

The phonetic description of the voice of teachers from public schools in São Paulo

Descrição fonética da voz de professoras da rede municipal de São Paulo

Descripción Fonética de la Voz de las profesoras de la red municipal de São Paulo

Ana Carolina Nascimento Fernandes¹ 

Zuleica Camargo² 

Susana Pimentel Pinto Giannini³ 

Luis Carlos Rusilo⁴ 

Léslie Piccolotto Ferreira² 

Abstract

Introduction: This study focuses on the phonetic assessment of voice quality using speech samples from public school teachers in São Paulo. **Purpose:** To describe vocal quality settings (VQS) in articulatory, phonatory, muscular tension aspects, and vocal dynamics elements (VDE) in speech samples from teachers in São Paulo. **Method:** A cross-sectional study with a retrospective collection of 60 speech samples analyzed using the Vocal Profile Analysis Scheme (BP-VPAS). All samples were edited with PRAAT software, and those with a signal-to-noise ratio above 2 were selected to minimize external noise interference. Researchers extracted 40-second semi-spontaneous speech excerpts and grouped them into the Experiment MFC 3.S script for auditory-perceptual tasks. One-third of the samples were repeated for

¹ Universidade de Brasília, Brasília, DF, Brazil.

² Pontifícia Universidade Católica de São Paulo - PUC-SP, São Paulo, SP, Brazil.

³ Fonoaudióloga, São Paulo, SP, Brazil.

⁴ Universidade Federal de Alenas - UNIFAL-MG, Alenas, MG, Brazil.

Funding: The researcher Ana Carolina Nascimento Fernandes received a CNPq scholarship for the elaboration of this work.

Authors' contributions:

ACNF: participated in the study design, analysis, discussion, writing and critical review of the manuscript and the submission of the article; data collection.

ZAC, LPF: participated in the study design, analysis, discussion, writing and critical review of the manuscript.

SPPG: participated in data collection, providing research material and correction of the article and dissertation.

LCR: participated in the statistical analysis and guidance of the analyzes for the elaboration of the discussion and statistical results.

Email for correspondence: fga.carolinafernandes@gmail.com

Received: 31/10/2024

Accepted: 03/03/2025

internal consistency, with the script playing them randomly. **Results:** Perceptual judgments showed high reliability (alpha coefficients: 0.777 and 0.814). Discriminant analysis identified vocal quality settings segregating groups based on vocal dynamics. Key variables included Rough Voice (71.7%), Creaky Voice (48.1%), Vocal Tract Hyperfunction (35.3%), Whisper (32.2%), and Pharyngeal Constriction (43.3%). These findings supported group classifications: Non-Case (no dysphonia), Case I (perceptual alteration only), Case II (laryngeal alteration only), and Case III (combined perceptual and laryngeal alterations). **Conclusion:** VQS related to muscular tension and supralaryngeal phonatory mechanisms were essential in distinguishing groups, enhancing the understanding of dysphonia among teachers.

Keywords: Voice; Auditory Perception; Voice Quality.

Resumo

Introdução: Este estudo investiga a avaliação fonética da qualidade vocal utilizando uma base de dados de amostras de fala de professores de escolas públicas de São Paulo. **Objetivo:** Descrever os ajustes de qualidade vocal (AQV) nos domínios articulatórios, fonatórios e de tensão muscular, bem como os elementos da dinâmica vocal (EDV) em amostras de fala de professores da rede municipal de ensino de São Paulo. **Método:** Estudo transversal com coleta retrospectiva e descritiva baseada na análise perceptivo-auditiva de 60 amostras de fala, utilizando o esquema Vocal Profile Analysis Scheme para o português brasileiro (VPAS-PB), extraído de uma base de dados de estudo de caso-controle. Todas as amostras foram editadas no software PRAAT. Foi calculada a relação sinal-ruído, selecionando amostras com valores acima de 2 para evitar interferências de ruído externo. Os pesquisadores extraíram trechos de 40 segundos de fala semi-espontânea de cada gravação e os agruparam no *script* Experiment MFC 3.S para tarefas de percepção da qualidade vocal. Um terço das 60 amostras foi repetido para avaliar a consistência interna, sendo reproduzidas de forma aleatória. **Resultados:** Os julgamentos perceptivos demonstraram alta confiabilidade (coeficientes alfa de 0,777 e 0,814). A análise discriminante revelou ajustes de qualidade vocal que segregaram grupos de falantes com base na dinâmica vocal. As variáveis mais influentes incluíram voz áspera (71,7%), voz crepitante (48,1%), hiperfunção do trato vocal (35,3%), escape de ar (32,2%) e constrição faríngea (43,3%). **Conclusão:** Os ajustes vocais relacionados à tensão muscular e aos mecanismos fonatórios supralaríngeos foram cruciais para distinguir grupos, avançando na compreensão da disfonia em professores.

Palavras-chave: Voz; Percepção Auditiva; Qualidade da Voz.

Resumen

Introducción: Este estudio se centra en la evaluación fonética de la calidad vocal utilizando muestras de habla de profesores de escuelas públicas de São Paulo. **Propósito:** Describir los ajustes de calidad vocal (AQV) en elementos articulatorios, fonatorios, de tensión muscular y elementos de dinámica vocal (EDV) en las muestras de profesores del sistema municipal de São Paulo. **Método:** Estudio transversal con recopilación retrospectiva de 60 muestras de habla analizadas mediante el esquema Vocal Profile Analysis Scheme (BP-VPAS). Las muestras se editaron con el software PRAAT, seleccionando aquellas con una relación señal-ruído superior a 2 para minimizar el ruido externo. Se extrajeron fragmentos de 40 segundos de habla semi-espontánea y se agruparon en el *script* Experiment MFC 3.S para tareas perceptivo-auditivas. Un tercio de las muestras se repitió para evaluar la consistencia interna, reproduciéndolas de forma aleatoria. **Resultados:** Los juicios perceptivos mostraron alta fiabilidad (coeficientes alfa: 0.777 y 0.814). El análisis discriminante identificó ajustes de calidad vocal que segregaron grupos según la dinámica vocal. Las variables clave incluyeron Voz Áspera (71.7%), Voz Crujiente (48.1%), Hiperfunción del Tracto Vocal (35.3%), Escape de Aire (32.2%) y Constricción Faríngea (43.3%). Estas variables respaldaron la clasificación de grupos: No Caso (sin disfonía), Caso I (alteración perceptiva), Caso II (alteración laríngea) y Caso III (alteraciones combinadas). **Conclusión:** Los ajustes vocales relacionados con la tensión muscular y los mecanismos fonatorios supralaríngeos fueron esenciales para distinguir grupos, contribuyendo a la comprensión de la disfonía en profesores.

Palabras clave: Voz; Percepción Auditiva; Calidad de la Voz.

Introduction

As a characteristic closely related to the need for interaction between human beings, the voice is the result of a sound produced by the vibration of the vocal folds and modified by the articulation. The process includes a functional harmony of five anatomical systems: pulmonary, phonation (larynx); resonance (labial, oral, nasal, pharyngeal, laryngeal and paranasal sinuses cavities); articulatory (tongue, lips, jaw, palate and teeth) and nervous systems. To better understand the characteristics of the voice, it is necessary to understand the anatomophysiological, psychological and sociocultural influences that manifest in the voice.¹

Therefore, in order for voice perception and production to be described, it is essential to evaluate the concepts of voice and speech from a perspective that integrates the perceptual-auditory and visual, physiological and acoustic levels. In this sense, given that the focus on the speech signal and its alterations is a subject of investigation in the fields of Speech-Language Pathology, Linguistics, Medicine, Psychology, Education, Engineering, among other areas of knowledge, this study highlights the importance of to understanding voice and speech from an integrated perspective. In a multidisciplinary context, Phonetic Sciences approach the interface of the complex relationships between perception and production of voice and speech, understanding the voice (or vocal signal) as an integral part of the speech signal.²

From a phonetic perspective, voice quality is defined by recurrent muscle tendencies (extrinsic factors) that are adopted during speech and that also undergo changes as a function of the physical structure of the vocal tract (intrinsic factors). Thus, these factors collaborate for the performance of the communicative and informative functions of voice quality. Translated as voice quality setting (VQS), the unit of analysis (setting) allows composing the Vocal Profile Analysis Scheme based on several audible combinations of characteristics of respiratory, laryngeal (phonatory), supralaryngeal (articulatory) and muscle tension activities.³

Since teachers are the group of professionals most affected by voice disorders, it is necessary to evaluate the entire work environment regarding environmental and organizational risk factors, collecting behavioral data and relevant habits that may impact on vocal production. In the field of

Speech-Language Pathology, there is an understanding that the voice must be evaluated from a multidimensional perspective that includes clinical, occupational and epidemiological history in the diagnostic process of a voice disorder. In this context, voice assessment should include clinical and laryngeal medical assessment, in addition to complementary exams and speech-language pathology assessment.^{4,5}

Therefore, the association of more than one investigation method is recommended in order to conduct a complete assessment of voice quality. Thus, it is possible to gather a more detailed knowledge of VQS in the phonatory and articulatory aspects and of muscle tension, in addition to aspects of vocal dynamics. From a perceptual point of view, the assessment can be performed through the application of the Vocal Profile Analysis Scheme-VPAS⁶ scheme, using the VPAS instrument adapted to Brazilian Portuguese.⁷

The researchers of this study decided to analyze voice quality from a descriptive perspective with phonetic motivation, aiming to understand how some voice and speech patterns are manifested in the daily life of a group of teachers from the public network of São Paulo.⁸ The proposal to adopt the BP-VPAS⁶ scheme is based on the search for expanding the description of vocal quality settings, in order to verify if there would be articulatory and muscle tension settings, in addition to vocal dynamics elements, compatible with certain laryngeal adaptations (presence of lesion or alteration irritative, functional or structural in the vocal fold) previously detected in a group of teachers, evaluated as having or not having voice and laryngeal disorders. This study aimed to describe the VQS in the articulatory, phonatory and muscle tension aspects, as well as vocal dynamics elements (VDE) in speech samples of teachers from the municipal public school system in São Paulo.

Material and methods

This is a descriptive cross-sectional study, with retrospective data collection. The study used the database from a case-control study carried out with 431 teachers from the municipal school system in São Paulo.⁸ All subjects who participated in this study signed an Informed Consent Form (ICF) of the Hospital do Servidor Público Municipal de São Paulo - HSPM (protocol No. 101/07), where data

collection took place. The project of this study was submitted to the Research Ethics Committee of the Pontifícia Universidade Católica de São Paulo (PUC-SP) together with a statement by the author, which explained the transfer of the database, having been approved under the opinion no. 330/2010.

The assessment of all study subjects included audio recordings of semi-spontaneous speech samples and data from the otorhinolaryngological evaluation, which included a telarlaryngoscopy examination. At the end of the otorhinolaryngological consultation, the subjects were qualified by the physician as: with laryngeal alteration (in the presence of lesion or irritative, functional or structural alteration in the vocal fold) and without laryngeal alteration (without lesion or alteration visible on examination).

The auditory-perceptual analysis of the voices in the database was performed by three speech-language pathologists, experts in voice, who reached a consensus using the GRBASI scale.⁹

Thus, four groups of samples were initially defined at the end of the evaluations: Non-case (107 samples): absence or slight alteration of vocal quality in speech-language and auditory-perceptual assessment and in otorhinolaryngological assessment; Case I (63 samples): presence of alteration in the auditory-perceptual assessment of the voice and no alteration in the otorhinolaryngological assessment; Case II (51 samples): no alteration in the perceptual-auditory assessment of the voice and presence of alteration in the otorhinolaryngological assessment; and, finally, Case III (135 samples): presence of alteration in the auditory-perceptual and otorhinolaryngological assessment. After this classification, the author⁸ defined the Case III group as the Case Group and the Non-case group as the Control Group.

Criteria for sample selection

The semi-spontaneous speech samples were edited using the open access software PRAAT¹⁰, using the measures of signal-to-noise ratio extracted from the intensity curve (from minimum, average and maximum values of signal and noise amplitude in the silence stretch). In this context, the author selected the recordings of values higher than 2.¹¹ Therefore, the author reached the number of 75 speech samples in the Non-case group; 50 in the Case I group; 40 in the Case II group and 97 samples in the Case III group. Subsequently, 15 speech

samples were drawn from each of the four groups, totaling 60. Then, these samples were edited into 40-second¹¹ snippets which were redistributed as follows: Group 1: Non-case: (S1 to S15); Group 2: Case I: (S16 to S30); Group 3: Case II: (S31 to S45); Group 4: Case III: (S46 to S60).

After that, the samples were grouped in the Experiment MFC 3.S script (for perception experiments), which is applicable to the PRAAT software¹⁰. To perform this script, one third of the 60 speech samples (totaling 20) were randomly selected and repeated to ensure the analysis of the reliability of the perceptual judgments. The script grouped the speech samples and redistributed them randomly, which allowed the judges not to notice the sequence of the presentation, or which samples had been presented in repetition, thus ensuring the judges did not have access to their own responses at the time of the repetitions. The 80 speech samples were recorded on a CD together with the perception¹² script to be run in the PRAAT software.¹⁰ The CD was provided along with a report with the reproduction of 80 BP-VPAS⁷ instrument and guidance to the judges on the correct way to carry out the auditory-perceptual analysis of voice quality. The judges had a Bachelor's Degree in Speech-Language Pathology, and a Specialization voice, with more than three years of experience in the application of the script used for analysis.

After the auditory-perceptual judgment was performed, the judges' assessments were entered into an Excel spreadsheet and submitted to descriptive statistical analysis and multivariate analysis.¹³ In a first step, the auditory-perceptual judgments were submitted to reliability and internal consistency analysis. Descriptive statistical analysis procedures allowed the definition of each group (Non-Case, Case I, Case II, and Case III) in relation to vocal quality settings and vocal dynamics elements.

The multivariate statistical analysis used the Discriminant Analysis and Agglomerative Hierarchical Cluster Analysis procedures, using the XL-Stat software (Addinsoft). In this sense, 54 evaluation parameters were used, which included 36 parameters representing the vocal quality settings and 18 parameters associated with vocal dynamics elements, which are part of the BP-VPAS script.

Results

From the perspective of perceptual judgments, the multivariate analysis showed high reliability and internal consistency (a 0.777 alpha coefficient for the analysis of degrees of VQS and a 0.814 alpha coefficient for analysis of absence or presence of VQS).

The results of the auditory-perceptual assess-

ment of the samples from the investigated database were presented based on the general profile of the group in terms of the combination of VQS (in the articulatory, phonatory and tension subgroups) and VDE.

In this context, Chart 1 shows information regarding the presence or absence of VQS and VDE in each group; while Chart 2 shows the degrees of manifestation of VQS and VDE

Chart 1. Distribution of speech samples in the four groups, according of voice quality settings (VQS) and voice dynamics elements (VDE)

VOCAL QUALITY SETTINGS (VQS) AND VOICE DYNAMICS ELEMENTS (VDE)	GROUP 1	GROUP 2	GROUP 3	GROUP 4
Rounded Lips				
Spread Lips				
Labiodentalization				
Minimized Lip Extension				
Extensive Lip Extension				
Closed Jaw				
Opened Jaw				
Protracted Jaw				
Minimized Jaw Extension				
Extensive Jaw Extension				
Advanced Tongue Tip				
Retracted Tongue Tip				
Advanced Tongue Body				
Retracted Tongue Body				
Raised Tongue Body				
Lowered Tongue Body				
Minimized Tongue Body Extension				
Extensive Tongue Body Extension				
Pharyngeal Constriction				
Pharyngeal Expansion				
Audible Nasal Air Leak				
Nasal				
Denasal				
Raised Larynx				
Lowered Larynx				
Vocal Tract Hyperfunction				
Vocal Tract Hypofunction				
Laryngeal Hyperfunction				
Laryngeal Hypofunction				
Modal				
Falsetto				
Creak				
Creaky Voice				
Breathy Voice				



VOCAL QUALITY SETTINGS (VQS) AND VOICE DYNAMICS ELEMENTS (VDE)	GROUP 1	GROUP 2	GROUP 3	GROUP 4
Rough Voice				
Extensive Habitual Pitch				
Minimized Habitual Pitch				
Minimized Pitch Range				
Extensive Pitch Range				
Minimized Pitch Variability				
Extensive Pitch Variability				
Extensive Habitual Loudness				
Minimized Habitual Loudness				
Minimized Loudness Range				
Extensive Loudness Range				
Minimized Loudness Variability				
Extensive Loudness Variability				
Interrupted Continuity				
Fast Speech Rate				
SlowSpeech Rate				
Adequate Respiratory Support				
Inadequate Respiratory Support				

Description: ☐ PRESENT ☐ ABSENT

Chart 2. Distribution of speech samples in the four groups, according to the degrees of manifestation of voice quality settings (VQS) and voice dynamics elements (VDE)

VOICE QUALITY SETTINGS (VQS) AND VOICE DYNAMICS ELEMENTS (VDE)	GROUP 1	GROUP 2	GROUP 3	GROUP 4
Rounded Lips				
Stretched Lips				
Labiodentalization				
Decreased Lip Extension				
Increased Lip Extension				
Closed Jaw				
Opened Jaw				
Protracted Jaw				
Decreased Jaw Extension				
Increased Jaw Extension				
Advanced Tongue Tip				
Recessed Tongue Tip				
Advanced Tongue Body				
Recessed Tongue Body				
Elevated Tongue Body				
Lowered Tongue Body				
Decreased Tongue Body Extension				
Increased Tongue Body Extension				
Pharyngeal Constriction				
Pharyngeal Expansion				
Audible Nasal Air Leak				
Nasal				
Denasal				
Elevated Larynx				



VOICE QUALITY SETTINGS (VQS) AND VOICE DYNAMICS ELEMENTS (VDE)	GROUP 1	GROUP 2	GROUP 3	GROUP 4
Lowered Larynx				
Vocal Tract Hyperfunction				
Vocal Tract Hypofunction				
Laryngeal Hyperfunction				
Laryngeal Hypofunction				
Modal				
Falsetto				
Crackling				
Crackling Voice				
Breathy Voice				
Rough Voice				
Increased Habitual Pitch				
Decreased Habitual Pitch				
Decreased Pitch Range				
Increased Pitch Range				
Decreased Pitch Variability				
Increased Pitch Variability				
Increased Habitual Loudness				
Decreased Habitual Loudness				
Decreased Loudness Range				
Increased Loudness Range				
Decreased Loudness Variability				
Increased Loudness Variability				
Interrupted Continuity				
Rapid Speech Rate				
Slow Speech Rate				
Adequate Respiratory Support				
Inadequate Respiratory Support				

Description: ☐ Absent ☐ 1st and 2nd Degree ☐ 1st to 3rd Degree ☐ 2nd and 3rd Degree ☐ 1st to 5th Degree

The discriminant analysis of VQS the Table 1 showed a segregation of groups of speakers at values higher than those of VDE in the Table 2.

Centroid charts of discriminant analysis allow estimating the degrees of segregation of groups based on VQS (Figure 1) and VDE (Figure 2).

Table 1. Confusion matrix for the estimation sample of groups of speakers based on voice quality settings (VQS)– BP-VPAS

From \ to	Group 1	Group 2	Group 3	Group 4	Total	% of correct
Group 1	12	1	2	0	15	80.00%
Group 2	1	11	1	2	15	73.33%
Group 3	2	3	10	0	15	66.67%
Group 4	1	2	0	12	15	80.00%
Total	16	17	13	14	60	75.00%

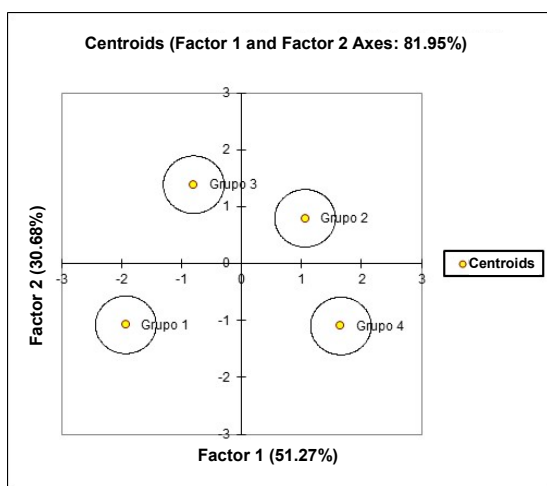


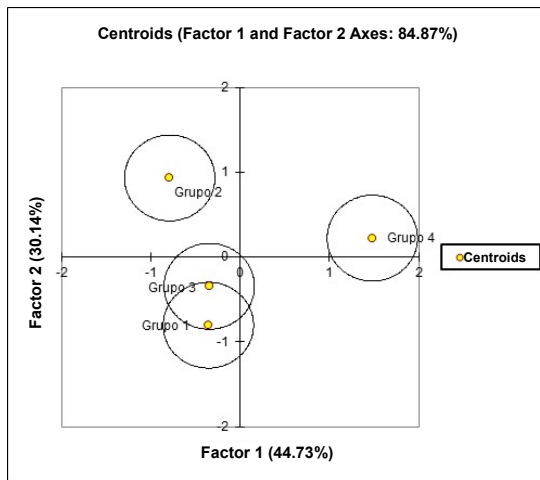
Figure 1. Centroid chart of the discriminant analysis of estimation of groups of speakers based on voice quality settings (VQS) using the BP-VPAS

The discriminant analysis showed that the perceptual judgments of the VQS segregated the samples of the four studied groups by 75%. The two main factors accounted for 81.95% (Figure 1). In descending order of correlation and percentage, the most influential variables in the segregation for each factor referred to: Factor 1:

Rough Voice (71.7%), Creaky Voice (48.1%), Vocal Tract Hyperfunction (35.3%), whisper (32.2%), Falsetto (33.9%) and raised Tongue Body (30.3%); Factor 2: Pharyngeal Constriction (43.3%), spread Lips (37.2%), retracted Tongue Body (29.4%), protruded Jaw (24.2%) and Closed Jaw (21.6%).

Table 2. Confusion matrix for the estimation sample of groups of speakers based on voice dynamics elements (VDE) – BP-VPAS

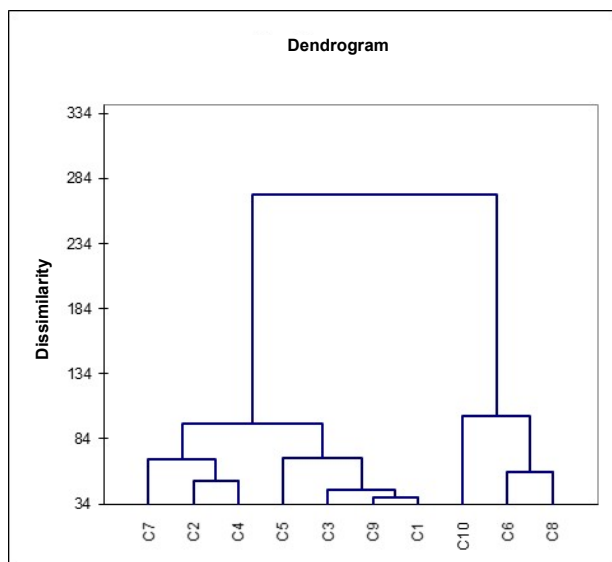
\ a	Group 1	Group 2	Group 3	Group 4	Total	% of correct
Group 1	9	1	4	1	15	60.00%
Group 2	4	8	2	1	15	53.33%
Group 3	3	2	8	2	15	53.33%
Group 4	1	2	2	10	15	66.67%
Total	17	13	16	14	60	58.33%

**Figure 2.** Centroid chart of the discriminant analysis of estimation of groups of speakers based on voice dynamics elements (VDE) using the BP-VPAS

The discriminant analysis showed that the perceptual judgments of the VDE partially segregated the samples of the 4 studied groups and at a lower total level than the VQS: 58.33%. The two main factors accounted for 84.87%. In descending order of correlation and percentage, the most influential variables in the segregation for each factor referred to: Factor 1: Decreased Habitual Pitch

(68.3%), Decreased Habitual Loudness (58.4%), and Decreased Pitch Variability (45.4%); Factor 2: Inadequate Respiratory Support (51.4%), Increased Speech Rate (40.98%), and Increased Pitch Variability (33.5%).

The hierarchical cluster analysis data allowed estimating the groupings of VQS (Figure 3) and VDE (Figure 4).

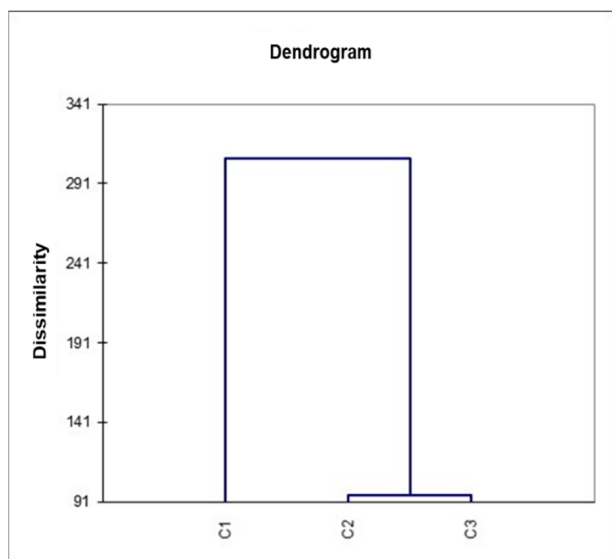


Description: C = class

Figure 3. Dendrogram of cluster hierarchical agglomerative analysis of perceptual judgments of voice quality settings (VQS)

The vocal VQS were grouped into 10 classes that were divided as follows: Class 1 (large number of articulatory and tension VQS not specified in subsequent classes); Class 2 (spread Lips); Class 3 (Labiodentalization, Pharyngeal and Nasal Constriction, creaky Voice and whisper); Class 4

(Decreased Jaw Extension); Class 5 (Advanced Tongue Tip and Body); Class 6 (raised Larynx); Class 7 (Vocal Tract Hyperfunction); Class 8 (Laryngeal Hyperfunction); Class 9 (Breathy Voice); and Class 10 (Rough Voice).



Description: C = class

Figure 4. Dendrogram of cluster hierarchical agglomerative analysis of perceptual judgments of voice dynamics elements (VDE)

The EDVs were grouped into three classes, configured as follows: Class 1 habitual pitch and loudness (increased and decreased), and increased pitch and loudness extension and variability, elocution rate (increased and decreased); Class 2 (decreased pitch range and variability) and Class 3 (decreased loudness extension and variability and inadequate respiratory support).

Discussion

A voice can be understood as normal when it is audible over a wide range of acoustic settings, even with relatively high levels of ambient noise. It should be noted that it is essential to consider the aspects of sex and age so that the speaker can fulfill the linguistic and paralinguistic functions during the speech. As there are several aspects involved in voice production, the discussion about healthy voice must start from the point of view that each human voice is unique, depending on anatomical, physiological, psychological, cultural, sociolinguistic and behavioral characteristics.¹⁴

In this context, the data analysis in the discussion of this study aimed to correlate sound characteristics in the voice that were manifested from the interaction of organic and phonetic factors. As only a few adults speak with all neutral settings, and a completely neutral type of vocal quality, it is crucial to understand the particularities of manifestations of vocal quality settings, their combinations and respective degrees of manifestation in order to obtain a broader understanding of voice clinical conditions, particularly in the group of teachers.

The database of speech samples used in this study included semi-spontaneous productions by female teachers from municipal schools. Therefore, this study did not intend to propose generalizations regarding the vocal findings, but rather to raise questions about vocal production and verify probable similarities of this group with other groups previously researched. The findings showed that the perceptual description of VQS, with phonetic origin, allowed to segregate the four groups originally defined (Table 1) and their particularities in terms of presence/absence of vocal quality alteration of a perceptual nature and presence/absence of laryngeal injury.

The samples of the four groups, respectively adopted in a previous study as control and case groups, were segregated in a proportion of 80%.

The combinations of some settings were relevant for this level of segregation (Figure 1), such as combinations of rough voice, creaky voice, vocal tract hyperfunction, whisper, falsetto and raised tongue body; in addition to the combination of spread lips settings, pharyngeal constriction, retracted tongue body, and protruded and closed jaw. This information shows that combinations of VQS of a phonatory, articulatory and muscle tension nature characterize and particularize the four groups of speakers outlined.

In this field, it should be noted that the combination of the whisper and rough voice VQS characterizes the quality known on other voice rating scales as 'hoarseness'. In this sense, the rough voice may indicate the presence of irregular vibration due to mucosal rigidity, lack of projection and may be characterized by system rigidity, laryngeal and pharyngeal tension and tension in articulatory structures.¹⁴ The whisper and rough voice VQS were the most found VQS in the speech samples of this study (68.3% and 80%, respectively). This can be explained by the study population composed only of teachers, since hoarseness is the most common aspect in this professional category, which may indicate probable voice alterations.^{17,18}

In isolation, the vocal dynamics elements showed a lower power of segregation of the groups compared to the vocal quality settings (Table 2). Figure 2 shows slight overlap between groups 1 and 3, with greater segregation of groups 2 and 4.

When evaluating the nature of the perceptual events detected, there are organic and phonetic characteristics that can cause similar effects on vocal quality and be perceived in the same way. Thus, articulatory effects can often be used to show the equivalence of configuration between situations of phonetic uses and (individual) organic characteristics of the vocal apparatus. Therefore, it is necessary to understand the nature of the interaction between the factors of phonetic uses and the (organic) nature of each speaker's vocal apparatus in order to understand the concept of phonetic analysis of voice quality.¹⁵

The pitch aspect of VDE is presented as a modifier of settings in the tract, since the system is organized so that there is a lengthening or shortening of the vocal folds, such as a faster vibration or an increase in the tension of the system, in order to produce a lower or higher sound. The pitch range is limited when there is a limitation in this



production.¹⁶ In this sense, this study found a co-occurrence between some VQS, such as whisper, rough voice and the decreased pitch range VDE, which reinforces the interdependence of these factors. Unlike what happens in the reduction of pitch range, the aspect of increased pitch range was found in the researched group to be associated with the VQS known as falsetto. This fact can be explained by the literature^{16,19,20}, which suggests greater skill in the phonatory system, when compared to more severe records. Higher pitch or falsetto production requires a thin configuration of vocal folds with minimal contact surface, in which phonation has an aspect of lightness in phonation and less tension in articulatory structures.

In this context, the data analysis in the discussion of this study aimed to correlate sound characteristics in the voice that were manifested from the interaction of organic and phonetic factors and allowed to expand the description of the groups of teachers in question.

In addition, VQS and VDE must be understood based on principles that govern their relationship^{15,21}, that is, the principle of interdependence, which shows that one adjustment interferes with the production of another. This interdependence can be found as an adjustment facilitating or impacting the execution of another. This principle is closely related to the interdependent physiological functioning of muscle mobilizations in the vocal tract. The second principle is the principle of compatibility, where one adjustment excludes the execution of another by antagonism.

The agglomerative hierarchical cluster analysis of perceptual judgments of voice quality (Figure 3) allows us to understand their main clusters, which refer to theoretical precepts of the phonetic model (compatibility and interdependence)¹⁵ and reinforce the multidimensional characteristic of voice quality. These aspects show that the combinations of simple VQS generate composite settings that allow expanding the understanding of certain populations and their vocal manifestations.

The laryngeal hyperfunction, rough voice and elevated larynx settings formed a grouping of classes, which concentrated the aspects of laryngeal tension and aperiodicity. There was also another grouping of classes that concentrated articulatory mobilizations and vocal tract tension, such as: stretched lips, decreased jaw extension and vocal tract hyperfunction. Mostly, the largest grouping of

settings included articulatory settings, with greater segregation of those related to labiodentalization, pharyngeal constriction, advanced tongue tip and body, in addition to some phonatory VQS, such as breathy and creaky voice.

As voice assessment is considered a measure that involves particular aspects of an evaluator, this aspect raises some questions. Some measures are taken to minimize the impressionistic issues of this assessment, such as proposals for protocols and scripts with scales that can measure vocal quality aspects, making them more descriptive.²¹ This explains the development of some studies based on the VPAS instrument.^{2,22,23,24}

The applicability of the VPAS script was investigated with the aim of describing the vocal quality of the Liverpool accent (differentiation of British accents), for example. In the aforementioned study, the author concluded that there was great variation in voice quality between different British accents when evaluating the vocal quality of four subjects was evaluated through the script. In addition, the author found that, although the VPAS was commonly used to assess altered voices, the tool could be used to describe voices outside the clinical context, and this could help both phoneticians and speech-language pathologists in the task of tracing the vocal profile of a speaker or even a population.²²

Some works have been developed to validate the adaptation of the VPAS script to Brazilian Portuguese^{3; 11; 12}. A study carried out with 60 speakers²⁵ considered without voice alteration, found a correlation between vocal quality and acoustic parameters and it was possible to divide the speakers into groups based on auditory-perceptual assessment using the BP-VPAS script. In the aforementioned study, the speakers were grouped in four different ways, considering the most present VQS during their speeches. The most prominent settings were modal voice (70.1%), laryngeal hyperfunction associated with rough voice (19.4%), lowered larynx associated with pharyngeal expansion and creaky voice (10.1%), and lips associated with advanced tongue tip, vocal tract hyperfunction, creaky voice, breathy voice and pharyngeal constriction (5.1%).

In turn, the most observed adjustment in this study was rough voice (80%) associated with the whisper adjustment (68.3%). The laryngeal hyperfunction was found in 36.6% of the analyzed samples and was also found to co-occur with the rough voice VQS. Modal voice was reported in



5% of speakers. The pharyngeal expansion and lowered larynx settings were also associated in this study. The differences between the settings found in the two studies can be explained by the different populations evaluated: the other study³ included speakers from different professions, who were considered without vocal changes or complaints, while this study evaluated teachers who, in terms of voice classification, are considered a heterogeneous group, and many of them had voice disorders.

After carrying out this study, it is possible to reflect that, in order to understand what happens to the speaker at the moment of speech production, the speech-language pathologist must be able to make associations between physiology and what is being heard (perception). Evaluation in the speech-language pathology clinic is essential to outline the therapeutic plan.

The BP-VPAS⁷ script contributes to speech-language pathology practice by allowing the description of vocal production phenomena to be expanded beyond the level of laryngeal activity, bringing together manifestations of the vocal tract and events of the nature of muscle tension. Even though they alone did not show the group segregation power as the vocal quality settings, the vocal dynamics elements showed particularities in their combinations and groupings that also make it possible to map the speech and voice events of the speakers and collaborate for a broader understanding of the clinical condition.

The agglomerative hierarchical cluster analysis of the perceptual judgments of VDE (Figure 4) showed classes grouped around habitual pitch and loudness, with increased loudness and speech rate range and variability. There was also another grouping of classes that combined pitch range and variability with inadequate respiratory support.

The breathy voice is characterized in the literature by inefficient coaptation of the vocal folds and there is a compensatory effort to try to phonate without audible noise, which can generate tension.^{16,19} In addition, the speech pattern in the breathy voice has a characteristic of an attempt at high intensity emission by the speaker to mask the breathiness of the voice. This description may explain the co-occurrence of breathy voice, tract hyperfunction and increased loudness variability in the analyzed groups.

VDE, such as inadequate respiratory support, lowered habitual pitch, increased speech rate,

increased habitual loudness, and increased pitch variability, show interdependence in relation to laryngeal hyperfunction tension adjustment. The speaker tends to exert excessive effort during phonation.¹⁴ In turn, inadequate respiratory support adjustment is the most reported aspect in the literature in studies carried out with teachers, followed by strong intensity. In the case of this category of voice professionals, it is important to emphasize the general muscle tension aspect. The studies mentioned above refer to laryngeal tension-fatigue syndrome or musculoskeletal tension syndrome.

This speech pattern with tension is common in the population of teachers due to the vocal behavior that is related to the so-called Work-Related Voice Disorder (WRVD), which consists of effort when speaking and stress after long periods of work.^{28,29}

After carrying out this study, it is possible to reflect that, in order to understand what happens to the speaker at the moment of speech production, the speech-language pathologist must be able to make associations between physiology and what is being heard (perception). This evaluation in the speech-language pathology clinic is essential to outline the therapeutic plan.

Thus, the BP-VPAS⁷ script contributes to the speech-language pathology clinic by allowing the description of what happens in the vocal tract and guiding clinical reasoning in order to understand what happens with a given speaker. In addition, this tool can be used in therapy to offer a visual model to patients, showing how each adjustment is performed and making them perceive, from the imitation of the settings, how their muscles are used during speech and production. By perceiving what happens in their body, patients can understand their own vocal behavior, which promotes the search for a vocal pattern that is more adequate to their physical structure and, also, adequate to their profession.

This study allowed us to verify the presence of VQS that are more compatible with certain laryngeal adaptations, also showing combinations of articulatory, phonatory and tension settings that indicate voice alterations in the researched group.

The auditory-perceptual analysis with phonetic origin showed that the phonatory elements and general muscle tension were the most evident settings in the analyzed group, composed of teachers.

In addition, the description of the VQS led to the understanding of the group's vocal pattern and allowed for an overview of the association

between the settings, which suggests attention when a speaker is evaluated and presents a certain non-neutral pattern. The study also made it possible to correlate VQS with VDE.

Conclusion

The VQS found in the perceptual analysis made it possible to segregate different groups of teachers as to the absence (Non-Case) or presence of the clinical condition of dysphonia (Case I = perceptual alteration only; Case II = laryngeal alteration only; and Case III = combination of perceptual and laryngeal alteration). Influential settings in segregation were related to the phonatory and muscle tension nature of the supralaryngeal vocal tract, followed by supralaryngeal VQS. The VDE had a slightly lower influence on the segregation of the mentioned groups, and the pitch and loudness VDE collaborated in the segregation of the groups, followed by the speech rate and respiratory support. These descriptions allowed the detailed analysis of perceptual voice findings in the previously searched database, reinforcing the previously outlined division of groups (Non-Case, and Cases I, II and III). The most influential variables in segregation referred to the following settings: rough voice, creaky voice, vocal tract hyperfunction, whisper, falsetto and raised tongue body; pharyngeal constriction, spread lips, retracted tongue body, and protruded and closed jaw. Based on phonetics, the perceptual findings made it possible to detail influential factors in the composition of clinical cases in the Voice specialty.

References

1. Cataldo E, Lucero JC, Sampaio R, Nicolato L. Comparison of some mechanical models of larynx in the synthesis of voiced sounds. *J Braz Soc Mech Sci Eng.* 2006; 28(4): 461–6. Available from: <https://doi.org/10.1590/S1678-58782006000400011>
2. Camargo ZA, Madureira S. Dimensões perceptivas das alterações de qualidade vocal e suas correlações aos planos da acústica e da fisiologia. *DELTA.* 2009; 25(2): 285–317. Available from: <https://doi.org/10.1590/S0102-44502009000200004>
3. Mackenzie-Beck J. Organic variation of vocal apparatus. In: Hardcastle WJ, Laver J, editors. *The handbook of phonetic sciences.* Oxford: Blackwell Publishers; 2005. p. 311–50.
4. Ferreira LP, Luciano P, Akutsu Lara M. Condições de produção vocal de vendedores de móveis e eletrodomésticos: correlação entre questões de saúde, hábitos e sintomas vocais. *Rev CEFAC.* 2008; 10(4): 528–35. Available from: <https://doi.org/10.1590/S1516-18462008000400013>
5. Cielo CA, Ribeiro VV, Hoffmann CF. Sintomas vocais de futuros profissionais da voz. *Rev CEFAC.* 2015;17(1): 34–43. Available from: <https://doi.org/10.1590/1982-0216201517013>
6. Laver J. *The phonetic description of voice quality.* Cambridge: Cambridge University Press; 1980.
7. Camargo ZA, Madureira S. Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP-VPAS). In: *Proceedings of the Fourth Conference on Speech Prosody*; 2008; Campinas. Campinas: Capes, Fapesp, CNPq; 2008. p. 57–60.
8. Giannini SP. *Distúrbio de voz relacionado ao trabalho docente: um estudo caso-controle [tese].* São Paulo: Faculdade de Saúde Pública; 2010.
9. Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier-Buchman L, Millet B. Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements. *Rev Laryngol Otol Rhinol (Bord).* 1996;117(3): 219–24.
10. Boersma P, Weenink D. Praat: a system for doing phonetics by computer. *Glott Int.* 2001; 5(9/10): 341–7.
11. Griffiths PR, Hsueh JA. *Fourier Transform Infrared Spectrometry.* New York: Wiley; 1986.
12. Barbosa PA. Detecting changes in speech expressiveness in participants of a radio program. In: *Proceedings of Interspeech 2009*; Brighton, UK. Brighton: International Speech Communication Association; 2009. p. 2155–8.
13. Anderson TA. *An introduction to multivariate statistical analysis.* New York: Wiley; 1984.
14. Coelho AC, Brasolotto AG, Bahmad F Jr. Development and validation of the protocol for the evaluation of voice in patients with hearing impairment (PEV-SHI). *Braz J Otorhinolaryngol.* 2019; 85(4):1–7. Available from: <https://doi.org/10.1016/j.bjorl.2019.05.007>
15. Laver J. Phonetic evaluation of voice quality. In: Kent RD, Ball MJ, editors. *Voice Quality Measurement.* San Diego: Singular Thomson Learning; 2000. p. 37–48.
16. Elsing JH, Harris JG. States of the glottis: an articulatory phonetic model based on laryngoscopic observations. In: Hardcastle WJ, Laver J, editors. *The handbook of phonetic sciences.* Oxford: Blackwell Publishers; 2005. p. 301–10.
17. Klasmeyer G, Sendlmeier WF. Voice and emotional states. In: Kent RD, Ball MJ, editors. *Voice Quality Measurement.* San Diego: Singular Thomson Learning; 2000. p. 330–58.
18. Bernardi JM, de Barros LN, Assunção LS, de Oliveira RS, Gambirásio YF, Medved DM, et al. Effect of the Finnish tube on the voice of a deaf musician: a case report. *J Voice.* 2019; 1:1–1.
19. Kreiman J, Gerratt B. Measuring vocal quality. In: Kent RD, Ball MJ, editors. *Voice Quality Measurement.* San Diego: Singular Thomson Learning; 2000. p. 73–102.
20. Coadou M. Voice quality and variation: a pilot study of the Liverpool accent. In: *Proceedings of Speech Prosody 2006*; Dresden, Germany. International Speech Communication Association; 2006. p. 1–5.
21. Camargo ZA, Madureira S. Avaliação vocal sob a perspectiva fonética: investigação preliminar. *Distúrbios da Comunicação.* 2008; 20(1): 77–96.

22. Camargo ZA, Rusilo LC, Madureira S. Evaluating speech samples designed for the Voice Profile Analysis Scheme for Brazilian Portuguese. In: Proceedings of the Fourth ISCA Tutorial and Research Workshop on Experimental Linguistics; 2011; Paris. Paris: University of Athens; 2011. p. 115–8.
23. Fuess VRL, Lorenz MC. Disfonia em professores do ensino municipal: prevalência e fatores de risco. *Rev Bras Otorrinolaringol.* 2003; 69(6): 807–12.
24. Rusilo LC, Camargo ZA, Madureira S. The validity of some acoustic measures to predict voice quality settings. In: Proceedings of the Fourth ISCA Tutorial and Research Workshop on Experimental Linguistics; 2011; Paris. Paris: University of Athens; 2011. p. 55–8.
25. Machado PG, Hammes MH, Cielo CA, Rodrigues AL. Os hábitos posturais e o comportamento vocal de profissionais de educação física na modalidade de hidroginástica. *Rev CEFAC.* 2011;13(2): 299–313.
26. Magalhães LPF, Alves CDA, Vieira RC. Comparação de análises de amostras de falantes por meio de diferentes protocolos perceptivo-auditivos. *Cad Saúde Pública.* 2019;18(3): 347–54. Available from: <https://periodicos.ufba.br/index.php/cmbio/article/view/34185>



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.