Using durational measures with non-native speech rhythm

Peggy Mok
Department of Linguistics and Modern Languages, The Chinese University of Hong Kong
peggymok@cuhk.edu.hk

Introduction

• This study investigates the reliability and effectiveness of various durational rhythmic measures by using non-native speech materials.

• Durational measures like ΔC and %V [1], VarcoC [2], raw consonantal and normalised vocalic PVIs [3] can classify languages into rhythmic clusters.

• These measures were developed using speech materials from native speakers of various languages.

• Some studies also used these measures to investigate non-native speech rhythm, e.g. Taiwan English [4], Dutch, English and Spanish [5].

• Native and non-native speech differ in many aspects. However, the reliability of these rhythmic measures on non-native speech was not clear.

• If the rhythmic measures are robust and reliable, they should classify both native and non-native speech equally well.

Cantonese and Beijing Mandarin

• Cantonese and Beijing Mandarin are syllable-timed languages [6], [7]. The accented English spoken by these speakers also sound syllable-timed [8].

• Both languages have a very simple syllable structure: Cantonese: CV, CVN, CVSM Mandarin: CV, CVN

• Both languages are tone languages.

• Cantonese has no lexical stress and no phonological vowel reduction.

• Mandarin has some phonologically unstressed syllables ("neutral tone").

Research question

• Can the durational rhythmic measures consistently classify Cantonese and Beijing Mandarin (native), Cantonese-accented English and Mandarin-accented English (non-native) as syllable-timed?

Method

• 6 native speakers of Cantonese and 6 of Beijing Mandarin (3 male, 3 female) were recorded reading the North Wing and the Sun story with a normal speech rate in their native languages and English.

• Duration of vocalic (V), consonantal (C) and syllabic (S) intervals were used with the following rhythmic measures:

  ΔC, ΔS, ΔV, VarcoC, VarcoS, rPVI_C, rPVI_S, nPVI_V, nPVI_S

• The data was compared with four languages in the BonnTempo Corpus [9]: German and British English (stressed-timed) and French and Italian (syllable-timed).

Results

Table 1. Values of four rhythmic measures using syllable duration.

<table>
<thead>
<tr>
<th>Lang</th>
<th>ΔS</th>
<th>VarcoC</th>
<th>ΔC</th>
<th>rPVI_C</th>
<th>nPVI_S</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEng</td>
<td>124.5</td>
<td>Eng</td>
<td>51.9</td>
<td>MEng</td>
<td>151.1</td>
</tr>
<tr>
<td>CEng</td>
<td>151.1</td>
<td>Eng</td>
<td>69.7</td>
<td>CEng</td>
<td>124.7</td>
</tr>
<tr>
<td>Eng</td>
<td>106.8</td>
<td>MEng</td>
<td>49.3</td>
<td>CEng</td>
<td>60.7</td>
</tr>
<tr>
<td>Man</td>
<td>57.8</td>
<td>Ger</td>
<td>43.5</td>
<td>Man</td>
<td>86.1</td>
</tr>
<tr>
<td>Ita</td>
<td>75.8</td>
<td>Ger</td>
<td>115.5</td>
<td>Man</td>
<td>54.8</td>
</tr>
<tr>
<td>Fr</td>
<td>124.7</td>
<td>MEng</td>
<td>57.7</td>
<td>Ita</td>
<td>49.5</td>
</tr>
<tr>
<td>Can</td>
<td>57.5</td>
<td>CEng</td>
<td>47.4</td>
<td>Can</td>
<td>106.8</td>
</tr>
<tr>
<td>Man</td>
<td>53.3</td>
<td>CEng</td>
<td>47.4</td>
<td>Ita</td>
<td>38.2</td>
</tr>
<tr>
<td>Ita</td>
<td>67.6</td>
<td>Man</td>
<td>38.2</td>
<td>Fr</td>
<td>82.7</td>
</tr>
<tr>
<td>Can</td>
<td>55.3</td>
<td>Can</td>
<td>63.6</td>
<td>Fr</td>
<td>36.2</td>
</tr>
<tr>
<td>Fr</td>
<td>55.3</td>
<td>Can</td>
<td>30.7</td>
<td>Can</td>
<td>63.6</td>
</tr>
<tr>
<td>Man</td>
<td>53.3</td>
<td>Can</td>
<td>34.3</td>
<td>Can</td>
<td>34.3</td>
</tr>
</tbody>
</table>

Conclusions

• All durational rhythmic measures confirm the syllable-timing of Cantonese and Beijing Mandarin.

• Except VarcoC and %V, the other measures all suggest that the two non-native English accents are numerically closer to stress-timed than syllable-timed languages, contrary to the auditory impression.

• A slower speaking rate and selective-lengthening in non-native speech contribute to this discrepancy.

• Native and non-native speech differ in many aspects. However, the reliability of these rhythmic measures on non-native speech was not clear.

• If the rhythmic measures are robust and reliable, they should classify both native and non-native speech equally well.

References


ACKNOWLEDGEMENTS

The author would like to thank Volker Dellwo for helpful discussion and generous sharing of his data.