An investigation of the parameters derived from inverse filtering of microphone and flow signals

Rosemary Orr, Bert Cranen, Felix de Jong
Are the parameters extracted from the flow signal comparable to those extracted from the microphone signal?
Subjects

• N = 61
• 16 males, age 18 to 44
• 45 females, age 17 to 41
• Logopedic and phoniatric examination established normal voice and healthy vocal folds
Phonation task

- /paepaepaepae/
  - Normal
  - High (frequency)
  - Low (frequency)
  - Loud (intensity)
  - Soft (intensity)
## Data

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mic.</th>
<th>Flow</th>
<th>EGG</th>
<th>Nasendo scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>●</td>
<td></td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>●</td>
<td></td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>3</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>
Parameters

• H1minH2
  – The difference in amplitude between the first and second harmonics
• Spectral Slope, dB/oct
• Open Quotient, OQ
• Speed Quotient, SQ
## Initial Results

<table>
<thead>
<tr>
<th>parameter</th>
<th>p (t-test)</th>
<th>R (pearson X moment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1minH2</td>
<td>0.01</td>
<td>0.56</td>
</tr>
<tr>
<td>spectral slope</td>
<td>0.00001</td>
<td>0.52</td>
</tr>
<tr>
<td>SQ</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>OQ</td>
<td>0.44</td>
<td>0.54</td>
</tr>
</tbody>
</table>
Initial Explanation

• Data processing
• Subjectivity of the experimenter
• Normal within-subject variation
• Psychological effect of the mask
• Acoustic distortion caused by the mask
• Naïve statistical approach
Data Processing

• Calibration
  – Check of manual marking
  – Check that the correct subject set was used

• Inverse filter settings
  – Second inspection and processing of inverse filtered waveforms

• Calculation of spectral parameters
  – Exclusion of frequency components above 1kHz
Subjectivity of the Experimenter

• Re-processing of ca. 25% of the usable mask signals by a second experimenter for comparison of inverse filtering results – there was no considerable difference
Normal within-subject variation

- Normal within-speaker variation can be large
- The dataset included only one full recording of each voicing condition
- Multiple recordings would not have been practical or affordable
Mask Effects

- Psychological effect of mask may have introduced different voice settings for voicing with and without a mask
- Different size of facial features
- Exclusion of high frequency components
- Effective lengthening of vocal tract
  - Lowering of formants
Naïve Statistical Approach

• Expectation of very clear correlation between mask and flow values
• The effect was not seen – nor was the cause of the obfuscation
• Repeated measures ANOVA shows main effects for three of the four parameters – SQ, Spectral Slope, and H1minH2
• Box plots illustrate a quite systematic effect
### ANOVA results – male group

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$F$ (df=1)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1minH2</td>
<td>9.15</td>
<td>0.003</td>
</tr>
<tr>
<td>spectral slope</td>
<td>7.95</td>
<td>0.006</td>
</tr>
<tr>
<td>SQ</td>
<td>12.13</td>
<td>0.0006</td>
</tr>
<tr>
<td>OQ</td>
<td>0.31</td>
<td>0.58</td>
</tr>
</tbody>
</table>
# ANOVA results – female group

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$F$ (df=1)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1minH2</td>
<td>43.66</td>
<td>$1.34 \times 10^{-10}$</td>
</tr>
<tr>
<td>spectral slope</td>
<td>0.03</td>
<td>0.86</td>
</tr>
<tr>
<td>SQ</td>
<td>0.172</td>
<td>0.68</td>
</tr>
<tr>
<td>OQ</td>
<td>1.113</td>
<td>0.29</td>
</tr>
</tbody>
</table>
Spectral Slope

Geneva 28th August, 2003
R. Orr, B. Cranen, F. de Jong

mask.female no mask.female mask.male no mask.male

−30 −20 −10 0 dB/oct

NS
Speed Quotient

The box plot shows the distribution of Speed Quotient (sq) for different conditions:
- **mask.female**
- **no mask.female**
- **mask.male**
- **no mask.male**

The plot indicates that there is no significant difference (NS) between the conditions.

Well?

• Data processing seems robust
• Subjectivity of the experimenter seems not to affect the results
• Normal within-subject variation remains a problem
• With this dataset, the mask may not be the most appropriate measurement instrument
So?

• Different experimental design
  – Experienced experimental voice
  – Controlled voice production
  – Repeated recordings
  – Different voice qualities with known relative parameter values
Definition of waveform parameters

\[ U_g \]

\[ \frac{dU_g}{dt} \]

\[ C_i \quad O_i \quad P_i \quad C_{i+1} \]

\[ AC \quad DC \]

MFDR

Geneva 28th August, 2003
R. Orr, B. Cranen, F. de Jong