

*Received 14 February 2016.*

*Accepted 31 March 2016.*

## **VOICE QUALITY AND SPEAKING STYLES**

Sandra MADUREIRA, Mario Augusto de Souza FONTES & Beatriz Coelho FONSECA

Pontifícia Universidade Católica de São Paulo

madusali@pucsp.br / fontes@pucsp.br/ beatrizcf21@gmail.com

### **Abstract**

The objective of this article is presenting an acoustic phonetic and perceptual experiment which explores voice quality and dynamic characteristics in the reading of prose texts, contrasting professional and non-professional voice speakers. In the task of perceptual evaluation, a semantic five-point rating differential scale was used. Two descriptors were introduced: “pleasantness of voice quality” and “professional voice style”. The results of the perceptual evaluation task indicated that the judges were able to differentiate between professional and non-professional voice speakers and that this differentiation could be explained in terms of vocal dynamic characteristics. In relation to the judgments of pleasantness of voice quality, the differentiation was guided by characteristics of vocal quality settings independently of them being professionals or not. Correlations between acoustic and perceptual parameters were based on multivariate analysis. Contributions to the discussion of the indexical power of voice quality for the attribution of characteristics to speakers are presented.

### **Keywords**

voice quality, vocal dynamics, speech style, professional voice, prosody

## **QUALIDADE DE VOZ E ESTILO DE FALA**

### **Resumo**

O objetivo deste artigo é apresentar um experimento fonético-acústico e perceptivo que explora características da qualidade e da dinâmica vocais em leitura de texto de prosa, contrastando locutores profissionais e não profissionais. Na tarefa de avaliação perceptiva, utilizou-se uma escala de diferencial semântico de cinco graus com os seguintes descritores: “agradabilidade de voz” e “estilo de locução profissional”. Os resultados do teste de avaliação perceptiva indicaram que os juízes foram capazes de

diferenciar entre locutores profissionais e não profissionais e essa diferenciação pode ser explicada em termos de características da dinâmica vocal. Em relação à agradabilidade da voz a distinção pautou-se por tipos de ajustes que caracterizavam a voz dos locutores, independentemente de serem profissionais ou não. Foram feitas correlações entre parâmetros acústicos e perceptivos por meio de análise multivariada. Contribuições para a discussão sobre o valor indexical da qualidade e dinâmica vocais na atribuição de características aos falantes são apresentadas.

### **Palavras-chave**

qualidade de voz, dinâmica vocal, estilo de fala, locução profissional, prosódia

## **1. Introduction**

The objective of this article is presenting an acoustic phonetic and perceptual experiment which explores voice quality and dynamic characteristics in the reading of prose texts, contrasting professional and non-professional voice speakers.

Voice quality is a prosodic element which has linguistic, paralinguistic and extralinguistic functions. For a long time in speech studies impressive labels were used to describe types of voice quality. Laver (1980) remarkably changed this panorama by introducing a descriptive phonetic model of voice quality analysis. The basic analytical unit in Laver's model, the setting, is defined as a long-term muscular tendency in the supralaryngeal and laryngeal parts of the vocal tract.

In the phonetic descriptive model of voice quality proposed by Laver (1980) fifty-three types of vocal quality settings and two principles governing them are introduced. The principle of susceptibility accounts for the fact that some speech segments are more susceptible to the effects of some voice quality settings than others. The principle of compatibility holds that some voice quality settings can co-occur while others can't. Oral sounds, for example are more susceptible to the nasal setting of voice quality than nasal sounds are as well as a nasal setting cannot co-occur with a denasal setting.

Based on the phonetic description of voice quality model by Laver (1980) a protocol, the Voice Profile Analysis Scheme (VPAS), was developed in 1981 to provide a graphic representation of speakers' vocal profiles (Mackenzie Beck 2005). The VPAS

was modified in 1991, 2000 and 2007. The 2007 version (Laver & Mackenzie Beck 2007) has two main sections, one for describing voice quality settings and another for describing the dynamic vocal aspects, that is, the prosodic aspects and temporal organization. Two further features are also included: respiratory support and diplophonia.

The section concerning the vocal quality settings is subdivided into three parts: vocal tract features, overall muscular tension and phonation features. Prosodic features comprise pitch and loudness and temporal organization includes speech rate and continuity.

The voice quality settings are described as variations from a reference setting, the neutral one, in which there are no constrictive or expansive effects in the vocal tract cavities; no shortening or lengthening of the extension of the vocal tract and no extreme variations in terms of muscular tension activity in the supralaryngeal and laryngeal parts of the vocal tract.

To apply the VPAS, judges are supposed to have a thorough phonetic background and experience on using it. Identification of the voice quality settings are made in two passes. In the first pass non-neutral settings are identified. In the second pass the judge is asked to evaluate in a scalar degree (from 1 to 6) the non-neutral settings which were identified in the first pass.

The VPAS has been used to support linguistic, paralinguistic and extralinguistic research. Mackenzie-Beck (2005) mentions the potential of VPAs for foreign language learning and teaching and drama teaching. It has been applied in clinical, educational and forensic contexts. Among the research studies which have been carried out, there are works on voice pathology (Camargo, Madureira & Tsuji 2003; Camargo & Madureira 2009), gender and regional characteristics (Bonfim et al. 2007; Camargo et al. 2012; Schaeffler, Mackenzie-Beck & Jannets 2015) mother-child interaction (Marwick et al. 1984), emotion expression (van Bezoooyen 1984; Fontes 2014 Fontes & Madureira 2015), social, personality and attributes of speakers (Camargo, Madureira & Schmitz 2013; Aragão et al. 2014), speech expressivity (Madureira 2008; Madureira & Camargo 2010; Madureira 2011, Viola & Madureira 2007), speaker identification

(French et al. 2015) and on the applicability of the profile (Laver 1991, 2000, Laver Mackenzie-Beck 2001; Mackenzie-Beck 2005; Rusilo, Camargo & Madureira 2011).

## 2. Methods

The corpus of the experiment consists of a three hundred eighty-eight-word prose text describing the characteristics of the Brazilian social structure. It was recorded in a studio at the Pontifical University of São Paulo with the help of a sound technician.

The text was read by eight female subjects aged 21 to 45 years old, four of them voice professionals (Subjects 2, 3, 5 and 6) two university teachers (subjects 1 and 7) and two undergraduates (Subjects 4 and 8). Subject 8 had some experience in amateur theater.

For the sake of acoustic analysis and perceptual tasks the text was divided into eight paragraphs. The following analytical methodological procedures were applied to each one of these paragraphs: acoustic analysis; perceptual analysis of voice quality settings, semantic differential scale questionnaire and multivariate statistical analysis. Therefore, two kinds of variables were concerned: qualitative and quantitative. The qualitative variables are the judgements of the VPAS settings and the judgments of the semantic differential scale. The quantitative variables are the measures extracted by the ExpressionEvaluator as explained in the following paragraph. Consequently, three groups of variables were formed: Gc1 (ExpressionEvaluator measures) and Gc2 (Judges' evaluation) and GC3 (judgements of the settings based on the VPAS). To correlate them, a non-parametric test applying Multiple Functional Analysis (MFA) methods was used.

Acoustic measures were automatically extracted by the ExpressionEvaluator Script developed by Barbosa (2009) which runs in PRAAT. The script extracts 12 measures: -f0 measures: f0 median (mdnf0), inter-quartile semi amplitude (sampquartisf0), skewness and 0,995 quantil (quan995f0); -f0 derivative: df0 mean (medderivf0), standard deviation (desvpaddf0), skewness (assimdf0div10); intensity

measures: intensity skewness (assimint); promptness (the difference between the acoustic energy of the integral signal and the intensity of the low pass filtered signal, upper band limit equal to  $1,5 * \text{average } f_0$  of the acoustic signal under analysis); spectral tilt: spectral tilt mean (medinclinespec), standard deviation (desvadinclinespec), skewness (assiminclinespec); and LTAS: LTAS frequency standard-deviation (desvapadltas).

The semantic differential scale questionnaire was built up containing two descriptors to be judged: pleasantness of the voice quality and professional speaking skills. A five-point-rating scale was used, introducing an adjective at one end and its antonym at the other end: rate 1 was meant to be very unpleasant/unprofessional and 5 very pleasant/professional. Along these lines, judges were asked to say if the speaker's voice quality was pleasant or unpleasant and if they considered the speaker to be a voice professional or not.

The semantic differential scale questionnaire test was applied to a group of 80 judges. The judges were undergraduate and graduate students attending courses at the Pontifical University of São Paulo. The stimuli were randomly presented.

In order to identify the vocal quality settings the Vocal Profile Analysis Scheme (VPAS) developed by Laver & Mackenzie-Beck (2007) was used. The settings were described by a phonetician with great expertise in the use of the scheme.

The number of utterances, stretches of speech between silent pauses as defined by Laver (1994) was also counted. The utterances were identified with the help of PRAAT.

In order to correlate the qualitative and quantitative data, an explorative multivariate analysis method, the Multiple Functional Analysis (MFA) was used. The data were analyzed with the software R, R-Commander and FactoMiner R (Husson et al. 2013).

Three groups of variables were taken into account: Gc1 refers to the Expression Evaluation measures; Gc2 to the judges' evaluation and Gc3 to the VPAS features.

### 3. Results

#### 3.1 The evaluation of the vocal quality settings

The application of the VPAS yielded the vocal profiles of the subjects. The identified settings are the ones which were judged as non-neutral. As mentioned before VPA establishes a six-point scale, being 1 the lowest and 6 the highest degree. Table 1 presents the settings which were found to characterize the subjects' voice qualities, specifying in brackets its degree. As mentioned before S2, S3, S5 and S6 are the professional voice speakers and the others the non- professional voice speakers.

Subjects	Voice quality settings	Vocal Dynamic Aspects
S1	Whispery voice [1]	Minimised pitch range [1]; Fast rate [1]
S2	Whispery voice [1]; Pharyngeal expansion [2]	Extensive loudness [1]; Extensive pitch range [1]
S3	Lowered larynx [3]; Harsh voice [1]	Low pitch [3]; Fast rate [1];
S4	Whispery voice [2]; Backed tongue body [1]	Minimised pitch range[2]
S5	Whispery voice [1]; Fronted tongue body [1]; Raised larynx [1]	Extensive pitch range [2]
S6	Harsh voice [2]	Low pitch [1]; Minimised loudness range [1]; Fast speech rate [3]; Minimised pitch range [1];
S7	Nasal voice [1]; Tense larynx [1]	High pitch [1]
S8	Fronted tip tongue [2]; Fronted tongue body [2]; Tense larynx [2]; Raised larynx [2]	High Pitch [3]; Minimised pitch range [1]; Fast rate[1]

Table 1. Results of the application of VPAS: speakers' vocal profiles. In brackets the degree of the perceived setting

Four subjects' voices (S1, S2, S4 and S5) were characterized as whispery, S4 with a higher grade than the others. Two subjects presented harsh voice settings and two subjects were characterized by extensive pitch range. Only two speakers were characterized with degree 2 of pitch range, one of them minimized (Subject 4) and the other maximized (S5).

For the sake of example, the fundamental frequency trace of a prosodic phrase as produced by S5 and S4 are contrasted in Figure 1a. In the stressed syllable of the

word “modo” (way) the  $f_0$  varies 102 Hz in S5’s speech and 28 Hz in S4’s speech. Figure 1b show the waveforms and the wideband spectrograms of these same productions. The relative difference between the duration of the production of the indefinite article “um’ in the productions of S5 (151 ms in a phrase whose total duration is 1115 ms) and S4 (59 ms in a phrase whose total duration is 1063 ms) is striking and may reflect different strategies used by these two subjects.

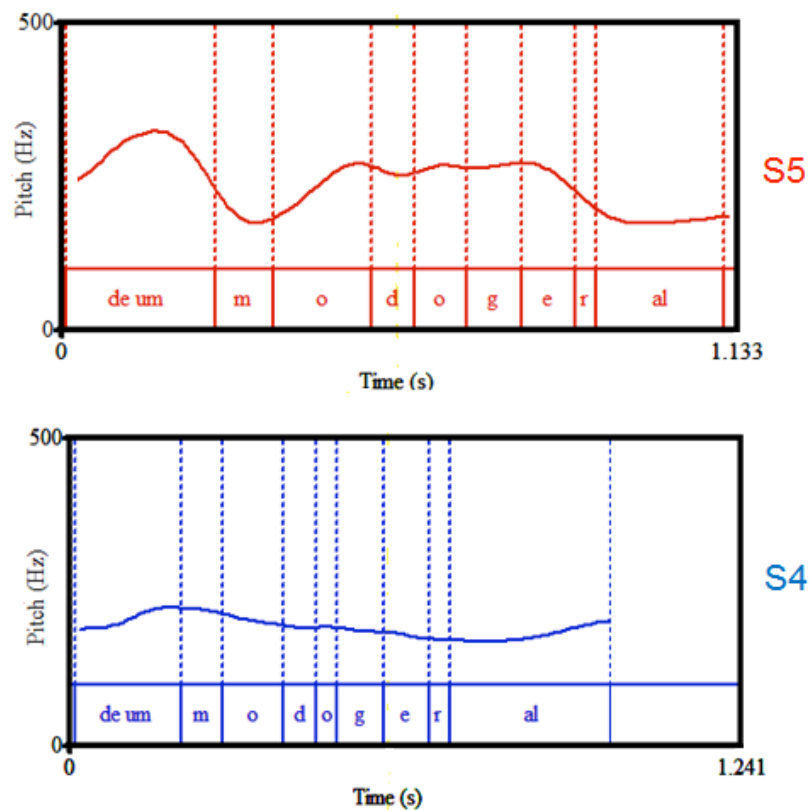


Figure 1a. Fundamental frequency traces of the same phrase (de um modo geral/ in a general way) as spoken by subjects 5 (S5) and 4 (S4).

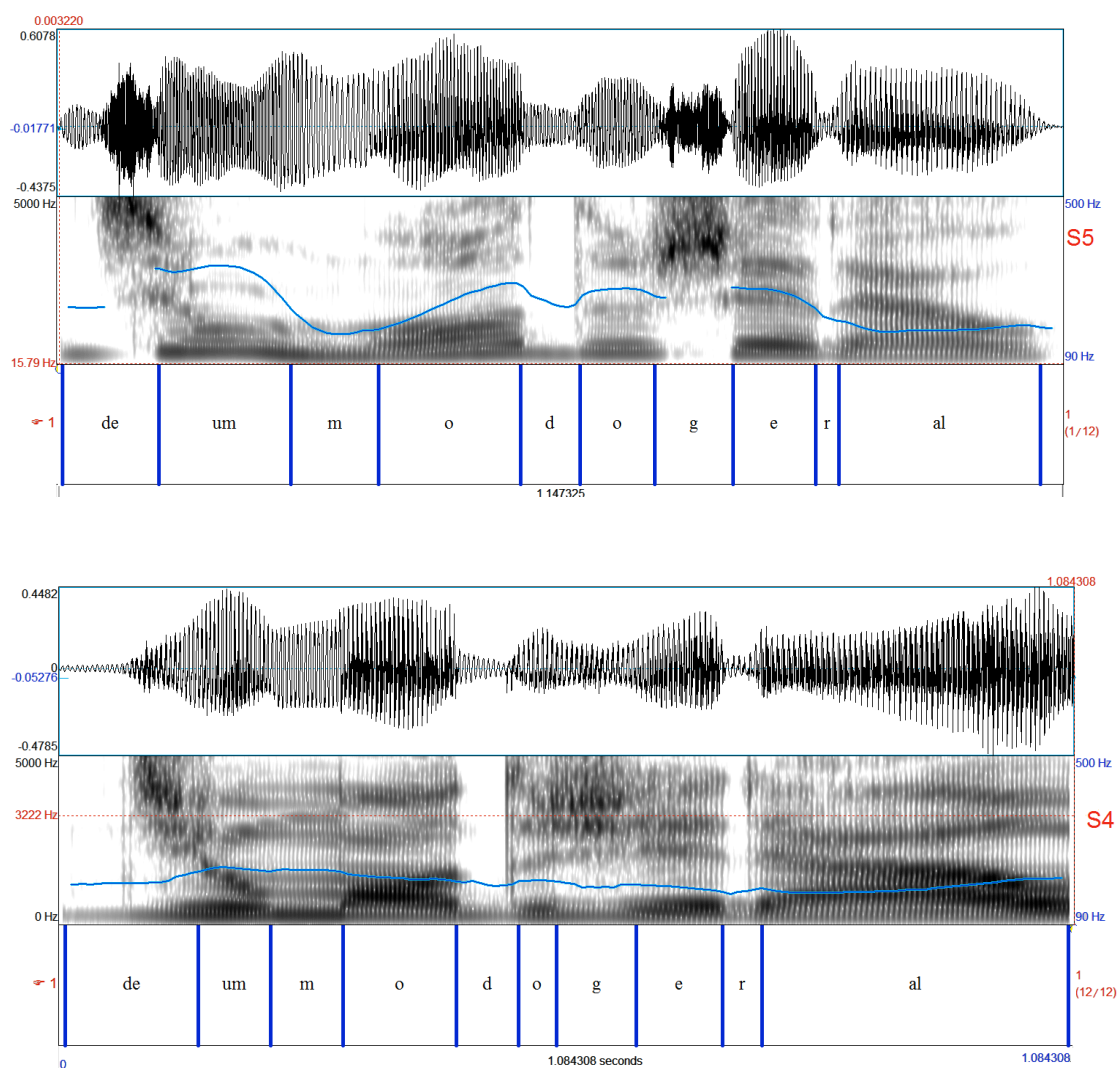


Figure 1b. Waveform, wideband spectrogram and fundamental frequency trace of the same phrase (de um modo geral/ in a general way) as spoken by subjects 5 (S5) and 4 (S4).

### 3.2 The number of utterances in the reading task

Speakers differed in the number of utterances produced in the reading task. This means that some speakers used more pauses than others. Utterance is defined here in the same way as Laver (1994) as the stretch of speech between silences. Professional voice speakers (S2, S3, S5, and S6) produced a larger number of utterances as compared with non-professional voice speakers (S1, S4, S7 and S8). Subject 8, who had some experience in amateur theater, produced a greater number of utterances than the other non-professional voice speakers. This might be related to her training in acting. Figure 2 displays the total number of utterances produced by each subject.



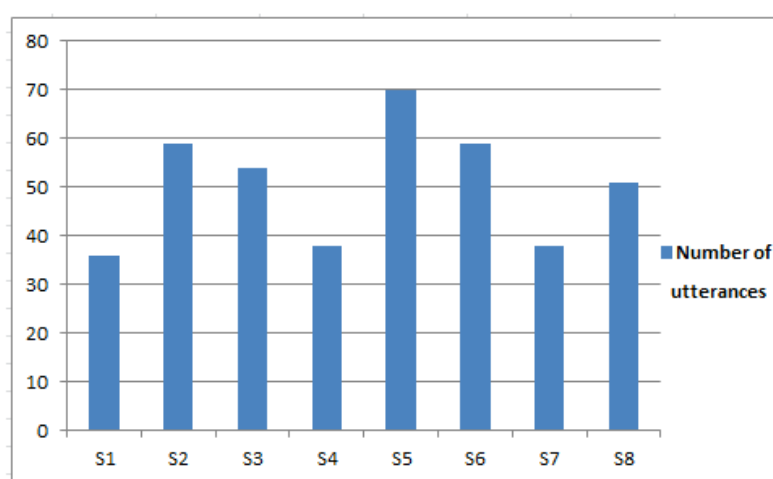


Figure 2. Number of utterances per subject in the reading task

### 3.3 The judges' evaluation task

The judges evaluated two descriptors: pleasantness of the voice quality and professional voice skills. Resultant scores took into account the grades, the weighted averages and the number of evaluations. The results of the evaluation of the subjects are showed in Figures 3, 4 and 5 as to the identification of professional voice skills, pleasantness of the voice and both professional skills and pleasantness of the voice.

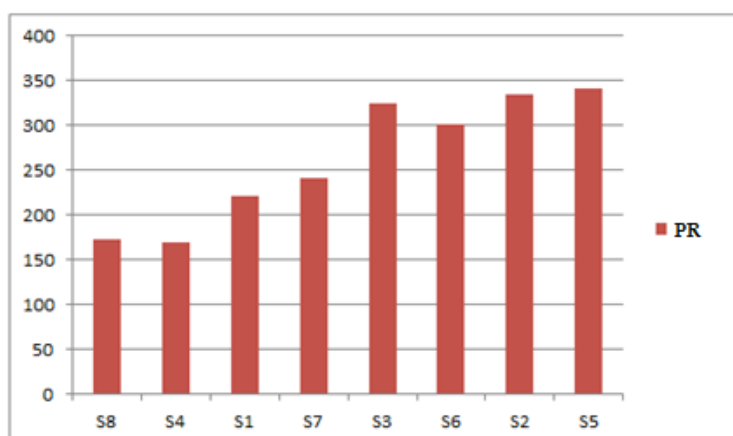


Figure 3. Global evaluation of the eight subjects concerning the descriptor professional voice

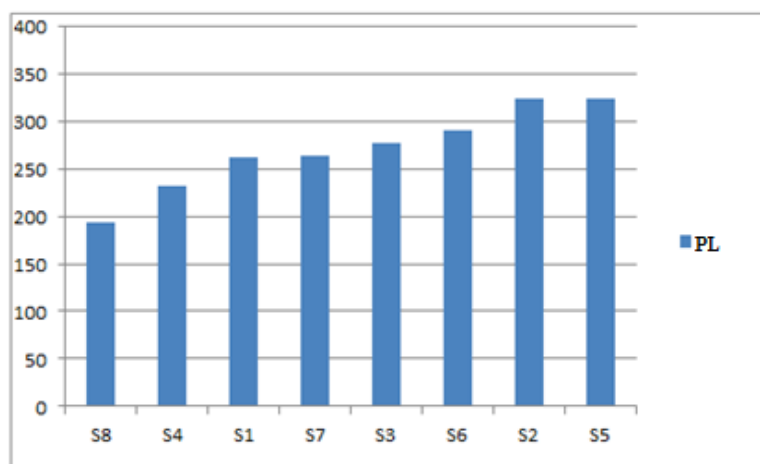


Figure 4. Global evaluation of the eight subjects concerning the descriptor pleasantness of voice quality

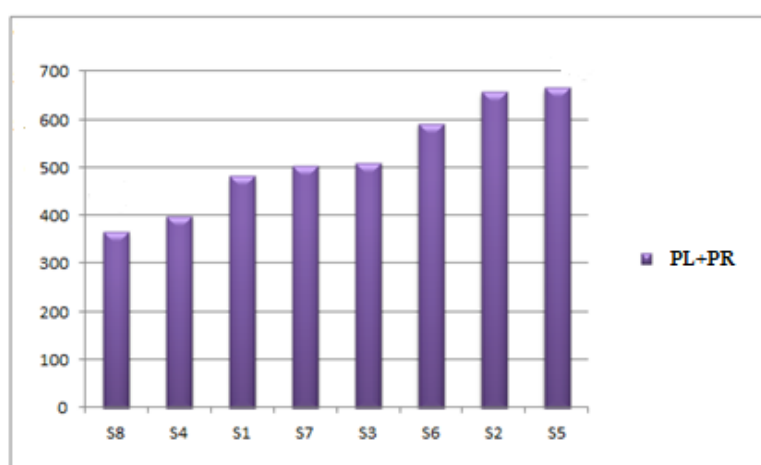


Figure 5. Global evaluation of the eight subjects concerning the two descriptors professional voice and pleasantness of voice quality

Subjects 5, 2 and 6 (professional voice speakers) got the highest overall scores and subjects 8 and 4 the lowest. Subjects 1 and 7 (non-professional voice speakers) and subject 3 (a professional voice speaker) formed an intermediate group.

Comparison in percentage values between the results of the judgements of pleasantness of voice quality and professional voice skills for the eight subjects indicate that S4 had the highest percentage values for pleasantness of voice quality and the lowest for professional voice skills. This can be interpreted in function of her whispery voice quality setting (grade 2) and the restricted number of prosodic utterances.

All professional subjects (S5, S2, S6 and S3) got higher percentage scores in relation to the judgements of professional voice skills and all non-professional voice speakers were better evaluated in relation to pleasantness of voice quality. Figure 6 displays the percentage values related to these judgments.

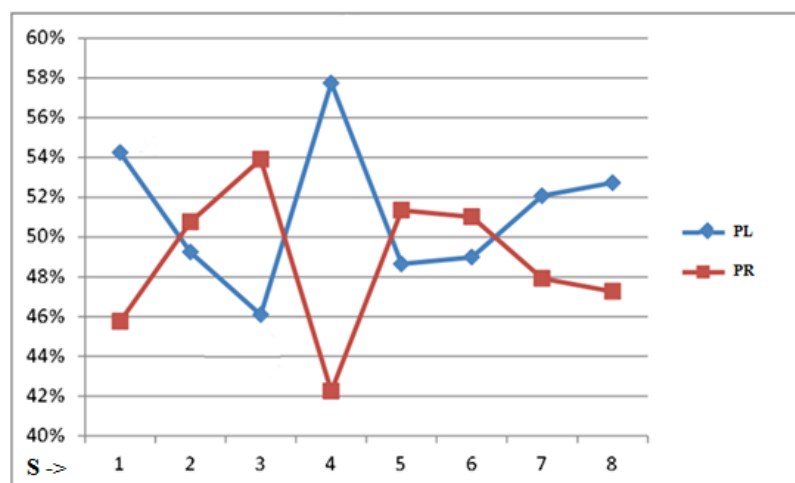


Figure 6. Comparison in percentage values between the results of the judgments of pleasantness in voice (PL) and professional voice skills (PR) for the eight subjects

### 3.4 Acoustic measures

The ExpressionEvaluator was applied to the audio files of the 8 paragraphs recorded by the 8 subjects (64 paragraphs). The following table shows the average values obtained for each subject's speech production.

Table 2 displays the average values concerning the 12 acoustic measures obtained by the application of the ExpressionEvaluator script to the 64 paragraphs produced by the 8 speakers (S1, S2, S3, S4, S5, S6, S7 and S8). Each subject read 8 paragraphs.

Measures	S1	S2	S3	S4	S5	S6	S7	S8
mednf0	-0,30375	-0,49875	-0,5575	-0,3125	-0,2575	-0,4325	-0,1675	-0,08
sampquartisf0	0,39375	0,545	0,5675	0,35625	0,635	0,68125	0,5075	0,445
quan995f0	1,10375	0,5375	1,80875	0,4125	0,855	0,8925	0,8375	1,4725
assimf0	0,0825	0,12525	0,13125	0,08375	0,0775	0,155	0,06325	0,105
medderivf0	-0,14375	-0,4875	0,012388	-0,19675	-0,47	-0,7225	-0,29875	-0,38375
desvpaddf0	0,73625	0,735	0,91875	0,3675	0,65875	0,5975	0,53625	0,8125
assimdf0div10	0,11	0,06875	0,0575	0,22	0,345	0,4275	0,2675	0,15875
assimint	0,19625	0,155	0,30875	0,2275	0,31	0,2875	0,26875	0,18375
medinclin-spec	0,25375	0,2475	0,22875	0,2475	0,26875	0,2525	0,24125	0,20625
desvpadinclin-spec	0,1475	0,1525	0,1775	0,1475	0,18625	0,16625	0,14375	0,15375
assiminclin-spec	-0,14625	-0,05	-0,17875	-0,1825	-0,24375	-0,32325	0,0525	-0,21403
slopeLTAS	1,85375	1,7825	1,46125	1,76625	1,94125	1,6375	1,33375	1,5175

Table 2. Average values concerning the 12 acoustic measures for the paragraphs read by the 8 subjects

These values were submitted to statistical multivariate analysis as considered in the following subsection.

### 3.5 The multivariate analysis

The application of the MFA method yielded two main clusters. In one of the clusters are the voice professionals and in the other the non-professionals. The grouping of the subjects reflected similarities in terms of their vocal performance. Figure 7 displays the dendrogram with the derived clusters.

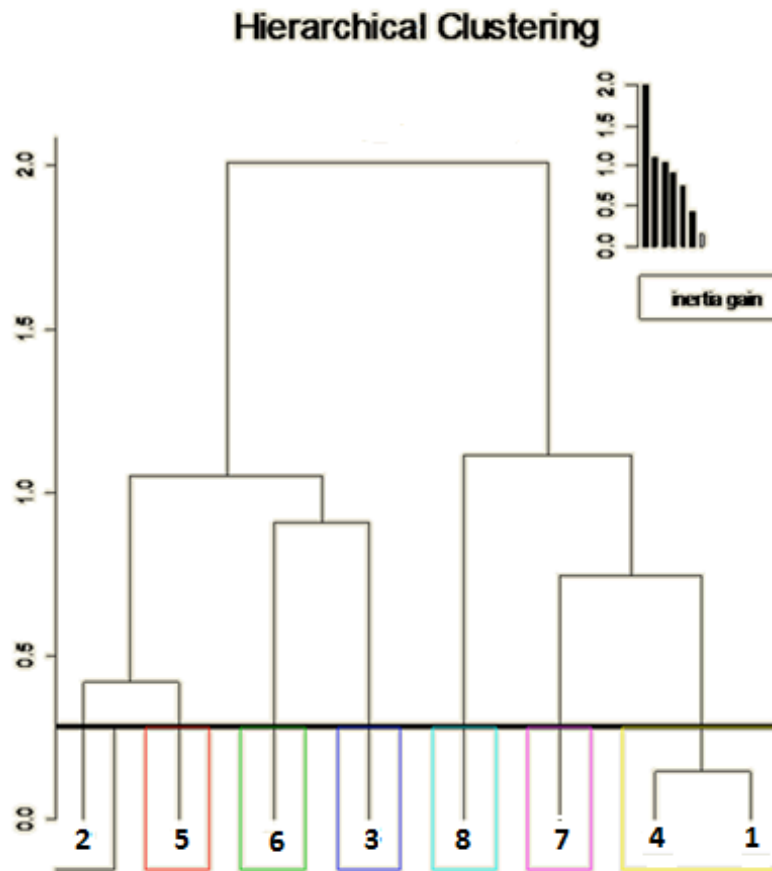


Figure 7. Dendrogram showing the clusters grouping the eight subjects. On the left cluster the voice professionals and on the right clusters the non-voice professionals are grouped.

The division into clusters took into account the 3 groups of analytical variables: the ExpressionEvaluator, the judges' evaluation and the VPAS. The projection of the groups of variables can be seen in Figure 8. They are quite equivalent as far as their distribution in the vector space is concerned.

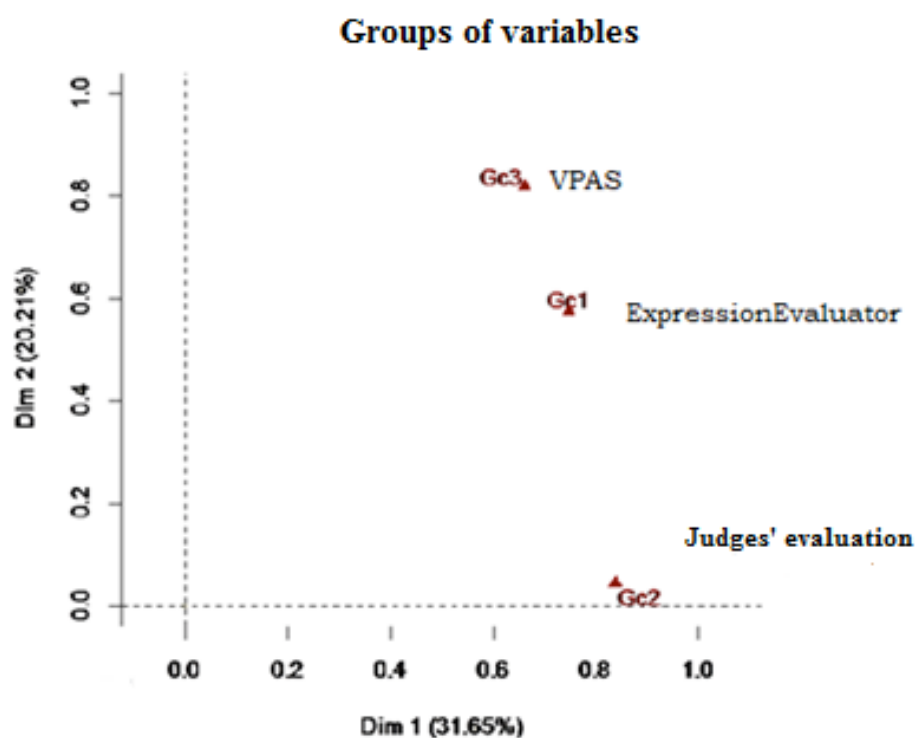


Figure 8. The distribution of the group of variables, Gc1 (ExpressionEvaluator) and Gc2 (Judges' evaluation) and GC3 (VPAS) in two dimensions (Dim 1 and Dim 2) of the vector space.

The distribution of the speakers in Dimensions 1 and 2 shows that speaker 8, the one who got the lowest score isolated in one of the quadrants. The non-professional voice speakers S1, S4 and S7 share one quadrant of the vector space. In one of the other two quadrants, the voice professional speakers S5 and S2 are together and in the other S6 and S3. This distribution equates with the results of the judges' evaluations. In Figure 9 the relative distribution of the subjects in the vector space is showed.

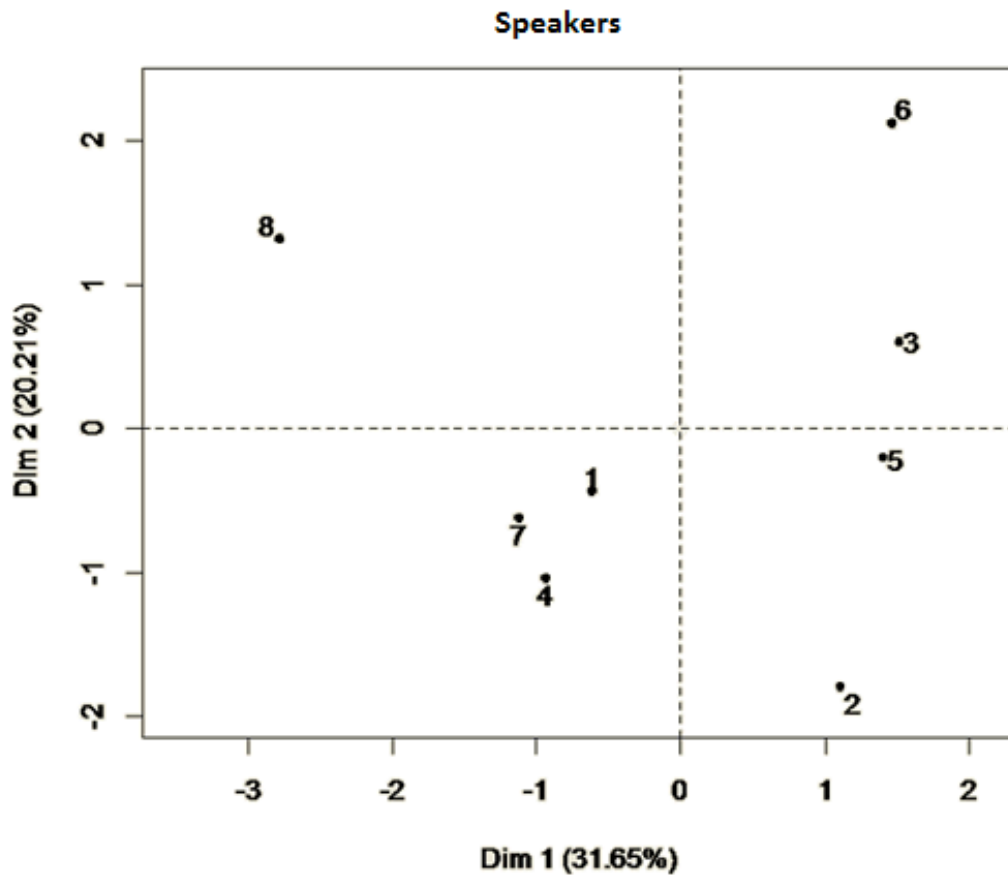


Figure 9. Factor map showing the relative distribution of the subjects in the vector space

The quantitative acoustic variables *sampquartisf0* and *mednf0* were found to be significant ( $p < 0,05$ ) in dimension one (DIM 1) of the vector space. In dimension two (DIM 2) the significant variables were the VPAS variables Harsh Voice and Minimized Loudness. They were also strongly correlated. The *f0* semi-amplitude quartiles (*sampquartisf0*) is a measure of the variation of the *f0* values, spurious values excluded. The *f0* median (*mednf0*) refers to both the variation of *f0* values as well as to the speed in which the *f0* changes.

Table 3 displays the correlation coefficients and the *p*-values concerning these variables.

• DIM 1	•	•	• DIM 2	•	•
• Variable	• correlation	• p,value	• variable	• correlation	• p, value
• Z,PR	• 0,9104	• 0,0017	• Harsh Voice	• 0,7275	• 0,0408
• Z,PL	• 0,8828	• 0,0037	• Minimized Loudness	• -0,815	• 0,0137
• Sampquartisf0	• 0,7473	• 0,0331	•	•	•
• Tense larynx	• -0,8046	• 0,016	•	•	•
• Mednf0	• -0,8171	• 0, 0133	•	•	•
• Pitch	• -0,8181	• 0.0131	•	•	•

Table 3. Correlation coefficients of the significant variables

The contributions of the group of variables can be seen in Table 3. In dimension one (DIM 1) the contributions of the three groups of variables are approximate. In the other dimensions the greater contributions are from the Expression Evaluator group (Gc1) in dimension 3 (DIM3) and the VPAS group (Gc3) in dimension 2 (DIM 2). Table 4 presents the numerical values.

contrib	Dim,1	Dim,2	Dim,3
Gc1	33,271	40,0431	71,3426
Gc2	37,2695	3,0057	2,1468
Gc3	29,4595	56,9512	26,5106

Table 4. Contribution of the groups of variables in three vector space dimensions

As we can tell from the values presented in Table 5, the group Gc1 (the judges' evaluation) is the best descriptor of the vector space as indicted by the MFA value (MFA = 1,81). The Lg coefficient measures the similarity of the groups of variables.

Lg	Gc1	Gc2	Gc3	MFA
Gc1	2,1716	0,6385	1,2705	1,8135
Gc2	0,6385	1,0019	0,4406	0,9248
Gc3	1,2705	0,4406	2,0763	1,6832
MFA	1,8135	0,9248	1,6832	1,9649

Table 5. Lg coefficient and MFA correlation value



As we can tell from the values presented in Table 6 the Gc3 (VPAS) and Gc1 (ExpressionEvaluator) are strongly correlated  $RV=0,60$ . We can also tell that Gc1 is closer to the other two groups  $RV=0,88$  according to the coefficient given by Escofier & Pagès (2008).

RV	Gc1	Gc2	Gc3	MFA
Gc1	1	0,4329	0,5983	0,8779
Gc2	0,4329	1	0,3055	0,6591
Gc3	0,5983	0,3055	1	0,8333
MFA	0,8779	0,6591	0,8333	1

Table 6. Degrees of correlation among the groups of variables

The strong correlations among the groups of variables indicate that these are capable of representing the phenomena under study.

#### 4. Conclusion

The results of the perceptual evaluation task indicated that the judges were able to differentiate between professional and non-professional speakers. This differentiation can be thought of in terms of vocal dynamic characteristics such as pitch and the use of pauses in prosodic phrasing rather than in terms of vocal quality setting characteristics.

In relation to the judgements of pleasantness of voice quality, the attribution was guided by characteristics of vocal quality settings independently of speakers being professionals or not: whispery voice and expanded pharynx settings were related to pleasantness and non-professional and professional speakers with those voice quality settings were better evaluated than the professional speakers whose voice profiles comprised harsh voice settings. This can be explained in terms of sound symbolism since sound meaning analogies may have played an impressive role.

Concerning the results of the acoustic analysis the acoustic measures *sampquartisf0* and *mednf0* were found to be significant ( $p < 0,05$ ). These two measures indicate that the varying fundamental frequency and the speed of its variation were relevant to explain the data. Professional speakers more than non-professional speakers tended to vary fundamental frequency more often and more rapidly assigning prominence to certain words.

The analysis also showed that the indexical attributes of the two groups of speakers were identified by the acoustic parameters and the perceptual evaluations. These were found to be strongly correlated. The application of the VPAS combined with the semantic differential scale questionnaire and the acoustic analysis provided by the Expression Evaluator script was found to be useful.

## References

- ARAGÃO, A.N., T.E. COUTO, Z.A. CAMARGO, M.A.R. SANTOS & A.C.C. GAMA (2014) "Análise da qualidade vocal antes e após o uso profissional e social da voz", *Audiology - Communication Research*, 19, 209-214.
- BARBOSA, P.A. (2009) "Detecting changes in speech expressiveness in participants of a radio program", *Proceedings of Interspeech 2009 - Speech and Intelligence*. Londres: Causal Productions, 2155-2158.
- BONFIM, M.F., Z.A. CAMARGO, L.P. FERREIRA & S. MADUREIRA (2007) "Qualidade Vocal e Formantes das Vogais de Falantes Adultos da Cidade de João Pessoa", *Revista CEFAC*, 9, 99-109.
- CAMARGO, Z., S. MADUREIRA & D.H. TSUJI (2003) "Analysis of disphonic voices based on the interpretation of acoustic, physiological and perceptual data", in *Proceedings of the 6th International Seminar on Speech Production*, Sidney, 1, 31-36.
- CAMARGO, Z.A. & S.MADUREIRA (2009) "Dimensões perceptivas das alterações de qualidade vocal e suas correlações aos planos da acústica e da fisiologia", *DELTA. Documentação de Estudos em Linguística Teórica e Aplicada*, PUCSP, 25(2), 285-317.
- CAMARGO, Z.A., S. MADUREIRA, A. PESSOA & L.C. RUSILO (2012) "Voice quality and gender: some insights on correlation between perceptual and acoustic dimensions", in *Proceedings of the International Conference on Speech Prosody*, Shanghai: Tongji University Press, 1, 115-118.

- CAMARGO, Z., S. MADUREIRA & J.R. SCHMITZ (2013) "Qualidade vocal e produções de fala em três línguas: um estudo de caso", *Revista Intercâmbio*, 27, 110-140.
- ESCOFIER, B. & J. PAGÈS (2008) *Analyses Factorielles Simples et Multiples*, Dunod: Paris, 29- 30.
- FONTES, M.A.S. (2014) "Gestualidade vocal e visual, expressão de emoções e comunicação falada", Tese de Doutorado inédita. Pontifícia Universidade Católica de São Paulo.
- FONTES, M.A.S. & S. MADUREIRA (2015) "Gestural prosody and the expression of emotions: a perceptual and acoustic experiment", The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK: the University of Glasgow <Paper number 0390 retrieved from <http://www.icphs2015.info/pdfs/Papers/ICPHS0390.pdf>>
- FRENCH, P., P. FOULKES, P. HARRISON, V. HUGHES, E. SAN SEGUNDO & L. STEVENS (2015) "The vocal tract as a biometric: output measures, interrelationships and efficacy", in The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow <Paper number 0817 retrieved from <http://www.icphs2015.info/pdfs/Papers/ICPHS0817.pdf>>
- HUSSON, F., J. JOSSE, S. LÊ & J. MAZET (2013) *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R. R package version 1.25*, New York: CRC Press, Taylor and Francis Group.
- LAVER, J. (1980) *The phonetic description of voice quality*, Cambridge: Cambridge University Press.
- LAVER, J. (1991) *The gift of speech*, Edinburgh, Scotland: Edinburgh University Press.
- LAVER, J. (1994) *Principles of Phonetics*, Cambridge: Cambridge University Press.
- LAVER, J. (2000) "Phonetic evaluation of voice quality", in R.D. Kent & M.J. Ball (eds.), *Voice quality measurement*, San Diego: Singular Thomson Learning, 37-48.
- LAVER, J. & J. MACKENZIE-BECK (2001) "Unifying principles in the description of voice, posture and gesture", in C. Cavé, I. Guaïtella & S. Santi (eds), *Oralité et Gestualité: Interactions et Comportements Multimodaux dans la Communication (Proceedings of ORAGE 2001)*, L'Harmattan: Paris, 15-24
- LAVER, J. & J. MACKENZIE-BECK (2007) *Vocal Profile Analysis Scheme – VPAS*, Edinburgh: Queen Margareth University College- QMUC, Speech Science Research Centre.
- 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*, Mahwah: Lawrence Erlbrum, 285-322.

- MADUREIRA, S. (2008) "Reciting a sonnet: production strategies and perceptual effects", in *Proceedings of the Speech Prosody 2008 Conference*, São Paulo: Editora RG, 1, 697-700.
- MADUREIRA, S., Z.A. CAMARGO, (2010) "Exploring sound symbolism in the investigation of speech expressivity", *Proceedings of ISCA, Tutorial and Research Workshop on Experimental Linguistics*, Athens, Greece, 105-108.
- MADUREIRA, S. (2011) "The Investigation of Speech Expressivity", in H. Mello, A. Panunzi & T. Raso (eds.), *Illocution, modality, attitude, information patterning and speech annotation*. Firenze: Firenze University Press, 1, 101-118.
- MARWICK, H., J. MACKENZIE-BECK, J. LAVER & C. TREVARTHEN (1984) "Voice quality as an expressive system in mother-to-infant communication: a case study", *Work in Progress*, University of Edinburgh, Department of Linguistics Work in Progress, 17, 85-97.
- RUSILO, L.C., Z.A. CAMARGO & S. MADUREIRA (2011) "The validity of some acoustic measures to predict voice quality settings: trends between acoustic and perceptual correlates of voice quality", in *Proceedings of the 4TH ISCA Tutorial and Research Workshop on Experimental Linguistics*, Athens: University of Athens, 115-118.
- SCHAEFFLER, F., J. MACKENZIE-BECK & S. JANNETTS (2015) "Phonation stabilization time as an indicator of voice disorder", The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK: the University of Glasgow <Paper number 0331 retrieved from <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0331.pdf>>
- VAN BEZZOIJEN, R. (1984) *The characteristics and recognizability of vocal expression of emotions*, Dordrecht: Foris.