Language Acquisition

Ben Weissman

COGS 4780

Language Acquisition

- Our ability to learn language so efficiently seems to have several bases
- Biological
- Cognitive
- Social

- A combination of factors from all of these groups is necessary

Contributing Factors

- Imitation – children will sometimes imitate what they hear around them
- This plays a role, but cannot be leaned on to explain much of language development
  - Children will produce sentences they've never heard
  - Children will make mistakes they've never heard
  - Children only imitate phrases that have grammatical constructions they've already learned

Contributing Factors

- Reinforcement – sometimes parents/caretakers will correct the child's productions
- Again, this does happen but cannot be used to explain most of language development
  - Parents do not always correct mistakes
  - Different types of mistakes are more likely to be corrected
    - lexical > grammatical

Contributing Factors

- Social interaction plays a large role in effective language learning
- **Child directed speech** features characteristics like:
  - Exaggerated pitch contours
  - Exaggerated emphasis
  - Exaggerated word separations
  - High pitch
  - Simplified morphological complexity
  - Repetition
  - Here-and-now
  - Basic category levels

Contributing Factors

- Social interaction plays a large role in effective language learning
- Children cannot learn a first language simply by watching TV, having language in socially meaningful situations is vital for language learning
- CDS capitalizes on **joint attention** – when child and caretaker both attend to the same thing
  - In order for children to learn, they need to be paying attention to things
  - CDS helps capture their attention better than adult-directed speech

Language Acquisition

- The fundamental, central question that needs addressing in any theory of language acquisition is **is language innate**?

- Some theories are strongly built on the premise of language being innate, like Universal Grammar

- Some are built on the premise of language not being innate

→ domain specific vs domain general

Universal Grammar

- In Universal Grammar, the ability to learn language fully and efficiently is innate
- One of Chomsky's main arguments for Universal Grammar is called **poverty of the stimulus**
  - We observe that children don't just imitate everything they hear, and they don't learn entirely through reinforcement
  - The incoming speech signal is messy, with errors and disfluencies and words that all merge together into an indiscriminable string
  - This *shouldn't* be enough information for the child to actually learn the language
  - And yet they're still able to learn language in full and create new sentences
  - THEREFORE, there must be some sort of innate internal supplement

Universal Grammar

- In Universal Grammar, the ability to learn language fully and efficiently is innate
- Another is called **convergence**
  - Children (even those exposed to the same languge) all get unique and different input
  - And yet they all learn the same language
  - Explanation: Universal Grammar
- Children also do not get any negative evidence about what constructions are ungrammatical, yet they wind up knowing what constructions are ungrammatical

Universal Grammar

- In Universal Grammar, the ability to learn language fully and efficiently is innate
- Everyone is born with a **Language Acquisition Device** that helps them achieve this
  - This works by adjusting parameters – this is what allows us to get from the abstract Universal Grammar to a concrete natural language grammar
  - Exposure to a certain language (or languages) will set the parameters of the internal grammar

Universal Grammar

- Example:

- In some languages, you don't need to include the subject pronoun
  - In Spanish: "yo hablo" = "I talk" … "hablo" = "I talk"
- This parameter is called pro-drop
  - Some languages are pro-drop – you can drop the pronoun
  - Some languages are not pro-drop – you cannot drop the pronoun

- This is a parameter that will get established with the right exposure to a certain language
  - Some categories themselves are innate (like noun and verb), and the ability to use induction to set the parameters is innate (Pinker 1984)

Universal Grammar

- Universal Grammar also relies on the existence of **universals** – constraints that are general enough to be in everyone's innate universal grammar and met by every natural human language
  - e.g., every language has nouns and verbs, every language has at least 3 pronominal person levels
- Why might these exist?
  - Because they're a part of Universal Grammar

  or

  - Because they're evolutionarily beneficial
  - Because they're innate to cognition

Universal Grammar

- Some doubt the existence of universals
- Perhaps languages do vary in every way they possibly can
- If we zoom all the way out to get our "universals" are these even meaningful anymore?
  - As we account for more and more languages, the evidence for universals becomes less and less convincing
- "Even more than for categories, features tend to be proposed ad hoc in the analysis of a particular language when some formal device is needed to distinguish one structure (or operation on a particular structure) from another. As a result, supplying even a provisional list of what the set of universal distinctive syntactic features might be seems quite hopeless." (Newmeyer 2008)

Universal Grammar

- Convergence claim: despite different inputs, people end up learning the same language
  - But research has shown that adults actually have considerable individual differences in grammatical knowledge of their language (e.g., Street & Dabrowska 2010)
- Negative evidence claim: despite not receiving any negative evidence about what is ungrammatical, children wind up learning what is ungrammatical
  - But maybe children do actually get indirect evidence about what is ungrammatical, in the form of clarification questions and reformulations
  - Negative evidence can be inferred from positive evidence

Universal Grammar

- Poverty of the stimulus claim: children aren't exposed to enough information to learn language as effectively as they do
    - But maybe they actually are
    - Even children with relatively poor language input environments may get several thousand examples of each structure to be learned (Pullum & Scholz 2002)
        - UG is never clear how many are needed, but several thousand seems like a lot

    - But maybe there are other phenomena that really do seem to be rare (Boxell, 2016)

Statistical Learning

- Those who don't agree with Universal Grammar have trended towards a different theory, built on **statistical learning**
- The response to the poverty of the stimulus argument is that the stimulus isn't actually impoverished at all because children are so efficient at learning from input
- Connectionist approaches treat the mind like a vast network that is sensitive to input and tracks probabilities implicitly
  - At a young age we have a tremendous amount of **neural plasticity**, and perhaps that's what makes us so exceptional at learning language
  - We'll look more into statistical learning in the next couple of classes
- Statistical learning is a **domain-general** theory
  - It isn't something specifically built for language that lets us learn language, but rather a more general cognitive skill for (any) learning

Compatibility?

- Are statistical learning and Universal Grammar compatible with each other?

Not incompatible

- The evidence of statistical learning isn't necessarily incompatible with the claims of Universal Grammar

- It could be the case that Universal Grammar exists *alongside* domain-general learning mechanisms
  - In fact, Chomsky has recently adopted this idea (Yang et al. 2017)

- Not as successful in the other direction
  - Some who believe statistical learning is everything, and no need for UG

Language and Genetics

- **Specific Language Impairment** – a language disorder characterized by difficulties producing and comprehending spoken language
  - No other cognitive disorders, no neurological damage, etc.
- Seems to run in families, suggesting a genetic basis
- Typically observed in children, though some effects stick around into adulthood

SLI

- There are grammatical constraints on reflexive pronouns in English

*Amanda said that Becca slapped herself*

Amanda said that Becca slapped her

SLI

- There are grammatical constraints on reflexive pronouns in English

*Amanda said that **Becca** slapped **herself***

Amanda said that Becca slapped her

SLI

- There are grammatical constraints on reflexive pronouns in English

*Amanda said that [**Becca** slapped **herself**]*

Amanda said that [Becca slapped her]

SLI

- Show children one of the following sentences:

*Amanda said that [**Becca** slapped **herself**]*

Amanda said that [Becca slapped her]

And ask them to identify the matching picture
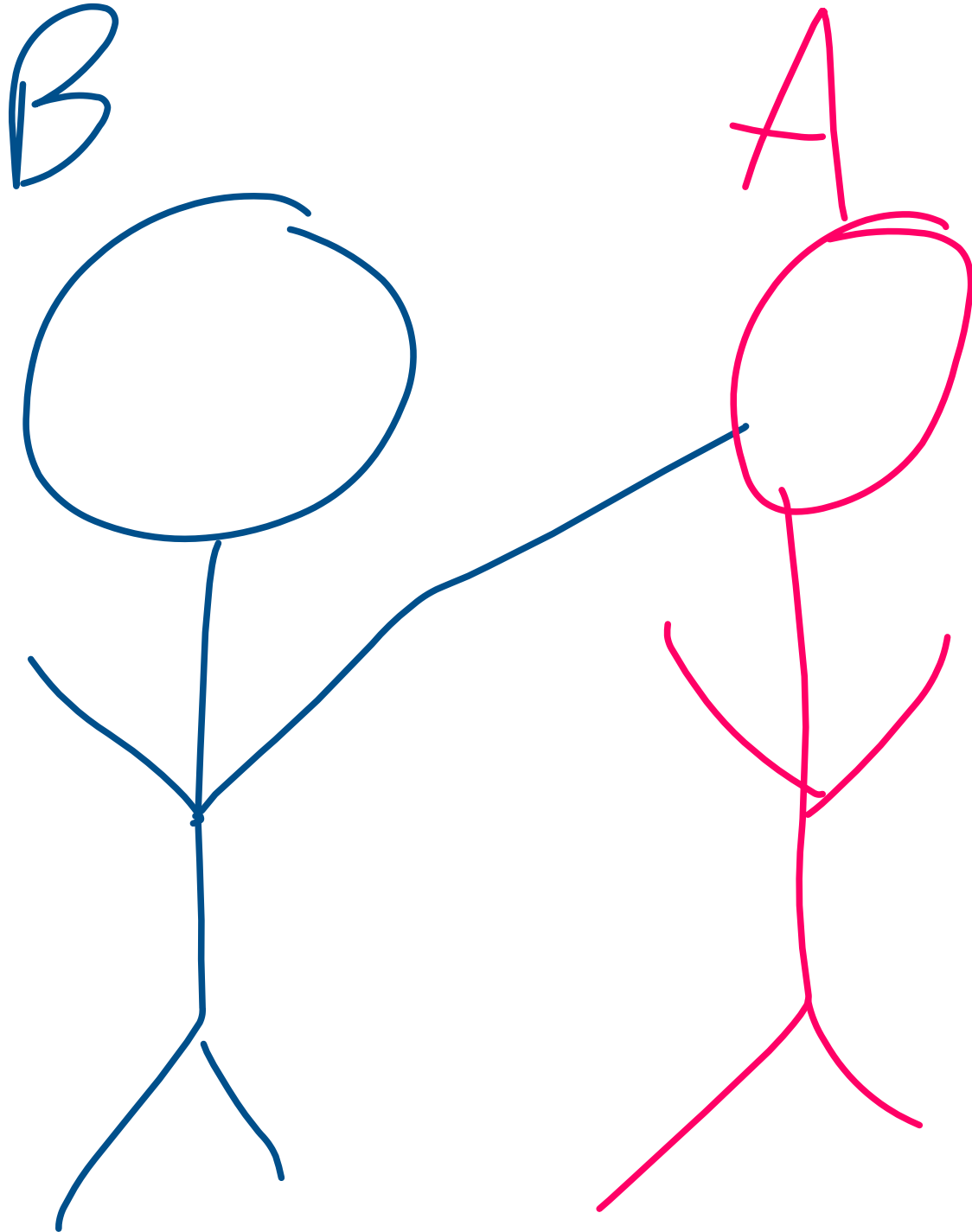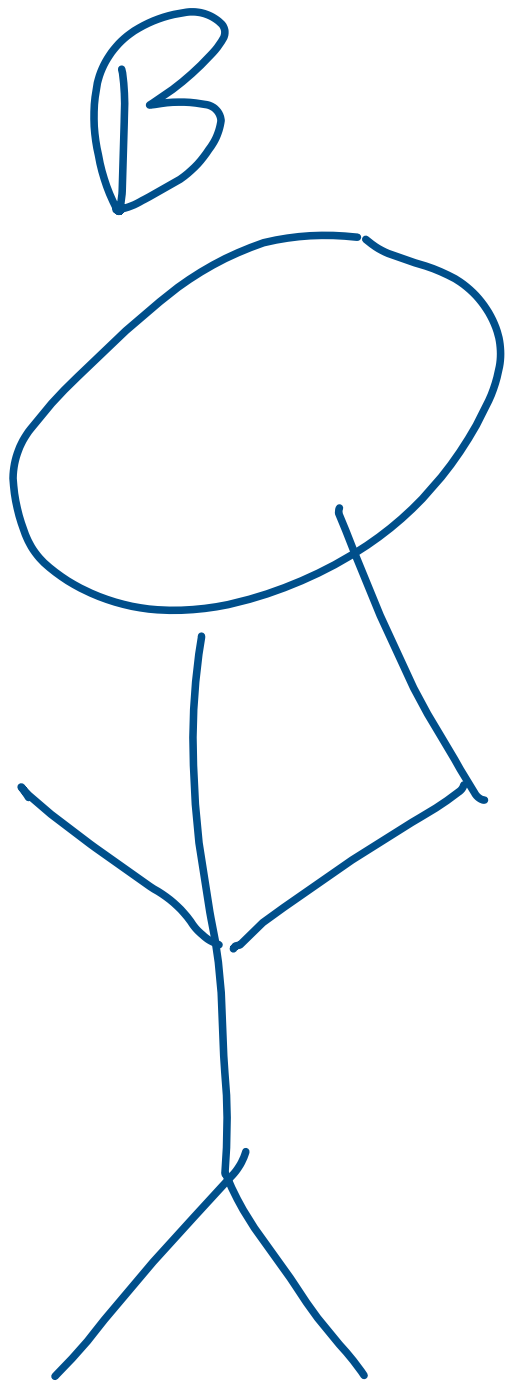
SLI

Result: children with SLI make significantly more matching errors than children without SLI

What could we conclude from this?

SLI

Result: children with SLI make significantly more matching errors than children without SLI

What could we conclude from this?

- Claim 1: children with SLI cannot represent the grammatical rules correctly
- Claim 2: children with SLI struggle with disambiguating referents
- Claim 3: children with SLI struggle in cases of increased information processing

SLI

Gabe said that [Becca slapped him]

If children with SLI can process this, maybe that's evidence that they do have the grammatical knowledge but there's interference in some scenarios (like ambiguity)

Language and Genetics

- **Williams syndrome** seems to be the opposite case – general learning disability but language functioning intact

- Comparison between adolescens with Down syndrome and Williams syndrome found similar IQ scores, similar cognitive functioning, but drastically different linguistic abilities (Bellugi et al. 2020)
  - Increased fluency, increased complexity in both production and comprehension among WMS group

Language and Genetics

- These two instances together may be taken to constitute evidence for a genetically specified language module
  - If language and general cognition seem to be independently genetically-specified, perhaps language has its own specific genetic basis
- Indeed, SLI at least has been linked to a specific gene – **FOXP2**
  - Within one UK family, there was a high prevalence of SLI and a discovered mutation in FOXP2
  - Same FOXP2 gene exists in chimpanzees, songbirds, and Neanderthals
    - Disrupting this gene in songbirds affects song development
  - Not a straightforward relationship
    - FOXP2 also affects plenty of other nonlinguistic things
    - and language comes from the interaction of many genes, not just FOXP2

Homesign

- Often deaf children will be born to hearing parents who don't know any sign language

- In these cases, children will often **homesign**, producing their own form of communication

- Homesigns have been found to contain semantics, complex ideas, morphology, maybe even recursion

Nicaraguan Sign Language

- In the late 1970s in Nicaragua, there was no existing community in which to learn a sign language

- Government attempts to bring deaf children together and have them learn Spanish, but it is largely unsuccessful

- Instead, the deaf children wound up interacting, piecing together their own individual homesigns

- This eventually turned into Nicaraguan Sign Language, which started out as a fairly piecemeal and clunky system but as newer students entered the community became more efficient

Nicaraguan Sign Language

- How might the emergence of Nicaraguan Sign Language relate to a discussion of Universal Grammar?

Critical Period

- The **Critical Period** hypothesis states that children must get the right linguistic experience during a specific time window in order to fully learn language
  - Biologically necessary to get the right input during this window or else the brain won't fully develop the right language organization
    - Typically believed to be hemisphere specialization
    - We aren't born with our hemispheres already specialized, this takes place over time
  - CPH says both hemispheres are equal at birth, undergo significant maturation between 2-5 years, fully established by the onset of puberty (Lenneberg 1967)
  - If adults experience language-hemisphere brain damage, they don't recover as well as children

Critical Period

- CPH used to explain why it's easy to learn a first language early but hard to learn a second language late
  - Learning language early, with a critical period for exposure, is innate and preprogrammed – **domain-specific**

Sensitive Period

- The decline seems to be gradual rather than discontinuous (Elman et al. 1996)
  - CPH would predict a discontinuity
- A weaker version: sensitive period hypothesis
  - Maybe there is a critical period for phonology, some elements of syntax
    - That's why late second language learners will have an accent, struggle with some elements of syntax
  - Otherwise, we see a gradual decline that correlates with age, perhaps related to neural plasticity and statistical learning