Frequency Effects in the Processing of Unbounded Dependencies

Marten van Schijndel, William Schuler, and Peter Culicover
Department of Linguistics
The Ohio State University

July 24, 2014
Background

Unbounded Dependency

A non-local dependency that potentially spans an unbounded number of lexemes.

e.g. That’s {the ball} John kicked ___.
e.g. That’s {the ball} Mary said John kicked ___.

This is hard because:

- Filler must be remembered
- Where is the gap?
  - Maybe people use subcategorization bias?
The preference for a lexical item to take a particular type of argument

The girl realized \( \left\{ \text{the house was on fire} \right\} \) \( \left\{ \text{her potential} \right\} \).

*Realized* prefers a sentential complement over a noun phrase.
Several studies have investigated subcat usage...

- Mitchell (1987)
- Pickering et al. (2000)
- van Gompel & Pickering (2001)
- Pickering & Traxler (2003)
Background

Pickering & Traxler (2003)

(1) That’s the plane that the pilot landed carefully behind in the fog at the airport.
(2) That’s the truck that the pilot landed carefully behind in the fog at the airport.
Background

Pickering & Traxler (2003)

(1) That’s the plane that the pilot landed carefully behind in the fog at the airport.
(2) That’s the truck that the pilot landed carefully behind in the fog at the airport.

Readers slow down at *landed* in (2)

Suggests they try to link *truck* as the object of *landed* despite:

- *landed* biased for PP complement (e.g. “landed on the ground”)
  - 40% PP complement
  - 25% NP complement

∴ Readers initially adopt a transitive interpretation despite subcat bias
Several studies have investigated subcat usage...

- Mitchell (1987)
- Pickering et al. (2000)
- van Gompel & Pickering (2001)
- Pickering & Traxler (2003)

Suggest subcat information actually isn’t used immediately for unbounded dependency processing.

Finding supports multistage models of sentence processing

- Garden Path Model (Frazier, 1987)
- Construal (Frazier & Clifton, 1996)
More recent studies have revisited this claim:

- Staub et al. (2006)
- Staub (2007)
- Arai & Keller (2013)

Provide indirect evidence that the previous set of results may be driven by frequency effects of larger syntactic structures.

Present Contribution:
The current work provides direct evidence in support of this hypothesis.
Alternative Explanation

Unbounded dependencies more often go to arguments than modifiers.
Model

Probabilistic context-free grammar (PCFG) for unbounded dependencies:

$\begin{align*}
P(NP \rightarrow D \ N) &= P(D \ N \mid NP) = 0.66 \\
P(NP \rightarrow NP \ RC-gNP) &= P(NP \ RC-gNP \mid NP) = 0.33
\end{align*}$
(1) That’s the plane that the pilot landed carefully behind in the fog at the airport.
(2) That’s the truck that the pilot landed carefully behind in the fog at the airport.

(a) Transitive

(b) Intransitive
Pickering & Traxler (2003)

(1) That’s the plane that the pilot landed carefully behind in the fog at the airport.
(2) That’s the truck that the pilot landed carefully behind in the fog at the airport.

(a) \[
\begin{array}{c}
\text{VP-gNP} \\
\text{VP-gNP} \\
\text{VP-gNP} \\
\text{TV} \quad t_i \\
\text{landed} \\
\end{array}
\quad \begin{array}{c}
\text{Adv} \\
P \\
\text{behind} \\
\text{carefully} \\
\text{Transitive} \\
\end{array}
\quad \begin{array}{c}
P \\
\text{NP} \\
\end{array}
\quad \begin{array}{c}
\text{PP} \\
\end{array}
\]

(b) \[
\begin{array}{c}
\text{VP-gNP} \\
\text{VP-gNP} \\
\text{VP} \\
\text{IV} \\
\text{landed} \\
\end{array}
\quad \begin{array}{c}
\text{Adv} \\
P \\
\text{behind} \\
\text{carefully} \\
\text{Intransitive} \\
\end{array}
\quad \begin{array}{c}
P \\
\text{t_i} \\
\end{array}
\]
(1) That’s the plane that the pilot landed carefully behind in the fog at the airport.
(2) That’s the truck that the pilot landed carefully behind in the fog at the airport.
Model

How probable is each subtree?
Wall Street Journal (WSJ) section of the Penn Treebank:

(a) VP-gNP
   VP-gNP
   VP-gNP  Adv  P  NP
   TV  ti  carefully  behind
   landed

(b) VP-gNP
   VP-gNP
   VP  Adv  P  t_i
   IV  carefully  behind
   landed

Transitive 0.17
Intransitive 0.01
Model

What is the probability of each interpretation?

\[ P(\text{syntactic configuration}) \cdot P(\text{generating the verb from that tree}) \]

\[ P(\text{Transitive}) = P(\text{VP-gNP} \rightarrow \text{VP-gNP PP}) \cdot P(\text{verb} | \text{TV}) \quad (1) \]

\[ P(\text{Intransitive}) = P(\text{VP-gNP} \rightarrow \text{VP PP-gNP}) \cdot P(\text{verb} | \text{IV}) \quad (2) \]
Model

What is the probability of each interpretation?
\[
P(\text{syntactic configuration}) \cdot P(\text{generating the verb from that tree}) \cdot \frac{P(\text{subcat bias})}{P(\text{preterminal prior})}
\]

\[
P(\text{Transitive}) = P(\text{VP-gNP} \rightarrow \text{VP-gNP PP}) \cdot P(\text{verb} \mid \text{TV}) \tag{1}
\]

\[
P(\text{Intransitive}) = P(\text{VP-gNP} \rightarrow \text{VP PP-gNP}) \cdot P(\text{verb} \mid \text{IV}) \tag{2}
\]
Model

What is the probability of each interpretation?

\[ P(\text{syntactic configuration}) \cdot P(\text{subcat bias}) / P(\text{preterminal prior}) \]

\[ P(\text{Transitive}) = P(\text{VP-gNP} \rightarrow \text{VP-gNP PP}) \cdot P(\text{verb} \mid \text{TV}) \quad (1) \]

\[ \propto P(\text{VP-gNP} \rightarrow \text{VP-gNP PP}) \cdot \frac{P(\text{TV} \mid \text{verb})}{P(\text{TV})} \]

\[ P(\text{Intransitive}) = P(\text{VP-gNP} \rightarrow \text{VP PP-gNP}) \cdot P(\text{verb} \mid \text{IV}) \quad (2) \]

\[ \propto P(\text{VP-gNP} \rightarrow \text{VP PP-gNP}) \cdot \frac{P(\text{IV} \mid \text{verb})}{P(\text{IV})} \]
Model

What are the preterminal priors?
Relative prior probability from the WSJ:

\[ P(TV) : P(IV) = 2.6 : 1 \]
Model

What is the probability of each interpretation?

\[ P(\text{syntactic configuration}) \cdot P(\text{subcat bias}) / P(\text{preterminal prior}) \]

\[ P(\text{Transitive}) \propto P(\text{VP-gNP} \rightarrow \text{VP-gNP PP}) \cdot \frac{P(\text{TV} \mid \text{verb})}{P(\text{TV})} \]

\[ = 0.17 \cdot \frac{P(\text{TV} \mid \text{verb})}{2.6} = 0.065 \cdot P(\text{TV} \mid \text{verb}) \tag{1} \]

\[ P(\text{Intransitive}) \propto P(\text{VP-gNP} \rightarrow \text{VP PP-gNP}) \cdot \frac{P(\text{IV} \mid \text{verb})}{P(\text{IV})} \]

\[ = 0.01 \cdot \frac{P(\text{IV} \mid \text{verb})}{1.0} = 0.01 \cdot P(\text{IV} \mid \text{verb}) \tag{2} \]

Pickering & Traxler (2003) experimentally determined subcat biases for a wide variety of verbs.
(1) That’s the plane that the pilot landed carefully behind in the fog at the airport.
(2) That’s the truck that the pilot landed carefully behind in the fog at the airport.

Using Pickering & Traxler’s (2003) subcat bias data:

\[
\begin{align*}
P(\text{Transitive} \mid \text{landed}) &\propto 0.17 \cdot \frac{0.25}{2.6} = 0.016 \\
P(\text{Intransitive} \mid \text{landed}) &\propto 0.01 \cdot \frac{0.40}{1.0} = 0.004
\end{align*}
\]

Transitive interpretation is 300% more likely!
## Evaluation

<table>
<thead>
<tr>
<th>Verb</th>
<th>Trans</th>
<th>Intrans</th>
<th>Str</th>
<th>Trans</th>
<th>Intrans</th>
<th>Str</th>
</tr>
</thead>
<tbody>
<tr>
<td>spoke</td>
<td>0</td>
<td>55</td>
<td>100</td>
<td>0</td>
<td>0.55</td>
<td>100</td>
</tr>
<tr>
<td>worried</td>
<td>0</td>
<td>50</td>
<td>100</td>
<td>0</td>
<td>0.50</td>
<td>100</td>
</tr>
<tr>
<td>pointed</td>
<td>10</td>
<td>90</td>
<td>90</td>
<td>0.65</td>
<td>0.90</td>
<td>58</td>
</tr>
<tr>
<td>fished</td>
<td>5</td>
<td>45</td>
<td>90</td>
<td>0.33</td>
<td>0.45</td>
<td>58</td>
</tr>
<tr>
<td>argued</td>
<td>11</td>
<td>64</td>
<td>85</td>
<td>0.72</td>
<td>0.64</td>
<td>53</td>
</tr>
<tr>
<td>searched</td>
<td>15</td>
<td>75</td>
<td>83</td>
<td>0.98</td>
<td>0.75</td>
<td>57</td>
</tr>
<tr>
<td>communicated</td>
<td>10</td>
<td>50</td>
<td>83</td>
<td>0.65</td>
<td>0.50</td>
<td>57</td>
</tr>
<tr>
<td>shouted</td>
<td>10</td>
<td>50</td>
<td>83</td>
<td>0.65</td>
<td>0.50</td>
<td>57</td>
</tr>
<tr>
<td>swore</td>
<td>6</td>
<td>17</td>
<td>74</td>
<td>0.39</td>
<td>0.17</td>
<td>70</td>
</tr>
<tr>
<td>travelled</td>
<td>20</td>
<td>40</td>
<td>67</td>
<td>1.31</td>
<td>0.40</td>
<td>77</td>
</tr>
<tr>
<td>landed</td>
<td>25</td>
<td>40</td>
<td>62</td>
<td>1.63</td>
<td>0.40</td>
<td>80</td>
</tr>
<tr>
<td>raced</td>
<td>35</td>
<td>55</td>
<td>61</td>
<td>2.29</td>
<td>0.55</td>
<td>81</td>
</tr>
<tr>
<td>blabbed</td>
<td>30</td>
<td>45</td>
<td>60</td>
<td>1.96</td>
<td>0.45</td>
<td>81</td>
</tr>
<tr>
<td>preached</td>
<td>30</td>
<td>45</td>
<td>60</td>
<td>1.96</td>
<td>0.45</td>
<td>81</td>
</tr>
</tbody>
</table>

Verbal bias percentages (Pickering & Traxler 2003)
Evaluation: Unaccusatives

Obligatorily intransitive verbs (e.g. erupt) do not cause such a slow down (Staub 2007)

Current model explains this via 0 transitive bias:

\[ P(\text{Transitive}) \propto 0.17 \cdot \frac{0.0}{2.6} = 0.0 \]
**Criticism**

**WSJ won’t generalize**

- Subcat biases determined experimentally by P&T (2003)
- Nguyen et al.’s (2012) results suggest WSJ unbounded dependency distribution may generalize
- The current work accounts for a variety of findings...
  - Pickering & Traxler (2003), etc. “lack” of subcat usage
  - Staub et al. (2006) heavy-NP shift processing heuristics
  - Staub (2007) unaccusative subcat usage
Conclusions

- Unbounded dependencies more often go to arguments than modifiers
- Previous studies of subcat bias confounded by syntactic frequency
  - Replication possible when subcat biases taken into immediate account
  - Weaken support for multistage models of sentence processing
- Shows the need to account for frequency at multiple levels of processing; not simply in terms of lexical bias
Questions?

Thanks to:

- Matthew Traxler
- Shari Speer
- OSU Syntactic Processing Seminar
- OSU Linguistics Targeted Investment for Excellence (2012-2013)


