



HAL
open science

The phonetic approach of voice qualities: challenges in corresponding perceptual to acoustic descriptions

Zuleica Camargo, Sandra Madureira, Nathalia dos Reis, Albert Rilliard

► To cite this version:

Zuleica Camargo, Sandra Madureira, Nathalia dos Reis, Albert Rilliard. The phonetic approach of voice qualities: challenges in corresponding perceptual to acoustic descriptions. *Subsidia. Tools and resources for speech sciences*, 2019. hal-02428197

HAL Id: hal-02428197

<https://hal.science/hal-02428197>

Submitted on 5 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The phonetic approach of voice qualities: challenges in corresponding perceptual to acoustic descriptions

Zuleica Camargo¹, Sandra Madureira¹, Nathalia dos Reis¹ and Albert Rilliard^{2, 3}

¹ Pontifical Catholic University of São Paulo

² LIMSIS, CNRS, Université Paris-Saclay

³ Federal University of Rio de Janeiro

e-mail: zcamargo@pucsp.br, fononana2@gmail.com, albert.rilliard@limsi.fr

Citation / Cómo citar este artículo: Camargo, Z., Madureira, S., dos Reis, N., & Rilliard, A. (2019). The phonetic approach of voice qualities: challenges in corresponding perceptual to acoustic descriptions. In J. M. Lahoz-Bengoechea, & R. Pérez Ramón (Eds.), *Subsidia. Tools and resources for speech sciences* (pp. 11–17). Málaga: Universidad de Málaga.

ABSTRACT: This study introduces an innovative approach to the phonetic investigation of voice qualities, comprising the application of the Vocal Profile Analysis Scheme (VPAS) to describe perceived voice quality settings, the extraction of acoustic measures, and a statistical link between perception and acoustics, weighting the relative proximity of controlled factors. The corpus was perceptually annotated by the VPAS and data from 44 speakers were grouped in terms of the most frequent combinations in the VPAS system, generating two vocal profiles; the “Wide” and the “Short” vocal tract kinds of profiles. Acoustic measures (f_0 , intensity, signal to noise ratio, spectral slope and the first formant) were extracted. Statistical analysis weighs the relative links between the voice quality profiles and the acoustic measures, compared to linguistic and gender constraints. f_0 measures were found to be the most relevant to establish perceptual and acoustic correlations. Some singularities of the correspondences detected are discussed.

Keywords: voice quality; auditory perception; speech acoustics; phonetics; statistical analysis.

RESUMEN: Este estudio presenta un enfoque innovador para la investigación de las cualidades de la voz. El método incluye la aplicación del Esquema de Análisis del Perfil Vocal (VPAS) para describir los ajustes de la cualidad de voz que se perciben, junto con la extracción de medidas acústicas y un análisis estadístico sobre la relación entre la percepción y la acústica, que permite calibrar la proximidad relativa de los factores controlados. Se anotó perceptivamente un corpus siguiendo el modelo VPAS y se agruparon los datos de 44 hablantes en función de las combinaciones más frecuentes del sistema VPAS, lo que dio lugar a dos grandes tipos de perfil vocal: el “Ancho” y el “Corto”. Se extrajeron ciertas medidas acústicas (f_0 , intensidad, ratio entre la señal y el ruido, declinación espectral y primer formante). El análisis estadístico mide el peso de la relación entre los perfiles de la cualidad vocal y las medidas acústicas, comparado con restricciones lingüísticas y de sexo. La f_0 resultó ser la medida más relevante a la hora de establecer correlaciones entre lo perceptivo y lo acústico. El artículo comenta algunos detalles de las correspondencias detectadas.

Palabras clave: cualidad de voz; percepción auditiva; acústica del habla; fonética; análisis estadístico.

1. INTRODUCTION

Voice quality descriptions (and / or evaluation scales) tend to focus on acoustic-perceptual correlations (Dejonckere et al., 1995; Hammarberg & Gauffin, 1995; Kreiman & Gerratt, 2000; Kreiman & Sidtis, 2011; Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995). The relevant literature is also plenty of descriptions of perceptual labels and their acoustic and / or physiologic counterparts, especially for phonatory adjustments, i. e. the voice source events (d’Alessandro, 2006; d’Alessandro, Darsinos, & Yegnanarayana, 1998; Dejonckere et al., 1995; Garellek, 2014; Hammarberg & Gauffin, 1995;

Rabinov et al., 1995; Sundberg & Gauffin, 1979). Some of them are based on voice quality settings described in the Vocal Profile Analysis Scheme (VPAS) system (Laver, Wirz, Mackenzie, & Hiller, 1981).

The concept of voice quality used in VPAS derives from the model of phonetic description of voice quality by Laver (1980). As a phonetically grounded model, voice quality is here considered as the result of phonatory and articulatory settings, that is, the result of specific adjustments of the vocal folds and of the articulators during speech.

To perform the voice quality evaluations based on the VPAS, judges need a phonetic background and experience on the use of the profile. The basic analytical unit is the Voice Quality Setting (VQS), a long-term muscular tendency in the vocal apparatus: supralaryngeal (articulators and resonators), laryngeal / phonatory (vocal folds vibrations) and muscular tension activity (laryngeal and supralaryngeal). The VQS are described as variations from a reference setting, the neutral one, in which no effect is found in longitudinal or transversal plans of the vocal tract, and there is no variation in terms of its muscular tension activity. For the neutral setting, the vibration of vocal folds must be periodic.

The VPAS is applied in two passes. The first pass comprises the identification of non-neutral VQS. The second pass involves the grading of non-neutral VQS in a scalar degree, generally from 1 to 6.

It is important to reinforce that the phonetic description of voice quality model (Laver, 1980) follows two principles: susceptibility and compatibility.

The susceptibility principle accounts for the fact that some speech segments are more susceptible to the effects of specific voice quality settings (Laver, 1980; Mackenzie-Beck, 1999, 2005). For example, oral speech segments are more susceptible to nasal settings than nasal segments and vice-versa; voiced sounds (vowels and some consonants) are susceptible to phonatory settings, like breathy and creaky voices. For the sake of describing phonetic voice quality settings, the corpus design must take into account the principle of susceptibility, making use of key speech segments and key sentences, containing the susceptible segments.

The compatibility principle states that some VQS can co-occur and others cannot (Laver, 1980; Mackenzie-Beck, 1999, 2005). Some settings can be easily combined, because they are physiologically compatible (lowered larynx and retracted body tongue; lowered larynx and pharyngeal expansion; raised larynx and pharyngeal constriction; raised larynx and laryngeal hyperfunction, for example). Other settings cannot be combined, since they are physiologically incompatible (lowered larynx and raised larynx; pharyngeal constriction and pharyngeal expansion).

So, the speakers' vocal profiles can be drawn in terms of one or more adjustments that can be combined during the time they are speaking, as a long-term composed VQS. The literature indicates recurrent tendencies of grouped voice quality events detected by different perceptual scales (d'Alessandro, 2006; d'Alessandro, Darsinos, & Yegnanarayana, 1998; Dejonckere et al., 1995; Hammarberg & Gauffin, 1995; Kreiman & Gerratt, 2000; Kreiman & Sidtis, 2011; Laver, 1980; Mackenzie-Beck, 2005; Mackenzie-Beck & Schaeffler, 2015; Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995). These findings are also applicable to the VPAS system, taking into account the compatibility principle of voice quality settings

(French, Harrison, Hughes, & Stevens, 2015; Laver, 1980; Mackenzie-Beck, 1999, 2005; Mackenzie-Beck & Schaeffler, 2015; Robieux & Meunier, 2015).

Furthermore, the discussion of the perceptual and acoustic counterparts of voice qualities is important in order to foster knowledge about the links between production and perception (Kreiman & Sidtis, 2011). Considering that voice quality is a prosodic element which has linguistic, paralinguistic and extralinguistic functions, the applications of voice quality analysis are multiple: language descriptions (cross-linguistic variations; Esling, 2000); expressivity investigations (voice expressivity; Barbosa, 2009; Fontes & Madureira, 2015); evaluation and rehabilitation of voice disorders (clinical procedures; Dejonckere et al., 1995; Gillespie, Dastolfo, Magid, & Gartner-Schmidt, 2014; Hammarberg & Gauffin, 1995; Maryn & Weenink, 2015), technological development (voice recognition and synthesis for many purposes, including Augmentative and Alternative Communication Systems) and forensic purposes (speaker recognition; French et al., 2015).

The answers to the following questions are pursued in this work: how can we investigate the perceptual and acoustic correspondences of voice qualities in a system offering so many parameters (VQS)? What about the relevance of some VQS (and their combinations) in the acoustic arena? How can we improve our voice quality descriptions systems?

This investigation aimed at addressing the correspondences between laryngeal and pharyngeal voice quality settings perceptually described by a phonetic grounded profile, the VPAS, and acoustic measures. The acoustic measures were chosen among a restricted set, so to address these phonetic and articulatory phenomena.

2. METHODS

2.1. Corpus description

The general voice quality database is composed of semi-spontaneous speech samples and repetitions of three key sentences (based on the susceptibility principle), read by 278 Brazilian Portuguese (BP) speakers. The voice quality database was perceptually annotated by means of the VPAS system.

Since susceptibility is an important issue for a phonetically grounded voice quality analysis, the use of key speech segments was proposed. For the present investigation, a sub-selection of [a] vowels was extracted from each of the three sentences of the corpus. Because of labeling costs, 44 subjects were selected (10 male and 34 female, ranging from 18 to 58 years old, with a mean age of 30), for a total of 826 vowel samples evaluated.

Four words were targeted from the three key sentences: *fala*, *Lara* and *cidade* (with two occurrences for the latter: at the beginning of the key sentence and at the middle of the key sentence). Each sentence was

repeated several times by each speaker; the number of repetition depends on the speaker.

As an open, lowered, backed and non-rounded vowel, [a] was found to be a susceptible segment for laryngeal and pharyngeal settings in perceptual evaluation (French et al., 2015; Laver, 1980; Mackenzie-Beck, 1999, 2005; Mackenzie-Beck & Schaeffler, 2015; Robieux & Meunier, 2015). For acoustic analysis, this vowel was chosen for its stability, and also because it is the vowel of choice for many investigations involving voice source analysis (Hanson, 1997), including intensity measures (Liénard & Barras, 2013) and periodic-aperiodic decomposition (d’Alessandro et al., 1998).

2.2. Perceptual and acoustic analysis

The perceptual (VPAS) parameters were estimated for each sentence in the dataset described by two expert raters and revised by one, with specific focus on the target vowels [a]. There were 37 parameters perceptually ranked on the profile (Figure 1), with degrees concentrated on a 0 to 4 range. Since VPAS judgments are componential, many zero scores (i. e. neutral voice *for that setting*) were generated in most vocal profiles.

Acoustic measures were extracted from the [a] vowels (the same that were perceptually analyzed along with the VPAS system), by means of scripts run in Praat (Boersma & Weenink, 2016), and thanks to a Matlab implementation of a periodic-aperiodic decomposition algorithm (d’Alessandro et al., 1998):

- Fundamental frequency (f_0), expressed in semitones (with a reference frequency of 1 Hz), a measure notably linked to pitch and register (d’Alessandro, 2006).
- A-weighted *intensity*, expressed in dBA, linked to the perception of voice strength (Liénard & Barras, 2013; Traunmüller & Eriksson, 2000).
- Harmonic-to-Noise Ratio (HNR), expressed in dB, and measured after a periodic-aperiodic decomposition (d’Alessandro et al., 1998), taking into account both additional and structural noises in voiced segments.
- The first formant (F1), expressed in Hz, that is linked to the size of the back cavity (Apostol, Perrier, & Bailly, 2004) and to jaw opening (Erickson, Suemitsu, Shibuya, & Tiede, 2012).
- Amplitude difference between the first harmonic and the third formant (H1-A3), expressed in dB, that is linked to tension and voice strength (Hanson, 1997).

2.3. Integrating perceptual and acoustic data

A multiple factorial analysis was run on the two datasets (VPAS settings and acoustic measures), taking into account all the data. This analysis did not extract links between both datasets, notably because of the scarcity of most VPAS annotations: the presence of zeros as soon as a setting is absent did not mean the

Figure 1: The *Vocal Profile Analysis Scheme VPAS 2007* version (Laver & Mackenzie-Beck, 2007).

Speaker:	Date of recording:	Judge:	Recording ID:					
FIRST PASS		SECOND PASS						
Neutral	Non-neutral	SETTING	moderate		extreme			
			1	2	3	4	5	
A. VOCAL TRACT FEATURES								
1. Labial		Lip rounding/protrusion						
		Lip spreading						
		Labiodentalization						
		Minimized range						
2. Mandibular		Extensive range						
		Closed jaw						
		Open jaw						
		Protruded jaw						
3. Lingual tip/blade		Extensive range						
		Minimized range						
		Advanced tip/blade						
4. Lingual body		Retracted tip/blade						
		Fronted tongue body						
		Backed tongue body						
		Raised tongue body						
		Lowered tongue body						
5. Pharyngeal		Extensive range						
		Minimized range						
6. Velopharyngeal		Pharyngeal constriction						
		Pharyngeal expansion						
7. Larynx height		Audible nasal escape						
		Nasal						
		Denasal						
8. Overall muscular tension		Raised Larynx						
		Lowered Larynx						
B. OVERALL MUSCULAR TENSION								
8. Vocal tract tension		Tense vocal tract						
		Lax vocal tract						
9. Laryngeal tension		Tense larynx						
		Lax larynx						
C. PRONATION FEATURES								
	SETTING	Present	Scalar Degree					
		Neutral	Non-neutral	Moderate		Extreme		
				1	2	3	4	
10. Voicing type	Voice							
	Falsetto							
	Creak							
	Creaky							
11. Laryngeal friction	Whisper							
	Whispery							
12. Laryngeal irregularity	Harsh							
	Tremor							

acoustic parameters won’t change, mostly because other settings may have an effect on them. It is thus difficult to match both datasets. To bypass that limitation, analyses of the links between each acoustic measure and the VPAS parameters were run. In order to address the limitation introduced by scarcity, VPAS parameters that show frequent correlations were combined, in order to have a more robust estimation of an aggregated perceptual dimension.

The voice quality settings detected by the VPAS were firstly categorized in four groups:

- Supralaryngeal (laryngeal height) VQS
- Supralaryngeal (pharyngeal) VQS
- Phonatory VQS
- (Supra)laryngeal tension VQS

In a second stage, for the sake of corresponding perceptual to acoustic descriptions, the compatibility principle was applied, generating the frequent vocal profiles, also based on some references (Camargo, Rusilo & Madureira, 2011; French et al., 2015; Laver, 1980; Mackenzie-Beck, 1999; Mackenzie-Beck & Schaeffler, 2015; Robieux & Meunier, 2015). The following vocal profiles were then generated from the aggregated VQS: “Short vocal tract”, “Wide vocal tract”. The vocal profile named “Short vocal tract” regroups VPAS parameters which cause vocal tract length and width reduction:

- Raised larynx VQS
- Laryngeal hyperfunction VQS
- Pharyngeal constriction VQS
- Closed jaw VQS
- Spread lips VQS

The vocal profile named “Wide vocal tract” regroups VPAS parameters which cause the vocal tract length and width expansion:

- Lowered larynx VQS
- Laryngeal hypofunction VQS
- Pharyngeal expansion VQS
- Creaky voice VQS

The “Wide” and “Short” vocal profiles were attributed grades according to the number of VQS (VPAS parameters) perceived. If no VQS was perceived these vocal profiles were assigned “0”. If just one VQS was perceived they were assigned “1” and if two or more VQS were perceived “+2” was assigned to the “Wide vocal profile” and “-2” to the “Short vocal profile”.

The combination of the two vocal profiles (Wide minus Short) created a single parameter, named Size. These vocal profiles, being linked to the vocal tract length (and width) and to the muscular tension of the vocal apparatus, are supposed to be correlated to acoustic data, such as f_0 , intensity, noise, first formant and spectral slope measures.

This project was approved by Ethics Committee (number 101/11).

2.4. Statistical analysis

Models of analysis of variance (ANOVA) were fitted to each acoustic parameter, so to explore which of the several factors controlled in the corpus and through the VPAS annotation do have an explanatory power for the measures’ variation. The controlled factors are the speakers’ *Gender*, the *Word* that contains the [a] vowel (4 levels), and the VPAS *Size* aggregate (5 levels), and their interactions. The alpha level was set at 5 %.

3. RESULTS

3.1. Fundamental frequency

The ANOVA model explains more than two thirds of the variance ($R^2 = 0.72$). The three main factors are significant, as well as the triple interaction. The factors that account for most of the explained variance are (in decreasing order): unsurprisingly the Gender (partial $\eta^2 = 0.52$), the Word (partial $\eta^2 = 0.31$) and the Size aggregate (partial $\eta^2 = 0.21$). The triple interaction accounts for only 3 % of the variance. Figure 2 shows the effect of both Gender and Size on f_0 .

For female voices, f_0 changes are observed immediately with differences in Size— f_0 rising with smaller sizes. Meanwhile, these changes do reach a ceiling rapidly, while females continue to lower their pitch for Size above 1. A tendency for such changes may be also true for males, but differences are not significant.

3.2. First formant

The ANOVA model for F1 explains less than a third of the variance ($R^2 = 0.27$). All factors have a significant effect, and the factors that account for most of the explained variance are, in decreasing order, Word

Figure 2: Boxplot showing the distribution of f_0 estimations for the 5 levels of Size (from -2 to +2), for females (grey) and males (white).

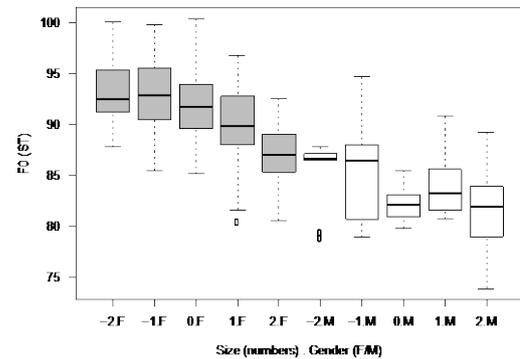
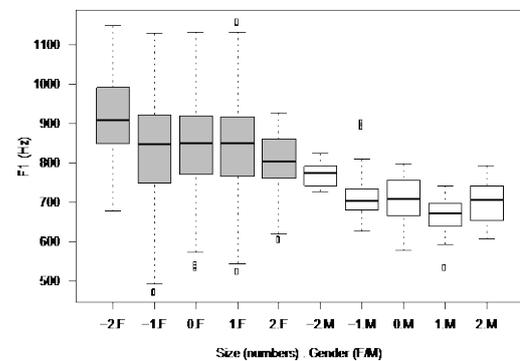


Figure 3: Boxplot showing the distribution of F1 estimation for the 5 levels of Size (from -2 to +2), for females (grey) and males (white).



(partial $\eta^2 = 0.15$), Size (partial $\eta^2 = 0.10$), and Gender (partial $\eta^2 = 0.05$). Figure 3 shows the effect of both Gender and Size on F1.

The first formant is linked to the back cavity (Apostol et al., 2004) and to jaw opening (Erickson et al., 2012); the observed changes are mostly linked to female voices, as for f_0 , but the effects are significant only for the higher levels of the Size factor (-2 or +2), respectively with a formant rise for “small” voice, and a formant fall for “wide” voices.

3.3. Intensity and spectral slope

The ANOVA model for intensity explains about half the variance ($R^2 = 0.43$); all the factors and their interactions are significant. The factors that account for most of the explained variance in intensity are, in decreasing order, Word (partial $\eta^2 = 0.35$), Size (partial $\eta^2 = 0.04$), and its interaction with Gender (partial $\eta^2 = 0.05$), while Gender accounts for about 1 % (partial $\eta^2 = 0.01$). Figure 4 reports the changes of intensity according to Size and gender. The ANOVA model for H1–A3 shows similar patterns, but the model is very messy ($R^2 = 0.27$). It is still linked mainly to the Word position (partial $\eta^2 = 0.15$), and then to Size (partial $\eta^2 = 0.10$) and Gender (partial $\eta^2 = 0.05$).

Changes in intensity are mainly explained by the position of word in the sentence, which is linked to the declination line. For changes that are linked to VPAS,

Figure 4: Boxplot showing the distribution of intensity for the 5 levels of Size (from -2 to +2), for females (grey) and males (white).

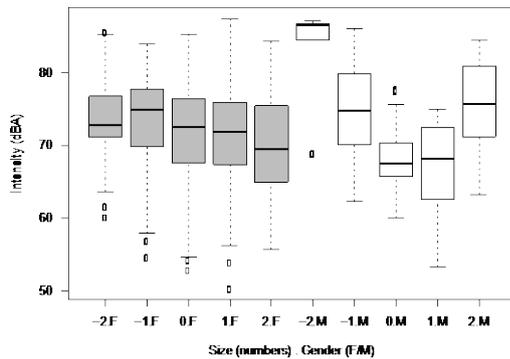
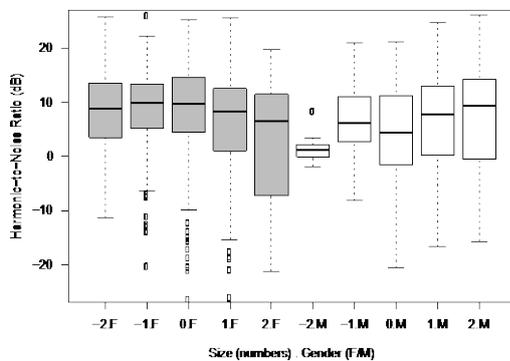


Figure 5: Boxplot showing the distribution of HNR estimation for the 5 levels of Size (from -2 to +2), for females (grey) and males (white).



they show an interaction with Gender: it seems that mostly males do vary intensity in correlation with the Size factor, increasing it when their voices depart from the neutral setting. A slight declination is observed in female voices characterized by lower pitches and wider tracts or laxer voice qualities and thus may be explained by changes in pitch (Titze & Sundberg, 1992).

3.4. Harmonic-to-Noise Ratio

The ANOVA model for HNR explains only a small part of the variance ($R^2 = 0.16$), if all factors have a significant effect. The factors that account for a part of the explained variance are, in decreasing order, Word (partial $\eta^2 = 0.07$) and Size (partial $\eta^2 = 0.02$). Figure 5 shows the effect of both Gender and Size on HNR.

One may observe on the graphs that only the most extreme value of the size factor do show more aperiodicities—typically for the lowest pitch voice (which are also at a low intensity) in female speakers, that could be related to creaky phenomenon.

This aggregate is not mostly related to this measure of noise, and one may rather observe potential relation with other possible aggregates.

4. DISCUSSION

To face the challenge of relating perceptual and acoustic data from voice qualities, speaker-standardized and non-standardized acoustic measures

were initially taken into account. Non-standardized measures provided more interesting results than standardized measures, as soon as the speaker's gender was considered in the statistical models. This is linked to the fact that standardization, when removing speaker-specific changes, also removes the specificities of voice quality that characterize the speaker's voice.

Among the various acoustic measures extracted from the corpus, f_0 was found to be the most relevant measure to relate the “Short” and “Wide” VPAS aggregates (cf. Figure 2). A point to consider in this discussion is the set of VPAS parameters that contribute to this general “size” vocal profile. Despite the fact that we found similar effects on f_0 measures, it is important to consider that many possible VQS combinations could be implemented in the speakers' vocal tract, leading to increasing (or decreasing) f_0 in distinct proportions. Some of these combinations of settings include notably a dimension of noise, that would probably be better expressed by measures such as harmonic-to-noise ratio (d'Alessandro, 2006; d'Alessandro et al., 1998).

Among the findings of this study is an effect of gender on the acoustic parameters which are relied on this “Short”/“Wide” vocal profiles. In this corpus, females clearly go for pitch as a primary cue, and then for F1; on the contrary males seems to rely on intensity rather than pitch.

This result is to be taken with caution, because of the relatively small set of male speakers included in the corpus (10 speakers, compared to the 34 female speakers). Meanwhile, it could be related to a social habit in BP for males to use a lower voice register, and/or for females to use a comparatively higher voice register. This could be related and explained in a similar way to the difference in pitch described between Japanese and Dutch women, that is linked by Van Bezooijen (1995) to the representations of gender in these two societies, and expected to be found in males also.

Yet, some factors influencing voice quality patterns have not been addressed in studies focusing acoustic-perceptual correlations. Some of them are related to speakers age and gender normalization, and intra-speaker variations. Other challenges are related to the overlapping of voice quality events and the degree of influence of the setting in the general vocal profile.

To address these limitations, we proposed to explore the vocal profiles, i.e. the combinations of voice quality settings that tend to be productive in daily communications, and even, in the voice disorder arena. In future explorations, the relevance of each VQS for the final vocal profile definition must also be addressed. The difficulty to focus on long-term events in relating them to intermittent or shot-term occurrences may also interfere in the study of acoustic correlates of VQS, like voice breaks and sudden voice quality changes. They were not frequent in the *corpus* analyzed, because this corpus was annotated in VPAS and vowel samples were revised to search for some

specific events in the key-speech segment (the vowel [a]).

To face all these limitations in describing voice qualities, a statistical model has been proposed and is meant to be improved in later works.

5. CONCLUSIONS

In this investigation, we depart from the phonetic approach of voice qualities, taking into account the susceptibility and compatibility theoretical principles. To fit the principles of the model by Laver et al. (1981), we had to consider the combinations of laryngeal and pharyngeal VQS and the inherent phonetic characteristics of the speech segments.

The methodological procedures made it possible to identify the relevance of f_0 measures as a main cue, and F1, intensity and H1–A3 as secondary cues, to describe perceptual data related to pharyngeal and laryngeal adjustments, generating the “Wide” and “Short” vocal tract kinds of profiles.

The findings provide evidence in favor of the relevance of the phonetic description of voice qualities.

6. REFERENCES

- Apostol, L., Perrier, P., & Bailly, G. (2004). A model of acoustic interspeaker variability based on the concept of formant–cavity affiliation. *The Journal of the Acoustical Society of America*, *115*(1), 337–351.
- Barbosa, P. A. (2009). Detecting changes in speech expressiveness in participants of a radio program. In *Proceedings of Interspeech* (pp. 2155–2158).
- Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.20, retrieved 3 September 2016 from <http://www.praat.org/>
- Camargo, Z., Rusilo, L. C., Madureira, S. (2011). Evaluating speech samples designed for the Voice Profile Analysis Scheme for Brazilian Portuguese. Presented at the Fourth ISCA Tutorial and Research Workshop on Experimental Linguistics, Paris, France.
- d’Alessandro, C. (2006). Voice source parameters and prosodic analysis. In S. Sudhoff, D. Lenertova, R. Meyer, S. Pappert, P. Augurzy, I. Mleinek, N. Richter, & J. Schlieer (Eds.), *Language, context, and cognition: Methods in empirical prosody research* (pp. 63–87). Berlin: Walter de Gruyter.
- d’Alessandro, C., Darsinos, V., & Yegnanarayana, B. (1998). Effectiveness of a periodic and aperiodic decomposition method for analysis of voice sources. *IEEE Transactions on Speech and Audio processing*, *6*(1), 12–23.
- Dejonckere, P. H., Remacle, M., Fresnel-Elbaz, E., Woisard, V., Crevier-Buchman, L., & Millet, B. (1995). Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. *Revue de laryngologie-otologie-rhinologie*, *117*(3), 219–224.
- Erickson, D., Suemitsu, A., Shibuya, Y. & Tiede, M. (2012). Metrical structure and production of English rhythm. *Phonetica*, *69*, 180–190.
- Esling, J. H. (2000). Crosslinguistic aspects of voice quality. In R. D. Kent, & M. J. Ball (Eds.), *Voice quality measurement*. San Diego: Singular Publishing Group.
- Fontes, M. A. S., Madureira, S. (2015). Gestural prosody and the expression of emotions: a perceptual and acoustic experiment. Presented at the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.
- French, P., Harrison, P., Hughes, V., & Stevens, L. (2015). The vocal tract as a biometric: output measures, interrelationships, and efficacy. Presented at the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.
- Garellek, M. (2014). Voice quality strengthening and glottalization. *Journal of Phonetics*, *45*, 106–113.
- Gillespie, A. I., Dastolfo, C., Magid, N., & Gartner-Schmidt, J. (2014). Acoustic analysis of four common voice diagnoses: moving toward disorder-specific assessment. *Journal of Voice*, *28*(5), 582–588.
- Hammarberg, B., & Gauffin, J. (1995). Perceptual and acoustic characteristics of quality differences in pathological voices as related to physiological aspects. In O. Fujimura, & M. Hirano (Eds.), *Vocal Fold Physiology: Voice Quality Control* (pp. 283–303). San Diego: Singular Publishing Group.
- Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America*, *101*(1), 466–481.
- Kreiman, J., Gerratt B. (2000). Measuring vocal quality. In R. D. Kent & M. J. Ball (Eds.), *Voice quality measurement* (pp. 73–101). San Diego: Singular Publishing Group.
- Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. Oxford: Wiley-Blackwell.
- Laver, J. (1980). *The phonetic description of voice quality*. Cambridge: Cambridge University Press.
- Laver, J., & Mackenzie-Beck, J. (2007). Vocal Profile Analysis Scheme-VPAS. *Queen Margaret University College-QMUC, Speech Science Research Centre, Edinburgh*.
- Laver, J., Wirz, S., Mackenzie, J., & Hiller, S. (1981). A perceptual protocol for the analysis of vocal profiles. *Edinburgh University Department of Linguistics Work in Progress*, *14*, 139–155.
- Liénard, J.-S. & Barras, C. (2013). Fine-grain voice strength estimation from vowel spectral cues. In *Proceedings of Interspeech* (pp. 128–132).
- Mackenzie-Beck J. (1999). Organic variation of the vocal apparatus. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 256–297). Malden, MA: Blackwell.
- Mackenzie-Beck, J. (2005). Perceptual analysis of voice quality: the place of vocal profile analysis. In

- W. J. Hardcastle, & J. Mackenzie-Beck (Eds.), *A figure of speech: a festschrift for John Laver* (pp. 285–322). Mahwah: Lawrence Erlbaum Associates.
- Mackenzie-Beck, J., & Schaeffler, F. (2015). Voice quality variation in Scottish adolescents: gender versus geography. Presented at the 18th International Congress of Phonetic Sciences, Glasgow, Scotland.
- Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice*, 29(1), 35–43.
- Rabinov, C. R., Kreiman, J., Gerratt, B. R., & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of Speech, Language, and Hearing Research*, 38(1), 26–32.
- Robieux, C., Meunier, C. (2015). *Phonetic considerations in vocal effort assessment*. Presented at the 11th Pan-European Voice Conference, Firenze, Italy.
- Sundberg, J., & Gauffin, J. (1979). Waveform and spectrum of the glottal voice source. *Frontiers of speech communication research*, 301–322.
- Titze, I. R. & Sundberg, J. (1992). Vocal intensity in speakers and singers. *The Journal of the Acoustical Society of America*, 91(5), 2936–2946.
- Traunmüller, H. & Eriksson, A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*, 107(6), 3438–3451.
- Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265.