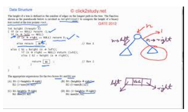
Lisa S. Pearl
Associate Professor
Department of Linguistics
Department of Cognitive Sciences
SSPB 2219, SBSG 2314
University of California, Irvine
lpearl@uci.edu

Computation of Language Laboratory
UC Irvine

Using Computational Modeling to Understand Language Acquisition

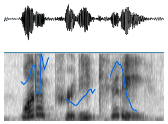
UC Computational Social Science



Noun

Kl tty

Who does… is pretty?



another one

Every kitty didn't …

Language acquisition: How humans learn language knowledge

Language acquisition: How humans learn language knowledge

First language acquisition = Learning native language(s)

Happens as a young child

Language acquisition: How humans learn language knowledge

Second language acquisition = Learning non-native/foreign language(s)

Happens as an older child or adult

First language acquisition

First language acquisition

Computation of Language Laboratory

UC Irvine

How do children acquire the knowledge about language that they do from the language data they have?

Why first language acquisition?

Babies are amazing at learning language

Babies are amazing at learning language

And they learn *a lot*!

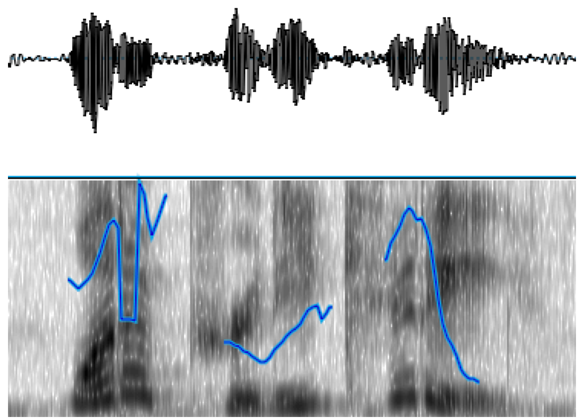
And they learn *a lot*!

Like what?

Like what?

Everything you know about your native language(s).

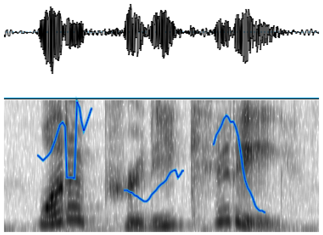
You know how to identify words in fluent speech (speech segmentation)



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!



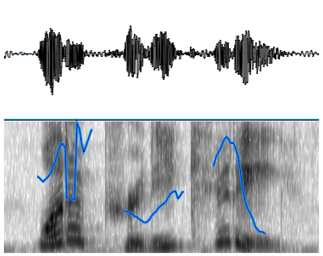
what a pretty kitty!

speech segmentation

You know how to pronounce words (metrical phonology)

✓ KI tty

✗ ki TTY



what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

metrical phonology

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

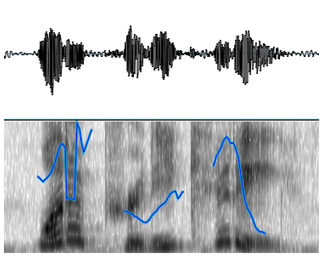
Noun

what a pretty ___!

owl

penguin

kitty



what a pretty kitty!

speech segmentation

KI tty

ki TTY

metrical phonology

Noun

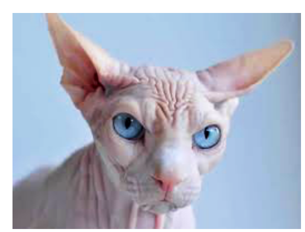
penguin owl

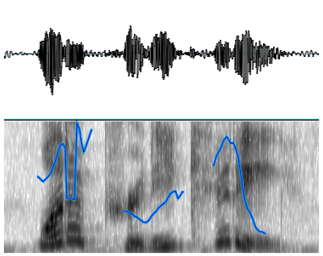
kitty

syntactic categorization

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

“Oh look — a pretty kitty!”
“Look — there’s another one!”





what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

metrical phonology

Noun

penguin owl

kitty

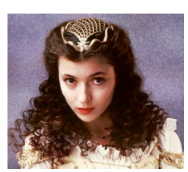
syntactic categorization

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

syntax, semantics

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

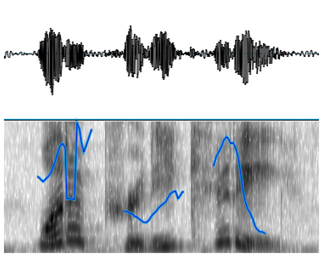
This kitty was bought as a present for someone.



Lily thinks this kitty is pretty.

☹

Who does Lily think the kitty for is pretty?



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!"

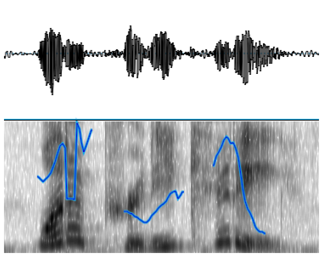
syntax, semantics

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.



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!"

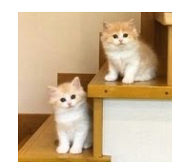
"Look — there's another one!"

syntax, semantics

"Every kitty didn't sit on the stairs"

Not all kitties sat on the stairs.

pragmatics



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.

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

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?

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

Given the available input, information processing done by human minds

Look at that kitty!
There's another one.

Input

Where did he hide?
What happened?

processing & generalization

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

words & morphemes

metrical phonology

syntactic categories

syntax

semantics

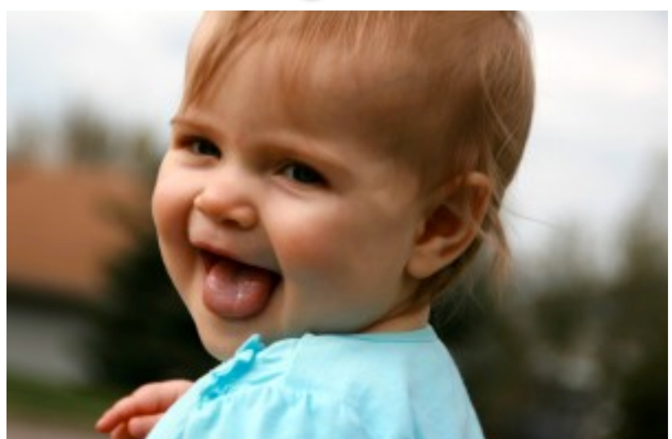
pragmatics

processing & generalization

Look at that kitty!
There's another one.

Input

Where did he hide?
What happened?



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

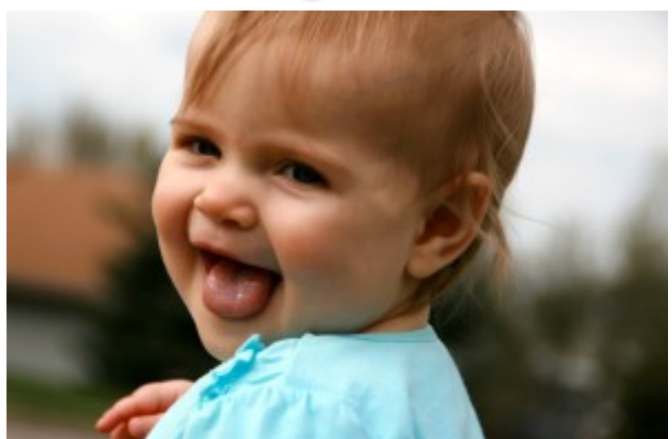
Given the available input, information processing done by human minds to build a system of linguistic knowledge whose output we observe

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.

To understand how children solve this acquisition task, we need to think more about all the components involved.

words & morphemes

metrical phonology

syntactic categories

syntax

semantics

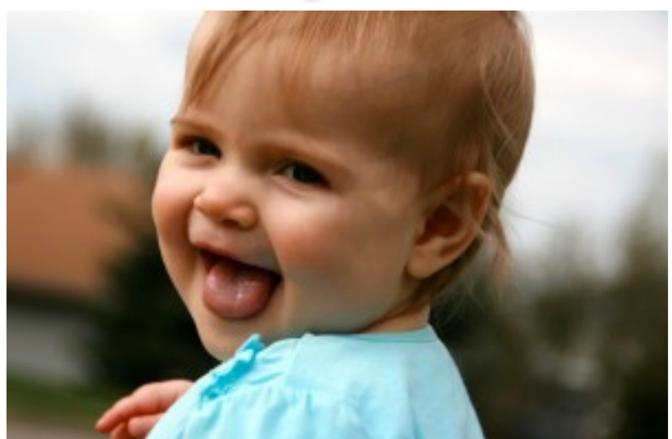
pragmatics

processing & generalization

Look at that kitty!
There's another one.

Input

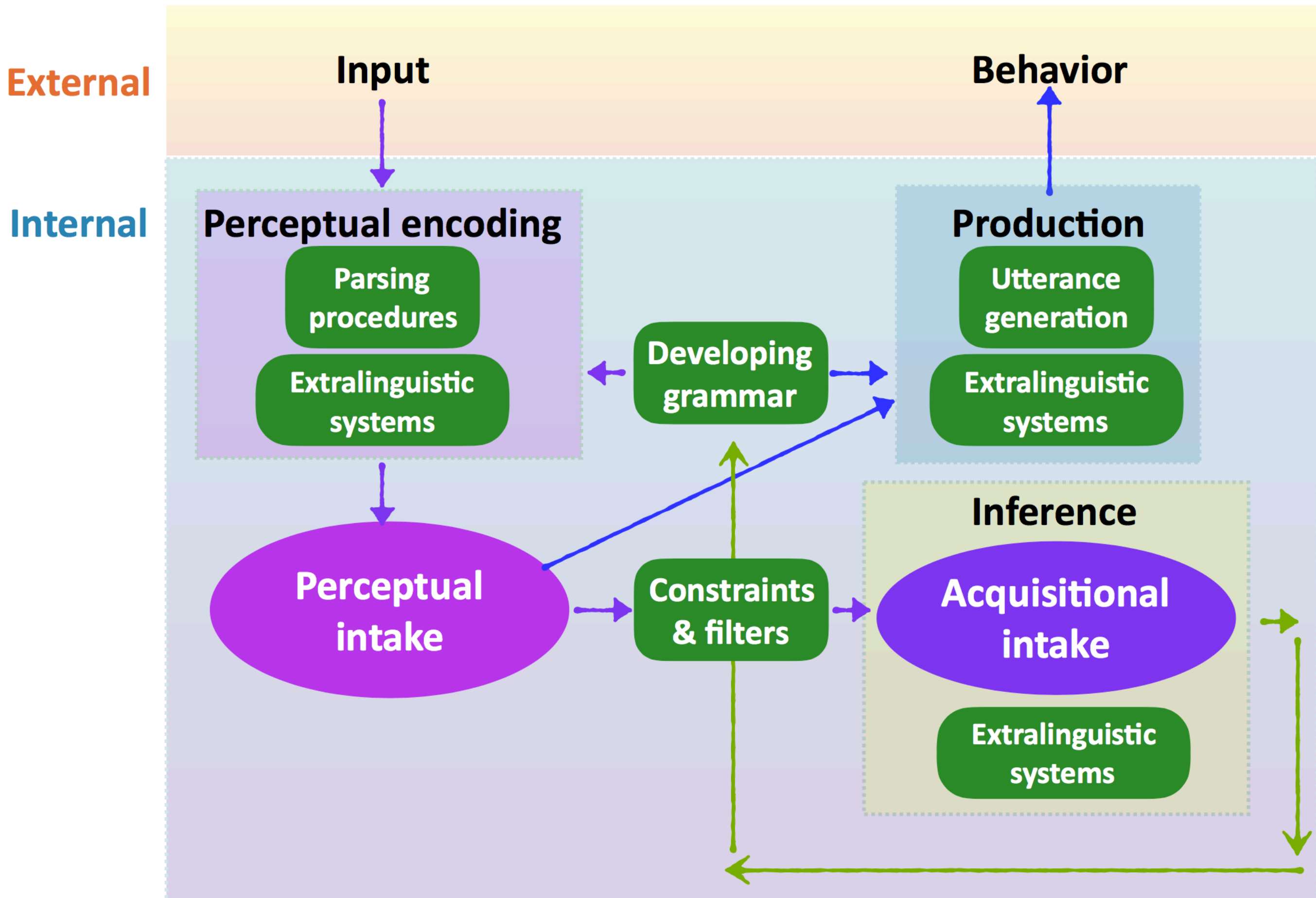
Where did he hide?
What happened?



Where's the kitty?

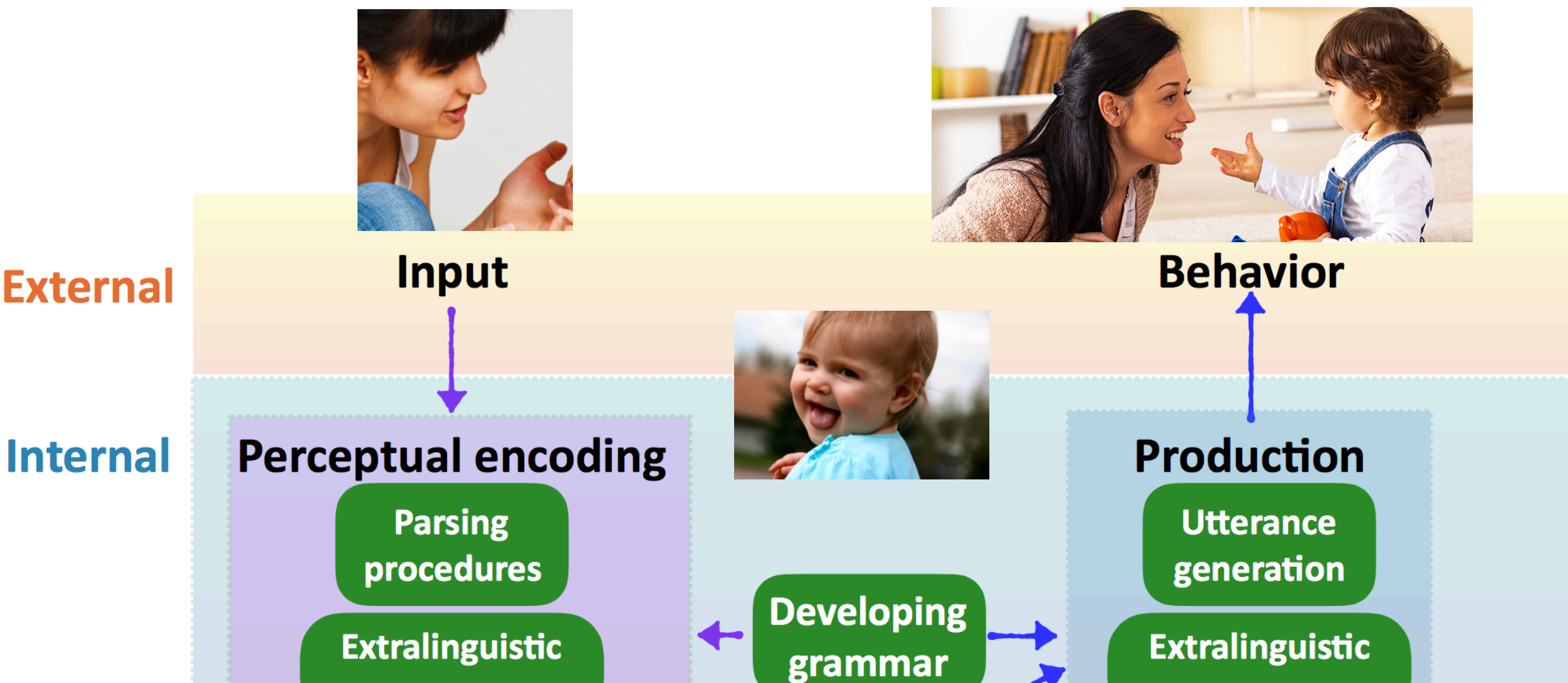
That one's really cute.

A framework that makes components of the acquisition task more explicit



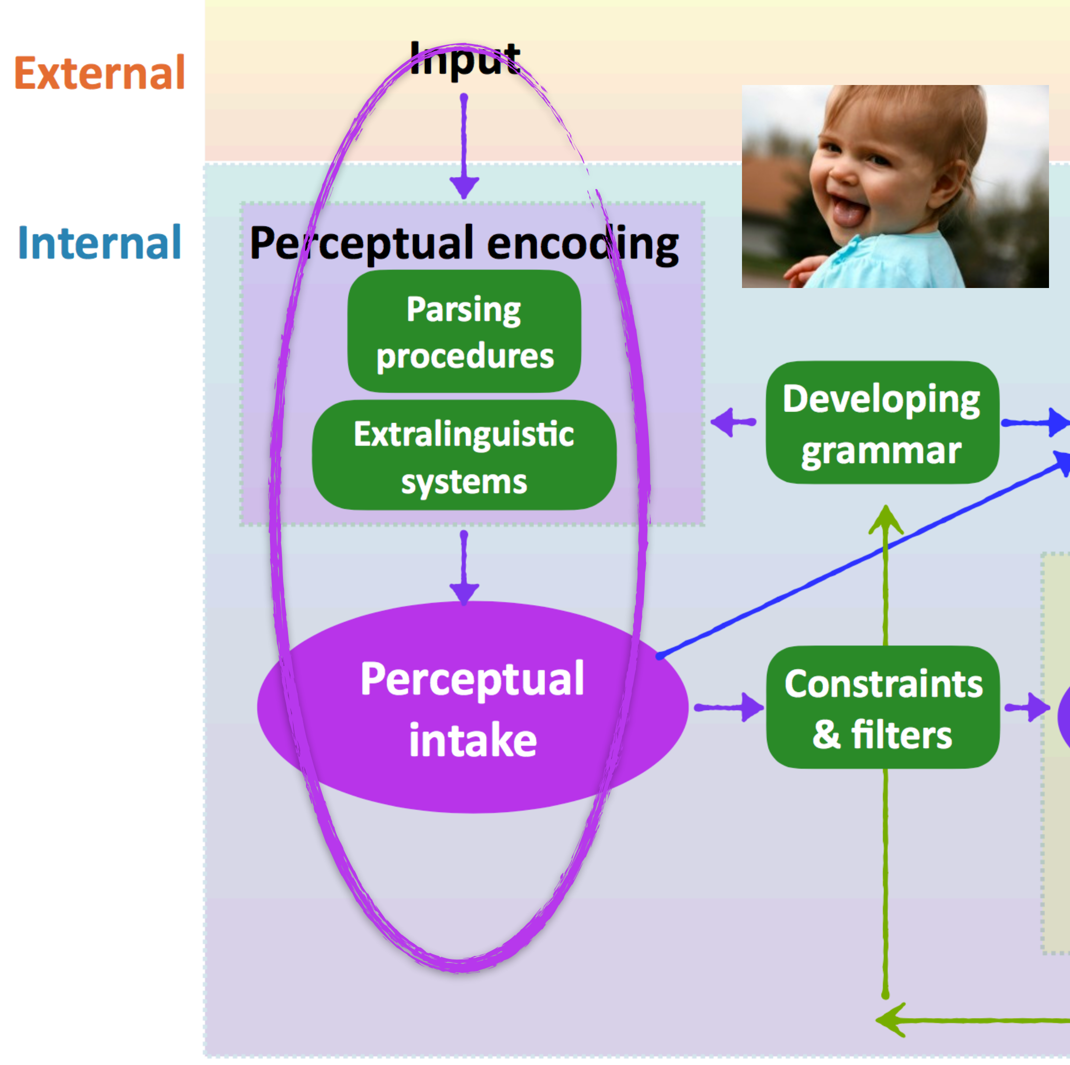
Adapted from Lidz & Gagliardi 2015

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



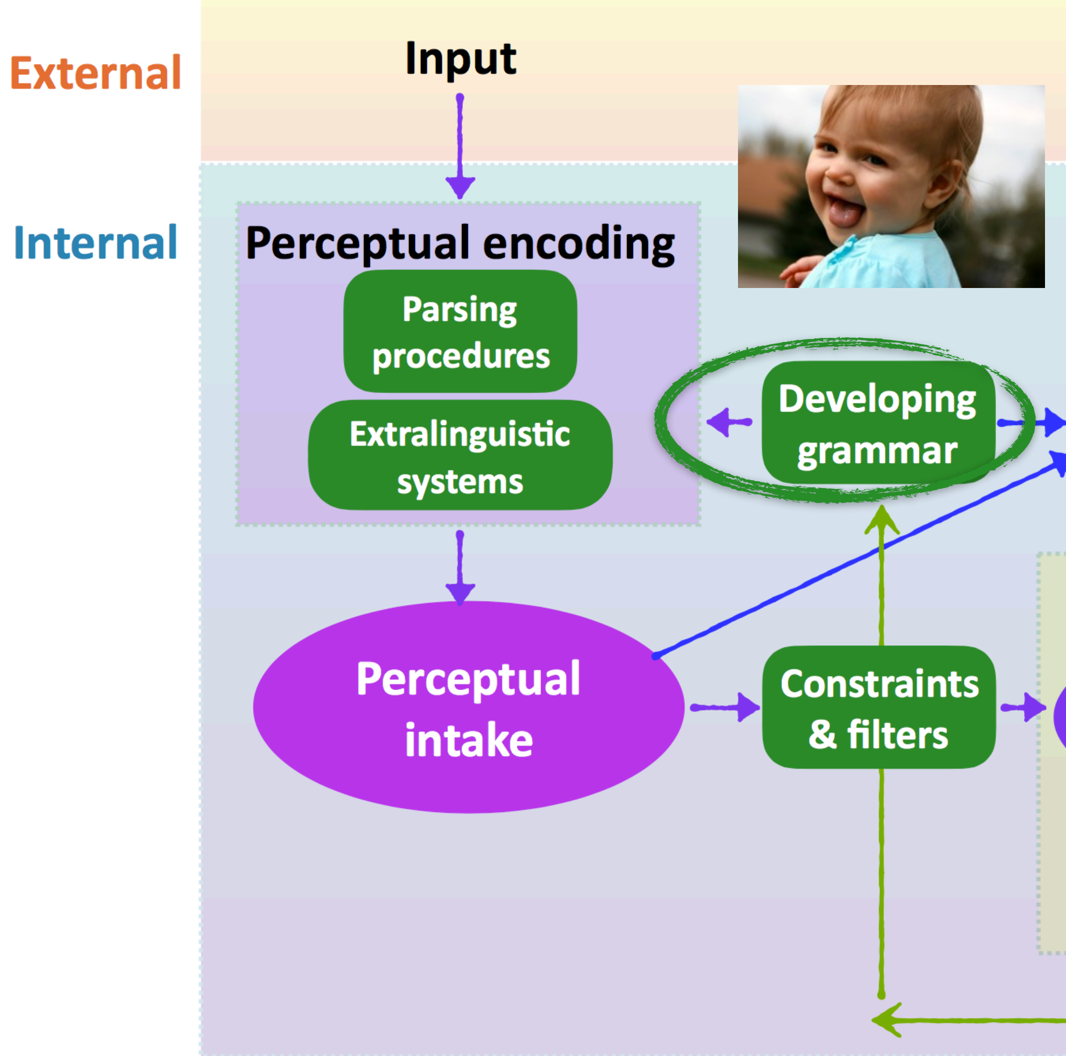
Perceptual encoding:

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



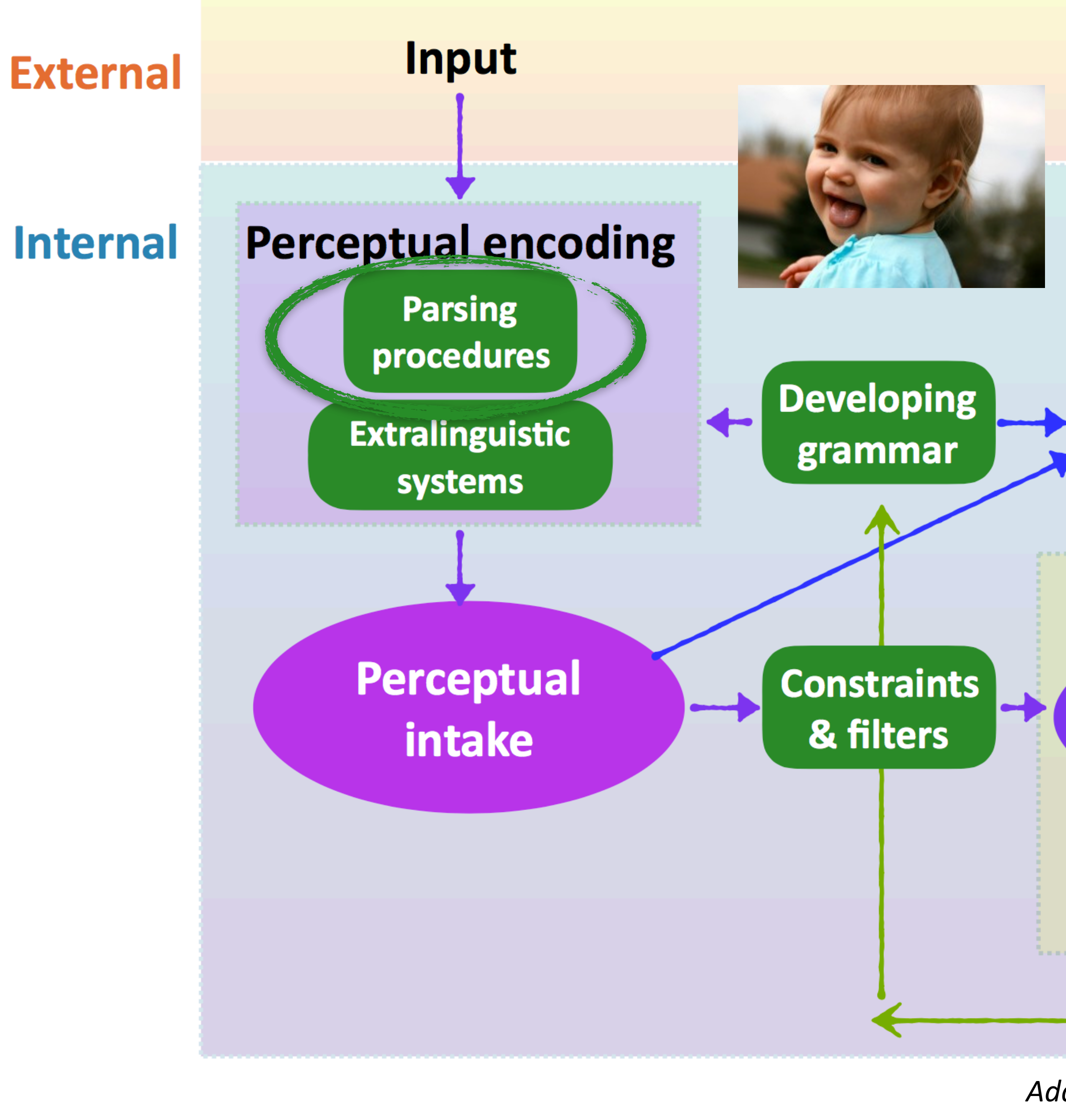
Perceptual encoding:

Involves using current knowledge of the language (the developing grammar)…



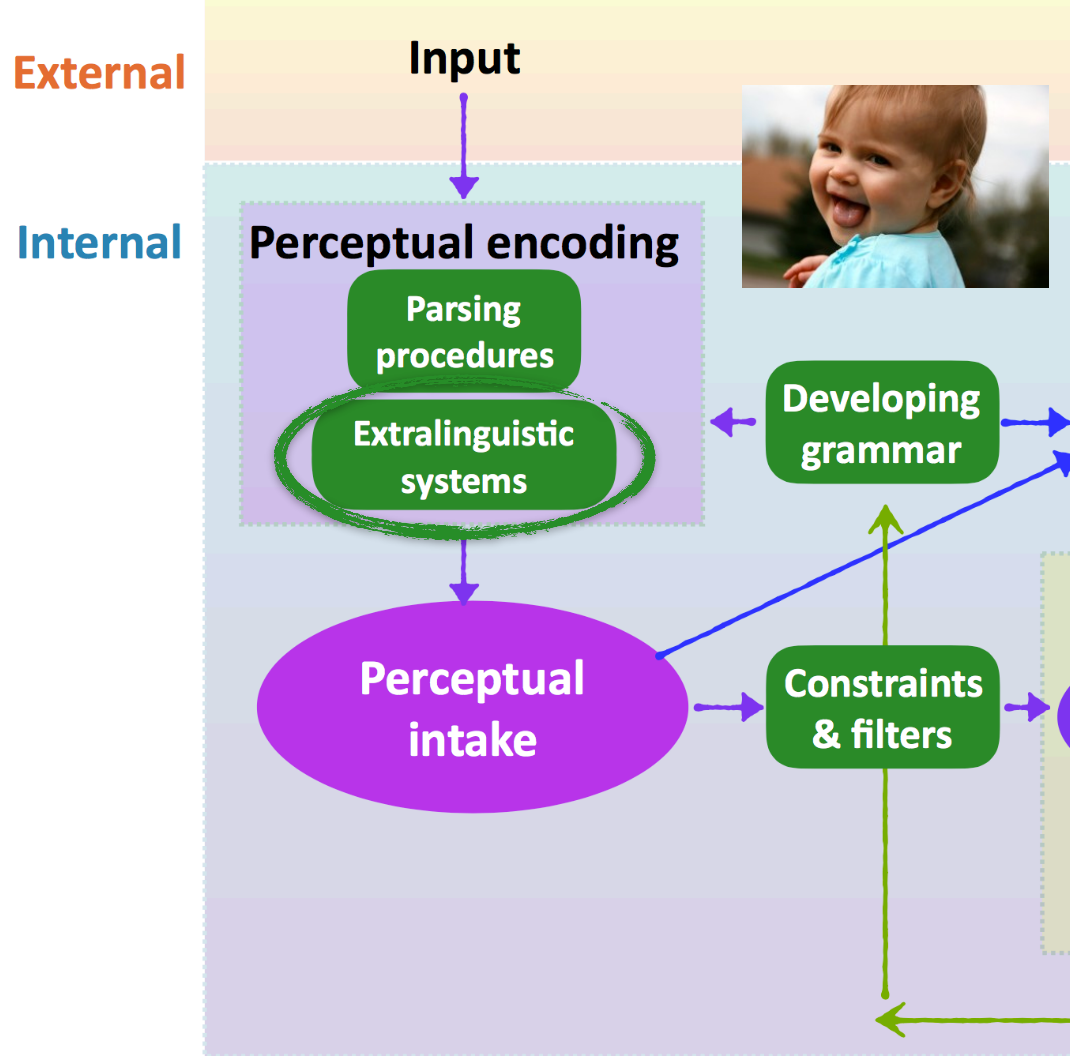
Perceptual encoding:

Involves using current knowledge of the language (the developing grammar) deployed in real time to parse the input…

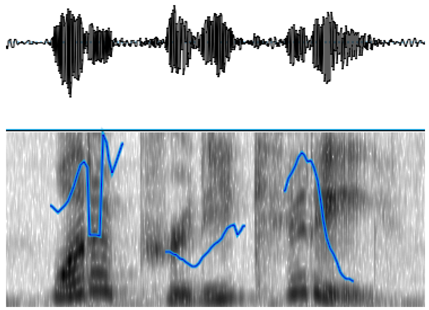


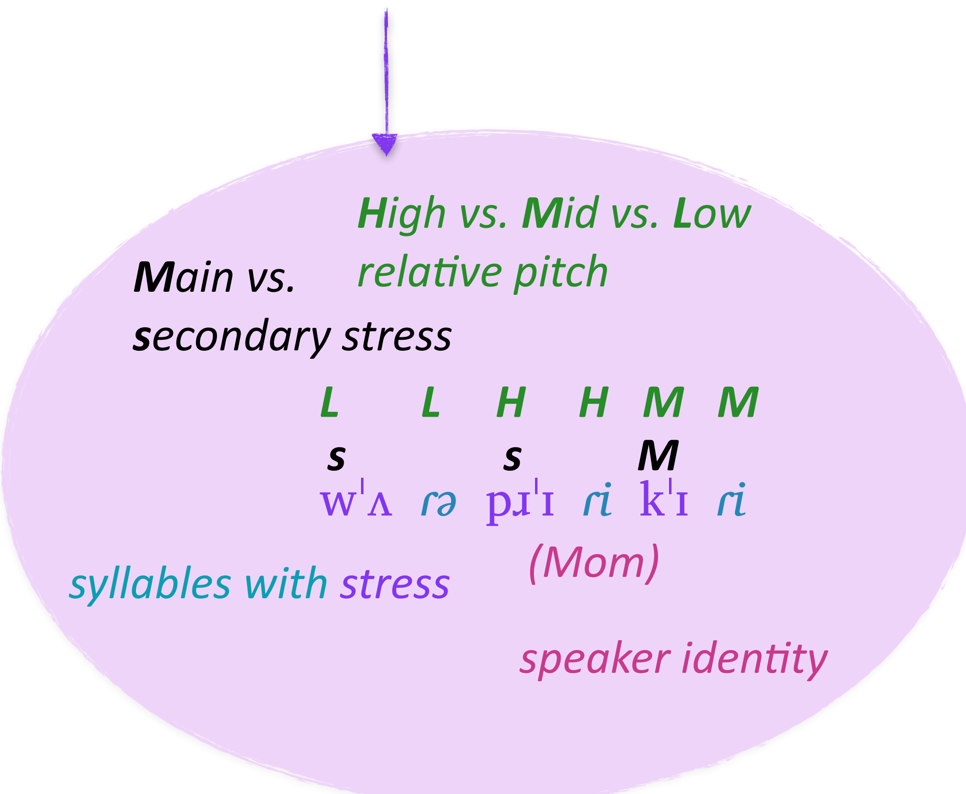
Perceptual encoding:

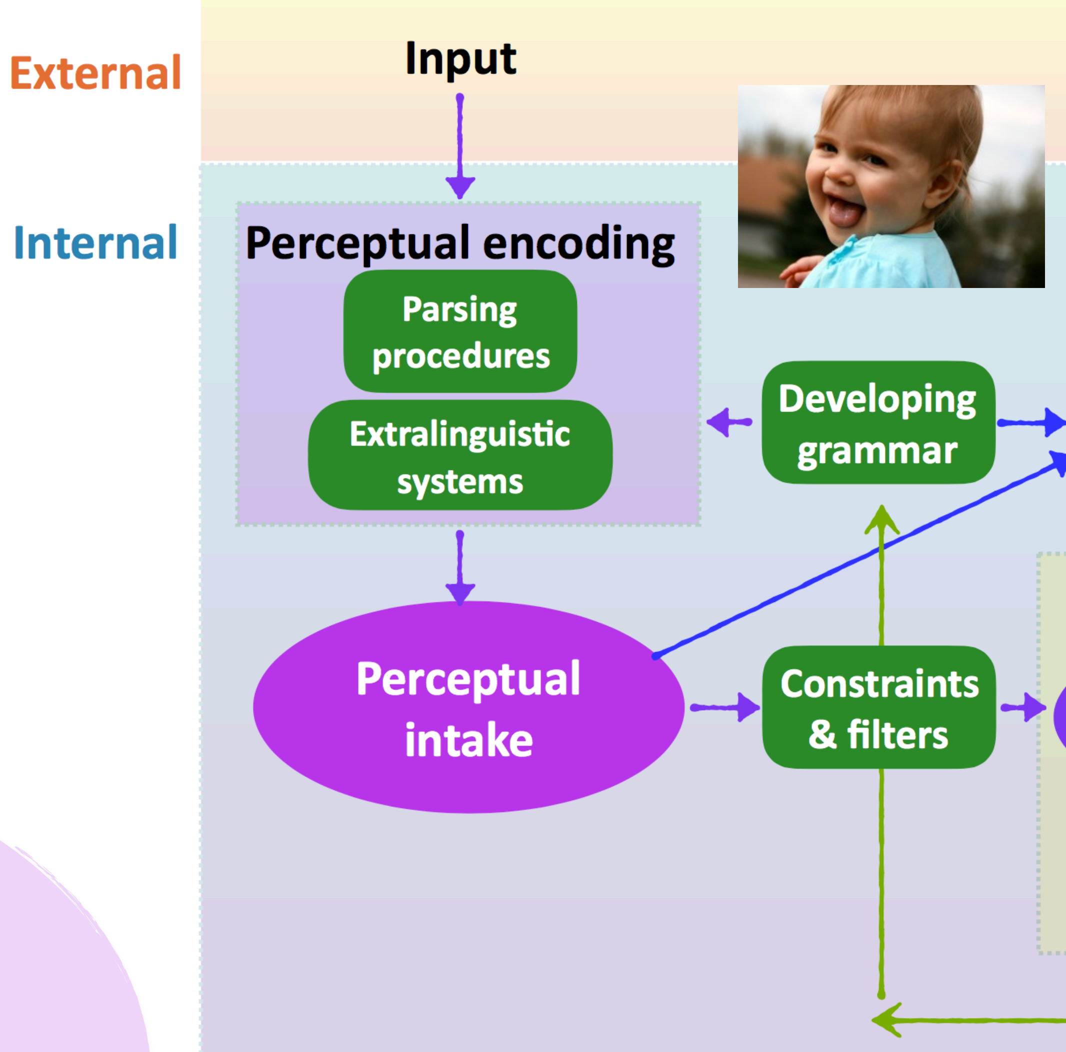
Involves using current knowledge of the language (the developing grammar) deployed in real time to parse the input, often drawing on extralinguistic systems (like working memory, auditory processing, etc.)



Perceptual encoding

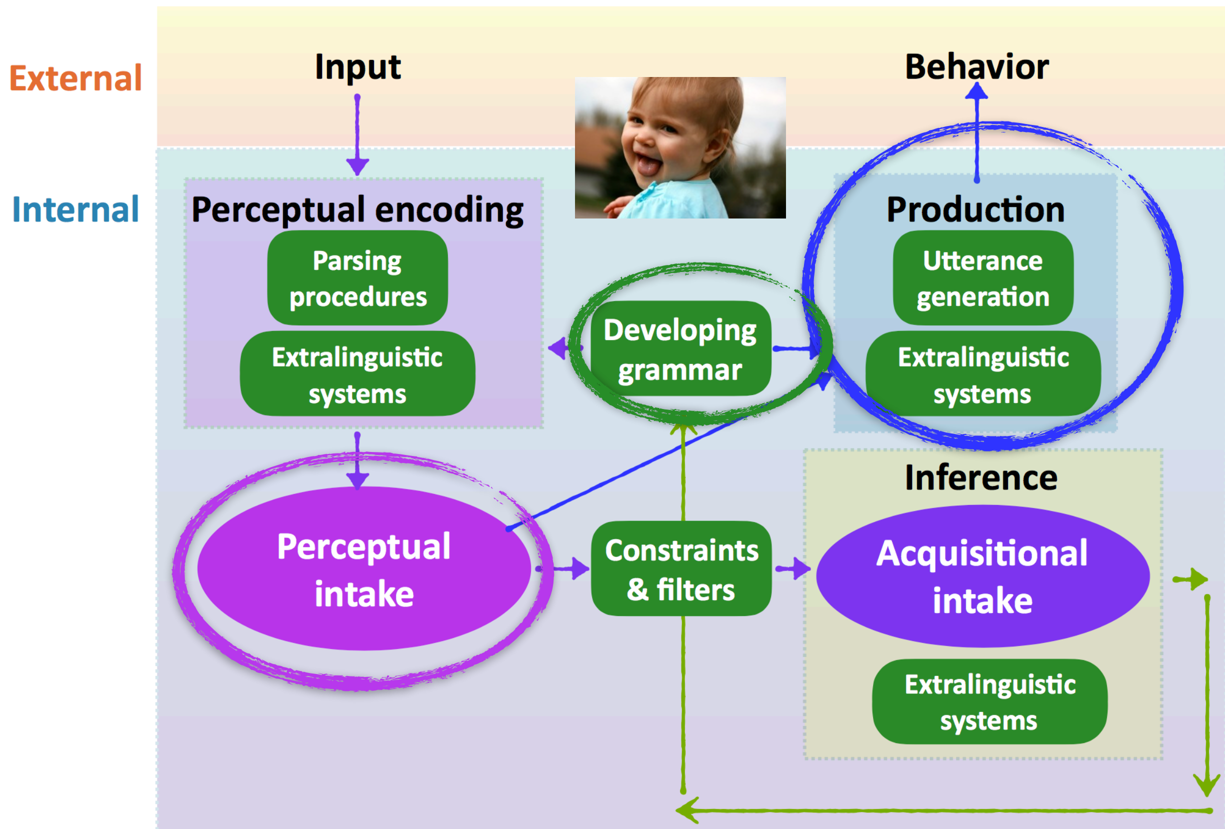






Generating observable behavior

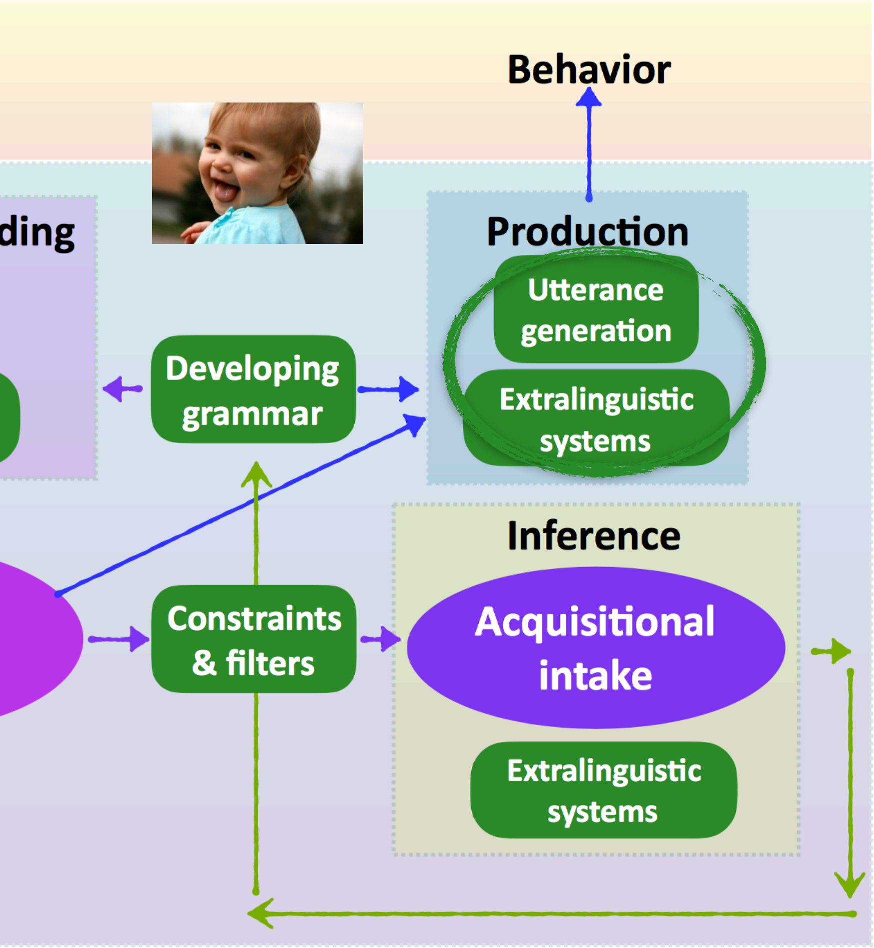
Involves the current linguistic representations and the developing grammar being used by the production system.



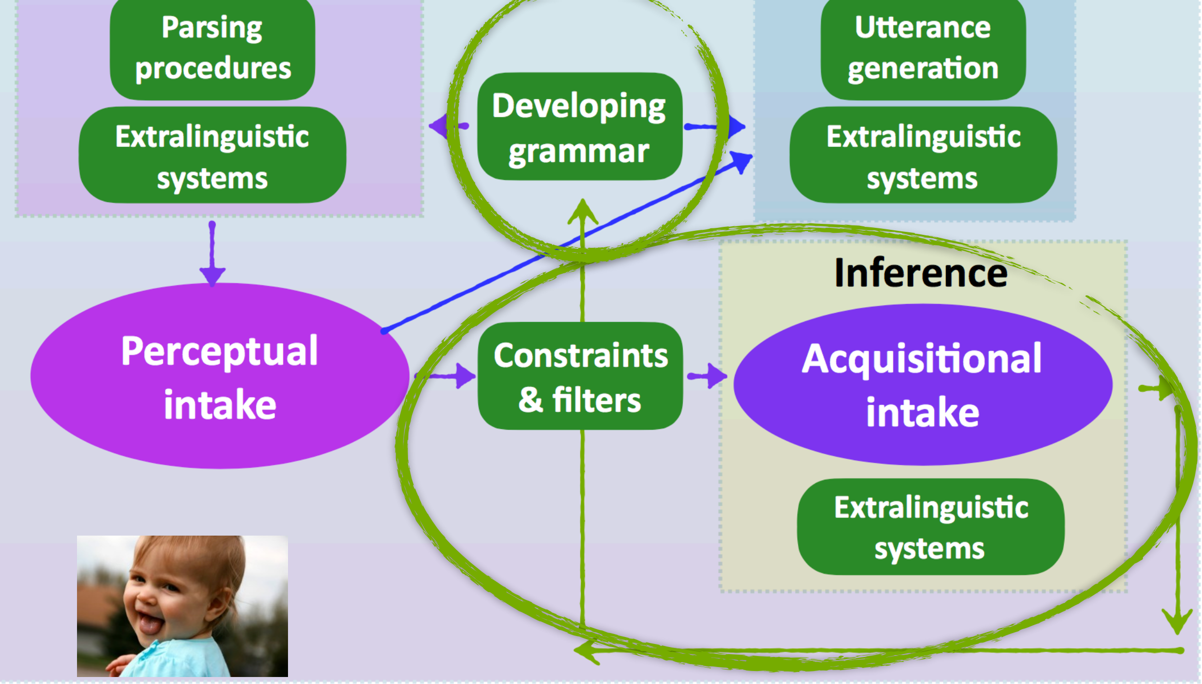
Adapted from Lidz & Gagliardi 2015

Generating observable behavior

These are used in real time to generate linguistic behavior (utterances) and non-linguistic behavior (pointing, looking, etc.). These behaviors require linguistic systems (utterance generation) and extralinguistic systems (motor control, attention, decision-making, etc.)



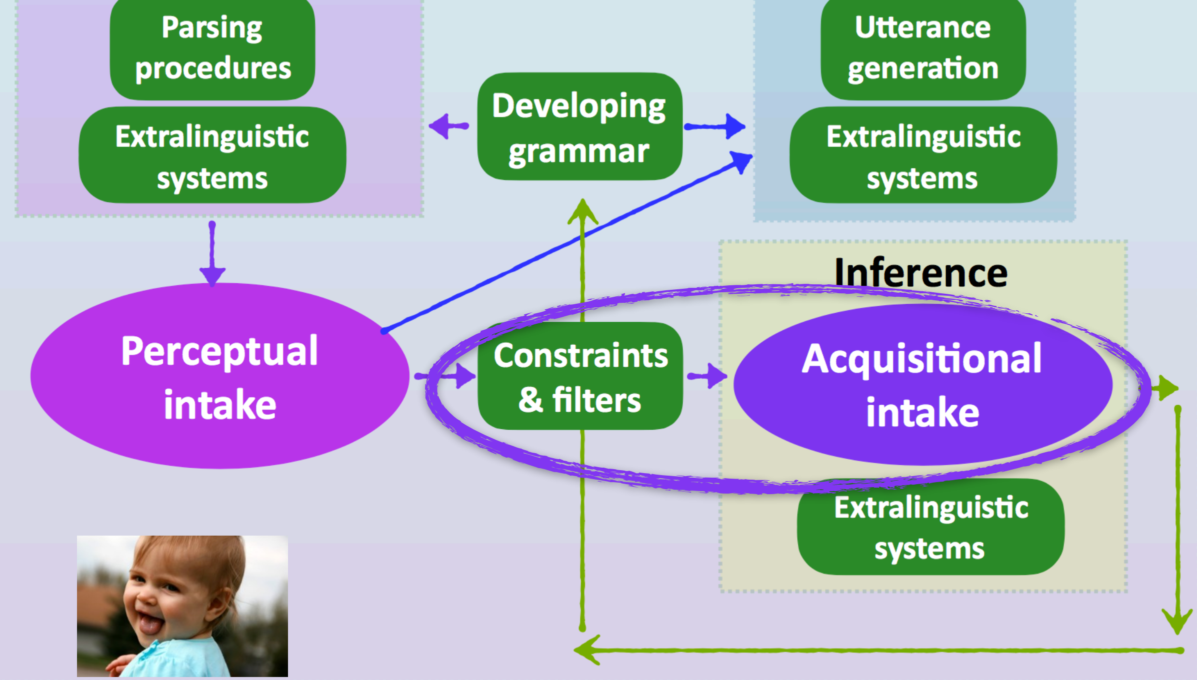
Adapted from Lidz & Gagliardi 2015



Adapted from Lidz & Gagliardi 2015

Inference = learning

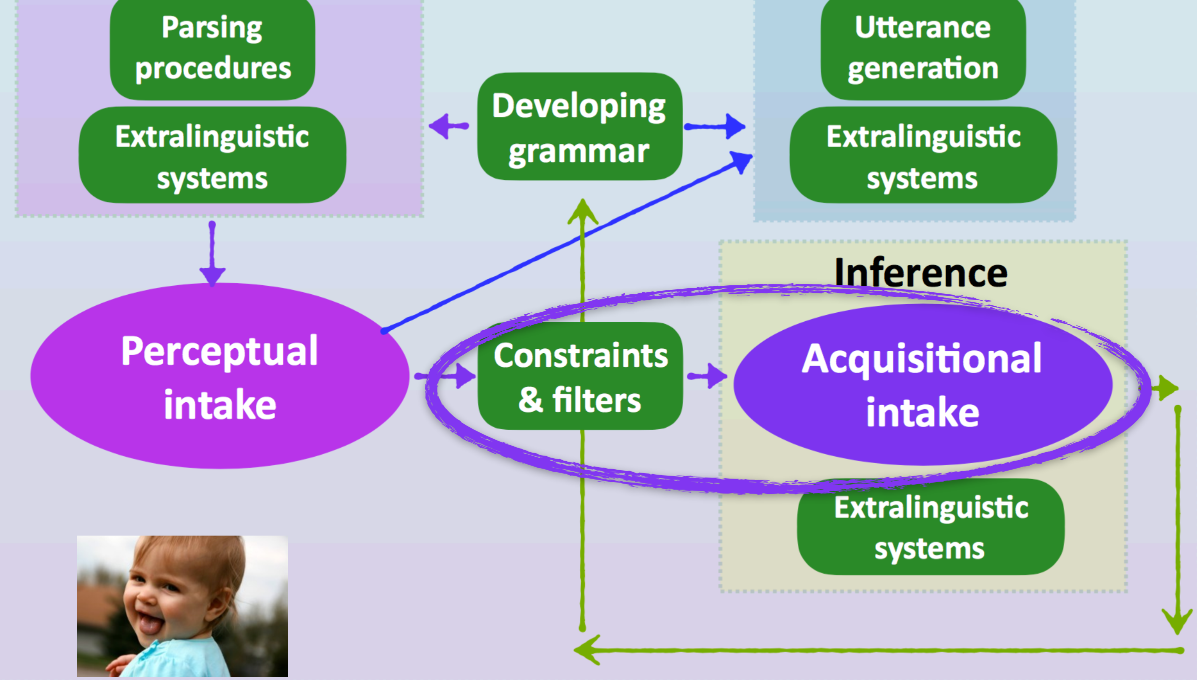
This is how children learn from the current data in order to update the developing grammar.



Adapted from Lidz & Gagliardi 2015

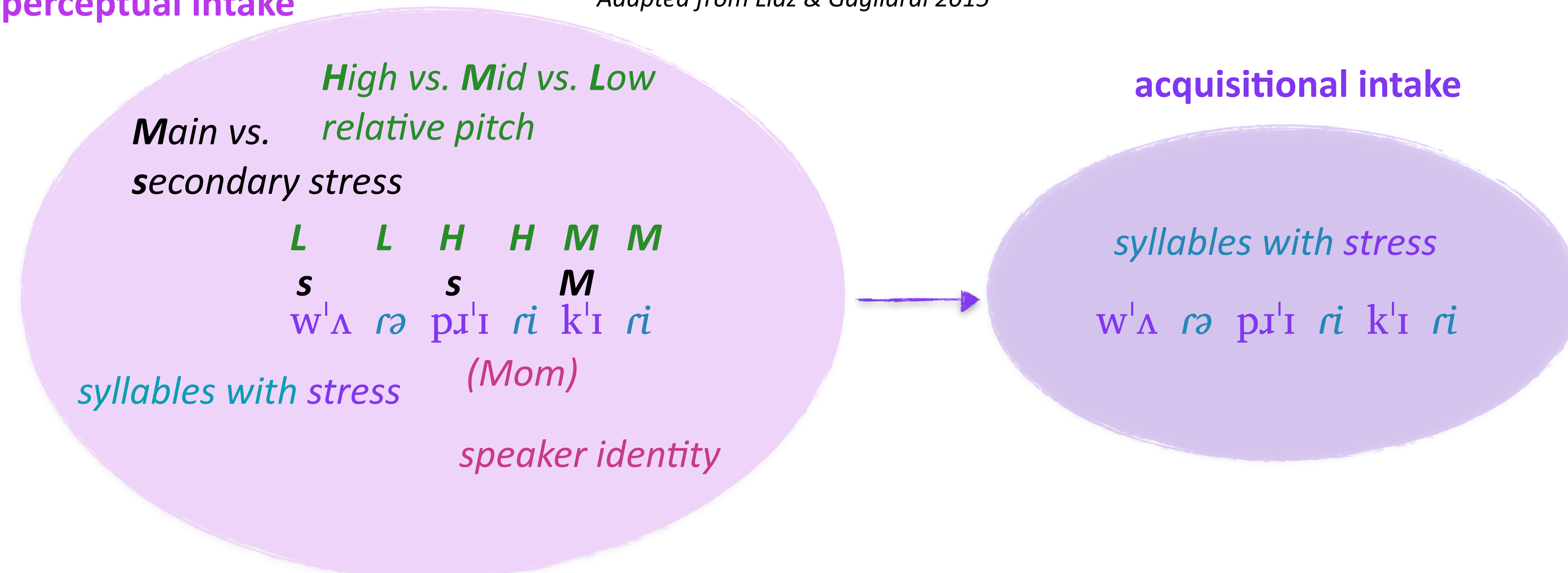
Inference = learning

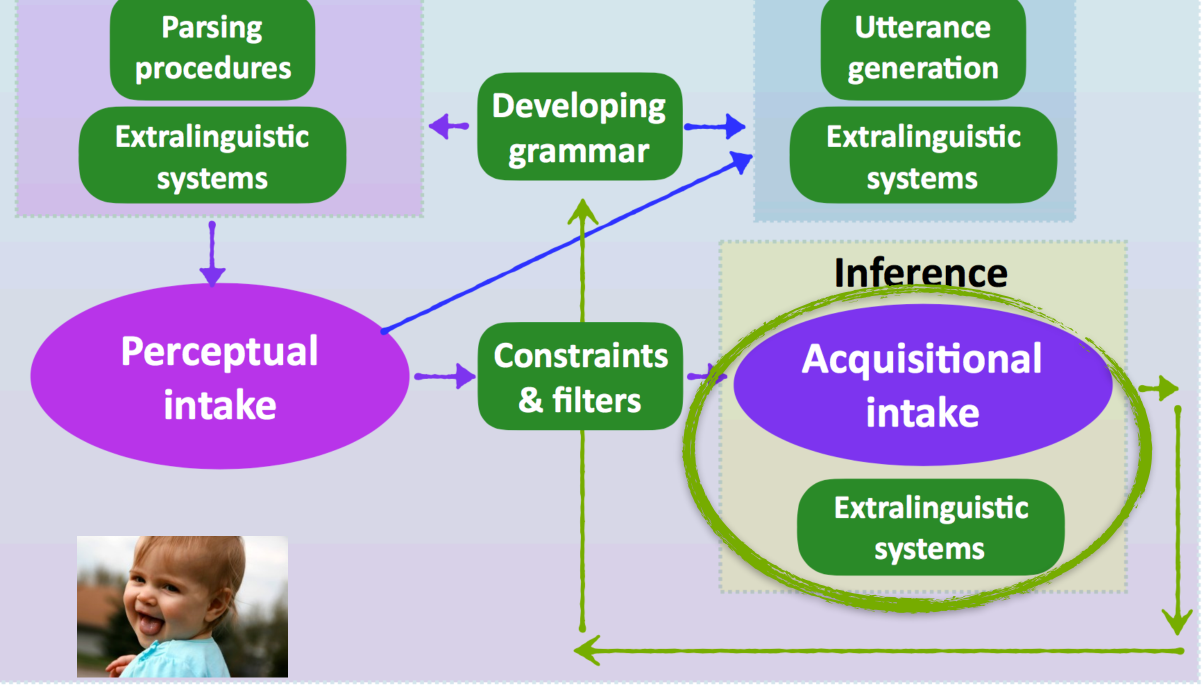
Constraints on children's hypotheses and filters on their attention cause them to heed a subset of the perceptual intake — this is the acquisitional intake.



Adapted from Lidz & Gagliardi 2015

perceptual intake

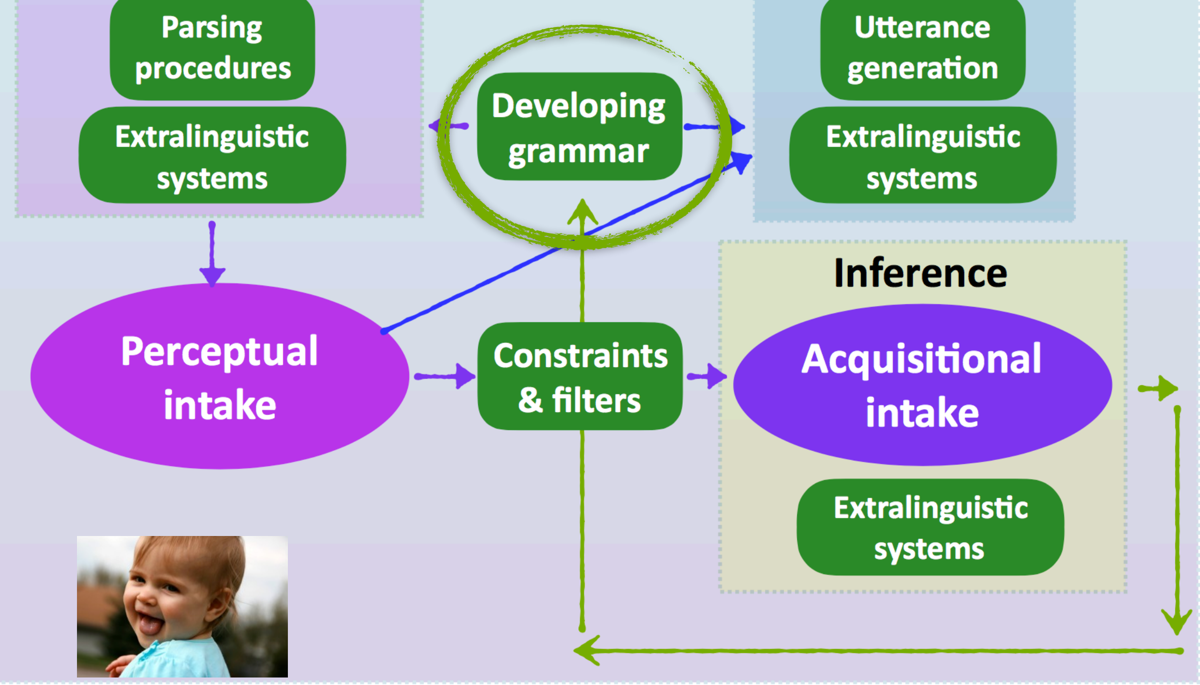




Adapted from Lidz & Gagliardi 2015

Inference = learning

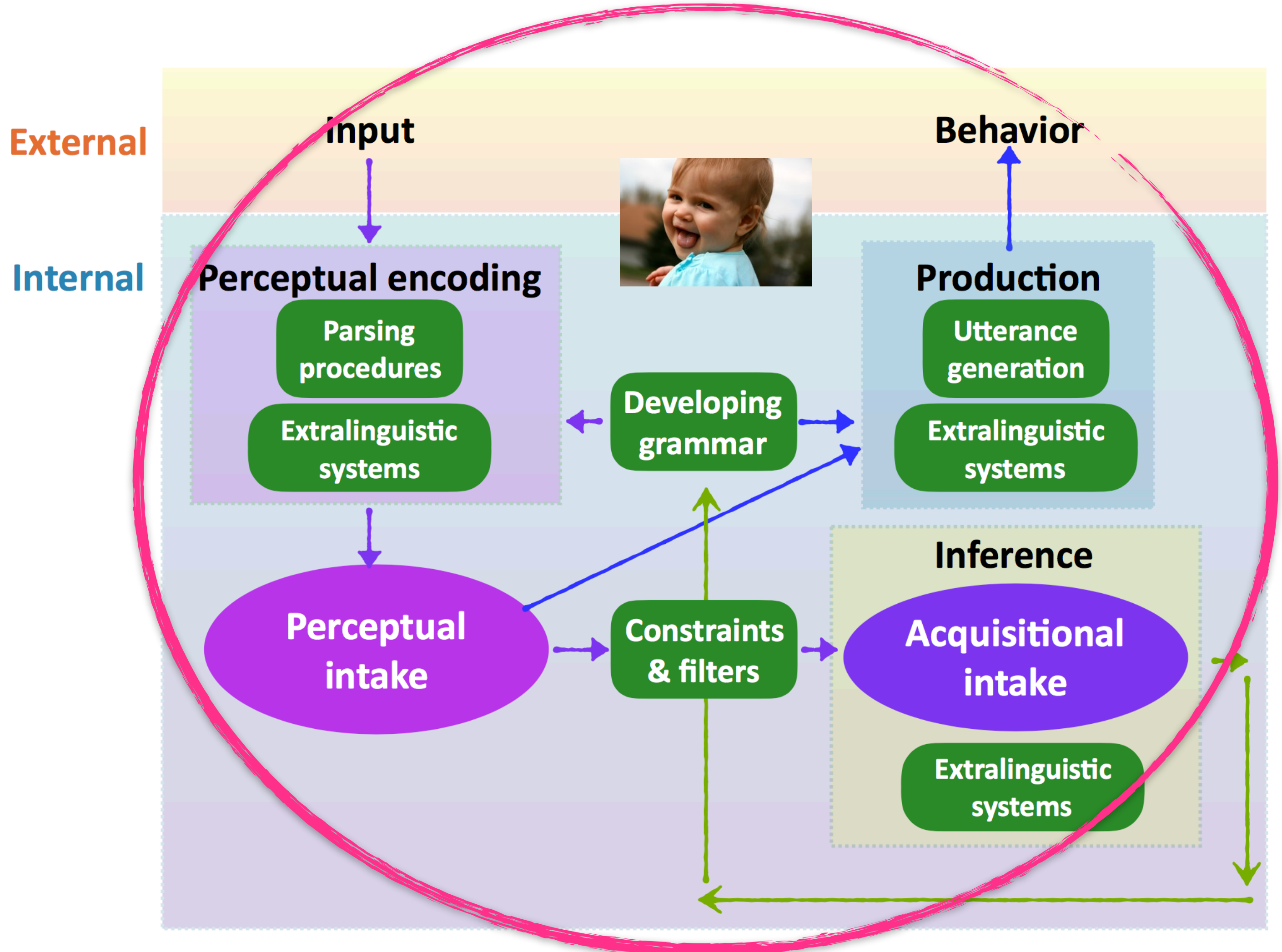
Inference happens over the acquisitional intake, using extralinguistic abilities (statistical learning, probabilistic inference, hypothesis testing, etc.) …



Adapted from Lidz & Gagliardi 2015

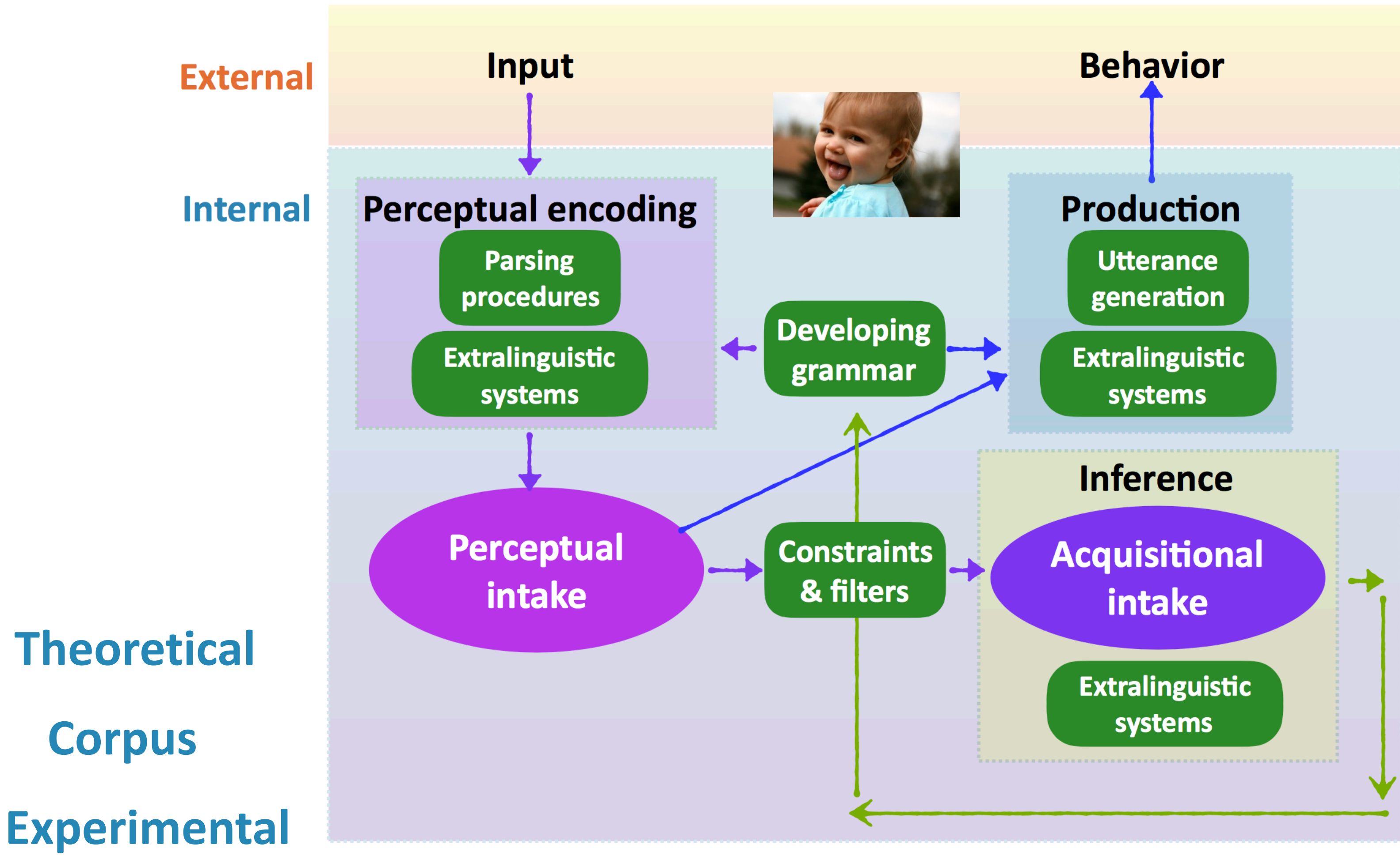
Inference = learning

Inference happens over the acquisitional intake, using extralinguistic abilities (statistical learning, probabilistic inference, hypothesis testing, etc.) to generate the most up-to-date ideas about the language's grammar.



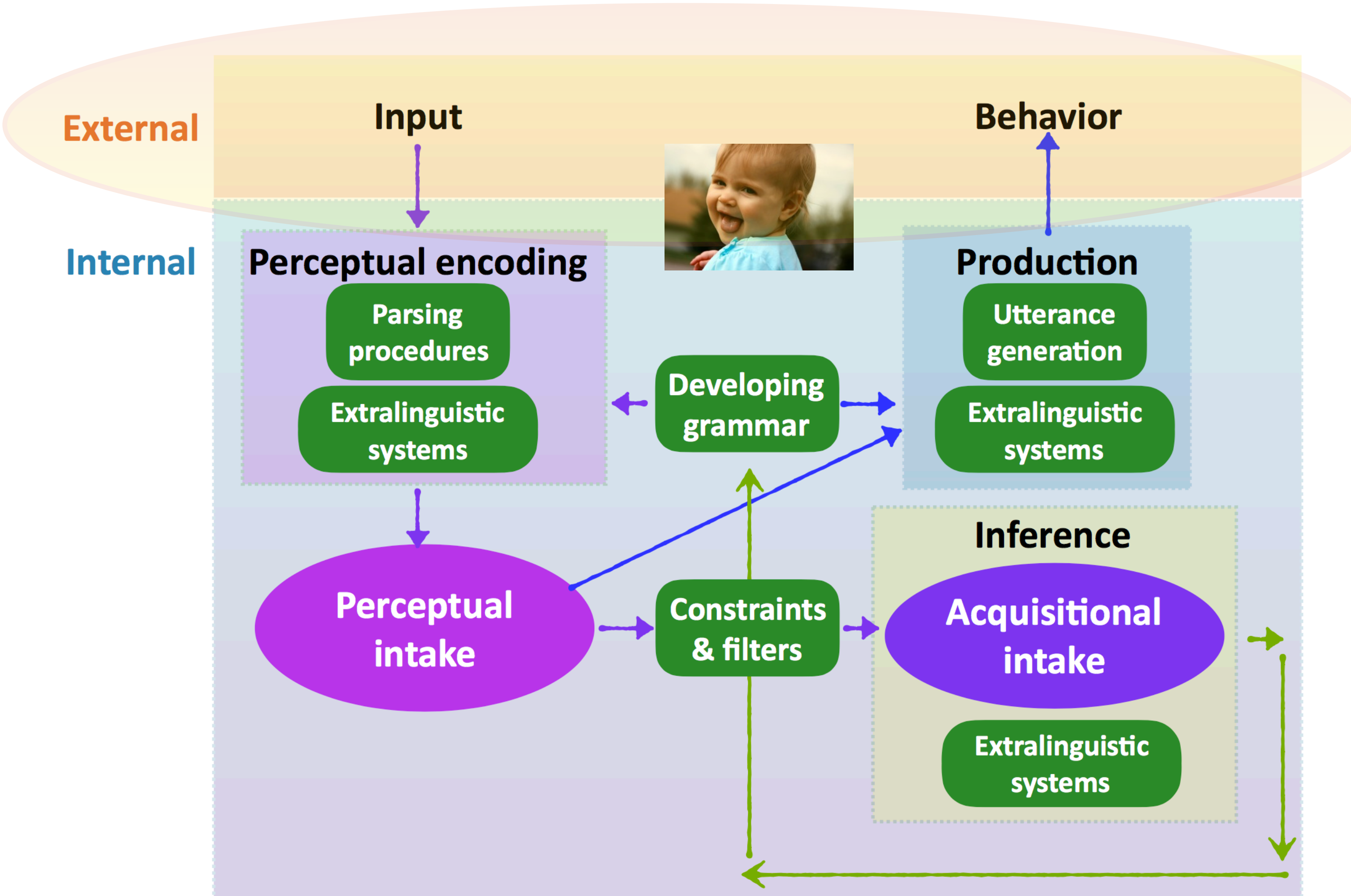
Adapted from Lidz & Gagliardi 2015

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



Adapted from Lidz & Gagliardi 2015

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



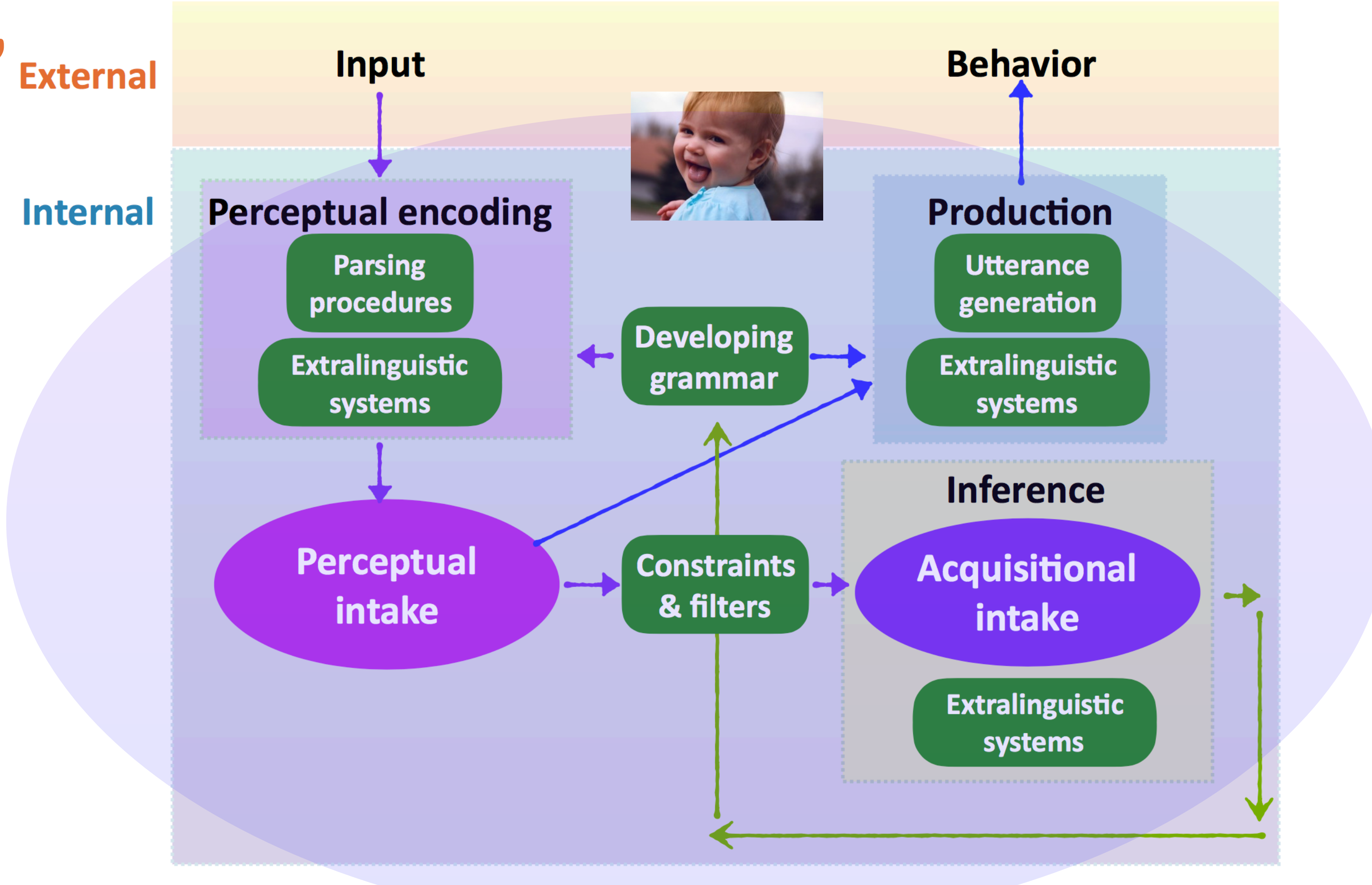
Adapted from Lidz & Gagliardi 2015

External ✔

When we have an informative computational model, it will connect the child's input to the child's output in just this way.

External

Internal



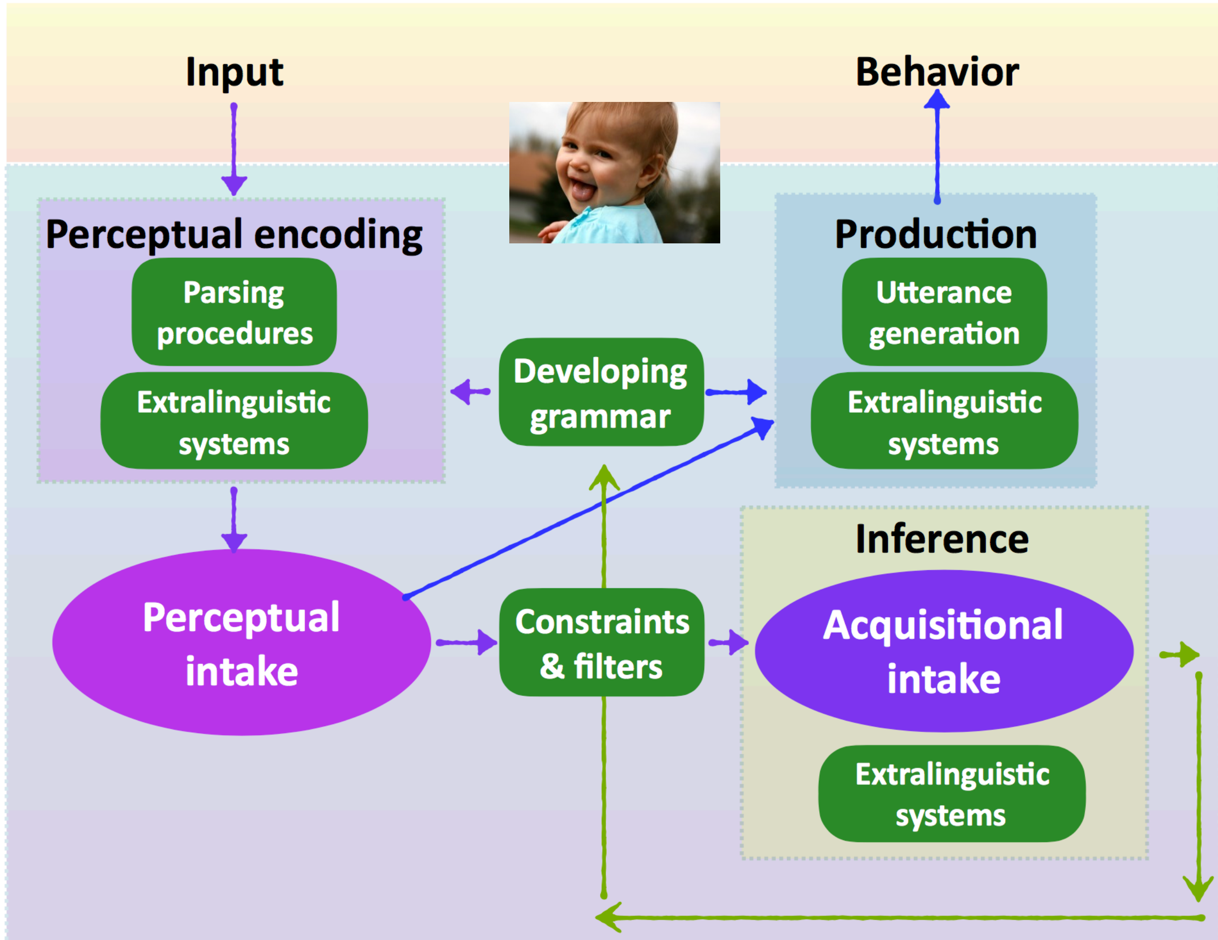
Adapted from Lidz & Gagliardi 2015

Internal

We can then look “under the hood” to see what internal pieces made that possible — this part is hard to do in real children’s minds!

External

Internal

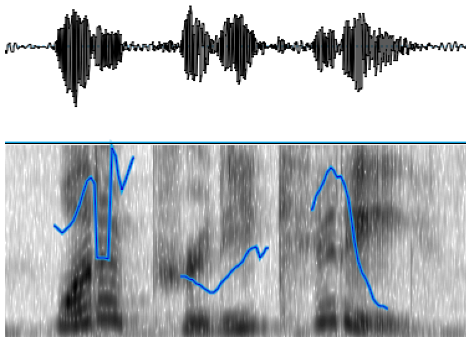


Adapted from Lidz & Gagliardi 2015

Upshot: With computational modeling, we can understand more precisely how the learning strategies that children use work.

Some things we've learned by model-building this way

speech segmentation



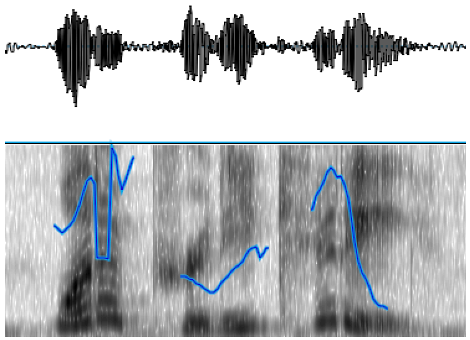
= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

Some things we've learned

speech segmentation



= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

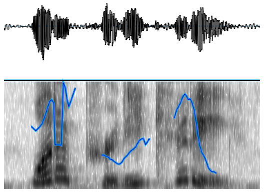
Investigating a Bayesian inference strategy for the very early stages of speech segmentation occurring around six months

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

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

Some things we've learned

speech segmentation



what a pretty kitty!

= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

The intuition of Bayesian inference (applied to speech segmentation)

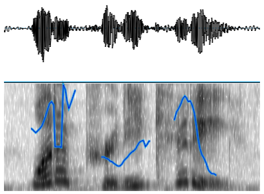
$$P(s|u) \propto P(s)P(u|s)$$

The best answer (based on the utterance you just heard) …

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

Some things we've learned

speech segmentation



what a pretty kitty!

= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

The intuition of Bayesian inference (applied to speech segmentation)

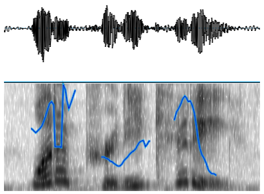
$$P(s|u) \propto P(s)P(u|s)$$

The best an**s**wer (based on the **u**tterance you just heard) depends on your prior beliefs about what good an**s**wers look like …

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

Some things we've learned

speech segmentation



what a pretty kitty!

= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

The intuition of Bayesian inference (applied to speech segmentation)

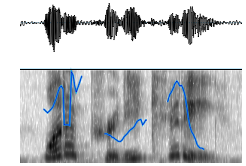
$$P(s|u) \propto P(s)P(u|s)$$

The best answer (based on the utterance you just heard) depends on your prior beliefs about what good answers look like and how easily an answer explains the data observed in the utterance.

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

Some things we've learned

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 list of word forms (= lexicon) that best generates the observable fluent speech utterances

Mathematically encoded preferences:

wʌɾə
pɹɪɾi
kɪɾi

wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

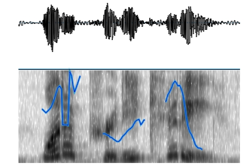
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

Some things we've learned

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 list of word forms (= lexicon) that best generates the observable fluent speech utterances

Mathematically encoded preferences:

(1) Prefer shorter words

wʌɾə
pɹɪɾi
kɪɾi

wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

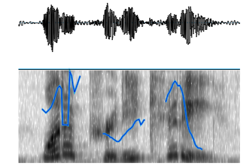
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips in press

Some things we've learned

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 list of word forms (= lexicon) that best generates the observable fluent speech utterances

Mathematically encoded preferences:

(1) Prefer shorter words

(2) Prefer lexicons with fewer words

wʌɾə
pɹɪɾi
kɪɾi

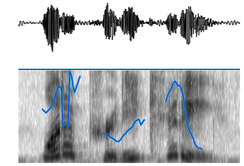
wʌ
ɾə
pɹɪɾikɪɾi

wʌɾə
pɹɪɾikɪɾi

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

Some things we've learned

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 list of word forms (= lexicon) 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

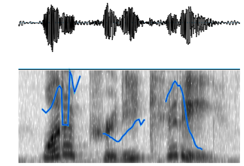
wʌɾə
pɹɪɾikɪɾi

Find the best segmentation

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

Some things we've learned

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 list of word forms (= lexicon) that best generates the observable fluent speech utterances

Mathematically encoded preferences:

(1) Prefer shorter words

(2) Prefer lexicons with fewer words

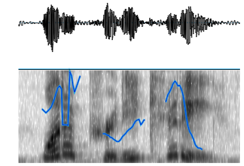
wʌɾə pɹɪɾi kɪɾi	wʌ ɾə pɹɪɾikɪɾi	wʌɾə pɹɪɾikɪɾi

Find the best segmentation that balances these preferences

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

Some things we've learned

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 list of word forms (= lexicon) 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

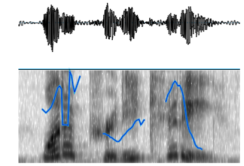
Find the best segmentation that balances these preferences

and can generate the observable fluent speech utterances

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

Some things we've learned

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 for children?**

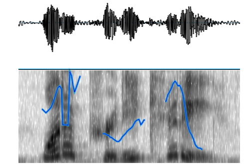
Modeled learners without cognitive limitations on their inference and memory can use this strategy to segment fairly well when given realistic English child-directed speech data to learn from.

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

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

Some things we've learned

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?

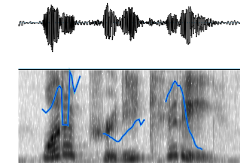
Is it useable by children?

Modeled learners with cognitive limitations 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

Some things we've learned

speech segmentation



Bayesian inference

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

= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

Is it useful?

Is it useable?

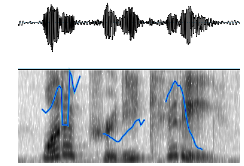
Does it work for different languages?

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

Some things we've learned

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

Recap

Language acquisition is an interesting area of research in human cognition because **it's really hard** and **little humans are really good at it**.

metrical phonology

speech segmentation

syntactic categorization

pragmatics

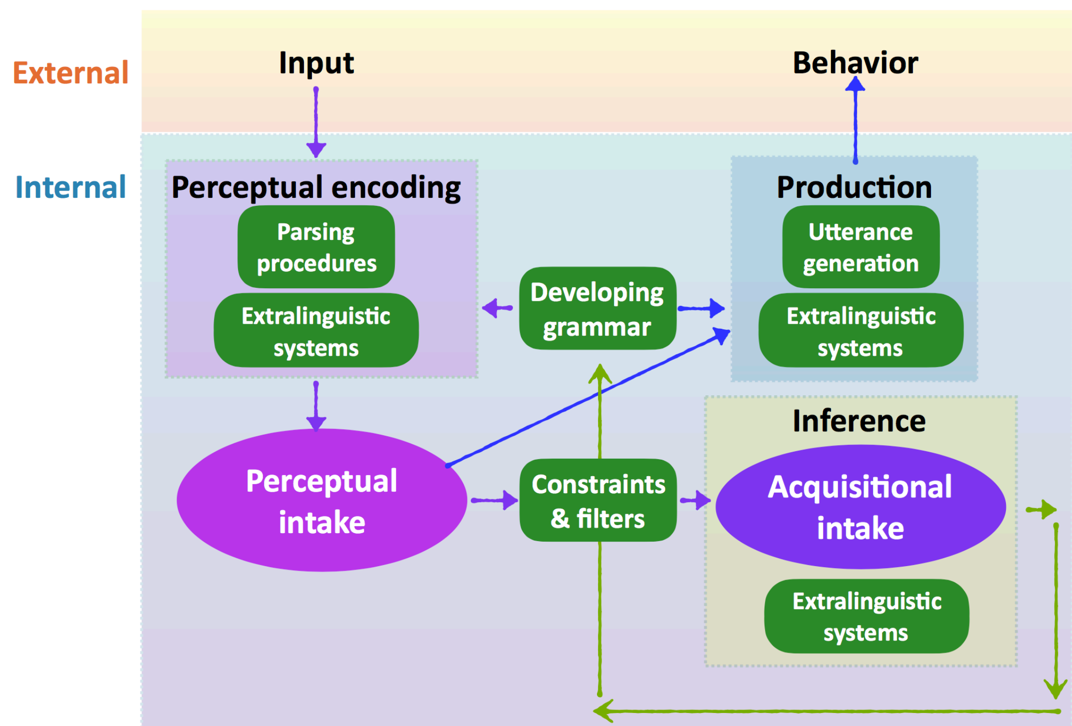
syntax

syntax, semantics

Recap

Language acquisition is interesting

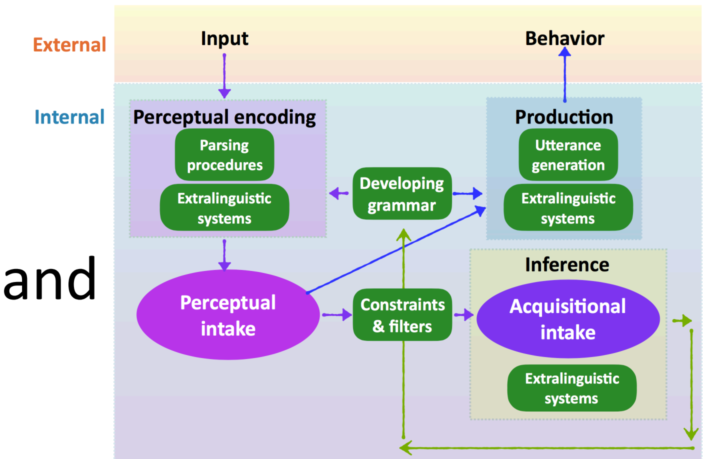
To understand how it works, we can build cognitive computational models that capture the **important components** of the process and then **look inside** to see exactly how they work



Recap

Language acquisition
is interesting

Models can capture **important components** and we can **look inside**



Some recent findings with this approach suggest **Bayesian inference** is a plausible early speech segmentation strategy that's useful, useable, and works for many languages

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

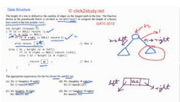
Lisa S. Pearl
Associate Professor
Department of Linguistics
Department of Cognitive Sciences
SSPB 2219, SBSG 2314
University of California, Irvine
lpearl@uci.edu

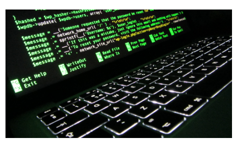
Computation of Language Laboratory

UC Irvine

Thank you!

UC Computational Social Science

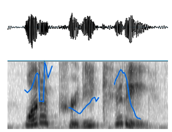




Noun

Kl tty

Who does... is pretty?



another one

Every kitty didn't ...