

# Psych 56L/ Ling 51: Acquisition of Language

## Lecture 2 The study of language acquisition

### Announcements

#### **Galia's office hours (starting this week):**

Wednesdays 10am-12pm, 2pm-3pm in SBSG 2221

Review questions for introductory material available on website

Homework 1 available (be working on it): due 10/13/15

Remember to look at the reference material in addition to downloading the lecture notes & listening to the podcasts (when available)

### Research methods



### Research methods

Important: do cross-linguistic and cross-cultural research. Even if language is universal, there are individual differences in language development and there may be more than one route to acquisition success. Also, there may be influence from different cultures on the language learning environment for children.



## Research methods

**Diary studies:** keeping diaries of children's development. Charles Darwin did this with his son (Darwin 1877), who seemed to follow the progression we now expect.



Other diary studies: Clara & Wilhelm Stern's 1907 *Die Kindersprache* and Werner Leopold's (1939-1949) four volume account of his daughter's acquisition of English & German.

Modern diary studies: Braunwald 1976; Bowerman 1985, 1990; Dromi 1987; A. Gopnik & Meltzoff 1987; L. Bloom, 1993; Naigles, Vear, & Hoff 2002

## A very modern diary study

[http://www.ted.com/talks/deb\\_roy\\_the\\_birth\\_of\\_a\\_word.html](http://www.ted.com/talks/deb_roy_the_birth_of_a_word.html)

Beginning through about 4:15 (full video is about 17 minutes total)



## Research methods

CHILDES Child Language Data Exchange System



<http://childes.psy.cmu.edu>

Video/audio recordings of spontaneous speech samples, along with transcriptions and some structural annotation. Extremely valuable resource to the language acquisition community.



```
@Loc: Eng-NA-MOR/Rollins/all2.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: all2, video
@Activities: Free Play
*MOT: you haven't seen this . .
*Mor: prɔjyɔu əux[həv-neg]not part[see@PASTP prɔ: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]2NCRROOT 3[4]2CT 4[2]XMOD 5[2]PUNCT
*MOT: do you know how to work that . .
*Mor: mod|do prɔjyɔu v|know adv:wh|how Inf|to v|work prɔ: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: cɔjɛs prɔjyɔu v|do .
*gra: 1[3]COM 2[3]SUBJ 3[0]ROOT 4[3]PUNCT
```

## Research methods

CHILDES Child Language Data Exchange System



<http://childes.psy.cmu.edu>

Difficulty: Have to transcribe recorded speech. May take between 5 and 20 hours to faithfully transcribe 1 hour of child speech.

Why?

Conversational speech does not often use complete sentences.

Child pronunciation is often not adult-like - and the non-adult-like parts are usually what researchers are interested in.

## Research methods

CHILDES Child Language Data Exchange System



<http://childes.psy.cmu.edu>

“In terms of its impact on the field of language development, CHILDES is a game-changer. It allows researchers with limited resources to test hypotheses using an extremely rich data set. It allows for comparison across many different languages, which makes it possible to look for universal cross-linguistic patterns in language development....because the transcripts also include language by the adults that the children are interacting with, it also allows researchers to test detailed quantitative predictions about the relationships between a child’s input and her language production.” — Sedivy 2014, p.224

## Research methods

CHILDES Child Language Data Exchange System



<http://childes.psy.cmu.edu>

Used to find out the nature of language children produce. Ideally, sample is representative of everything child says - but hard to do in practice. (Deb Roy’s work is a notable exception.)

Because of this, it is hard to make claims that children don’t use/know a particular structure based on its absence in spontaneous speech samples. It could be that they simply didn’t say that structure when they were being recorded.

## Research methods

Getting standardized assessments of children’s performance

Use coding systems like Mean Length of Utterance (MLU), which correlates with measures of children’s grammatical and phonological development. This is done by tracking the average number of meaning-bearing units (morphemes) in the child’s speech.

Ex: “He likes me” = 4 morphemes (“he”, “like”, “-s”, “me”)

Use estimates that caregivers provide of children’s performance, such as the MacArthur-Bates Communicative Development Inventories (CDIs): 8-16 months, 16-30 months, 30-36 months. These include checklists of words, gestures, and word combinations children produce or comprehend.

## Research methods

Some ways to assess children’s comprehension abilities:

(1) Use examiner-administered tests like the Peabody Picture Vocabulary Test, where the child points at a picture matching the word(s).

(2) Act-out tasks: The child is given toys and a linguistic description, and must make the toys act out the appropriate scenario.

“The wolf is happy to bite the lion.”



<https://www.youtube.com/watch?v=UY04SEjZJSw&list=PL95604CD0326F659A&index=2>

## Research methods

Some ways to assess children's **comprehension** abilities:

(3) **Pointing tasks**: The child points at the picture that matches the linguistic description (words or sentences).

(4) **Grammaticality judgment tasks**: Child indicates whether spoken utterance sounds "okay" or "silly".

Grammaticality: Is this a silly thing to say?



Every penguin ate two fish. 😊

Every penguin went two fish. 😞

## Research methods

Some ways to assess children's **production** abilities:

(1) **elicited production**:

"What's Ernie doing?" "What happened to the ball?"

(2) **repetition/imitation elicitation**:

"Say this: 'After she ate the peach, Sarah feel asleep.'"

(3) **syntactic priming**: Modeling a syntactic construction with one utterance, and having the child produce a novel utterance that uses that same construction

Passive example:

"...the ball is being bounced by Ernie...Oh look! What's happening to that peach?"

(Intended response: "The peach is being eaten by Sarah.")

## Research methods

### Computational modeling (Digital children)

Create a computer program that takes the data children hear as input and see if it can learn the same knowledge children do from that input.

Usually, the program will **implement some learning theory's assumptions about how learning works** (ex: what learning strategies children might use), and therefore test that theory empirically.

Ex: Learning to identify words in fluent speech (speech segmentation): Swingley 2005, Gambell & Yang 2006, Pearl, Goldwater, & Steyvers 2011, Phillips & Pearl 2012, 2014a, 2014b, 2015, in press

húwzəfɹéjdəvðəbíg bædwɹlf



húwz əfɹéjd əv ðə bíg bæd wɹlf  
who's afraid of the big bad wolf

## Theoretical viewpoints



## The question

“It is obvious that children have some quality of mind that explains why they learn to talk but kittens, for example, do not” – Hoff 2008, p.254

Not obvious what this quality is.

Idea 1: Children have specialized (domain-specific) knowledge about how language works.

Idea 2: Children’s domain-general cognitive processes allow them to acquire language while a kitten’s do not.

## Chomskyan revolution

Chomsky 1957: *Syntactic Structures*

Innovation: What speakers do is not as interesting as the mental grammar that underlies what speakers do



So, if adults have a mental grammar that explains what they do when they talk, children must have a mental grammar that explains what children do when they talk.



New formation of language development: What are children’s grammars like and how do they eventually achieve adult grammars?

## Some current approaches

Language as a complex cognitive system that maps sounds to meaning

One idea for the mechanism behind this process: Language Acquisition Device

Information from the environment



→ Language Acquisition

Language Acquisition Device  
(unconscious process inside child’s mind, used only for learning language)

## Some current approaches

### Linguistic approach

Premise: LAD contains some domain-specific knowledge about the structure of language (this is often called Universal Grammar).

Knowledge specifically about human language

Focus: description of children’s prior (innate) linguistic knowledge and how that knowledge interacts with the data from the native language to produce knowledge of the native language

## Some current approaches

### LAD + information from the environment

Basic premise: The language acquisition device provides a **little bit of knowledge about how human languages work to get the child started**. This allows the child to **use her language input more effectively** – to notice certain things more easily and to entertain only certain hypotheses about how language works.



## Innate linguistic knowledge?

Why do children need this kind of head start?

Proposal: Input is too impoverished for children to converge on the right language rules without it. This is sometimes called the **Poverty of the Stimulus**.

So, children need something else besides just the data in the input to help them decide what the right rules are.

## Some current approaches

Another idea for the mechanism behind this process: **general learning abilities**

### Domain-general cognitive approach

Premise: Language acquisition is no different from any other kind of knowledge acquisition; children can solve this problem in the same way that they solve other problems (such as perception)

Focus: description of **domain-general** learning capacities that serve language development, and the sources of input those capacities use

Useful for all kinds of learning (ex: grouping things together into larger units)

## Some current approaches

### Domain-general cognitive approach

Basic premise: Abilities that are useful for other kinds of input besides language input are used to learn language. **There is no knowledge or ability that is unique to language learning.**

## Domain-general response to Poverty of the Stimulus

Maybe children don't need domain-specific knowledge to learn language. *Maybe they just use the data available to them more cleverly than some researchers think they do.*

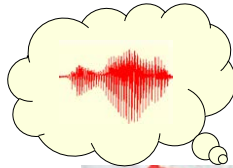
Example:

Saffran, Aslin, & Newport (1996): 8-month-olds can (unconsciously) track probabilities between syllables in order to identify words in fluent speech in an artificial language

*tu pi ro go la bu bi da ku pa do ti go la bu tu pi ro pa do ti...*

Sample audio input

[http://whyfiles.org/058language/images/baby\\_stream.aiff](http://whyfiles.org/058language/images/baby_stream.aiff)

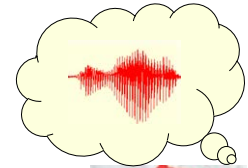
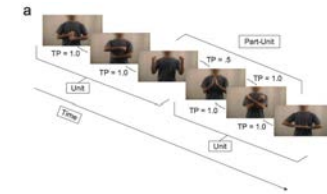


## Domain-general response to Poverty of the Stimulus

Maybe children don't need domain-specific knowledge to learn language. *Maybe they just use the data available to them more cleverly than some researchers think they do.*

Example:

Roseberry, Richie, Hirsh-Pasek, Golinkoff, & Shipley (2012): 8-month-old infants are able to (unconsciously) track probabilities between dynamic events, such as a series of hand motions.

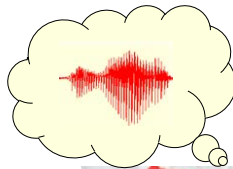
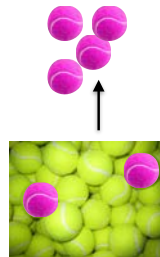


## Domain-general response to Poverty of the Stimulus

Maybe children don't need domain-specific knowledge to learn language. *Maybe they just use the data available to them more cleverly than some researchers think they do.*

Example:

Denison, Reed, & Xu (2011): 6-month-old infants are able to create probabilistic expectations about their environment, based on their observations of their environment. For example, after seeing that a box is mostly filled with yellow balls, they are surprised when someone pulls four pink balls in a row out of the box.

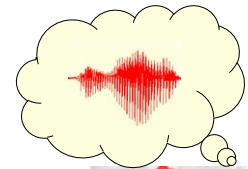


## Domain-general response to Poverty of the Stimulus

Maybe children don't need domain-specific knowledge to learn language. *Maybe they just use the data available to them more cleverly than some researchers think they do.*

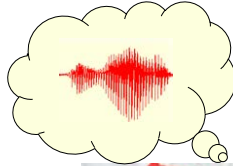
Example:

Denison, Bonawitz, Gopnik & Griffiths (2013): 4- and 5-year-olds select a hypothesis to evaluate against the data based on how probable a hypothesis is (called sampling a hypothesis). For example, when guessing which color block fell into a container from a box where 5 blue and 20 red blocks were, children guess blue 20% of the time (5/25) and red 80% of the time (20/25).



## Domain-general response to Poverty of the Stimulus

Maybe children don't need domain-specific knowledge to learn language. **Maybe they just use the data available to them more cleverly than some researchers think they do.**

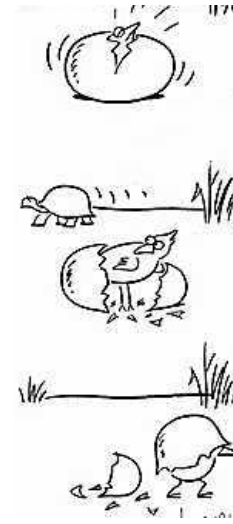


Example:

Kidd, Piantadosi, & Aslin (2012): 7- to 8-month-old infants have a tendency to learn only from data whose informational complexity is neither too high nor too low (the "Goldilocks Effect").



## Nature vs. Nurture



## The debate in a nutshell

Is the development of language in children the result of humans' innate endowment (like upright posture & bipedal locomotion)? Or is it the result of circumstances in which children are nurtured (like table manners and formal math, which depend on particular experiences)?

**Empiricism: all knowledge and reason come from experience**



**Nativism: mind has some pre-existing structure it imposes to interpret experience**

## Nativism: Why believe it?

- (1) Children acquire language rapidly
- (2) Children acquire language with very little conscious effort
- (3) Children acquire language without explicit instruction for most of it



**Nativism: mind has some pre-existing structure it imposes to interpret experience**



## Nativism: Why believe it?

“Language learning is not really something that the child does; it is something that happens to a child placed in an appropriate environment, much as the child’s body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation.” - Chomsky, 1973



Nativism: mind has some pre-existing structure it imposes to interpret experience

## Nativism: Why believe it?

Arguments for Nativism (and Universal Grammar in particular)  
Up through ~2:38 for general intro, 7:37 - 8:34 for summary

<http://www.thelingspace.com/episode-1>

<https://www.youtube.com/watch?v=MLNFGWJQjA>



## Constructionist View

“We on the other side think that learning language is a long slog, which requires from the child a lot of work. And the child is working as hard as he can, fifteen, sixteen hours a day. We think it requires a relationship with an adult, and a whole set of cognitive abilities.” - Snow, 1993



Constructionist: language is constructed by the child from experience, and the input is crucial - but there may still be some innate knowledge contributing

## Back to nativism: the nature of nature

There are different ways for something to be innate:

Knowledge itself is innate

Procedures for learning are innate (knowledge is the result from these procedures)

## Back to nativism: the nature of nature

There are different ways for something to be innate:

Knowledge itself is innate: children have inborn knowledge of the general form of language (domain-specific knowledge)

Procedures for learning are innate (knowledge is the result from these procedures)

## Why do we think knowledge could be innate?

Common properties of human languages: all languages of the world share structural properties. This could be due to innate biases about how languages are structured.

Evolution has equipped the human mind with other useful knowledge (ex: world is 3D, even though retinas process only 2D) - why not prior knowledge about language?



## Back to nativism: the nature of nature

There are different ways for something to be innate:

Knowledge itself is innate: children have inborn knowledge of the general form of language (domain-specific capacities)

Procedures for learning are innate (knowledge is the result from these procedures): children have domain-general capacities that all contribute to language acquisition, such as symbolic representation, memory, chunking input into smaller parts, and probabilistic analysis.

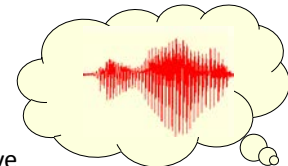
## Why do we think some learning procedures are innate?

Babies as statistical learners

Statistical learning: keeping track of the relative frequency of two things (ex: how often they occur together)

Evidence that infants (6-month-olds, 8-month-olds) are capable of statistical learning and probabilistic reasoning abilities:

Saffran et al. 1996, Denison et al. 2011, Roseberry et al. 2012

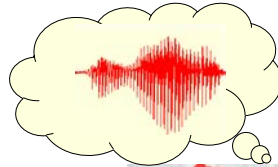


## Why do we think some learning procedures are innate?

Babies as statistical learners

Statistical learning is domain-general.

Saffran, Johnson, Aslin, & Newport (1999): babies can track the probabilities between tones (not just between language stimuli like syllables)

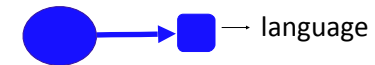


Denison et al. (2011): Infants can create probabilistic expectations about their environment (such as the color of balls in boxes), not just about language.

Roseberry et al. (2012): Infants can track probabilities between dynamic events.

## Back to nativism: the nature of nature

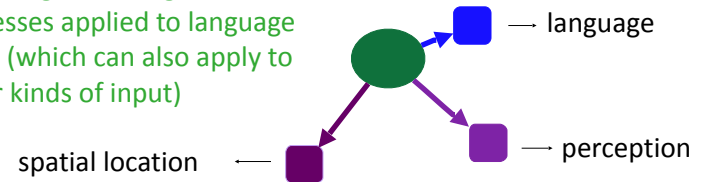
There are different ways for language acquisition to work:



One domain-specific module



Domain-general cognitive processes applied to language input (which can also apply to other kinds of input)



## Back to nativism: the nature of nature

There are different ways for language acquisition to work:



Currently this debate between domain-specific and domain-general is going on for many areas of cognition, not just for language acquisition.

## Viewpoint comparison

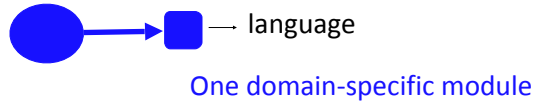
Generativist

Constructionist

## Viewpoint comparison

Generativist: Universal Grammar, which contains biases for language structure, is innate. Language experience triggers prior knowledge and/or language-specific learning abilities (domain-specific).

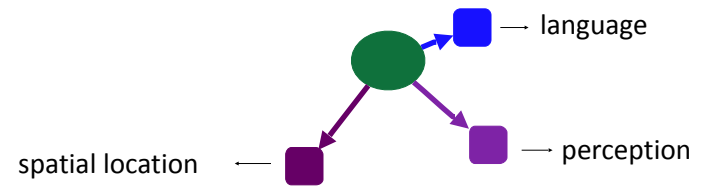
Constructionist



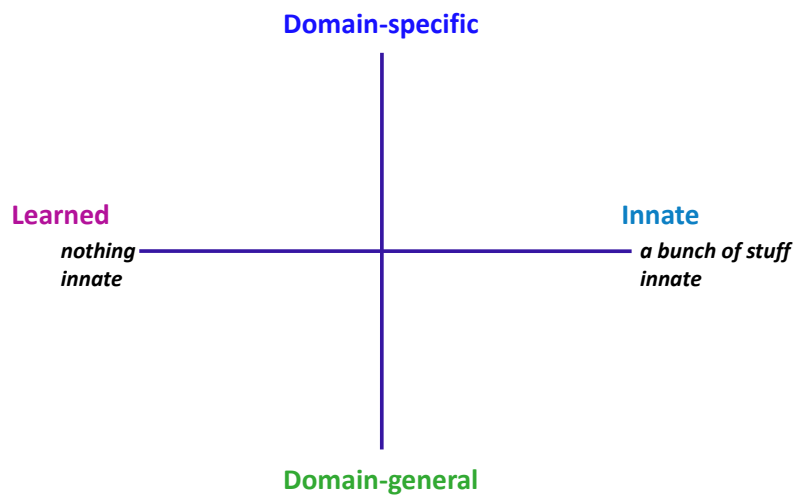
## Viewpoint comparison

Generativist

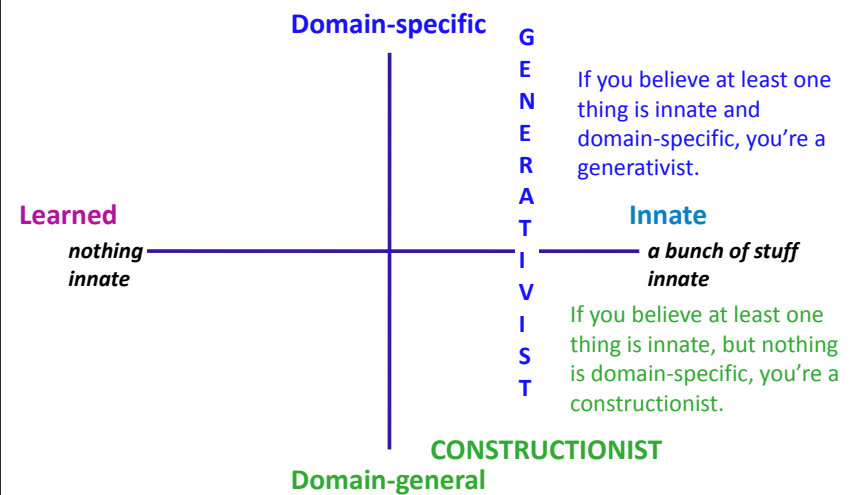
Constructionist: language is constructed by the child using general cognitive learning procedures applied to language input. These are domain-general abilities used for language learning.



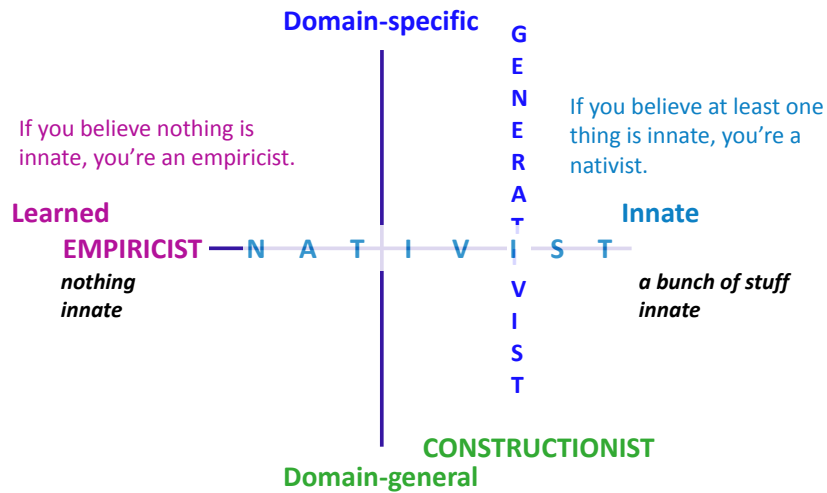
## An important division



## An important division



## An important division



## Another way to think about it

	<i>nothing innate</i>	<i>at least one thing innate</i>
<i>at least one thing domain-specific</i>	<i>empiricist</i>	<i>nativist, generativist</i>
<i>nothing domain-specific</i>	<i>empiricist</i>	<i>nativist, constructionist</i>

## Recap

There are different methods for investigating questions in language acquisition, most of which involve using child-directed input and child-produced output. One research method gaining prominence in the field is computational modeling, which tends to look at specific implementations of how the process of language acquisition could work.

Some current approaches to how language acquisition works include the generativist approach and the constructionist approach. Both believe in innate knowledge, though only the generativist approach believes at least some of that knowledge is domain-specific.

## Questions?



You should now be able to answer all of the review questions for the introductory material, up through question 4 on HW1.