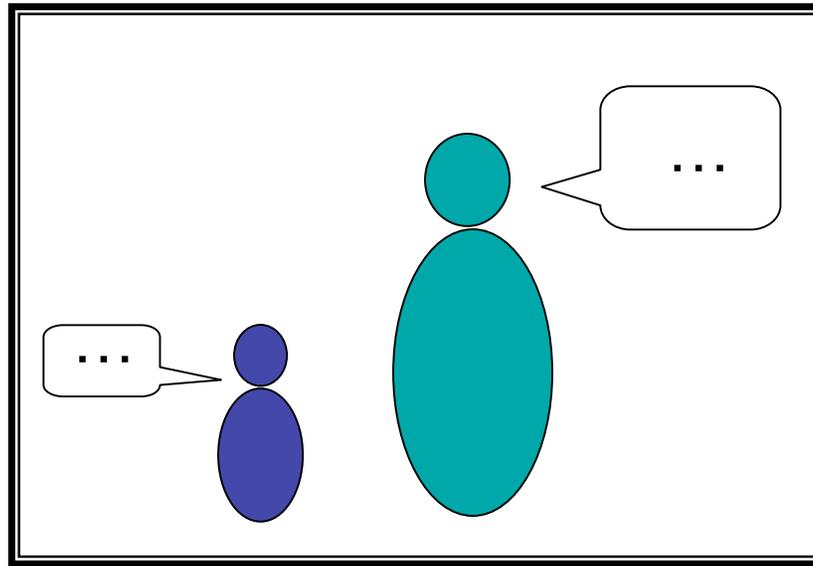


# Language Acquisition



**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

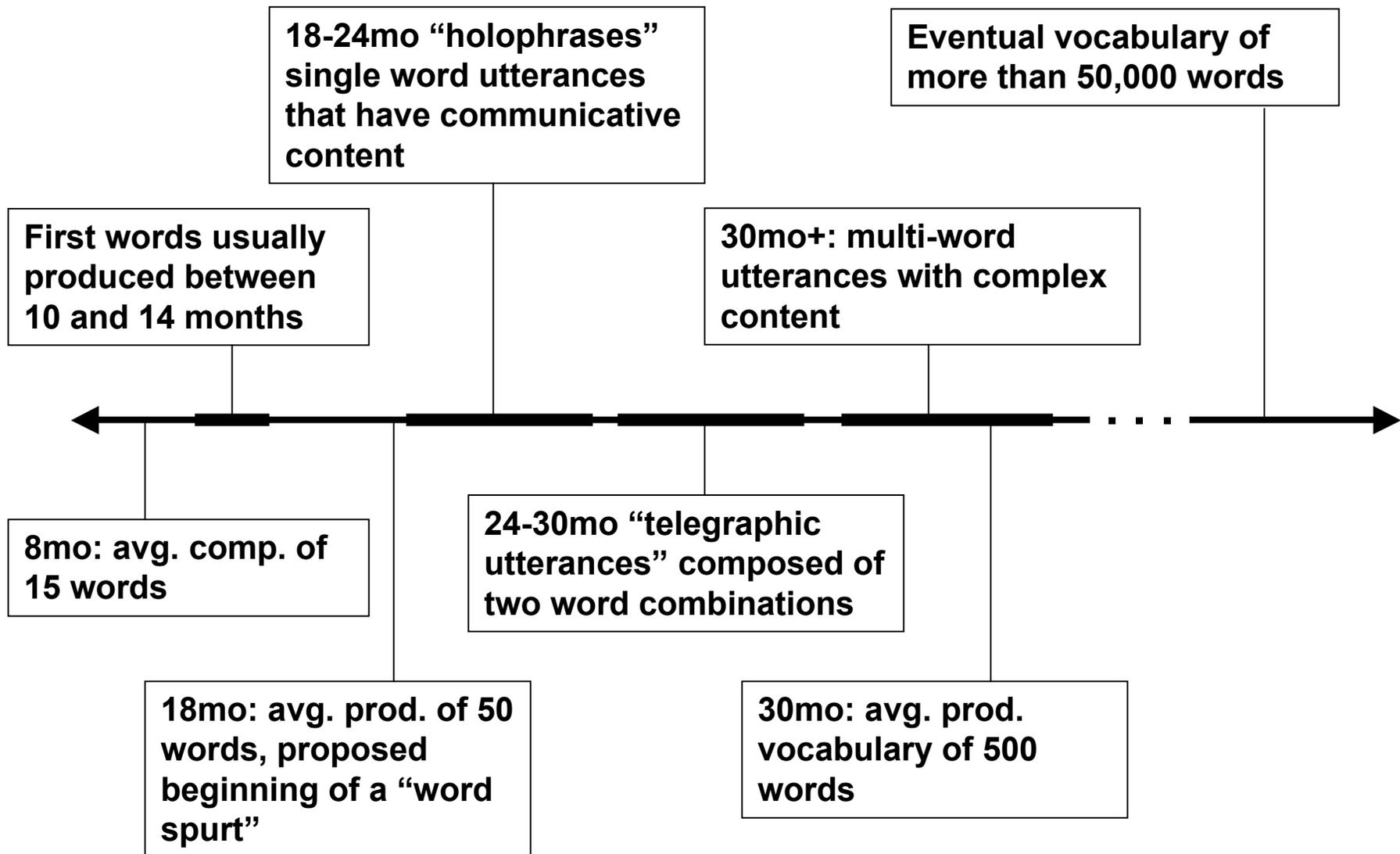
# Word Learning: Propaganda

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.”

Image of chimpanzee examining objects with trainer. Image removed due to copyright restrictions.

Morgan’s Canon (1894): “In no case may we interpret an action as the outcome of the exercise of a higher psychological 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



# Word Learning: Comprehension

Figure removed due to copyright restrictions.

Please see:

Figure from Fenson, L., P. S. Dale, J. S. Reznick, E. Bates, D. J. Thal, S. J. Pethick, M. Tomasello, C. B. Mervis, and J. Stiles. "Variability in Early Communicative Development." *Monographs of the SRCD* 59, no. 5 (1994).

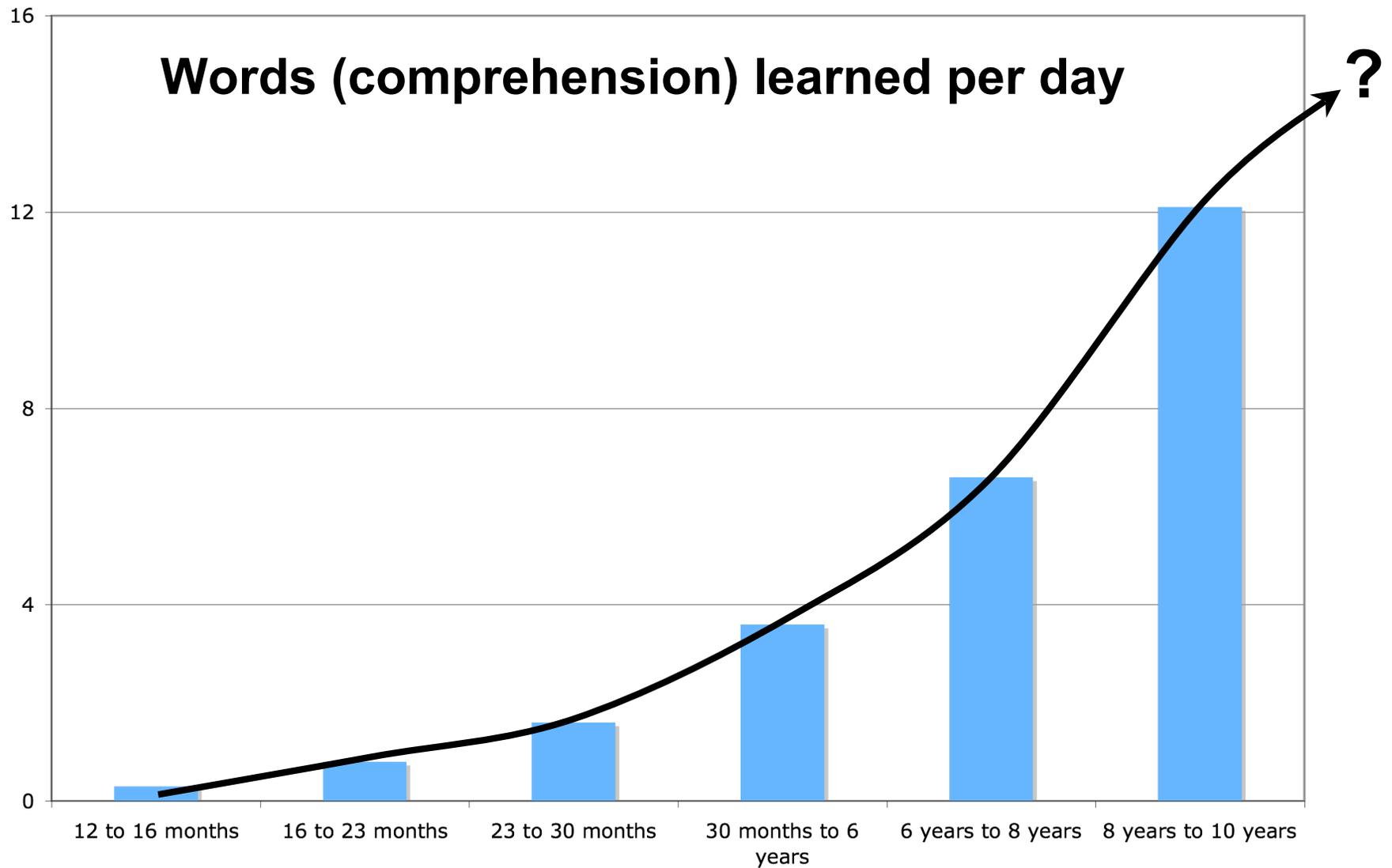
# Word Learning: Production

Figures removed due to copyright restrictions.

Please see:

Figures from Fenson, L., P. S. Dale, J. S. Reznick, E. Bates, D. J. Thal, S. J. Pethick, M. Tomasello, C. B. Mervis, and J. Stiles. "Variability in Early Communicative Development." *Monographs of the SRCD* 59, no. 5 (1994).

# Word Learning: Rate of acquisition



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

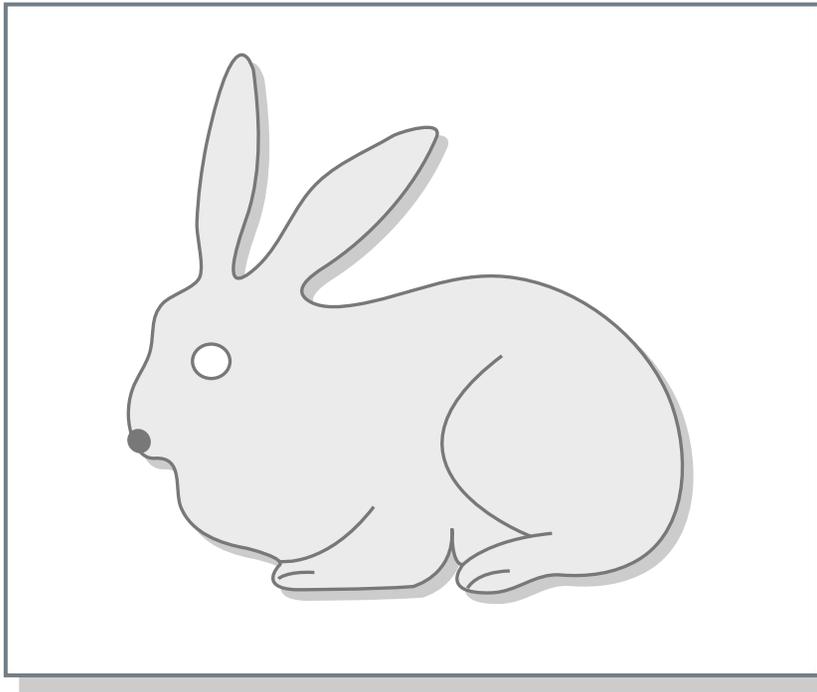
Figure removed due to copyright restrictions.

Please see:

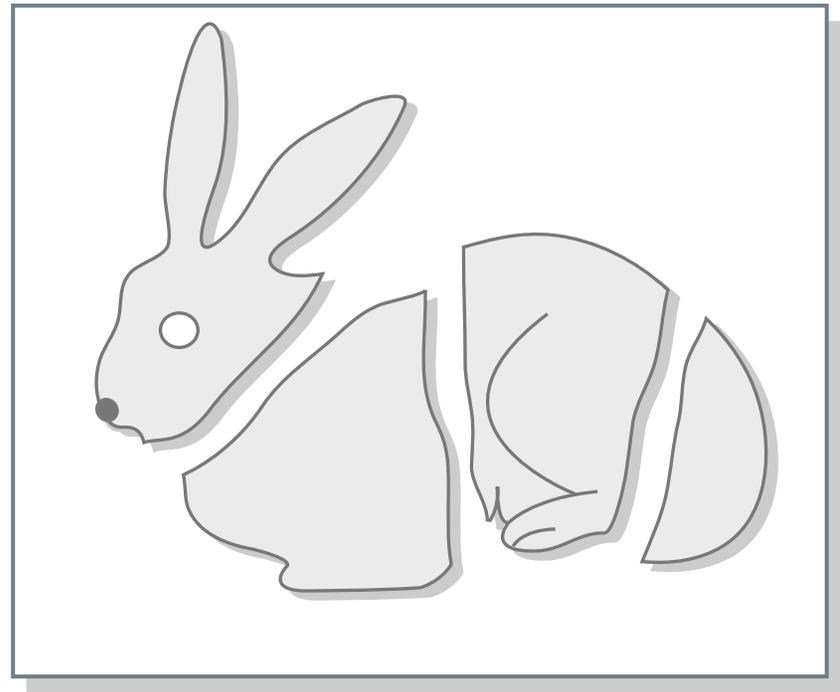
Figure from Gillette, J., H. Gleitman, L. Gleitman, and A. Lederer. "Human Simulations of Vocabulary Learning." *Cognition* 73 (1999): 135-176.

# And you thought verbs were hard?

Rabbit?



Or group of undetached  
rabbit parts?



Figures by MIT OCW.

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

W.V.O. Quine, *Word and Object*, MIT Press, Cambridge Mass., 1960.

# 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**

Markman, E.M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. In Gunnar, M.R. & Maratsos, M.P. (eds): Modularity and constraints in language and cognition. Hillsdale, NJ: Erlbaum .

# 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:

Figure 1 from Smith, L. B., S. S. Jones, B. Landau, L. Gershkoff-Stowe, and L. Samuelson. "Object Name Learning Provides On-the-Job Training for Attention." *Psychological Science* 13 (2002): 13-19.

# Shape bias as a tool for word learning

## Vocabulary growth

Figure removed due to copyright restrictions.

Please see:

Figure from Smith, L. B., S. S. Jones, B. Landau, L. Gershkoff-Stowe, and L. Samuelson. "Object Name Learning Provides On-the-Job Training for Attention." *Psychological Science* 13 (2002): 13-19.

## Training items

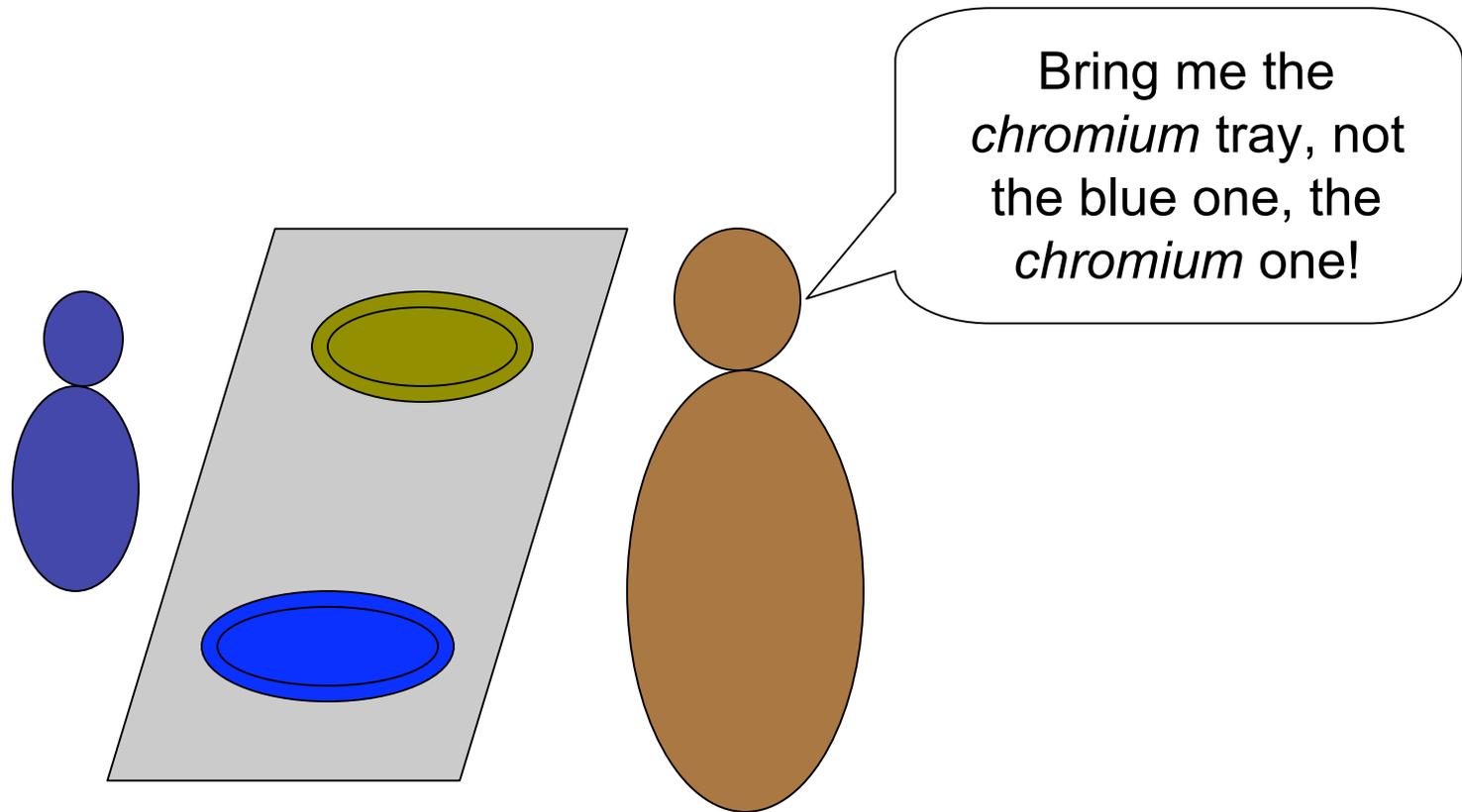
Figure removed due to copyright restrictions.

Please see:

Figure 4 from Smith, L. B., S. S. Jones, B. Landau, L. Gershkoff-Stowe, and L. Samuelson. "Object Name Learning Provides On-the-Job Training for Attention." *Psychological Science* 13 (2002): 13-19.

# Abilities for word learning: “fast mapping”

3 & 4 year olds learn words after a single exposure



Carey, S. (1978). The child as word learner. In M. Halle, G. Miller & J. Bresnan (Eds.), *Linguistic Theory and Psychological Reality*, 264-293. Cambridge, MA: MIT Press.

# Fast mapping: Domain-general ability?

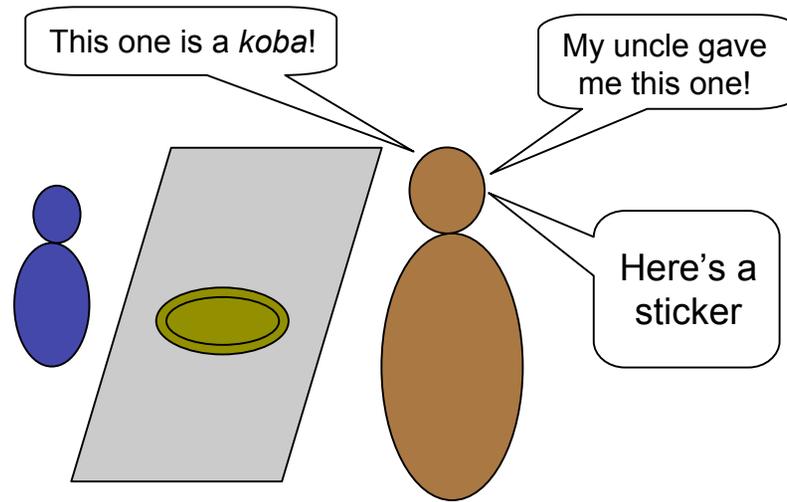


Figure removed due to copyright restrictions.

Please see:

Figures from Markson, L. and P. Bloom. "Evidence against a dedicated system for word learning in children." *Nature* 385 (1997): 813-815.

# Pragmatic Cues for Word Learning

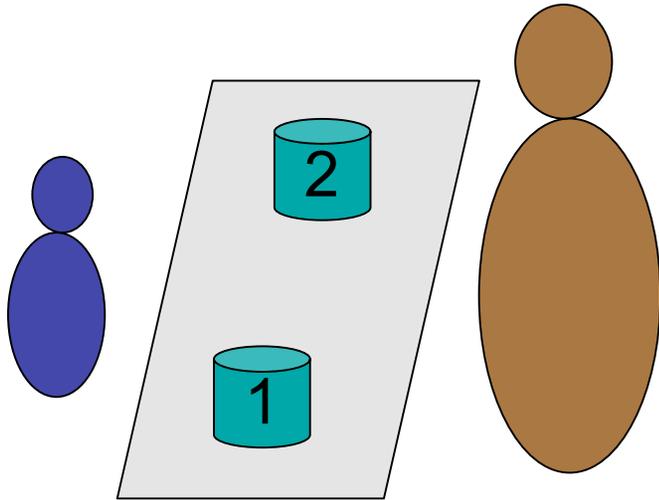


Figure removed due to copyright restrictions.

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.

## Two conditions

1. Coincide: name matches toy shown
2. Conflict: name does not match toy shown

# 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:

Table 2 from Tomasello, M., and N. Akhtar. "Two-year-olds use pragmatic cues to differentiate reference to objects and actions." *Cognitive Development* 10 (1995): 201-224.

# 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 blicket” to objects of the same shape, extended “some blicket” 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?

# Acquisition of Morphology: Introduction

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.

Image of Stephen Anderson, Linguistics Professor at Yale University. Image removed due to copyright restrictions.

Anderson, S. & Lightfoot, D. (2000). The Human Language Faculty as an Organ  
Annual Review of Physiology.

# Gradual development of inflection

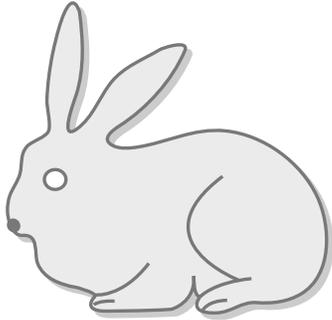
Figure removed due to copyright restrictions.

Please see:

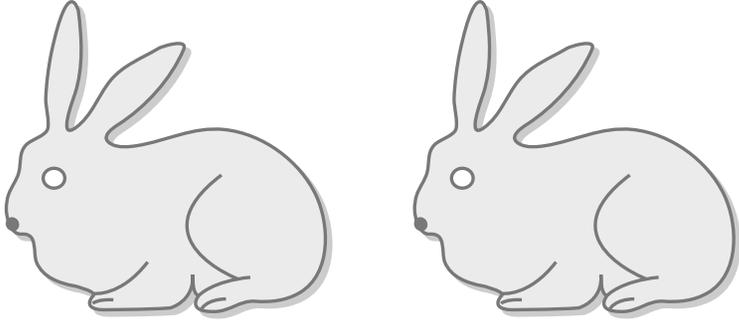
Figure from Fenson, L., P.S. Dale, J. S. Reznick, E. Bates, D. J. Thal, S. J. Pethick, M. Tomasello, C. B. Mervis, and J. Stiles. "Variability in Early Communicative Development." *Monographs of the SRCD* 59, no. 5 (1994).

# Generalization of Morphology

THE WUG TEST



This is a wug.



Now there is another one.  
There are two of them.  
There are two \_\_\_\_ .

Figure by MIT OCW.

# 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

### **Stage 1**

went

saw

looked

walked

**(unanalyzed)**

### **Stage 2**

goed

seed

looked

walked

**(over-regularized)**

### **Stage 3**

went

saw

looked

walked

**(irregulars learned)**

# Time-course of acquisition

Figure removed due to copyright restrictions.

Please see:

Figure from Fenson, L., P.S. Dale, J. S. Reznick, E. Bates, D. J. Thal, S. J. Pethick, M. Tomasello, C. B. Mervis, and J. Stiles. "Variability in Early Communicative Development." *Monographs of the SRCD* 59, no. 5 (1994).

# More on the time course of acquisition

## Past tense over-regularization rates

Figure removed due to copyright restrictions.

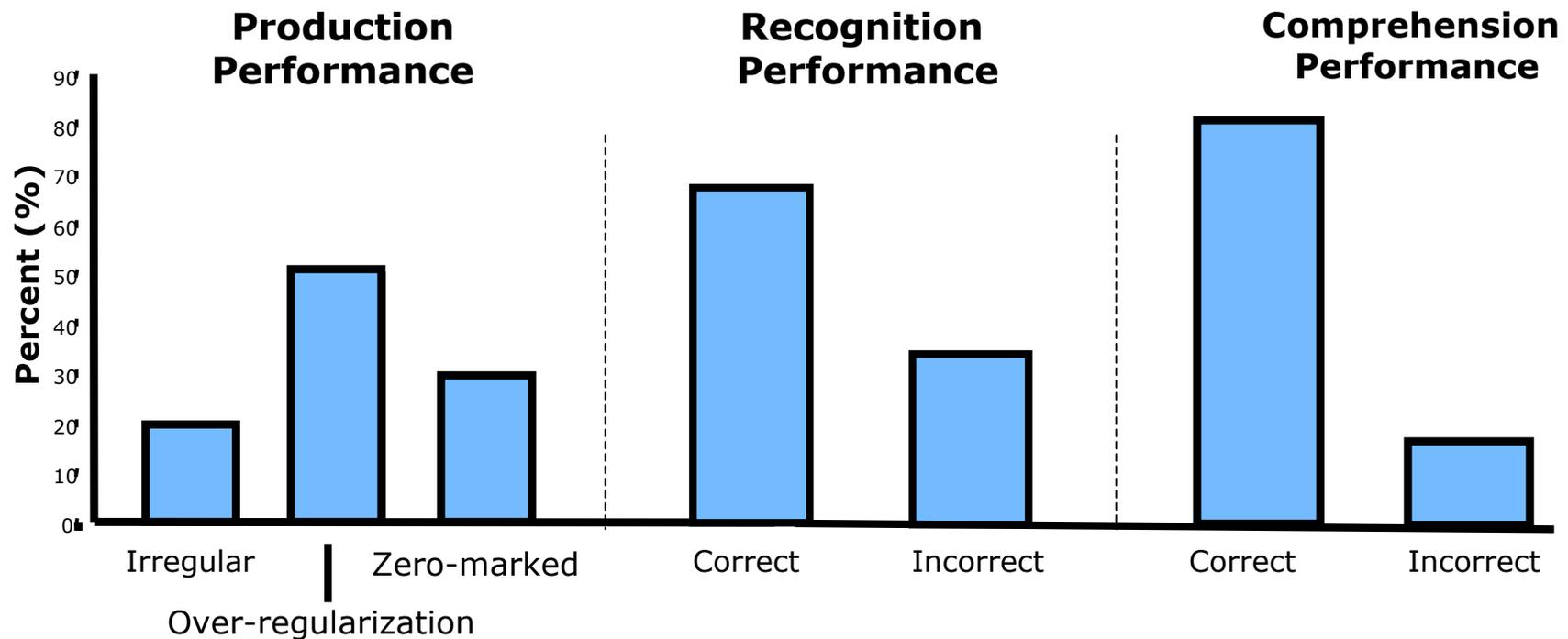
Please see:

Marcus, G. F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen, and F. Xu.  
"Overregularization in Language Acquisition." *Monographs of the Society for Research in Child Development* 57, nos. 4, 228 (1992).

**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



Ramscar, M. & Yarlett, D. (2006). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition.

# 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 holded 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 holded the baby rabbits and we petted them.

**Parent:** Did you say she held them tightly?

**Child:** No, she holded them loosely.

Brown, R. (1973). *A First Language: The Early Stages*.

**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?

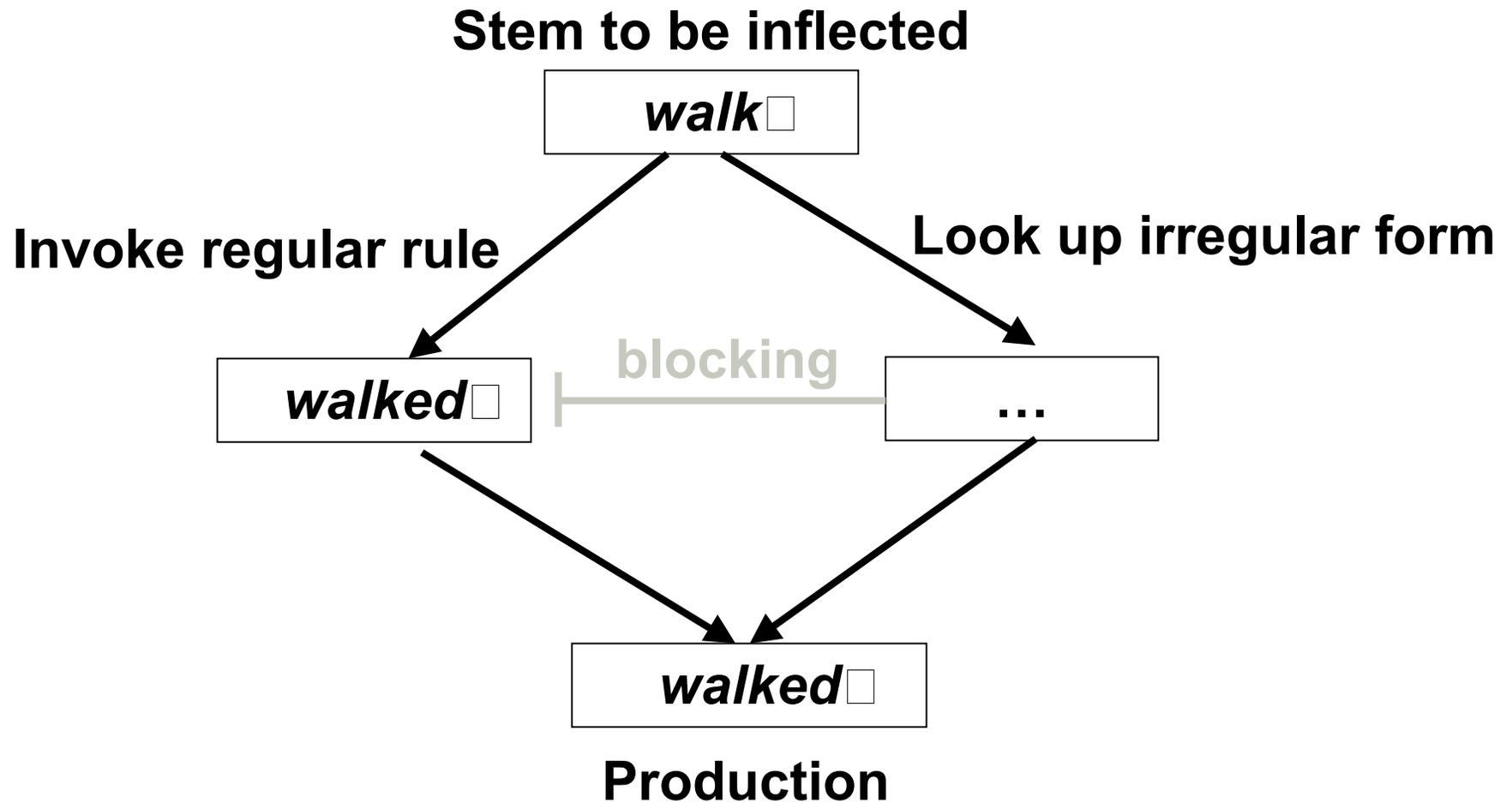
M. Braine, "On Two Types of Models of the Internalization of Grammars," In D. Slobin, editor, *The Ontogenesis of Grammar*. Academic Press, 1971.

# Theoretical accounts

(how do we avoid learnability paradoxes?)

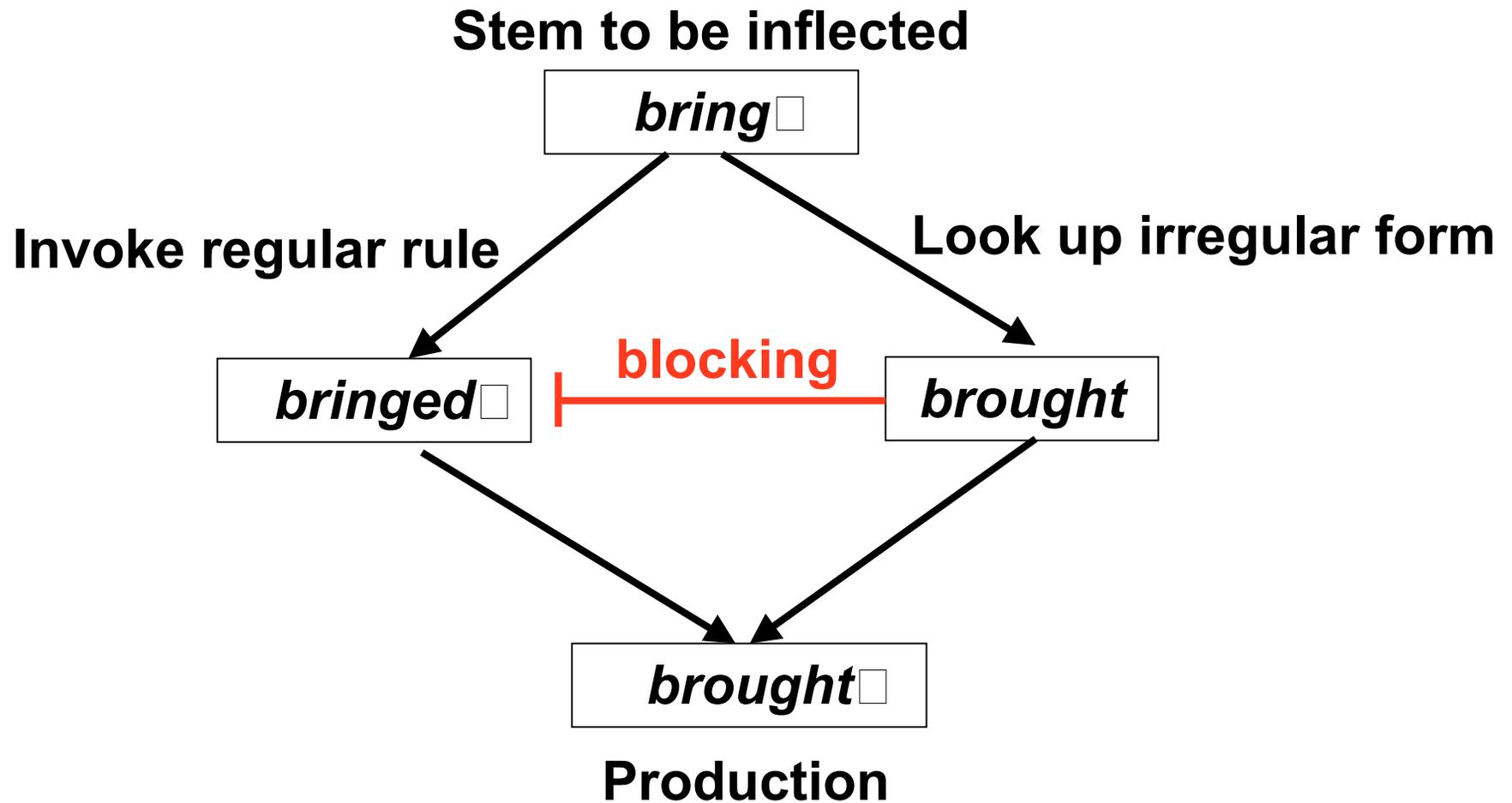
- Two answers:
  - Rule learning mechanisms or
  - Statistical generalization
- Words and rules (Pinker, 1991; 1999)
- Single-route theories (Rumelhart & McClelland, 1986)
- Alternative accounts
  - Probabilistic rules (Albright & Hayes, 2004)
  - Competition and spreading activation (Ramscar & Yarlett, 2006)

# The Words and Rules Account

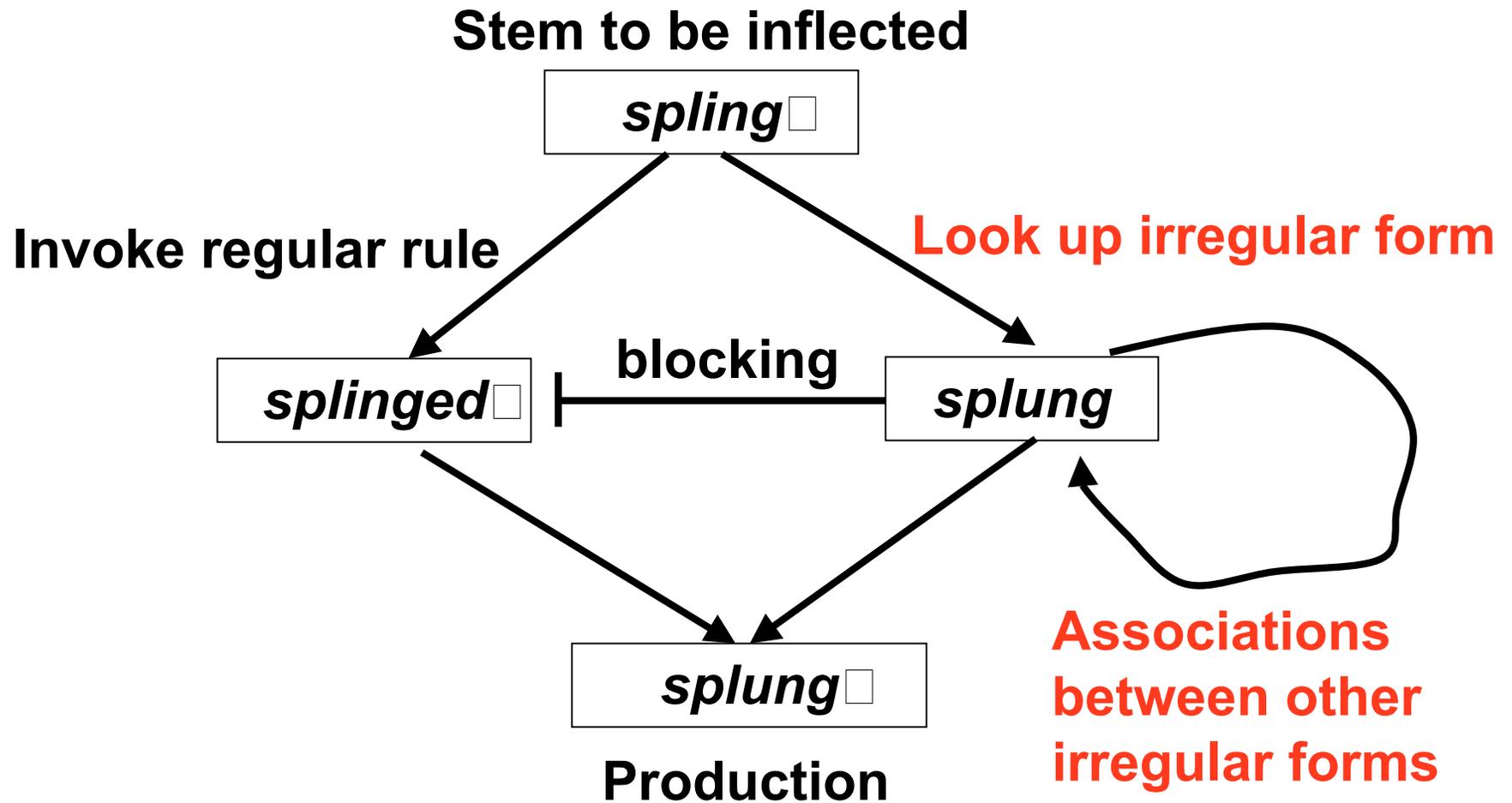


Pinker, S. (1991). Rules of Language. Science, 253, 530-535.

# The Words and Rules Account



# The Words and Rules Account



Pinker, S. (1991). Rules of Language. *Science*, 253, 530-535.

# 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."  
*PDP Research Group 2* (1986): 216-271.

# 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

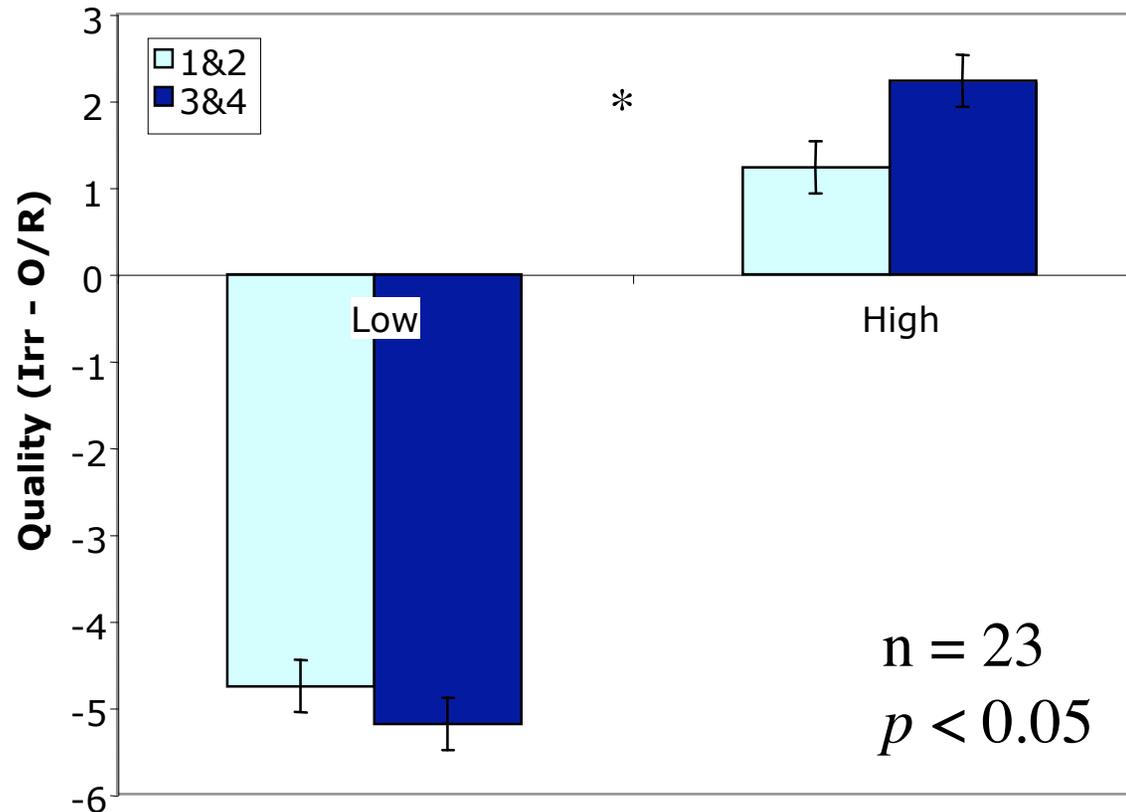
Pinker, S. & Prince, A. (1988) On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73-193. Reprinted in S. Pinker & J. Mehler (Eds.) (1988) *Connections and symbols*. Cambridge, MA: MIT Press.

# Back and forth some more

- PDP models without changes in learning
  - Plunkett, K. & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 48(1), 21-69.
- Broad arguments against PDP modeling in general
  - Marcus, G. F. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. Cambridge, MA: MIT Press.
- Non-PDP single-route models
  - Hahn, U., & Nakisa, R.C. (2000). German Inflection: Single or Dual Route? *Cognitive Psychology*, 41, 313-360.

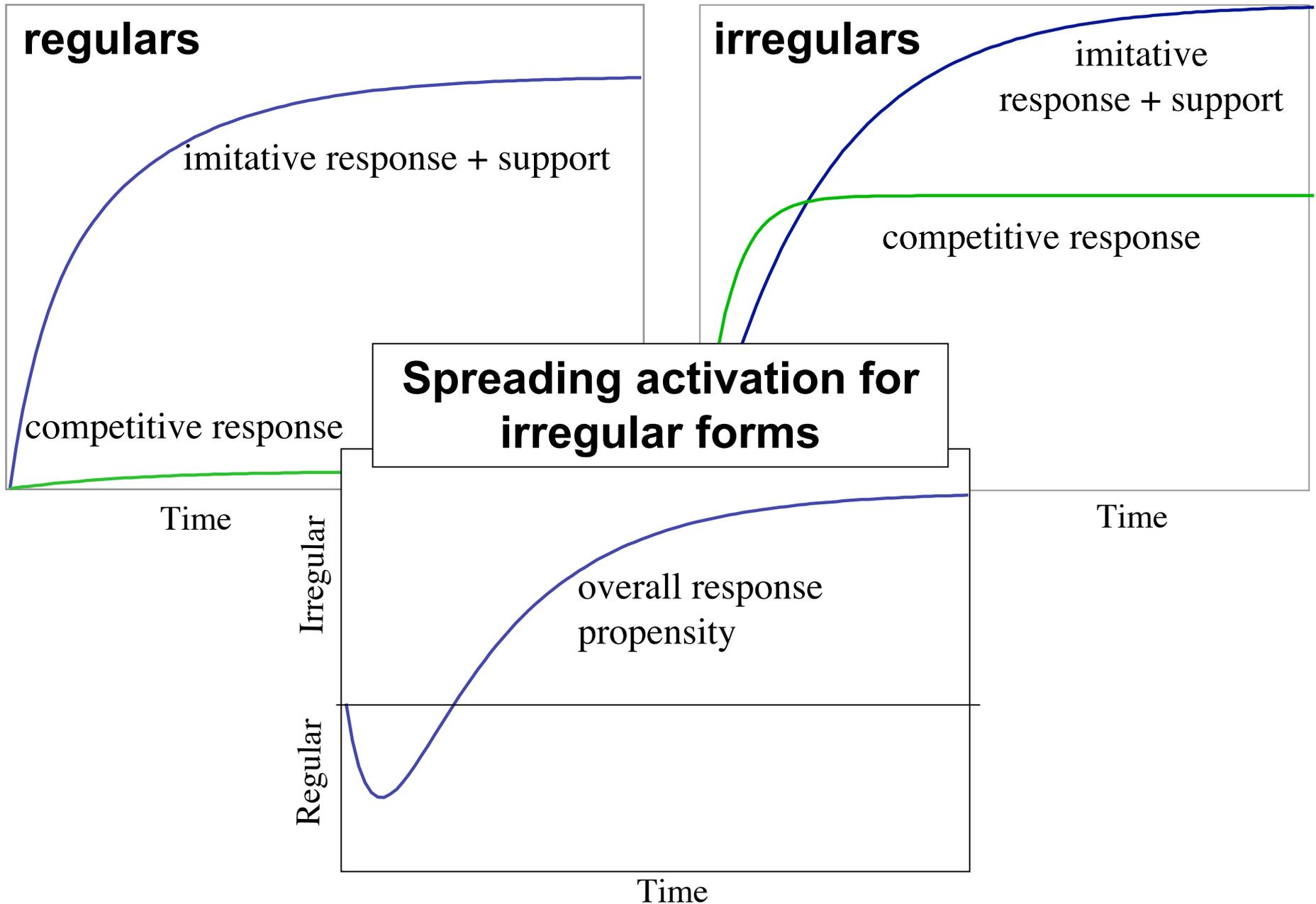
# Alternatives: Ramscar & Yarlett (2006)

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



Ramscar, M. & Yarlett, D. (2006). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition.

# Alternatives: Ramskar & Yarlett (2006)



# Albright & Hayes (2004)

## Rules model

Figure removed due to copyright restrictions..

Please see:

Table 1 from Albright, A., and B. Hayes. "Rules vs. analogy in English past tenses: A computational/experimental study." *Cognition* 90 (2004): 119-161.

## Analogy model

Figure removed due to copyright restrictions..

Please see:

Figure 1 from Albright, A., and B. Hayes. "Rules vs. analogy in English past tenses: A computational/experimental study." *Cognition* 90 (2004): 119-161.

# Albright & Hayes (2004)

“...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:

Figure 1 from Albright, A., and B. Hayes. "Rules vs. analogy in English past tenses: A computational/experimental study." *Cognition* 90 (2004): 119-161.

# 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

# Language Acquisition: Epigraph

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*