Prosodic Variability and Stylistic Coherence

1) Why rhythm?
   a) The rhythm class hypothesis continues to generate an impressive amount of research:
   b) Rhythmic information helps listeners identify dialects they are familiar with (Szakay 2008). It appears to be salient to newborns (Nazzi and Ramus 2003) and non-human primates (Tincoff et al. 2005).
   c) There has been a focus on dialectal variation in rhythm (Low, Grabe, and Nolan 2000; Thomas and Carter 2006).

2) How to do rhythm
   a) PVI (Grabe and Low 2002; Low, Grabe, and Nolan 2000)
      i) My version (individual scores, normalized):
         \[ 100 \times \frac{|d_i - d_{i+1}|}{(d_i + d_{i+1})/2} \]
   b) % V, ΔV, ΔC (Ramus, Nespor, and Mehler 2000)
   c) "VI" Variability index (Deterding 2001)
   d) varcoΔC, varcoΔS, etc. (Dellwo 2006; Mok and Dellwo 2008)

3) Problems with rhythm
   a) Rhythm may be epiphenomenal, an (imperfect) reflection of phonological structure (e.g., phonotactics, vowel reduction) (Dauer 1983)
   b) Popular measures of rhythm are prone to high degrees of individual variability, or may be determined by other factors (e.g. speech rate), or may not really capture "rhythm"
   c) Everyone (including me) implements PVI differently, usually with consequences for the shape of the measurements.
      (1) Low et al. take a mean score for each utterance; Grabe and Low do not divide passages into utterances, and do include final syllables.
      (2) Thomas and Carter do segment utterances, do not include final feet or take utterance means.
      (3) Some compare populations of PVI scores with statistics (Low et al., Grabe and Low). Others (Mok and many others) use summary statistics, e.g. mean. Others (Thomas and Carter) do both.
   d) Final lengthening is also important: it may be a confound for PVI studies, and it may be related to the organization of isochronous prosodic units (Wagner and Dellwo 2004).
   e) Evaluation of rhythm measures take an ambivalent stance toward variability.
      i) Principled fretting about comparability, replicability, etc.
      ii) Demonstrated interest in sociolinguistic dimensions of rhythm (Thomas and Carter 2006; Grabe 2002)

4) My study
   a) Mandarin Chinese
   b) Data from Nüren hua 'Women flowers,' Chinese historical TV drama (2008)
      i) 6 speakers, 3 women, 3 men:
Table 1. Speakers

<table>
<thead>
<tr>
<th>Character</th>
<th>Description</th>
<th>Actor</th>
<th>Birthplace</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huang Mei’er</td>
<td>Female, opera performer</td>
<td>Liu Tao</td>
<td>Nanchang</td>
</tr>
<tr>
<td>Li Qingquan</td>
<td>Male, army captain</td>
<td>Zhang Chao</td>
<td>Beijing</td>
</tr>
<tr>
<td>Lin Xuelian</td>
<td>Female, teahouse operator</td>
<td>Qin Hailu</td>
<td>Dalian</td>
</tr>
<tr>
<td>Liu Jianxiong</td>
<td>Male, profligate/businessman</td>
<td>Cheng Sihan</td>
<td>Heilongjiang</td>
</tr>
<tr>
<td>Ouyang Xiu</td>
<td>Female, reformer/schoolteacher</td>
<td>Zhang Tong</td>
<td>Xinjiang</td>
</tr>
<tr>
<td>Wu Yusheng</td>
<td>Male, lawyer/judge</td>
<td>Feng Zhaofeng</td>
<td>Shanghai</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measures</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i)</td>
<td>Look at whole syllable durations, rather than vocalic/intervocalic intervals</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>This may be more appropriate for conversational speech (Deterding 2001).</td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>Intonational phrases segmented according to criteria in Peng, Tseng, Huang, Lee &amp; Beckman (2006); breaks &quot;3&quot; and stronger counted.</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>Syllabification decisions made on the basis of citation forms of characters.</td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>Segmentation standards based on presence of formant structure, intensity levels, and then auditory criteria.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measures</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i)</td>
<td>Normalized syllable PVI (nSPVI)</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>Take utterance medians.</td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>I do not use comparisons involving the final two syllables of an utterance.</td>
<td></td>
</tr>
<tr>
<td>ii)</td>
<td>varco∆S</td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>Unlike Dellwo, compute per-utterance scores.</td>
<td></td>
</tr>
<tr>
<td>(a)</td>
<td>[ \frac{s}{d} ] where ( s ) is the sample standard deviation of syllable duration ( d ) in the IP</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measures</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>i)</td>
<td>(Separate) ANOVAs for nSPVI, varco∆S by gender, speaker.</td>
<td></td>
</tr>
<tr>
<td>ii)</td>
<td>For final lengthening, smoothing spline ANOVA (Gu 2009; Davidson 2006; Koops 2009) with syllable duration as response, speaker or gender as grouping factor, syllable position as continuous variable.</td>
<td></td>
</tr>
</tbody>
</table>

5) Gender and rhythm

a) differences for both nSPVI and varco∆S:

Table 2. ANOVA: nSPVI and gender

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>0.2215</td>
<td>0.2215</td>
<td>4.669</td>
<td>0.03154</td>
</tr>
<tr>
<td>Residuals</td>
<td>286</td>
<td>13.5696</td>
<td>0.0474</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

i) estimated effect = 5.55 points

Table 3. ANOVA: varco∆S and gender

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>4444</td>
<td>4444</td>
<td>27.618</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Residuals</td>
<td>286</td>
<td>46022</td>
<td>161</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ii) estimated effect = 7.86 points
6) Speaker and rhythm

Figure 1. Median varco$\Delta S$ and nSPVI scores. Right chart w/ 95% confidence intervals from ANOVAs with speaker as response. Reference values from Mok (2008)

- a) By nSPVI, more syllable-timed on average than in Mok. No apparent difference for varco$S$.
- b) Variability in speakers parallels variability in languages.
- c) Women (Huang, Ouyang, Lin) more stress-timed than men (Li, Liu, Wu).
- d) There are significant differences among speakers for varco$\Delta S$, but not for nSPVI:

Table 4. ANOVA: nSPVI and speaker

<table>
<thead>
<tr>
<th>DF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>5</td>
<td>0.3317</td>
<td>0.0663</td>
<td>1.3897</td>
</tr>
<tr>
<td>Residuals</td>
<td>282</td>
<td>13.4595</td>
<td>0.0477</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. ANOVA: varco$\Delta S$ and speaker

<table>
<thead>
<tr>
<th>DF</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaker</td>
<td>5</td>
<td>5299</td>
<td>1060</td>
<td>6.6168</td>
</tr>
<tr>
<td>Residuals</td>
<td>282</td>
<td>45167</td>
<td>160</td>
<td></td>
</tr>
</tbody>
</table>

i) Tukey tests reveal that the women (Ouyang, Huang, Lin) are different from Li and Liu but not Wu.
7) Final lengthening

Figure 2. Left: SSANOVA estimated final syllable durations, by speaker with 95% CIs. Right, estimated syllable durations by gender, with 95% CIs.

a) Gender difference—women characters do more final lengthening.

b) No significant differences in lengthening among women. Among men, Wu and Li significantly different.

8) Individual differences in rhythm and prosody?

a) Female characters more syllable-timed, do more final lengthening than men.

b) Wu apparently more syllable-timed, does more final lengthening than Li.

i) Can you spot the difference?

ii) Wu and Li both Nationalists, modernizers, young, foreign-educated

(1) wen/wu: refinement vs. martiality

(a) Perhaps, in an instance of fractal recursivity (Irvine and Gal 2009), the distinction between Wu (wen) and Li (wu) is constructed as similar to the difference between women and men.

(i) This would explain the direction of the difference and its reproduction in varcoΔS and final lengthening.

1 Source: http://a2.att.hudong.com/62/09/01300000013054121328098587955.jpg

2 Source: http://ent.people.com.cn/mediafile/200802/20/F200802201337211173117007.jpg
9) Conclusions
   a) More speakers would be needed for a proper sociolinguistic survey, and it would be
      interesting to apply the more traditional vocalic nPVI/intervocalic rPVI analysis.
   b) There are, nevertheless, indications that the study of rhythm, using PVI or other
      measures, can be brought from the realm of typology and dialectology into more
      traditional variationist sociolinguistic territory, such as gender.
   c) Final lengthening also has a sociolinguistic dimension, which perhaps only Kiesling
      (2005) has before recognized.
   d) Individual variability on culturally specific dimensions like wen/wu may help explain
      some of the insalutary variability that makes PVI (much like mean F0 or range) such a
      "volatile" measure.
   e) If so, it remains unclear how sociolinguistically meaningful rhythmic variability interacts
      with variability due to speech rate or other performance factors, or style shifts based on
      elicitation task (Mok 2008).

References
Davidson, Lisa. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of
Frankfurt am Main, Germany: Peter Lang.
project.org/web/packages/gss/index.html.
*English World-Wide* 26:1-42.
Koops, Christian. 2009. /uw/-monophthongization and Southern style. *New Ways of Analyzing Variation 38*.
Ottawa, ON, Canada.
Low, Ee Ling, Esther Grabe, and Francis Nolan. 2000. Quantitative Characterizations of Speech Rhythm: Syllable-
Mok, Peggy, and Volker Dellwo. 2008. Comparing native and non-native speech rhythm using acoustic rhythmic
Communication* 41:233-244.
Peng, Shu-hui, Marjorie K M Chan, Chiu-yu Tseng, Tsan Huang, Ok Joo Lee, and Mary E Beckman. 2006. Towards a
Pan-Mandarin System for Prosodic Transcription. *Prosodic Typology: The Phonology of Intonation and
Tincoff, R., M. Hauser, F. Tsao, G. Spaepen, Franck Ramus, and Jacques Mehler. 2005. The role of speech rhythm in
Wagner, Petra S., and Volker Dellwo. 2004. Introducing YARD (Yet Another Rhythm Determination) and re-