Overview

- HW 5
- The BIG picture
- Ling 567 preview
HW 5

• 8.5: Conjoined Conjunctions
• 8.6: Arguments in Japanese
• 8.7: Japanese causatives
Big picture: Our model

- Describes a set of strings
- Associates semantic representations (and trees) with well-formed strings
- Is stated in terms of declarative constraints
- ... which are order-independent
- Locates most constraints ‘in the lexicon’
- Is stated in a precise fashion
Parts of our model

- Type hierarchy (lexical types, other types)
- Phrase structure rules
- Lexical rules
- Lexical entries
- Grammatical principles
- Initial symbol
Pause for reflection

• What have you learned about the nature of human language?
• What have you learned about how linguists think about language?
• How does this model/type of model differ from CFG (with atomic categories)?
• In what applications might (atomic category) CFG be sufficient?
• What applications might benefit from something linguistically more motivated?
Ling 567: What is Grammar Engineering?

• Implementation of NL grammars in software

• ... which relate surface strings to semantic representations

• Practical or theoretical focus

• Knowledge-engineering approach to parsing
Ling 567: Goals

• Master tfs formalism
• Hands-on experience with grammar engineering
• Different perspective on NL syntax
• Practice building and debugging extensible system
• Contribute to on-going research
Ling 567: Course requirements

- Build implemented grammar for a language we haven’t worked on before, on the basis of the Grammar Matrix customization system

- Weekly labs (group work): including
  - Test suite development
  - Grammar customization
  - Grammar extension
Ling 567: Grammar coverage

- Major constituent word order
- PNG, pronouns, agreement (if applicable)
- Case
- Tense/aspect
- Negation
- Argument optionality
- Yes-no questions
- Clausal complements
- Coordination, modification, ...
What is the Grammar Matrix?

Elicitation of typological information

Questionnaire definition

Questionnaire (accepts user input)

Choices file

Validation

Grammar creation

Core grammar

Stored analyses

Customization

Customized grammar

HTML generation
What is the Grammar Matrix?

<table>
<thead>
<tr>
<th>Language</th>
<th>Family</th>
<th>Lg-specific types</th>
<th>Matrix types</th>
<th>Head types</th>
<th>Lex rules</th>
<th>Phrasal rules</th>
<th>Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERG</td>
<td>Germanic</td>
<td>3654</td>
<td>N/A</td>
<td>N/A</td>
<td>71</td>
<td>226</td>
<td>N/A</td>
</tr>
<tr>
<td>Breton</td>
<td>Celtic</td>
<td>220</td>
<td>413</td>
<td>510</td>
<td>57</td>
<td>49</td>
<td>1692</td>
</tr>
<tr>
<td>Cherokee</td>
<td>Iroquoian</td>
<td>182</td>
<td>413</td>
<td>510</td>
<td>95</td>
<td>27</td>
<td>985</td>
</tr>
<tr>
<td>French</td>
<td>Romance</td>
<td>137</td>
<td>413</td>
<td>510</td>
<td>29</td>
<td>22</td>
<td>740</td>
</tr>
<tr>
<td>Jamamadjf</td>
<td>Arauan</td>
<td>188</td>
<td>413</td>
<td>510</td>
<td>87</td>
<td>11</td>
<td>1151</td>
</tr>
<tr>
<td>Lushootseed</td>
<td>Salish</td>
<td>95</td>
<td>413</td>
<td>510</td>
<td>20</td>
<td>8</td>
<td>391</td>
</tr>
<tr>
<td>Nishnaabemwin</td>
<td>Algonquian</td>
<td>289</td>
<td>413</td>
<td>510</td>
<td>124</td>
<td>50</td>
<td>1754</td>
</tr>
<tr>
<td>Pashto</td>
<td>Iranian</td>
<td>234</td>
<td>413</td>
<td>510</td>
<td>86</td>
<td>19</td>
<td>1839</td>
</tr>
<tr>
<td>Pali</td>
<td>Indo-Aryan</td>
<td>237</td>
<td>413</td>
<td>510</td>
<td>92</td>
<td>55</td>
<td>1310</td>
</tr>
<tr>
<td>Russian</td>
<td>Slavic</td>
<td>190</td>
<td>413</td>
<td>510</td>
<td>56</td>
<td>35</td>
<td>993</td>
</tr>
<tr>
<td>Shona</td>
<td>Bantu</td>
<td>136</td>
<td>413</td>
<td>510</td>
<td>51</td>
<td>9</td>
<td>591</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>Austro-Asiatic</td>
<td>105</td>
<td>413</td>
<td>510</td>
<td>2</td>
<td>26</td>
<td>362</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td><strong>182.9</strong></td>
<td><strong>413</strong></td>
<td><strong>510</strong></td>
<td><strong>63.5</strong></td>
<td><strong>28.3</strong></td>
<td><strong>1073.5</strong></td>
</tr>
</tbody>
</table>

Table 1: Grammar sizes in comparison to ERG

Bender et al 2010 (ACL posters)
Grammar Matrix: Languages

Bender et al 2010 (ACL posters)

http://depts.washington.edu/uwcl/twiki/bin/view.cgi/Main/LanguagesList
Overview

• HW 5
• The BIG picture
• Ling 567 preview