Learning Metrical Stress Unsanitized
Julie Anne Legate & Charles Yang

Why Stress?
- Major theoretical interests (SPE, Hayes, HV87, H98)
- Longitudinal data (Fikkert 1994), rare in phonology
- How to handle exceptions: all grammars leak

What’s Different?
- Child-directed data, not hand-picked informative examples (Dresher & Kaye, Tesar & Smolensky)
- A principled approach to exceptions

More exceptions⇒Slower processing
- *kick faster* than *lift the bucket*
- German irregular p.p. (-n) *faster* than regular (-t)
- English irregular verbs *faster* than regulars
  ➡ Lexical lookup precedes rule computation
  ➡ Too many exceptions delay rule computation
  ➡ Optimization of time complexity: Rule vs. Lexicaliation

Productive or Lexical?
- *Exception 1*  
- *Exception 2*  
- *Exception 3*  
- *...*  
- *Rule*

A rule applicable to N items can only be productive if the number of exceptions M

Tolerance Principle (TP, Yang 2005)

\[ M < \frac{N}{\ln N} \]

Productivity needs filibuster-proof majority

Learning English Metrical Stress

Acquisition Model

\[ G_1 \rightarrow G_2 \rightarrow G_3 \rightarrow \ldots \rightarrow G_T \]

- A parametric system: quantity, footing directionality ...
- If \( G_i \) fails TP threshold, the learner moves on \( G_{i+1} \)
- If no grammar is productive, then lexicalize everything

Methods
- 4.5 million words from child directed English (CHILDES)
- A Part-of-Speech tagger and CMUDICT to extract stress patterns of words
- Only learning placement of primary stress

Theories of Stress
- Nouns
  - Halle & Vergnaud (1987): Stress final \( \sigma \) if VV; otherwise heavy (VV or VC) penult; otherwise antepenult
  - Halle (1998): Stress final \( \sigma \) if heavy; otherwise antepen
- Verbs (HV1987 and H1998 identical)
  - Stress final \( \sigma \) if VV or VCC\(^+\); otherwise penult.

Early Stage of Stress Learning
- Words at least once 10K: materials for early language
  - 402 words: *Quantity insensitive* (QI)
  - System has only 10 exceptions, well below 402/ln402 = 67
  - Initial stage of stress acquisition is Quantity Insensitive (Fikkert 1994, Kehoe & Stoel-Gammon 1997)
  - Children often mis-stress (balloon balloon): trochaic bias⇒

Later Stage of Stress Learning
- Words at least once per million
- 5763 words, requiring 5097 to be productive
- Even if the initial stress pattern is thoroughly dominant (4960): 80% fails to be productive
- No evidence for a QI stage after 2;0 (Fikkert, Baker & Smith 1976, Guion et al. 2003)

Looking for Productivity: two choices
- **Lex**: Divide words into words and verbs
- **Stem**: Remove inflectional endings

Meeting the TP Standard

<table>
<thead>
<tr>
<th></th>
<th>stem</th>
<th>HV87</th>
<th>H98</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>

515 exceptions vs. 355 exceptions

Stress Shifting Derivations

<table>
<thead>
<tr>
<th></th>
<th>shifting</th>
<th>N</th>
<th>m</th>
<th>valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ary</td>
<td>no</td>
<td>41</td>
<td>8</td>
<td>yes</td>
</tr>
<tr>
<td>-ous</td>
<td>yes</td>
<td>90</td>
<td>30</td>
<td>no</td>
</tr>
</tbody>
</table>

Conclusion
- A principled approach to learning with exceptions
- **kangaroo** vs. **creamery** as core case (HV vs. H)
- Halle 1998: near perfect coverage of English stress

What’s Different?
- Child-directed data, not hand-picked informative examples (Dresher & Kaye, Tesar & Smolensky)
- A principled approach to exceptions