Language Acquisition

Michael Frank

9.012
Why study language acquisition?

• Language constitutes a major difference between ourselves and other primates
• Looking at children may give us insight into adult performance
• Learning about language may tell us about the nature of mental representation (c.f. Fodor, 1971)
• Language acquisition can settle philosophical issues like rationalism vs. empiricism (c.f. Pinker, 1994).
• Universals of language may tell us about the structure of the mind (c.f. Chomsky, 1965)
Breaking down language into pieces

• We will be talking about:
  – **Word learning**
  – **Inflectional morphology** (pieces of words)
  – **Syntax** (word order) - Ted

• We won’t be talking (much) about:
  – Intonation and prosody
  – Phonology
    • Veronica will talk about this
    • Speech production is a whole different topic
  – Gesture
  – Pragmatics (social use of language)
Outline

• Word-learning
  – The phenomena
  – Proposed constraints on word learning
  – Abilities for word learning

• Acquisition of morphology
  – The phenomena
  – Theoretical Positions
    • Words & Rules
    • Single-route theories
  – Novel approaches
Müller (1864): “The one great barrier between man and brute is *Language*. Man speaks, and no brute has ever uttered a word. Language is our Rubicon, and no brute will dare to cross it.”

Morgan’s Canon (1894): “In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale.”
Timeline for language learning

- First words usually produced between 10 and 14 months
- 8mo: avg. comp. of 15 words
- 18mo: avg. prod. of 50 words, proposed beginning of a "word spurt"
- 24-30mo "telegraphic utterances" composed of two word combinations
- 18-24mo "holophrases" single word utterances that have communicative content
- 30mo+: multi-word utterances with complex content
- 30mo: avg. prod. vocabulary of 500 words
- Eventual vocabulary of more than 50,000 words
Word Learning: Comprehension

Figure removed due to copyright restrictions.
Please see:
Word Learning: Production

Figures removed due to copyright restrictions.
Please see:
Word Learning: Rate of acquisition

Words (comprehension) learned per day

From Bloom (2000), estimates to 30mo from Fenson (1994), above 30mo from Anglin (1993)
The composition of early vocabulary

American children

Figure removed for copyright restrictions.
Please see:
Figure from Caselli, M. C., P. Casadio, and E. Bates. "A comparison of the transition from first words to grammar in English and Italian." Journal of Child Language 26 (1999): 69-111.

Italian children

Figure removed for copyright restrictions.
Please see:
Figure from Caselli, M. C., P. Casadio, and E. Bates. "A comparison of the transition from first words to grammar in English and Italian." Journal of Child Language 26 (1999): 69-111.
Why the noun bias?

Figure removed due to copyright restrictions.
Please see:
Figure from Gentner, D., and L. Boroditsky.
"Individuation, relativity and early word learning." In
Language acquisition and conceptual development. Edited
by M. Bowerman and S. Levinson. Cambridge, UK: Cambridge
University Press, pp. 215-256.

Figure removed due to copyright restrictions.
Please see:
Figure 8.1, from Gentner, D., and L. Boroditsky.
"Individuation, relativity and early word learning." In
Language acquisition and conceptual development. Edited
by M. Bowerman and S. Levinson. Cambridge, UK: Cambridge
University Press, pp. 215-256.
A different account

Figure removed due to copyright restrictions.
Please see:
And you thought verbs were hard?

Rabbit?

Or group of undetached rabbit parts?

How do you figure out the referent even for a simple object noun?

Solving the paradox with constraints

- Motivated by the “word spurt”
  - Change in learning rate potentially caused by the application of constraints (not supported by current evidence)
- The **whole-object assumption**
  - Words refer to objects, not parts
- The **taxonomic constraint**
  - Words refer to objects of same kind, not having same theme
  - e.g., *dax* = dog, kids choose cat as another *dax*, not dogfood (Markman & Hutchinson, 1984)
- The **mutual exclusivity assumption**
  - Each object has only one name
  - e.g., *dax* rejected as name of object with known name, used for a part instead.
- Also: **basic-level** and **equal-detail assumptions**

Solving the paradox with principles

1. **Conventionality**: For certain meanings, speakers assume that there is a conventional form that should be used in the language community.

2. **Contrast**: Speakers assume that any difference in form signals a difference in meaning.

Note that these do roughly the same things as Markman’s constraints, but in a domain-general, pragmatic way rather than a domain-specific, modular way.

Word learning as Bayesian inference

Figure removed due to copyright restrictions.
Please see:
Figure from Tenenbaum, J., and F. Xu. "Word learning as Bayesian inference." Proceedings of the 22nd Annual Conference of the Cognitive Science Society.
Shape bias as a tool for word learning

Figure removed due to copyright restrictions.
Please see:
Shape bias as a tool for word learning

Vocabulary growth

Training items

Figure removed due to copyright restrictions.
Please see:
Abilities for word learning: “fast mapping”

3 & 4 year olds learn words after a single exposure

Fast mapping: Domain-general ability?

This one is a *koba*!

My uncle gave me this one!

Here’s a sticker

Figure removed due to copyright restrictions.
Please see:
Two conditions

1. Coincide: name matches toy shown

2. Conflict: name does not match toy shown

Figure removed due to copyright restrictions.
Please see:
Figure from D. A. Baldwin. "Early referential understanding: Infants’ ability to recognize referential acts for what they are." *Developmental Psychology* 29, no. 5 (1993): 832-843.
Discourse cues for word learning

• Experimenter presents novel action and novel object
• Action is performed multiple times with different objects (novel object) OR object performs multiple actions (novel action), experimenter says “Modi!”
• Test: “Can you show me modi!”

Table removed for copyright restrictions.
Please see:
Syntactic cues for word learning

- Brown (1957)
  - Showed kneading confetti in a bowl and asked: ‘do you know what it means to sib?’, ‘do you know what a sib is?’, or ‘have you seen any sib?’
  - Pick ‘sibbing,’ ‘a sib,’ or ‘sib’ from an array depicting several actions, substances, and containers
  - Children identified ‘sibbing’ as a depiction of kneading, ‘a sib’ as a picture of the bowl, and ‘sib’ as a picture with confetti

- Soja, Carey, & Spelke, (1991)
  - Two year olds extended “a bicket” to objects of the same shape, extended “some bicket” to portions of the same substance, regardless of shape
Recap: Word learning

• Phenomena
  – Focus on nouns in early vocabulary
  – Rapid growth, but probably no “word spurt”

• Approaches
  – Social/cognitive: Tomasello, Bloom, Clark
  – Innate constraints: Markman, Waxman
  – Learned constraints: Golinkoff, Hirsh-Pasek, Smith

• Limitations (among many)
  – No discussion of other kinds of words, e.g. Wednesday, fair, bat, the, to, etc.
Conclusions: Word learning

• Constraints on word learning
  – Probably not hard, domain-specific constraints
  – Correct inference more likely to be driven by general social & cognitive inferences

• Abilities for word learning
  – Conceptual
    • Similarity, category generalization
  – Social/pragmatic
    • Referential intent, eye-gaze, discourse, etc.
  – Linguistic
    • Syntactic form and POS help infer meaning

• Cue integration (?!)
  – My own personal favorite: how do learners put all of this information together?
Every human child exposed in even limited ways to the triggering experience of linguistic data develops a full, rich capacity which is essentially homogeneous with that of the surrounding community.

Gradual development of inflection

Figure removed due to copyright restrictions.
Please see:
Generalization of Morphology

Case study: irregular and regular plurals

• English past tense
  – Mostly regular: walk -> walked
  – Occasionally (~100 forms) irregular: go -> went, run -> ran, sing -> sang

• English plural
  – Almost entirely regular: book -> books
  – Very few irregulars (~10 forms): mice, geese, teeth, feet, cacti, children, men, etc.
The past-tense debate: phenomena

Putative time course of acquisition

1. Some irregulars learned by rote
   - Most irregulars tend to be very high frequency
2. Then over-generalization of predominant pattern
3. Finally, correct performance on regulars

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>went</td>
<td>goed</td>
<td>went</td>
</tr>
<tr>
<td>saw</td>
<td>seed</td>
<td>saw</td>
</tr>
<tr>
<td>looked</td>
<td>looked</td>
<td>looked</td>
</tr>
<tr>
<td>walked</td>
<td>walked</td>
<td>walked</td>
</tr>
<tr>
<td>(unanalyzed)</td>
<td>(over-regularized)</td>
<td>(irregulars learned)</td>
</tr>
</tbody>
</table>
Time-course of acquisition

Figure removed due to copyright restrictions.
Please see:
More on the time course of acquisition

Past tense over-regularization rates

Figure removed due to copyright restrictions.
Please see:

Conclusions:
Generalization of morphology is in general very good, although older children are still not perfect
Experimental evidence

- Yoked test of 12 children aged 3;4 - 5;0 (avg. 4;2) in:
  - Production: name picture of single; plural
  - Recognition: 2AFC between puppets saying correct and incorrect forms
  - Comprehension: Select singular and plural from a 3 * 3 array

Learnability theory

How do you recover from overgeneralization without negative feedback?

Figure removed due to copyright restrictions.
Please see:
Figure 1 from Pinker, S. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press, 1984.
Feedback in morphology learning

Child: My teacher held the baby rabbits and we petted them.
Parent: Did you say your teacher held the baby rabbits?
Child: Yes.
Parent: What did you say she did?
Child: She held the baby rabbits and we petted them.
Parent: Did you say she held them tightly?
Child: No, she held them loosely.


Child: Want other spoon, Daddy.
Father: You mean you want THE OTHER SPOON?
Child: Yes, I want other one spoon please Daddy.
Father: Can you say “the other spoon”?
Child: Other… one… spoon.
Father: Say… “other”.
Child: Other.
Father: “Spoon.”
Child: Spoon.
Father: “Other… spoon.”
Child: Other… spoon. Now give me other one spoon?

Theoretical accounts

(how do we avoid learnability paradoxes?)

• Two answers:
  – Rule learning mechanisms or
  – Statistical generalization

• Words and rules (Pinker, 1991; 1999)

• Single-route theories (Rumelhart & McClelland, 1986)

• Alternative accounts
  – Probabilistic rules (Albright & Hayes, 2004)
  – Competition and spreading activation (Ramscar & Yarlett, 2006)
The Words and Rules Account

- Stem to be inflected
- Walk
- Look up irregular form
- Blocking
- Walked
- ... 

Production

The Words and Rules Account

Stem to be inflected

Invoke regular rule

Look up irregular form

Production

The Words and Rules Account

Stem to be inflected

Invoke regular rule

Look up irregular form

Associations between other irregular forms

Rumelhart & McClelland (1986) Model

Figure removed due to copyright restrictions.
Please see:
Figure from Rumelhart, D. E., and J. L. McClelland. "On Learning the Past Tenses of English Verbs."
Criticisms of the R&M1986 model

1. it cannot represent certain words
2. it cannot learn many rules
3. it can learn rules found in no human language
4. it cannot explain morphological and phonological regularities
5. it cannot explain the differences between irregular and regular forms
6. it fails at its assigned task of mastering the past tense of English
7. it gives an incorrect explanation for two developmental phenomena: stages of overregularization of irregular forms such as bringed, and the appearance of doubly-marked forms such as ated, and
8. it gives accounts of two others (infrequent overregularization of verbs ending in t/d, and the order of acquisition of different irregular subclasses) that are indistinguishable from those of rule-based theories

Back and forth some more

- PDP models without changes in learning

- Broad arguments against PDP modeling in general

- Non-PDP single-route models
Alternatives: Ramscar & Yarlett (2006)

Task: what is this? (mouse), what are these? (mice/mouses)

![Graph showing quality differences between low and high conditions]

- Low Condition: 
  - Group 1&2: Plotted at -3
  - Group 3&4: Plotted at -2

- High Condition: 
  - Group 1&2: Plotted at -2
  - Group 3&4: Plotted at -1

The graph indicates a significant difference between the low and high conditions, with a p-value of less than 0.05.

$n = 23$

Alternatives: Ramscar & Yarlett (2006)

Spreading activation for irregular forms

- Regulars
  - Competitive response
  - Imitative response + support

- Irregulars
  - Competitive response
  - Imitative response + support

Overall response propensity
Figure removed due to copyright restrictions.
Please see:

---

Figure removed due to copyright restrictions.
Please see:
“...we infer that analogy, in its most basic form, is too powerful a mechanism to account for how morphological systems in human languages work; and that a multiple-rule approach is a more accurate model of how speakers create novel forms.”

Figure removed due to copyright restrictions..
Please see:
Conclusions: Inflectional morphology

• Generalization
  – Pervasive (so language is not imitation)
  – Gradual (doesn’t reflect binary processes)

• Acrimonious debate
  – Pattern of data is more subtle than previously thought
  – Any complete account should include:
    • graded generalization
    • probabilistic competition

• Learnability
  – Many computational level models demonstrate in principle learnability of morphological patterns (syntax is another story)
  – No proof: no validated process models yet
Returning to where we began

• Language constitutes a major difference between ourselves and other primates
• Looking at children may give us insight into adult performance
• Learning about language may tell us about the nature of mental representation (c.f. Fodor, 1971)
• Universals of language may tell us about the structure of the mind (c.f. Chomsky, 1965)
• Language acquisition can settle philosophical issues like rationalism vs. empiricism (c.f. Pinker, 1994)
Returning to where we began

- Language constitutes a major difference between ourselves and other primates
  - Word learning has a huge social component (!)
  - So maybe we should look at social capabilities also
- Learning about language may tell us about the nature of mental representation (c.f. Fodor, 1971)
  - Proposals are getting more sophisticated
  - Challenge #1: incorporating fast, symbolic generalization into graded, probabilistic processes
  - Challenge #2: integrating multiple information sources (e.g., social, linguistic, & conceptual)
End
We do not want to know languages, we want to know language; what language is, how it can form a vehicle or an organ of thought; we want to know its origin, its nature, its laws; and it is only in order to arrive at that knowledge that we collect, arrange, and classify all the facts of language that are within our reach.

—Max Müller, *Lectures on the Science of Language*