LSci 51/Psych 56L: Acquisition of Language

Lecture 2

Children's input &

Research methods

Announcements

Be working on review questions & comments for intro material

Be working on HW1
(due 8/8/25 at 12:50pm, submitted through Canvas)

Please note that you can only submit HW assignments once.
Only submit when you've completed the entire assignment.

So how exactly do children learn all this?

We know they do it relatively quickly.

speech segmentation

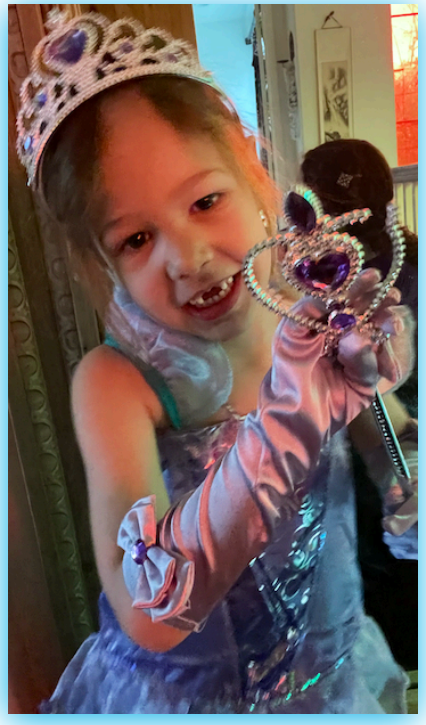
phonology

syntactic categorization

syntax

syntax, semantics

pragmatics



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

Interesting: They do this mostly without explicit instruction.

What about learning by explicit correction?

Even if the knowledge is subconscious, couldn't parents teach children these rules of language by explicitly correcting them when they say something wrong?

What about learning by explicit correction?

Even if the knowledge is subconscious, couldn't parents teach children these rules of language by explicitly correcting them when they say something wrong?

The problem: parents don't correct their children that often about the form of the language. Instead, they tend to correct when the meaning is incorrect.

Child: "Her curl my hair."

Parent: "Uh huh."

Child: "There's an animal farmhouse."

Parent: "No, that's a lighthouse."

What about learning by implicit correction?

Parents may provide implicit correction by offering alternative language forms when a child has said something incorrect. In effect, the parents provide a good example of language use for children without explicitly correcting them. This is called a recast (or reformulation).

Child: The dog runned really fast, Daddy.

Parent: Yeah, he ran really fast, didn't he?

Recasts?

https://www.youtube.com/watch?v=a7Un06tDOn0&feature=youtu.be

4:33-5:31

What about learning by implicit correction?

However, parents don't provide recasts all the time or all that consistently (Waller, Nozari, & Yurovsky 2022): mothers only made recasts after 27.3% of incorrect forms. The rest of the time, they didn't bother.

Also, sometimes parents will repeat children's incorrect utterances (e.g., 16.1% from Waller et al. 2022) if they agree with the meaning of the meaning! This would seem to reinforce the incorrect language usage.

Child: Read book.

Mother: Alright, you read book.

(instead of *read the book*)

What about learning by implicit correction?

Still, recasts can be helpful when they offer a direct and immediate contrast between the child's way of saying something and the correct way (Saxton et al. 1998, Waller et al. 2022).

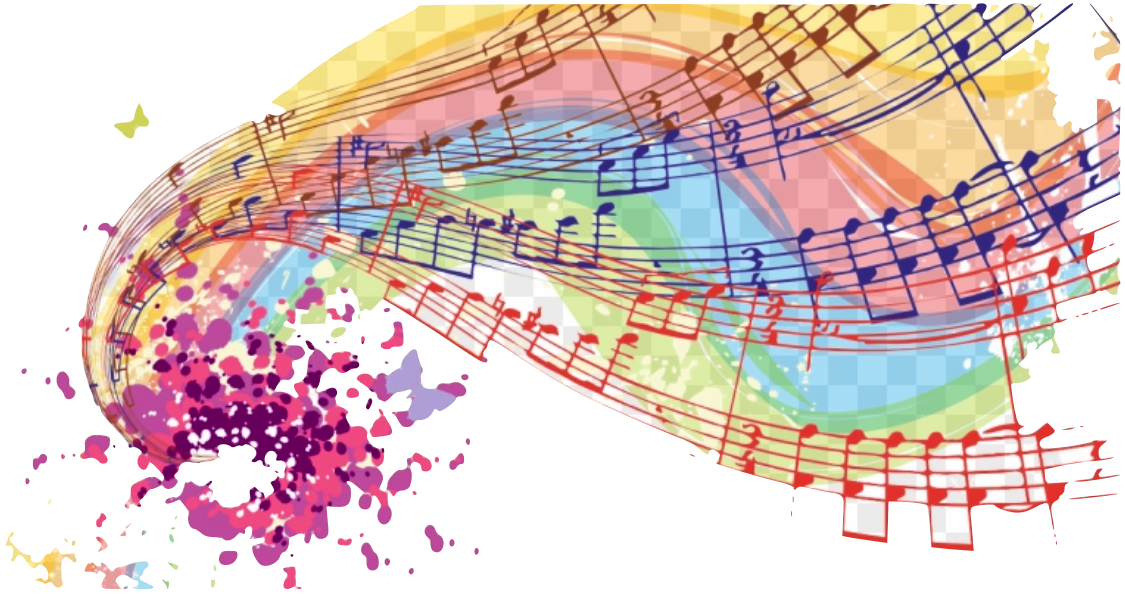
Taumoepeau 2016: The percentage of utterances caretakers expanded (and recast) when their children were between the ages of 24 and 33 months had a strong impact on children's vocabulary development.

Recasts may help speed up learning, but probably aren't responsible for learning all knowledge about language.

The nature of the input

About the input

"Motherese has interpretable melodies: a rise-and-fall contour for approving, a set of sharp, staccato bursts for prohibiting, a rise pattern for directing attention, and smooth, low legato murmurs for comforting." – Pinker, *The Language Instinct*



About the input

https://www.sciencedaily.com/releases/2017/10/171012143326.htm

"We use **timbre**, the tone color or unique quality of a sound, all the time to distinguish people, animals, and instruments…We found that mothers alter this basic quality of their voices when speaking to infants, and they do so in a highly consistent way across many diverse languages." — Elise Piazza, Princeton (about the findings of Piazza et al. 2017)

"Timbre is the reason it's so easy to discern idiosyncratic voices -- the famously velvety sound of Barry White, the nasal tone of Gilbert Gottfried, and the gravelly sound of Tom Waits -- even if they're all singing the same note."

About the input

Properties of motherese (speech adults use with children):

prosodic features are **exaggerated**, and pauses tend to occur at phrase boundaries (helping to identify how words cluster together into larger units like phrases)

"The brave older *sister* (pause)
went to *rescue* (pause)
her *little baby brother* Toby."

"The brave older sister" = noun phrase
"her little baby brother Toby" = noun phrase

Noun phrase indicator: Can replace with pronoun
"The brave older sister" = *she*
"her little baby brother Toby" = *him*

About the input

Properties of motherese (speech adults use with children):

prosodic features are **exaggerated**

Rosslund, Mayor, Mundry, Singh, Cristia, & Kartushina 2024: "Parents used higher pitch, wider pitch range, slower articulation rate, longer vowel duration, and more variable and less distinct vowels in IDS than in ADS…featured wider pitch range, larger vowel space areas, and shorter vowel duration in older as compared to younger infants…"

About the input

Properties of motherese (speech adults use with children):

prosodic features are **exaggerated**

https://www.sciencedaily.com/releases/2022/10/221011105727.htm

Cox, Bergmann, Fowler, Keren-Portnoy, Roepstorff, Bryant, & Fusaroli 2022

cross-linguistic: "…certain features of [motherese], such as pitch, melody, and articulation rates have the same properties across most of the world's languages"

About the input

Properties of motherese (speech adults use with children):

prosodic features are **exaggerated**

Räsänen, Kakorous, & Soderstrom 2017:

pitch contours (one very key prosodic feature) are far less predictable — and are therefore far more surprising and attention-getting — in motherese, compared to adult-directed speech

About the input

Properties of motherese (speech adults use with children):

prosodic features are **exaggerated**

Panneton, Cristia, Taylor, & Moon 2023: "hyperarticulation may also result from more positive valence (e.g., speaking with positive vocal emotion) often found in mothers' speech to infants"

About the input

Properties of motherese (speech adults use with children):

prosodic features are **exaggerated**…and maybe sound more like the child's own productions

On Polka, Masapollo, & Ménard 2021:
"By mimicking the sound of a smaller vocal tract…we're cluing babies in to how the words should sound coming out of their own mouths….babies' dawning ability to control their voices and make words out of babble could be what makes the infant-like sounds more appealing"

https://www.sciencedaily.com/releases/2021/12/211210121848.htm

What about “fatherese”?

VanDam, DePalma, & Strong (2015):

Fatherese may serve as a bridge intonation-wise

“…the mothers used higher pitch and varied their pitch more when interacting with their child than with adults. The fathers, on the other hand, did not show the same pattern, and instead talked to their children **using intonation patterns more like when they talked to other adults**…The data support what VanDam refers to as the bridge hypothesis -- that fathers, by speaking to their children more like adults, **might act as a link to the outside world by helping them to deal with unfamiliar speech**.”

http://www.sciencedaily.com/releases/2015/05/150519083257.htm

What about “fatherese”?

A fatherese demo

https://www.youtube.com/watch?v=6OUGNgTZATw

About the input

Properties of motherese (speech adults use with children):

topics are about the here and now (easier to link words to meanings) (Hills 2013)

Note: There is considerable individual variation in how well and how much caretakers do this, but children of caretakers who do this more learn vocabulary faster (Cartmill et al. 2013).

About the input

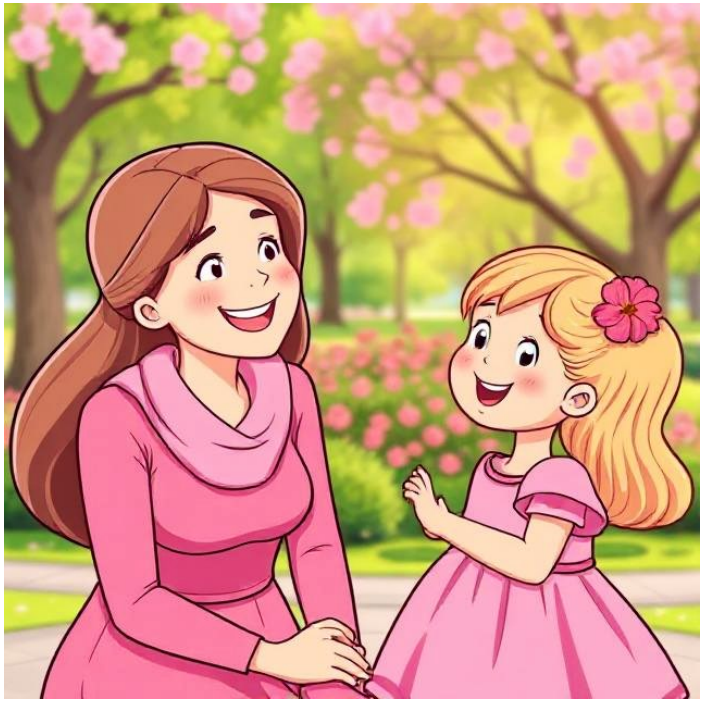
Properties of motherese (speech adults use with children):

When talking about objects, English adults tend to say the name of the object last ("*this is the [object]*") and precede it with a small set of reliable cues (ex*: the, a*) (Yurovsky et al. 2013).

About the input

Properties of motherese (speech adults use with children):

very few grammatical errors (good example of correct grammar usage)



About the input

Properties of motherese (speech adults use with children):

adults tend to use gestures to secure children's attention (easier to link words to meanings) — in general, engaging children socially is very important for the input to have an impact

About the input

Properties of motherese (speech adults use with children):

speech is repetitious (easier to remember when you have a short attention span) (Hills 2013)

example of repetitious speech:

"What a pretty sunhat! I love your sunhat! Your sunhat looks so nice!"

About the input

Properties of motherese (speech adults use with children):

adults will often expand children's utterances (learning how to convey the meaning they want by example)

example of expanding an utterance:

"Sunhat!"

"Your sunhat is so pretty!"

About the input

Properties of motherese (speech adults use with children):

child-directed speech is tuned to the level of linguistic complexity (vocabulary, syntax, function words) the child can handle until around age five (Yurovsky, Doyle, & Frank 2016, Denby & Yurovsky 2019, Leung, Tunkel, & Yurovsky 2019, 2021) — it's easiest to absorb information if it's neither too simple or too complex. The better tuned the child-directed speech is, the better the child's linguistic development is (Denby & Yurovsky 2019).

https://www.sciencedaily.com/releases/2021/07/210702114538.htm

Helpful motherese

Motherese can help jumpstart the language parts of the brain:

Just 24 hours after birth, the sound of a **mother's voice specifically activates the language processing and motor circuits of the brain**, moreso even than another female voice.

(Beauchemin et al. 2010)

Helpful motherese

Motherese can help jumpstart the language parts of the brain:

"…infants who heard more words…the structure of their white matter was slower to develop. The children went on to have better linguistic performance when they began to talk….slower maturation of white matter confers a cognitive advantage."

(Estrada, Govindaraj, Abdi, Moraglia, Wolff, Meera, Dager, McKinstry, Styner, Zwaigenbaum, Piven, & Swanson 2023)

https://www.sciencedaily.com/releases/2023/06/230605181307.htm

Helpful motherese

Children who attend day care centers with more **one-on-one contact with an adult** acquire language more rapidly than children who get less one-on-one adult contact (Hoff 2006).

Older children (who receive all of their parents' child-directed speech) generally develop language earlier than later-born children, who have to share it with their siblings (Hoff-Ginsberg 1998).

Helpful motherese

Parents who were coached to use more motherese with their 6- and 10-month-olds ended up with 14-month-olds who produced significantly more words (Ferjan Ramírez, Roseberry Lytle, Fish, & Kuhl 2018) and 18-month-olds who had more vocalizations & conversational turn-taking as well as larger vocabularies (Ferjan Ramírez, Roseberry Lytle, & Kuhl 2020)

https://www.sciencedaily.com/releases/2018/11/181126123348.htm
https://www.sciencedaily.com/releases/2020/02/200203151158.htm

Helpful motherese

What this means: “…[motherese] sounds happy and conveys total engagement with the child. Spoken directly to the child -- and used across many languages -- [motherese] resonates with infants…and helps babies tune in socially to their parents, and motivates them to talk back, even if that just means babbling”

https://www.sciencedaily.com/releases/2018/11/181126123348.htm
https://www.sciencedaily.com/releases/2020/02/200203151158.htm

Helpful motherese

21-month-olds learn new words better from child-directed speech, as compared to adult-directed speech (Ma et al. 2011).

There's something special about words specifically directed at children, compared to words children simply overhear – words that are simply overheard have very little impact on vocabulary acquisition (Schneidman et al. 2013)….at least until children are preschool age (Foushee & Xu 2016). This may have to do with the relative complexity — overheard speech is more complex than child-directed speech until children are 30 months old (Foushee, Griffiths, & Srinivasan 2016, Lakoutou, Le Normand, & Cristia 2019).

Helpful motherese

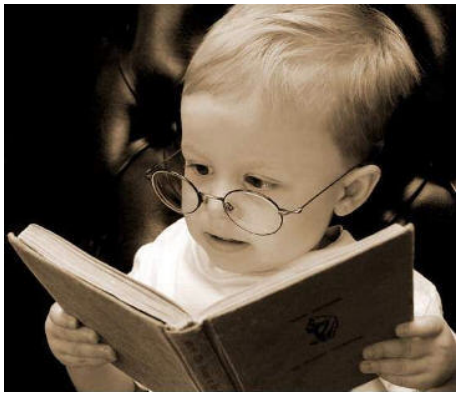
Bergelson, Soderstrom, Schwarz, Rowland, Ramírez-Esparza, Hamrick, Marklund, Guez, Casillas, Benetti, van Alphen & Cristia 2023:

“1,001 2- to 48-month-olds from 12 countries spanning six continents across urban, farmer-forager, and subsistence-farming contexts….children who heard more talk from adults produced more speech…”

But not really necessary…

Motherese seems very helpful in a variety of ways, but is unlikely to be necessary for successful language acquisition. This is because children in some cultures (Tseltal Mayan: Casillas et al. 2019, Foushee & Srinivasan 2024; Inuit: Crago et al. 1997) don't receive much (or any) child-directed speech, and yet still manage to acquire their native languages just fine.

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

[Extra]

A very modern diary study

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.talkbank.org

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

Research methods

CHILDES — Child Language Data Exchange System

http://childes.talkbank.org

Also, it's really important to be able to test learning theories on realistic data that comes from natural environments, and not just experimental settings. Sometimes you get different results (Lavechin, de Seyssel, Métais, Metze, Mohamed, Bredin, Dupoux, & Cristia 2024).

Research methods

CHILDES

Child Language Data Exchange System

http://childes.talkbank.org

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

Why?

Conversational speech doesn't often use complete sentences.

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

http://childes.talkbank.org/browser/index.php?url=Eng-NA/Braunwald/1-05-09.cha

Research methods

CHILDES

Child Language Data Exchange System

http://childes.talkbank.org

Example from the Braunwald corpus

http://childes.talkbank.org/browser/index.php?url=Eng-NA/Braunwald/1-05-09.cha

```
62    *CHI: no . [+ SR]
63  %mor:   co|no .
64  %gra:   1|0|INCROOT 2|1|PUNCT
65    *CHI: <get down> [?] .
66  %mor:   v|get adv|down .
67  %gra:   1|0|ROOT 2|1|JCT 3|1|PUNCT
68    *MOT: what do you want to do , Laura ?
69  %mor:   pro:int|what mod|do pro:per|you v|want in
70          n:prop|Laura ?
71  %gra:   1|4|LINK 2|4|AUX 3|4|SUBJ 4|0|ROOT 5|6|IN
72          9|4|PUNCT
73    *MOT: you wanna [: want to] go night_night ?
```

Research methods

CHILDES

Child Language Data Exchange System

http://childes.talkbank.org

"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.talkbank.org

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:

Use examiner-administered tests like the Peabody Picture Vocabulary Test, or other pointing task, where the child points at a picture matching the linguistic description (a word or sentence). Less cognitively-demanding for child than other tasks (e.g., an act-out task).

Research methods

Some ways to assess children's comprehension abilities:

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:

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:

elicited production: Try to get the child to produce a linguistic expression that demonstrates some linguistic knowledge of interest.

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

Research methods

Some ways to assess children's production abilities:

repetition/imitation elicitation: Children find it easier to say things that they find acceptable in their language.

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

Research methods

Some ways to assess children's production abilities:

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

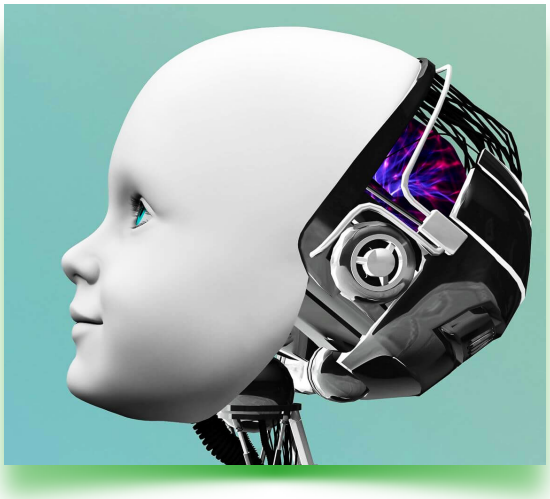
Passive example:

"…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

Some ways to assess how children might learn:
computational cognitive modeling

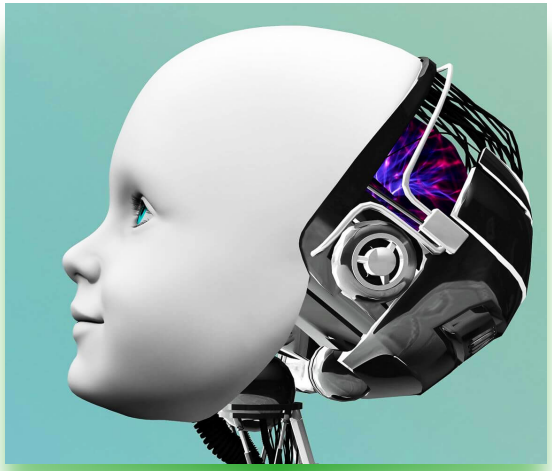


Computation of Language Laboratory

UC Irvine

Research methods

Some ways to assess how children might learn:
computational cognitive modeling



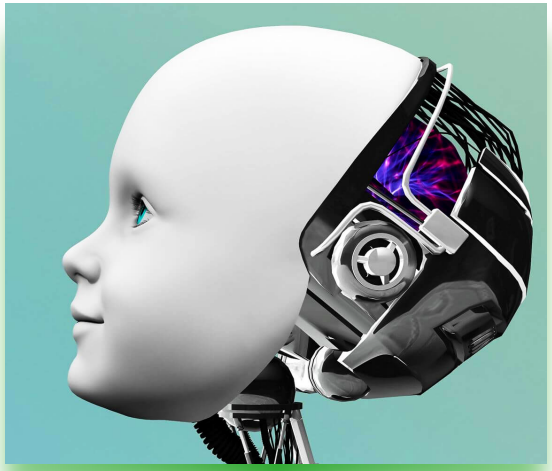
Computational cognitive modeling lets us explore theoretical ideas precisely, and evaluate how well any particular theory can explain empirical data on children's language acquisition.

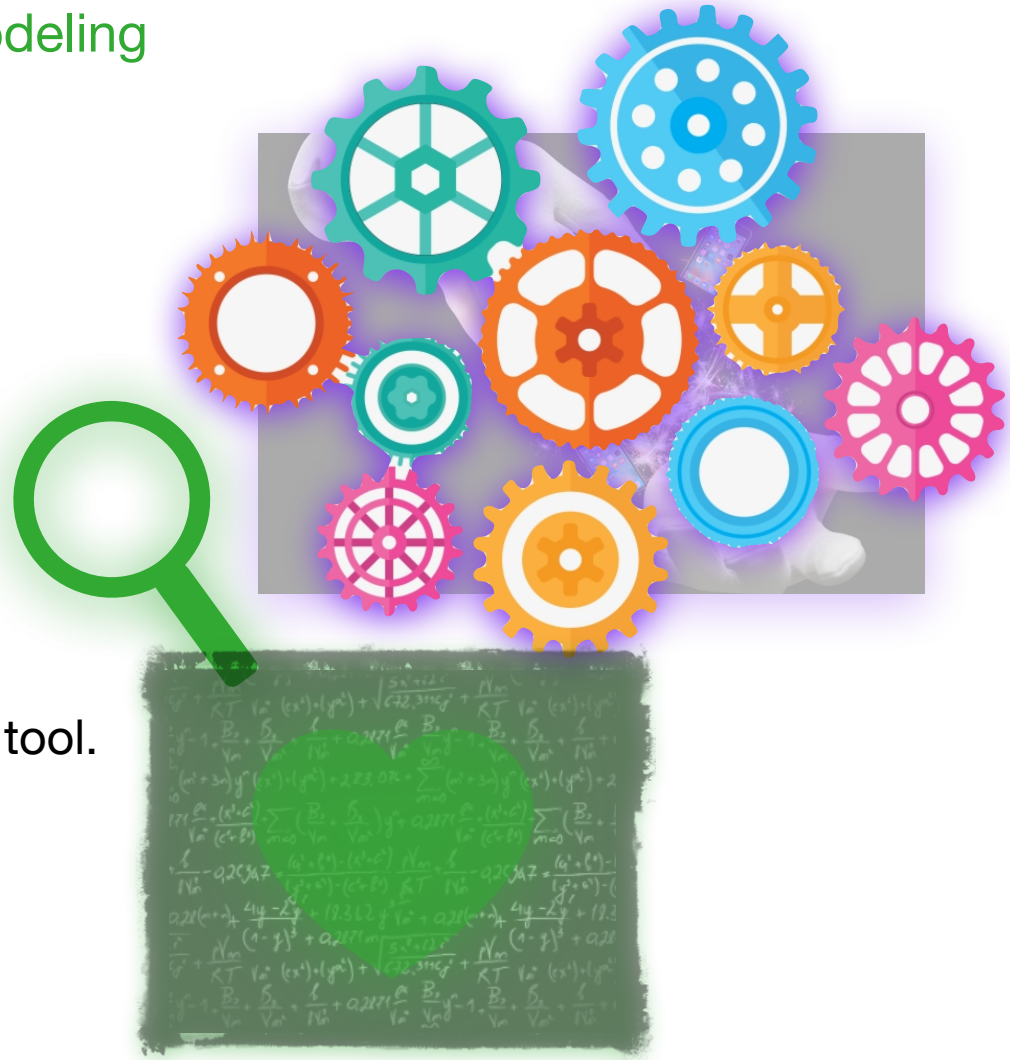
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

Some ways to assess how children might learn:

computational cognitive modeling





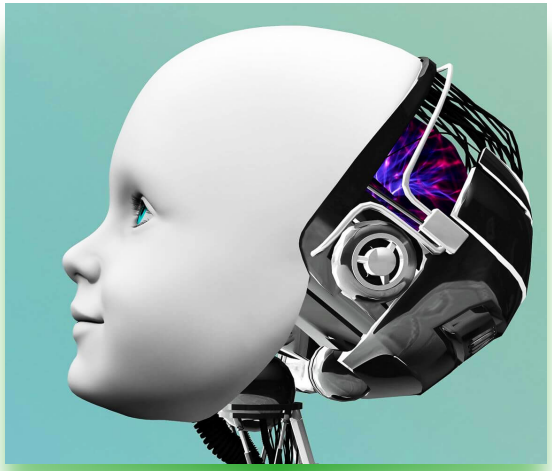
Math is at the heart of this tool.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

Some ways to assess how children might learn:

computational cognitive modeling

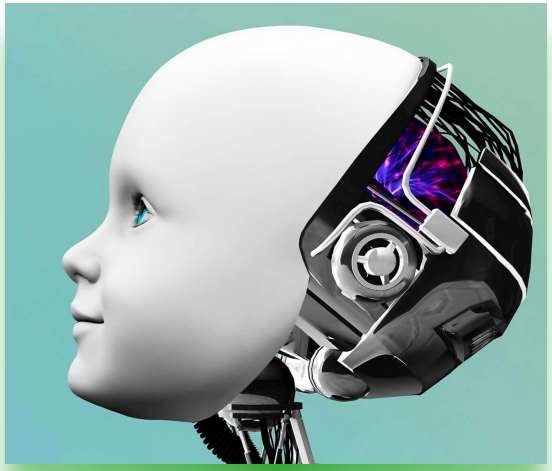


One main part: Counting things

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

Some ways to assess how children might learn:
computational cognitive modeling

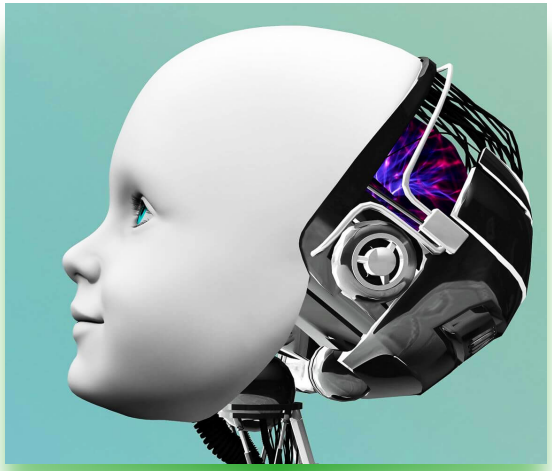


One main part: Counting things
(sometimes we count a lot of things)

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

Some ways to assess how children might learn:
computational cognitive modeling



counting

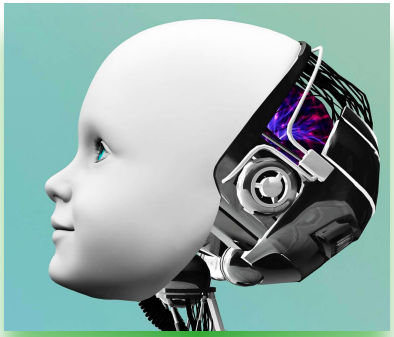
Another part: principled reasoning
based on those counts

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

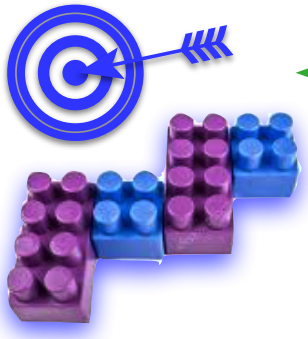
Research methods

Some ways to assess how children might learn:

computational cognitive modeling



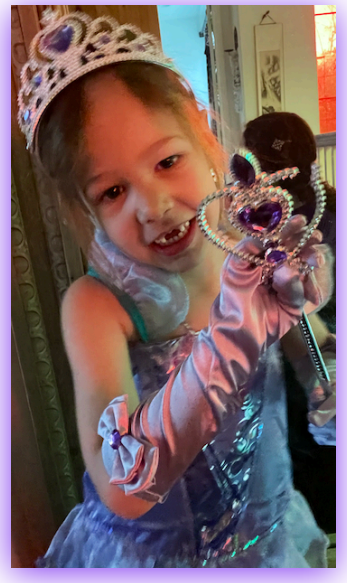
But what do we count and reason over? How do we connect that information to language acquisition?



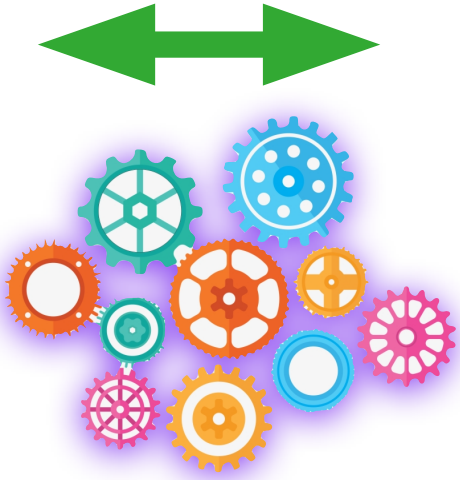
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

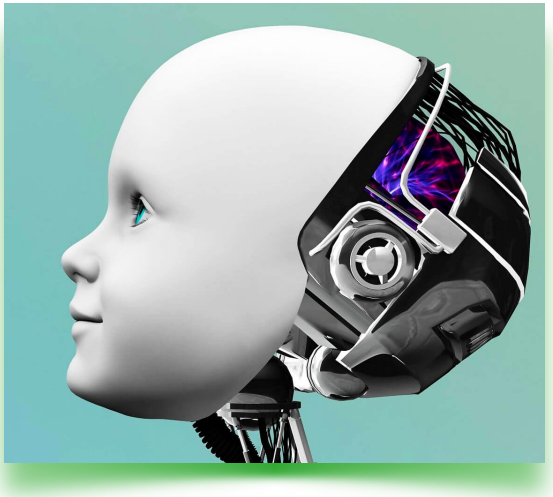
Research methods

child language acquisition



computational cognitive modeling





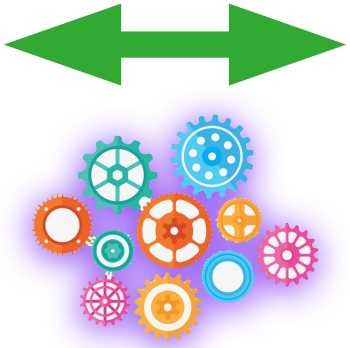
We typically using computational cognitive modeling to encode a child's acquisition process very precisely.

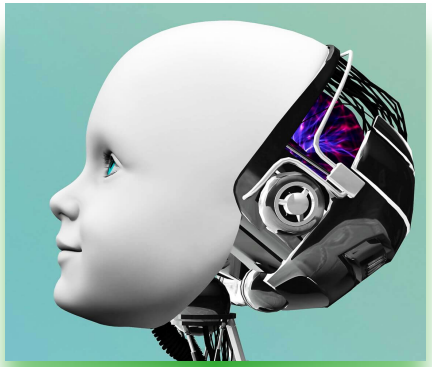
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

child language acquisition

computational cognitive modeling





We think the child is learning by counting different parts of her input and reasoning over those counts in a sensible way.

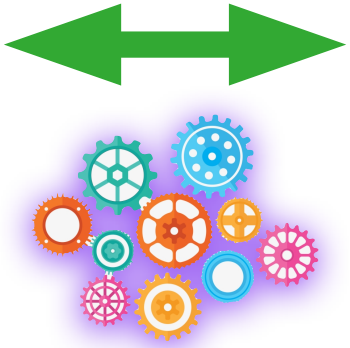
So, the modeled learner will count those same things and learn about language by doing principled reasoning over those counts.

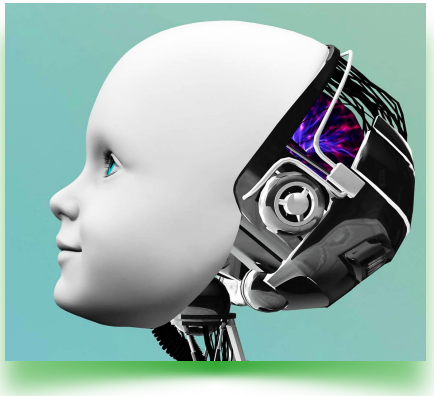
Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

child language acquisition

computational cognitive modeling





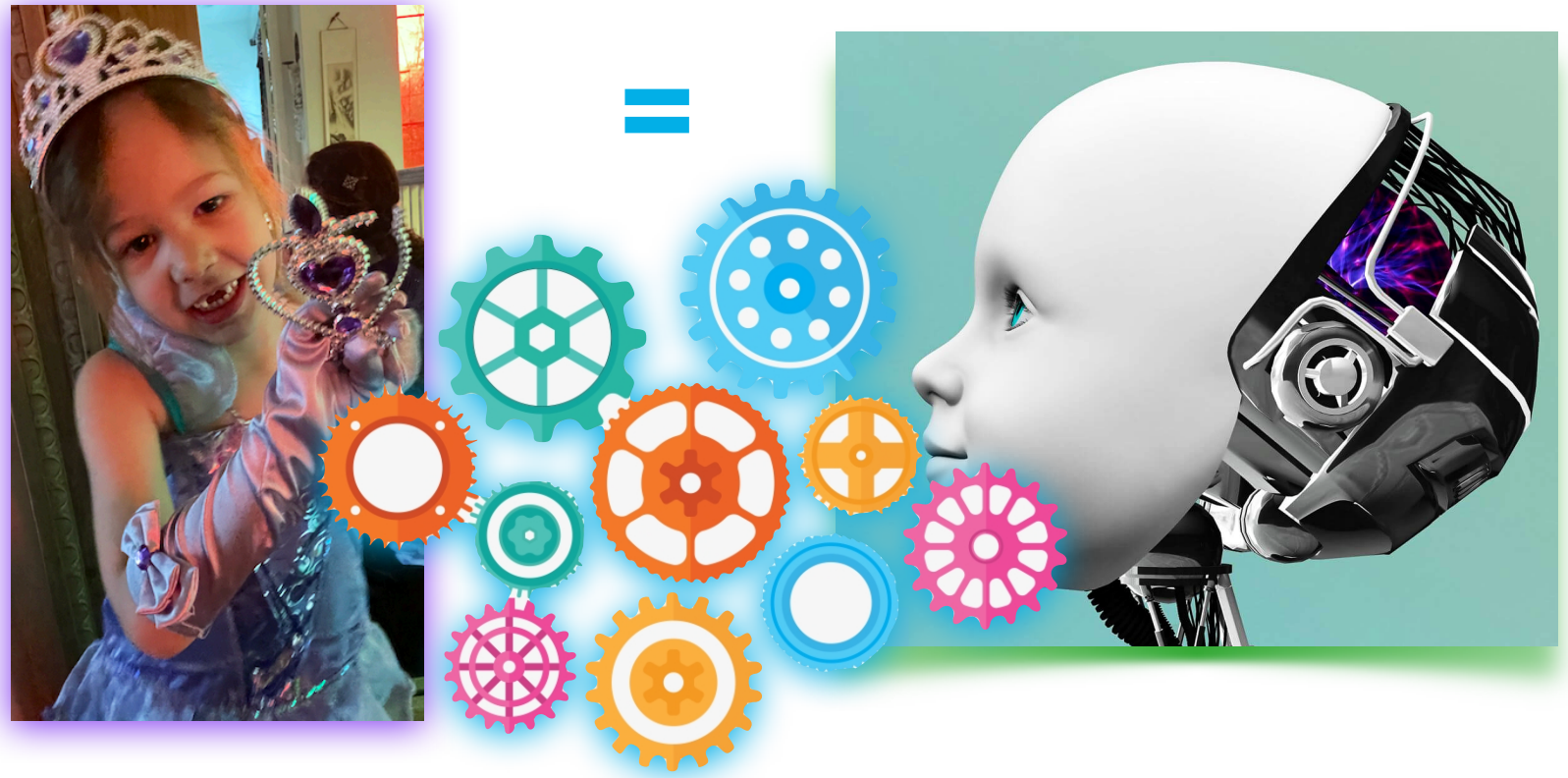
And then we see if the modeled learner behaves the way a child behaves.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Research methods

child language acquisition

computational cognitive modeling



If so, then the learning theory implemented by the modeled learner is one possible way that children could succeed at acquisition.

Pearl, 2010, 2014, Pearl and Goldwater, 2016, Pearl, 2017, 2019, 2020, 2021,a,b, 2023a,b, under review

Recap

Even though children rarely get explicit correction, they can get some help about what the correct forms are from implicit correction (like recasts).

Children's input often consists of caretaker speech, which has many properties that may aid language acquisition.

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 cognitive modeling, which tends to look at specific implementations of how the process of language acquisition could work.

Questions?

You should be able to do up through 16 on the introductory review questions and up through 8 on HW1.