Psych229: Language Acquisition

Lecture 18

Poverty of the Stimulus & Modeling

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

The Standard Theory, according to Chomsky

Big Questions of Language Acquisition:

What constitutes knowledge of language?

How is this knowledge acquired?

How is this knowledge used?

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Knowledge of language, according to Chomsky

Knowledge of language = grammar

Grammar = complex set of rules and constraints that gives speakers intuitions that some sentences belong in the language while others do not

Competence Hypothesis: Grammar is separate from "performance factors", like dysfluencies (she said…um..wrote that), errors (I bringed it), memory capacity (The boy that the dog that the cat chased bit ran home.), and statistical properties of language (frequency of transitive (Sarah ate the peach) vs. intransitive use (Sarah ate))

"I think we are forced to conclude that…probabilistic models give no particular insight into some of the basic problems of syntactic structure" - Chomsky, 1957

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Properties of language, according to Chomsky

Grammar is generative: it can be used to produce and comprehend an infinite number of sentences

Grammar involves abstract structures: information that speakers unconsciously used is not overtly available in the observable data

Grammar is modular: there are separate components with different types of representations governed by different principles

Grammar is domain-specific: language exhibits properties not seen in other areas of cognition, so it cannot be the product of our general ability to think and learn

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Language acquisition, according to Chomsky

How does a child acquire a grammar that has those properties (generative, involving abstract structures, modular, domain-specific)?

Poverty of the stimulus problem: Available data insufficient to determine all these properties of the grammar. Therefore, children must bring innate knowledge to the language learning problem that guides them to the correct instantiation of grammar.

Available data properties leading to this inductive problem:
- noisy (degenerate): sometimes there are incorrect examples in the input
- variable: no child's input is the same as another's, but all converge
- no reliable negative evidence: no labeled examples of what's not in the language
- no positive evidence for some generalizations: yet children still converge on them

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

The induction problem, according to Chomsky

The input is too "poor": what people know extends far beyond the sample of utterances in their input

The input is too "rich": the available data can be covered by a number of generalizations, but only some of them are the right ones (yes/no questions auxiliary inversion)

Conclusion: Without innate biases, generalizations of language are unlearnable from the available data.

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

How language is used, according to Chomsky

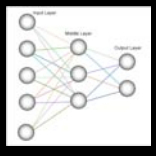
How is the grammar used to produce and comprehend utterances in real time?

Not the focus of the generative theory.

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Other developments regarding the nature of language and learning

Neural networks



Designed to solve tasks, provide input-output mapping based on data

Learning: gradual changes to the weights between units in the network that determine patterns of activation

Parameters: learning rule that adjusts weights, network structure

Not a grammar

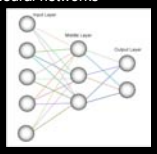
Grammar = higher level generalization about network behavior, abstracts away from actual implementation

Grammar = computational level, network = algorithmic + implementational level

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Other developments regarding the nature of language and learning

Neural networks



Property: Can derive structural regularities from relatively noisy input. (This comes from the gradual learning capability.) Realistic learning input.

Property: A network that has learned can then process novel forms. It has generative capacity. (Ex: word pronunciation)

Implication: Poverty of the stimulus may not be the induction problem originally thought?

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

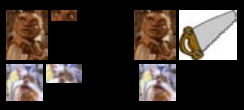
Other developments regarding the nature of language and learning

Data resources: corpora of adult and child-directed speech
Accurate estimation of the data available.

Psycholinguistic resource: sentence processing
Statistical properties of language influence ease/difficulty of processing, especially when there is ambiguity.

Ambiguity

We saw



Less probable

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

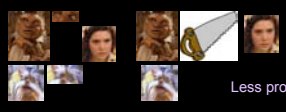
Other developments regarding the nature of language and learning

Data resources: corpora of adult and child-directed speech
Accurate estimation of the data available.

Psycholinguistic resource: sentence processing
Statistical properties of language influence ease/difficulty of processing, especially when there is ambiguity.

Ambiguity

We saw her



Less probable

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Other developments regarding the nature of language and learning

Data resources: corpora of adult and child-directed speech
Accurate estimation of the data available.

Psycholinguistic resource: sentence processing
Statistical properties of language influence ease/difficulty of processing, especially when there is ambiguity.

Ambiguity

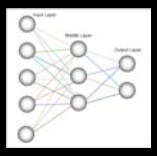
We saw her duck

Less probable

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Other developments regarding the nature of language and learning

Seidenberg's point: Statistical properties determine language use and neural nets provide a way to explicitly encode, acquire, and exploit this information.



Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Other developments regarding the nature of language and learning

Children can encode statistical properties of language (Jusczyk 1997 = properties of sounds, Saffran et al. 1996 = transitional probabilities of syllables)

Seidenberg's point: Acquisition is about learning to use the language, which means paying attention to its statistical properties and learning from them.

Another point: Connectionist networks formalize the implementation of bootstrapping - extracting regularity from the data (used for word segmentation, word meaning, grammatical category, syntactic constructions)

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Big point of Seidenberg:

"…[connectionism] attempts to explain language in terms of how is it acquired and used rather than an idealized competence grammar. The idea is not merely that competence grammar needs to incorporate statistical and probabilistic information; rather it is that the nature of language is determined by how it is acquired and used and therefore needs to be explained in terms of these functions and the brain mechanisms that support them. Such performance theories are not merely the competence theory plus some additional assumptions about acquisition and processing; the approaches begin with different goals and end up with different explanations for why languages have the properties they have."

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Connectionism in Action: An example where it could help
Correlations between verb meaning and verb usage

Hoggle loaded jewels into his bag.
Hoggle loaded his bag with jewels.

Hoggle poured jewels into his bag.
*Hoggle poured his bag with jewels.

*Hoggle filled the jewels into his bag.
Hoggle filled his bag with jewels.

Input is irregular - children do not get explicit examples of all of these, yet somehow come to know this paradigm.

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Clue
clusters of verbs with similar properties (if children realize this, learning is easier)
load, pile, cram, spray, scatter
pour, drip, slop, slosh
fill, blanket, cover, coat

Problem: How would the child know to cluster these verbs together if they never hear all the verbs in all the possible syntactic frames? Semantically, they're very similar.

However…

This is a constraint satisfaction problem, which neural nets are really good at solving.

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

Information available on groupings

load, pile, cram, spray, scatter
pour, drip, slop, slosh
fill, blanket, cover, coat

1) How much the semantics of each verb overlaps with any other verb
2) Correlations between syntactic frames verbs appear in and the exact semantics of the verb
3) Item-specific idiosyncracies (due to language change)

Connectionist net can learn the right subgroups (Allen 1997) from this information

…and then much easier to notice that there are syntactic usage generalizations for the groups. Therefore, this can be learned. Which is good, since it's a language-specific property.

Seidenberg (1997): Innate Biases ≠ Grammatical Knowledge

But what about learning more abstract things (like syntax) and language-independent things that are hard (or impossible) to observe?

…future work for connectionist models.

And innate knowledge?

"Innate capacities may take the form of biases or sensitivities toward particular types of information inherent in environmental events such as language, rather than a priori knowledge of grammar itself."

"Brain organization therefore constrains how language is learned, but the principles that govern the acquisition, representation, and use of language are not specific to this type of knowledge"