

Ling 566  
Oct 31, 2013  
Grammar and Processing

# Overview

- Psycholinguistics and grammar design
  - What grammar has to say
  - What psychological evidence has to say
    - Acquisition
    - Production
    - Comprehension
- Universals

# What does grammar have to do with psychology?

Three ways it could be relevant:

- It provides insight into how children acquire language.
- It provides insight into how speakers produce utterances.
- It provides insight into how listeners understand utterances.

# Our model: Key characteristics

- Surface-oriented
- Constraint-based
- Lexicalist

# Chomsky's position:

- Grammar represents knowledge of language (“competence”).
- This is distinct from use of language (“performance”).
- We can draw a strong conclusion about language acquisition, namely, most grammatical knowledge is innate and task-specific.
- Serious study of language use (production and comprehension) depends on having a well-developed theory of competence.

# Brief remarks on language acquisition

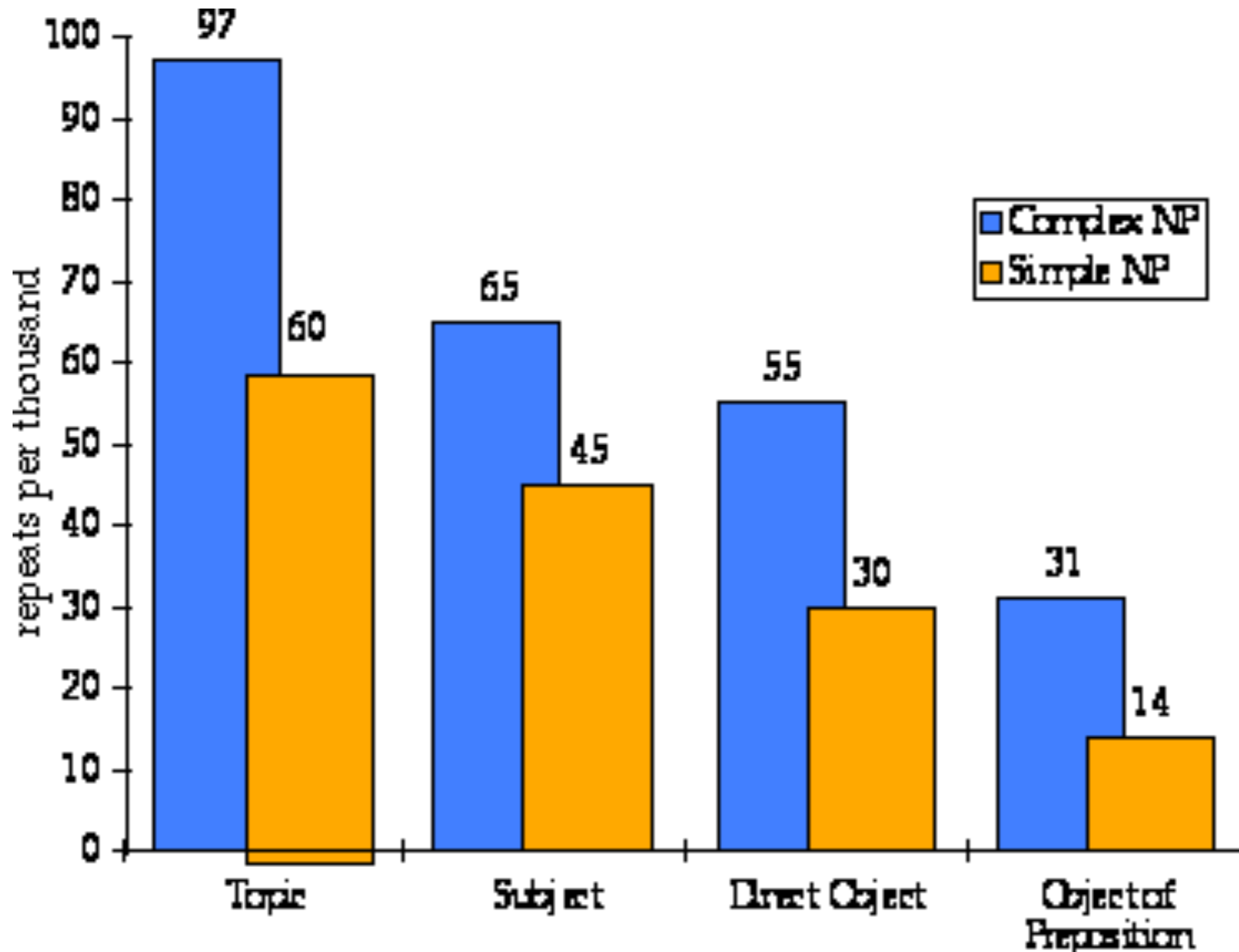
- Chomsky's nativism is very controversial
  - It is based on the “poverty of the stimulus” argument, and a model of learning as hypothesis testing.
  - The environment may be more informative than he assumes.
  - There may be more powerful learning methods than he assumes.
- There has not been much work on language acquisition using constraint-based lexicalist theories like ours; **but**
  - Explicit formulation is a prerequisite for testing learning models
  - Our feature structures could model richer context information.
- We're neutral with respect to this controversy.

# Production and Grammar

- Evidence for left-to-right effects
- Evidence for top-down planning

# Disfluencies are sensitive to structure:

Repeat rate of *the* varies with position and complexity of the NP it introduces:





# Production errors are sensitive to syntactic structure

Agreement errors are more common with PP complements than sentential complements: errors like (2) are significantly more common than errors like (1).

(1) *\*The claim that the wolves had raised the babies were rejected.*

vs.

(2) *\*The claim about the newborn babies were rejected.*

# So why?

- Speculation: Clauses are their own agreement domains, so people don't mistake an NP in a lower clause as a trigger for agreement
- Original work: Kay Bock (1980s).

## Some high-level sentence planning is necessary, too

- *Ich habe dem Mann, den ich gesehen habe geholfen.*  
I have the-dat man who-acc I seen have helped  
“I helped the man I saw”
- *Ich habe den Mann, dem ich geholfen habe gesehen.*  
I have the-acc man who-dat I helped have seen.  
“I saw the man I helped ”
- The choice between *dem* and *den* depends on the choice of verbs several words later.

# A production model should allow interaction of top-down and left-to-right information

- Grammar plays a role in production.
- Partial grammatical information should be accessible by the production mechanism as needed.
- This argues against grammatical theories that involve sequential derivations with fixed ordering.
- Our theory of grammar has the requisite flexibility.

# Comprehension

- Early work tried to use transformational grammar in modeling comprehension
- The Derivational Theory of Complexity: The psychological complexity of a sentence increases with the number of transformations involved in its derivation.
- Initial results seemed promising, but later work falsified the DTC.

# Some relevant quotes

- “The results show a remarkable correlation of amount of memory and number of transformations”  
– Chomsky, 1968
- “[I]nvestigations of DTC...have generally proved equivocal. This argues against the occurrence of grammatical derivations in the computations involved in sentence recognition”  
– Fodor, Bever, & Garrett, 1974

# Another quote

- “Experimental investigations of the psychological reality of linguistic structural descriptions have...proved quite successful.”  
– Fodor, Bever, & Garrett, 1974
- In particular, they concluded that “deep structures” and “surface structures” were psychologically real, but the transformations relating them weren’t.

# Early Evidence for the Psychological Reality of Deep Structures

- The proposed DS for (2) had three occurrences of *the detective*, while the proposed DS for (1) had only two:
  - (1) *The governor asked the detective to prevent drinking.*
  - (2) *The governor asked the detective to cease drinking.*
- In a recall experiment, *detective* was significantly more effective in prompting people to remember (2) than (1)



# Typical Problem Cases for the DTC

- (1) *Pat swam faster than Chris swam.*
- (2) *Pat swam faster than Chris did.*
- (3) *Pat swam faster than Chris.*

- The DTC predicts that (1) should be less complex than (2) or (3), because (2) and (3) involve an extra deletion transformation.
- In fact, subjects responded more slowly to (1) than to either (2) or (3).

# What should a psychologically real theory of grammar be like?

- The “deep structure” distinctions that are not evident on the surface should be represented.
- The transformational operations relating deep and surface structures should not be part of the theory.
- Our information-rich trees include all of the essential information in the traditional deep structures, but without the transformations.

# Jerry Fodor claims the human mind is “modular”

“A module is...an informationally encapsulated computational system -- an inference-making mechanism whose access to background information is constrained by general features of cognitive architecture.”

-- Fodor, 1985

A central issue in psycholinguistics over the past 20 years has been whether language is processed in a modular fashion.

# Tanenhaus's Eye-Tracking Experiments

- Participants wear a device on their heads that makes a videotape showing exactly what they're looking at.
- They listen to spoken instructions and carry out various tasks.
- They eye-tracking provides evidence of the cognitive activity of participants that can be correlated with the linguistic input.

# Non-linguistic visual information affects lexical access

- Participants' gaze settled on a referent before the word was completed, unless the initial syllable of the word was consistent with more than one object.
- For example, participants' gaze rested on the pencil after hearing *Pick up the pencil* more slowly when both a pencil and a penny were present.

# Non-linguistic visual information affects syntactic processing

- Eye movements showed that people hearing (1) often temporarily misinterpreted *on the towel* as the destination.  
(1) *Put the apple on the towel in the box.*
- When *on the towel* helped them choose between two apples, such misparses were significantly less frequent than when there was only one apple.

# General Conclusion of Eye-Tracking Studies

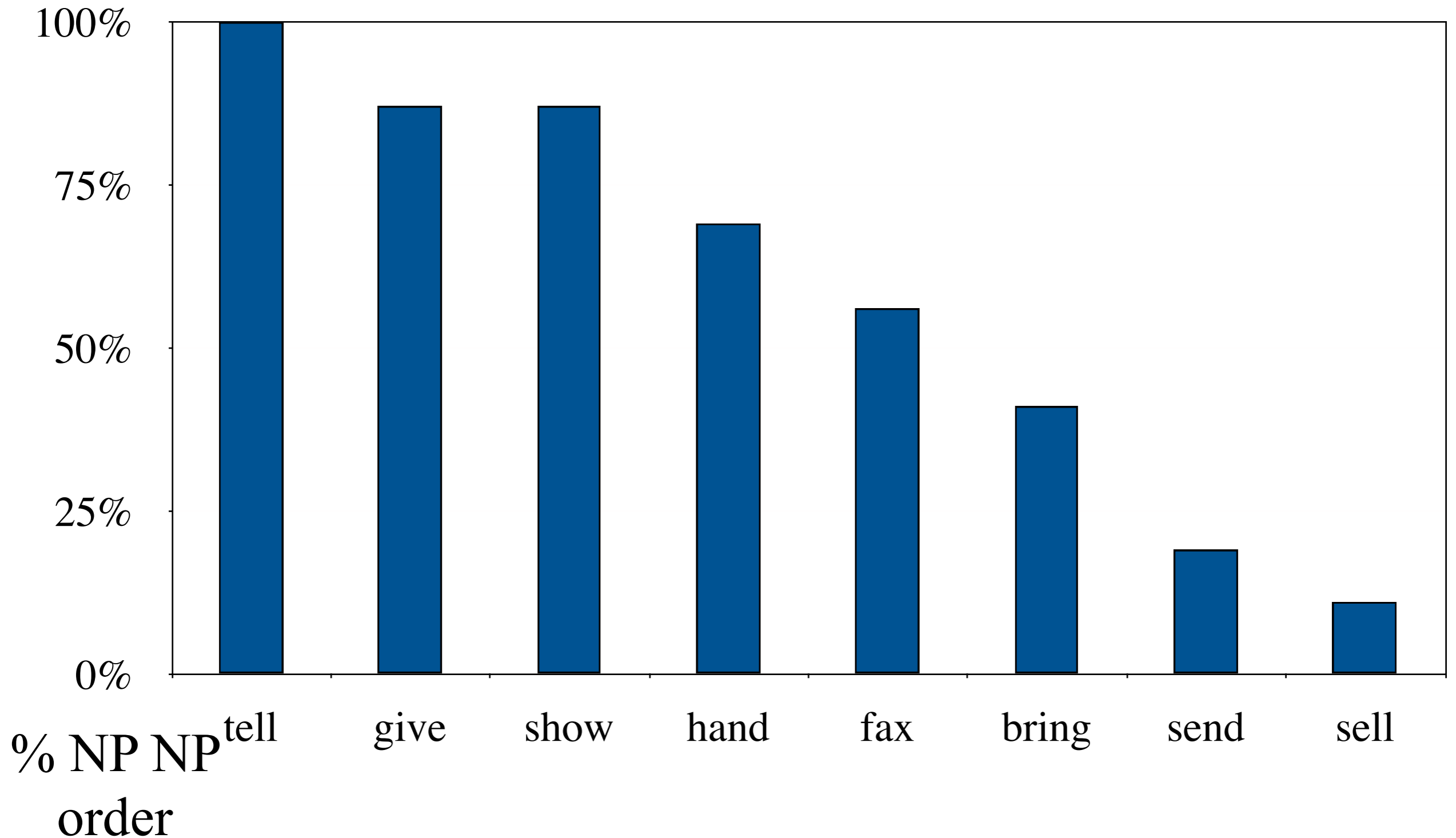
- People use whatever information is available as soon as it is useful in interpreting utterances.
- This argues against Fodorian modularity.
- It argues for a model of language in which information is represented in a uniform, order-independent fashion.

# Speakers know a great deal about individual words

- Individual lexical items have many idiosyncrasies in where they can occur, and in where they tend to occur.
- For example, the verb *behoove* occurs only with the subject *it* (and only in certain verb forms), and the verb *beware* has only the base form.
- We also know that the transitive use of *walk* is much rarer than the intransitive.



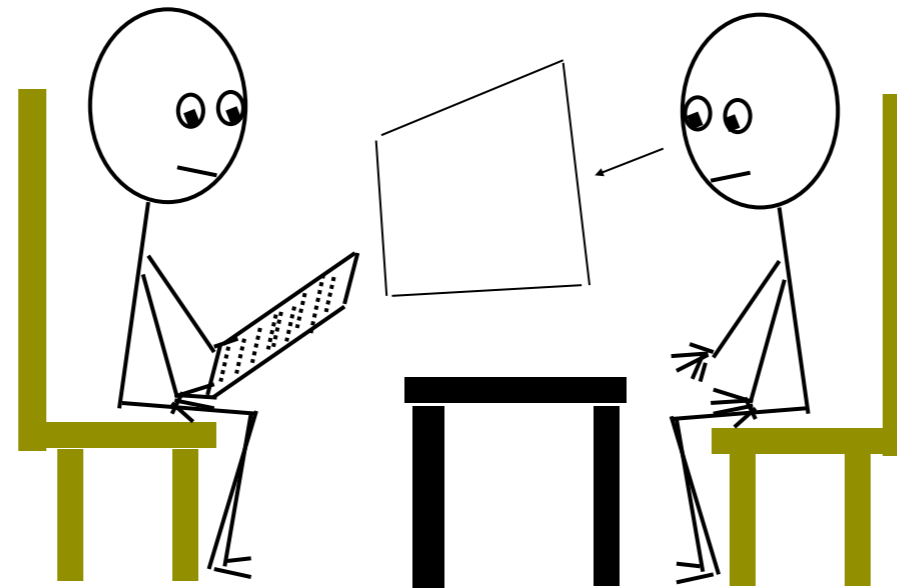
# V-NP-NP vs. V-NP-PP Frequency in the *NYT*



# Lexical biases influence processing

- Wasow et al ran a production experiment to test whether ambiguity avoidance would influence speakers' choice between (1) and (2):
  - (1) *They gave Grant's letters to Lincoln to a museum.*
  - (2) *They gave a museum Grant's letters to Lincoln.*
- Lexical bias of the verbs turned out to be a significant predictor of which form speakers used (and ambiguity avoidance turned out not to be).

# Experimental Method



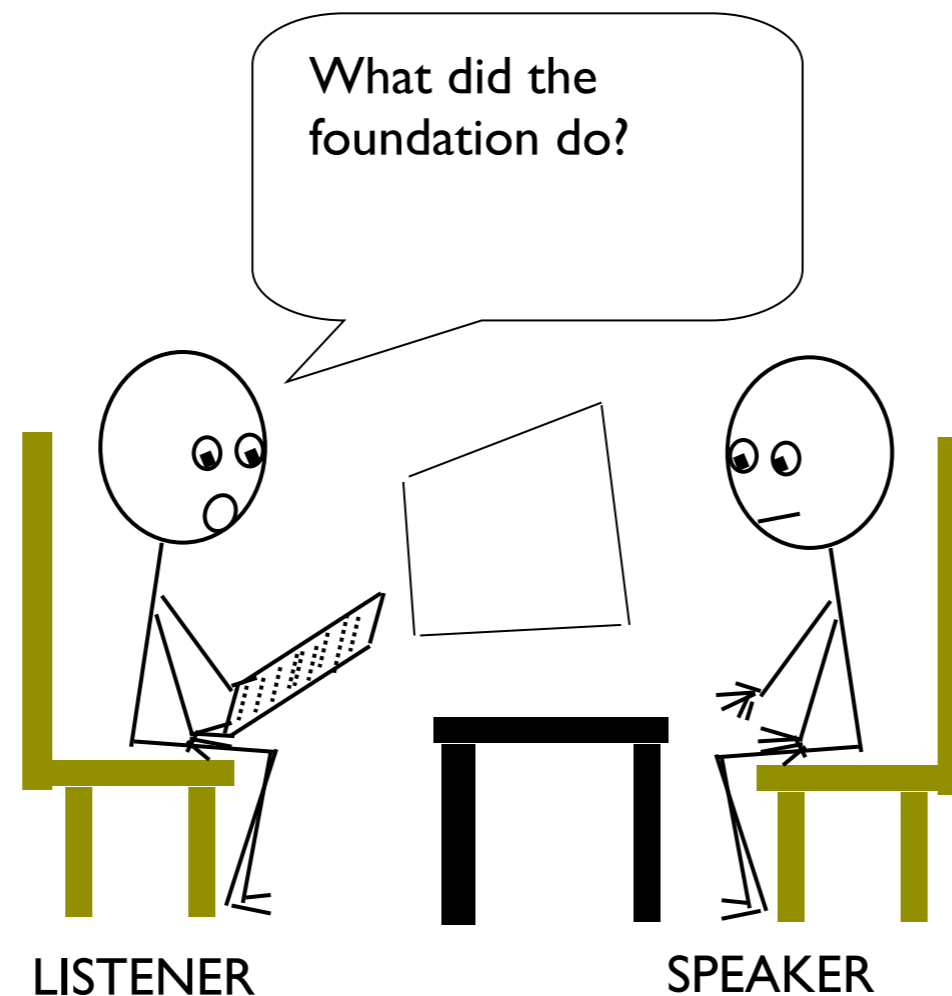
LISTENER

SPEAKER

1. Speaker silently reads a sentence:

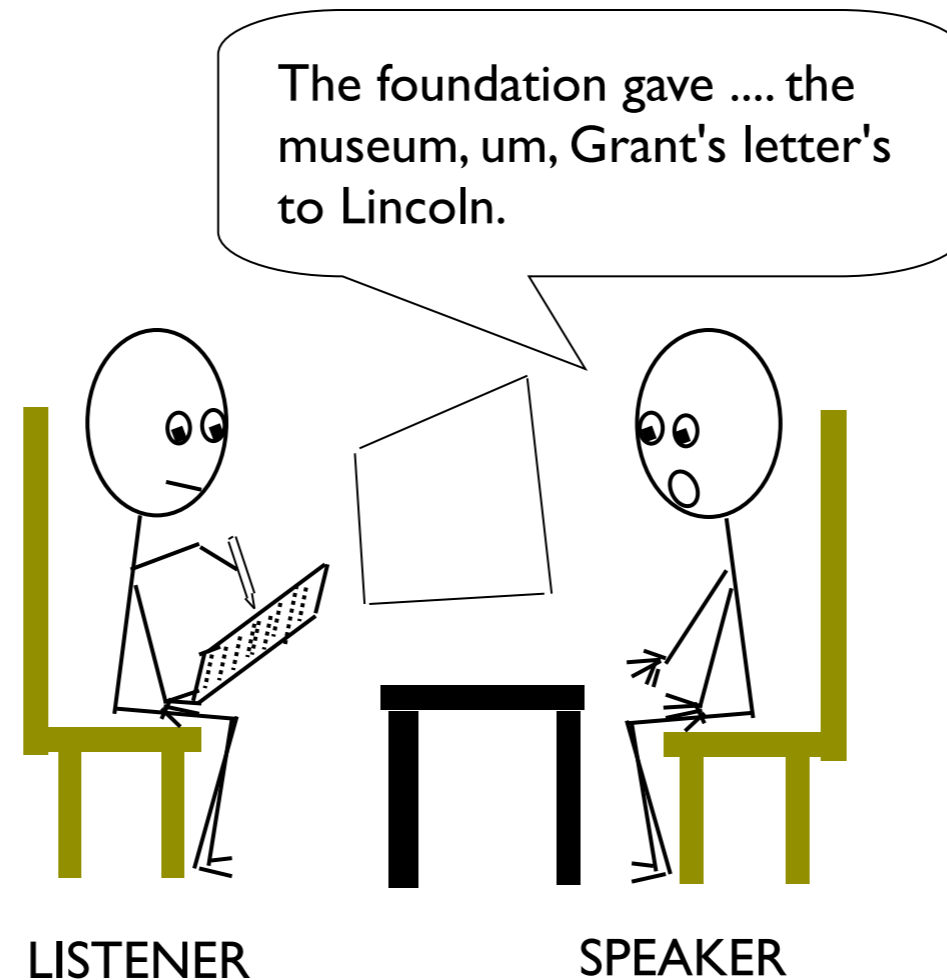
*A museum in Philadelphia received Grant's letters to Lincoln from the foundation.*

# Experimental Method, continued



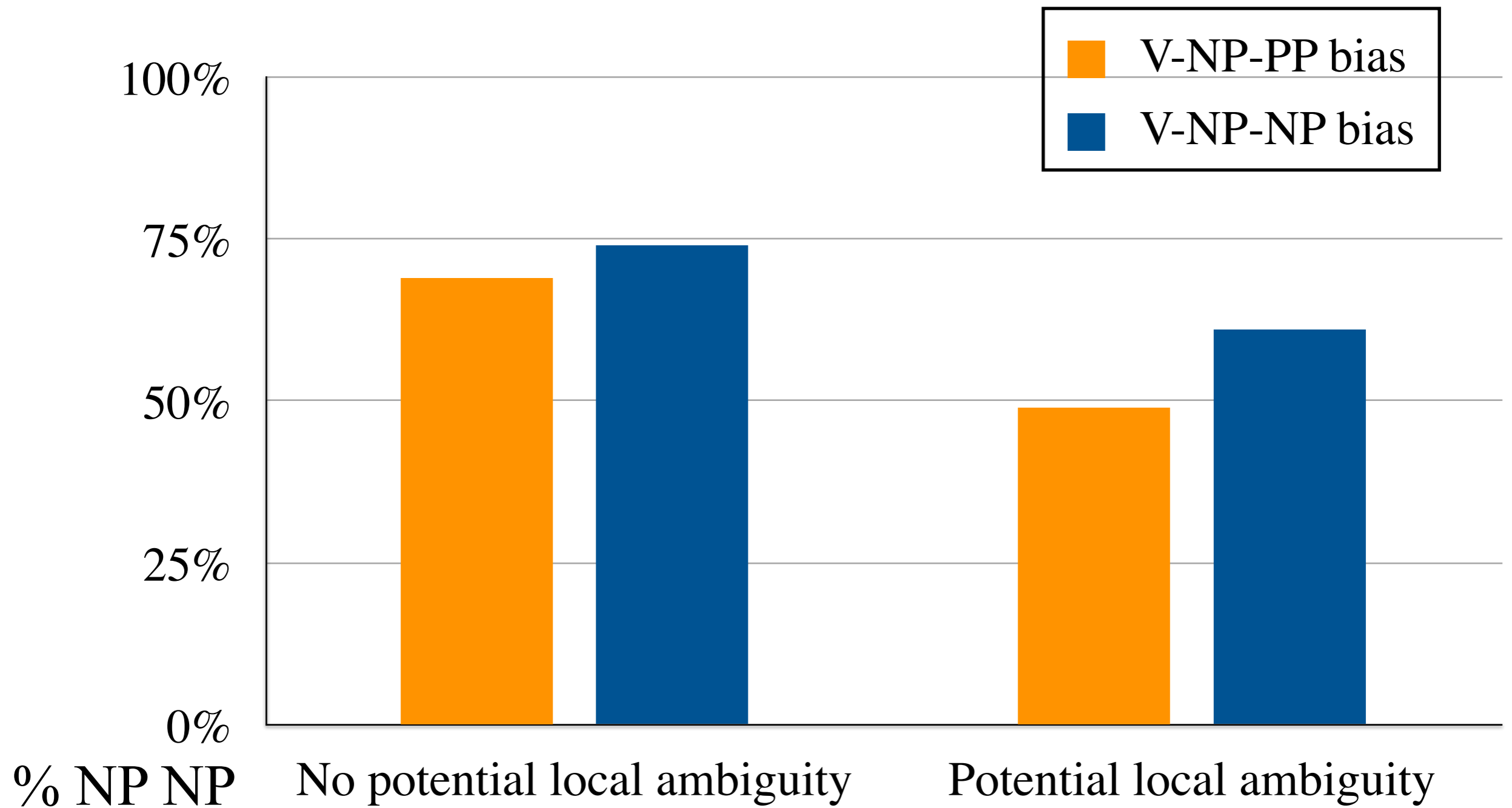
2. The sentence disappears from the screen.  
The listener reads the next question from a list.

# Experimental Method, continued



3. The speaker answers the listener's question.  
The listener chooses the correct response on a list (from two choices).

# Experimental Results on Local Ambiguity



# *Reverse ambiguity effect*

- Arnold, Wasow, Asudeh & Alrenga 2004  
*Journal of Memory & Language*
- Re-ran the experiment with slightly better methodology and found a *stronger* reverse ambiguity effect.

# A psychologically real grammar should be lexicalist

- Early generative grammars downplayed the lexicon.
- Now, however, the importance of the lexicon is widely recognized.
- This aspect of grammar has been developed in greater detail in our theory than in any other.
- It would be easy to add frequency information to our lexicon, though there is debate over the wisdom of doing so.



# Conclusion

- Grammatical theory should inform and be informed by psycholinguistic experimentation.
- This has happened less than it should have.
- Existing psycholinguistic evidence favors a constraint-based, lexicalist approach (like ours).

# Universals?

- P&P (top-down): attempts to relate multiple typological properties to single parameters.
- Grammar Matrix (bottom-up): attempts to describe many languages in a consistent framework and then takes stock of common constraints.

# Universals?

- Case constraint
- SHAC
- Binding theory
- Head-complement/-specifier/-modifier
- Head Feature Principle
- Valence Principle
- Semantic Compositionality Principle
- ...

# Overview

- Psycholinguistics and grammar design
  - What grammar has to say
  - What psychological evidence has to say
    - Acquisition
    - Production
    - Comprehension
- Universals

# Reading Questions

- How could we test empirically whether language skill is independent of other brain functions?
- As for modularity, what about aphasia and the like? I thought these disorders, caused by damage to specific parts of the brain, proved that there is at least some modularity in our production and comprehension of language.
- I am having trouble associating the features of our grammar with the broader issue of the competence / performance dichotomy (or the lack of it).

# Reading Questions

- I'm a bit confused by the discussion in section 9.5.2 about the constraint-based nature of our grammar. The section starts off by saying that our constraints are order-independent, and later argues that such order-independence is a necessary property of our constraints in order for performance-compatibility to be achieved. However, immediately after the first mention of order-independence, there is a reference to the previous discussion in this chapter on incremental processing and how the order of presentation of the words in an utterance affects how the listener comprehends it. How do you separate the constraints from the words in a plausible model? In other words, why are constraints order-independent while the way we process language is order-dependent?

# Reading Questions

- What is meant by "the partial grammatical analyses have to be keyed in to partial linguistic meanings"
- "I really like your hat... not!" (ha ha) What would that look like in HPSG online processing?
- In section 9.5.1, page 301, I'm having trouble seeing how the grammar we are developing could handle the situation described where Speaker A knows what Speaker B is asking, possibly, after Speaker B utters only one word. All of the analysis we've done so far only goes as far as a sentence with no consideration of the preceding utterance/sentence.

# Reading Questions

- Why does the fact that phrases are built from components "in a context-free manner" facilitate incremental processing? (I understand why transformational grammar won't work, but are all surface-oriented approaches necessarily context-free, and if not, why is a context-free approach more desirable for modeling incremental processing?)



# Reading Questions

- So, what are the arguments for transformational grammar in the face of the seemingly damning evidence in this chapter?
- Are you convinced that it is necessary to understand/simulate those processes to build a grammar that generates grammatical utterances and does not generate ungrammatical utterances? Is understanding/simulating those processes what will allow this grammar to get an "advantage" on something based more purely on large amounts of linguistic data?

# Reading Questions

- All the observations and experiments made about language, do we know if that holds for all languages? Do we know what languages such experiments were carried on?
- We learn a language we are exposed to growing up regardless of background (e.g. parents' language...) Why is it that humans cannot learn animal languages?

# Reading Questions

- Section 9.6 makes mention of universalizing the grammar in the book such that it applies to all human languages. Has any serious attempt been made to create a formal constraint based lexicalist universal grammar?
- The section also mentions how the grammar rules could be generalized to allow for different word orderings. How would you write a single rule to allow different orderings?

# Reading Questions

- What is major opinion about how written language represents language performance? In my syntax classes thus far we have focused on diagramming sentences that are written (but ostensibly could have been spoken). So is it our goal to parse written sentences of ideal form, without the repetition and mistakes of spoken language?

# Reading Questions

- How can we use these lexical entries? What are they for? How are they different from lexemes? Can we get an example of how the feature structure for the lexeme for "him" would look any different from the lexical entry?
- Also, on page 291 in the description of a tree structure (36) says "each terminal node is labeled by a phonological form." A terminal node is simultaneously (or could be any of) a lexical entry, a lexeme, or a word?
- I guess I don't understand what a "family of lexemes" is. Lexical entries seem way too specific (with their phonological form and synsem information) to describe an entire family of lexemes, which are still not even words themselves.