LSci 51/Psych 56L: Acquisition of Language

Lecture 2
Children’s input & Research methods
Announcements

Be working on review questions for intro material

Be working on HW1
(due 10/4/19 at 2:50pm, submitted through Canvas EEE)

Please note that you can only submit HW assignments once.
Only submit when you’ve completed the entire assignment.
Last time: How do children acquire the rules of their language?

We know they do it relatively quickly.

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.”
[Extra] Explicit correction problems

https://www.youtube.com/watch?v=a7Un06tDOon0&feature=youtu.be

1:33-4:33
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.

Child: The dog runned really fast, Daddy.
Parent: Yeah, he ran really fast, didn’t he?
Recasts?

https://www.youtube.com/watch?v=a7Un06tDOon0&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. One study looking at interactions between 2-year-olds and their mothers showed that they only made recasts after 26.3% of incorrect sentences. The rest of the time, they didn’t bother.

Also, sometimes parents will repeat children’s incorrect utterances if they agree with the meaning of them! 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 very 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) found that children learned more quickly when they were given recasts.

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
"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
"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):

(1) 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
Properties of motherese (speech adults use with children):

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

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.
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
Properties of motherese (speech adults use with children):

(2) 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).

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):

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

(4) 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

More on securing children’s attention

“Screen time” interactions
About the input

More on **securing children’s attention**

“Screen time” interactions

[https://www.sciencedaily.com/releases/2016/07/160715115023.htm](https://www.sciencedaily.com/releases/2016/07/160715115023.htm)

“…children paid attention and responded to their on-screen partners, but **only children who experienced interactive video chat** responded in sync with the partner, such as clapping to imitate after the partner had clapped.”
“Screen time” interactions

https://www.sciencedaily.com/releases/2016/07/160715115023.htm

“…learning new words and patterns… occurred from video chat only when children talked to an on-screen ‘partner’ who responded to them in real time.”
“Screen time” interactions

https://www.sciencedaily.com/releases/2016/07/160715115023.htm

“....starting at about 17 months, children begin to get something out of live video interaction with real people”

- Lauren J. Myers

At 9 months, infants can learn some sound information from a screen when they learn with another 9-month-old present.

Sarah Roseberry Lytle, Adrian Garcia-Sierra, Patricia K. Kuhl. Two are better than one: Infant language learning from video improves in the presence of peers. Proceedings of the National Academy of Sciences, 2018; 115 (40): 9859 DOI: 10.1073/pnas.1611621115.
More on securing children’s attention

“Screen time” interactions

https://www.sciencedaily.com/releases/2018/10/181016132000.htm

“What this study introduces for the first time is that part of the reason we learn better when we learn collaboratively is that a social partner increases arousal, and arousal in turn increases learning. Social partners not only provide information by showing us how to do things, but also provide motivation for learning.”

- Patricia Kuhl

Sarah Roseberry Lytle, Adrian Garcia-Sierra, Patricia K. Kuhl. Two are better than one: Infant language learning from video improves in the presence of peers. Proceedings of the National Academy of Sciences, 2018; 115 (40): 9859 DOI: 10.1073/pnas.1611621115.
Properties of motherese (speech adults use with children):

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

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

“Milk.” “You want some milk?”
About the input

Properties of motherese (speech adults use with children):

(7) child-directed speech is tuned to the level of linguistic complexity the child can handle until around age five (Yurovsky, Doyle, & Frank 2016) — it’s easiest to absorb information if it’s neither too simple or too complex.
The importance of speech directed at children

Vouloumanos & Waxman (2014):
Child-directed speech scaffolds lots of knowledge

Vouloumanos: “…listening to speech promotes the babies' acquisition of the fundamental cognitive and social psychological capacities that form the foundation for subsequent learning.”

What kinds of things?

“…noticing patterns or regularities among the sounds or objects that surround them, recognizing partners with whom they can communicate, and establishing coherent categories of objects and events…”

http://www.sciencedaily.com/releases/2015/01/150105141707.htm
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)
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).
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).

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).
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.

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

http://childes.talkbank.org
Research methods

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.

Research methods

[CHILDES logo]

http://childes.talkbank.org

Example from the Braunwald corpus


```
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 ?
```
“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

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:

(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 fell 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 implements some learning theory’s ideas about how acquisition works (ex: what learning strategies children might use), and therefore test that theory empirically.

Ex: Learning to segment units in fluent speech (speech segmentation):

= wəɾəmɪɾɪkɪɾɪ
 wəɾ ĝə mpɪɾɪ kɪɾɪ

speech segmentation

what a pretty kitty!
Research methods

Computational modeling (Digital children)
An example with speech segmentation, implementing a particular learning strategy

= wλɜðriŋkɪri
wɜ ə pəri kiri

speech segmentation
what a pretty kitty!
Research methods

(1) Decide what kind of learner the model represents

For the first stages of speech segmentation:
Typically developing 6- to 8-month-old child learning first language
Research methods

Computational modeling (Digital children)

speech segmentation

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

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

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

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

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

syllables with stress

\[ w'_{\Lambda} \ r_{\alpha} \ p_{\Gamma} \ r_{i} \ k_{I} \ r_{i} \]
(4) Decide what hypotheses the child has

Example hypotheses: what the words are

\[ w' \Lambda r\emptyset p'r'i k'i r'i \]

what a pretty kitty!
(5) Decide how belief in different hypotheses is updated

Example: based on **transitional probability** between syllables
Computational modeling (Digital children)

Research methods

(5) Decide how belief in different hypotheses is updated

Example: based on transitional probability between syllables
(6) Decide what the measure of success is

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

(6) Decide what the measure of success is

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

Example developing knowledge

Proto-lexicon of word forms

$w'\Lambda r\ e\ \text{pretty k'iri}$

$w'\Lambda r\ e\ \text{a k'iri}$

$\text{what a pretty kitty!}$
(6) Decide what the measure of success is

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

Example behavior indicating developed knowledge:

- Recognizing useful units (such as words) in a fluent speech stream, as indicated by **looking time behavior**
Computational modeling (Digital children)

Research methods

speech segmentation

what a pretty kitty!

This is the heart of the model
Computational modeling (Digital children)

Research methods

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

Computational modeling (Digital children)

speech segmentation

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

Example developing knowledge
Proto-lexicon of word forms

w^l_Ar what
ə a
pr^liri pretty
k'iri kitty
Recognizing useful units (such as words) in a fluent speech stream, as indicated by looking time behavior.

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

what a pretty kitty!
(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 strategy implemented concretely in the model did.
Research methods

Computational modeling (Digital children)

speech segmentation

(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!”
Research methods

Computational modeling (Digital children)
speech segmentation

(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!”
Recap

Even though children rarely get explicit correction, they can get some help on what the correct forms are by listening to recasts in the input.

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 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 18 on the introductory review questions and up through 2 on HW1.