Chunks and Dependencies and Human Sentence Processing

Steven Abney
University of Tübingen
What are Chunks

- Introspection about sentence processing in reading
  
  *the old man in the park was feeding the pigeons*

  ![Sentence Diagram]

  - Not constituents
• Prosody

[Diagram of prosody with peaks marked on the words "the old man," "in the park," "was feeding," and "the pigeons." The peaks indicate stress and timing in the sentence.]
• Martin 1970—Naive parsing
• Transition error probabilities
• Parsing and pausing
Finite-State Cascade

- Regular expressions defining chunks and clauses
  
  **Clause:** \( S \rightarrow \text{Comp? \,(PP|Adv)* \,NP} \,(PP|Adv)^+ \,\text{VP} \,\text{NP? \,(PP|Adv)^*} \)
  
  **Chunk:** \( \text{NP} \rightarrow \text{Det? \,Adj}^+ \,\text{N} \)
  
  \( \text{VP} \rightarrow \text{Mod? \,Aux}^+ \,\text{V} \)

- Finite-state machines
  
  **Clause:**
  
  \[
  \begin{array}{c}
  \text{S} \rightarrow \text{Comp, PP, Adv} \\
  S_0 \rightarrow \text{NP} \\
  S_1 \rightarrow \text{NP, PP, Adv} \\
  S_2 \rightarrow \text{NP} \\
  S_3 \rightarrow \text{PP, Adv} \\
  S_4 \rightarrow \text{PP, Adv}
  \end{array}
  \]

  **Chunk:**
  
  \[
  \begin{array}{c}
  \text{NP} \rightarrow \text{D, A} \\
  \text{C}_0 \rightarrow \text{N} \\
  \text{C}_1 \rightarrow \text{V} \\
  \text{C}_2 \rightarrow \text{Aux}
  \end{array}
  \]
• Longest match within level

• Punt if no match

• No center-embedding

• Special level for nested RC’s, participles
Longest Match

- Chunks:

- Clauses:
Similar to Late Closure

• Chunks: the emergency crews hate most is domestic violence

• Clause: because John always jogs a mile seems short to him

• Compounds: a hundred pound bags
Effectiveness

- Corpus data
- Hand tagged
- Disambiguate via longest match at each level, punt if no match
- Effectiveness

<table>
<thead>
<tr>
<th></th>
<th>recall</th>
<th>precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>chunks:</td>
<td>96%</td>
<td>91%</td>
</tr>
<tr>
<td>NP’s:</td>
<td>92%</td>
<td>87%</td>
</tr>
</tbody>
</table>
Nested Relative Clauses

- Nested relative clause

  S
  
  NP  RC  VP  NP
  
  NP  Comp  NP  VP  VP  NP
  the man  that  you  saw  was  a crook

- Bever 1970

  S
  
  NP  PrtC
  
  NP  VP  PP  VP
  the boat  floated  down the river  sank
Attachment

- Lexical head-head associations

- Attachment preferences:
  - Attach as argument (lexical selection)
  - Attach to (anticipated)
  - Attach to nearest site

- Argument:

  I thought about his interest in the Volvo

- Head:

  I sang to the cat in the kitchen
Anticipated Head

- Subject:

- Head-final (German):
• Level-by-level is completely implausible

\textit{I think Bill said Mary left}

• One pass is no problem
- Slow-down in reading and decision times when it is realized that order is OVS

- Slow-down is relative to:
  - nom V acc
  - nom V ambig
  - ambig V acc
  - ambig V ambig
Problem for Chunking

• OK:

\[
\text{NP acc} \quad \text{D A N V D A N}
\]

• But: if case anomaly not recognized till NP is constructed:

\[
\text{NP nom?} \quad \text{NP nom} \quad \text{D A N V D A N}
\]
Finer-Grained Alternative

• Case features are unified when arc is transversed: 

\[
\begin{array}{c}
\text{Case X} \\
\text{Case X} \\
\text{Case X}
\end{array}
\]

\[
\begin{array}{c}
A \\
D \\
Mod, Aux
\end{array}
\]

NP
VP

\[
\begin{array}{c}
C_0 \\
C_1 \\
C_2 \\
C_3 \\
C_4
\end{array}
\]

\[
\begin{array}{c}
N \\
V
\end{array}
\]

\[
\begin{array}{c}
\text{Mod, Aux}
\end{array}
\]

• Assumption: expectation [Case $\neg$nom] from $S_3$ is available to chunk state 

\[
\begin{array}{c}
S_0 \\
S_2 \\
S_3
\end{array}
\]

\[
\begin{array}{c}
\text{NP} \\
\text{VP}
\end{array}
\]

\[
\begin{array}{c}
C_0 \\
C_1 \\
C_1 \\
C_2 \\
C_3 \\
C_4 \\
C_0 \\
C_1
\end{array}
\]

\[
\begin{array}{c}
D \\
A \\
N \\
V \\
D \\
A \\
N
\end{array}
\]

• Phrase construction = commitment. GP if wrong.
- Why delayed slowdown for first NP?

- Hemforth et al.: Left-corner parsing

- Chunk analogue: levels started on demand

\[ \begin{align*}
S_0 \\
C_0 & D & A & C_0 \\
\cdots & S_2 & S_3 \\
C_2 & C_0 & C_4 & C_0 & C_1 \\
\cdots & V & D & A & \end{align*} \]
Problem for Left Corner

- Left-corner, arc-standard: create parent and predict sisters, but don’t attach.
- Left-corner, arc-eager: attach parent to grandparent as soon as parent is created.
- Eager is too eager:

```
  S'   
    /   
   S   V
  /    /
D NP D NP
   /    /
D   N'  D   N'
```
- Makes right predictions for first NP
• Makes wrong predictions for second NP
Not Just a Technicality
Discriminating Hypotheses

- Mixed LC: arc-standard if stack is empty, else arc-eager
- ‘Interactive’ chunk model and Mixed LC both get Hemforth et al. data right
- Different predictions on multi-chunk first NP’s
Multi-Chunk NP

- Mixed LC:

```
NP  NP  NP  NP
  den  den  N'  N'
 Vortrag  Vortrag  PP  PP
    PP
  ueber  ueber  NP
       NP
     Kausalitaet
```
• Chunks & dependencies: finite-state cascade, attachment within levels
• Longest match disambiguates chunks
• Corpus evaluation
• Addressing incrementality:
  – Levels on demand
  – Feature expectations from context, unified during transitions
• Only mixed LC consistent with Hemforth et al. data
• Differing predictions on multi-chunk NP’s