Quick Recap: Syntactic Accounts

- **Frazier**: early parsing decisions driven by purely syntactic heuristics
- **MA** and **LC** were argued to be by-products of a race mechanism
- **Pritchett**:
  - **Theta-attachment** determines preferred parse
  - **Theta-reanalysis** constraint determines when reanalysis of difficult
- Eager dependency-formation plays a strong role in driving parsing decisions:
  - What about **long-distance dependencies**?
Long Distance Dependencies

- Wh-Fillers:
  - Who \(i\) did Fred tell Mary \(e_i\) left the country? dispreferred
  - Who \(i\) did Fred tell \(e_i\) Mary left the country? preferred

- Subject-Relative preference:
  - I met the man \(i\) that John likes \(e_i\). dispreferred
  - I met the man \(i\) that \(e_i\) likes John. preferred

- Active Filler Strategy: (“Gap as a first resort”)
  - When a filler has been identified, rank the possibility of a assigning it to a gap above all other options.

Further observations ...

- Filled-Gap effect:
  - My brother wanted to know who \(i\) Ruth will bring (*\(e_i\)) us home to \(e_i\) at Christmas
  - My brother wanted to know if Ruth will bring us home to Mom at Christmas

- Found an increased reading time at us, interpreted this as surprise

- Intuitively easy:
  - Who \(i\) \((e_i)\) did you want \((e_i)\) Mother to bake \((e_i)\) a cake for \(e_i\)?
  - ... despite 3 possible earlier gap locations
Gaps versus Dependencies

• Consider:

  • [In which tin]i did you put the cake ei?  
    Gap account

  • [In which tin]i did you put the cake?  
    Dependency account

• If keeping the filler in memory causes difficulty, we can compare:

  [In which tin]i did you put the cake that your little sister baked for you ei?  
  Easy

  [Which tin]i did you put the cake that your little sister baked for you in ei?  
  Hard

• Intuitive support for the dependency account, and against gaps.

Other evidence

• Implausibility detected immediately at the verb, as shown by increased reading times.

  • That's the [pistol/garage]i with which the heartless killer shoti the hapless man ei yesterday afternoon.

• Garnsey et al (1998) found an N400 at the verb, for the implausible condition

  • The businessman knew which [customer/article]i the secretary called ei at home

• Pickering and Barry (1996) argued that a dependency-based account was preferable to a trace-based account.
Parsing in 2 dimensions

- Gaps don’t exist in the input, so we needn’t wait until they are found
- We can associate a filler & gap as soon as the structure licenses it:

Consider: *Den Hund* sah Maria eji eji.

Deterministic Parsing

- Alternative to reanalysis accounts: 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:

  - *John knows Mary*

    ![Tree Diagram]

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

  \[ \forall w, x, y, z \cdot \text{prec}(x,y) \land \text{dom}(x,w) \land \text{dom}(y,z) \rightarrow \text{prec}(w,z) \]

- *dom* and *prec* are transitive relations

- *dom* is reflexive, *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”

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

\[\{\text{dom}(S,\text{NP1}), \text{dom}(S,\text{VP}), \text{dom}(S,\text{V}), \text{dom}(S,\text{NP2}), \text{prec}(\text{NP1},\text{VP}), \text{dom}(\text{VP},\text{V}), \text{dom}(\text{VP},\text{NP2}), \text{prec}(\text{V},\text{NP2}), \text{dom}(\text{VP},S2), \text{dom}(S2,\text{NP2}), \text{prec}(\text{NP2},\text{VP2}) \ldots}\]

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.
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]] [the race 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]]”
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”

<table>
<thead>
<tr>
<th></th>
<th>NP/S (a)</th>
<th>NP/S (u)</th>
<th>NP/Z (a)</th>
<th>NP/Z (u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>700</td>
<td>350</td>
<td>1,400</td>
<td>1,050</td>
</tr>
<tr>
<td>duration</td>
<td>(ms)</td>
<td>(ms)</td>
<td>(ms)</td>
<td>(ms)</td>
</tr>
</tbody>
</table>

- Main effects of construction type, ambiguity, and a significant interaction

- GP effect: NP/Z (400ms) vs. NP/S (87ms)
(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”