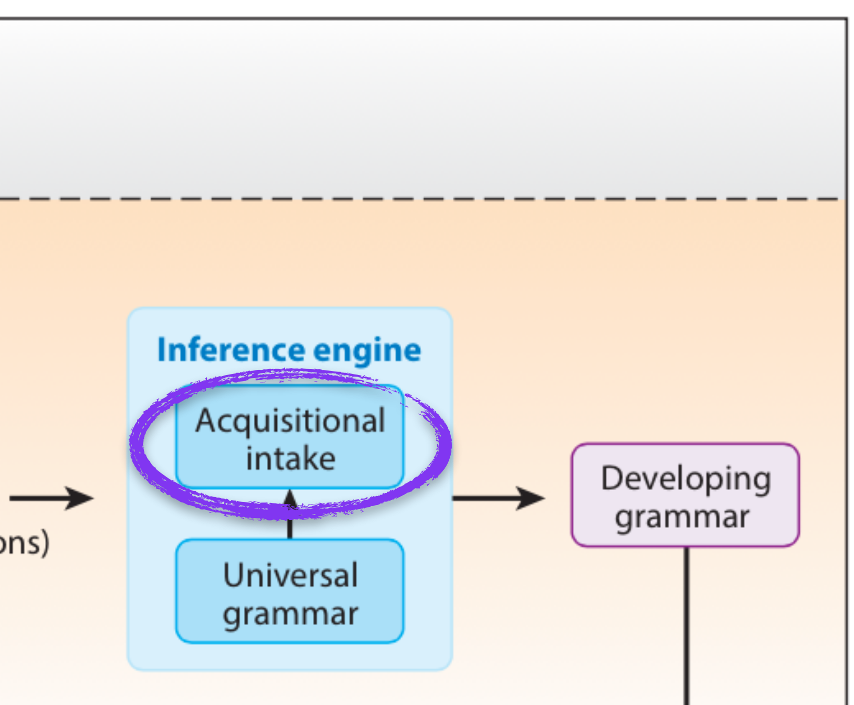


constraints



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

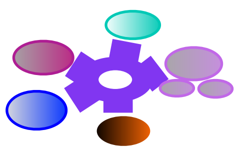
metrical phonology

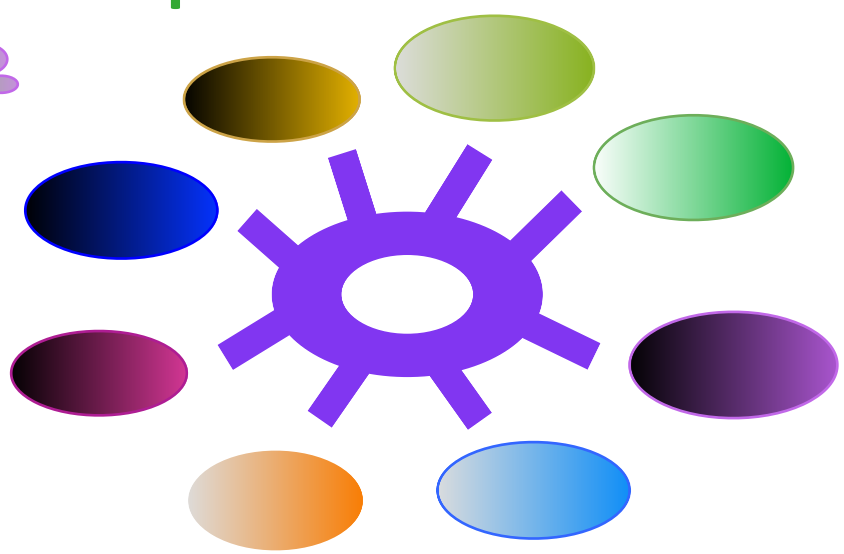
✓ a DO ra ble
✗ A do RA ble
✗ a DO ra BLE

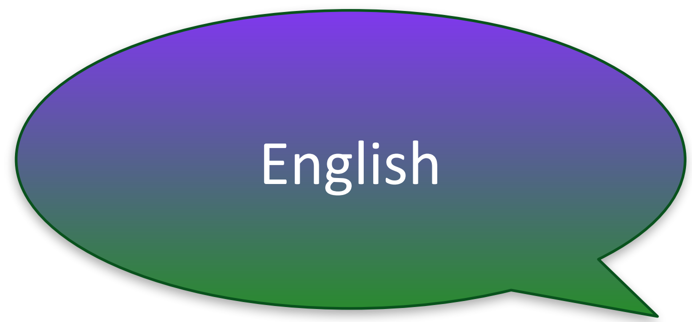
✓ KI tty
✗ ki TTY

knowledge representation options

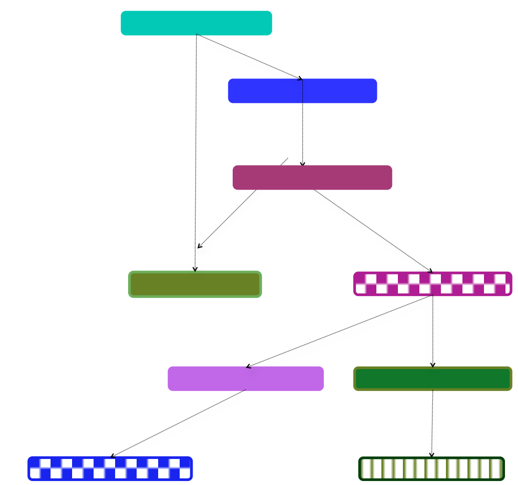
parameters







constraints



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

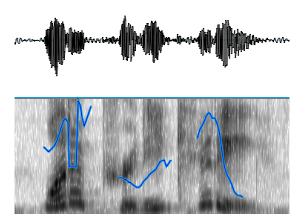
Pearl et al. 2014, Pearl 2017 , Pearl et. al in press

What we can learn

metrical phonology

syntactic categorization

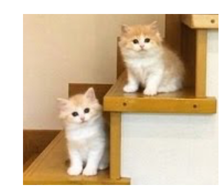
speech segmentation



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

Noun

penguin

kitty

owl

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 owl
penguin
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

didn’t
not
Negation
wouldn’t
can’t
won’t

might
would
could
will
Auxiliary
can
should

surprise
stand
find
Verb
adore

idea
glitter
unicorn
Noun
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

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

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

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

There's significant debate on when these categories develop.

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

Verb ???: surprise, adore, stand, find, dance

What we can learn

syntactic categorization

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

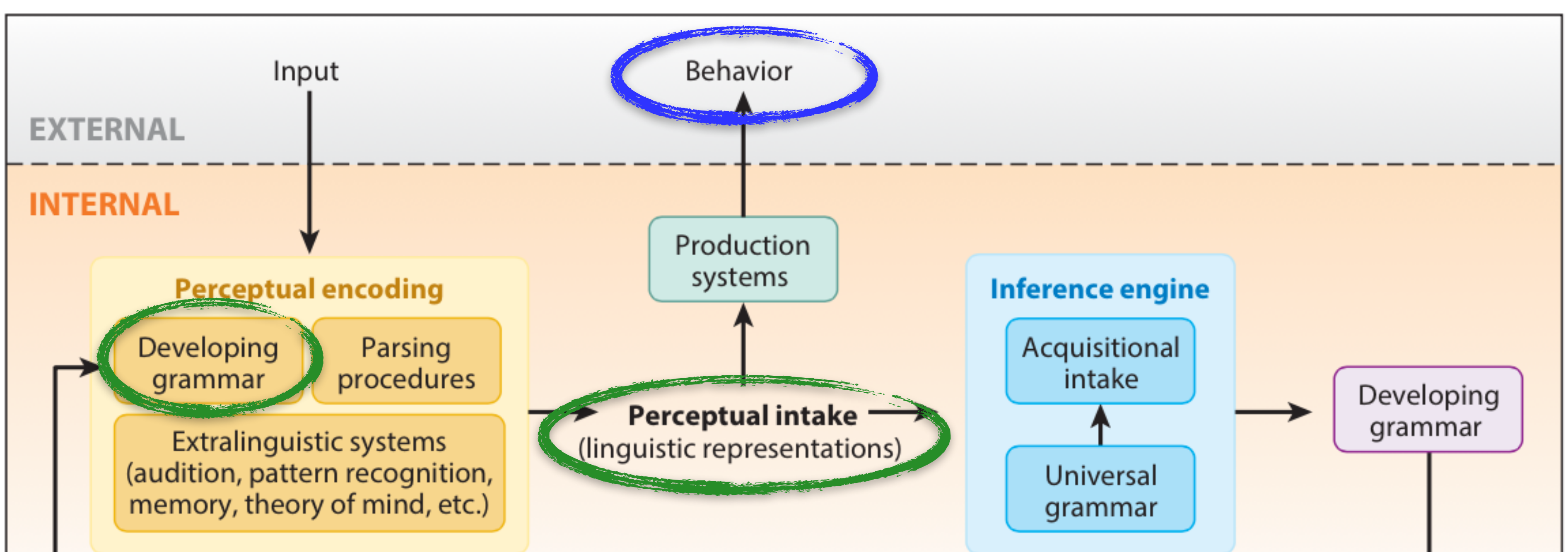
might would could
will **Auxiliary**
should can

glitter idea unicorn
Noun owl
penguin kitty

surprise stand
find **Verb** dance
adore

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
should can

idea
glitter unicorn
Noun owl
penguin kitty

surprise stand
find Verb dance
adore

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

✓
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
penguin ~~**Noun**~~ 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

Auxiliary

might would could will can should

Negation

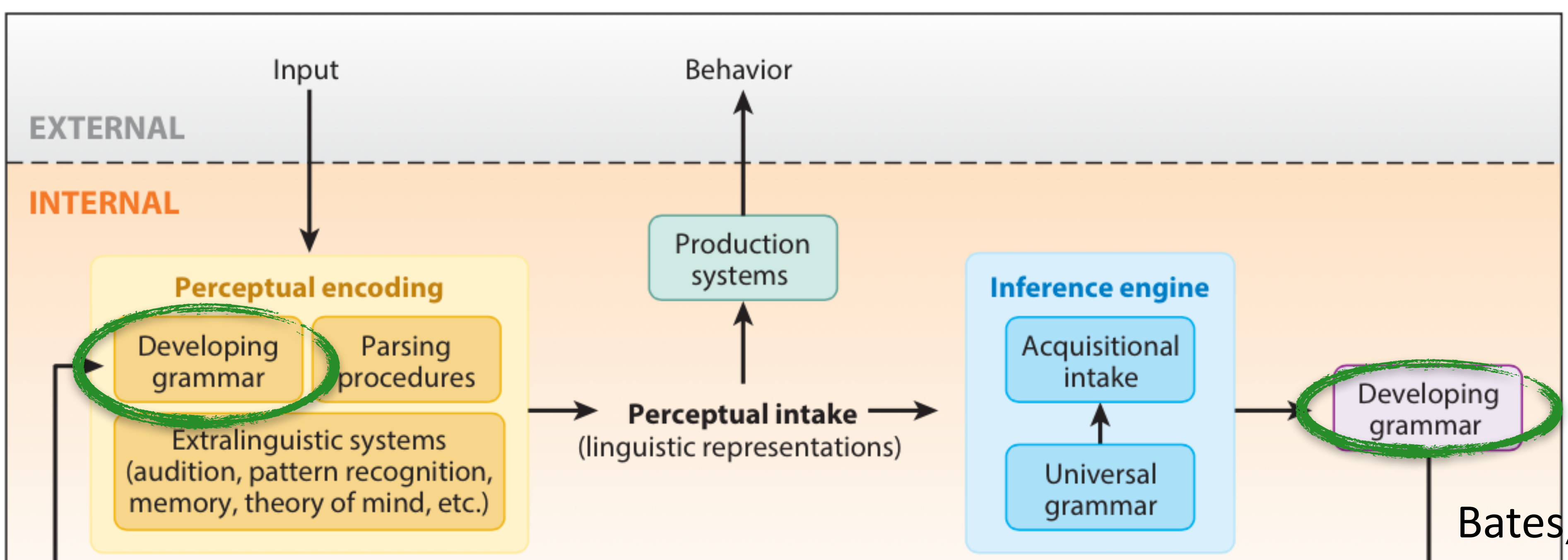
didn't not wouldn't can't won't

surprise stand find Verb dance adore

idea glitter unicorn Noun penguin owl 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.



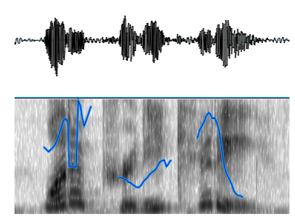
Bates, Pearl & Braunwald, in prep.

What we can learn

metrical phonology

syntactic categorization

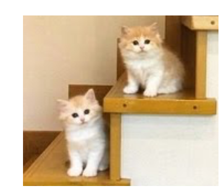
speech segmentation



Noun

syntax

✔KI tty



Who does … is pretty? *another one*

Every kitty didn't …

pragmatics

syntax, semantics

What we can learn

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

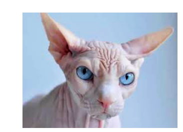
What we can learn

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!"

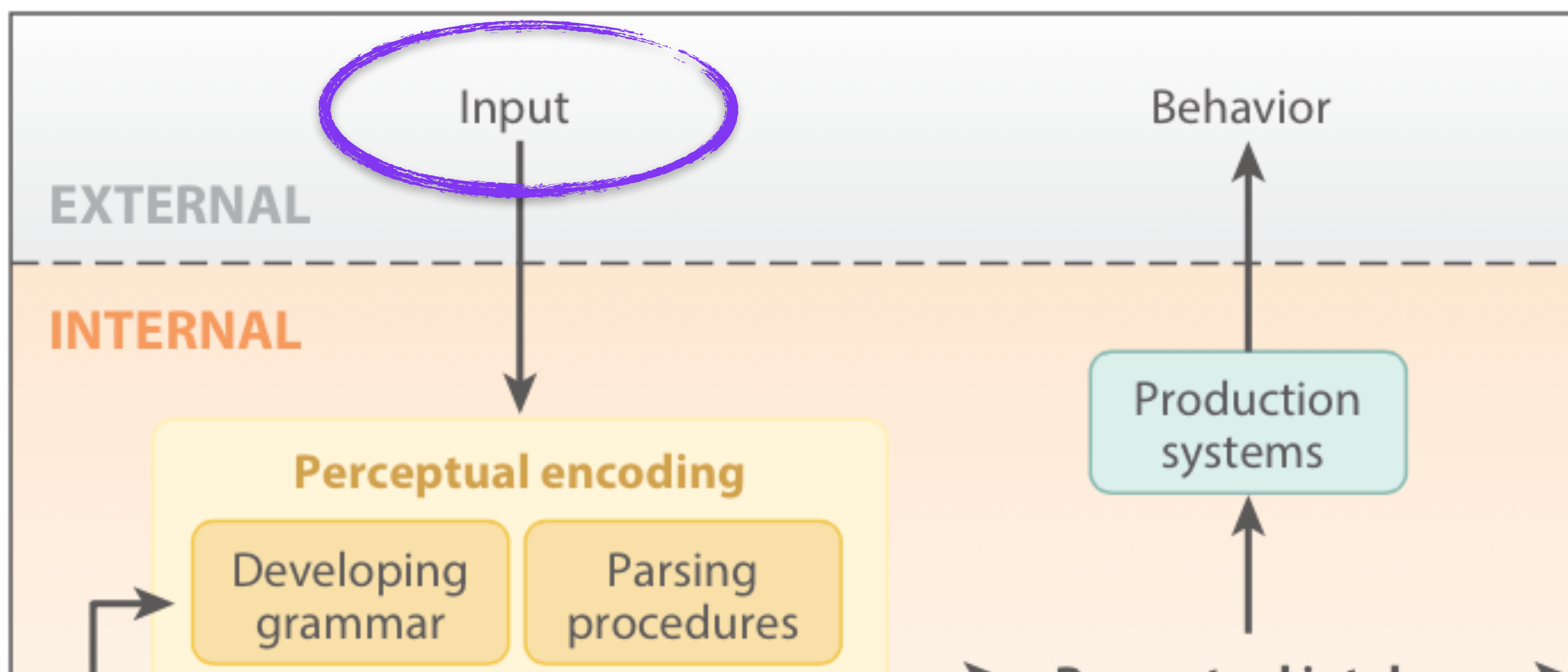
Noun'

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!"

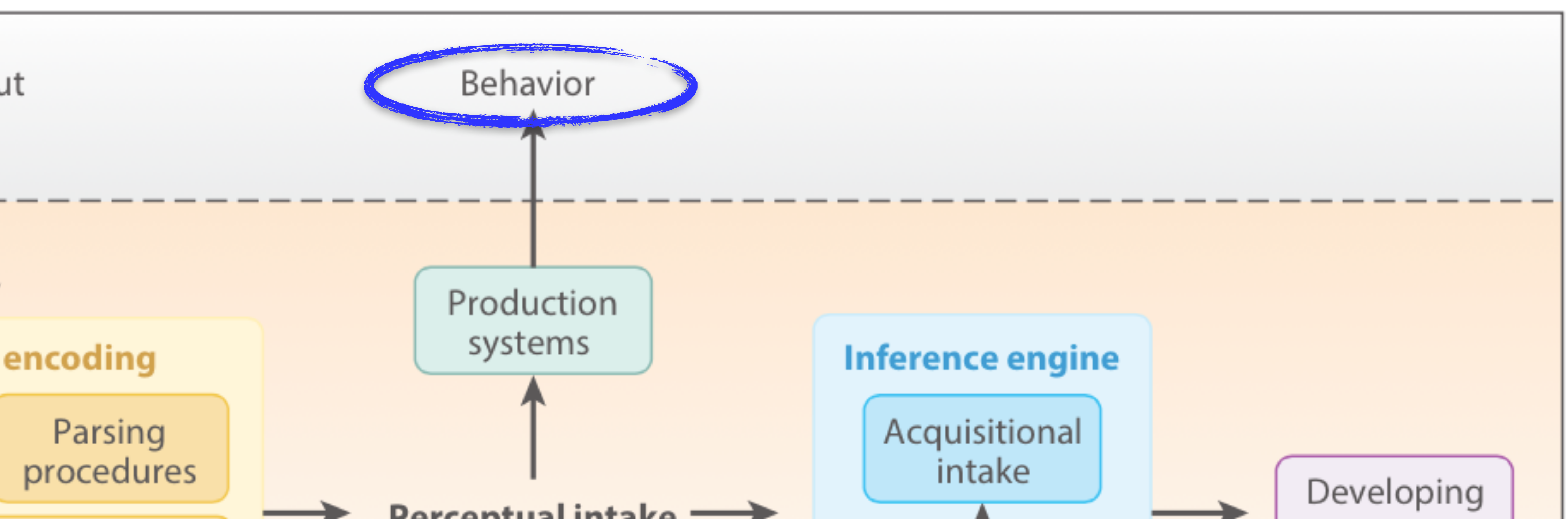
Noun'

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

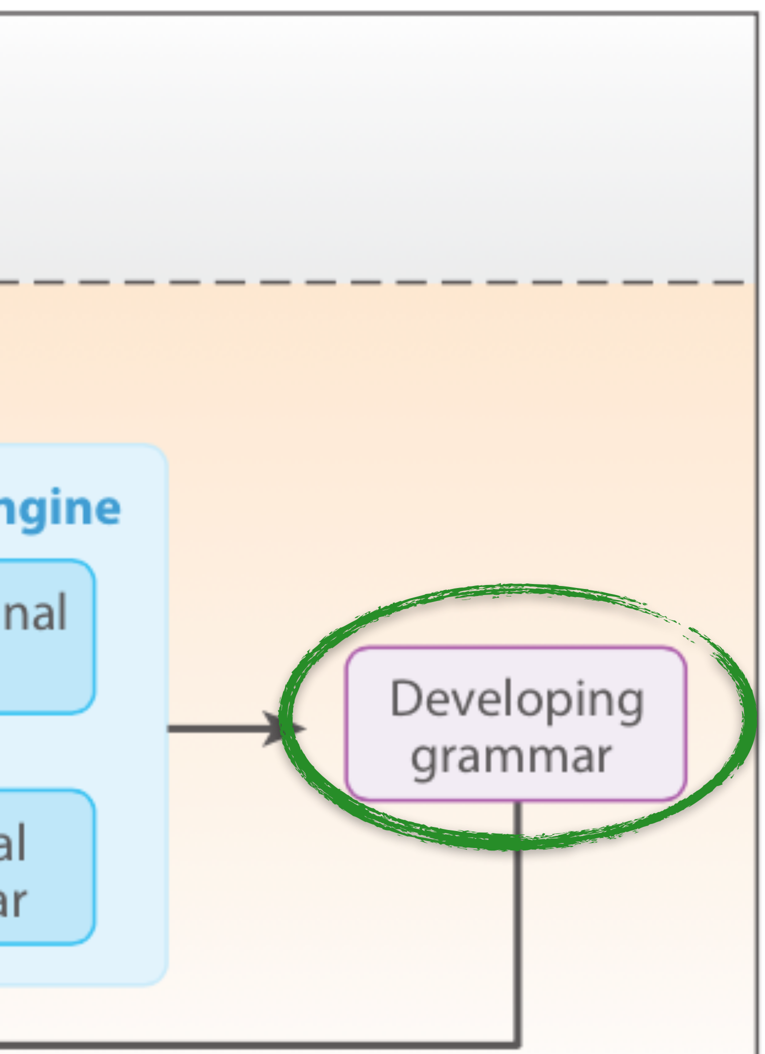
another one

"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!"

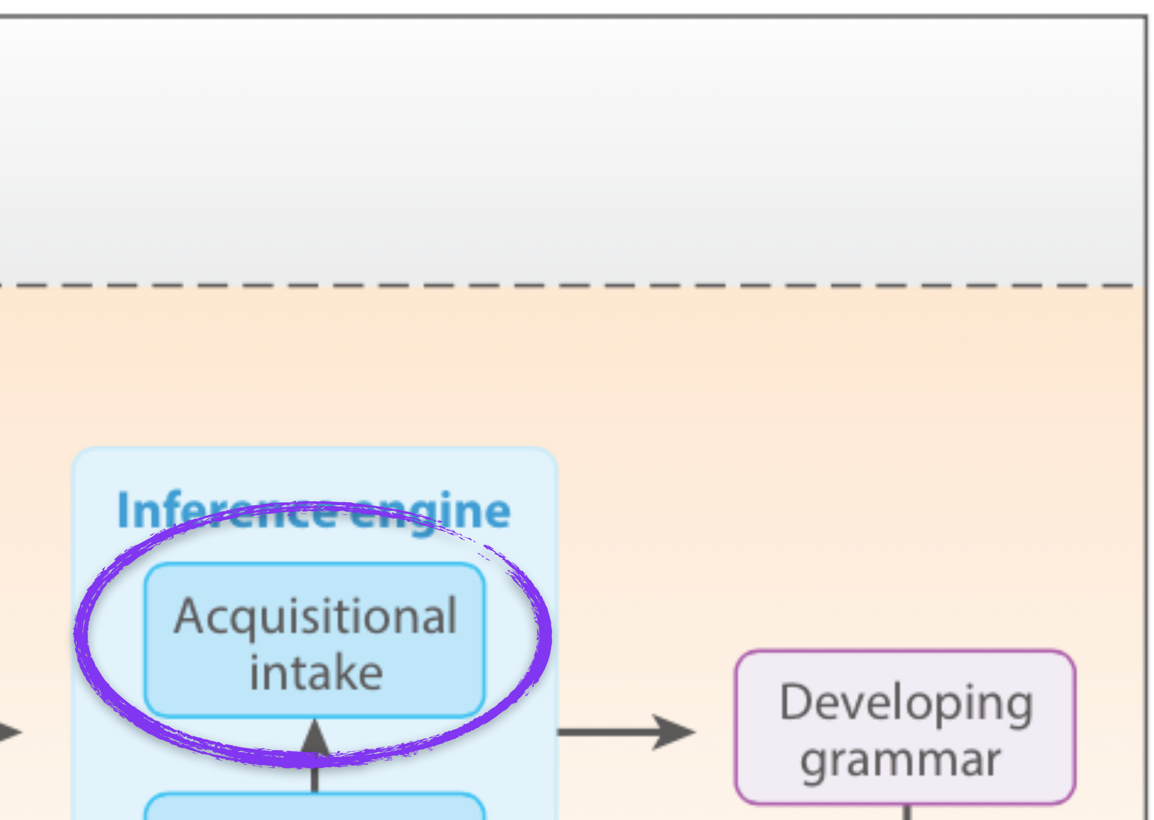
Algorithmic-level

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

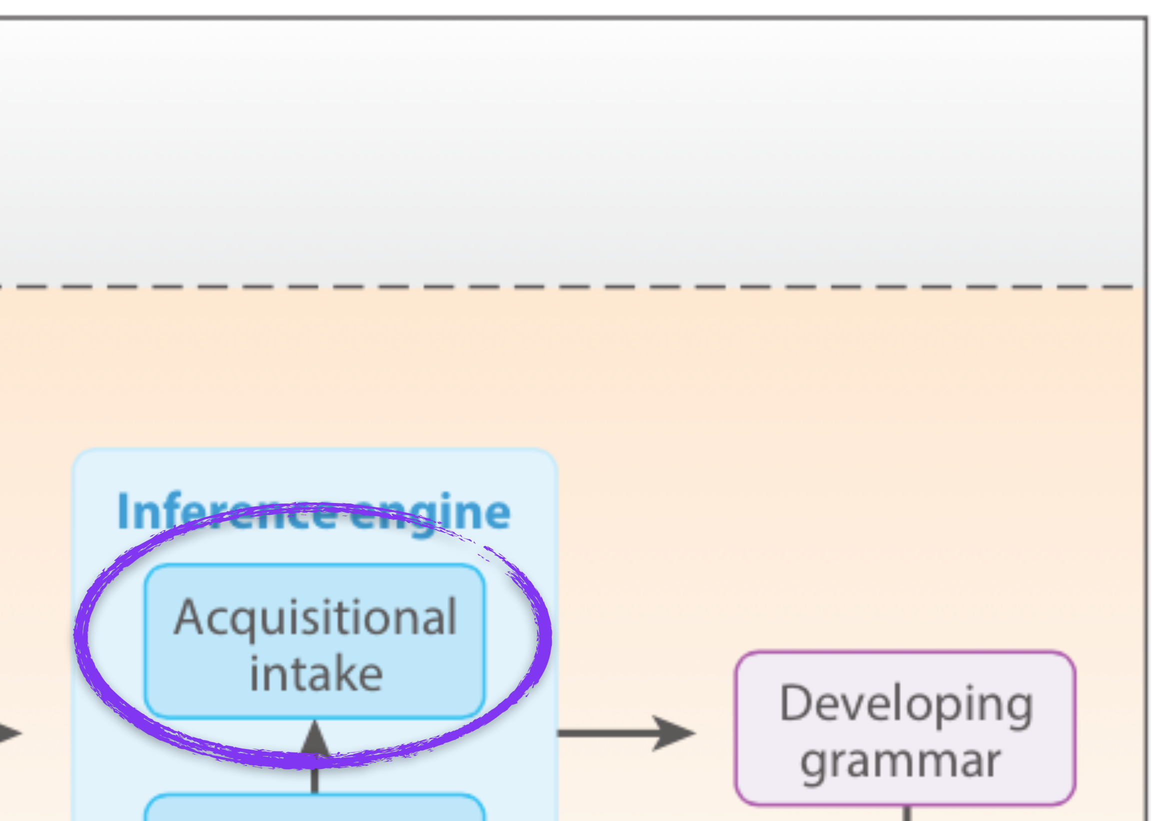
another one

"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 ?"

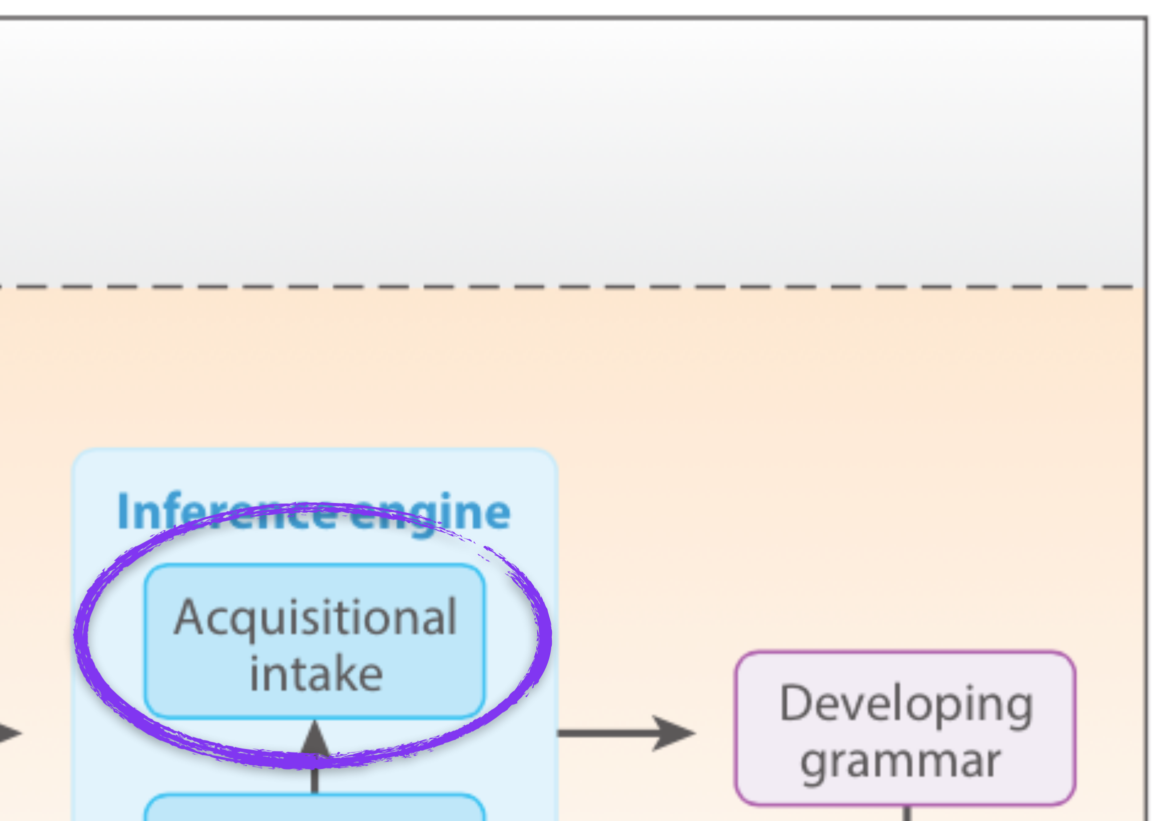
Algorithmic-level

✓

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!”

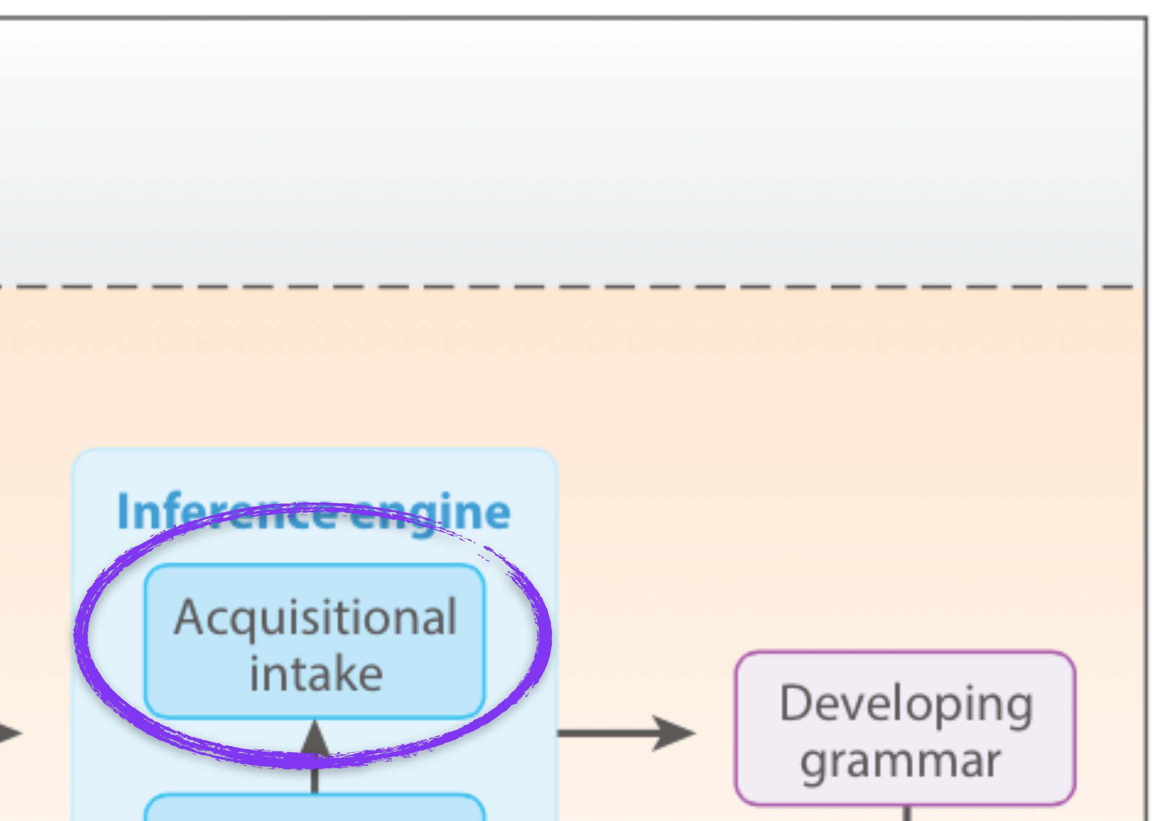
“Do you see another one ?”

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!"

Algorithmic-level

Strategy 1: Ignore ✓

Less robust

Strategy 2: Include other ✗

More robust

Noun'

pretty kitty

"Do you see another one ?"

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.

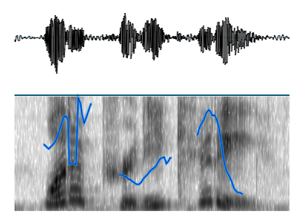
Pearl & Mis 2016

What we can learn

metrical phonology

syntactic categorization

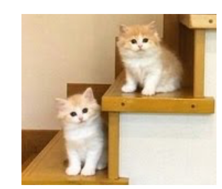
speech segmentation



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___ ?

Jack is somewhat tricksy.

He claimed he bought something.

What did Elizabeth wonder whether Jack bought ___?

Elizabeth worried it was something dangerous.

What did Elizabeth worry if Jack bought ___ ?

What we can learn

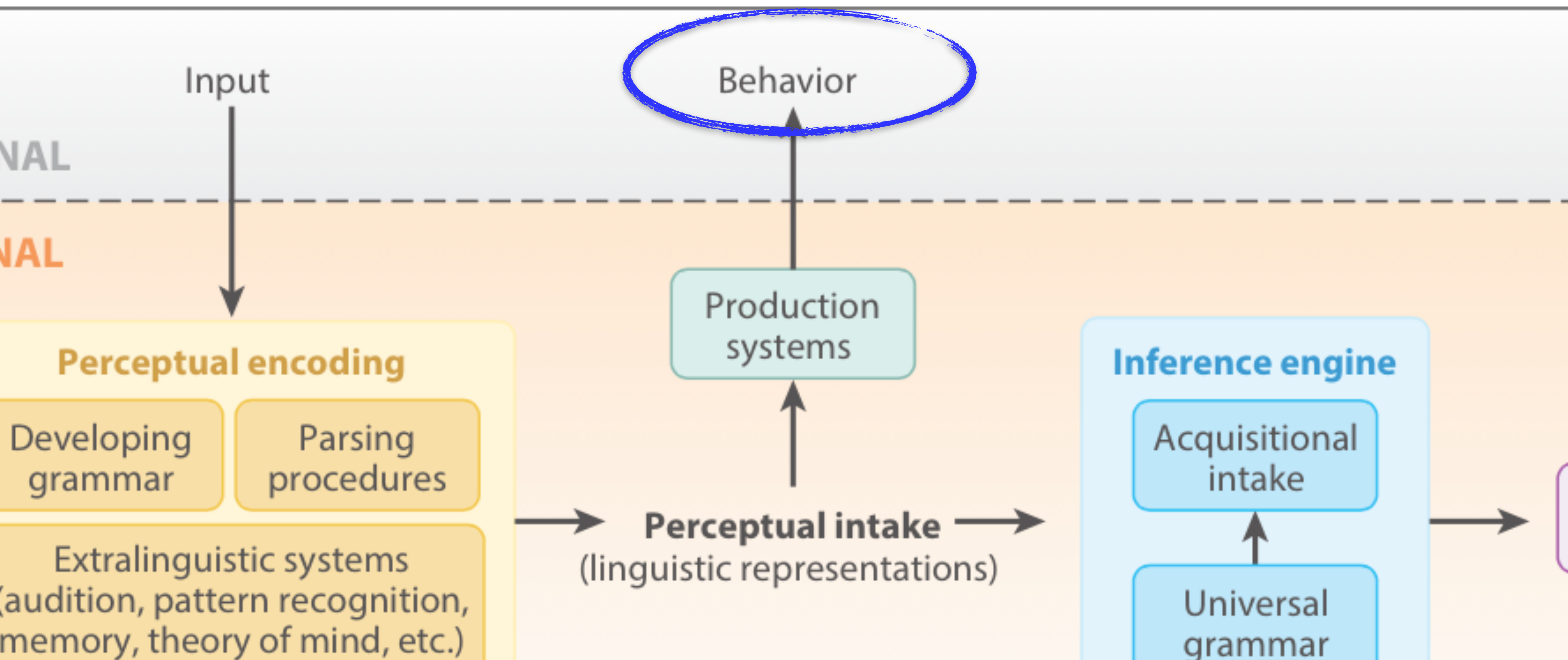
syntax

Who does Lily think the kitty for is pretty?

syntactic island

Who does Lily think the kitty for ___ is pretty?

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

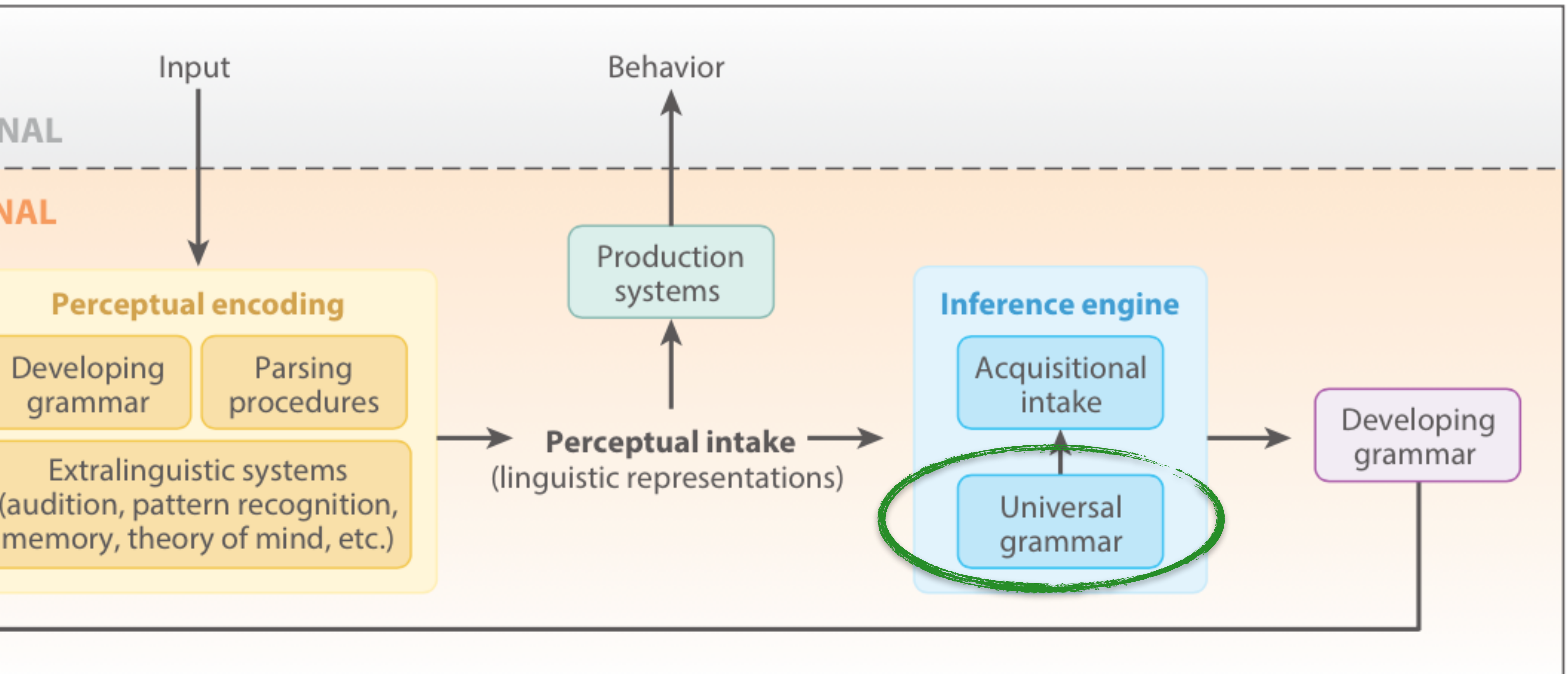
syntactic island

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

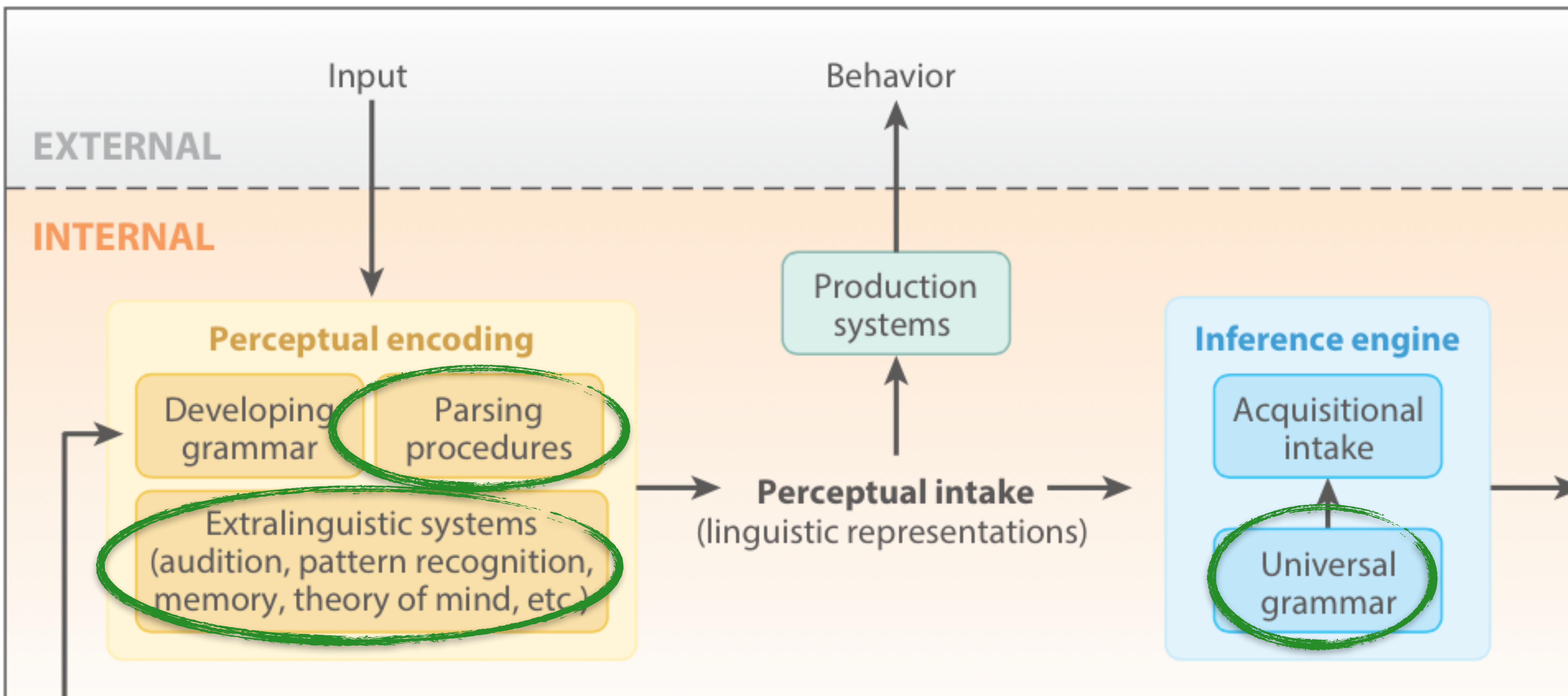
syntactic island

Who does Lily think the kitty for is pretty?

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

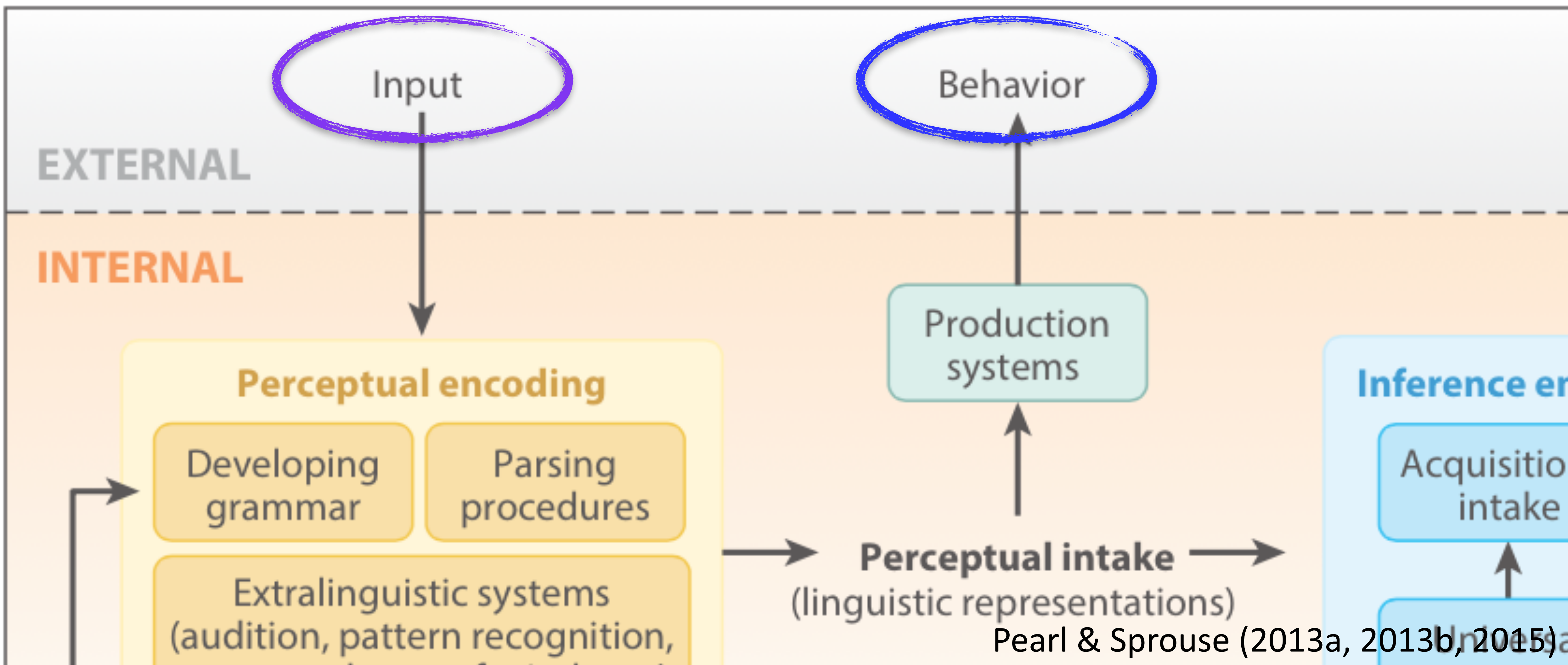
syntax

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

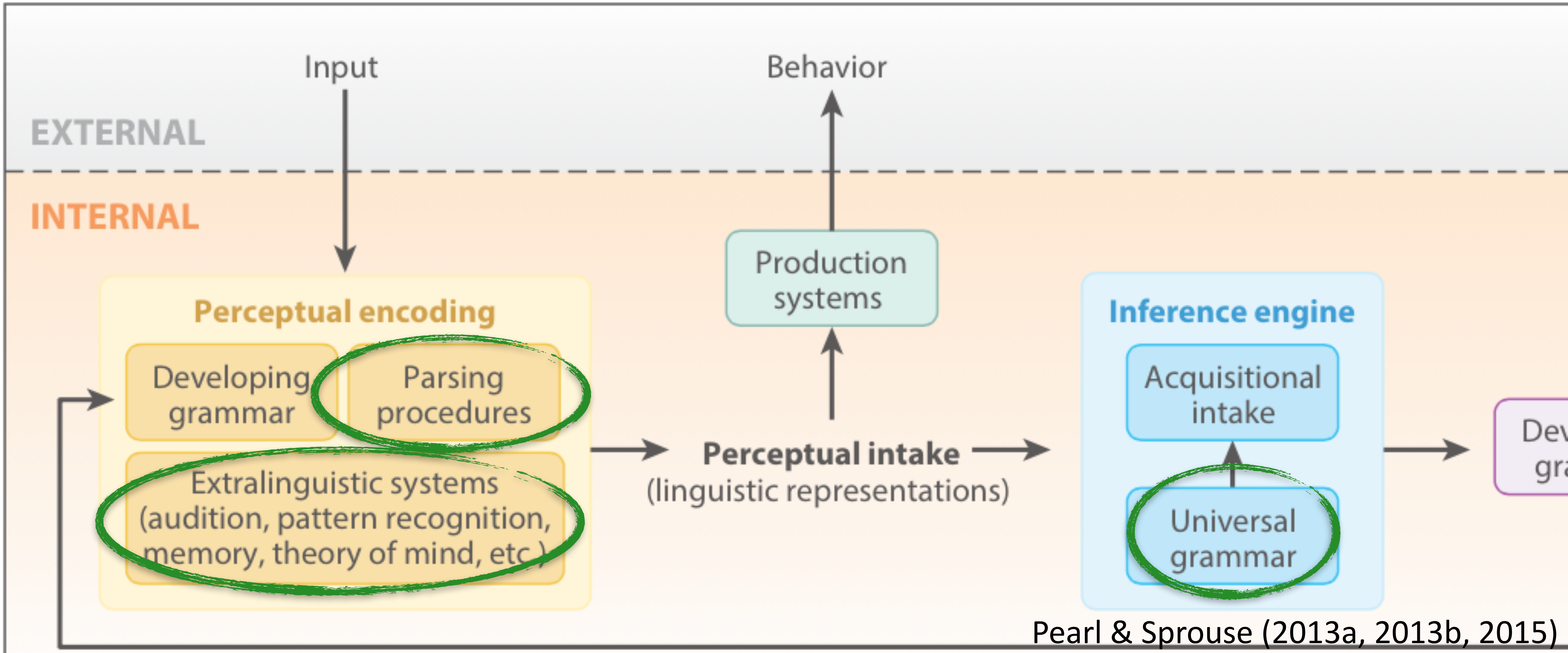
syntactic island

Who does

Lily think the kitty for is pretty?

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.



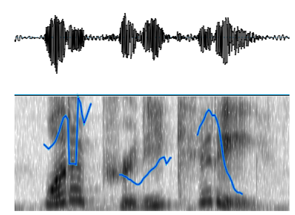
Pearl & Sprouse (2013a, 2013b, 2015)

What we can learn

metrical phonology

syntactic categorization

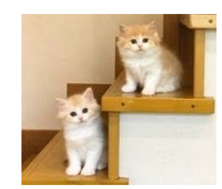
speech segmentation



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”

✗ 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"

$\forall$ $\neg$

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

“**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.

5-year-olds

But why?

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

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

X " **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"

$\forall$ $\neg$

??

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”

$\forall$ $\neg$

??

inverse $\neg$ $\forall$ 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 didn**’t** sit on the stairs”

$\forall$ $\neg$

??

Not all kitties sat on the stairs.

inverse $\neg$ $\forall$

QUD

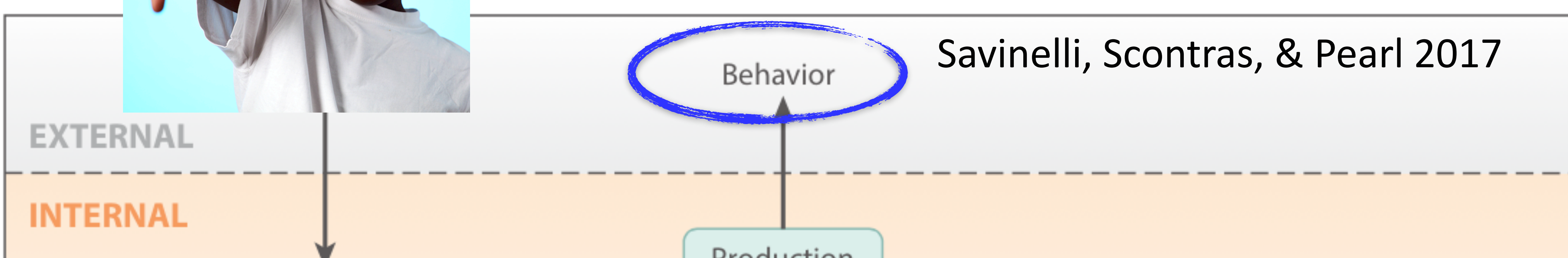
grammatical processing

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

X "**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

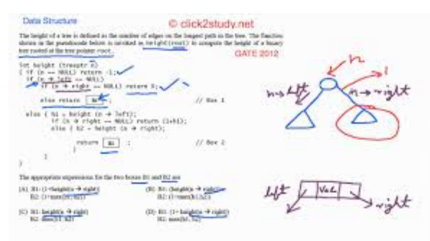
The pragmatic factors seem to be the driving force behind children's behavior. This suggests that 5-year-olds are still developing their ability to manage the pragmatic context of a conversation as well as adults do.

Savinelli, Scontras, & Pearl 2017

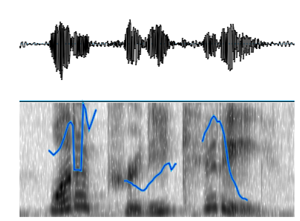
Today's Plan:
Computational models of language acquisition

I. Why

II. How



III. What we can learn



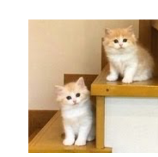
Noun

✔KI tty

Who does… is pretty?

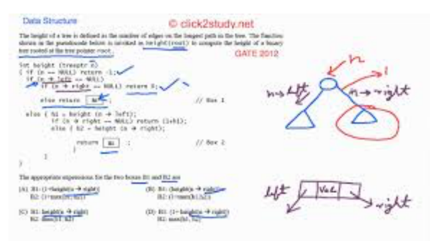
another one

Every kitty didn't …

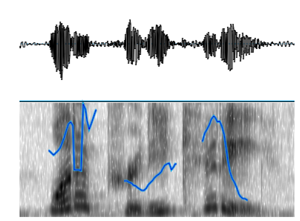


Today's Plan:
Computational models of language acquisition

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



II. How



III. What we can learn

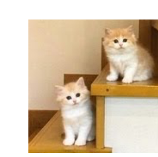
Noun

✓KI tty

Who does… is pretty?

another one

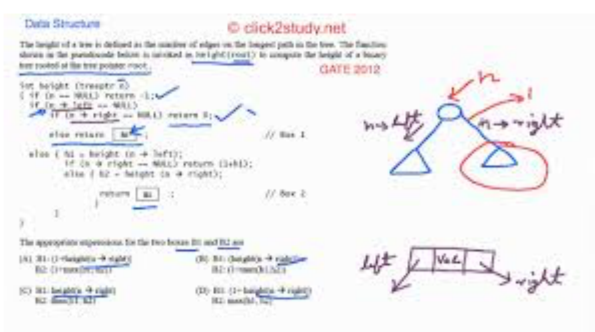
Every kitty didn't …



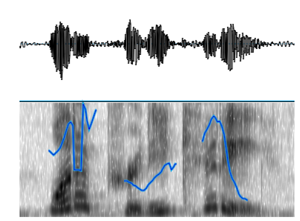
Today's Plan:
Computational models of language acquisition

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



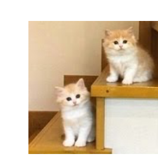
Noun

KI tty

Who does… is pretty?

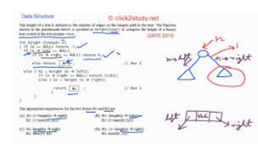
another one

Every kitty didn't …



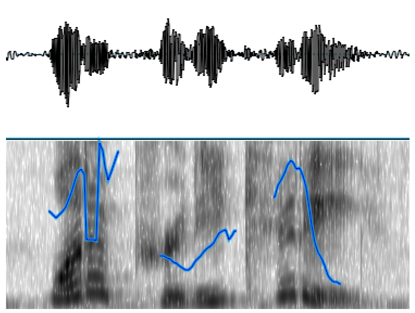
Today's Plan:
Computational models of language acquisition

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

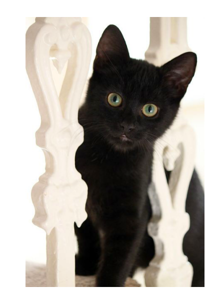
Who does… is pretty?

syntax

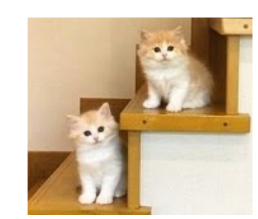
Noun

syntactic categorization

another one



syntax, semantics

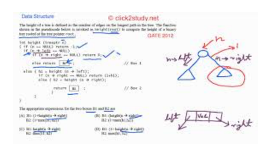


pragmatics

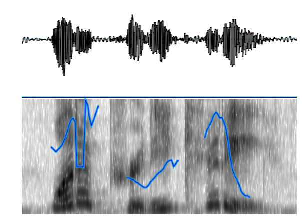
Every kitty didn't …

Today's Plan:
Computational models of language acquisition

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



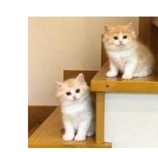
II. How : By building informative computational models

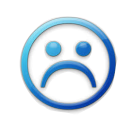


Noun

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

✓KI tty





another one

Who does… is pretty?

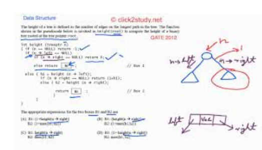
Every kitty didn't …

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

Thank you!

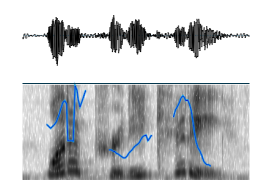
Lawrence Phillips Timothy Ho Zephyr Detrano
Alandi Bates Sue Braunwald Galia Bar-sever
Jon Sprouse Ben Mis Greg Scontras K.J. Savinelli
Jeff Lidz Members of CoLaLab

Audiences at: CSUF 2016, GLEEFUL 2016, GALANA 2015



Computation of Language Laboratory

UC Irvine



Noun

Who does... is pretty?

KI tty

another one

Every kitty didn't ...

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.