Psych 215L: Language Acquisition

Lecture 2
The Language Learning Mechanism

Children's Language Learning

Stage 1 (first few months): “cooing” vocalization
“goo goo ga ga”

Stage 2 (~6 months): “babbling”
strings of syllables using a wide range of sounds
(some sounds aren’t even those used in native language)
general consensus: baby playing with the vocal tract

deaf babies do it (in the absence of auditory input)
dead babies exposed to sign language babble with their hands, too
not all babies do it, though:

after a few months, babbling takes on intonation patterns of native language

Stage 3 (~10-20 months): single word utterances
“Mommy!” “Juice!” “Up!”
(surprisingly communicative)
Within 6 months: child’s vocabulary grows to ~50-100 words

Stage 4 (~24 months): two word utterances
“Mommy sock” “Drink soup” “No eat”
Consistent use of word order, even though not all words are used

“Mommy should throw the ball”
Mommy throw
Not “throw Mommy”

“Throw ball”
Throw ball
Not “ball throw”

Stage 4 continued (~24 months): vocabulary spurt
Parents can’t keep track of all the words their child knows

(estimate: 10,000 words by 5 years old)
This works out to about 1 per waking hour from ages 2 to 5 years old! (Child likely working on multiple words at once, too.)

Stage 5 (~30 months): grammatical growth
Child constructs longer and more grammatically complex sentences

(If age 5. Very good approximation of adult word order rules, though there are still some wrinkles to be worked out)

Knowing more than they say

Phonology (sound system):
Children often simplify the sounds of words.
Ex: “spoon” becomes “poon”
“bus” becomes “buh”
“duck” becomes “guck”
“truck” becomes “guck”

But children comprehend more than they can produce sound-wise. A child who says “guck” for “duck” and “truck” will have no problem distinguishing ducks from trucks when asked.

“If you deliberately pronounce a word the way your child does, he or she will get mad at you and tell you to say it right. If you tell your child to say ‘duck’, not ‘guck’, most of the time you’ll get ‘guck’ and a blank stare.” - Jackendoff (1994)
Knowing more than they say

Syntax (word order system):
Can test children who are in the 1-word stage on their understanding of word order rules (which involve more than 1 word).
(Hirsh-Pasek & Golinkoff: 17-month-olds)

“Big Bird is tickling Cookie Monster.”

Knowing more than they say

“Cookie Monster is tickling Big Bird.”

Knowing more than they say

Word categorization knowledge: from word patterns
1-word stage children (17 months old)
“This is DAX.”

“Could you give me DAX?”

Knowing more than they say

Word categorization knowledge: from word patterns
1-word stage children (17 months old)
“This is a DAX.”
Knowing more than they say
Word categorization knowledge: from word patterns
1-word stage children (17 months old)
"This is a DAX."
"Could you give me a DAX?"

Getting to children’s knowledge
Using novel test items (since children will not have heard these before)
"This is a wug."
"Now there is another one. There are two..."
Kids don’t seem to have this figured out till about age 6.

Getting to children’s knowledge
Observe patterns of mistakes
From Edward Klima & Ursula Bellugi
Wh-questions
Stage 1
What book name?
Why you smiling?
What soldier marching?
Stage 2
What he can ride in?
Which way they should go?
Why kitty can’t stand up?
Stage 3
Where will you go?
Why can’t kitty see?
Why don’t you know?

Getting to children’s knowledge
Observe patterns of mistakes
From Edward Klima & Ursula Bellugi
Use of negative elements (not, n’t)
Stage 1
No the sun shining.
No a boy bed.
No sit there.
Stage 2
He no bite you.
I no want envelope.
I no taste them.
Stage 3
I didn’t did it.
You didn’t caught me.

Getting to children’s knowledge
Observe patterns of mistakes
From Edward Klima & Ursula Bellugi
Use of past tense verbs (U-shaped curve of performance)
Stage 1
walked
played
came
went
Stage 2
walked
played
came
gored
held
Stage 3
walked
played
came
wented
Stage 4
walked
played
came
went
held

Getting to children’s knowledge
Observe patterns of mistakes
From Edward Klima & Ursula Bellugi
Use of past tense verbs (U-shaped curve of performance)
Stage 1
walked
played
came
went
Stage 2
walked
played
came
holded
Stage 3
walked
played
came
wented
Stage 4
walked
played
came
went
held
Main points

Children understand more than they can imitate. (Comprehension greater than production)

Children don’t just imitate what they’ve heard - they’re trying to figure out the patterns of their native language.

The patterns they produce during learning are often stripped-down versions of the adult pattern, but they make mistakes that cannot be attributed directly to the input.

Levels of Representation
Marr (1982)

“…it gradually became clear that something important was missing that was not present in either of the disciplines of neurophysiology or psychophysics. The key observation is that neurophysiology and psychophysics have as their business to describe the behavior of cells or of subjects but not to explain such behavior….What are the problems in doing it that need explaining, and what level of description should such explanations be sought?” - Marr (1982)

On Explaining (Marr 1982)

“…[need] a clear understanding of what is to be computed, how it is to be done, the physical assumptions on which the method is based, and some kind of analysis of the algorithms that are capable of carrying it out.”

“This was what was missing - the analysis of the problem as an information-processing task. Such analysis does not usurp an understanding at the other levels - of neurons or of computer programs - but it is a necessary complement to them, since without it there can be no real understanding of the function of all those neurons.”

On Explaining (Marr 1982)

“…But the important point is that if the notion of different types of understanding is taken very seriously, it allows the study of the information-processing basis of perception to be made rigorous. It becomes possible, by separating explanations into different levels, to make explicit statements about what is being computed and why and to construct theories stating that what is being computed is optimal in some sense or is guaranteed to function correctly. The ad hoc element is removed…”

Our goal: Substitute “language acquisition” for “perception”.

On Explaining (Marr 1982)

“…But the important point is that if the notion of different types of understanding is taken very seriously, it allows the study of the information-processing basis of perception to be made rigorous. It becomes possible, by separating explanations into different levels, to make explicit statements about what is being computed and why and to construct theories stating that what is being computed is optimal in some sense or is guaranteed to function correctly. The ad hoc element is removed…”

Our goal: Substitute “language acquisition” for “perception”.

Describing vs. Explaining

“…it gradually became clear that something important was missing that was not present in either of the disciplines of neurophysiology or psychophysics. The key observation is that neurophysiology and psychophysics have as their business to describe the behavior of cells or of subjects but not to explain such behavior….What are the problems in doing it that need explaining, and what level of description should such explanations be sought?” - Marr (1982)
The three levels

Computational

What is the goal of the computation? What is the logic of the strategy by which it can be carried out?

Algorithmic

How can this computational theory be implemented? What is the representation for the input and output, and what is the algorithm for the transformation?

Implementational

How can the representation and algorithm be realized physically?

The three levels:

An example with the cash register

Computational

What does this device do? Arithmetic.

Task: Master theory of addition.

Algorithmic (Addition)

Addition: Mapping of a pair of numbers to another number.

\[(3,4) \rightarrow 7\] (often written \((3+4=7)\))

Properties: \((3+4) = (4+3)\) [commutative], \((3+4)+5 = 3+(4+5)\) [associative], \((3+0) = 3\) [identity element], \((3+ -3) = 0\) [inverse element]

True no matter how numbers are represented: this is what is being computed.

Implementational

How does cash register implement this? A series of mechanical and electronic components.

The three levels

Marr (1982)

“Although algorithms and mechanisms are empirically more accessible, it is the top level, the level of computational theory, which is critically important from an information-processing point of view. The reason for this is that the nature of the computations that underlie perception depends more upon the computational problems that have to be solved than upon the particular hardware in which their solutions are implemented. To phrase the matter another way, an algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and the hardware) in which it is embodied.”

Mapping the Framework:

Algorithmic Theory of Language Learning

Goal: Understanding the “how” of language learning

First, we need a computational-level description of the learning problem.

Computational Problem: Divide sounds into contrastive categories

Diagram: 

```
C1 -- C2
  |  |
  |  |
  C4  C3
```
First, we need a computational-level description of the learning problem.

Computational Problem: Divide spoken speech into words

who’s afraid of the big bad wolf

Computational Problem: Map word forms to speaker-invariant forms

friends

Computational Problem: Identify grammatical categories

"This is a DAX."

DAX = noun

Computational Problem: Identifying the rules of word order for sentences.

Jareth juggles crystals

Kannada

Input = sounds, syllables, words, phrases, ...

Output = sound categories, words, words with affixes, grammatical categories, ...

Process the can take us from input to output: statistical learning, algebraic learning, ...

Considerations: input available to child, psychological plausibility of learning algorithm, hypotheses child considers
Framework for language learning (algorithmic-level)

What are the hypotheses available (for generating the output from the input)?

Ex: general word order patterns

Input: words (adjective and noun)
Output: ordered pair

- Adjective before noun (ex: English)
  red apple

- Noun before adjective (ex: Spanish)
  manzana roja
  apple red

Framework for language learning (algorithmic-level)

What are the hypotheses available (for generating the output from the input)?
Ex: general word order patterns

What data are available, and should the learner use all of them?
Ex: exceptions to general word order patterns

Ignore special use of adjective before noun in Spanish
Special use: If the adjective is naturally associated with the noun:
la blanca nieve
the white snow

Why not usual order? Snow is naturally white

Framework for language learning (algorithmic-level)

What are the hypotheses available (for generating the output from the input)?
Ex: general word order patterns

What data are available, and should the learner use all of them?
Ex: exceptions to general word order patterns

How will the learner update beliefs in the competing hypotheses?
Ex: shifting belief in what the regular word order of adjectives and nouns should be

This usually will involve some kind of probabilistic updating function.