

Original articles

Ultrasonography applied to the description of voice quality settings in adult speakers of Brazilian Portuguese

Tamar Vieira de Jesus¹<https://orcid.org/0000-0002-8495-3057>Aline Neves Pessoa Almeida²<https://orcid.org/0000-0001-5892-3957>Zuleica Camargo³<https://orcid.org/0000-0001-8725-2419>

¹ Pontifícia Universidade Católica de São Paulo - PUC-SP, São Paulo, São Paulo, Brasil.

² Universidade Federal do Espírito Santo – UFES, Departamento de Fonoaudiologia, Vitória, Espírito Santo, Brasil.

³ Pontifícia Universidade Católica de São Paulo - PUC-SP, Faculdade de Filosofia, Comunicação, Letras e Artes (FAFCLA), Departamento de Ciências da Linguagem e Filosofia; Pontifícia Universidade Católica de São Paulo - PUC-SP, Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem - PEPG em LAEL; Fundação Armando Álvares Penteado – FAAP, São Paulo, São Paulo, Brasil.

Work conducted at the Pontifical Catholic University of São Paulo / PUC-SP, Integrated Laboratory of Acoustic Analysis and Cognition - LIAAC and Graduate Program in Applied Linguistics and Language Studies Program - LAEL.

Research support source: Edital Universal CNPq. PROCESS: 459178/2014-8.

Conflict of interests: Nonexistent



Received on: June 23, 2021

Accepted on: October 12, 2021

Corresponding address:

Prof^a Dr^a Aline Neves Pessoa Almeida
Avenida Marechal Campos, nº 1468 –
Maruípe
CEP: 29047-105 – Vitória, Espírito Santo,
Brasil
E-mail: aline.n.almeida@ufes.br

ABSTRACT

Purpose: to relate ultrasound images with auditory-perceptual data on vocal quality settings in adult speakers of Brazilian Portuguese.

Methods: the corpus consisted of speech samples (sentences contained in the instructional material of the *Vocal Profile Analysis Scheme – VPAS-PB*) from seven adult subjects of both genders, recorded simultaneously by acoustic and ultrasonographic means. Data analysis was based on auditory-perceptual judgments of vocal quality and ultrasound images generated by the AAA software.

Results: vocal quality settings related to the position of the tongue body and the extension of the tongue and jaw found correspondences to the contours of ultrasound images of the tongue in selected key segments (oral vowels), especially those with greater degrees of manifestation.

Conclusion: there were correspondences between vocal quality settings detected in the perceptual sphere and their respective tongue body and jaw ultrasound images.

Keywords: Voice Quality; Tongue; Ultrasonics; Phonetics; Speech Perception; Adults

INTRODUCTION

Studies on speech have benefited from constant technological development¹, which allows for the broadening of the scope of investigations in the circuit between speech production and perception. With due support from the Phonetic Sciences, such technological resources can play a role in the perceptual, acoustic, and physiological aspects. Physiologically, information regarding the respiratory, phonatory, articulatory, and resonant mechanisms are particularly relevant. In this set of possible information, technological resources aimed at generating images of the vocal tract, using non-invasive techniques such as ultrasonography, are invaluable. This resource has been gradually consolidating itself as an instrument applicable to both assessment and therapy in the field of speech sound disorders²⁻⁴.

Its contributions focus on plotting tongue contour images in the sagittal and coronal planes, resulting in several lines of investigation required for modeling of such images and in the approaches to ultrasound data using quantitative and qualitative means for the functions of speech and swallowing⁵⁻¹⁰. In the possible applications of ultrasound to the speech language pathology field, however, vocal quality approaches are scarce¹⁰.

In the field of voice specialty, the speech therapist works in contexts of rehabilitation of a wide range of disorders,¹¹⁻¹⁶ as well as in coaching voice and singing professionals. Therefore, detailed knowledge of the mechanisms developed by the various segments of the vocal tract, and their intrinsic plasticity, is essential.

The phonetic model for the description of vocal quality, proposed by Laver (1980), includes descriptions of vocal quality settings according to characteristics of the anatomy of the vocal tract (intrinsic factor) and mobilizations performed during speech (extrinsic factor)¹⁷. Vocal quality is defined as the combination of supralaryngeal (articulatory), laryngeal (phonatory), and muscle tension settings that characterize the subject's speech most of the time, in a recurrent way.

Laver e col (1981)¹⁷ goal was to match the articulatory descriptions of the segments (units) of speech (vowels and consonants) and apply them to the field of vocal quality, analyzing how speakers adjust their vocal tracts during the production of the speech chain⁴.

For the phonetic analysis of vocal quality, Laver's (1980)¹⁸ framework involved a unit of analysis: the setting, which refers to a long-lasting muscle tendency, that is, a recurrent muscular tendency. In order to

define the various possible vocal quality settings, the author proposes a reference setting, which he calls the neutral setting,^{17,18} defined by lips neither projected nor stretched; larynx neither lowered nor raised; constancy of the diameter of the supralaryngeal tract cavities; anterior oral articulation with the tongue body neither advanced nor retracted; absence of constrictions in the supralaryngeal vocal tract; jaw position neither markedly open nor closed; nasality present only in segments where it is phonologically relevant; periodic vibration of vocal folds, with moderate longitudinal and adductor tension, and efficient air current without noise; and, finally, intermediate tension of the vocal tract musculature.

The neutral setting, fundamentally, is characterized by the absence of marked action of the lips, jaw, tongue, or pharynx during the vocal production process and muscle stability so as not to change the extension of the vocal tract.

As a result, there are no drastic variations in the total extension of the vocal tract or in the areas of its resonator cavities. In other words, the vocal tract is neither excessively elongated nor shortened, nor does it present cavities in states of constriction or expansion.

Furthermore, muscle tension is maintained at an intermediate level (without extremes of hyper or hypofunction) and the vocal folds vibrate regularly (with periodic mobility and absence of noise generated by irregularities). It is noteworthy that the reference setting – neutral – does not signal a state of rest or a proposal for a parameter of vocal normalcy.

To describe the variety of vocal settings, Laver (1980)¹⁸ therefore characterized the possible variations from the neutral setting in supralaryngeal (articulatory), laryngeal (phonatory)¹¹⁻¹², and muscle tension levels in the laryngeal and supralaryngeal planes.

From the variations of the neutral setting, several possible mobilizations, contemplated in the Vocal Profile Analysis Scheme (VPAS)¹⁷ manifested themselves in chosen speech samples (Figures 1 and 2). This proposal inspired an adaptation to Brazilian Portuguese, VPAS-PB¹⁹ used to identify different clinical demands for speech and voice^{2,12,20-22}.

Commonly, the scales proposed for the auditory-perceptual assessment of vocal quality focus on descriptions of vocal manifestations at the laryngeal activity (phonatory)^{11,12} level, with resonance and articulation events (primarily supralaryngeal) still posing challenges to the establishment of a consensus on parameters and descriptors supported by acoustic

and physiological parameters. Given the demand for expanding instruments to assess vocal quality in the area of supralaryngeal settings, this article focuses on the proposal of the VPAS¹⁸ instrument, precisely because of the possibility of its application to samples of speakers with and without voice alterations^{21,22}.

This approach allows for an analysis using elements related to features rarely systematized in terms of clinical assessment²⁰⁻²², since some resonance and articulation descriptors are still impressionistic.

In this vocal quality analysis plan, imaging tools can be applied to dissect the dynamics of the articulators in the oral and pharyngeal regions, leading to associations between the analysis of the visible and audible spheres of speech and voice^{1,5-10}.

Imaging resources can be utilized in the training of speech therapists, as well as in clinical settings for assessment, therapy, and coaching of speech and voice, in the form of integrated approaches of perceptual, articulatory, and acoustic analyses¹¹. Ultrasonography, in particular, has the advantages of not exposing the speaker to radiation and providing non-invasive access to the vocal tract. Furthermore, it can be used in clinical settings and its increasing portability translates into cost savings that can facilitate its broader use.

Some approaches have been identified as promising in the literature for the clinical monitoring of speech^{2-10,13-16,23-25} and in the process of learning a second language, such as English²⁶. Studies refer to the advantages of ultrasound as being a non-invasive, safe, low-cost method, free from exposure of speakers to radiation and without the generation of a magnetic field, with the possibility of analyzing the movement of the tongue during speech production and swallowing in real time, and finally as allowing for the generation of images in different planes: sagittal, coronal, and oblique^{2-10,13,14-16,23-28}.

Such proposals emphasize that the incorporation of real-time feedback using images of the vocal tract contributes to stable and lasting speech results, both in the clinical and language learning contexts.

This intersection makes it possible to characterize the speech production process by observing the arrangement of articulators in the process of its emission and their correspondences to auditory sensations¹¹. This possibility affords clinicians more accurate descriptions of speech and voice manifestations, complementing the descriptions of sometimes

unpleasant auditory sensations caused by a speech and/or voice disorder^{11-13,20-22,29}.

Studies involving ultrasonography have been gaining ground in different areas of speech language pathology, demonstrating significant benefits in relation to other instruments^{1,27,28}. Different modalities of vocal tract imaging analysis are essential to speech therapy because they can reveal the dynamics of the speech organs in the production of sounds and help in the investigation of how the resulting sounds are perceived, leading to better diagnostic and prognostic accuracy^{27,28}.

In addition, ultrasonography makes it possible to observe the tongue during the production of some speech sounds that are characterized by a period of silence (such as during the occlusion of consonants like bilabials), leading to a broad characterization of speech and voice.

In describing supralaryngeal aspects of vocal quality, the use of ultrasonography is promising. Supported by the difficulty in characterizing the aspects of resonance and articulation when describing voices, this study leverages the possibilities offered by ultrasonography to highlight contributions in the perceptive detailing of supralaryngeal vocal tract settings and enable future training of professionals with greater ability to make perceptive judgments of vocal quality with phonetic motivation, as well as to expand the incorporation of ultrasonography to the practice of speech therapy. It, thus, aims to relate ultrasound images of tongue contours, in the sagittal plane, to auditory-perceptual data from vocal quality settings in adult speakers of Brazilian Portuguese.

METHODS

This is an exploratory study, in which all participants were selected using convenience sampling and signed the Informed Consent Form, according to the norms of the Research Ethics Committee of the Pontifical Catholic University of São Paulo, PUC-SP, Brazil, to which this project was submitted and approved under CAEE nº 16400319.7.0000.5482. Participants were invited to participate in the data collection session through university message groups. The inclusion criteria were: being an adult speaker of Brazilian Portuguese, with a minimum level of high school education. The exclusion criteria referred to the absence of hearing disorders, previous treatments for major disorders of the head and neck, such as tumor resection, radiotherapy or chemotherapy, or major orthognathic surgery.

The data collection procedures involved simultaneous recordings of audio and image samples (ultrasonography) of the vocal tract (tongue contour) of seven adult Brazilian Portuguese speakers, of both genders – one male and six females – aged between 20 and 59 years old. Six were university students (Speech Language Pathology and Language majors) and one Information Technology graduate. One speaker worked as an amateur radio broadcaster and was in speech therapy due to the demand for the use of a professional voice. One speaker was trained as an actress and worked for dubbing studios.

Using the Concept M6 Digital Ultrasonic Diagnostic Imaging System™ Ultrasound System, 120-degree angulation transducer, and an adjustable aluminum headset for mounting the transducer in the submandibular region, the data collection sessions took place in the Radio Laboratory of the home institution. After collections, a synchronization correction factor of 0.034 seconds was applied to compensate for the the capture system's temporal difference in the recording of image and audio stimuli, using the Advanced Assistant Articulate software (AAA from Articulate Instruments™).

Speakers were instructed to remain seated in front of a computer containing the AAA software, whose interface presented the speech tasks in a sequenced manner, through a visual stimulus for reading and an audible alarm indicating the beginning of the recording of the audio stimulus.

Once positioned, subjects who wore glasses or other head accessories were instructed to remove them for better adjustment of the headset containing the ultrasound probe (transducer) and the headset-style microphone (Laboratory Quality Behringer™ ECM-8000 omnidirectional).

The aluminum head support was used for head stabilization, with the intention of fixating the equipment's transducer, avoiding interference from head and jaw movements. The support was adjusted according to the length and width of the speaker's face and skull, taking into account speakers' responses regarding comfort and attachment points.

The recording included the steps of swallowing water, briefly reporting on the city where the speaker was born, and reading (5 randomized repetitions) 4 sentences proposed in the instructional material for the use of the VPAS-PB,^{11,19} referred to as key sentences since they contain key segments for detecting vocal quality settings^{11,17-19}.

1. *“O objeto de estudos da fonética é esta complexa, variável e poderosa face sonora da linguagem: a fala.”*

2. *“O garoto tirou muitas fotografias do tucano, da coruja, do pombo e do jaburu”.*

3. *“Soube que a Casa dos Bispos é visitada por turistas todos os dias e que o roteiro de visita dura cerca de duas horas para ser percorrido”.*

4. *“Detesto ir à casa dele, pois fica do outro lado da cidade e o acesso é difícil”.*⁽¹⁾

Data analysis procedures were based on three stages: perceptive-auditory analysis of each voice, analysis of the body of the tongue shapes from the ultrasound images, and matching of perceptual and acoustic data through a qualitative approach.

The first stage of data analysis began simultaneously with the collection of speech samples (audio and images), when a speech therapist examiner, experienced in the use of the VPAS-PB¹⁹, performed the auditory-perceptual assessment of vocal quality based on the reading of the key sentences contained in the VPAS-PB¹⁹ instructional materials. Such examiner remained in the collection room without, however, having access to the real-time images generated during collection, monitored by two other researchers.

In the second stage of data analysis, the image samples referring to the studied subjects were identified and sorted according to the vocal quality settings (detected through the auditory-perceptual analysis) of the tip and body of the tongue (tip advanced or retracted, body advanced, retracted, or lowered). In the collected samples, no settings representative of high tongue body were detected. As a reference parameter, images were selected from a female speaker who presented neutral settings of tongue tip and body.

The collected images were then analyzed using the AAA software in two approaches. The first was the identification and labeling of vowels that were considered key segments, precisely because they reveal the greatest prominence of vocal quality settings in the speech sequence, being considered susceptible segments to vocal quality settings in the Laver's (1980)^{17,18} framework. For each sentence, 3 repetitions

(1) Sentences used in Brazilian Portuguese:

1. “The object of study in phonetics is this complex, variable, and powerful sound facet of language: speech.”
2. “The boy took many photographs of the toucan, the owl, the pigeon, and the jabiru.”
3. “I learned that the House of Bishops is visited by tourists every day and that the tour takes about two hours.”
4. “I hate going to his house because it's on the other side of town and access is difficult.”

were selected that represented the highest degree of manifestation of vocal quality settings referring to the tip and body of the tongue.

Then, using the AAA software, sagittal plane plots of the tongue contour (splines) in a stationary period of the vowel were generated, representing the moment of support of the articulators' posture during emission. Such periods coincide, in a broader way, with the intermediary step of the emission, comprised between the transitions to adjacent sounds (anterior and posterior consonants). The vowels [i], [e], [o], and [u] (as closed and semi-closed vowels) were considered key segments for the detection of lower tongue body settings; [a] (central and open vowel) was considered a key segment for the detection of advanced or retracted and raised tongue body settings; [i] and [e] (front vowels) for retracted tongue body settings; and [u] and [o] (back vowels) for advanced tongue body settings. Therefore, the vowels adopted as the key segments in the following words in sentences 1 to 4 were analyzed: est[u]dos, f[a]la, gar[o]to, cor[u]ja, tur[i]stas, d[e]le, and cid[a]de.

In the third stage of data analysis, from each of the 3 repetitions of the vowels selected in the listed words, a tongue contour plot (spline) was generated. Such plots were considered in association with the vocal quality settings detected in the first stage of analysis. Qualitatively, tongue contour plots were considered with a view to matching neutral and non-neutral vocal quality settings in the supralaryngeal region. This stage

of analysis also included, in the case of non-neutral vocal quality settings, the comparative analysis of the seven oral vowels in the set of productions of each speaker, according to the expectations of articulation of the vocalic sounds of Brazilian Portuguese³⁰. This stage of intra-speaker analysis was followed by inter-speaker analysis, noting that the information from image plots is considered relative to the set of movements revealed by the various vowels by the same speaker, since there are inherent anthropomorphic issues involved including physical type, size and measurement of cervical circumference, head size, and facial type.

Finally, the findings were analyzed in light of the theoretical precepts of the phonetic model for the description of voice quality.

RESULTS

Results are presented below based on correspondences between perceptual findings of vocal quality and ultrasonographic images of tongue contours.

Perceptual judgments of vocal quality

Auditory-perceptual judgments of the subjects' speech samples are summarized in Figures 1 and 2.

Chart 1 contains the vocal quality settings identified in the participants.

Chart 2 contains vocal quality settings specific to the tip of the tongue, as observed in the seven participants.

Chart 1. Vocal quality settings detected in the auditory-perceptual analysis of the studied subjects' speech samples

| Vocal quality settings | | Grade 1 | Grade 2 | Grade 3 |
|------------------------|------------------------|---------|---------|---------|
| | | N= 63 | N= 262 | N= 102 |
| Lips | Rounded | | | |
| | Spread | | | |
| | Labiodentalization | | | |
| | Reduced extension | | | |
| | Overextension | | | |
| Jaw | Closed | | | |
| | Open | | | |
| | Protracted | | | |
| | Reduced extension | | | |
| | Overextension | | | |
| Tongue tip/blade | Advanced | | | |
| | Retracted | | | |
| Tongue body | Advanced | | | |
| | Retracted | | | |
| | Elevated | | | |
| | Lowered | | | |
| | Reduced extension | | | |
| | Overextension | | | |
| Pharynx | Constriction | | | |
| | Expansion | | | |
| Velopharynx | Audible nasal emission | | | |
| | Nasal | | | |
| | Denasal | | | |
| Larynx height | Elevated | | | |
| | Lowered | | | |
| Vocal tract tension | Hyperfunction | | | |
| | Hypofunction | | | |
| Larynx tension | Hyperfunction | | | |
| | Hypofunction | | | |
| Mode of phonation | Modal | | | |
| | Falsetto | | | |
| | Creak / vocal fry | | | |
| | Creaky voice | | | |
| Laryngeal friction | Air escape | | | |
| | Breathy voice | | | |
| Irregularity | Harsh voice | | | |

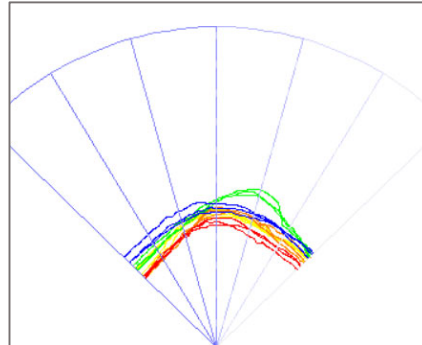
Chart 2. Identification of vocal quality settings of tongue tip and tongue body displayed by each speaker in the studied group (grade identified based on most occurrences)

| Identification | Vocal quality settings – tongue tip and body |
|----------------|---|
| Speaker 1 | Neutral setting of tongue tip and body. |
| Speaker 2 | Advanced tongue tip (grade 1-2) and tongue body (grade 1-2) settings. |
| Speaker 3 | Advanced tongue tip (grade 2-3) and tongue body (grade 2-3) settings. |
| Speaker 4 | Advanced tongue tip (grade 2-3) and tongue body (grade 2-3) settings. |
| Speaker 5 | Advanced tongue tip (grade 2-3) and lowered tongue body (grade 2) settings, with reduced extension. |
| Speaker 6 | Lowered tongue body (grade 2) and open jaw (grade 2-3) settings. |
| Speaker 7 | Retracted (grade 3) and lowered (grade 3) tongue body settings, with reduced extension. |

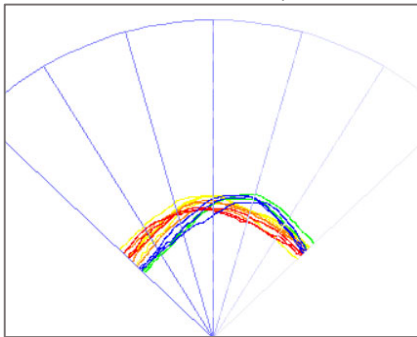
Ultrasonographic images of tongue contours

The analysis of the tongue body contours of the vowel productions of all seven speakers is in Figures

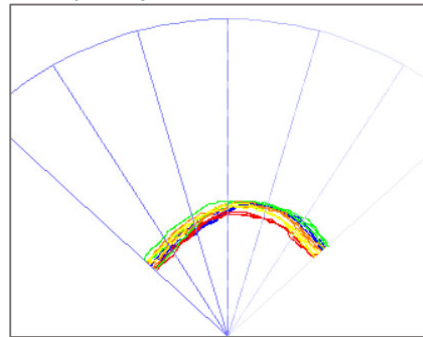
1-4, with an overall view of all vowel productions in Figure 1 and differentiated key segments in Figures 2-4.



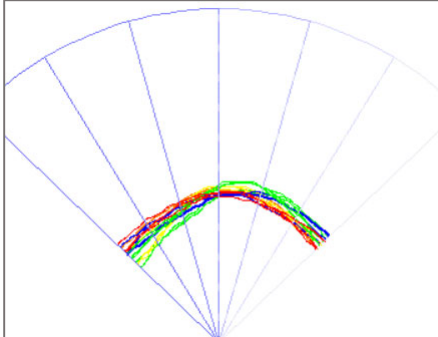
Speaker 1 – neutral tongue tip and body settings.



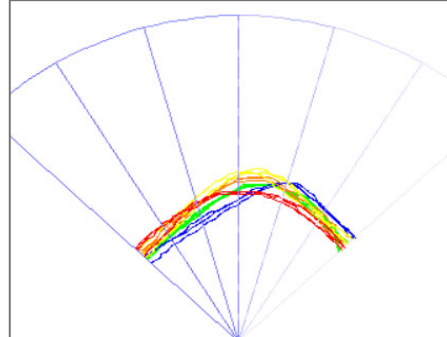
Speaker 2 – advanced tongue tip and body settings.



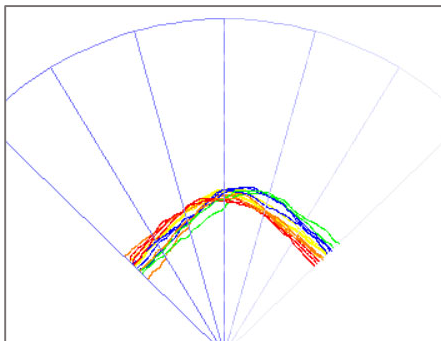
Speaker 3 – advanced tongue tip and body settings.



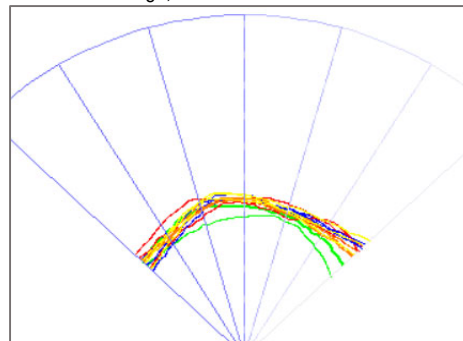
Speaker 4 – advanced tongue tip and body settings.



Speaker 5 – advanced tongue tip and lowered tongue body settings, with reduced extension.



Speaker 6 – lowered tongue tip and open jaw settings.



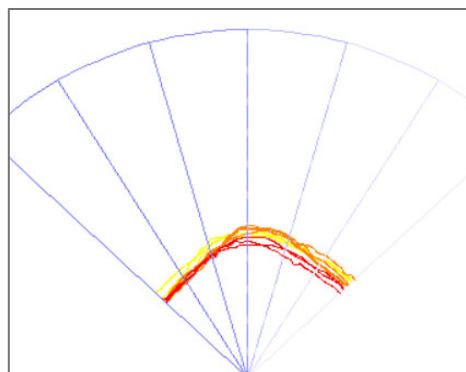
Speaker 7 – retracted and lowered tongue body settings, with reduced extension.

Captions: colors – red (vowel [a]), blue (vowel [e]), green (vowel [i]), orange (vowel [o]), yellow (vowel [u]). Orientation: Right – tongue tip (anterior); Left – pharyngeal cavity (posterior).

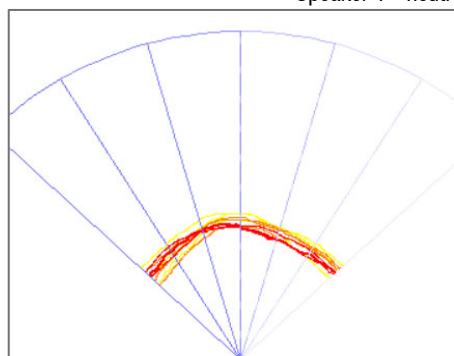
Figure 1. Tongue body splines, in sagittal view, of [a], [e], [i], [o], and [u] vowel samples of each speaker in the studied group

Based on the principle of susceptibility of segments to vocal quality settings, in Laver's (1980)^{11,22} view, Figures 2-4 show tongue contours of the key segments as related to the vocal quality settings detected in distinct tongue movement axes.

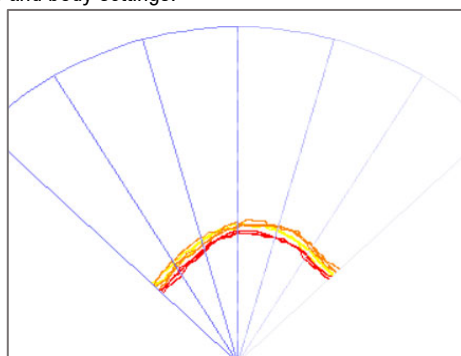
Figure 2 shows tongue contours of the productions of the low central vowel [a] and high back vowels [u] and [o] of the speaker with the neutral setting and speakers with advanced tongue tip and body settings.



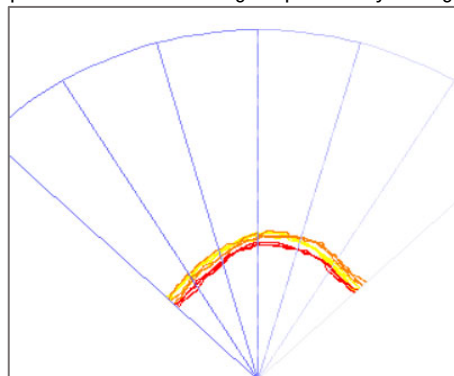
Speaker 1 - neutral tongue tip and body settings.



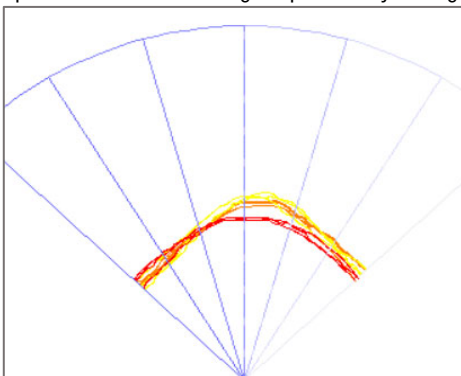
Speaker 2 – advanced tongue tip and body settings.



Speaker 3 – advanced tongue tip and body settings.



Speaker 4 – advanced tongue tip and body settings.



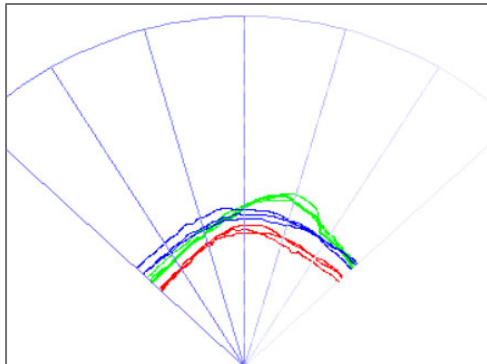
Speaker 5 – advanced tongue tip and lowered tongue body settings, with reduced extension.

Captions: colors – red (vowel [a]), orange (vowel [o]), and yellow (vowel [u]). Orientation: Right – tongue tip (anterior); Left – pharyngeal cavity (posterior).

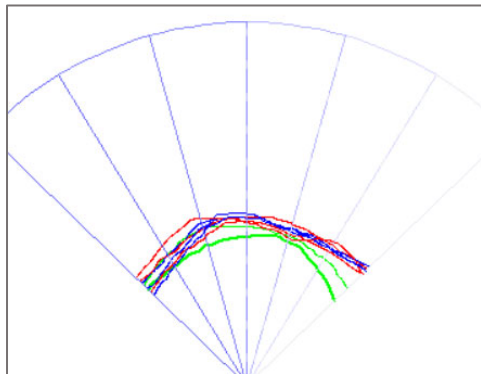
Figure 2. Tongue splines of the productions of low central vowel [a] and closed back vowels [u] and [o] of speaker with neutral setting and speakers with advanced tongue tip and body settings

Figure 3 shows tongue contours of the productions of the low central vowel [a] and high front vowels [i] and [e] of the speaker with the neutral setting and

the speaker with retracted and lowered tongue body settings.



Speaker 1 - neutral tongue tip and body settings.



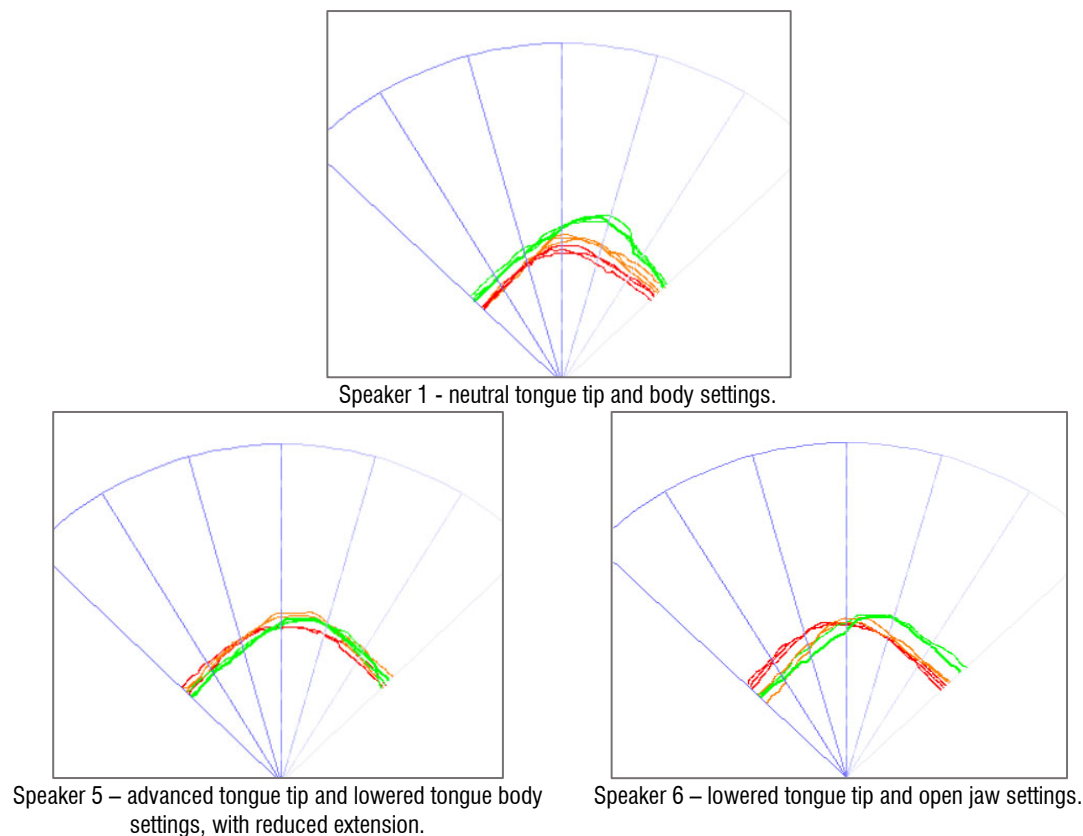
Speaker 7 – retracted and lowered tongue body settings, with reduced extension.

Captions: colors – red (vowel [a]), blue (vowel [e]), and green (vowel [i]). Orientation: Right – tongue tip (anterior); Left – pharyngeal cavity (posterior).

Figure 3. Tongue splines of the productions of vowel [a] and high front vowels [e] and [i] of speaker with neutral setting and speaker with retracted and lowered tongue body settings

Figure 4 shows tongue contours of the productions of the low central vowel [a] and high vowels, front and back [i] and [o] respectively, of the speaker with the

neutral setting and speakers with the lowered tongue body setting.



Captions: colors – red (vowel [a]), green (vowel [i]), and orange (vowel [o]). Orientation: Right – tongue tip (anterior); Left – pharyngeal cavity (posterior).

Figure 4. Tongue splines of productions of vowel [a] and high vowels [i] and [o] of speaker with neutral setting and speakers with lowered tongue body setting

DISCUSSION

Based on the auditory-perceptual analysis, we observe a greater occurrence of long-term advanced tongue body and tip settings, as per Tables 1 and 2, followed by lowered tongue body and retracted tongue body.

The analysis of intra-subject data points to specific muscle characteristics and trends (of the tongue as the central articulator) of vocal quality settings regarding more advanced or retracted or even lowered positioning, with greater or lesser extension of the tongue and jaw. Such information was collected from plots extracted from a vowel's stationary point (Figures 3 to 4).

The key segments (vowels) have characteristics that allow for better visualization of the extension of the tongue in speech production, as they present movements in the anterior, middle, and posterior regions of the oral cavity²⁷.

One of the articulatory classification criteria of vowels refers to the tongue position in the horizontal

(front, central, or back) and vertical (low, middle, or high) planes, with associated jaw positions to characterize the openings. For Brazilian Portuguese, the classification is as follows: [a] is a low, central (with tendencies to posteriorization), and open vowel; [e] is mid-high, front, and semi-closed; [i] is high, front, and closed; [o] is mid-high, back, and semi-closed; and [u] is high, back, and closed³⁰.

In this way, it is possible to match the vocal quality settings detected audibly with the judgments of the ultrasound plots of the vowels, since, in a tonic context, the tongue mobilizations of each speaker tend to highlight the setting detected by auditory-perceptual analysis.

The vocal quality settings related to the body of the tongue (lowered, advanced, and retracted), in addition to the reduction in the extension of the tongue and jaw, contained in the VPAS-PB, found correspondences to the contours of ultrasonographic images of the tongue in the chosen key-segments – oral vowels – especially those with greater degrees of manifestation, as described below.

In the case of speaker 1, representative of the neutral tongue tip and body setting, different tongue postures were detected for vowels in extreme positions of the vowel chart, such as [a], [i] and [o]^{19,30}.

It is noteworthy that the plotting of [u] was not clear and easily observable in the various emissions from different speakers, as previously observed in the literature⁵⁻¹⁰.

In settings with tendencies to tongue body advancement, back vowels were the ones that best revealed long-term muscle trends and vice-versa. The sample did not contain representative instances of high tongue body.

Speakers 2, 3, and 4 (and 5 for the tip of the tongue only), who presented advanced tongue tip and body settings, demonstrate, by the tongue contours of their vocal productions, tongue posture slanted to the right both in the full set of vowels (Figure 1) and in the set of susceptible vowels, such as the back [o] and [u] and the central [a], which, by definition, would have their settings in the back region. Of these, the clearest mobilizations occurred in speaker 5.

As for mobilizations that tended to lower the body of the tongue, the high and mid-high (closed) vowels revealed these trends, with little variation in the height axis, as observed in speakers 6 and 7.

Speaker 7 revealed settings composed of retracted and lowered tongue body, with decreased extension of the tongue and jaw. The ultrasonographic images (Figure 3) revealed the position of the retracted and lowered tongue in the emission of front high vowels such as [e] and [i] in relation to the back low vowel [a].

Finally, there were correspondences between the perceptual descriptions of the vocal quality settings and images, based on the tongue contour plots of the vocal utterances of speakers 5 and 6 (Figure 4). With advanced tongue tip and tongue body with decreased extension, lowered tongue body and open jaw settings, their articulation contours of high vowels [o] and [u] were fronted, even though these are effectively back high vowels.

The findings reinforce the principle of susceptibility of segments to vocal quality settings of the phonetic model for the description of vocal quality and point to the applicability of ultrasonography in the field of speech language pathology^{4-10,13-16,23-25,27,28}. The data presented in this study, in particular, expand this applicability to voice therapy and voice coaching. At this point, ultrasonography endorses a fine-grained look at the tongue articulator and contemplates fertile interpretations as

a complementary tool to diagnose disorders involving the supralaryngeal region and for therapeutic follow-up, through the use of real-time images, which provide visual feedback to patients, who can reconfigure their tongue mobilizations.

The aim is to continue this kind of exploration to advance the capture and reading of ultrasound image data of the tip of the tongue, as well as to explore images beyond the sagittal plane (addressed in this study), such as the coronal and oblique planes. Furthermore, we seek to expand this approach to the variety of settings characteristic of voice resonance, broadening the focus of events of the laryngopharyngeal region and incorporating procedures for extracting measures of tongue displacement¹⁰.

In addition to its academic and technological contributions, the findings of this study point to the efficiency of the associated use of ultrasound techniques for clinical processes and coaching in speech and voice, highlighting the articulators as worthy of attention in the exploration of vocal quality, mainly through ultrasound images, which can provide evidence as to the use of and relationships between the laryngeal, supralaryngeal, and general muscle tension spheres.

CONCLUSION

The vocal quality settings referring to the body of the tongue (lowered, advanced, and retracted), in addition to the reduction in the extension of the tongue and jaw, contained in the VPAS-PB, found correspondences in the contours of ultrasonographic images of the tongue in selected key segments (oral vowels), especially those with greater degrees of manifestation.

In the case of settings with tendencies to tongue body advancement, back vowels were the ones that best revealed long-term muscle tendency and vice-versa. In the case of mobilizations that tended to lower the body of the tongue, upper and mid (closed) vowels revealed such trends. The sample did not have representative instances of high tongue body.

The findings reinforce the principle of susceptibility of segments to vocal quality settings of the phonetic model of voice quality description and highlight the applicability of ultrasonography to clinical settings and voice coaching.

REFERENCES

- Lima JPH, Julio A, Daniel CG. Technological resources in the field of speech and hearing: review. In: Camargo C, editor. *Fonética Clínica - Vinte anos de LIAAC*. São Paulo: Pulso Editorial; 2016. p.167-79.
- Pereira LCK, Madureira S, Camargo Z, Lima JPH. Proposal for multi-instrumental investigation of the speech of subjects with hearing impairment: acoustic, perceptual and ultrasonographic analysis. In: Camargo C, editor. *Fonética Clínica - Vinte anos de LIAAC*. São Paulo: Pulso Editorial; 2016. p. 183-95.
- Zharkova N, Gibbon FE, Lee A. Using ultrasound tongue imaging to identify covert contrasts in children's speech. *Clinical Linguistics & Phonetics*. 2017;31(1):21-34.
- Wein B, Bockler R, Klajman S, Obrebowski IA. Ultrasonography of the tongue in the rehabilitation of speech articulation disorders. *Otolaryngol Pol*. 1991;45(2):133-40.
- Zharkova N. Using ultrasound to quantify tongue shape and movement characteristics. *The Cleft Palate-Craniofacial Journal*. 2013;50(1):76-81.
- Stone M. A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *Journal of the Acoustical Society of America*. 1981;87(5):2207-17.
- Silva LM, Vassoler AMO, Marino VCC, Berti LC. Quantitative analysis of tongue movement in 14 phonemes of Brazilian Portuguese. *CoDAS [journal on the internet]* 2017 [accessed 2021 Jun 12]; 29(4):1-8. Available at: <https://www.scielo.br/j/codas/a/NTMV6ZvNCbXDY9FGQF7wNWc/?lang=en>
- Wertzner HF, Francisco DT, Pagan-Neves LO. Tongue contour for /s/ and /ʃ/ in children with speech sound disorder. *CoDAS [journal on the internet]* 2014 [accessed 2021 Jun 12]; 26(3):248-51. Available at: <https://www.scielo.br/j/codas/a/ZFqp5XDM7F4QLpztVx9CChR/?lang=en>
- Barberena LS, Simoni SN, Souza RCS, Moraes DAO, Berti LC, Keske-Soares M. Ultrasound analysis of tongue contour for the sound [j] in adults and children. *CoDAS [journal on the internet]* 2017 [accessed 2021 Jun 12]; 29(6): [about 8 p] Available at: <https://www.scielo.br/j/codas/a/nWWQ6FYDMmxQ6WSG44SFSFx/?lang=en>
- Bressmann T, de Boer G, Marino VC de C, Fabron ELG, Berti LC. Influence of voice focus on tongue movement in speech. *Clinical Linguistics & Phonetics*. 2017;31(3):212-21.
- Camargo Z, Madureira S. Vocal evaluation from the phonetic perspective: preliminary investigation. *Disturb. Comum*. 2008;20(1):77-96.
- Freitas SV. Correlation between acoustic and perceptual assessment in the characterization of pathological voices [dissertation]. Porto (Portugal): Universidade do Porto; 2010.
- Vassoler AMO, Berti LC. Quantitative ultrasound analysis of the production of consonant clusters performed by typical and atypical children. In: 23^o Nacional Brazilian Congress e 9^o International Congress of Fonoaudiologia. 2015. Salvador. Bahia. p. 1-7.
- Barberena LS, Brasil BC, Mezzomo CL, Mota HB, Keske-Soares M. Ultrasound applicability in Speech Language Pathology and Audiology. *CoDAS [journal on the internet]* 2014 [accessed 2021 Jun 12]; 26(6):520-30. Available at: <https://www.scielo.br/j/codas/a/GsBMpMGcdD8YwxVX3dy5DxB/?lang=en>
- Rocha SG, Silva RG, Berti LC. Qualitative and quantitative ultrasound analysis of oropharyngeal swallowing. *CoDAS [journal on the internet]* 2015 [accessed 2021 Jun 12]; 27(5):437-45. Available at: <https://www.scielo.br/j/codas/a/Z7RZkK7hQrmnMYfXkKwbcxK/?lang=pt&format=pdf>
- Steidi CMS, Gonçalves BFT, Bilheri D, Brancher EC, Pasqualoto AS, Mancopes R. Application of ultrasound method in the evaluation of biomechanical swallowing – literature review. *Disturb. Comum. [journal on the internet]* 2016 [accessed 2021 Jun 12]; 28(2):219-28. Available at: <https://pesquisa.bvsalud.org/portal/resource/pt/biblio-1657>
- Laver J, Wirz S, Mackenzie J, Hiller S. A perceptual protocol for the analysis of vocal profiles. In: Edinburgh University Press. Edinburgh University Department of Linguistics, Edinburgh (United Kingdom); V. 14, 1981. p.139-55.
- Laver J. The phonetic description of voice quality. 1^a ed, Cambridge: Cambridge University Press, 1980.
- Camargo Z, Madureira S. Material instrutivo para uso do roteiro vocal profile e analysis scheme para o português brasileiro (VPAS-PB). In: Camargo C, editor. *Fonética Clínica - Vinte anos de LIAAC*. São Paulo: Pulso Editorial; 2016. p.225-35.

20. Isolan-Cury R, Scalissi N, Monte O, Cury AN, Camargo Z, Freitas AB. Qualidade vocal e acromegalia: dados do trato vocal. In: Camargo C, editor. *Fonética Clínica - Vinte anos de LIAAC*. São Paulo: Pulso Editorial; 2016.p.33-58.
21. Gomes PC, Oliveira LR, Camargo Z. Respiração oral na infância: parâmetros perceptivo-auditivos e acústicos de qualidade. In: Camargo C, editor. *Fonética Clínica - Vinte anos de LIAAC*. São Paulo: Pulso Editorial; 2016. p.121-8.
22. Freitas T, Oliveira LR, Camargo Z. Ajustes de qualidade vocal nos tipos faciais: análise perceptivo-auditiva e acústica. In: Camargo C, editor. *Fonética Clínica - Vinte anos de LIAAC*. São Paulo: Pulso Editorial; 2016.p.129-40.
23. Brum-de-Paula MR, Ferreira-Gonçalves G, editors. *A ultrassonografia e os gestos da fala*. Pedro & João Editores, São Carlos: 2021. [accessed 2021 jun 14]. Available at: <https://pedroejoaoeditores.com.br/site/a-ultrassonografia-e-os-gestos-da-fala/>.
24. Balch P, Wrench A. Towards a 3D tongue model for parameterising ultrasound data. 18th International Congress of Phonetic Sciences ICPhSc Proceedings, 2015, Glasgow, v. 1, paper 0263.
25. Correa BT, Gonçalves GF. Parâmetros em análises qualitativa e quantitativa: dados ultrassonográficos de língua. In: Brum-de-Paula MR, Ferreira-Gonçalves G, editors. *A ultrassonografia e os gestos da fala*. Pedro & João Editores, São Carlos: 2021. [Accessed 2021 ago 28]. Available at: <https://pedroejoaoeditores.com.br/site/a-ultrassonografia-e-os-gestos-da-fala/>
26. Silva AF. Contrastes na produção de vogais da língua inglesa: um estudo com o uso de técnicas de fonética acústica e ultrassonografia. *CBTecLe*. 2018;1(1):1-18.
27. Wiethan F, Ceron MI, Marchetti P, Giachchini V, Mota GB. The use of electroglottography, electromyography, spectrography and ultrasound in speech research - theoretical review. *Rev. CEFAC [journal on the internet]* 2015 [accessed 2021 jun 12]; 17(1):115-25. Available at: <https://www.scielo.br/j/rcefac/a/bn3QHbq3JR4BP378JRjrxCK/?lang=en>
28. Barberena LS, Portalete CR, Simoni SN, Prates ACM, Keske-Soares M, Mancopes R. The use of electroglottography, electromyography, spectrography and ultrasound in speech research - theoretical review. *CoDAS [journal on the internet]* 2017 [acessado em 12 de Junho de 2021]; 29(2):115-25. Available at: <https://www.scielo.br/j/rcefac/a/bn3QHbq3JR4BP378JRjrxCK/?lang=en>
29. Erickson D. Voice quality. In: Verbetes LBASS. 2021 [accessed 2021 ago 28]. Available at: <http://www.letras.ufmg.br/lbass/>
30. Barbosa PA, Albano EC. Illustrations of the IPA. *JIPA*. 2004;34(2):227-32.