Chapter 9

And now for something completely different…..

Grammar & Psychology
A Little Background: Transformational Grammar

Early arguments against using CFG for natural language concerned sentences in which some constituent seemed to be out of place or missing.

- Passives: No NP complement in *The steak was devoured*, even though *devour* usually requires one.

- Questions: In *What did the dragon devour?* the complement of *devour* precedes the subject, as does a verb.

- Elliptical constructions: Most of a VP is missing in *I will put the book on reserve, if I can _____.*
The Transformational Approach

- Every sentence is associated with a sequence of trees.
- The first one (sometimes called “deep structure”) has all the displaced or missing constituents in their canonical locations.
- Rules called “transformations” permute, delete, and insert elements in trees, to get the observable forms (“surface structures”) from the deep structures.
- The lexicon can just specify the pre-transformational contexts for words, which is simpler.
- Semantic relations are more transparent at deep structure.
- CAVEAT: This is an oversimplified and dated characterization of TG
For Example

• Actives/passive pairs were originally derived from the same deep structure.

  \textit{The dog chased the cat} ⇒ \textit{The cat was chased by the dog}

• Likewise declaratives and questions.

  \textit{You will do what} ⇒ \textit{What you will do} ⇒ \textit{What will you do}

• Some material was deleted under identity

  \textit{I will taste the mushroom if you will taste the mushroom} ⇒ \textit{I will taste the mushroom if you will} ∅.
More Caveats...

- The earliest TG work (ca. 1955-1963) was formulated quite precisely, so the consequences of analyses could be tested empirically.
- Subsequent versions of TG have become less and less explicit, making it very hard to compare theories these days.
- I’m indicating the effects of transformations by showing strings of words, but the rules have always been conceived of as operations on trees.
- Early TG formulated complex, language-specific tree operations.
- More recent work has assumed very general rules, with their operations constrained by general principles, interacting with lexical information.
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.
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.
People plan their utterances to some extent, but typically don’t have sentences fully formulated before they start uttering.

This is evident from the existence of disfluencies, such as *uh*, *um*, repetitions, false starts, etc.
because you see I, uh, some of our people, [pause and clears throat] who are doing LEss, uumm, have to consider which paper [pause] to do,
Disfluencies are sensitive to structure:

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

<table>
<thead>
<tr>
<th>Position</th>
<th>Complex NP</th>
<th>Simple NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>97</td>
<td>60</td>
</tr>
<tr>
<td>Subject</td>
<td>65</td>
<td>45</td>
</tr>
<tr>
<td>Direct Object</td>
<td>55</td>
<td>30</td>
</tr>
<tr>
<td>Object of Preposition</td>
<td>31</td>
<td>14</td>
</tr>
</tbody>
</table>
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.
Some high-level sentence planning is necessary, too

• *Ich habe dem Mann, den ich gesehen habe geholfen.*
  I have the-dat man whom-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 whom-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 idiosyncracies 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

- tell: 100%
- give: 100%
- show: 100%
- hand: 75%
- fax: 50%
- bring: 25%
- send: 0%
- sell: 0%
Lexical biases influence processing

- We 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).
1. Speaker silently reads a sentence:

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

The foundation gave .... the museum, um, Grant's letter's to Lincoln.
Experimental Results on Local Ambiguity

No potential local ambiguity

Potential local ambiguity

V-NP-PP bias
V-NP-NP bias
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).