Onset Weight Effects in Stress/Accent Systems
Exhibiting Variation
UD Conference on Stress and Accent

Kevin Ryan
Harvard University

Saturday, December 1, 2012

Overview

- Can onsets contribute to syllable weight? (cf. e.g. Gordon 2005, Topintzi 2010)
- This talk: stress systems exhibiting lexical variation
- Onsets consistently and productively affect stress placement
- P-center theory of weight

Overview

English
- Lexicon
- Productivity
- Analogy

Russian
Etc.
Analysis

English primary stress

- Predictable, but not fully deterministic
  - própane vs. cocaine
  - ptomaine ("compound associated with putrefaction")
  - Mélanie vs. Tennessee

- Generative accounts generally ignore onsets
  (though cf. Nanni 1977 on Eng. adjs. in -ative; also Kelly 2004, infra)

English stress: lexicon

- Simplex disyllables in CELEX (Baayen et al. 1993)
- Longer onset ⇒ greater incidence of primary stress (with Kelly 2004)

Rimes aren’t covertly driving the effect (see last row)

- Even holding both rimes at their modes (˘V.C˘VC), Ø < C and C < CC persist (with Tukey’s HSD p < .001; n = 1,399)

Medial onsets

- Medial onset structure (unaccounted by Kelly 2004) behaves similarly
- Both contrasts persist when initial onset is held at C (Tukey’s HSD < .001)
Overview English Russian Etc. Analysis

Logistic model

- Data
  Simplex English disyllables from CELEX

- Dependent variable
  Initial (1) or final (0) primary stress

- Predictors
  - Initial onset size (0 to 3)
  - Final onset size (1 to 3)
  - Initial coda size (0 to 3)
  - Final coda size (0 to 3)
  - Initial vowel identity (23 levels)
  - Final vowel identity (24 levels)
  - CELEX part of speech (9 levels)
  - log (frequency+1)

Logistic model: results

- All eight main effects significant \( p < .0001 \) in ANOVA

  - Initial onset
    - \( \emptyset < C \) (Tukey's \( p < .0001 \))
    - \( C < CC \) (Tukey's \( p < .0001 \))
    - \( CC < CCC \) (Tukey's \( p < .05 \))

  - Medial onset
    - \( C > CC \) (Tukey's \( p < .0001 \))
    - \( CC > CCC \) (Tukey's \( p < .0001 \))

*Since initial stress is being predicted, the coefficients are reversed

- Even with fully specified rimes as crossed random factors, all but one of the above contrasts persists

Summary: English lexicon

- Stress propensity \( \sim \) onset size broadly supported
  - Robust across various divisions of the lexicon
  - Unconfounded by rime structure
  - Found independently in initial & medial syllables
  - Found independently in disyllables & trisyllables

- Monotonic
  - \( \emptyset < C \) and \( C < CC \) initially
  - \( C < CC \) and \( CC < CCC \) medially

Productivity: previous work

- Is *brontoon* more likely to be initially stressed than *bontoon*?

  - Kelly (2004)
    - Orthographic stimuli pronounced aloud, stresses logged
  
  - Ryan (2011)
    - Self-reported judgments of orthographic stimuli
    - Screening based on performance on real words

- Both: initial \( C < CC \) in disyllables

Perception experiment

- Auditory wug-test

- Addresses possible problems with orthographic stimuli
  (e.g. “visual syllable” confound)

Stimuli

- Critical items: 8 wugs in 3 nested conditions (crossed by participant)

<table>
<thead>
<tr>
<th>C-</th>
<th>CC-</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ummorm</td>
<td>ummorm</td>
</tr>
<tr>
<td>2. izzoof</td>
<td>izzzoof</td>
</tr>
<tr>
<td>3. izzoof</td>
<td>izzzoof</td>
</tr>
<tr>
<td>4. illawm</td>
<td>illawm</td>
</tr>
<tr>
<td>5. izzooit</td>
<td>izzzooid</td>
</tr>
<tr>
<td>6. izzooit</td>
<td>izzzooid</td>
</tr>
<tr>
<td>7. evvain</td>
<td>devvain</td>
</tr>
<tr>
<td>8. izzotaistrow</td>
<td>listrow</td>
</tr>
</tbody>
</table>
  
- Fillers: 8 real disyllables with C onsets, 50% trochaic

- Critical items and fillers randomized
  (except first two item fillers and then no two adjacent fillers)

- All items recorded and processed in Praat (Boersma and Weenink 2011)
  - Pitch \( \rightarrow 150 \) Hz
  - Intensity \( \rightarrow 65 \) dB
  - Onsets spliced onto same completion across conditions
    (a fixed portion of the vowel was also replaced for natural-sounding transitions)
    
    izztw; listrow; slistrow

- Amazon’s Mechanical Turk
  - Analyzed if US location, native speaker, and 7+/8 on fillers
  - 38 usable participants (from 166)
Results

- Φ- 43% trochaic, C- 63%, CC- 79% (ANOVA F(2) = 9.8, p < .0001)
- Φ < C and Φ < CC both significant in a mixed model (Tukey’s p < .05)
- Neighbors of pseudoword plizzoof

Russian stress/accent

- Like English, not fully predictable in roots
  - 100+ minimal pairs (e.g. milka ‘torment’ vs. runka ‘flour’) (Calhoun 2002)
- But non-deterministic ≠ unpredictable (e.g. Zarea 2010)
- Corpus
  - 32,616-lemma frequency list (Russel 2002)
  - Excluded monosyllables & compounds
  - Accents supplied from online dictionary (starling.rinet.ru)
  - Excluded items with mobile stress
  - Result: 11,757 nouns; 5,258 adjectives; 7,399 verbs

Hayes (2012, handout):

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Pseudo/Temp</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress Pseud when its vowel is Low</td>
<td>7335/699</td>
<td>3.16</td>
</tr>
<tr>
<td>Stress Pseud when it has a triple onset (CCC)</td>
<td>180</td>
<td>0.92</td>
</tr>
<tr>
<td>Stress Pseud when it has (at least) a double onset (CC)</td>
<td>2596178</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Russian initial onset size & accent

- In trisyllabic lemmata (modal word length):

Russian: logistic regression

- Model set up as for English
- Trisyllables (all Tukey’s p < .0001)
  - Φ < C
  - C < CC
  - CC < CCC+
- Disyllables (plot not shown) (both Tukey’s p < .0001)
  - Φ < C
  - C < CC

- As in English, onset structure & stress propensity covary significantly & systematically
Convergent evidence from poetic meter, e.g. Sanskrit

- All else equal, syllables with longer onsets are more skewed towards strong positions

![Convergent evidence from poetic meter, e.g. Sanskrit](image)

### Proposal: P-center theory of weight

- The domain over which weight percept is assessed begins not with the rime, but with the perceptual center of the syllable (on p-centers, e.g. Patel et al. 1999, Villing et al. 2003, Soraghan et al. 2005, Barbosa et al. 2005, Tilsen 2006, Port 2007, Wright 2008, Villing 2010)

- Near beginning of rime, but perturbed by onset structure

![Proposal: P-center theory of weight](image)

- As onset size increases, p-centers increasingly anticipate the rime, but only by a small fraction of the duration of the onset

![Proposal: P-center theory of weight](image)

### Conclusion

- Stress & meter are rhythmic phenomena
- Timing/isochrony studies (e.g. op. cit.) suggest linguistic rhythm is not anchored to (sub)syllabic structure per se
- Auditory recovery (Geocha 2005): unlikely to be the whole story
  - Ceiling ~40 ms (Dolens 1982/1983), Whereas events well outside of this window affect stress attraction (e.g. slide 27)
  - Null onset problem (Geocha 2005)
  - Geminate onset problem (Rispens 2010/2011)

- Onsets contribute to syllable weight
  - Their influence is clearest in, but not exclusive to, non-deterministic systems
- But they contribute less than codas
  - Less likely to be invoked by categorical criteria
  - Smaller coefficient in non-deterministic systems
- The p-center hypothesis can explain this asymmetry
  - Onsets parsed only partially into weight domain
  - Assuming categorization optimizes both perceptual dispersion & formal simplicity (Geocha 2003), criteria are expected to favor codas