

# HG4041 Theories of Grammar

## Grammar and Processing

Francis Bond

**Division of Linguistics and Multilingual Studies**

<http://www3.ntu.edu.sg/home/fcbond/>

[bond@ieee.org](mailto:bond@ieee.org)

Lecture 7

Location: HSS SR3

HG4041 (2013)

# Schedule

---

Lec.	Topic		Reading	Problems
1	Introduction (HPSG)		SWB 1–2	1:1
2	Feature Structures		SWB 3	3:1, 3
3	Complex Feature Values		SWB 4	4:1, 5, 6
4	Semantics		SWB 5–6	5:1; 6:1, 3, 4, 5
5	Binding		SWB 7	7:1, 2
6	The Structure of the Lexicon	<b>Mid-term</b>	SWB 8	8:1, 2, 6
7	Realistic Grammar		SWB 9	9:1
8	Passive		SWB 10	10: 1, 3
9	Dummies and Idioms		SWB 11	11:1, 3, 4
10	Raising and Control		SWB 12	12:1, 2, 4, 6
11	Long Distance Dependencies	<b>Final</b>	SWB 14	14: 1, 2, 3
12	Wrap-up	<b>Project Presentations</b>	SWB 16	
	<b>Research Paper</b>			
		due two weeks after presentations		

# 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:

- More common in front of complex NPs
- More common with prominent NPs  
Topic > Subject > Direct Object > Preposition Object

(1) The *the book I told the student about on the train*

## Production errors are sensitive to syntactic structure

---

- Agreement errors are more common with PP complements than sentential complements: errors like (3) are significantly more common than errors like (2).

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

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

- 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

---

(4) *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”

(5) *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.

## 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)
- “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.

## Evidence for the Psychological Reality of Deep Structures

---

- The proposed Deep Structure for (7) had three occurrences of *detective*, while the proposed DS for (6) had only two:
  - (6) *The governor asked the detective to prevent drinking.*
  - (7) *The governor asked the detective to cease drinking.*
- In a recall experiment, *detective* was significantly more effective in prompting people to remember (7) than (6).

## Typical Problem Cases for the DTC

---

- (8) *Pat swam faster than Chris swam.*
  - (9) *Pat swam faster than Chris did.*
  - (10) *Pat swam faster than Chris.*
- The DTC predicts that (8) should be less complex than (9) or (10), because (9) and (10) involve an extra deletion transformation.
- In fact, subjects responded more slowly to (8) than to either (9) or (10).



# 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 (11) often temporarily misinterpreted *on the towel* as the destination.

(11) *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.

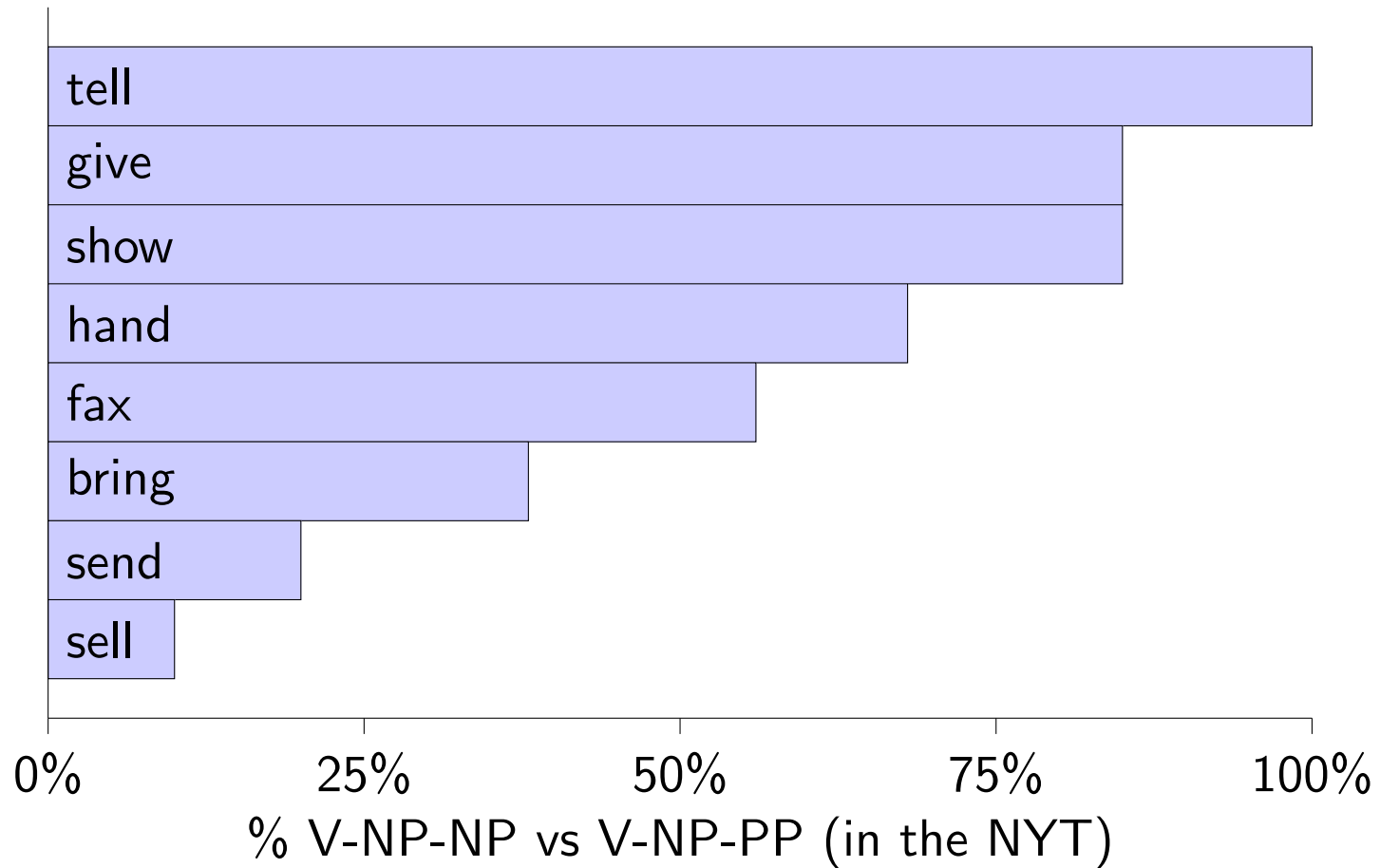
## Speakers know a great deal about individual words

---

- 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.

## Different verbs favor different COMPS lists

---





## Lexical biases influence processing

---

- Wasow et al. ran a production experiment to test whether ambiguity avoidance would influence speakers' choice between (12) and (13):

(12) *They gave Grant's letters to Lincoln to a museum.* NP-PP

(13) *They gave a museum Grant's letters to Lincoln.* NP-NP

- Ambiguity avoidance predicts that you should prefer (13)
- 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).

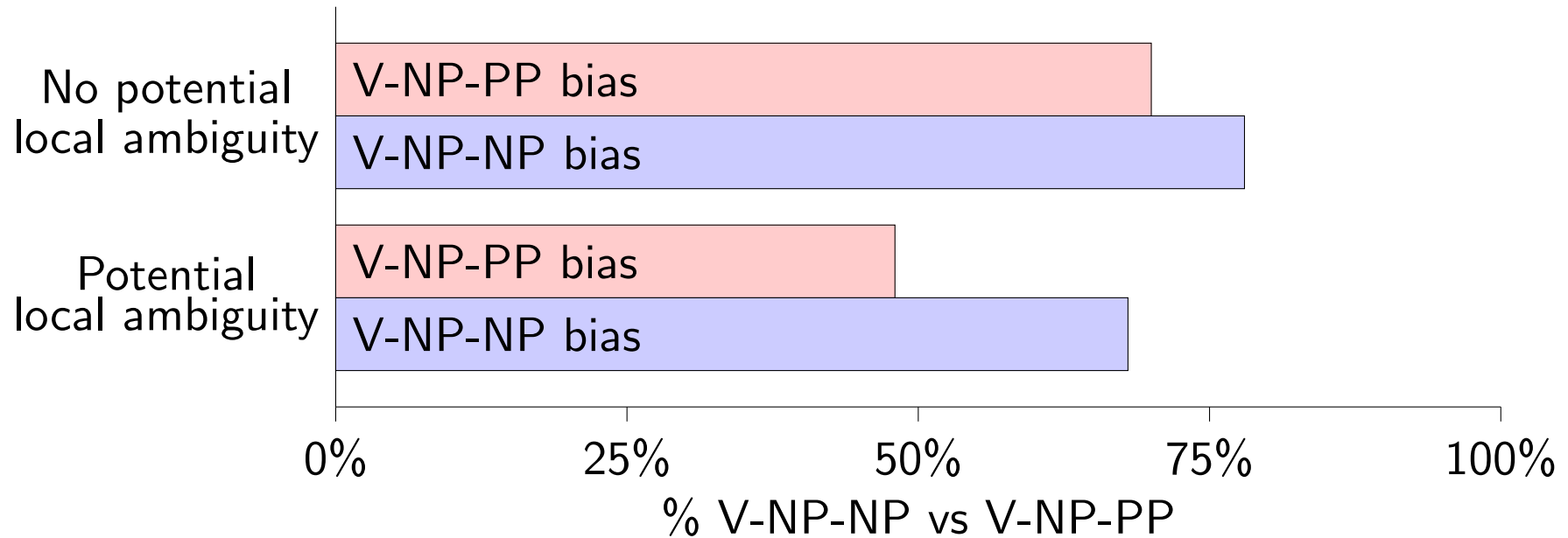
## Experimental Method

---

- Speaker and Listener sit next to each other. Speaker can see a screen.
- Speaker silently reads a sentence shown on the screen  
*A museum in Philadelphia received Grant's letters to Lincoln from the foundation.*
- The sentence disappears from the screen.
- Listener asks a question:  
*What did the foundation do?*
- The speaker answers the listener's question.  
*The foundation gave . . . . the museum, um, Grant's letter's to Lincoln.*
- The listener records which kind of response on a list (from two choices).

## Experimental Results on Local Ambiguity

---



- Arnold, Wasow, Asudeh & Alrenga (2004: *Journal of Memory & Language*) re-ran the experiment with slightly better methodology and found an even 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 HPSG than in any other theory.
- It would be easy to add frequency information to the lexicon, though there is debate over the wisdom of doing so.
- Frequency is currently recorded as part of the **parse ranking models** which select the most plausible out of all possible interpretations.

## 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 HPSG).

# Universals?

---

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

# 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



## P1: Constant Rules

---

The Singular Noun Lexical Rule, the Non-3rd-Singular Verb Lexical Rule, and the Base Form Lexical Rule are all inflectional lexical rules (*i-rule*) which have no effect on the shape (i.e. the phonology) of the word.

- A. Explain why we need these rules anyway.
- B. Each of these rules have lexical exceptions, in the sense that there are lexemes that idiosyncratically don't undergo them. Thus, there are some nouns without singular forms, verbs without non-third-person singular present tense forms, and verbs without base forms. List any you can think of.