Effect of Telephone-Line Transmission and Digital Audio Format on Formant Tracking Measurements

Christoph Meinerz
Landeskriminalamt Brandenburg, Germany
christoph.meinerz@gmx.de

Herbert Masthoff
Department of Phonetics, University of Trier, Germany
masthoff@uni-trier.de
Introduction - Formants, Speaker ID and Audio Compression

Method - Experimental Setup, Hardware, Software

Results - Formant Shift

Conclusion - What to do

Christoph Meinerz
Landeskriminalamt Brandenburg, Germany
christoph.meinerz@gmx.de

Herbert Masthoff
Department of Phonetics, University of Trier, Germany
masthoff@uni-trier.de
• (revival of) reports of formant measurements for speaker identification (i.e. Nolan/Grigoras, 2005; Becker et al., 2007; Jessen et al., 2010; Simpson/French, 2010)

• reports of effects of telephone and lossy compression on acoustic parameters (Künzel, 2001; Köster/Grasmück, 2004; Gonzalez et al., 2003)

• the problem is real: telephone-intercepts in low-Bit .mp3!

➡ results of preliminary study: effects of telephone-line and lossy low-Bit audio compression on LPC-based formant-measurement and no intra-speaker variation
Experimental set-up - „The Plan“
### Method II

<table>
<thead>
<tr>
<th>Format</th>
<th>Codec</th>
<th>Sampling Rate</th>
<th>Bit Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>mike.wav</td>
<td>PCM</td>
<td>44.1 kHz</td>
<td>705 kbps</td>
</tr>
<tr>
<td>mike.wma</td>
<td>CBR</td>
<td>22. kHz</td>
<td>20 kbps</td>
</tr>
<tr>
<td>mike.mp3</td>
<td>CBR</td>
<td>8 kHz</td>
<td>8 kbps</td>
</tr>
</tbody>
</table>

**Tech-Specs:**
- Sound Studio UoT
- Mike: Neumann M147 Tube
- Soundcard: RME Hammerfall

<table>
<thead>
<tr>
<th>Format</th>
<th>Codec</th>
<th>Sampling Rate</th>
<th>Bit Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>tel.wav</td>
<td>PCM</td>
<td>44.1 kHz</td>
<td>705 kbps</td>
</tr>
<tr>
<td>tel.wma</td>
<td>CBR</td>
<td>22. kHz</td>
<td>20 kbps</td>
</tr>
<tr>
<td>tel.mp3</td>
<td>CBR</td>
<td>8 kHz</td>
<td>8 kbps</td>
</tr>
</tbody>
</table>

**Tech-Specs:**
- „Re-Tel“ – Tel. Rec. Adapter 157
- Soundcard: MBox 2 Pro

### Audio Formats and Hardware
Shift of average formant frequency according to format (males)
Shift of average formant frequency according to format (females)
Mean shift of average formant frequency according to format % (all)
Sonographic symptoms (top mike .wav, bottom mike .mp3)
Sonographic symptoms (top tel .wav, bottom tel .mp3)
Results VI

Sonographic symptoms (top mike .wav, bottom mike .mp3)
Sonographic symptoms (top tel .wav, bottom tel .mp3)
Summary

- shift of formant frequencies (all)
  - F3: downward $\approx 2 - 23\%$
  - F2: downward $\approx 1 - 17\%$
  - F1: mike downward $\approx 1 - 16\%$
    - tel upward $\approx 2 - 4\%$, .wav + .wma
    - tel downward $\approx 1\%$, .mp3

- highest amount of shift in tel .mp3, 8 kbps

- telephone-line alone produces shift of F2, F3 $\approx$ mike .mp3

- sonagraphic and auditory symptoms
  - spectral cancellations - „the moth“
  - „musical noise“ effect
• results confirm those already reported (i.e. Becker et al., 2011!)

• consider shifting effects when doing formants and formant-related ASR (LPC)

• include larger population for statistical significance - possibly detect “critical” Bit-rate

• possibly cross-check with FFT-based measurements
Thank you for your attention!
References

Harrison, P.: Formant measurement errors for multiple synthetic speakers. IAFPA Annual Conference 2010, Trier.