Structural Expectations in Chinese Relative Clause Comprehension

Zhong Chen, Kyle Grove, John Hale
Department of Linguistics
Cornell University
Outline

Introduction
Chinese RC
Modeling
Conclusion
Relative clause comprehension

Subject relative clause (SR)

The senator \(i\) [who \(e_i\) attacked the reporter] admitted the error.

Object relative clause (OR)

The senator \(i\) [who the reporter attacked \(e_i\)] admitted the error.

Comprehension difficulty: SR << OR
Two major accounts

Memory-based
(Gibson 1998, 2000; Grodner & Gibson 2005)

Experience-based
Memory-based approach

Subject Relative clause (SR)

The senator\textsubscript{i} [who e\textsubscript{i} attacked the reporter] admitted the error.

Object Relative clause (OR)

The senator\textsubscript{i} [who the reporter attacked e\textsubscript{i}] admitted the error.

Memory-based: SR $<<$ OR
Memory-based approach

Subject Relative clause (SR)

The senator\textsubscript{i} [who \textit{e}\textsubscript{i} attacked the reporter] admitted the error.

Object Relative clause (OR)

The senator\textsubscript{i} [who the reporter attacked \textit{e}\textsubscript{i}] admitted the error.

Memory-based: SR $\ll$ OR
Memory-based approach

Subject Relative clause (SR)

The senator \( i \) [who \( e_i \) attacked the reporter] admitted the error.

Object Relative clause (OR)

The senator \( i \) [who the reporter attacked \( e_i \)] admitted the error.

Memory-based: SR << OR
Memory-based approach

Subject Relative clause (SR)

The \textbf{senator}_i [who \textbf{e}_i attacked the reporter] admitted the error.

Object Relative clause (OR)

The \textbf{senator}_i [who \textbf{the reporter} attacked \textbf{e}_i] admitted the error.

Memory-based: SR << OR
Experience-based approach

English Penn Treebank: 86% SR vs 13% OR (Hale 2001)

German NEGRA: 74% SR vs 26% OR (Skut et al. 1997)

Experience-based: SR << OR
Outline

• Introduction
• Chinese RC
• Modeling
• Conclusion
Hsiao & Gibson (2003)

SR: [ei yaoqing fuhao de] guanyuan_i da-le jizhe  
    invite     tycoon DE official             hit     reporter  
    ‘The official who invited the tycoon hit the reporter.’

OR: [fuhao yaoqing ei de] guanyuan_i da-le jizhe 
    tycoon invite     DE official             hit     reporter  
    ‘The official who the tycoon invited hit the reporter.’

Memory-based: SR >> OR

Experience-based: SR << OR
Hsiao & Gibson (2003)

SR: [e_i yaoqing fuhao de] guanyuan_i da-le jizhe
   invite    tycoon DE official    hit    reporter
   ‘The official who invited the tycoon hit the reporter.’

OR: [fuhao yaoqing e_i de] guanyuan_i da-le jizhe
   tycoon invite    DE official    hit    reporter
   ‘The official who the tycoon invited hit the reporter.’

Memory-based: SR >> OR

Experience-based: SR << OR
Hsiao & Gibson (2003)

SR: [e_i yaoqing fuhao de] guanyuan_i da-le jizhe

invite tycoon DE official hit reporter

‘The official who invited the tycoon hit the reporter.’

OR: [fuhao yaoqing e_i de] guanyuan_i da-le jizhe

tycoon invite DE official hit reporter

‘The official who the tycoon invited hit the reporter.’

Memory-based: SR >> OR

Experience-based: SR << OR
Hsiao & Gibson (2003)

SR: \[\text{invite tycoon DE official} \quad \text{hit reporter}\]

‘The official who invited the tycoon hit the reporter.’

OR: \[\text{tycoon invite DE official} \quad \text{hit reporter}\]

‘The official who the tycoon invited hit the reporter.’

Memory-based: SR >> OR

Experience-based: SR << OR
Lin & Bever (2006)

• Subject-modifying SR (SR-S)
   • [e_i yaoqing fuhao de] guanyuan_i da-le jizhe

• Subject-modifying OR (OR-S)
   • [fuhao yaoqing e_i de] guanyuan_i da-le jizhe
Lin & Bever (2006)

- Subject-modifying SR (SR-S)
  - [e_i yaoqing fuhao de] guanyuan_i da-le jizhe
- Subject-modifying OR (OR-S)
  - [fuhao yaoqing e_i de] guanyuan_i da-le jizhe
- Object-modifying SR (SR-O)
  - jizhe da-le [e_i yaoqing fuhao de] guanyuan_i
Lin & Bever (2006)

- Subject-modifying SR (SR-S)
  - \([e_i \text{ yaoqing fuhao de}] \text{ guanyuan}_i \text{ da-le jizhe}\)
- Subject-modifying OR (OR-S)
  - \([\text{fuhao yaoqing } e_i \text{ de}] \text{ guanyuan}_i \text{ da-le jizhe}\)
- Object-modifying SR (SR-O)
  - \(\text{jizhe da-le } [e_i \text{ yaoqing fuhao de}] \text{ guanyuan}_i\)
- Object-modifying OR (OR-O)
  - \(\text{jizhe da-le } [\text{fuhao yaoqing } e_i \text{ de}] \text{ guanyuan}_i\)

Figure 3. Mean reading times in milliseconds for each condition in Experiment 2a.

Figure 4. Mean reading times in milliseconds for each condition in Experiment 2b.
Lin and Bever (2006)

- Subject-modifying SR (SR-S)
  - \([e_i \text{ yaoqing fuhao de}] \text{guanyuan}_{i} \text{ da-le jizhe}\)
- Subject-modifying OR (OR-S)
  - \([\text{fuhao yaoqing } e_{i} \text{ de}] \text{guanyuan}_{i} \text{ da-le jizhe}\)
- Object-modifying SR (SR-O)
  - jizhe da-le \([e_{i} \text{ yaoqing fuhao de}] \text{guanyuan}_{i}\)
- Object-modifying OR (OR-O)
  - jizhe da-le \([\text{fuhao yaoqing } e_{i} \text{ de}] \text{guanyuan}_{i}\)

reporter hit tycoon invite
Lin and Bever (2006)

- Subject-modifying SR (SR-S)
  - [e_i yaoqing fuhao de] guanyuan_i da-le jizhe

- Subject-modifying OR (OR-S)
  - [fuhao yaoqing e_i de] guanyuan_i da-le jizhe

- Object-modifying SR (SR-O)
  - jizhe da-le [e_i yaoqing fuhao de] guanyuan_i

- Object-modifying OR (OR-O)
  - jizhe da-le [fuhao yaoqing e_i de] guanyuan_i

**reporter hit tycoon**
Analyses of reading times Table 1 of the Appendix lists the word-by-word mean RTs of the region that includes two words before and after “DE”. Figure 2 plots RTs across four RT types that refer to conditions in (10). A direct reading time comparison at the head noun is presented in Figure 4. Table 4 reports the results of statistical analyses.

Over the five words taken as a whole, subject extractions (a and c) were read 88.0 ms faster than object extractions (b and d) (p < 0.05). At the head noun, we see a marginal SR advantage in subject-modifying conditions (a vs b) (t = 0.79, n.s.). If RCs modified the matrix object, this SR preference was intensified (c vs d) (t = 2.32, p < 0.05).

Adding a spillover predictor in the model will not change the SR advantage in either contrast. At the next word of the head noun, the SR preference persisted.

Table 2: Mean reading times in milliseconds for each condition in Experiment 2a.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Region</th>
<th>Contrast/Predictor</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a head noun</td>
<td>av sb</td>
<td>0.012</td>
<td>0.0153</td>
<td>0.79</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>2a head noun</td>
<td>cv sd</td>
<td>0.037</td>
<td>0.0159</td>
<td>2.35</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>2a head noun</td>
<td>av sb</td>
<td>0.006</td>
<td>0.0148</td>
<td>0.41</td>
<td>&lt;0.05</td>
<td></td>
</tr>
<tr>
<td>2a head noun</td>
<td>cv sd</td>
<td>0.033</td>
<td>0.0153</td>
<td>2.13</td>
<td>&lt;0.05</td>
<td></td>
</tr>
</tbody>
</table>

Experiment 2b: Method Experiment 2b was also conducted in Dalian, China. However, the participants were 60 college students who did not take Experiment 2a. Each participant...
Outline

- Introduction
- Chinese RC
- Modeling
- Conclusion
Surprisal (Hale, 2001)
Surprisal (Hale, 2001)

- is a model of sentence processing difficulty
Surprisal (Hale, 2001)

• is a model of sentence processing difficulty

• is built on a language model, such as a PCFG
Surprisal (Hale, 2001)

- is a model of sentence processing difficulty
- is built on a language model, such as a PCFG
- quantifies the “unlikelihood” (surprise) of integrating a word in the sentence
Surprisal (Hale, 2001)

• is a model of sentence processing difficulty
• is built on a language model, such as a PCFG
• quantifies the “unlikelihood” (surprise) of integrating a word in the sentence
• is used to predict word-by-word reading times
Surprisal: an example

A six-word string:

0 The 1 horse 2 raced 3 past 4 the 5 barn 6
Surprisal: an example

A six-word string:

0 The 1 horse 2 raced 3 past 4 the 5 barn 6

Saturday, April 23, 2011
Surprisal: an example

A six-word string:

0 The 1 horse 2 raced 3 past 4 the 5 barn 6

main-clause reading >> reduced RC reading
Surprisal: an example

The next word leads to a structural reanalysis:

0 The 1 horse 2 raced 3 past 4 the 5 barn 6 fell 7
Surprisal: an example

The next word leads to a structural reanalysis:

0 The 1 horse 2 raced 3 past 4 the 5 barn 6 fell 7

main-clause-reading
Surprisal: an example

The next word leads to a structural reanalysis:

0 The 1 horse 2 raced 3 past 4 the 5 barn 6 fell

main-clause reading

reduced-relative reading
Surprisal: the calculation

\[
0 \text{ The 1 horse 2 raced 3 past 4 the 5 barn 6 fell 7 }
\]

\[
surprisal = \log_2 \left( \frac{\alpha(0,6)}{\alpha(0,7)} \right)
\]
Surprisal: the calculation

\[ \text{surprisal} = \log_2 \left( \frac{\alpha(0,6)}{\alpha(0,7)} \right) \]
Surprisal: the calculation

\[ \text{surprisal} = \log_2 \left( \frac{\alpha(0,6)}{\alpha(0,7)} \right) \]
Surprisal: the calculation

\[ \text{surprisal} = \log_2 \left( \frac{\alpha(0,6)}{\alpha(0,7)} \right) \]
Surprisal: the calculation

\[ \text{surprisal} = \log_2 \left( \frac{\alpha(0,7)}{\alpha(0,6)} \right) \]
Table r: A richer PCFG for Chinese Relative Clauses with pro drop rules

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
<th>Rule</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun subject</td>
<td>10870</td>
<td>$S \rightarrow \text{NPSBJ VP}$</td>
<td>0.9261</td>
</tr>
<tr>
<td>pro subject</td>
<td>14921</td>
<td>$S \rightarrow \text{NPRC VP}$</td>
<td>0.0739</td>
</tr>
<tr>
<td>RC subject</td>
<td>2057</td>
<td>$\text{NPRC} \rightarrow \text{CPSR NP}$</td>
<td>0.5935</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{NPRC} \rightarrow \text{CPOR NP}$</td>
<td>0.4065</td>
</tr>
<tr>
<td>noun object</td>
<td>14041</td>
<td>$\text{VP} \rightarrow \text{V NPOBJ}$</td>
<td>0.8615</td>
</tr>
<tr>
<td>pro object</td>
<td>109</td>
<td>$\text{VP} \rightarrow \text{V NPRC}$</td>
<td>0.1385</td>
</tr>
<tr>
<td>RC object</td>
<td>2273</td>
<td>$\text{CPSR} \rightarrow \text{VP DEC}$</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{CPOR} \rightarrow \text{S/NP DEC}$</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{S/NP} \rightarrow \text{NPSBJ VP/NP}$</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{VP/NP} \rightarrow \text{V NP/NP}$</td>
<td>1.0</td>
</tr>
<tr>
<td>SRC</td>
<td>895</td>
<td>$\text{NSBJ} \rightarrow \text{NP}$</td>
<td>0.4215</td>
</tr>
<tr>
<td>ORC</td>
<td>613</td>
<td>$\text{NP} \rightarrow \text{N}$</td>
<td>0.5785</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{NPSBJ} \rightarrow \text{pro}$</td>
<td>0.9928</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{NPOBJ} \rightarrow \text{pro}$</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{NP} \rightarrow \text{N}$</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table s: Attestation count from Chinese Treebank

4.2 Parser
An statistical prefix parsing system was used to obtain surprisals for the prefix in each of the four examples. It employs a bottom-up chart parsing strategy in the style of Shieber et al. The parser constructs a probabilistic model at each prefix calculating the inside probability of each non-terminal. For example, if more than one parse is available, the total probability should be the summation of the probability of all parses.

5 Results
Starting from the simpler grammar Subject-modifying Relative Clauses as in Figure...
## Probabilistic Grammar

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun subject</td>
<td>10870</td>
</tr>
<tr>
<td>pro subject</td>
<td>14921</td>
</tr>
<tr>
<td>RC subject</td>
<td>2057</td>
</tr>
<tr>
<td>noun object</td>
<td>14041</td>
</tr>
<tr>
<td>pro object</td>
<td>109</td>
</tr>
<tr>
<td>RC object</td>
<td>2273</td>
</tr>
<tr>
<td>SRC</td>
<td>895</td>
</tr>
<tr>
<td>ORC</td>
<td>613</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.9261</td>
</tr>
<tr>
<td>NPSBJ VP</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.0739</td>
</tr>
<tr>
<td>NPRC VP</td>
<td></td>
</tr>
<tr>
<td>NPSBJ VP</td>
<td>0.5935</td>
</tr>
<tr>
<td>NPRC CPOR NP</td>
<td></td>
</tr>
<tr>
<td>NPRC CPOR NP</td>
<td>0.4065</td>
</tr>
<tr>
<td>VP</td>
<td>0.8615</td>
</tr>
<tr>
<td>V NPOBJ</td>
<td></td>
</tr>
<tr>
<td>VP</td>
<td>0.1385</td>
</tr>
<tr>
<td>V NPRC</td>
<td></td>
</tr>
<tr>
<td>CPSR VP DEC</td>
<td>1.0</td>
</tr>
<tr>
<td>CPOR S/NP DEC</td>
<td>1.0</td>
</tr>
<tr>
<td>S/NP VP/NP NP</td>
<td>1.0</td>
</tr>
<tr>
<td>NPSBJ VP</td>
<td>0.4215</td>
</tr>
<tr>
<td>NP</td>
<td></td>
</tr>
<tr>
<td>NPSBJ pro</td>
<td>0.5785</td>
</tr>
<tr>
<td>NPSBJ pro</td>
<td>0.9928</td>
</tr>
<tr>
<td>NPOBJ pro</td>
<td>0.0072</td>
</tr>
<tr>
<td>NPOBJ pro</td>
<td>1.0</td>
</tr>
<tr>
<td>NPOBJ pro</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>0.5</td>
</tr>
<tr>
<td>V</td>
<td></td>
</tr>
<tr>
<td>→</td>
<td>yaoqing</td>
</tr>
<tr>
<td>N</td>
<td>0.5</td>
</tr>
<tr>
<td>V</td>
<td></td>
</tr>
<tr>
<td>→</td>
<td>dale</td>
</tr>
<tr>
<td>N</td>
<td>0.3333</td>
</tr>
<tr>
<td>→</td>
<td>fuhao</td>
</tr>
<tr>
<td>N</td>
<td>0.3333</td>
</tr>
<tr>
<td>→</td>
<td>guanyuan</td>
</tr>
<tr>
<td>N</td>
<td>0.3334</td>
</tr>
<tr>
<td>→</td>
<td>jizhe</td>
</tr>
<tr>
<td>NP</td>
<td>1.0</td>
</tr>
<tr>
<td>→</td>
<td>de</td>
</tr>
<tr>
<td>NP</td>
<td>1.0</td>
</tr>
<tr>
<td>→</td>
<td>ε</td>
</tr>
<tr>
<td>NP</td>
<td>1.0</td>
</tr>
<tr>
<td>→</td>
<td>ε</td>
</tr>
</tbody>
</table>

---

Table R: A richer PCFG for Chinese Relative Clauses with pro drop rules

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun subject</td>
<td>10870</td>
</tr>
<tr>
<td>pro subject</td>
<td>14921</td>
</tr>
<tr>
<td>RC subject</td>
<td>2057</td>
</tr>
<tr>
<td>noun object</td>
<td>14041</td>
</tr>
<tr>
<td>pro object</td>
<td>109</td>
</tr>
<tr>
<td>RC object</td>
<td>2273</td>
</tr>
<tr>
<td>SRC</td>
<td>895</td>
</tr>
<tr>
<td>ORC</td>
<td>613</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.9261</td>
</tr>
<tr>
<td>NPSBJ VP</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.0739</td>
</tr>
<tr>
<td>NPRC VP</td>
<td></td>
</tr>
<tr>
<td>NPSBJ VP</td>
<td>0.5935</td>
</tr>
<tr>
<td>NPRC CPOR NP</td>
<td></td>
</tr>
<tr>
<td>NPRC CPOR NP</td>
<td>0.4065</td>
</tr>
<tr>
<td>VP</td>
<td>0.8615</td>
</tr>
<tr>
<td>V NPOBJ</td>
<td></td>
</tr>
<tr>
<td>VP</td>
<td>0.1385</td>
</tr>
<tr>
<td>V NPRC</td>
<td></td>
</tr>
<tr>
<td>CPSR VP DEC</td>
<td>1.0</td>
</tr>
<tr>
<td>CPOR S/NP DEC</td>
<td>1.0</td>
</tr>
<tr>
<td>S/NP VP/NP NP</td>
<td>1.0</td>
</tr>
<tr>
<td>NPSBJ VP</td>
<td>0.4215</td>
</tr>
<tr>
<td>NP</td>
<td></td>
</tr>
<tr>
<td>NPSBJ pro</td>
<td>0.5785</td>
</tr>
<tr>
<td>NPSBJ pro</td>
<td>0.9928</td>
</tr>
<tr>
<td>NPOBJ pro</td>
<td>0.0072</td>
</tr>
<tr>
<td>NPOBJ pro</td>
<td>1.0</td>
</tr>
<tr>
<td>NPOBJ pro</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td></td>
</tr>
</tbody>
</table>

Table S: Attestation count from Chinese Treebank

4.2 Parser
An statistical prefix parsing system Grovek was used to obtain surprisals for the prefix in each of the four examples. It employs a bottom-up chart parsing strategy in the style of Shieber et al. The parser constructs a probabilistic model at each prefix calculating the inside probability of each non-terminal. For example, if more than one parse is available, the total probability should be the summation of the probability of all parses.
Probabilistic Grammar

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun subject</td>
<td>10870</td>
</tr>
<tr>
<td>pro subject</td>
<td>14921</td>
</tr>
<tr>
<td>RC subject</td>
<td>2057</td>
</tr>
<tr>
<td>noun object</td>
<td>14041</td>
</tr>
<tr>
<td>pro object</td>
<td>109</td>
</tr>
<tr>
<td>RC object</td>
<td>2273</td>
</tr>
<tr>
<td>SRC</td>
<td>895</td>
</tr>
<tr>
<td>ORC</td>
<td>613</td>
</tr>
</tbody>
</table>

Table r: A richer PCFG for Chinese Relative Clauses with pro drop rules

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun subject</td>
<td>10870</td>
</tr>
<tr>
<td>pro subject</td>
<td>14921</td>
</tr>
<tr>
<td>RC subject</td>
<td>2057</td>
</tr>
<tr>
<td>noun object</td>
<td>14041</td>
</tr>
<tr>
<td>pro object</td>
<td>109</td>
</tr>
<tr>
<td>RC object</td>
<td>2273</td>
</tr>
<tr>
<td>SRC</td>
<td>895</td>
</tr>
<tr>
<td>ORC</td>
<td>613</td>
</tr>
</tbody>
</table>

### Table r: A richer PCFG for Chinese Relative Clauses with pro drop rules

<table>
<thead>
<tr>
<th>Type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun subject</td>
<td>10870</td>
</tr>
<tr>
<td>pro subject</td>
<td>14921</td>
</tr>
<tr>
<td>RC subject</td>
<td>2057</td>
</tr>
<tr>
<td>noun object</td>
<td>14041</td>
</tr>
<tr>
<td>pro object</td>
<td>109</td>
</tr>
<tr>
<td>RC object</td>
<td>2273</td>
</tr>
<tr>
<td>SRC</td>
<td>895</td>
</tr>
<tr>
<td>ORC</td>
<td>613</td>
</tr>
</tbody>
</table>

Table s: Attestation count from Chinese Treebank

5 Results

Starting from the simpler grammar

Subject-modifying Relative Clauses as in Figure...
Parser

- bottom-up chart-parsing algorithm (Goodman, 1999)

- the incremental parser considers multiple parses at each position.
Results: Subj-modifying RC

Total:  
SR  8.89  
OR 10.47
Results: Subj-modifying RC

Total:
SR 8.89
OR 10.47
Results: Subj-modifying RC

Total:
SR  8.89
OR 10.47
Calculate surprisals

- In SR-S, before “de”: V N “invite tycoon …”

(1) pro-drop main clause

(2) SR
Calculate surprisals

- In **SR-S**, after “de”: V N de “invite tycoon de ...”

![Diagram showing sentence structures and surprisal values](image)

(1) **pro-drop main clause**

(2) **SR**
Calculate surprisals

- In SR-S, after "de": V N de "invite tycoon de ..."

(1) pro-drop main clause

(2) SR
Calculate surprisals

• In OR-S, before “de”: N V “tycoon invite ...”

(1) main clause

(2) OR
Calculate surprisals

- In OR-S, after “de”: N V “tycoon invite de ...”

(1) main clause

(2) OR
Calculate surprisals

• In OR-S, after “de”: N V “tycoon invite de ...”

(1) main clause

(2) OR
Results: Obj-modifying RC

Total:
SR 8.04
OR 9.61

Surprisal

V1(SR)N1(OR) 4.17 1.82
N1(SR)V1(OR) 2.25 6.21
N2(head) 1.58 1.58
Results: Obj-modifying RC

Total:
SR  8.04
OR  9.61
Results: Obj-modifying RC

Total:
SR  8.04
OR  9.61

V1(SR) N1(OR)
SR  4.17
OR  1.82

N1(SR) V1(OR)
SR  2.25
OR  6.21

DE
0.04
0

N2(head)
SR  1.58
OR  1.58
Calculate surprisals

- In SR-O, NVV “reporter hit invite ...” → NVV N

SR reading has already been recognized.
Calculate surprisals

• In OR-O, NV N “reporter hit tycoon ...”

(1) main clause

(2) OR
Calculate surprisals

• In OR-Ø, NV NV “reporter hit tycoon invite ...”

(1) main clause

(2) OR
Calculate surprisals

- In OR-O, NV NV “reporter hit tycoon invite ...”
Outline

• Introduction
• Chinese RC
• Modeling
• Conclusion
Conclusion
Conclusion

Surprisal

• uses **structural frequencies** as a reflection of readers’ **linguistic experience**

• models the resolution of **incremental ambiguity**
Conclusion

Surprisal

• uses structural frequencies as a reflection of readers’ linguistic experience

• models the resolution of incremental ambiguity

Results are consistent with recent empirical data

• argue against the memory-based account

• support the experience-based account
Acknowledgment

- Chien-Jer Charles Lin (Indiana)
- Shravan Vasishth (Potsdam)
- Jiwon Yun (Cornell)
- This work is supported by
  - Cornell Cognitive Science Program
  - Cornell East Asian Program
  - a NSF CAREER Award (0741666) to JTH
谢谢！
Thank you!
References


Lin and Bever (2011)

No strong effect

![Graph showing reading times for different conditions in Experiment 2a. The graph includes lines for SR-S, OR-S, SR-O, and OR-O conditions, with reading times measured in milliseconds.]