Learning Words

Chapter 5 (pp. 145-158 + 168-180)
C & C (pp. 37-43)

Dr. Morten H. Christiansen

© Morten H. Christiansen 2017
Previous Class:
Learning to Perceive Speech

- Categorical Perception
  - A perceptual discontinuity across a continuously varying physical dimension
  - VOT: /ba/ vs. /pa/

- Phonological Development
  - Basic Phoneme Distinctions: Innate universal speech sound discrimination capacity
    - not species-specific.
  - Perceiving Sounds in Context: Speaker normalization, vowel prototype representation.
Lexical Development (I)

- Children start producing their first words around 12 months.
- Words are used holophrastically: A word stands for an entire sentence.
- By 24 months they have an expressive vocabulary of between 50 to 600 words.
- Experience matters for vocabulary growth.
- Privileged children hear about 2,100 words/hour.
- Disadvantaged children hear only about 600 words/hour.
Lexical Development (II)

- Children often assign different meanings to words compared with adults:
  - Underextension: Narrow assignment of word meaning
    - *cat* → a particular cat (*Tigger*) and not other cats.
  - Overextension: Broad assignment of word meaning
    - *dog* → all four-legged creatures.
Constraints on Lexical Meaning Assignment (I)

- The indeterminacy of reference (Quine, 1960):
  - For any finite data set there an infinite number of possible interpretations.
  - \( gavagai = \text{rabbit, hopping, long ears, rabbit holes} \).
- Potential constraints on meaning assignment:
  - Whole Object Bias: New words label whole objects (not object parts).
    - \( gavagai \rightarrow \text{rabbit} \) (not rabbit parts).
  - Taxonomic Bias: New words refer to categories of objects (not thematic relations).
    - \( gavagai \rightarrow \text{rabbit} \) (not rabbit holes).
Constraints on Lexical Meaning Assignment (II)

- Potential constraints on meaning assignment (cont.):
  - Mutual Exclusivity Bias: Objects only have a single label (new words label unlabeled objects).
    - gavagai → rabbit (no label)
  - These biases reflect tendencies in the way children assign meaning.

- Syntactic Bootstrapping Hypothesis:
  - Children use syntax to infer word meaning.
    - “Here’s someone who is nazzing”
    - “nazzing” is a verb → action reference.
Morphology: Knowledge of Word Structure (I)

- Morpheme:
  - The smallest meaningful unit in a language.
    - Free Morphemes can stand alone as words.
    - Bound Morphemes must combine with other morphemes to form a word.

- Inflectional Morphology:
  - Constraints on changing a word given its grammatical role.
    - N plural: stem + -s.
      - cat + -s = cats
Morphology: Knowledge of Word Structure (II)

Derivational Morphology:

Constraints on the combination of morphemes to create new words, new meaning or grammatical category:

- re- [prefix] + live [verb] = relive
- eat [verb] + -er [suffix] = eater [noun]
How to Study Language Acquisition?

- Methods for Studying Expressive Language Ability:
  - Diary Studies: Sample recordings of children's speech
    - → CHILDES corpus.
  - Useful for corpus analyses and computational modeling
  - Elicited Production: Elicit productions in response to stimuli to test for sensitivity to specific grammatical regularities.
The Acquisition of English Past Tense

- **U-shaped Learning:**
  - Early Correct Stage: Few verbs used, mostly high-frequency irregular verbs, mostly correct usage
    - $\text{come}_{\text{irregular}} \rightarrow \text{came}$ - $\text{walk}_{\text{regular}} + \text{-ed} \rightarrow \text{walked}$.
  - Overgeneralization Stage: Large number of verbs, many regular verbs, overgeneralization of irregular verbs
    - $\text{come}_{\text{irregular}} + \text{-ed} \rightarrow \text{comed}$ - $\text{walk}_{\text{regular}} + \text{-ed} \rightarrow \text{walked}$.
  - Late (Adult) Correct Stage: Correct performance on both regular and irregular verbs.
    - $\text{come}_{\text{irregular}} \rightarrow \text{came}$ - $\text{walk}_{\text{regular}} + \text{-ed} \rightarrow \text{walked}$. 
Children acquire an -\textit{ed} rule to generate the past tense of regular verbs

\[ \text{walk}_{\text{regular}} + \text{-ed} \rightarrow \text{walked}. \]

Past tenses of irregular verbs are stored in the lexicon.

The Blocking Principle:

Block the application of the -\textit{ed} rule if another past tense form of the word is in the lexicon.

Overgeneralization occurs because irregular words are not strongly stored in memory.

Problem: No account of frequency effects.
The Dual-Route Account of Past Tense Generation

Irregular Route

Lexicon

Past Tense of Verb

Verb Stem

Regular Route

Past Tense Rule (+ -ed)
Children acquire associations between verb roots and their past tenses.

Account of U-Shaped Learning:

- Early Correct Period:
  - Rote memorization of verb root → past tense combinations.

- Overgeneralization:
  - The acquisition of a large number of mainly regular verbs forces the building of a “system” initially biased towards the regular past tense.

- Late Correct Period:
  - Irregular past tenses are reinforced through repeated experience with them.
The Rumelhart & McClelland (1986) Past Tense Model
Rumelhart & McClelland (1986): Training Issues (I)

- Training Regime Problem:
  - Stage 1: Early correct period.
    - 10 high-frequency verbs
      - 8 irregulars + 2 regulars.
  - Stage 2: Overgeneralization.
    - Early in training with 420 new medium-frequency verbs added to the vocabulary
      - 80% regulars.
  - Stage 3: Late correct period.
    - Late in training with the 430 verbs.
Rumelhart & McClelland (1986): Training Issues (II)

- Problem: U-shaped learning pattern created by the sudden increase in number of verbs (Pinker & Prince, 1988).

- Solution: U-shaped learning obtained both with training on
  - a constant vocabulary of 500 verbs (Plunkett & Marchmann, 1991)
  - and with a gradual increase in vocabulary size (20-500 verbs) (Plunkett & Marchmann, 1993).
# English Inflections

**The Appearance of English Inflections** (Belugi, 1964)

<table>
<thead>
<tr>
<th>Inflection</th>
<th>Adam</th>
<th>Eve</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present progressive, -ing</td>
<td>28</td>
<td>19.5</td>
</tr>
<tr>
<td>Plural on nouns, -s</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>Past on regular verbs, -ed</td>
<td>39</td>
<td>24.5</td>
</tr>
<tr>
<td>Possessive on nouns, -s</td>
<td>39.5</td>
<td>25.5</td>
</tr>
<tr>
<td>Third person on verbs, -s</td>
<td>41</td>
<td>26</td>
</tr>
</tbody>
</table>
Next Class
Learning to Combine Words
Chapter 6 (pp. 185-194)

• Milestones of acquisition
• Syntactic development
  • MLU growth
• Cross-linguistic differences in speed of acquisition