Automatically identifying characteristic features of non-native English accents

Jelke Bloem, Martijn Wieling and John Nerbonne
Outline

- Background
- Measure
- Material
- Results: Hungarian
- Results: Dutch
- Discussion
Background: Dialectometry

- The measurement of dialect differences
- Aggregates over features
  - Phonetic distance
- Tool: GabMap  [www.gabmap.nl](http://www.gabmap.nl)
Dialectometry: Characteristic features
Prokić, Çöltekin & Nerbonne (2012)

- Which word is typical of each dialect area?
- Based on edit distance between transcriptions of the same word
- Apply to accents?
  Levenshtein distance correlates with human native-likeness judgements ($r = -0.81$) (Wieling et al. 2014)
Accents: Characteristic features? Errors?

- Background of pronunciation training
- Errors likely to stem from L1 interference
- Some errors are ‘worse’ than others
- Certain features on which to concentrate in pronunciation training (Abercrombie, 1956):
  - Hierarchy of errors
- Severe is not necessarily characteristic
Dutch error hierarchy
(Van den Doel, 2006)

- Judgement task for native speakers
- 32 sentences, each contains 1 error
- Over 500 native speakers judged severity of errors
- From this, a hierarchy of errors was compiled
- Closest thing to ranking of characteristic features
Measure

- **Representativeness**
  - Difference within the L1 group is small

- **Distinctiveness**
  - Difference between L1 and native groups is large

Mean difference over speaker pairs for some feature

- Final score is the difference between the z-scores of distinctiveness and representatives

From Prokić et al. (2012)
Material: Speech Accent Archive
http://accent.gmu.edu/

The speech accent archive uniformly presents a large set of speech samples from a variety of language backgrounds. Native and non-native speakers of English read the same paragraph and are carefully transcribed. The archive is used by people who wish to compare and analyze the accents of different English speakers.

last updated: 20 November 2014  2023 samples
Can observe some characteristics from examples
But: individual variation
Production errors possible
-> Better to study aggregate data
Results

- Hungarian
  - “Background to English pronunciation.” (Nádasdy, 2006)

- Dutch
  - Perception study (Van den Doel, 2006)

- French
  - French sound structure (Walker, 2001)
Hungarian

- Has dental sounds and fricatives, but no dental fricatives
- Perceived /v/, produced /d/

<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>these [1]</td>
<td>2.06</td>
<td>díz (2/7 : 0/181)</td>
<td>óiːz (0/7 : 35/181) óiz (0/7 : 19/181)</td>
</tr>
<tr>
<td>2</td>
<td>please</td>
<td>1.70</td>
<td>plis (4/7 : 1/181)</td>
<td>phliːz (0/7 : 39/181) phliːz(0/7 : 31/181)</td>
</tr>
<tr>
<td>3</td>
<td>big</td>
<td>1.69</td>
<td>bik (5/7 : 0/181)</td>
<td>big(0/7 : 77/181)</td>
</tr>
<tr>
<td>4</td>
<td>these [2]</td>
<td>1.55</td>
<td>díz (1/7 : 1/181)</td>
<td>óiːz (0/7 : 59/181) óiːz (0/7 : 38/181)</td>
</tr>
<tr>
<td>5</td>
<td>the [1]</td>
<td>1.52</td>
<td>də (1/7 : 0/181)</td>
<td>də (0/7 : 97/181) nə (0/7 : 64/181)</td>
</tr>
</tbody>
</table>

Identifying characteristic features:

- Has dental sounds and fricatives, but no dental fricatives
- Perceived /v/, produced /d/

---

University of Amsterdam
### Hungarian

*cf. Nádasdy (2006)*

#### Regressive assimilation
- Take voicedness of next C, even across word boundaries

<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>these [1]</td>
<td>2.06</td>
<td>ḍiːs (5/7 : 3/181) ḍiːz (2/7 : 0/181)</td>
<td>ḍiːz (0/7 : 35/181) ḍiːz (0/7 : 19/181)</td>
</tr>
<tr>
<td></td>
<td>“these things”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>please</td>
<td>1.70</td>
<td>pʰliːs (4/7 : 1/181) pʰliː (2/7 : 5/181)</td>
<td>pʰliːz (0/7 : 39/181) pʰliː (0/7 : 31/181)</td>
</tr>
<tr>
<td></td>
<td>“please call”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>big</td>
<td>1.69</td>
<td>bɪk (5/7 : 0/181) bɪk (1/7 : 1/181)</td>
<td>bɪg (0/7 : 77/181)</td>
</tr>
<tr>
<td></td>
<td>“big toy”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>these [2]</td>
<td>1.55</td>
<td>ḍiːs (4/7 : 1/181) ḍiːz (1/7 : 1/181)</td>
<td>ḍiːz (0/7 : 59/181) ḍiːz (0/7 : 38/181)</td>
</tr>
<tr>
<td></td>
<td>“these things”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>the [1]</td>
<td>1.52</td>
<td>ɚ (6/7 : 3/181) ɚ (1/7 : 0/181)</td>
<td>ɚ (0/7 : 97/181) ɚ (0/7 : 64/181)</td>
</tr>
</tbody>
</table>

**red bags: 100% /d/**

- Identifying characteristic features
## Dutch: Error hierarchy

**Van den Doel (2006)**

<table>
<thead>
<tr>
<th>Severity</th>
<th>British (RP)</th>
<th>American English</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 3.5</td>
<td>Fortis/lenis neutralization Use of uvular-(r) /(R)/</td>
<td>Fortis/lenis neutralization Use of uvular-(r) /(R)/</td>
</tr>
<tr>
<td>2.2-3.5</td>
<td>Fortis/lenis neutralization Use of uvular-(r) /(R)/ /(θ,ð)/ ~ /t,d/ Glottalization of final /d/ Epenthetic [(ə)] in /lm/ /(v) ~ w/ confusion /(æ) ~ e, (ʌ) ~ o, (ʊ) ~ u:/ Unaspirated [t]</td>
<td>/(θ,ð)/ ~ /t,d/ Glottalization of final /d/ Epenthetic [(ə)] in /lm/ /(v) ~ w/ confusion /(æ) ~ e/ confusion Inappropriate post-vocalic (r)</td>
</tr>
</tbody>
</table>
### Use of /r/

**Dutch**


<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>big</td>
<td>1.92</td>
<td>bik (13/16 : 1/181)</td>
<td>big (0/16 : 77/181) big` (0/16 : 41/181)</td>
</tr>
<tr>
<td>2</td>
<td>to</td>
<td>1.22</td>
<td>tu (10/16 : 11/181)</td>
<td>re (0/16 : 112/181)</td>
</tr>
<tr>
<td>3</td>
<td>slabs</td>
<td>1.12</td>
<td>slæps (5/16 : 0/181)</td>
<td>slæbz (1/16 : 66/181) slæ:bz (0/16 : 38/181)</td>
</tr>
<tr>
<td>4</td>
<td>bags</td>
<td>1.08</td>
<td>bæks (4/16 : 0/181)</td>
<td>bægz (1/16 : 39/181) bæ:gz (0/16 : 33/181)</td>
</tr>
<tr>
<td>5</td>
<td>of [1]</td>
<td>1.06</td>
<td>of (7/16 : 7/181)</td>
<td>əv (0/16 : 58/181) əf (7/16 : 46/181)</td>
</tr>
</tbody>
</table>

---

Identifying characteristic features
Dutch


19. for [2]: Only /r/ or Ø

- In fact, /R/ is never used
Dutch  


<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>big</td>
<td>1.92</td>
<td>bɪk (13/16 : 1/181)</td>
<td>bɪg (0/16 : 77/181) bɪg̴ (0/16 : 41/181)</td>
</tr>
<tr>
<td>2</td>
<td>to</td>
<td>1.22</td>
<td>tu (10/16 : 11/181)</td>
<td>ɾə (0/16 : 112/181)</td>
</tr>
<tr>
<td>3</td>
<td>slabs</td>
<td>1.12</td>
<td>slæps (5/16 : 0/181)</td>
<td>slæbz (1/16 : 66/181) slæːbz (0/16 : 38/181)</td>
</tr>
<tr>
<td>4</td>
<td>bags</td>
<td>1.08</td>
<td>bæks (4/16 : 0/181)</td>
<td>bæːgz (1/16 : 39/181) bæːgz (0/16 : 33/181)</td>
</tr>
<tr>
<td>5</td>
<td>of [1]</td>
<td>1.06</td>
<td>of (7/16 : 7/181)</td>
<td>əv (0/16 : 58/181) əf (7/16 : 46/181)</td>
</tr>
</tbody>
</table>

- **Fortis/lenis neutralization** (in Dutch, basically devoicing of final consonant cluster)
# Dutch


## Dental fricatives: /θ, ð/ ~ /t, d/  

<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>big</td>
<td>1.92</td>
<td>bɪk (13/16 : 1/181)</td>
<td>bɪɡ (0/16 : 77/181) bɪɡ̆ (0/16 : 41/181)</td>
</tr>
<tr>
<td>2</td>
<td>to</td>
<td>1.22</td>
<td>tu (10/16 : 11/181)</td>
<td>rə (0/16 : 112/181)</td>
</tr>
<tr>
<td>3</td>
<td>slabs</td>
<td>1.12</td>
<td>slæps (5/16 : 0/181) slæbps (3/16 : 1/181)</td>
<td>slæbz (1/16 : 66/181) slæːbz (0/16 : 38/181)</td>
</tr>
<tr>
<td>4</td>
<td>bags</td>
<td>1.08</td>
<td>bæks (4/16 : 0/181) bæːgs (3/16 : 2/181)</td>
<td>bægz (1/16 : 39/181) bæːgz (0/16 : 33/181)</td>
</tr>
<tr>
<td>5</td>
<td>of [1]</td>
<td>1.06</td>
<td>əf (7/16 : 7/181)</td>
<td>əv (0/16 : 58/181) əf (7/16 : 46/181)</td>
</tr>
</tbody>
</table>
Dental fricatives: /θ, ð/ ~ /t, d/

Dutch

17. the [2]:

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>[d]</td>
<td>7</td>
</tr>
<tr>
<td>[g]</td>
<td>4</td>
</tr>
<tr>
<td>[ð]</td>
<td>3</td>
</tr>
</tbody>
</table>

Not representative, but distinctive!
## Dutch


### Dental fricatives: /θ, ð/ ~ /t, d/
Sorted by distinctiveness

<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Distinctiveness score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>slabs</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>the [2]</td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>the [3]</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>big</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>these [2]</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dutch

- Epenthetic [ə] in /lm/: No /lm/ in paragraph
- /v ~ w/ confusion: very rare, 1 or 2 speakers
- Glottalization of final /d/: not in *red bags*
- /æ ~ e/ confusion (*slabs*):

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>[æ]</td>
<td>12</td>
</tr>
<tr>
<td>[ɛ]</td>
<td>1</td>
</tr>
<tr>
<td>[a]</td>
<td>1</td>
</tr>
</tbody>
</table>
Dutch: Error hierarchy
Van den Doel (2006)

<table>
<thead>
<tr>
<th>Severity</th>
<th>British (RP)</th>
<th>American English</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 3.5</td>
<td>Fortis/lenis neutralization&lt;br&gt;Use of uvular-r /ʁ/&lt;br&gt;Glottalization of final /d/&lt;br&gt;Epenthetic [ə] in /lm/&lt;br&gt;/v ~ w/ confusion&lt;br&gt;/æ ~ e, ʌ ~ ɒ, ʊ ~ u:/&lt;br&gt;Unaspirated [t]</td>
<td>Fortis/lenis neutralization&lt;br&gt;Use of uvular-r /ʁ/</td>
</tr>
<tr>
<td>2.2-3.5</td>
<td>Fortis/lenis neutralization&lt;br&gt;Use of uvular-r /ʁ/&lt;br&gt;Glottalization of final /d/&lt;br&gt;Epenthetic [ə] in /lm/&lt;br&gt;/v ~ w/ confusion&lt;br&gt;/æ ~ e, ʌ ~ ɒ, ʊ ~ u:/&lt;br&gt;Unaspirated [t]</td>
<td>/θ, ð/ ~ /t,d/&lt;br&gt;Glottalization of final /d/&lt;br&gt;Epenthetic [ə] in /lm/&lt;br&gt;/v ~ w/ confusion&lt;br&gt;/æ ~ e/ confusion&lt;br&gt;Inappropriate post-vocalic r</td>
</tr>
</tbody>
</table>

- “Severe” and “characteristic” are different things
- Characteristic features measure provides new information
Conclusions

- Statistical evidence for characteristic features
- Characteristic features are L1 interference
- Measure of characteristicness only somewhat overlaps with perception of severity
- Distinctiveness can also identify some errors
- New source of information on accents
- Can be applied to other languages (if transcribed)
## Homework: Polish

<table>
<thead>
<tr>
<th>R</th>
<th>Word</th>
<th>Score</th>
<th>Characteristic forms</th>
<th>Native forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>kids</td>
<td>1.24</td>
<td>kĭts (5/13 : 7/181)</td>
<td>kĭdz (0/13 : 24/181)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>kĭds (2/13 : 7/181)</td>
<td>kʰĭdz (0/13 : 16/181)</td>
</tr>
<tr>
<td>2</td>
<td>please</td>
<td>1.11</td>
<td>plis (5/13 : 1/181)</td>
<td>pʰliːz (0/13 : 39/181)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>əplis (2/13 : 0/181)</td>
<td>pʰliːz (0/13 : 31/181)</td>
</tr>
<tr>
<td>3</td>
<td>these [1]</td>
<td>0.87</td>
<td>ɡîs (3/13 : 0/181)</td>
<td>dioːz (0/13 : 35/181)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ɡis (3/13 : 3/181)</td>
<td>ɡiz (1/13 : 19/181)</td>
</tr>
<tr>
<td>4</td>
<td>call</td>
<td>0.80</td>
<td>kɔl (6/13 : 12/181)</td>
<td>kʰal̞ (0/13 : 48/181)</td>
</tr>
<tr>
<td>5</td>
<td>for [1]</td>
<td>0.78</td>
<td>fɔr (7/13 : 24/181)</td>
<td>fə (0/13 : 60/181)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fɔ (2/13 : 7/181)</td>
<td>fə (0/13 : 26/181)</td>
</tr>
</tbody>
</table>
References


Van den Doel, Rias. 2006. An evaluation of native-speaker judgements of foreign accented British and American English. LOT.


Measures: detail

Representativeness

\[
\bar{d}_f^l = \frac{2}{|l|^2 - |l|} \sum_{s,s' \in l} d_f(s,s')
\]

Distinctiveness

\[
d_f^{\sim l} = \frac{1}{|l|(|G| - |l|)} \sum_{s \in l, s' \notin l} d_f(s,s')
\]

Score

\[
\frac{\bar{d}_f^{\sim l} - \bar{d}_f}{sd(d_f)} - \frac{\bar{d}_f^l - \bar{d}_f}{sd(d_f)}
\]

I = native language under consideration, consisting of |l| speaker samples
G = larger group of languages, |G| speaker samples
s = speaker
f = feature
d = measure of between-speaker difference with respect to a feature f