Quick Recap

- Frazier: early parsing decisions driven by purely syntactic heuristics
- MA and LC were argued to be by-products of a race mechanism
- Eager dependency-formation plays a strong role in driving parsing decisions:
  - Pritchett’s theta-attachment
  - Local coherence trumps global syntactic parsing constraints
  - Active-Filler Hypothesis
  - Pickering & Barry’s Dependency Association account
Well-known local ambiguities

NP/VP Attachment Ambiguity:

“The cop [saw [the burglar] [with the binoculars]]”
“The cop saw [the burglar [with the gun]]”

NP/S Complement Attachment Ambiguity:

“The athlete [realised [his goals]] last week”
“The athlete realised [[his goals] were unattainable]”

Clause-boundary Ambiguity:

“Since Jay always [jogs [[a mile] doesn’t seem very long]]”
“Since Jay always jogs [[a mile] doesn’t seem very long]”

Reduced Relative-Main Clause Ambiguity:

“[[The woman [delivered the junkmail on Thursdays]]”
“[[The woman [delivered the junkmail]] threw it away]”

Relative/Complement Clause Ambiguity:

“The doctor [told [the woman] [that he was in love with her]]”
“The doctor [told [the woman [that he was in love with]] [to leave]]”

Deterministic Parsing

- Alternative: ensure only one possible parsing action at any point in parsing:
  - Avoid/delay rule selection until it is fully determined. How?
    - bottom up (e.g. S/R or LR) plus lookahead [Marcus; Berwick & Weinberg; Abney]

- Advantage: very fast, clear predictions

- Disadvantages:
  - not fully incremental (up to 3 constituents of look-ahead)
  - unsuccessful for head-final languages
  - wrong predictions: e.g. no gradedness of processing difficulty
Monotonic Parsing

- Inspired by determinism and reanalysis:
  - many local ambiguities seem to cause little difficulty
  - contra predictions of naive backtracking
  - Gorrell; Weinberg; Sturt & Crocker
- Provides a richer set of ‘tree-building’ operations which means destructive backtracking is not always required
- Predicts ‘reanalysis’ outwith these operations to be difficult

Talking about Talking about Trees

- Traditional theories locate reanalysis in the parser: “re-parsing”
- Can local ambiguity be handled using underspecified representations?
  - Representations which allow some ambiguity to remain, and be later removed without (destructive) re-parsing.
- Description-Theory: (Marcus, Hindle & Fleck, 1983)
  - Uses tree descriptions, not trees: e.g. dominance and precedence
  - Permits immediate interpretation, but allows insertion of nodes & branches
Monotonic Parsing  [Gorrell; Sturt & Crocker]

- Trees are described as a set of nodes, and a set of precedence and dominance relations:
  
  - \textit{John knows Mary}

  \[
  \begin{array}{c}
  \text{S} \\
  \text{NP}_1 \ \\ 
  \text{VP} \\
  \text{John} \\
  \text{V} \ \\ 
  \text{NP}_2 \\
  \text{knows} \ \\ 
  \text{Mary}
  \end{array}
  \]

  \{\text{dom}(S,\text{NP}_1), \text{dom}(S,\text{VP}), \text{dom}(S,\text{V}), \text{dom}(S,\text{NP}_2), \text{prec}(\text{NP}_1,\text{VP}), \text{dom}(\text{VP},\text{V}), \text{dom}(\text{VP},\text{NP}_2), \text{prec}(\text{V},\text{NP}_2) \ldots}\}

Properties of Trees

- Single root condition: a single root node dominates all nodes
  \[\exists x \forall y \cdot \text{dom}(x, y)\]

- Exclusivity condition: no two nodes can stand in dom & prec relations
  \[\forall x, y \cdot \text{prec}(x, y) \lor \text{prec}(y, x) \iff \neg \text{dom}(x, y) \land \neg \text{dom}(y, x)\]

- Inheritance: nodes inherit precedence properties of their ancestors
  \[\forall w, x, y, z \cdot \text{prec}(x, y) \land \text{dom}(x, w) \land \text{dom}(y, z) \rightarrow \text{prec}(w, z)\]

- \text{dom} and \text{prec} are transitive relations

- \text{dom} is reflexive, \text{prec} is irreflexive
Constraints on the Model

- **Strict incrementality**: words are connected to the tree description as they are encountered.

- **Coherence**: tree properties must always be satisfied, the tree must be grammatically licensed.

- **Full specification of nodes**: no features on nodes (e.g. bar-level) can be left unspecified.

- **Informational monotonicity**: the tree description at state $n$ is a subset of the description at state $n+1$.

- **Obligatory assertion of precedence**: precedence must be specified for sisters.

Monotonic Parsing and Reanalysis

- **Easy**: Monotonic reanalysis, “John knows Mary is smart”

- **Hard**: Non-monotonic reanalysis, “While John walked the dog barked”

- How does the parser actually work?
  - The monotonic parser uses precomputed tree descriptions (e.g. for lexical items), and “macro” operations which meet the constraints.
  - Non-monotonic operations are not permitted.
  - Easy reanalysis parsing is “monotonic”.
  - Difficult reanalysis, when tree-descriptions are changed non-monotonically.
Monotonic Parsing Operation

- Tree-lowering: “John knows Mary …”       “John knows Mary is smart”

\[
\begin{align*}
\text{NP1} & \quad \text{VP} \\
\text{John} & \quad \text{knows} \quad \text{NP2} \\
\text{VP} & \quad \text{NP2} \\
\text{Mary} & \quad \text{is smart}
\end{align*}
\]

\{\text{dom}(S,\text{NP1}), \text{dom}(S,\text{VP}), \text{dom}(S,V), \text{dom}(S,\text{NP2}), \text{prec}(\text{NP1},\text{VP}), \text{dom}(\text{VP},V), \text{dom}(\text{VP},\text{NP2}), \text{prec}(V,\text{NP2}), \text{dom}(\text{VP},S_2), \text{dom}(S_2,\text{NP2}), \text{prec}(\text{NP2},\text{VP2}) \ldots\}\}

Theta-Reanalysis: Easy

- Reanalysis to a position within the original theta-domain is easy.
Non-Monotonic Parsing

• Predicting difficult reanalysis: While John walked the dog … barked.

Theta-Reanalysis: Difficult

• Reanalysis to a position outside the original theta-domain is difficult.
Parsing Operations: Attachment

• Left attachment:

• Right attachment:

TAG Adjunction

• The operations of the monotonic parser resemble those of Tree Adjoining Grammar.
Tree Lowering

Psycholinguistic Evidence

- Are there really two types of reanalysis?
  - **NP/S (A):** “The woman saw the famous doctor had been drinking”
  - **NP/Z (A):** “Before the woman visited the famous doctor had been drinking”
  - **NP/S (U):** “The woman saw that the famous doctor had been drinking”
  - **NP/Z (U):** “Before the woman visited, the famous doctor had been drinking”

- All verbs are biased (BNC) towards NP complement
  - To make sure the object attachment is initially adopted, forcing reanalysis

- Plausibility of the direct object analysis is similar (pre-test).

Sturt, Pickering & Crocker, JML, 1999
Results

- Reading times:
  Region 3

```
“The woman saw the famous doctor had been drinking all day”
```

- Main effects of construction type, ambiguity, and a significant interaction
- GP effect: NP/Z (400ms) vs. NP/S (87ms)

Search in Parsing/Reanalysis

- (Some) reanalysis is simply monotonic attachment: e.g. tree-lowering.
- What if there are multiple such “lowering” attachments
- Consider a double NP/S ambiguity:
  - “I know the man who believes the countess killed herself”
  - “I know the man who believes the countess killed himself”
- Which is easier?
- How does the parser search for an attachment?
The Trees

Search

- English appears to use a bottom-up search strategy for attachment
  - late closure, recency
  - *Someone shot the servant of the actress who was on the balcony*

- Possibly other influences:
  - Verb bias? Predicate proximity (Gibson)?
  - Japanese seems to be top-down
    - Head final, left-branching language (Sturt & Crocker, 1996)
Summary of Reanalysis

• Frazier: no clear account

• Pritchett: cost determined by syntactic nature of reanalysis (TRC)

• Monotonic Parsing:
  • Representations allow for some kinds of local ambiguity
    • Some reanalysis is monotonic (easy), some is destructive (difficult)
  • Similar in some respects to Pritchett's theory (of reanalysis)
  • Search mechanism still required when multiple reanalyses is possible

Summary of Syntactic Models

• Syntactic Parsing Theories:
  • Frazier: emphasis on syntactic structure/form
  • Pritchett: emphasis on syntactic dependencies/content

• Assume serial, incremental parsing. Reanalysis causes difficulty

• Preference to associate fillers with role-assigners immediately

• Monotonic models enable some local ambiguities to be revised without destructive reanalysis
  • distinguish easy and difficult “garden paths”
A Puzzle

• Sometimes local thematic assignment appears to violate global parse:
  
  • The coach smiled at the player tossed a frisbee by the ...
  
  • The coach smiled at the player thrown a frisbee by the ...
  
  • The coach smiled at the player who was tossed a frisbee by the ...
  
  • The coach smiled at the player who was thrown a frisbee by the ...


A Puzzle

• Sometimes local thematic assignment appears to violate global parse:

  • [A/R] The coach smiled at the player tossed a frisbee by the ...
  
  • [U/R] The coach smiled at the player thrown a frisbee by the ...
  
  • [A/U] The coach smiled at the player who was tossed a frisbee by the ...
  
  • [U/U] The coach smiled at the player who was thrown a frisbee by the ...

• Do people consider the locally coherent, but not globally licensed parse?
  
  • This structure shouldn’t even be considered by incremental parsers
A Puzzle

• Sometimes local thematic assignment appears to violate global parse:

  • [A/R] The coach smiled at the player tossed a frisbee by the ...
  • [U/R] The coach smiled at the player thrown a frisbee by the ...
  • [A/U] The coach smiled at the player who was tossed a frisbee by the ...
  • [U/U] The coach smiled at the player who was thrown a frisbee by the ...

• We might expect to see:

  • Main effect of verb ambiguity: if ambiguous verbs are difficult
  • Main effect of structure ambiguity: if ambiguous RRCs are difficult

Results:

Table 1: Word-by-word significances for the Ambiguity x Reduction interaction of Experiment 1. IS = Interaction Strength = (HR –HU) – (NR – NU).

<table>
<thead>
<tr>
<th></th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>smiled</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>player</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tossed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>frisbee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>by the</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F1(1,39)</td>
<td>&lt; 1</td>
<td>2.467</td>
<td>1.264</td>
<td>&lt; 1</td>
<td>6.591*</td>
<td>3.320</td>
<td>1.874</td>
<td>4.389*</td>
<td>1.227</td>
</tr>
<tr>
<td>MSE</td>
<td></td>
<td>35996</td>
<td>23706</td>
<td>-</td>
<td>52890</td>
<td>58689</td>
<td>40456</td>
<td>20317</td>
<td>50012</td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>0.117</td>
<td>0.261</td>
<td>-</td>
<td>0.010</td>
<td>0.069</td>
<td>0.171</td>
<td>0.037</td>
<td>0.268</td>
</tr>
<tr>
<td>F1(1,19)</td>
<td>&lt; 1</td>
<td>1.693</td>
<td>0.370</td>
<td>&lt; 1</td>
<td>4.804*</td>
<td>0.820</td>
<td>0.570</td>
<td>2.200</td>
<td>2.000</td>
</tr>
<tr>
<td>MSE</td>
<td></td>
<td>53806</td>
<td>28829</td>
<td>-</td>
<td>76359</td>
<td>75369</td>
<td>58585</td>
<td>31091</td>
<td>62238</td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>0.194</td>
<td>0.543</td>
<td>-</td>
<td>0.029</td>
<td>0.365</td>
<td>0.447</td>
<td>0.138</td>
<td>0.157</td>
</tr>
<tr>
<td>IS (ms)</td>
<td>16</td>
<td>35</td>
<td>15</td>
<td>-23</td>
<td>73</td>
<td>56</td>
<td>31</td>
<td>41</td>
<td>-42</td>
</tr>
</tbody>
</table>

Notes: **p < .05, ***p < .01, ****p < .001.

But: they found an interaction!

• Implies that an “impossible” parse influences the verb’s difficulty
These results are problematic for theories requiring global contextual consistency (e.g. Frazier, 1987; Gibson, 1991, 1998)