FAIC – Foreign Accent Imitation Corpus

Sara Neuhauser

University of Jena, Germany

IAFPA 2011
Vienna, 24.–28.07.2011
1 Background
   - Preliminary study

2 Corpus
   - Speakers and accents
   - Texts and recordings
   - Data processing

3 Summary

4 References
Background

- imitation of foreign accents as a possible form of voice disguise
- alleged aims: conceal identity and mislead investigation
Experimental studies have focused on

1. authenticity judgements by listeners
   (e.g. Tate, 1979; Markham, 1999; Neuhauser & Simpson, 2007)

2. production of imitated foreign accents / different regional dialects
   (e.g. Torstensson et al. 2004; Neuhauser, 2008)
FAIC – Foreign Accent Imitation Corpus

Background

_corpus collected within my PhD project (Neuhauser, 2011)_
- investigating foreign accent imitation from a production and a perception perspective within a German context
  - accent authenticity, accent identification
  - VOT and voicing in plosives
  - syllable initial glottalisation and glottal friction
  - realisation of final /-ən/

_published via an internet platform_
- examples of the recordings, all texts, description of data acquisition, recording conditions and data processing
The background of the study was deliberately withheld to exclude stereotypes. The main investigation will involve the acquisition of German subjects (potential imitators) for accent imitations.
2 groups of native German speakers were asked to name the two accents they think would be the easiest to imitate

1. 121 subj. (60 male, 61 female; mean age = 22.7 years)
2. 70 subj. (27 male, 43 female; mean age = 21.9 years)

1st group: a list of accents were given (AE, BE, French, Ital., Russ., Swiss Germ., Slav., Span., Turk., Vietn.)

- choose up to 2 accents from this list or name another accent

Control group: no accents were given
FAIC – Foreign Accent Imitation Corpus

Background

Preliminary study

![Bar chart showing number of votes for different accents](chart.jpg)
35 native German speakers imitating foreign accents

<table>
<thead>
<tr>
<th>Accent imitations</th>
<th>No of speakers</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>French accent</td>
<td>22</td>
<td>(15 female, 7 male)</td>
</tr>
<tr>
<td>Am.-English accent</td>
<td>11</td>
<td>(5 female, 6 male)</td>
</tr>
<tr>
<td>Russian accent</td>
<td>2</td>
<td>(1 female, 1 male)</td>
</tr>
<tr>
<td>Turkish accent</td>
<td>1</td>
<td>(female)</td>
</tr>
<tr>
<td>Total</td>
<td>35(^1)</td>
<td>(22 female, 13 male)</td>
</tr>
</tbody>
</table>

- mean age = 23.8 years
- mostly from eastern central Germany (Thuringia, Saxonia)

\(^{1}\) 1 male speaker produced a French and an Am.-Engl. accent
12 non-native German speakers speaking German

<table>
<thead>
<tr>
<th>Mother tongue</th>
<th>No of speakers</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>4</td>
<td>(2 female, 2 male)</td>
</tr>
<tr>
<td>Am.-English</td>
<td>8</td>
<td>(4 female, 4 male)</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>(6 female, 6 male)</td>
</tr>
</tbody>
</table>

- mean age = 23 years
- subjects differ in their L2 competence level
- noticeable foreign accent in German

\(^2\)recordings of 1 Russian, 1 Czech, 1 Italian and 1 Am.-Engl. speaker were excluded
Texts

- 2 prepared texts (about 220-240 words) ca. 1 week before the recordings
- composition of an own text with given words and phrases
- immediately before the recordings: 1 unprepared text
- spontaneous speech

- native Germans: disguised and undisguised productions (12 texts/speaker)
- non-native Germans: German productions and 1 text in their mother tongue (10 texts/speaker)
each subj. attended 2 recording sessions with a time-lag of 2 weeks

- recording conditions and order of reading the texts were the same in both recording sessions

- speakers were informed about the forensic context immediately before the 1st recording session

- instruction to speak as naturally as possible rather than trying to produce standard German in the undisguised text versions
Recordings

- sound treated room directly to PC
- digitizing at a sampling rate of 16 kHz and an amplitude resolution of 16 bits
- speaking distance to the microphone = ca. 40 cm
Examples

- French accent:
  - female native German speaker reading text 1
  - female native French speaker reading text 1

- American-English accent:
  - male native Am.-English speaker – spontaneous speech
Data processing
French and Am.-Engl. accents

- transliteration of the self-composed texts and the spontaneous speech
- annotation at the word level (506 of 518 signals)
  - using praat and xassp (converted to praat .TextGrid)
- anonymisation of spontaneous speech
  - most significant personal details of each speaker (names, address, place of birth etc.)
  - using a script and the annotation
  - 400 Hz low pass filter, transition bandwidth 100 Hz – preserving pitch and temporal aspects
Example
Annotation & Anonymisation

- male native German speaker imitating Am.-Engl. accent (spontaneous speech)
Summary

FAIC offers annotated speech material of 47 native and non-native German speakers producing...

- imitated and authentic foreign accents
- read and spontaneous speech
- self-composed texts
FAIC is of forensic interest...

- detailed phonetic analyses of imitated and authentic accents produced in different linguistic activities (e.g. read vs. spontaneous)
- non-phonetic linguistic analyses of self-composed texts and spontaneous speech (e.g. syntactical or lexical aspects)
- analysis of speakers’ consistency during accent imitation (2 recording sessions)
Projects using FAIC

- Neuhauser (2011) etc.
  Production and perception of imitated foreign accents
  Epiphenomenal sound production in foreign accent disguise
- Witecy (2010)
  Foreign Accent Syndrome and the perception of accents
- Hellmann (2008)
  German as a foreign language for native American-English speakers
FAIC [fɛɪk] – Foreign Accent Imitation Corpus

http://www.uni-jena.de/~x0nesa/faic.html

or mail to

sara.neuhauser@uni-jena.de

Thanks to Adrian P. Simpson and to the student assistants
Susanne Friedrich, Judith Hellmann, Anne-Katrin Leich and Toni Linke


