Perception of overall voice quality: combining continuous speech and sustained vowels

Youri Maryn
Sint-Jan General Hospital, Bruges, Belgium
Department of Speech-Language Pathology and Audiology
Measurement of overall voice quality

- Perceptual
- Acoustic
## Introduction

### Speaking task

<table>
<thead>
<tr>
<th></th>
<th>Sustained vowel (SV)</th>
<th>Continuous speech (CS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glottal and supraglottal time-invariance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitch detection and extraction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized production</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-voiced phonemes, voice onsets/terminations, prosodic changes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech rate, pauses, phonetic context</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dialect, region, language, cognition, ...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

→ Preference for SV
### Speaking task

<table>
<thead>
<tr>
<th>Significant difference between SV and CS</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Krom (1994)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Revis et al. (1999)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolfe et al. (1995)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zraick et al. (2005)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Examples**

<table>
<thead>
<tr>
<th>CS+SV</th>
<th>CS</th>
<th>SV</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Sound Icon]</td>
<td>![Sound Icon]</td>
<td>![Sound Icon]</td>
</tr>
<tr>
<td>![Sound Icon]</td>
<td>![Sound Icon]</td>
<td>![Sound Icon]</td>
</tr>
<tr>
<td>![Sound Icon]</td>
<td>![Sound Icon]</td>
<td>![Sound Icon]</td>
</tr>
</tbody>
</table>
**Speaking task**

Maryn et al. (in press): «... it seems essential for perceptual and instrumental analyses to be based upon both sample types if it is to be considered ecologically valid ...»
Study question

Q1 – Is there a statistically significant difference between SV and CS?

Q2 – What determines the final rating of a global auditory-perceptual evaluation based on both SV & CS?

- SV
- CS
- Worst of SV and CS
- Best of SV and CS
- Mean of SV and CS
**Methods**

**Subjects**

- Gender: 25 f – 14 m
- Age: 47 y (16-86 y)
- Voice clinic: 39 patients with various voice disorders
Methods

Auditory-perceptual ratings of G (0-3)

- 5 experienced raters
- At random ratings of:
  - CS+SV
  - CS
  - SV
Methods

Statistics

- Statistical difference: Wilcoxon signed rank test
- Correlation coefficient: $r_s$
- Coefficient of determination: $r_s^2$
Results

Q1 – CS versus SV

- Wilcoxon: \( p = 0.002 \)
- \( r_s = 0.769 \)
- \( r_s^2 = 0.591 \)
### Q2 - Determination of CS+SV

<table>
<thead>
<tr>
<th></th>
<th>$r_s$</th>
<th>$r_s^2$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>0.831</td>
<td>0.691</td>
<td>0.017</td>
</tr>
<tr>
<td>CS</td>
<td>0.929</td>
<td>0.863</td>
<td>0.030</td>
</tr>
<tr>
<td>Worst</td>
<td>0.881</td>
<td>0.776</td>
<td>0.000</td>
</tr>
<tr>
<td>Best</td>
<td>0.920</td>
<td>0.874</td>
<td>0.001</td>
</tr>
<tr>
<td>Mean</td>
<td>0.934</td>
<td>0.872</td>
<td>0.131</td>
</tr>
</tbody>
</table>
Results

CS+SV - Mean

![Graph showing the relationship between G and CS+SV with mean values.]
Q1

- CS and SV: significantly different

- Wolfe et al. (1995) and Zraick et al. (2005)

- Important to include both CS and SV in perceptual rating of dysphonia severity (Maryn et al., in press)

- Further investigation → Q2
Discussion

Q2

- CS+SV is predominantly determined by averaging CS and SV
  - G of worst stimulus is decreased
  - G of best stimulus is increased

- CS+SV is most influenced by CS
  - Not only SV in auditory-perceptual ratings of G


