A CROSS-LANGUAGE VOWEL NORMALISATION PROCEDURE

Geoffrey Stewart Morrison*
&
Terrance M. Nearey
University of Alberta

Research supported by:
Social Sciences and Humanities Research Council of Canada
Conseil de recherches en science humaines du Canada

*now at Boston University
Single language/dialect
Single language/dialect

vocal-tract length differences

\ln(F1)

\ln(F2)

\ln(F1)
Log-mean normalisation
Nearey (1978)

deviation from speaker mean

\[
\ln(F1) = \ln(F2)
\]
Log-mean normalisation

slide so speaker means have same reference value
Log-mean normalisation

deviation from language/dialect reference value
Making a number of simplifying assumptions about language and dialect differences:
Multiple languages/dialects

differences in inventory pattern number and distribution of phonemes (size & skew) affect speaker means
Ideal bilingual

$G_L$ due to inventory differences, not vocal tract differences
Estimate $G_L$ from balanced samples of speakers from each language.
Cross-Language Vowel Normalisation:

perception of an instance of a vowel from language B (Spanish)
in terms of vowel categories from language A (English)
log-mean normalise all English speakers' vowels, train model
normalise a single token of a Spanish vowel
Within-language normalised

within-language normalised token of a Spanish vowel
Cross-language normalised

add/subtract $G_L$
Evaluation data:

**English:** /i/, /ɪ/, /e/, /ɛ/

**Spanish:** /i/, /ei/, /e/

Acoustic variables:

F1, F2 at 25% duration of vowel

ΔF1, ΔF2 (difference from 25-75% duration of vowel)

duration
Statistical model:

discriminant analysis
trained on English vowels
used to classify instances of Spanish vowels
  a posteriori probabilities (APPs)

3 versions:
  • non-normalised
  • within-language normalised
  • cross-language normalised
Monolingual English listeners: classified instances of Spanish vowels in terms of English vowel categories proportions (pooled across listeners)

Test value: correlation between model APPs and listener proportions
Results:

<table>
<thead>
<tr>
<th>Model</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-normalised</td>
<td>$r = .848$</td>
</tr>
<tr>
<td>within-language normalised</td>
<td>$r = .853$</td>
</tr>
<tr>
<td>cross-language normalised</td>
<td>$r = .869$</td>
</tr>
</tbody>
</table>
Conclusion:

The cross-language vowel normalisation procedure increased the correlation between the classification of Spanish vowels by a model trained on L1-English vowel productions and L1-English listeners’ perception of Spanish vowels.
Canonical Discriminant Function 1

Canonical Discriminant Function 2

Eng /ɪ/

Eng /i/

Eng /ɛ/

Eng /e/

Eng /ɛ/
<table>
<thead>
<tr>
<th></th>
<th>Produced</th>
<th>Perceived</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eng /i/</td>
<td>Eng /i/</td>
</tr>
<tr>
<td>Sp /i/</td>
<td>.951</td>
<td>.036</td>
</tr>
<tr>
<td>Sp /ei/</td>
<td>.005</td>
<td>.003</td>
</tr>
<tr>
<td>Sp /e/</td>
<td>.004</td>
<td>.275</td>
</tr>
</tbody>
</table>

### Model

<table>
<thead>
<tr>
<th></th>
<th>Produced</th>
<th>Perceived</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eng /i/</td>
<td>Eng /i/</td>
</tr>
<tr>
<td>Sp /i/</td>
<td>.997</td>
<td>.001</td>
</tr>
<tr>
<td>Sp /ei/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp /e/</td>
<td>.014</td>
<td>.583</td>
</tr>
</tbody>
</table>