Introduction

• Studies in the area of voice assessment still lack objectivity in the description and evaluation of vocal quality.

• In therapeutic practice in Portugal, the assessment of vocal pathologies is not uniform, because each Speech and Language Therapist (SLT) uses individual and diverse assessment procedures to evaluate.
Introduction

• The aim of this project was to develop the first standardised protocol for voice assessment in EP: Universidade de Aveiro’s Voice Evaluation Protocol.
• Creating a working tool for SLTs.
• A pilot study to test the reliability of the protocol has been reported elsewhere (Jesus et al. 2009, Interspeech 2009, Brighton).

Method

• 40 patients (9 males and 31 females) admitted to the Department of Otolaryngology of the Hospital de São João, Porto, Portugal.
• The sample included several clinical diagnoses: nodules, polyps, hypotonia of the vocal folds, Reinke’s oedema, musculo-skeletal syndrome and dysfunctional dysphonia.
• The diagnoses were made by an experienced SLT and an Otorinolaringology consultant.
Method

- Recordings were made using a desktop computer, a microphone connected to a soundcard, set to 16 bits and 22050 Hz or 44100 Hz sampling frequency.
- During the recordings, the microphone was held on a tripod placed 25-30 degrees to the left of the patient's mouth, at a distance of 30-40 cm.
- Acoustic parameters extracted using Praat 5.1.10:
  - F0 Hz (median), F0 Hz (mean) and F0 (std deviation);
  - Jitter% (ppq5 – five-point frequency perturbation quotient equivalent to MDVP’s PPO);
  - Shimmer% (apq11 – eleven-point amplitude perturbation quotient equivalent to MDVP’s APO);
  - HNRdB (mean Harmonics-to-Noise Ratio).

Method

- A single sustained sample selected from each speaker’s productions of vowel /a/ was used to extract the acoustic parameters.
  - 100 consecutive cycles taken 200 ms after phonation onset were used for analysis.
- Analysis of the correlation between corresponding parameters of the CAPE-V and GRBAS scales was performed.
- The perceptual parameters grade (global in CAPE-V), roughness and breathiness were also compared individually with the objective acoustic parameters.
Method

- Acoustic parameter data and scale evaluation scores were compared to find statistically significant differences between males and females using the Mann-Whitney U Test.
- Correlation analysis (Spearman correlation test) between the perceptual parameters in CAPE-V and GRBAS scales were also evaluated with the acoustic parameters.

Method

- Finally, analysis of the correlation (Spearman correlation test) between corresponding parameters of the CAPE-V and GRBAS scales was performed.
- All statistical analyses were conducted using SPSS 13.0 and a p value of less than 0.05 was considered significant.
Results

- Statistical significances were found between the perceptual subscale grade from GRBAS and subscales global and roughness from CAPE-V, roughness in GRBAS and global in CAPE-V, and breathiness in GRBAS and in CAPE-V.
- The correlation values are good, ranging from 0.60 to 0.87, with the exception of the correlation value between the subscale roughness in GRBAS and the subscale global in CAPE-V for the total value.
- The results found for males alone can be ascribed to the smaller sample size (N=9).

<table>
<thead>
<tr>
<th></th>
<th>CAPE-V</th>
<th>GRBAS</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grade</td>
<td>Roughness</td>
<td>Breathiness</td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.00</td>
<td>0.86*</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.65*</td>
<td>0.20</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.60*</td>
<td>0.42*</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Roughness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.23</td>
<td>-0.14</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.79*</td>
<td>0.14</td>
<td>-0.19</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.75*</td>
<td>0.26</td>
<td>-0.20</td>
<td></td>
</tr>
<tr>
<td>Breathiness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-0.18</td>
<td>0.67*</td>
<td>0.63*</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.09</td>
<td>-0.10</td>
<td>0.87*</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.04</td>
<td>0.08</td>
<td>0.80*</td>
<td></td>
</tr>
</tbody>
</table>
Results

- **Statistically significance** differences between males and females (as expected) were found in:
  - age;
  - median F0;
  - mean F0;
  - jitter ppq5;
  - mean HNR;
  - global and roughness parameters in the CAPE-V scale;
  - grade and roughness parameters in the GRBAS scale.

Results

- **Statistical significances** were found for median F0 and mean F0 with the perceptual subscales global and roughness for CAPE-V and grade and roughness for GRBAS.
- However, the correlations were weak, with values ranging from -0.38 to -0.60.
- No significant differences were found when we considered either only the male sample or the other acoustic parameters.
Conclusions

- The two scales (GRBAS and CAPE-V) have been previously used simultaneously (Karnell et al. 2007), with results showing a strong correlation between the two rating systems.
- Our results have also shown a good correlation except for roughness, because the term used in EP and Brazilian Portuguese for Grade is “grau de rouquidão”, which, appears to have been erroneously related to the CAPE-V EP term “rouquidão” (roughness).

Future Work

- This issue will be addressed in the future with a further validation of the EP version of CAPE-V that will use the procedures presented by Zraick (2007).
- This will include the production of a CD-ROM with voice samples to be evaluated, voices used for training and samples of voices that represent specified grades of severity.
Future Work

• The lack of success in finding the hypothesised correlations between acoustic and perceptual measures have long been known (Kreiman and Gerratt 2000, pp. 75-80), and do not seem to be related to language specific characteristics, as our results have shown, even when we limit our set of acoustic and perceptual parameters as in Dejonckere (2007).
• Different factors contribute to a failure to find consistent correlations: deficiencies in the theoretical framework; incoherencies in the definitions of parameters; limitations in estimation techniques (Kreiman and Gerratt 2000, pp. 75-80).

Acknowledgements

This work was supported by Fundação para a Ciência e a Tecnologia, Portugal (PTDC/SAU-BEB/67384/2006)