Ling 566
Nov 4, 2021

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.

Where do you currently stand wrt the hypothesis of innate UG?

Sounds plausible, no prev UG class

Sounds plausible, studied UG before

Not sure, no prev UG class

Not sure, studied UG before

Seems unlikely, no prev UG class

Seems unlikely, studied UG before

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 grammar in processing

- Evidence for top-down planning

Disfluencies are sensitive to structure:

Repeat rate of *the* varies with position and complexity of the NP it introduces:
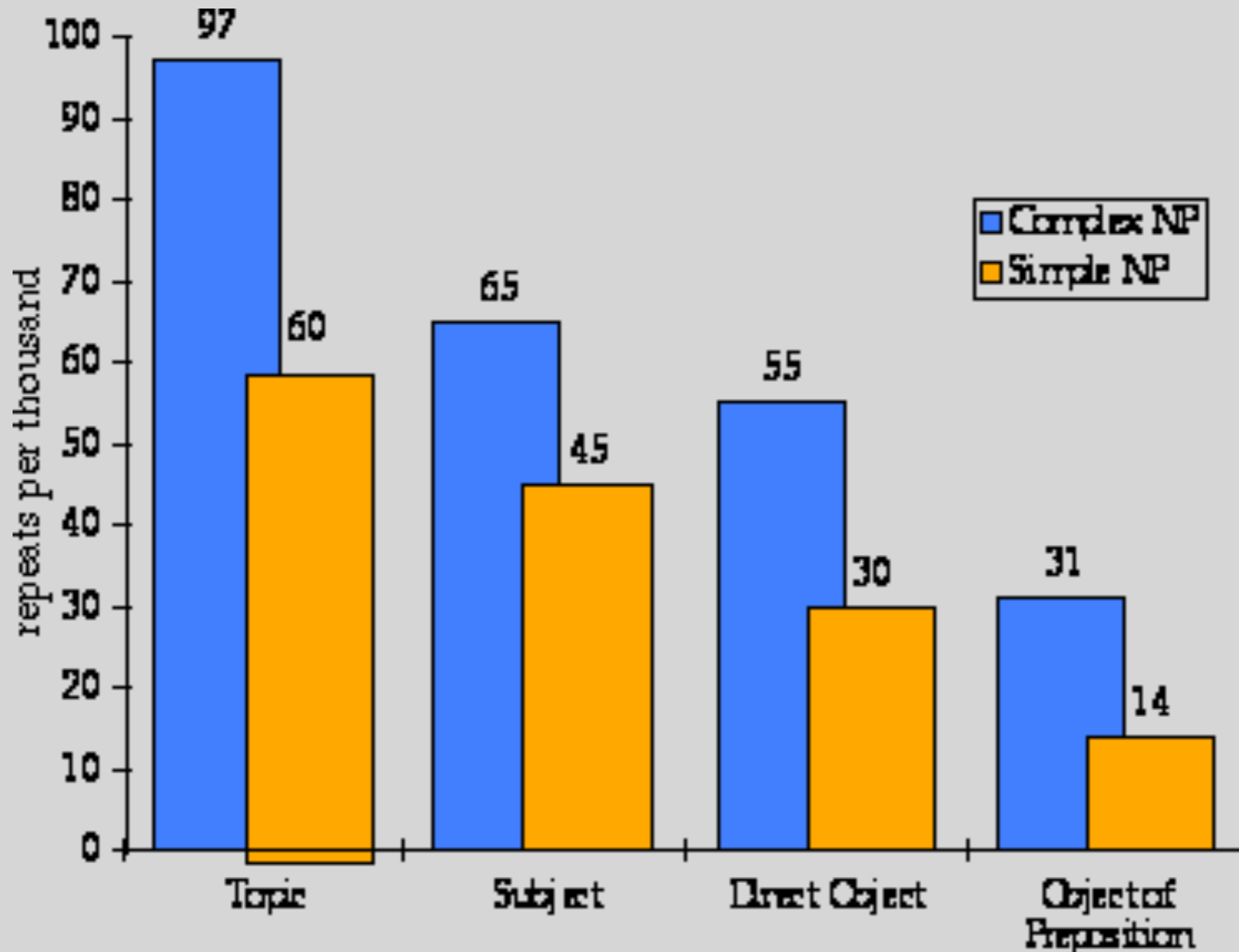
Production errors are sensitive to syntactic structure

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*

% NP NP order
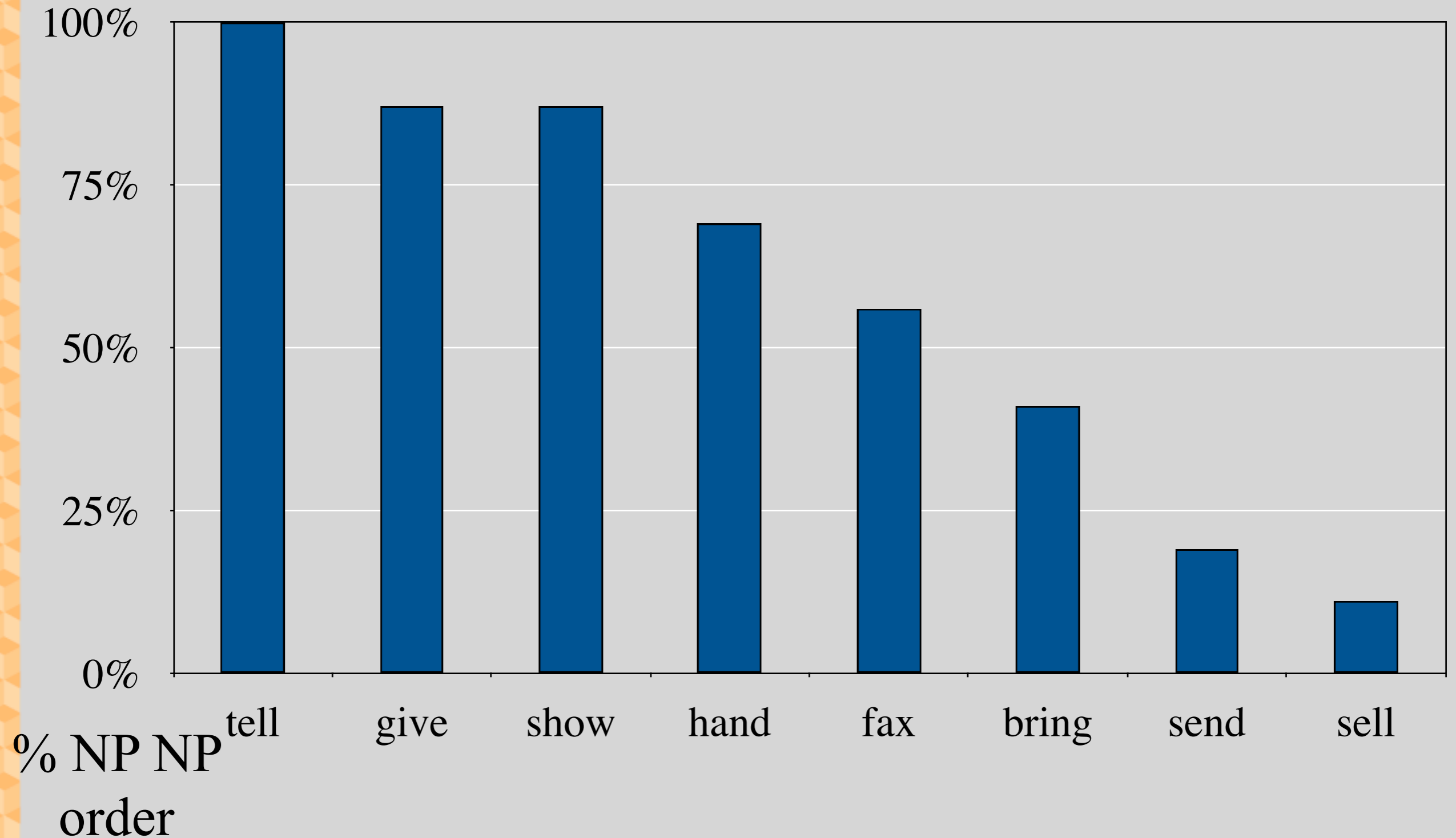
Lexical biases influence processing

- 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).
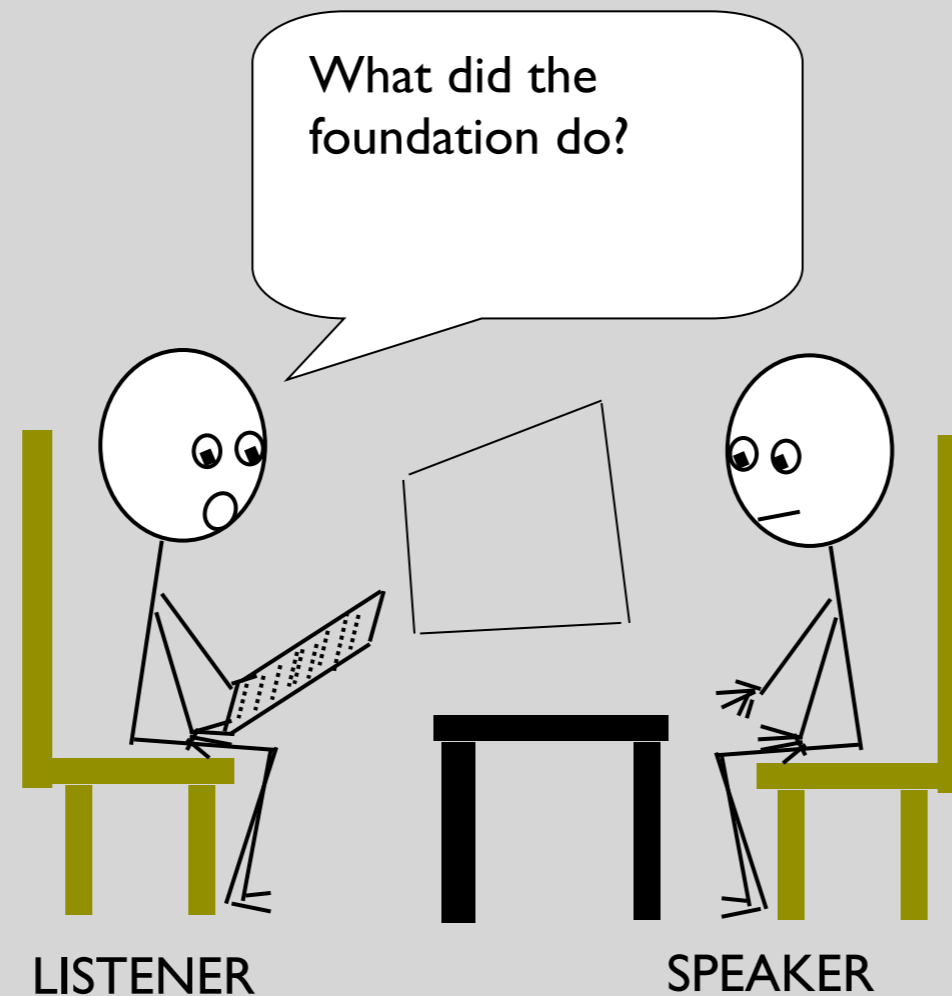
Experimental Method
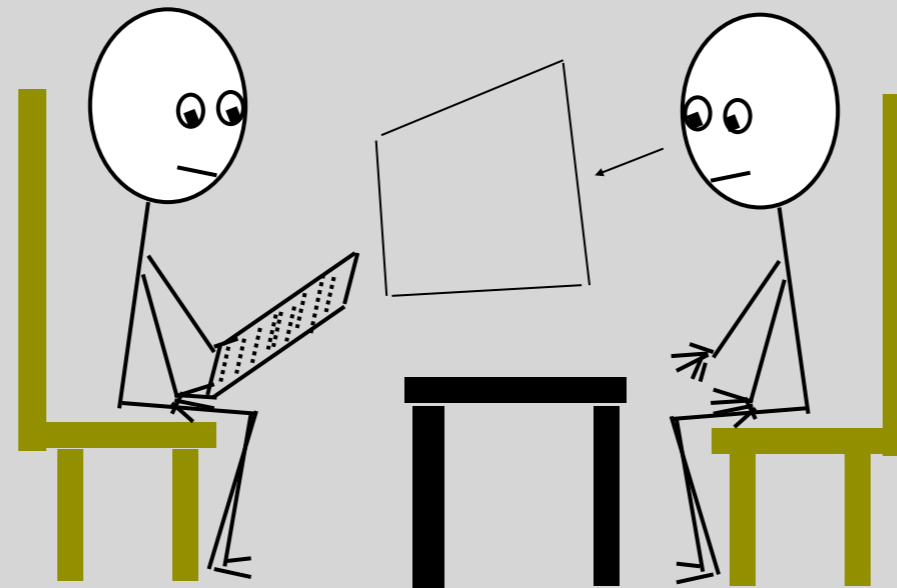
LISTENER   SPEAKER
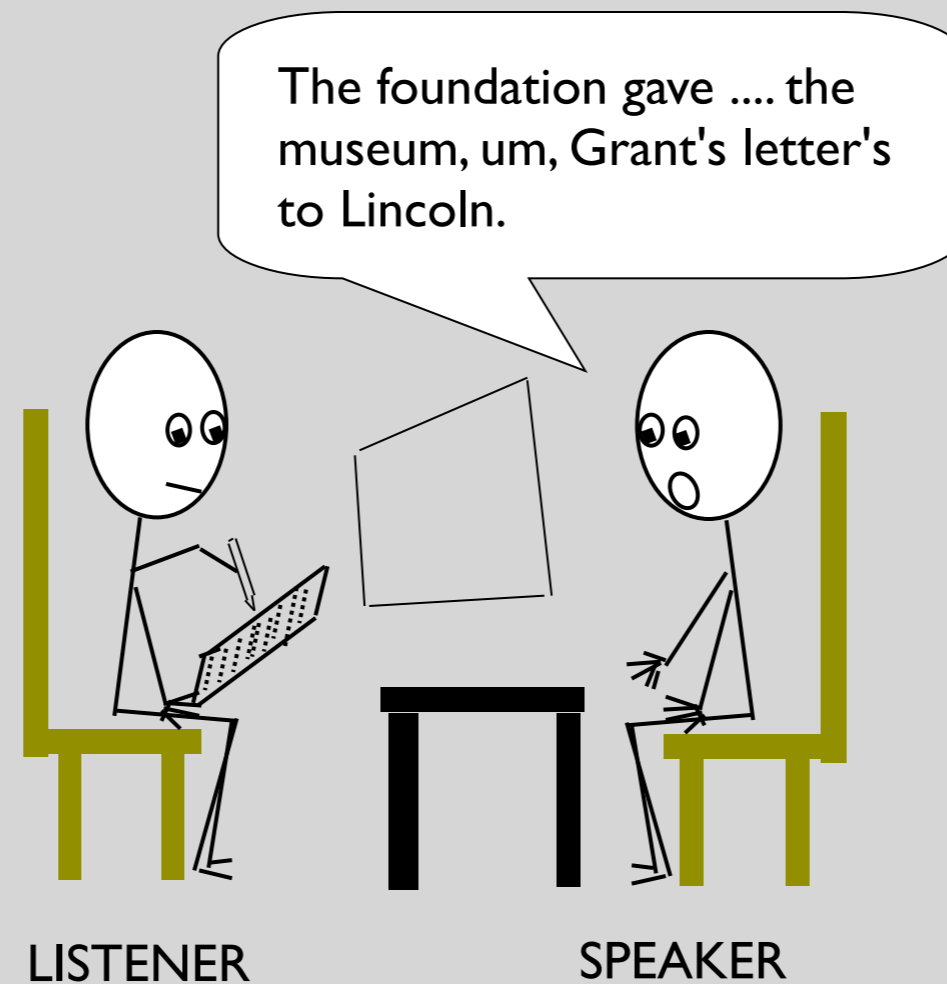
1. Speaker silently reads a sentence:

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

Experimental Method, continued

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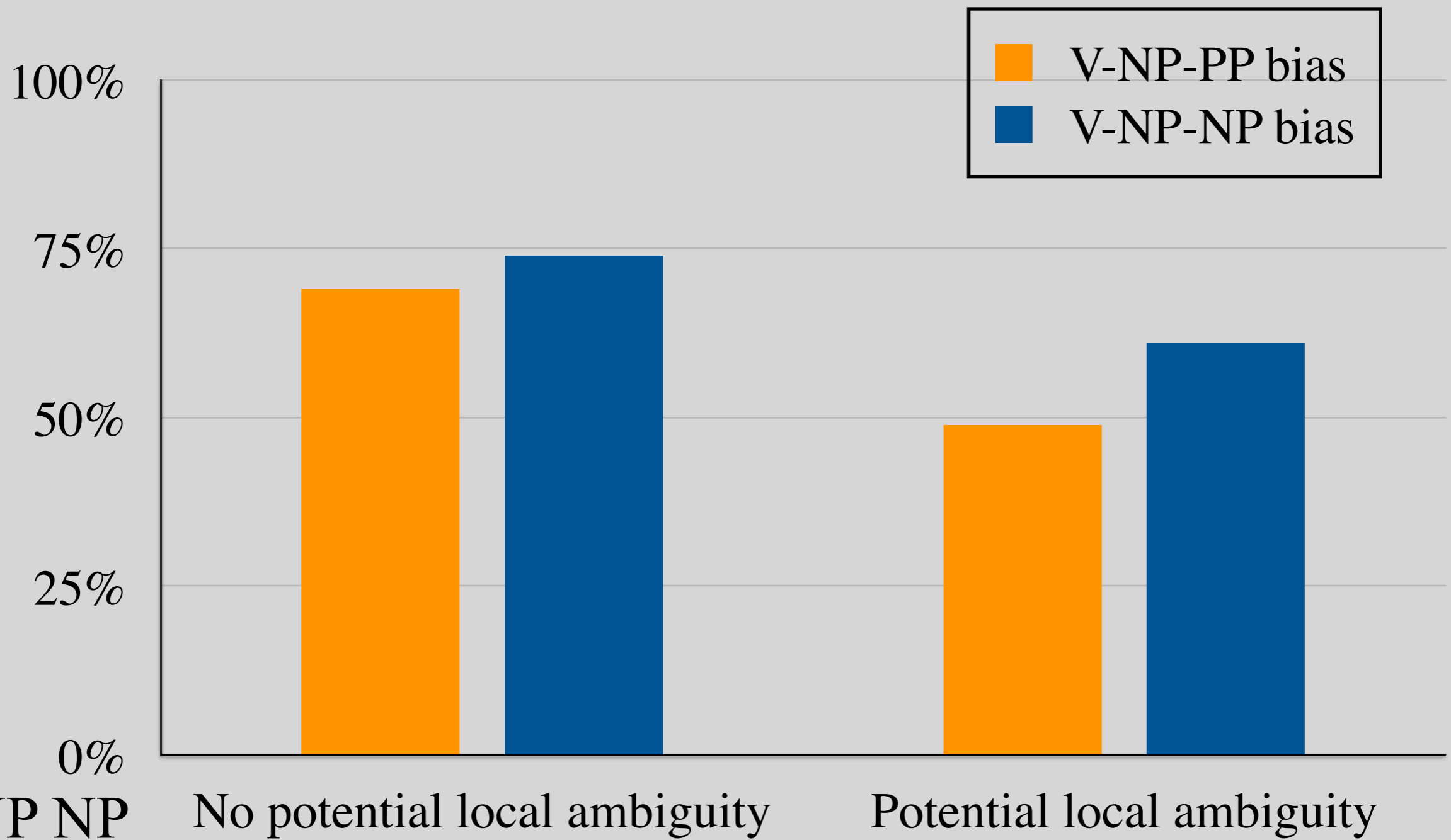   The listener reads the next question from a list.

Experimental Method, continued

The foundation gave .... the museum, um, Grant's letter's to Lincoln.

LISTENER   SPEAKER

Poll!

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

© 2003 CSLI Publications

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(-ish)): attempts to describe many languages in a consistent framework and then takes stock of common constraints.

What aspects of our grammar fragment so far seem English-specific?

What aspects of our grammar fragment so far seem like they might be crosslinguistically useful or even universal?

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

Midterm

- Posted tomorrow

- Due next Friday (11/12)

- No collaboration

- Send questions to me by email

- Check Canvas read-only midterm Q&A discussion for Q&A :)

Reading Questions

- What is the purpose of the type atom?

- When was the feature structure list [FIRST, REST] introduced to the type hierarchy? How does it work?

- How does it relate to list(tau)?

- How does it relate to l-sequence?

	$\Big\langle$ $\Big[$ SYN $\quad$ $\boxed{}$ $\Big]$ $\Big\rangle$ $\Big]$	
list		*feat-struc*
$list(\tau)$	$\begin{bmatrix} \text{FIRST} & \tau \\ \text{REST} & list(\tau) \end{bmatrix}$	*list*
l-sequence	$\begin{bmatrix} \text{FIRST} & atom \\ \text{REST} & \langle word \rangle \mid \langle lexeme \rangle \end{bmatrix}$	*list*
synsem	$\begin{bmatrix} \text{SYN} & syn\text{-}cat \\ \text{SEM} & sem\text{-}cat \end{bmatrix}$	*feat-struc*

Reading Questions

- On page 307 the book mentions that "it would not be hard to factor out the ordering (of rules) so that versions of these rules could be posited as part of universal grammar." I am curious what rules without an ordering would look like.

Reading Questions

- What would the alternative of constraint-based would be? What would it mean for an operation to "destructively modify" a representation?

- On p295, one of the basic observations of our grammar is strongly lexicalist. Since a great part of the constraints information is stored in the lexical entry, I'm starting to wonder what is the boundary of the constraint. Are the constraints only within one sentence? This might be beyond syntax, but is there any way to make constraints for more than one sentence?

Reading Questions

- I am confused with Surface Orientation saying our grammar is quite compatible with how a speaker does the "partial" linguistic analyses to interpret a complete utterance. If I have understood this right, then I would like to know how we know if a speaker's partial interpretation and our grammar pick the same increments make a complete interpretation. How can we ensure that our grammar represents the same "parts" for a complete interpretation as a speaker would?

Reading Questions

- Of the modern grammatical frameworks, does HPSG have the strongest integration of semantic features due to its emphasis on strong lexicalism? Are there cases where the grammar's emphasis on strong lexicalism (as opposed to complex syntax rules) is disadvantageous?

Reading Questions

- I was really interested by the psycholinguistic studies mentioned in this chapter that discuss the increased frequencies of disfluencies such as um and uh when referring to complex situations or situations that require more planning. On page 299, the study even goes as far as to discuss which of the two words occurs more frequently in certain situations. Is this type of evidence of psycholinguistic analysis enough to posit rules in our grammar for such disfluencies? Does it warrant rules?

Reading Questions

- If we want our grammar to reflect the process going on in our minds, how important is it for our grammar to make accurate predictions about common syntactic errors? Does our current grammar predict common errors? If so, how?

- In any case going back to the idea that our speech is like predictive text. It's almost like our speech really is like that except our brain is a model trained on every utterance we've ever heard, every thought we've ever had, and of course our knowledge of the language. If that is the case, could it be that theoretically all we need to model a person's speech is a record of all sentences they've heard, their thoughts, and a very good model of English?

Reading Questions

- How do we comprehend semi-structured language (for instance spoken by L2 speakers that are not very fluent with the language) or ambiguous sentences and are still able to make sense of it?

- I do not understand what footnote 16 is trying to say, why would listeners who allow grammatical principles to be relaxed would understand more?

Reading Questions

- Why does it become difficult to learn a second language after a certain age? How does this relate to language learnability in general?

Reading Questions

- What has happened in the last 20 years of psycholinguistics?