Universidade do Porto

FEUP Faculdade de Engenharia

Acoustic and Audio Perceptual Evaluation in the Characterization of the Human Voice



Susana Vaz Freitas

19th of January of 2013

Organization of the presentation

- Evaluation of the Human Voice
- Starting questions
- Objectives of the PhD Research
- Chronogram of PhD Work
- Main Studies/Research
- Main Results
- Voice-PE (Voice – Perceptual Evaluation)
- Papers and Presentations



The Evaluation of the Human Voice



- Voice is produced by the vocal tract, starts at the larynx, and results from the movement of air passing thru the vocal folds and the movement of the supra-laryngeal articulatory structures (Akpinar et al., 2011)

- The definition of vocal quality is related to physiological, perceptual and acoustic findings (Behlau, 2001; Guimarães, 2007; Camargo e Madureira, 2010; Akpinar et al., 2011)

- The evaluation of the human voice depends on a complex analysis of factors which involves a deep and multidisciplinary knowledge of the laryngeal functions and vocal quality (Behrman, 2005; Wertzner et al., 2005; Guimarães, 2007; Behlau, 2010)

Starting Questions



- Which scales exist?
- Which ones better describe voice and disturbed voice signals? And why?

Audio Perceptual Evaluation

- What measures are usually used?
- Which ones better describe voice and disturbed voice signals? And why?

Acoustic Evaluation

- What is the relation between results of voice evaluation?
- Are results independent or does meaningful associations exist?
- Any domain predicts others`?

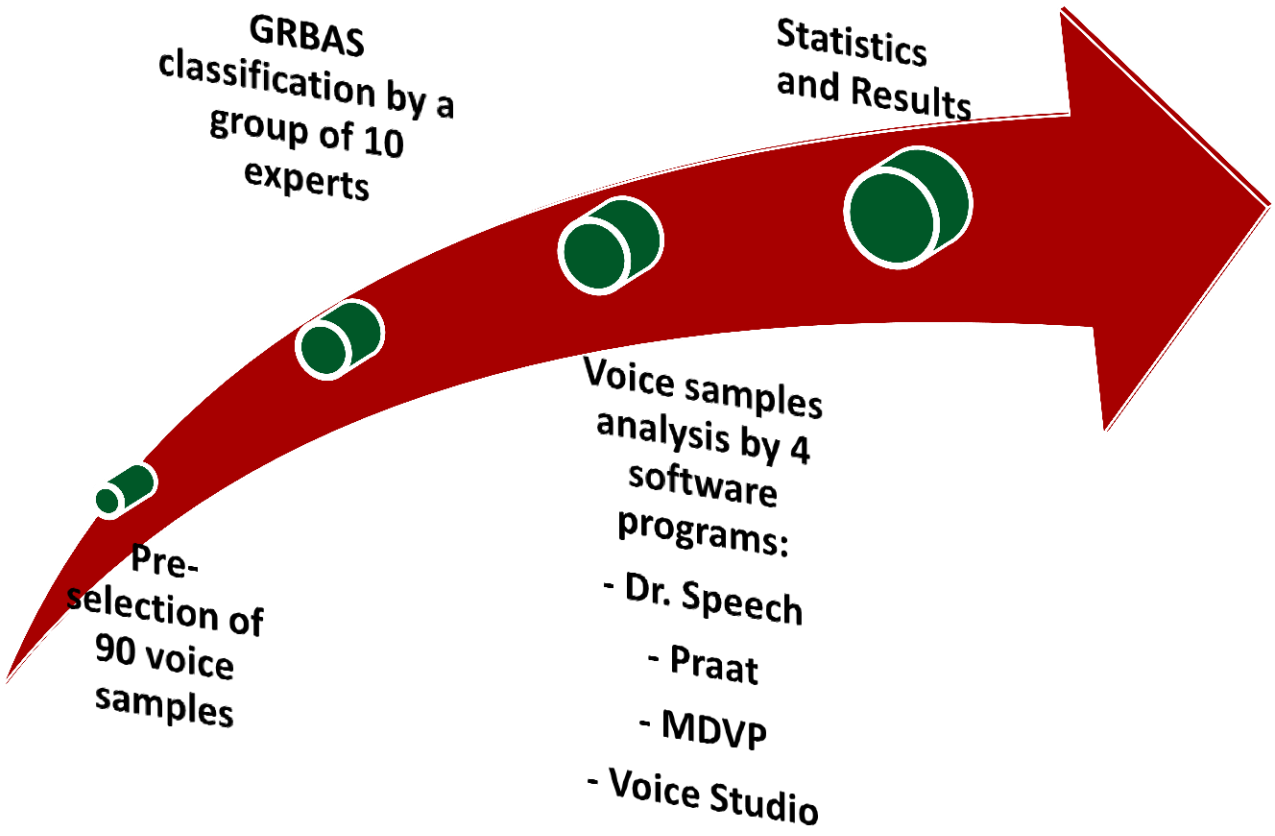
Correspondence between the results

Objectives of the PhD



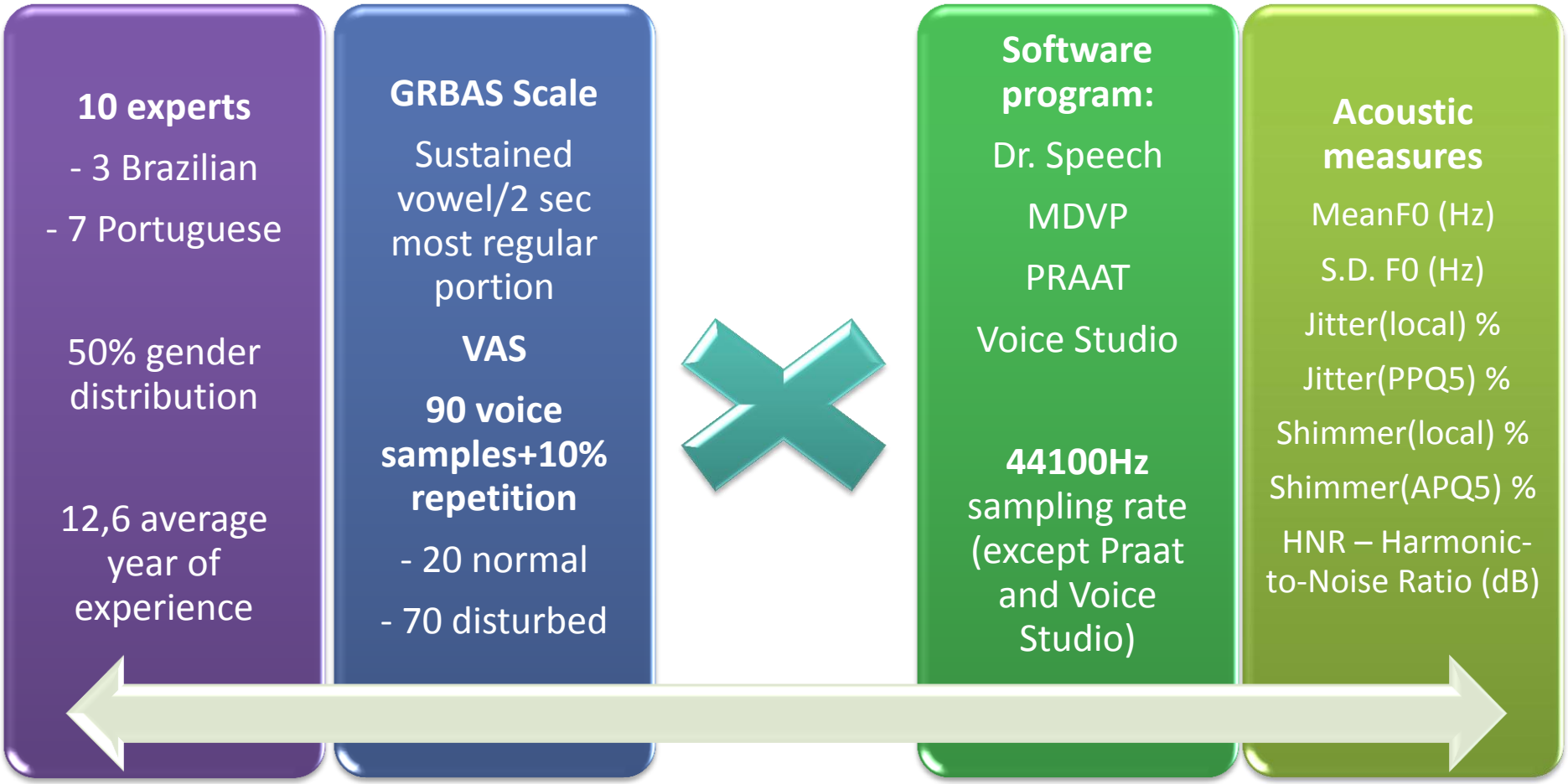
- Describe and analyze two of the most used methods for the characterization of the human voice: **audio perceptual scales** and the **acoustic measures**
- Develop a database of **voice samples of Portuguese speakers**, based on the first 10 years of professional clinical experience of the PhD Student
- Collect, describe and analyze the results of the **audio perceptual evaluation** of the database samples made by a **group of experts**
- Collect, describe and analyze the results of the **acoustic analyzes** of the database samples made by different **software programs**
- Analyze the **intra and inter-rater consistency** between the experts evaluation and the differences among measures obtained by the acoustic programs

Chronogram of the PhD Work





Main Studies/research



Main Studies/research



Audio Perceptual Evaluation

-Description of each **audio perceptual parameter (GRBAS)** for each **expert**
-**Inter-observer consistency (ICCC)**
-**Intra-observer consistency**

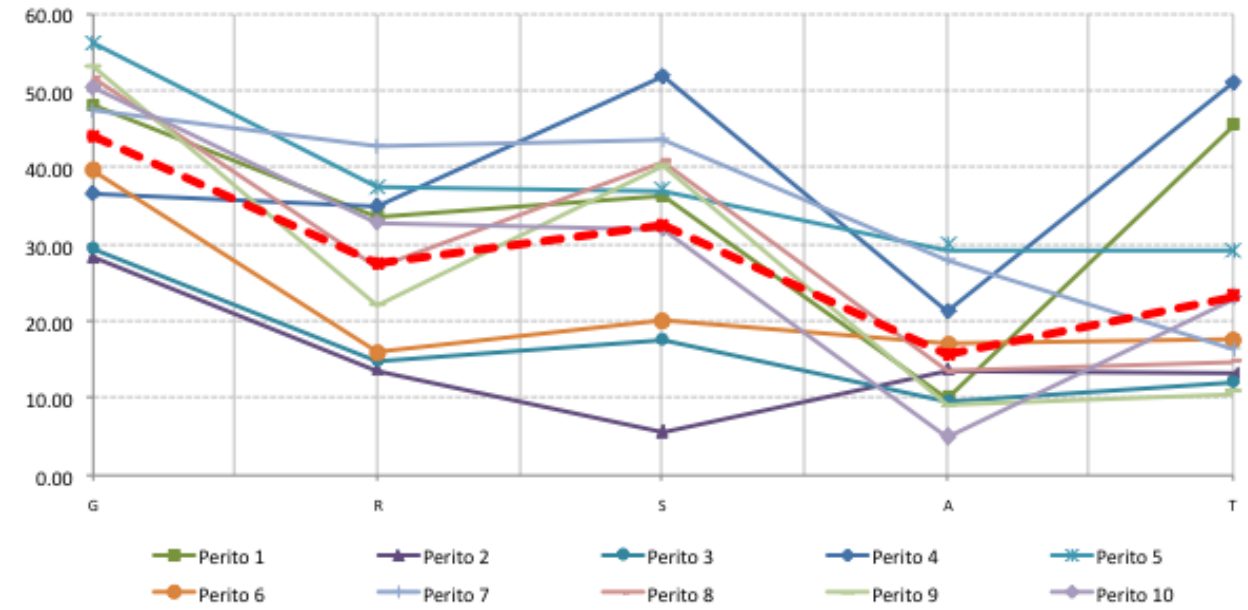
Linear Regression

- Multiple regression analysis to determine the association between the **acoustic measures** and the **audio perceptual parameters**, for each of the software programs

Acoustic Analysis

-Description of each **acoustic measure** obtained by each **software**
-**Correlation Coefficient (Spearman Ro)** between **software programs** for each acoustic measures (inter-rater consistency)

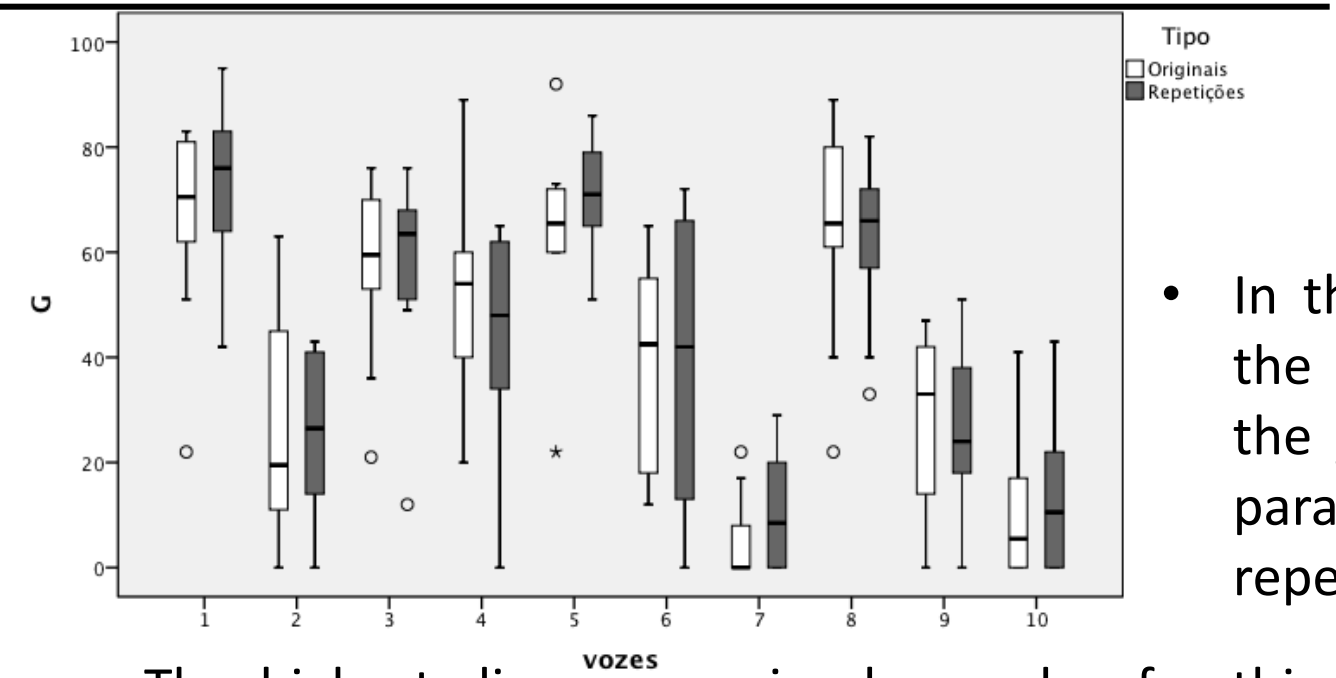
Main Results – AUDIO PERCEPTUAL EVALUATION



- The average classification given by the experts in the GRBAS scale is statistically different (p values < 0.05)

- But the parameters are consistent – the 95% confidence intervals of the mean correlations are between **0.838 e 0.966**. The experts use the same evaluation criteria, but quantify them differently
- G(Grade), R(Roughness) and B(Breathiness) parameters are the most consistent. The A(Asthenia) and S(Strain) parameters had an inferior consistency

Main Results – AUDIO PERCEPTUAL EVALUATION

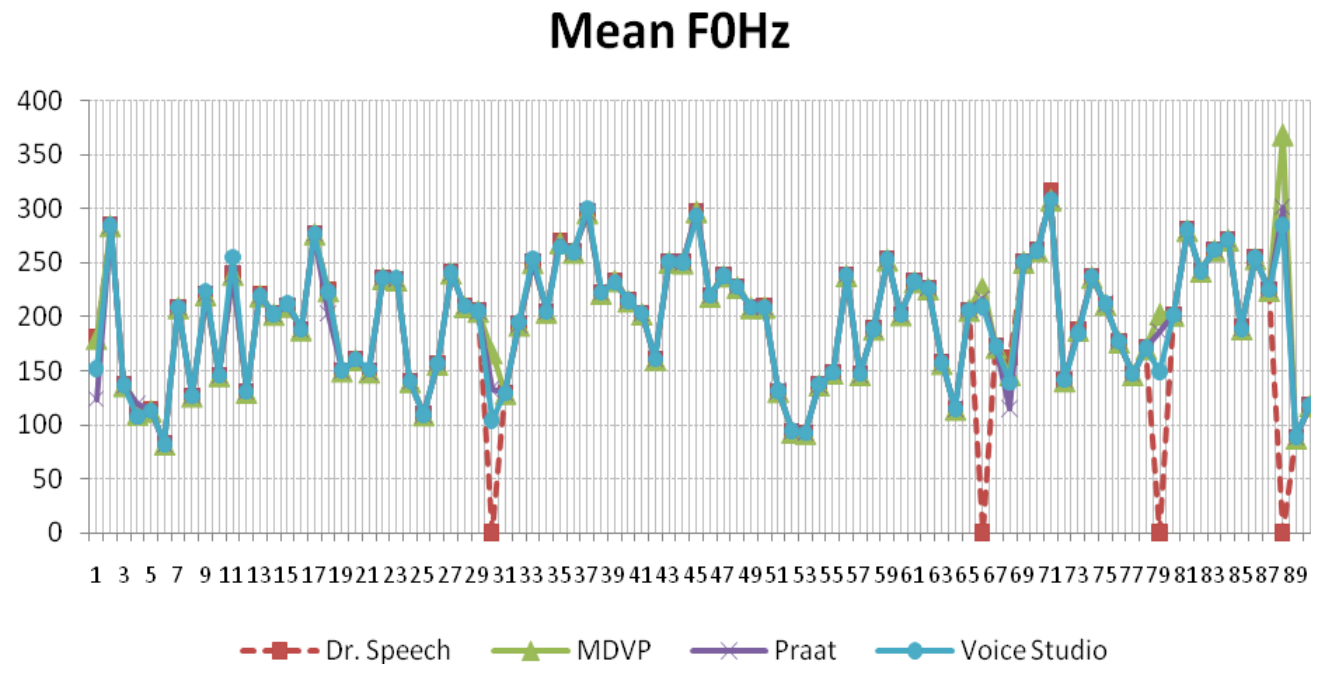


- In this study we observed the consistencies between the judges in each GRBAS parameter, for the 10% repeated voices

- The highest discrepancy is observed – for this grade parameter – in voices 2 and 6.
- But, in the end we concluded that voices 1, 5 and 6 were less consistent, so they were excluded
- The analysis of the intra-rater consistency demonstrated statistically significant differences ($p<0.05$) for the experts 1, 6, 9 and 10 ← excluded

Main Results – ACOUSTIC MEASURES





- Mean F0 was the only measure were differences were not significant (p>0,05)

- Voice Studio presented acoustic measures significantly higher then the other software programs, except for HNR (Dr. Speech)
- The strength of the correlation between the software programs is positive, moderate to strong, for all of the measures (with some exceptions), and statistically significant (p values < 0.01).

Main Results – MULTIPLE REGRESSION ANALYSIS



	G	R	B	A	
MeanF0Hz		-0.234***	0.372***		
SDF0Hz		0.298***			0.673***
HNRdB	-0.838***	-0.580***	-0.836***	-0.875***	
Jitterlocal				-0.514***	
JitterPPQ5					
Shimmerlocal					
ShimmerAPQ5					
Modelo	***	***	***	***	***
R^2_a	0.70	0.68	0.71	0.33	0.45

*** $p < 0.001$

- The acoustic measures predictive value for the audio perceptual parameters assumed a magnitude from weak (R^2_a=0,17) to moderate (R^2_a =0,71), with discrepancies from a software to another. Higher magnitudes obtained by Praat.

Voice-PE – *Voice – Perceptual Evaluation*

- Voice-PE (Voice – Perceptual Evaluation) is a training and testing software based in the average GRBAS classification of the six most consistent experts

- Voice is, essentially, an auditory phenomena so it depends on training (Bassich & Ludlow, 1986; De Bodt et al., 1997). The software was based upon the work of Oates & Russel (1997) and Awan & Lawson (2009)

- Voice-PE will be distributed by SLP and graduate students or other professionals who also need to develop skills in audio perceptual evaluation

- It's a way of recalibrating the judges internal audio perceptual standards (Gerratt et al., 1993; Chan & Yiu, 2002; Eadie & Baylor, 2006; Kreiman et al., 2007; Awan & Lawson, 2009; Brasolotto & Rehder, 2011; Gould et al., 2012; Iwarsson & Petersen, 2012; Sofranko & Prosek, 2012)



Papers and Oral Presentations



PEVOC 10 (Praha, August)

National Congress of the SPORL (Oporto, May)

IALP Congress (Turin, August)

2013



Papers

- Journal of Voice*
- Folia Phoniatrica et Logopaedica*
- Annals of Biomedical Engineering*

*in construction

Oral Presentations

- 4º Simpósio Luso-Brasileiro TF, 2009, São Paulo
- I Curso Teórico-Prático de Voz, 2010, Coimbra
- II Congresso Saúde, 2011, UA
- Pensar a Voz, 2011, ESTSP

Final Considerations



- The description and analysis of the human voice is not an easy or consensual task.

- The multi disciplines and dimensions to be considered in the capture of the voice samples and their posterior classification reflects its complexity.

- It's the clinician/voice scientist responsibility to learn and master the different instruments to collect, analyze and classify the voice signal → design estimations about the voice production and perception.