Transforming F0 Contours

Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Eurospeech 2003 - Interspeech 2003

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Transforming F0 Contours

Ben Gillett, now with Camel Audio (www.camelaudio.com)
Simon King, Centre for Speech Technology Research, University of Edinburgh, UK

ben@camelaudio.com, Simon.King@ed.ac.uk

**Goal**

To transform the F0 contour of some speech from a source speaker, such that listeners believe it to have been uttered by some target speaker.

**Applications**

- voice transformation
  - also need a method for transforming voice quality - see poster by King & Gillett in session PWeBe
- speech synthesis
  - as a way of adapting existing intonation models (trained on one speaker) to a new speaker, without having to annotate much more data

**Evaluation**

25 subjects presented with speech from target speaker and speech with transformed F0 contours ("imitator speech") in XABX format. Asked to judge which of A or B was spoken by the target speaker.

**Results**

<table>
<thead>
<tr>
<th>Preference</th>
<th>Mean (%)</th>
<th>Std. Dev.</th>
<th>α</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>preference for ( M_{PL} ) over ( M_{PL} ) for ( S_{\text{same}} )</td>
<td>67</td>
<td>10</td>
<td>(&lt;1 \times 10^{-6})</td>
<td>-8.71</td>
</tr>
<tr>
<td>preference for ( M_{PL} ) over ( M_{PL} ) for ( S_{\text{same}} )</td>
<td>54</td>
<td>8</td>
<td>(~0.02)</td>
<td>-2.49</td>
</tr>
<tr>
<td>preference for target over mapped contours</td>
<td>73</td>
<td>9</td>
<td>(&lt;1 \times 10^{-6})</td>
<td>-13.8</td>
</tr>
</tbody>
</table>

**What next?**

- Apply the method to full voice transformation or speech synthesis

**See also...**

- at this conference:
  - poster by Gillett & King in session PWeBe (voice quality transformation)
- www.cstr.ed.ac.uk for latest progress on voice transformation and speech synthesis
- www.camelaudio.com for musical instrument transformation and morphing