VOT as a parameter in identification of voice disguise

Sara Neuhauser

University of Jena

IAFPA 2009

(3.–5.8.2009, Cambridge)
Introduction

- differences between fortis and lenis stops are realised through various articulatory and acoustic patterns
- e.g. closure duration, force of plosive release, VOT, duration of the preceding vowel
- language specific differences
- various contextual variables within a particular language
Voice Onset Time (VOT)

- since the 1960s measurement of VOT has been adopted for categorisation of plosives
  - Lisker & Abramson (1964) most cited and discussed study
- Braun (1983) showed that the timing relation between onset of voicing and the release of occlusion as a phonetic correlate of the phonological fortis-lenis-distinction was already discussed in the late 19th century
Voice Onset Time (VOT): Timing relation between the release of occlusion and onset of glottal vibration (i.e. voicing).

- onset of voicing before release (negative VOT) → voicing lead
- onset of voicing after release (positive VOT) → voicing lag

- e.g. Lisker & Abramson (1964), Keating (1984) showed language specific character of this parameter
- e.g. English, French, German contrast two categories of plosives
German /p, t, k/ and /b, d, g/ 

- initial /p, t, k/ in stressed syllables are articulated with strong aspiration
- no or less aspiration for /p, t, k/ after /ʃ/ or /s/, or before a syllabic sonorant
- /b, d, g/ usually realised voiceless (short lag VOT), may be voiced in intervocalic position
French /p, t, k/ and /b, d, g/ 

- /p, t, k/ are described as normally unaspirated, i.e. vocal cord vibration starting immediately after the plosive release (short lag)
- in regional variants, emphasis and before high vowels aspiration or friction is possible
- /b, d, g/ usually fully voiced (voicing lead)
Introduction & Motivation

- contrast between German and French plosives (Künzel, 1977)
  - duration of aspiration (VOT) for fortis plosives
  - presence / absence of voicing for lenis plosives

- a native French speaker might be expected to produce German /p, t, k/ with a shorter VOT associated with the realisation of the corresponding fortis plosives in French

- native German speakers might aspirate French fortis plosives
VOT in the forensic context

- VOT might be a crucial parameter in speaker identification
  - speaker specific (Künzel, 1987)
  - dependent of regional origin within a particular language (Braun, 1996)
  - possible characteristics for bilingual speakers (Gurski, 2006)
How about voice disguise imitating foreign accents?
Hypothesis

Native German speakers reduce VOT in voiceless plosives during imitation of a French accent.
8 native German speakers imitating a French accent reduced VOT in initial fortis plosives

some of them were consistently able to deaspirate voiceless plosives (within the text and also in the 2nd recording session)

also produced other features consistent with the phonetic correlates of German lenis plosives, i.e. shorter closure, weaker release
Aims of current study

- further testing of the hypothesis (statistical significance?)
- comparison of the French accent imitations to German productions of native French speakers
- comparison of native French speakers’ productions in French to their productions in German
Method

- VOT measurements of 25 initial fortis plosives in German spoken by
  - 22 native German speakers (15♀, 7♂) in 2 modes, i.e. undisguised and imitated French accent
  - 4 native French speakers (2♀, 2♂ speaking German)
  - /p, t, k/ in stressed syllables, except Poli’zei (police)

- VOT measurements of 5 initial fortis plosives in French spoken by 4 native French speakers
  passez, toujours, tout (droit), coute, commissariat
VOT as a parameter in identification of voice disguise

Experiment

<table>
<thead>
<tr>
<th>C</th>
<th>VOT</th>
<th>PR</th>
<th>OV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0265</td>
<td>26 msec</td>
<td>0.1867</td>
<td></td>
</tr>
</tbody>
</table>

Time (s)
Results

- Native German speakers reduced VOT in initial fortis plosives significantly during imitation of a French accent.
- Reduction is highly significant: $p < 0.001$ (Wilcoxon Signed Rank Test for matched-pairs).
- Observed mean VOT of initial fortis plosives in:
  - Undisguised German: 54 msec (sd=9)
  - French accent imitation: 37 msec (sd=12)
Summary of mean VOT in undisguised and disguised German

VOT in msec

<table>
<thead>
<tr>
<th>undisguised</th>
<th>disguised (French accent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>30</td>
</tr>
<tr>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>60</td>
<td>70</td>
</tr>
</tbody>
</table>

VOT as a parameter in identification of voice disguise

Experiment
closer inspection shows high interspeaker variability

mean reduction of VOT ≈ 18 msec

varies between 39 msec reduction – 4.5 msec lengthening (both outliers)
VOT as a parameter in identification of voice disguise

Experiment

Variation of mean VOT

Speakers

VOT in msec

un disguise
disguise

| Speakers | f01 | f02 | f03 | f04 | f05 | f06 | f07 | f08 | f09 | f10 | f11 | f12 | f13 | f14 | f15 | m01 | m02 | m03 | m04 | m05 | m06 | m07 |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| VOT in msec |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
Are the VOT values of the imitated French accents comparable to the VOT values of authentic French accent in German?
VOT as a parameter in identification of voice disguise

Experiment

<table>
<thead>
<tr>
<th></th>
<th>authentic accent</th>
<th>undisguised German</th>
</tr>
</thead>
<tbody>
<tr>
<td>imitated accent</td>
<td>20, 30, 40, 50, 60, 70</td>
<td>20, 30, 40, 50, 60, 70</td>
</tr>
<tr>
<td>VOT in msec</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Box plot showing the distribution of VOT in msec for imitated accent, authentic accent, and undisguised German accents.
lower VOT values for imitated French accents than for authentic French accents in German

- exaggeration by the native Germans during their accent imitation?
- good competence of the native French speakers in their German productions?
Comparison of French speakers' productions

- minor changes in mean VOT for the 4 native French speakers speaking French or German
- increased standard deviation, slightly higher maximum values in German

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>min</th>
<th>max</th>
<th>N (token)</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>43</td>
<td>4</td>
<td>38</td>
<td>47</td>
<td>5</td>
</tr>
<tr>
<td>German</td>
<td>45</td>
<td>10</td>
<td>36</td>
<td>59</td>
<td>25</td>
</tr>
</tbody>
</table>

- Fr: low interspeaker variability, high intraspeaker variability (low values for /p/, high values for /t, k/)
- Ger: higher interspeaker variability, high intraspeaker variability (place of articulation)
most of the native German speakers reduced VOT significantly during imitation of a French accent

suggest that native German speakers were aware of reduced VOT as a possible interference characteristic of native French speakers

indicate that native German speakers were able to conform to the French pattern of unaspirated voiceless fortis plosives
reduction of VOT leads to lower VOT values for the accent imitations than for authentic accents

native French speakers seem not to vary the VOT in their German productions significantly

suggests an “exaggeration of accent imitation” in this parameter

problematic for forensic context: high intra-speaker variability for some speakers, especially for German speakers imitating French accent but also for native French speakers in German and French
Outlook

What we have not done yet:

- consistency of VOT reduction: compare productions in 1\textsuperscript{st} and 2\textsuperscript{nd} recording session
- voicing in lenis plosives
- other phonetic correlates of fortis-lenis-distinction, e.g. contextual f0, i.e. are some German speakers just employing correlates of German lenis for French fortis?
References


References


