Ling 151/Psych 156A: Acquisition of Language II

Lecture 3

Intro III

Announcements

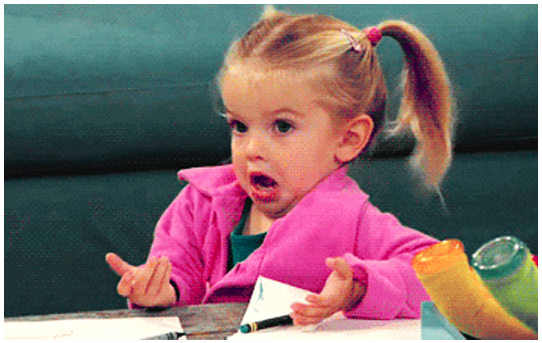
Be working on HW1 (due: 1/17/18)

Be looking over the review questions for introduction

Computational models of language acquisition

Why model language acquisition?

What does it mean to model something?



What does it mean to model something?

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

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

So what questions should we be using models for?

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

The importance of theory

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

The importance of theory

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

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

(2) evaluate whether it actually works, and

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

- Pearl 2017

Making the components explicit

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

- Pearl 2017

Example: Learning linguistic structure from signals ("triggers") in the input

Specific example:

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

Making the components explicit

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

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

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

Making the components explicit

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

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

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

The importance of theory

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

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

(2) evaluate whether it actually works, and

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

- Pearl 2017

Evaluating whether the theory works and determining what makes it work

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

- Pearl 2017

Two basic outcomes:

- the model predictions match children's data

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

Evaluating whether the theory works and determining what makes it work

The model predictions match children's data

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

Note: Doesn't rule out alternative acquisition theories

Two basic outcomes:

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

Evaluating whether the theory works and determining what makes it work

The model predictions match children's data

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

The model predictions don't match children's data

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

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

Evaluating whether the theory works and determining what makes it work

The model predictions match children's data

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

The model predictions don't match children's data

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

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

Evaluating whether the theory works and determining what makes it work

What did the penguin do ___*what*?

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

Evaluating whether the theory works and determining what makes it work

What did the penguin do ___*what*?

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

~~*The penguin did what?!*~~ ✓

default = don't move

Evaluating whether the theory works and determining what makes it work

What did the penguin do ___what?

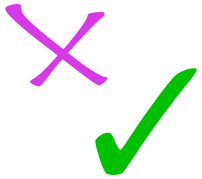
The penguin did what?! ✗

✓ *default = don't move*

This is **useful**!

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

Modeling as a useful tool



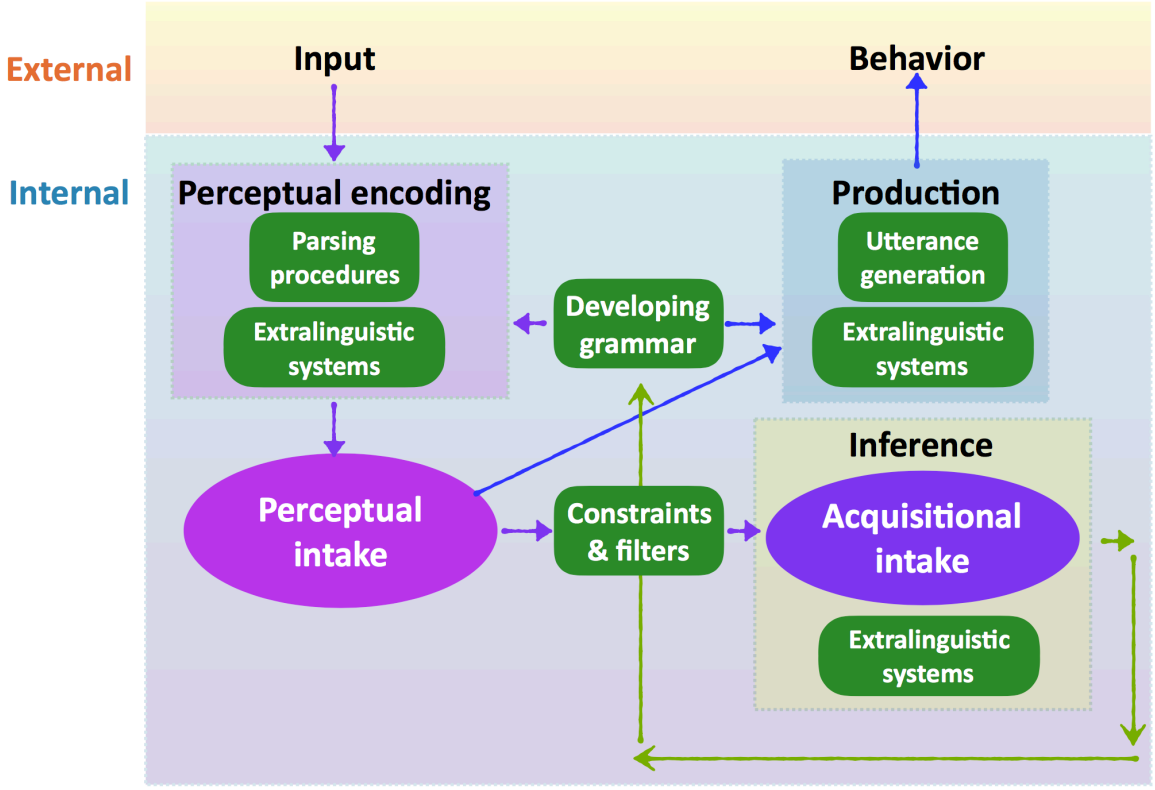
Modeling can be used as a tool for both developing and refining acquisition theories.

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

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

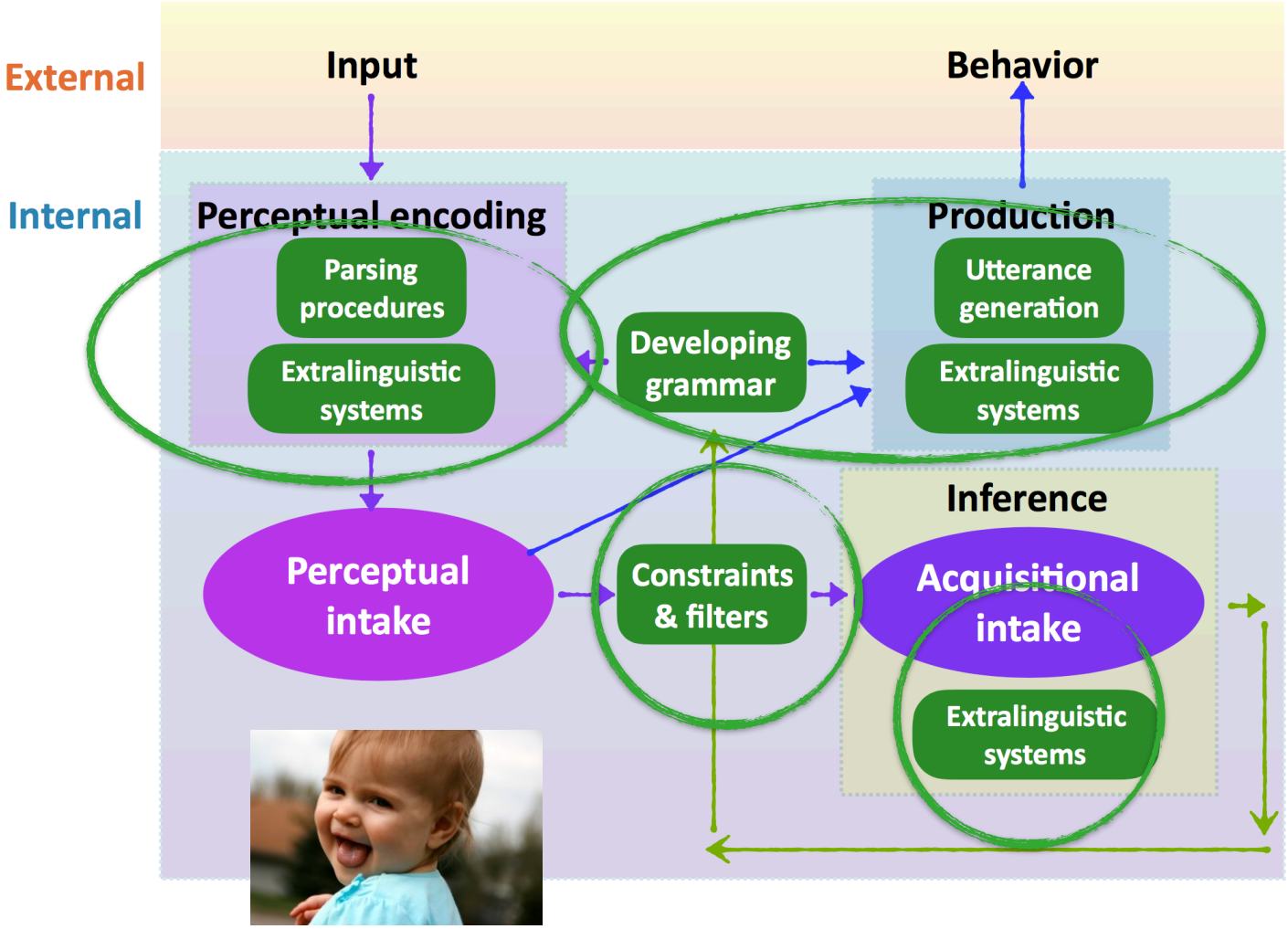
An informative model incorporates both.

When building a specific model, it can be helpful to think about the acquisition process in five main parts



Initial state

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

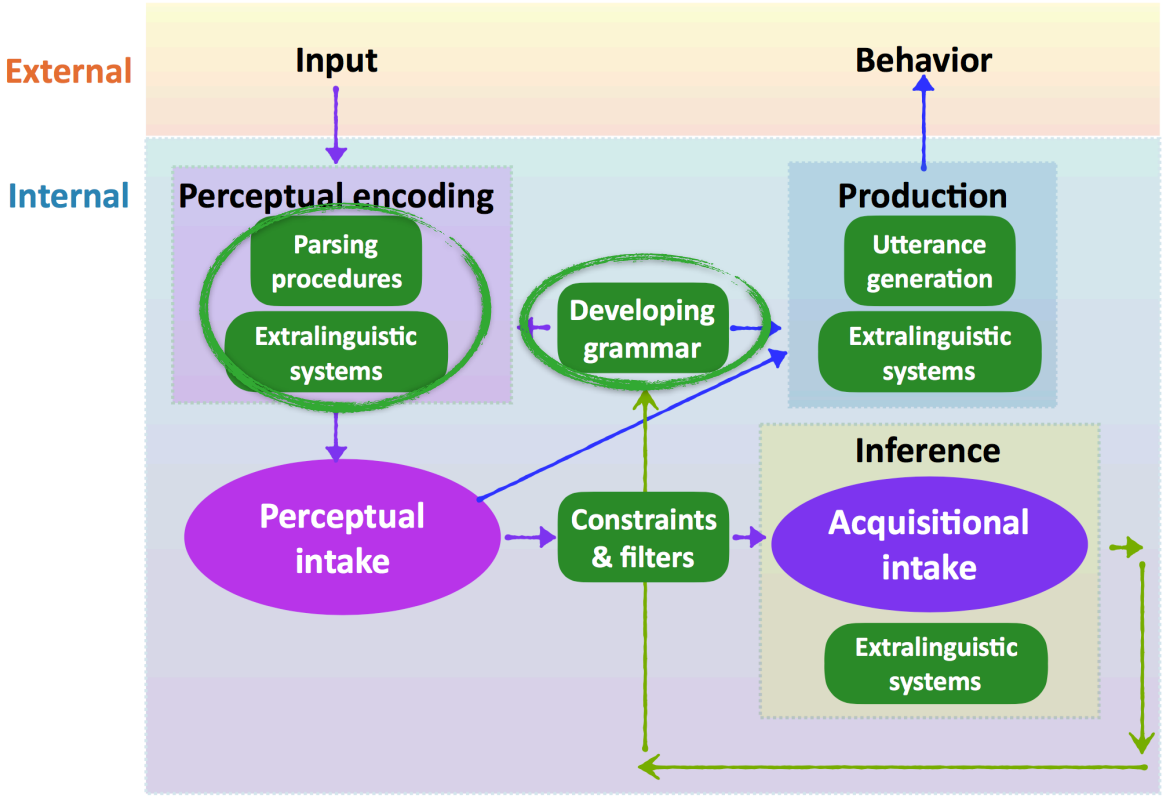


Initial state

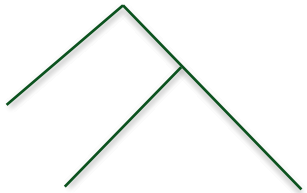
What does the child start with?
What **knowledge**, abilities, and learning biases does the child already have?

Example knowledge:

- syntactic categories exist and can be identified
- phrase structure exists and can be identified
- participant roles can be identified



N, V, Adj, P, …



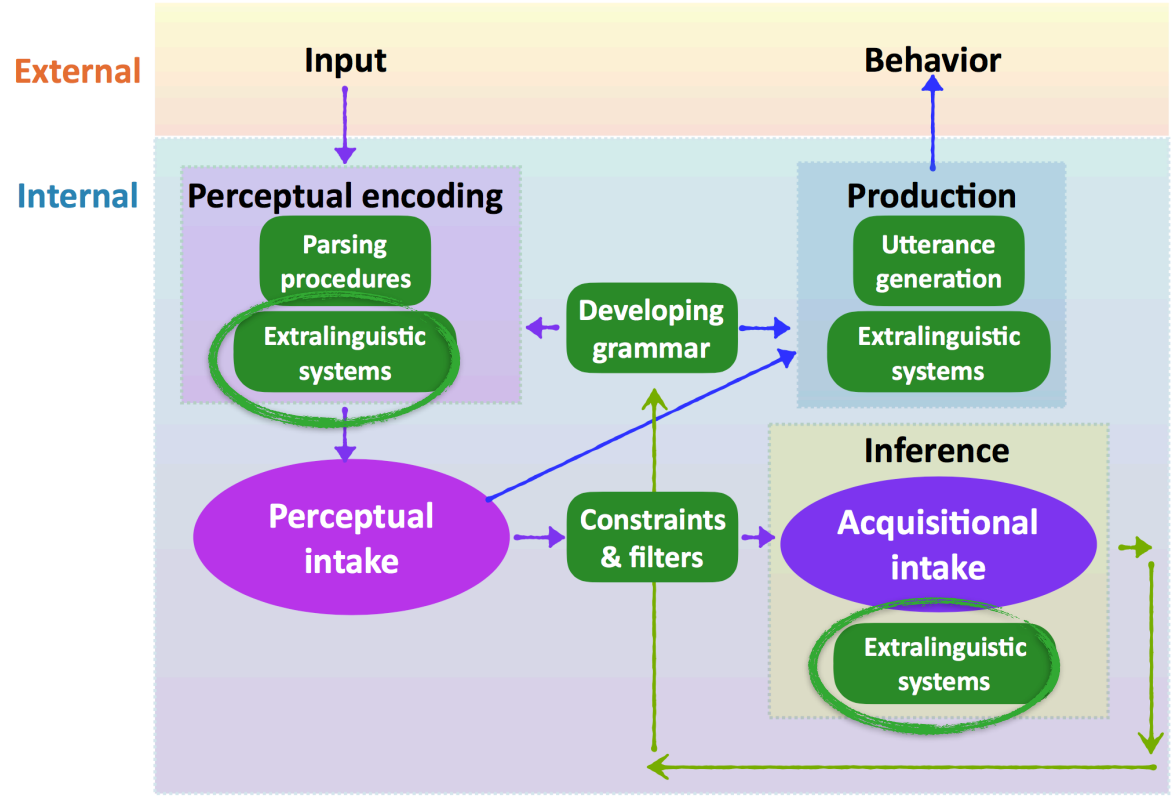
Agent, Patient, Goal, …

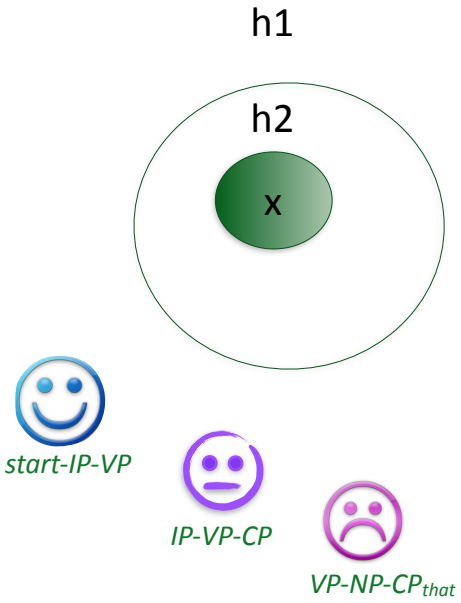
Initial state

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

Example abilities & biases:

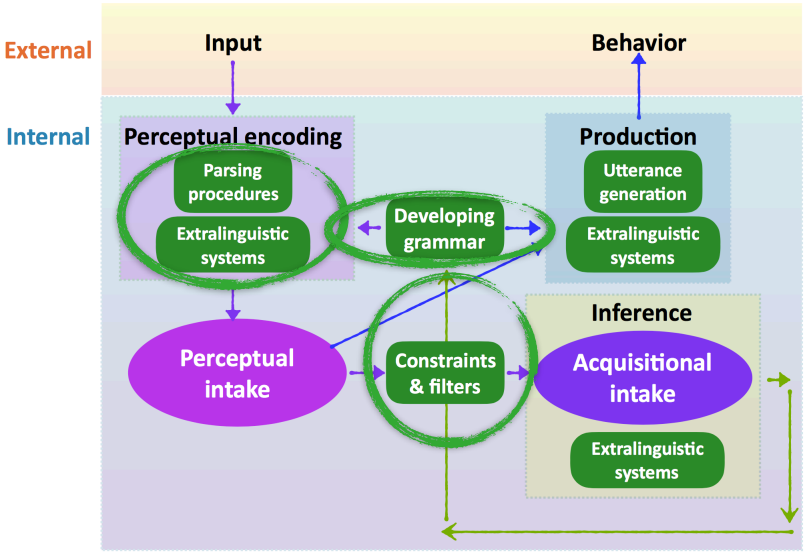
- frequency information can be tracked
- distributional information can be leveraged

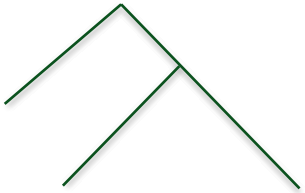




Initial state

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

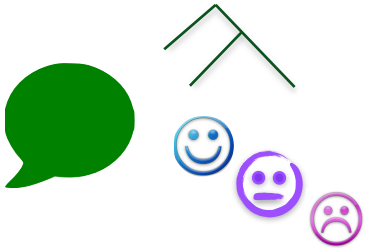




Example initial state: A strategy that depends on the frequency of certain syntactic structures would need the child to know about that syntactic structure already, recognize it in the input via the developing language processing abilities, and be able to track the frequency of that structure.

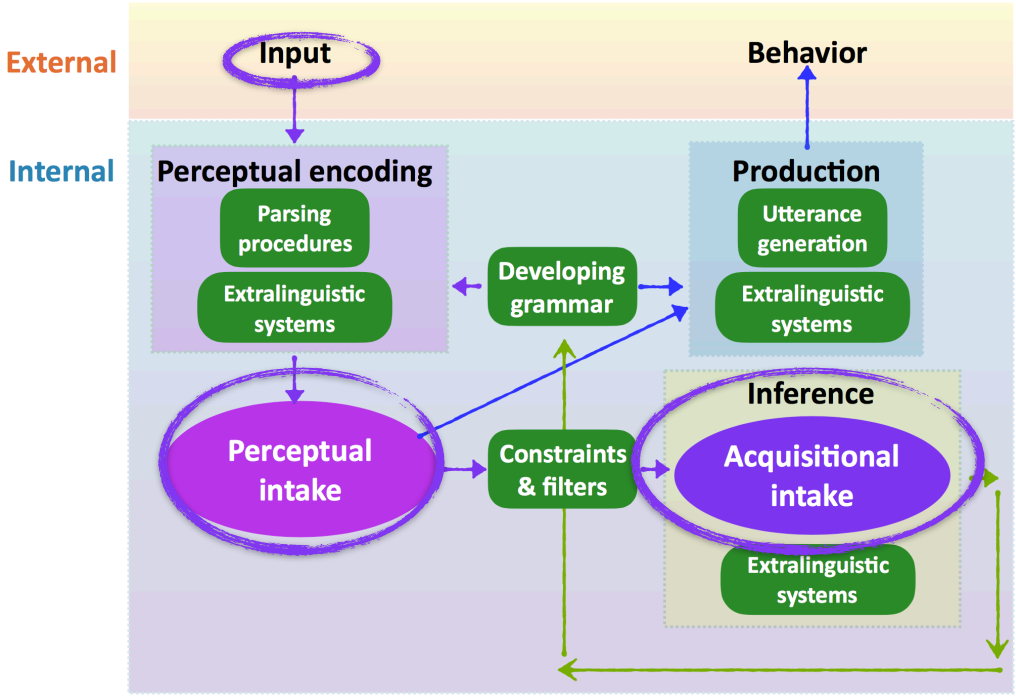
Initial state

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



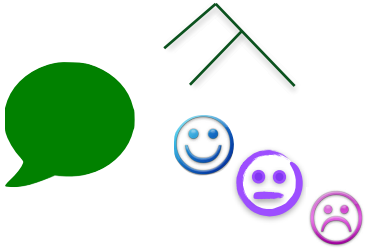
Data intake

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



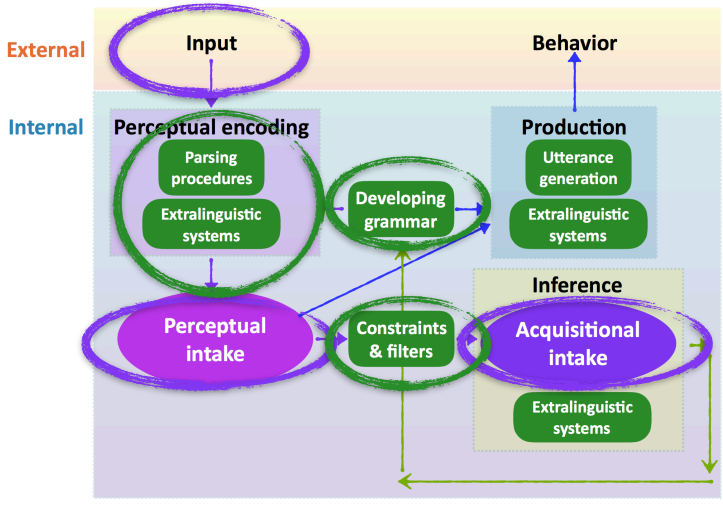
Initial state

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



Data intake

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



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

ex: all pronoun data when learning about *one* used as a pronoun

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

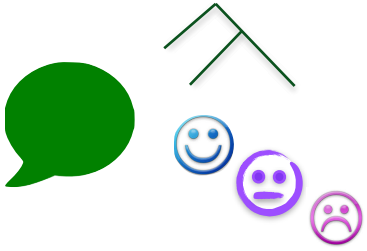
[defined by knowledge & biases/capabilities in the initial state]

Data intake

What is the acquisitional intake?

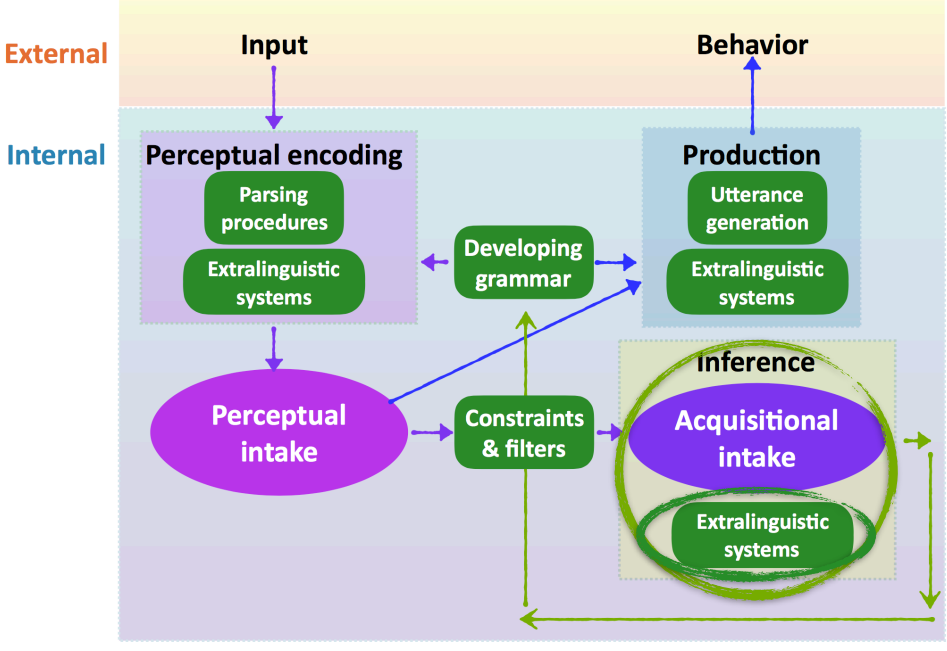
Initial state

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



Inference

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



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

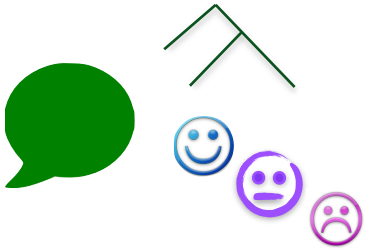
[defined by knowledge & biases/ capabilities in the initial state]

Data intake

What is the acquisitional intake?

Initial state

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

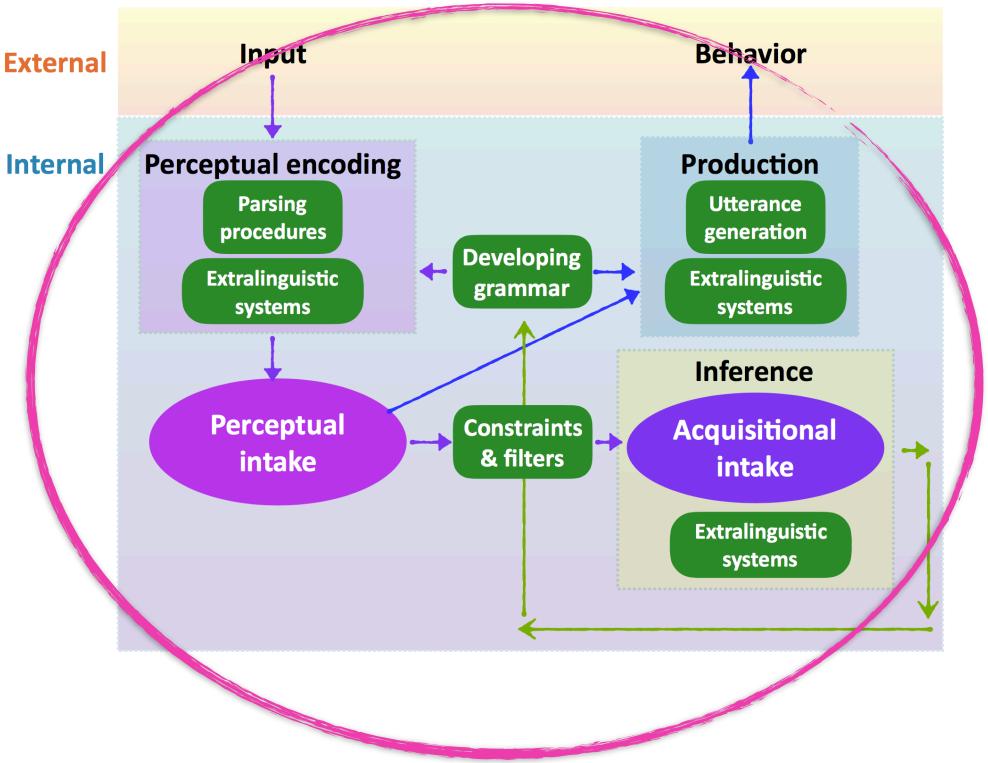


Inference

How are updates made?

Learning period

How long does the child have to learn?

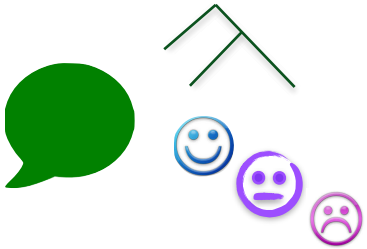


Data intake

What is the acquisitional intake?

Initial state

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



Inference

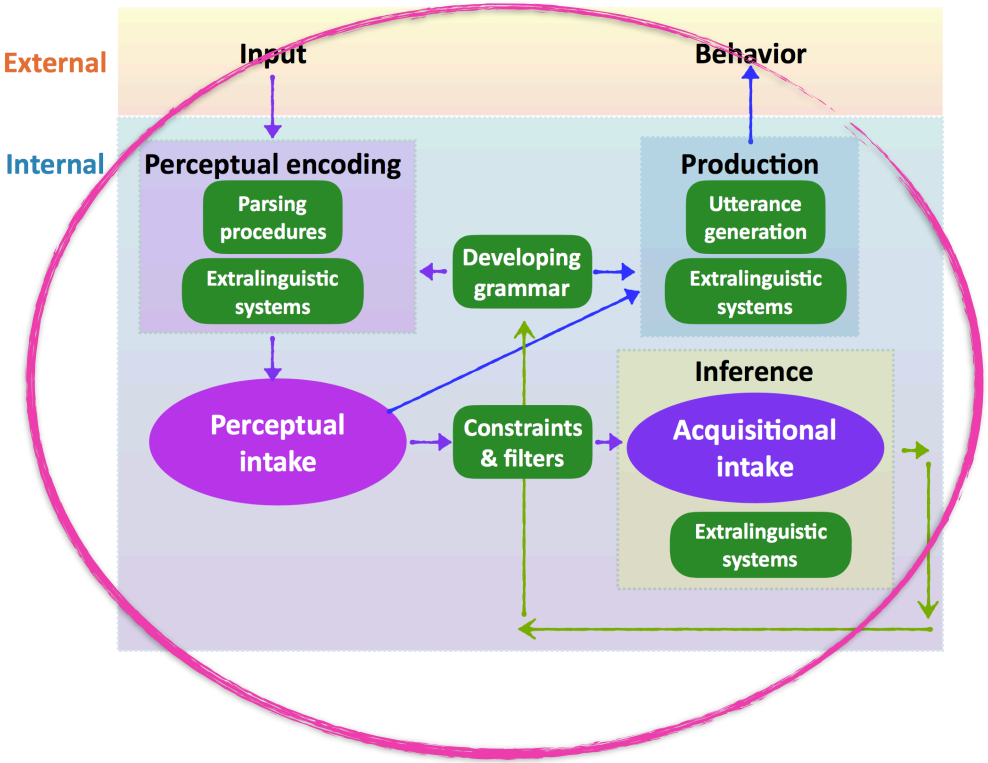
How are updates made?

Learning period

How long does the child have to learn?

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

ex: 4 months, ~36,500 data points

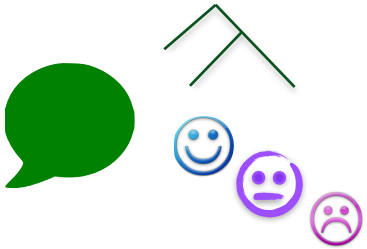


Data intake

What is the acquisitional intake?

Initial state

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



Inference

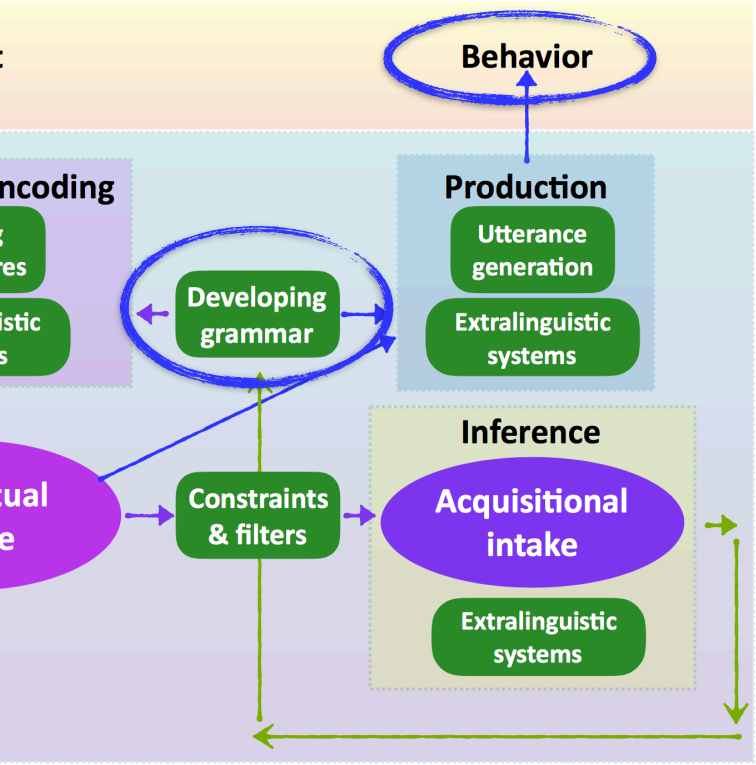
How are updates made?

Learning period

How long does the child have to learn?

Target state

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

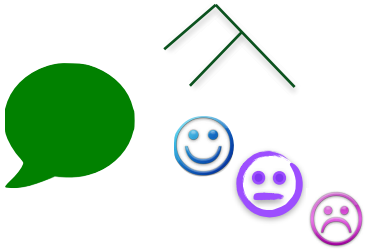


Data intake

What is the acquisitional intake?

Initial state

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

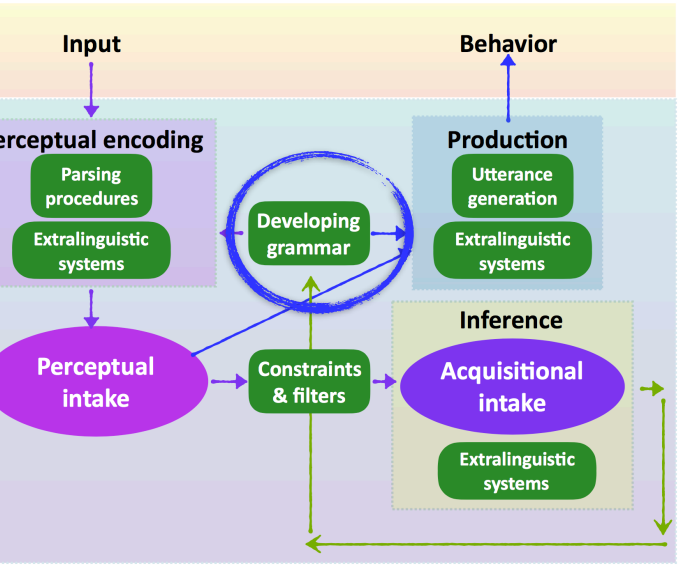


Inference

How are updates made?

Learning period

How long does the child have to learn?



Target state

What does successful acquisition look like?

knowledge

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

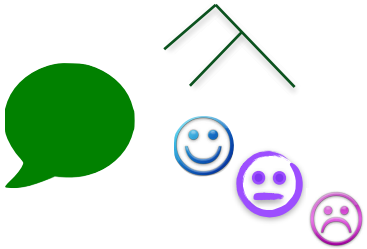
ex: Where did Jack buy a necklace from __ for Lily for her birthday?

Data intake

What is the acquisitional intake?

Initial state

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



Inference

How are updates made?

Learning period

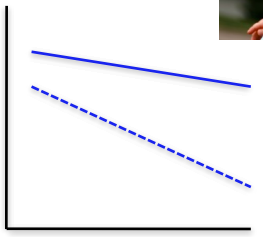
How long does the child have to learn?

Target state

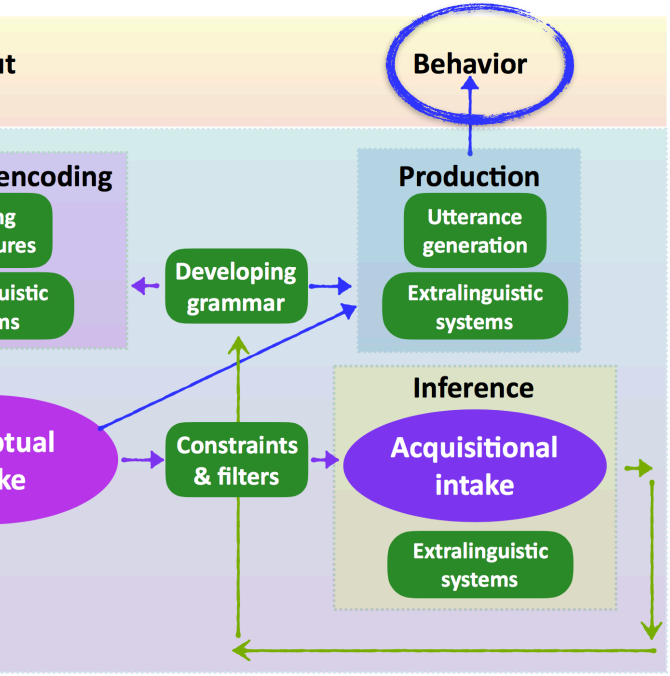
What does successful acquisition look like?

behavior

looking time preferences

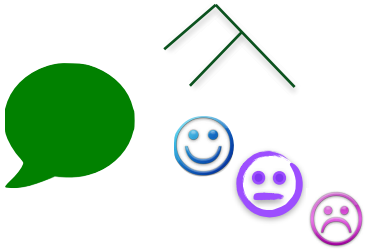


z-score rating



Initial state

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



Data intake

What is the acquisitional intake?

Inference

How are updates made?

Target state

What does successful acquisition look like?

knowledge

behavior

Learning period

How long does the child have to learn?

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

Building an informative model about…

Which learning strategies could children be using?

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

Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

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

Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

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

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

Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

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

When do children learn different aspects of the linguistic system?

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

Building an informative model about…

Which learning strategies could children be using?

Which learning biases are necessary?

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

When do children learn different aspects of the linguistic system?

What factors affect children's observable behavior?

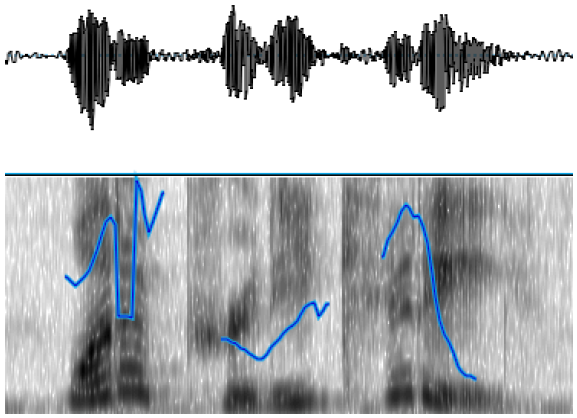
(Nguyen & Pearl in prep., Savinelli, Scontras, & Pearl 2018, Nguyen & Pearl 2017, Savinelli, Scontras, & Pearl 2017)

Model-building

Let's look at an example with speech segmentation

Model-building

An example with speech segmentation



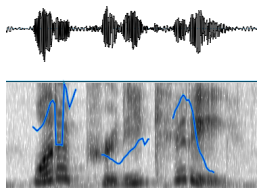
= wʌɾəpɹɪɾikɪɾi

wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

Model-building

An example with speech segmentation



what a pretty kitty!

(1) Decide what kind of learner the model represents

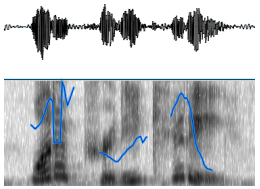
This depends on what task you're modeling

For the first stages of speech segmentation: **Experimental**

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

Model-building

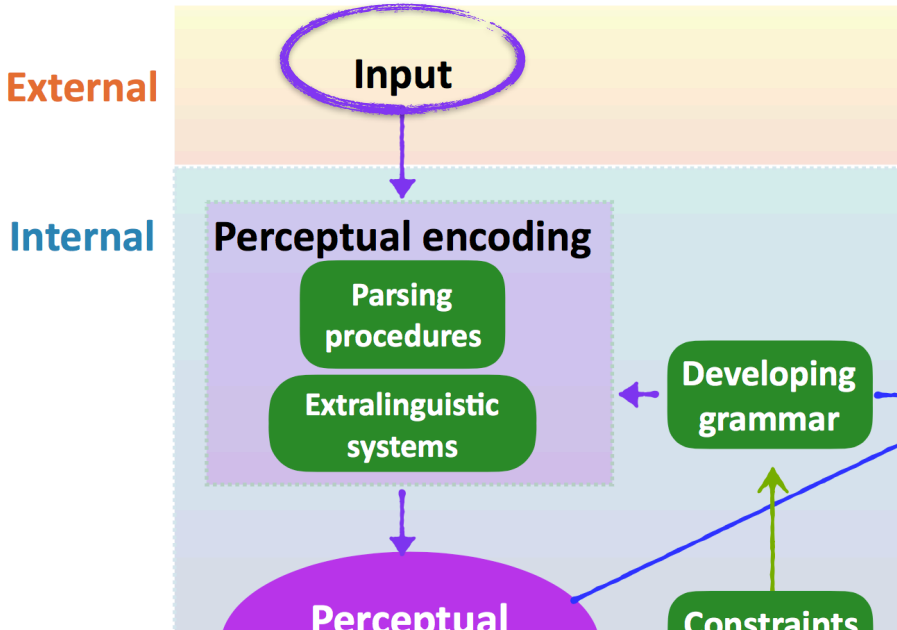
An example with speech segmentation



what a pretty kitty!

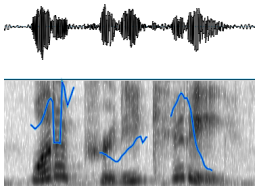
(2) Decide what data the child learns from (input)

This depends on your acquisition theory and the empirical data available



Model-building

An example with speech segmentation



what a pretty kitty!

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

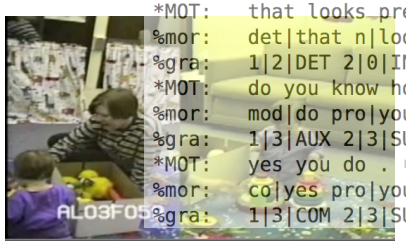
Corpus

Example empirical data: CHILDES database

http://childes.talkbank.org

CHILDES Child Language Data Exchange System

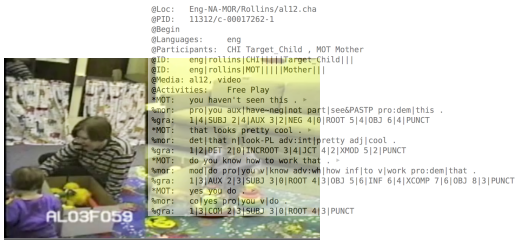
Video/audio recordings of speech samples, along with transcriptions and some structural annotations.

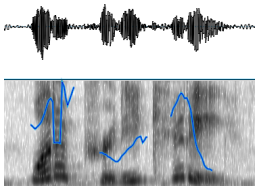


```
@Loc:    Eng-NA-MOR/Rollins/al12.cha
@PID:    11312/c-00017262-1
@Begin
@Languages:     eng
@Participants:  CHI Target_Child , MOT Mother
@ID:     eng|rollins|CHI|||||Target_Child|||
@ID:     eng|rollins|MOT|||||Mother|||
@Media:  al12, video
@Activities:     Free Play
*MOT:    you haven't seen this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

Model-building

An example with speech segmentation



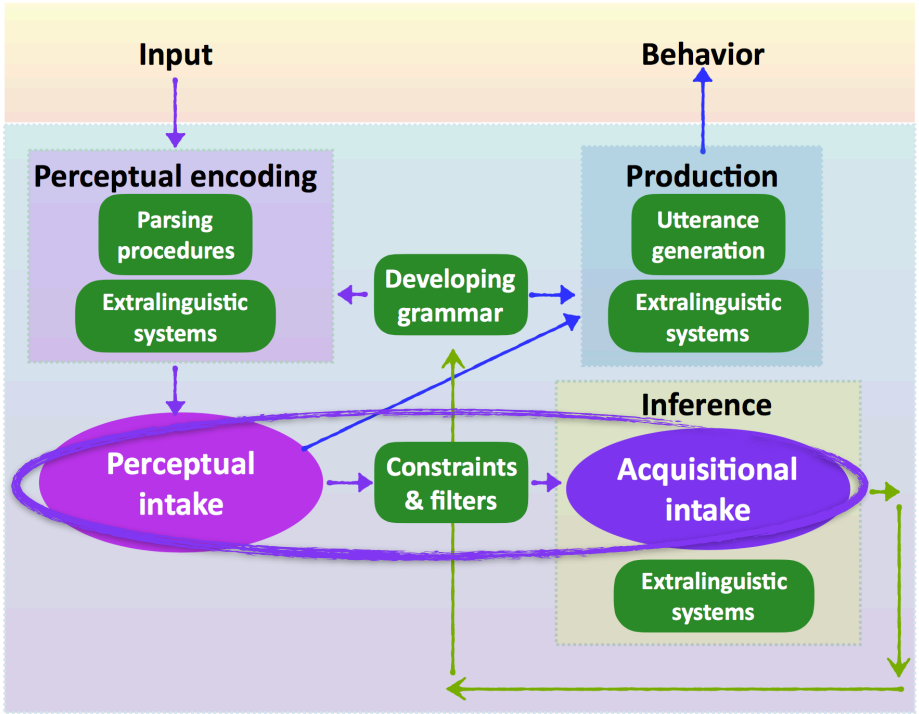


what a pretty kitty!

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

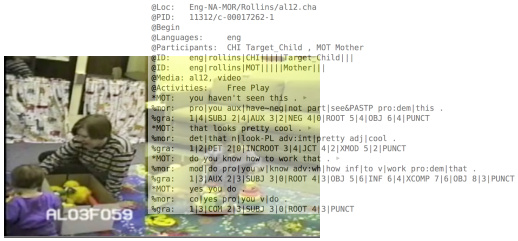
This depends on what children are known to be capable of perceiving at that age and your acquisition theory

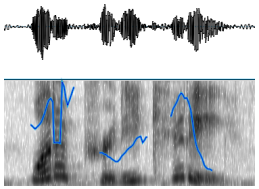
Experimental **Theoretical**



Model-building

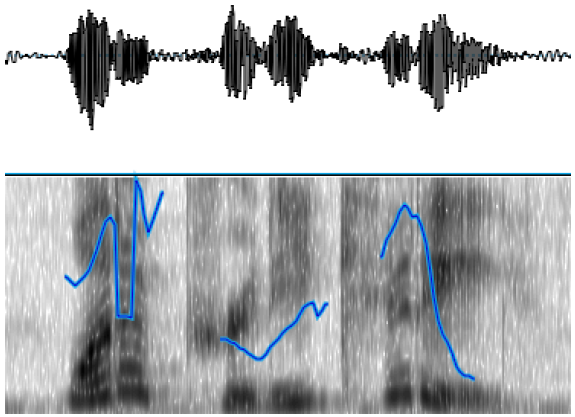
An example with speech segmentation





what a pretty kitty!

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

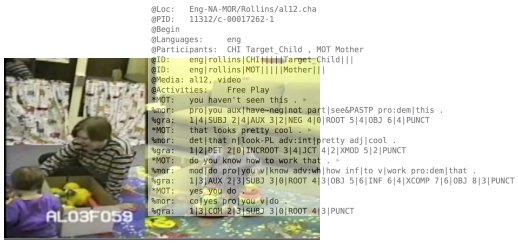


syllables with stress

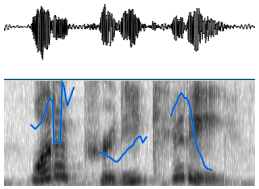
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

Model-building

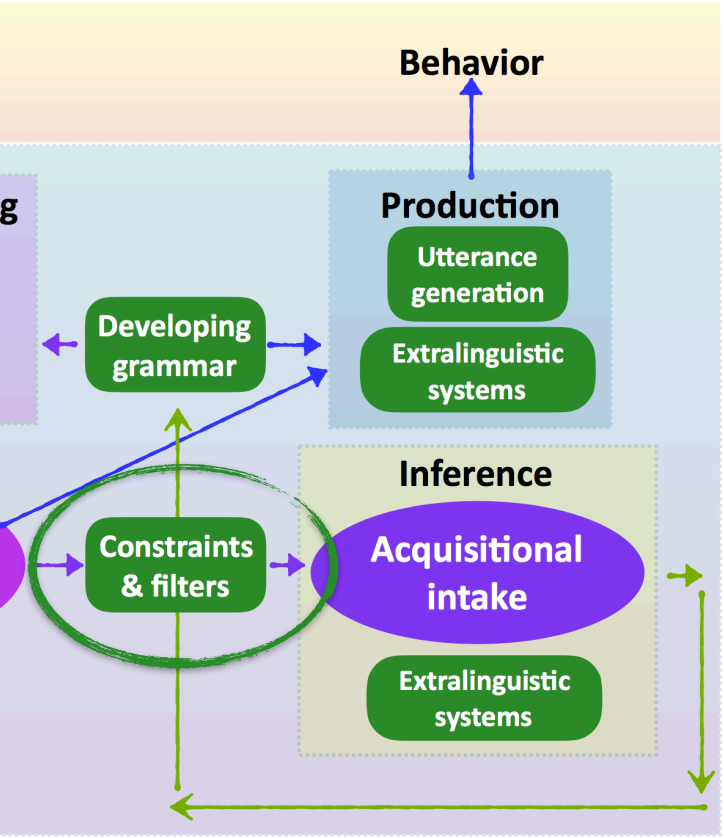
An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!



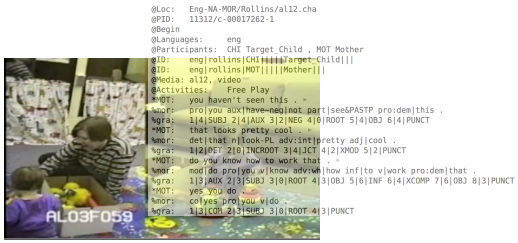
(4) Decide what hypotheses the child has and what information is being tracked in the input

This depends on your acquisition theory

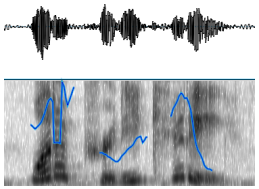
Theoretical

Model-building

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

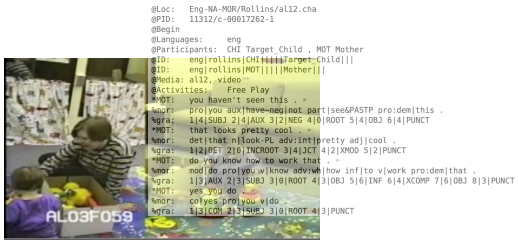
(4) Decide what hypotheses the child has and what information is being tracked in the input

Example hypotheses: what the words are

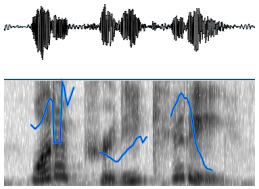
wˈʌɾə	wˈʌ	wˈʌɾə
pɹˈɪɾi	ɾə	pɹˈɪɾikˈɪɾi
kˈɪɾi	pɹˈɪɾikˈɪɾi	

Model-building

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



what a pretty kitty!

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi

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

Example information:

transitional probability between syllables,

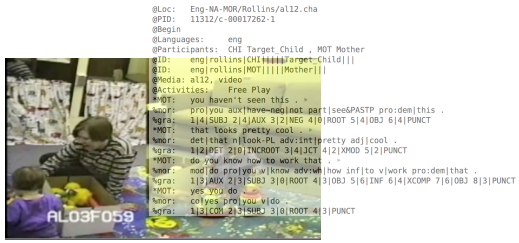
stress on syllables

Experimental

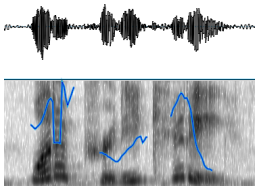
wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

Model-building

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



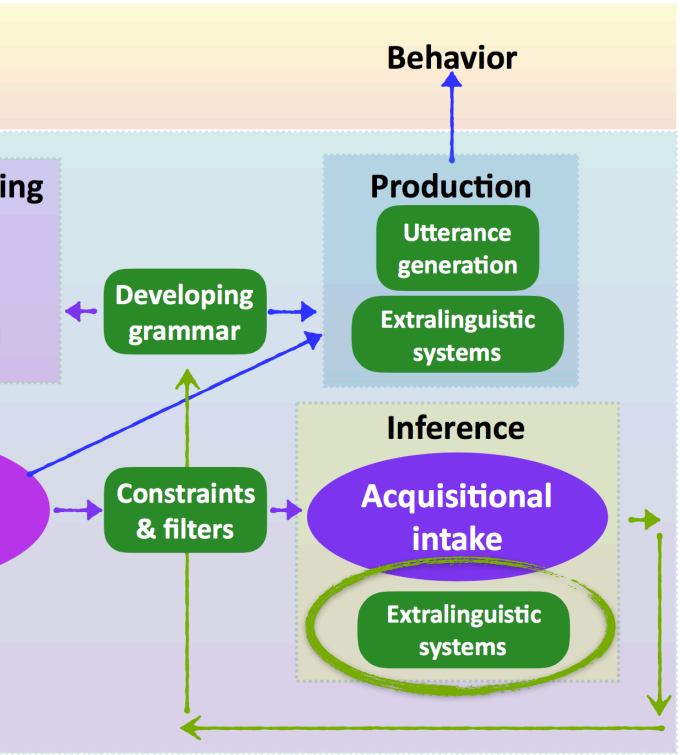
what a pretty kitty!

wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi



(5) Decide how belief in different hypotheses is updated

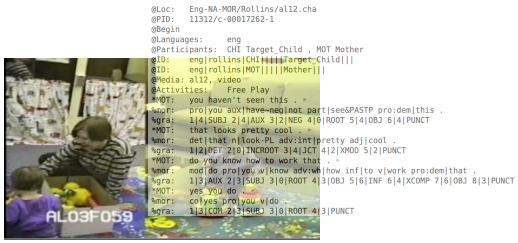
This depends on your acquisition theory and what we know about children's abilities at that age

Theoretical **Experimental**

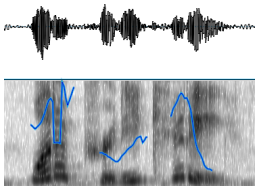
Example: based on transitional probability between syllables

Model-building

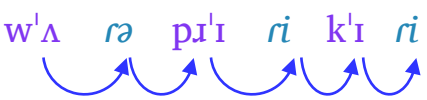
An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



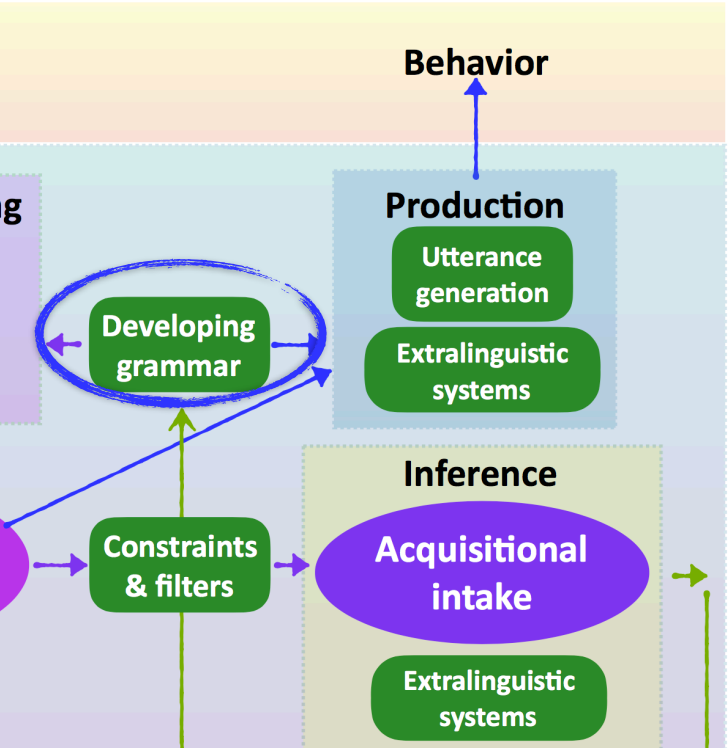
what a pretty kitty!



wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi



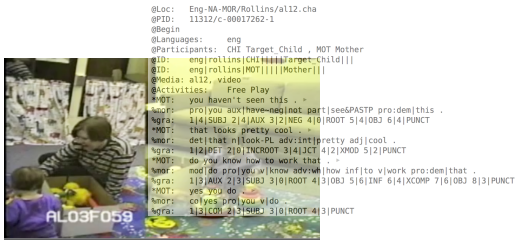
(6) Decide what the measure of success is

This can be based on your theory of what the developing grammar is like…

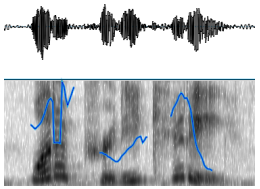
Theoretical

Model-building

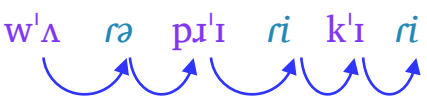
An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



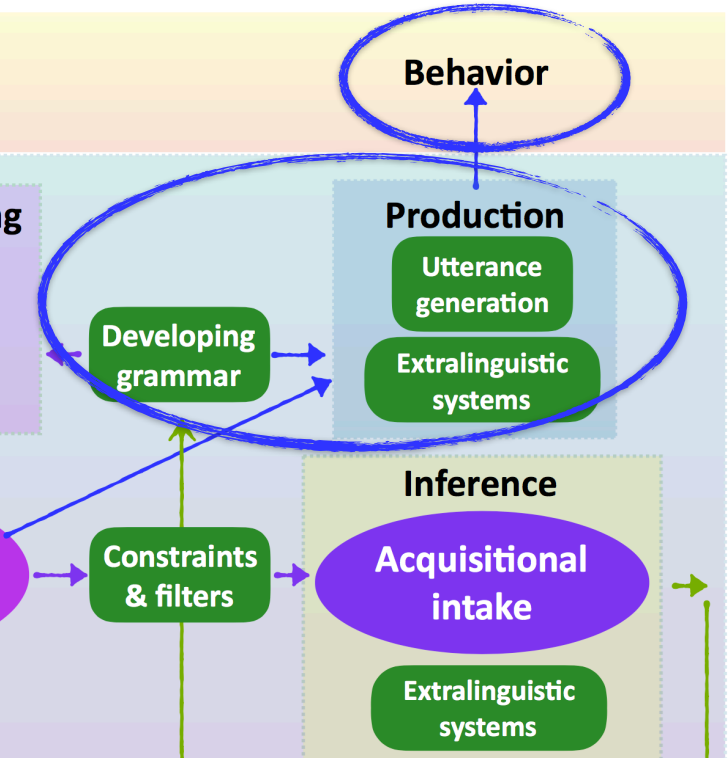
what a pretty kitty!



wˈʌɾə
pɹˈɪɾi
kˈɪɾi

wˈʌ
ɾə
pɹˈɪɾikˈɪɾi

wˈʌɾə
pɹˈɪɾikˈɪɾi



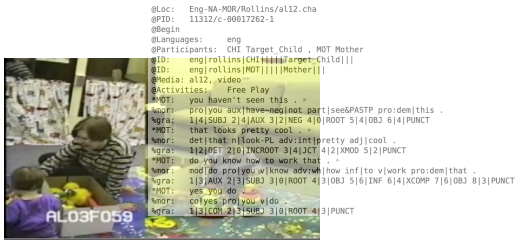
(6) Decide what the measure of success is

This can be based on your theory of what the developing grammar is like or data about children's behavior in different scenarios that demonstrates knowledge of that developing grammar.

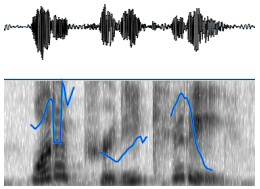
Experimental

Model-building

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

Example developing knowledge

Proto-lexicon of word forms

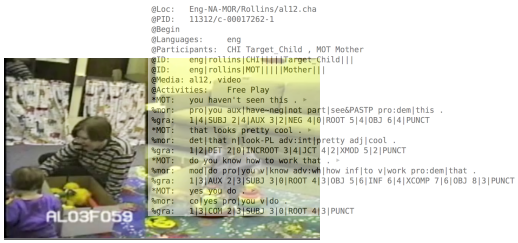
wˈʌɾ	*what*
ə	*a*
pɹˈɪɾi	*pretty*
kˈɪɾi	*kitty*

(6) Decide what the measure of success is

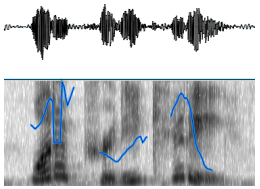
This can be based on your theory of the developing grammar or data about children's behavior

Model-building

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

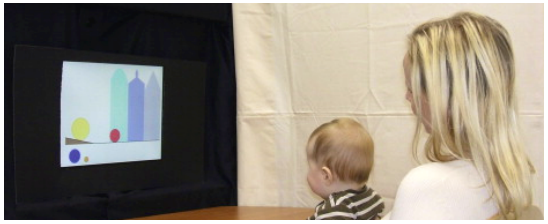
wˈʌɾ	*what*
ə	*a*
pɹˈɪɾi	*pretty*
kˈɪɾi	*kitty*

(6) Decide what the measure of success is

This can be based on your theory of the developing grammar or data about children's behavior

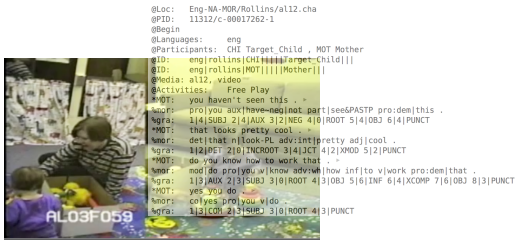
Example behavior indicating developed knowledge:

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

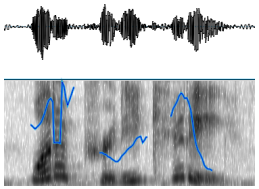


Model-building

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

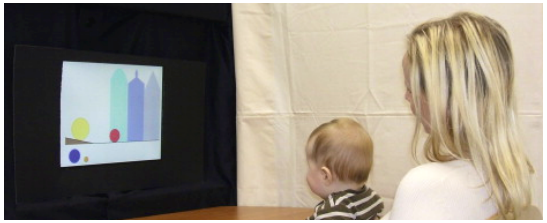
This is the heart of the model

wˈʌɾ — *what*

ə — *a*

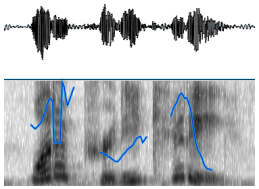
pɹˈɪɾi — *pretty*

kˈɪɾi — *kitty*



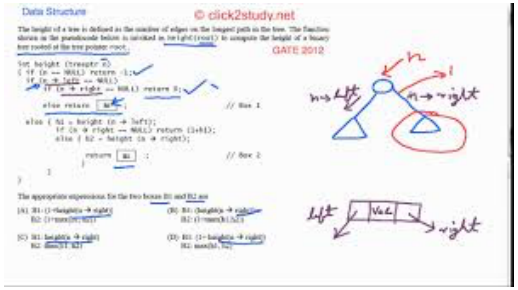
Model-building

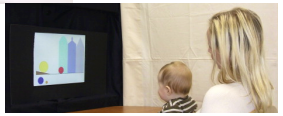
An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

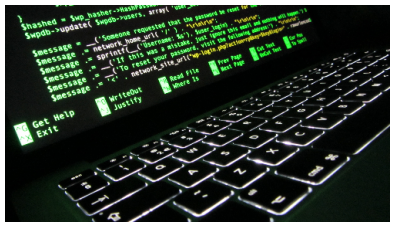
(7) Implement the model in a programming language of choice



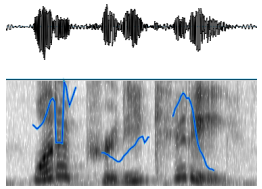


Model-building

An example with speech segmentation

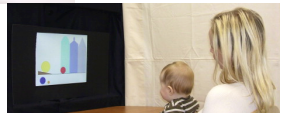


= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

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



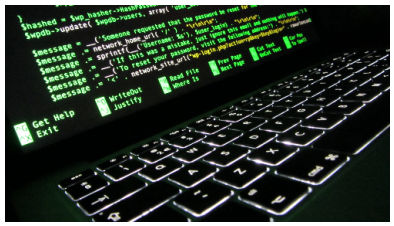
Example developing knowledge

Proto-lexicon of word forms

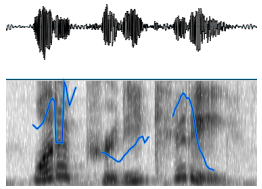
	wˈʌɾ	*what*
	ə	*a*
???	pɹˈɪɾi	*pretty*
	kˈɪɾi	*kitty*

Model-building

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



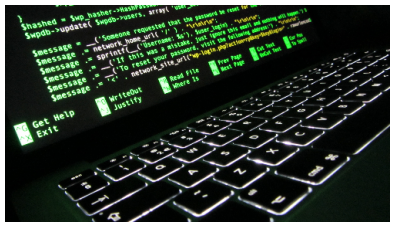
(8) See how well the model did w.r.t. the measure of success

Matching infant looking time behavior

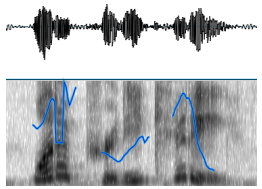
???

Model-building

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



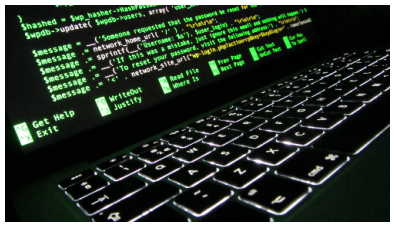
???

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

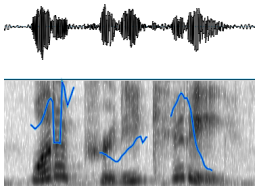
From this, we can determine how well the model did — and more importantly, how well the learning strategy implemented concretely in the model did.

Model-building

An example with speech segmentation



= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



???

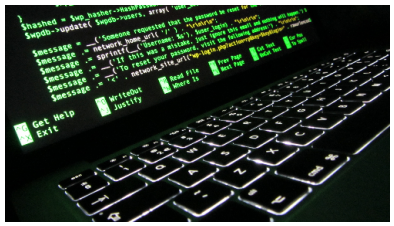
(9) Interpret the results for other people who aren't you so they know why they should care

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

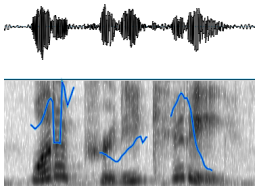
wˈʌɾ	*what*
ə	*a*
pɹˈɪɾi	*pretty*
kˈɪɾi	*kitty*

Model-building

An example with speech segmentation



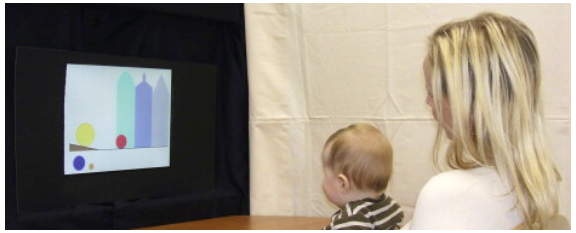
= wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi



???

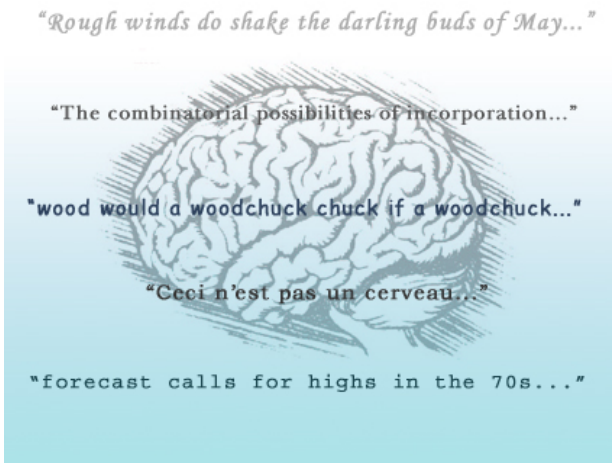
wˈʌ ɾə pɹˈɪ ɾi kˈɪ ɾi

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



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

Levels of explanation (Marr 1982)



Describing vs. explaining in vision

"…it gradually became clear that something important was missing …neurophysiology and psychophysics have as their business to *describe* the behavior of cells or of subjects but not to *explain* such behavior….What are the problems in doing it that need explaining, and what level of description should such explanations be sought?"
- Marr (1982)

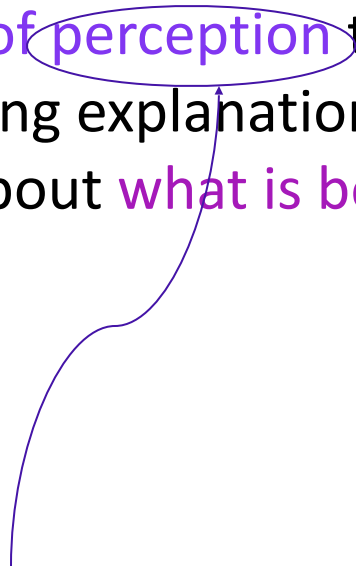
Describing vs. explaining

"This is a common trick of psychologists, to pretend they solved a riddle of the human mind by giving it a name, when all they've done is invented an agreed upon name for the mystery rather than solved it." - Tom Stafford, "The Psychology of Tetris"

http://www.bbc.com/future/story/20121022-the-psychology-of-tetris/1

On explaining (Marr 1982)

"But the important point is that if the notion of different types of understanding is taken very seriously, it allows the study of the information-processing basis of perception to be made *rigorous*. It becomes possible, by separating explanations into different levels, to make explicit statements about what is being computed and why…"



Our goal: Substitute "language acquisition" for "perception"

The three levels

Computational

What is the goal of the computation?

Algorithmic

What is the representation for the input and output, and what is the algorithm for the transformation?

Implementational

How can the representation and algorithm be realized physically?

The three levels:
An example with the cash register

Computational

What does this device do?

Arithmetic (ex: addition).

Addition: Mapping a pair of numbers to another number.

(3,4) → 7 [often written (3+4=7)]

Properties:

(3+4) = (4+3) [commutative]

(3+4)+5 = 3+(4+5) [associative]

(3+0) = 3 [identity element]

(3+ -3) = 0 [inverse element]

True no matter how numbers are represented: this is what is being computed

The three levels:
An example with the cash register

Computational

What does this device do?

Arithmetic (ex: addition).

Addition: Mapping a pair of numbers to another number.

Algorithmic

What is the input, output, and method of transformation?

Input: arabic numerals (0,1,2,3,4…)

Output: arabic numerals (0,1,2,3,4…)

Method of transformation: rules of addition, where least significant digits are added first and sums over 9 have their next digit carried over to the next column

$$\begin{array}{r} 99 \\ +\ \ 5 \\ \hline \end{array}$$

The three levels:
An example with the cash register

Computational

What does this device do?

Arithmetic (ex: addition).

Addition: Mapping a pair of numbers to another number.

Algorithmic

What is the input, output, and method of transformation?

Input: arabic numerals (0,1,2,3,4…)

Output: arabic numerals (0,1,2,3,4…)

Method of transformation: rules of addition, where least significant digits are added first and sums over 9 have their next digit carried over to the next column

$$\begin{array}{r} 99 \\ +\ 5 \\ \hline 14 \end{array}$$

The three levels:
An example with the cash register

Computational

What does this device do?

Arithmetic (ex: addition).

Addition: Mapping a pair of numbers to another number.

Algorithmic

What is the input, output, and method of transformation?

Input: arabic numerals (0,1,2,3,4…)

Output: arabic numerals (0,1,2,3,4…)

Method of transformation: rules of addition, where least significant digits are added first and sums over 9 have their next digit carried over to the next column

$$\begin{array}{r} {\scriptstyle 1} \\ 99 \\ +\ \ 5 \\ \hline 104 \end{array}$$

The three levels:
An example with the cash register

Computational

What does this device do?

Arithmetic (ex: addition).

Addition: Mapping a pair of numbers to another number.

Algorithmic

What is the input, output, and method of transformation?

Input: arabic numerals (0,1,2,3,4…)

Output: arabic numerals (0,1,2,3,4…)

Method of transformation: rules of addition, where least significant digits are added first and sums over 9 have their next digit carried over to the next column

$$\begin{array}{r} {\scriptstyle 1} \\ 99 \\ +\ \ 5 \\ \hline 104 \end{array}$$

The three levels:
An example with the cash register

Computational

What does this device do?

Arithmetic (ex: addition).

Addition: Mapping a pair of numbers to another number.

Algorithmic

What is the input, output, and method of transformation?

Input: arabic numerals (0,1,2,3,4…)

Output: arabic numerals (0,1,2,3,4…)

Method of transformation: rules of addition

Implementational

How can the representation and algorithm be realized physically?

A series of electrical and mechanical components inside the cash register.

The three levels:
An example with a sandwich

Computational

What is the goal?

Make a peanutbutter and jelly sandwich.

Properties:

- slices of bread containing both peanutbutter and jelly
- number of bread slices: 2
- sandwich is sliced in half
- crusts are left on
- jelly type: grape
- peanutbutter type: crunchy

etc.

The three levels:
An example with a sandwich

Computational

What is the goal?

Make a peanutbutter and jelly sandwich.

Algorithmic

What is the input, output, and method of transformation?

Input: ingredients (peanutbutter, jelly, bread slices), tools (knife, spoon)

Output: completed, edible sandwich with the required properties

Method: Use the spoon to put jelly on one slice & spread it with the knife. Use the spoon to put peanutbutter on the other slice & spread it with the knife. Put the two slices of bread together, with the spread sides facing each other. Cut the joined slices in half with the knife.

The three levels:
An example with a sandwich

Computational

What is the goal?

Make a peanutbutter and jelly sandwich.

Algorithmic

What is the input, output, and method of transformation?

Input: ingredients (peanutbutter, jelly, bread slices), tools (knife, spoon)
Output: completed, edible sandwich with the required properties
Method: PBJ-making steps.

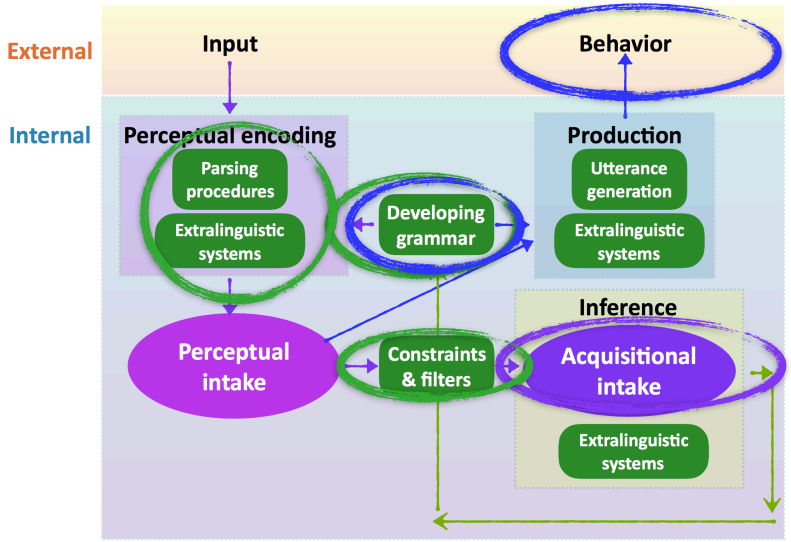
Implementational

How can the representation and algorithm be realized physically?
Directing your younger sibling to follow the steps above to make you a sandwich.

Levels when modeling 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?



Computational-level (Marr 1982)

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

Levels when modeling 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?

Levels when modeling 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.)

Levels when modeling 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).

Levels when modeling language acquisition

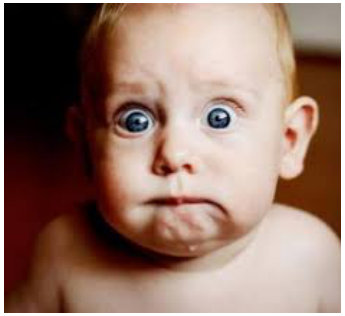
What level of model do you want to build?

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?

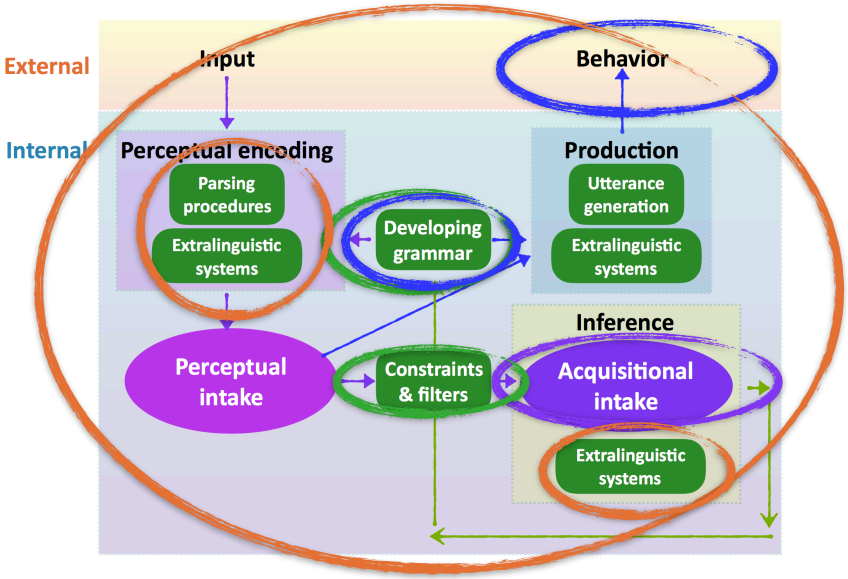


Because they're solving the wrong acquisition task…oops.

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?



Levels when modeling 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.

Levels when modeling language acquisition

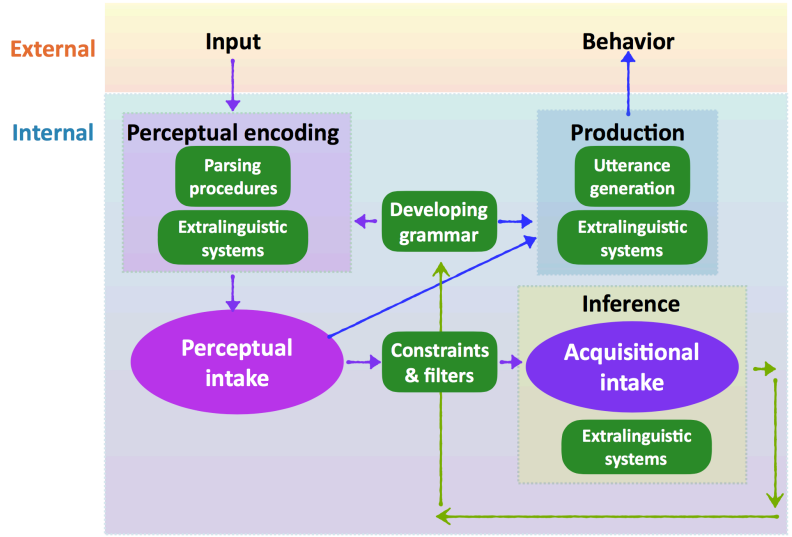
What level of model do you want to build?

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



Levels when modeling language acquisition

What level of model do you want to build?

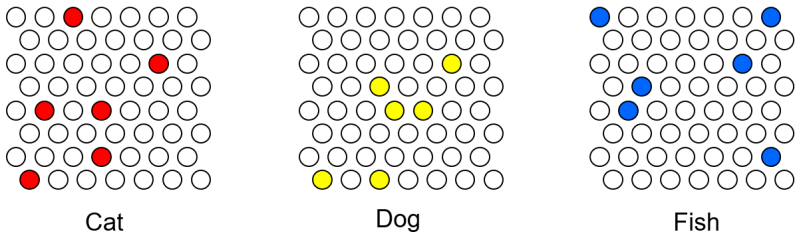
Computational-level

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

Algorithmic-level

Implementational-level

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



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

Levels when modeling language acquisition

What level of model do you want to build?

The types I generally work with

Computational-level

Algorithmic-level

Implementational-level

Recap

Computational modeling is a technique for understanding how language acquisition works — in particular, understanding the learning strategies children use to solve different acquisition tasks.

Computational modeling is useful for both specifying and testing theories of language acquisition.

One of the main goals of the study of language acquisition is to explain it, rather than just describe it.

There are three different levels of explanation, according to Marr: the computational level, the algorithmic level, and the implementational level. When modeling language acquisition, it's important to know what level of model you're trying to build in order to know how to interpret the model results.

Questions?

You should be able to do all the introductory review questions and all of the questions on HW1.