Original Article

https://doi.org/10.1590/2317-6431-2022-2655en

Audiology Communication Research
ISSN 2317-6431

Auditory-perceptual and acoustic measures in women with and without vocal nodules

Medidas perceptivo-auditivas e acústicas de mulheres com e sem nódulos vocais

Késsia Cecília Fernandes Conserva[1], Fernanda Pereira França[1], Leonardo Wanderley Lopes[2]

ABSTRACT

Purpose: To compare the acoustic and auditory-perceptual measures of the voice of women with and without vocal nodules. **Methods:** Twelve women with vocal nodules (MNV group) and 12 without vocal nodules (MSNV group) participated in the study. They were submitted to the recording of their sustained /a/ vowel, in order to extract the mean of the fundamental frequency (f0), standard deviation of f0 (SD f0), jitter, shimmer, GNE, cepstral measure of CPPS, and spectral measures of differences of the first and second harmonics (H1-H2); and recording of the carrier phrases: "I say papa baixinho", "I say pipa baixinho" and "I say pupa baixinho", to extract the first (F1) and second formant (F2) of the vowel segments /a, i, u/. For auditory-perceptual assessment, the visual-analog scale (VAS) was used. **Results:** The comparative analysis between the groups shows higher values for the MNV in the parameters general degree, roughness and breathiness, and for the shimmer acoustic measure. The F1 values for the vowels /a/ and /u/, and the F2 values for the vowel /a/ were higher in the same group. **Conclusion:** According to the data observed in the investigated sample, women with nodules have more deviated voices, with the presence of roughness and breathiness, and changes in vocal tract adjustments, with a possible reduction in the amplitude of the articulators, when compared to women without vocal nodules.

Keywords: Acoustics; Voice; Voice Disorders; Voice Quality; Speech Therapy

RESUMO

Objetivo: Comparar as medidas acústicas e perceptivo-auditivas de mulheres com e sem nódulos vocais. **Métodos:** Participaram do estudo 12 mulheres com nódulos vocais (grupo MNV) e 12 sem nódulos vocais (grupo MSNV). Foram submetidas à gravação da vogal /a/ sustentada, com o objetivo de extrair a média da frequência fundamental (f0), desvio padrão de f0 (DP f0), *jitter, shimmer, Glottal Noise Excitation*, medidas relacionadas ao *Cepstral Peak Prominence-Smoothed,* medidas espectrais das diferenças do primeiro e segundo harmônico (H1-H2) e gravação das frases-veículo: "Digo papa baixinho", "Digo pipa baixinho" e "Digo pupa baixinho", para extração do primeiro (F1) e segundo formante (F2) dos segmentos vocálicos /a, i, u/. Para avaliação perceptivo-auditiva, utilizou-se a Escala Visual Analógica. **Resultados:** A análise comparativa entre os grupos evidenciou maiores valores para o grupo MNV nos parâmetros grau geral, rugosidade e soprosidade e para a medida acústica *shimmer.* Os valores de F1 para as vogais /a/ e /u/ e os valores de F2 para a vogal /a/ apresentaram-se mais elevados no mesmo grupo. **Conclusão:** mulheres com nódulos apresentaram vozes mais desviadas, com presença de rugosidade e soprosidade e modificações nos ajustes do trato vocal, com possível redução na amplitude dos articuladores, quando comparadas às mulheres sem nódulos vocais.

Palavras-chave: Acústica; Voz; Distúrbios da voz; Qualidade da voz; Fonoaudiologia

Study carried out at Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.
[1]Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.
[2]Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.
Conflict of interests: No.
Authors' contribution: KCFC, FPF, and LWL participated in the study construction and development; KCFC and FPF participated especially in data collection and tabulation; KCFC participated in data interpretation and article writing; LWL participated in study design, supervision, and writing of the final version of the article.
Funding: None.
Corresponding author: Fernanda Pereira França. E-mail: fgafernandap@hotmail.com
Received: March 28, 2022; **Accepted:** October 24, 2022

INTRODUCTION

The purpose of integrating auditory-perceptual analysis, otorhinolaryngological analysis, self-assessment, and acoustic analysis is to reach an effective diagnosis and open the way for voice rehabilitation[1-3].

The auditory-perceptual assessment considers the description of the voice signal through human hearing. Even though it is a subjective analysis, it is pointed out as the main method to assess voice disorders[4,5]. Hence, the effort to find an objective voice assessment led to progress in voice analysis, helping to understand the relationship between physiology (input) and auditory acoustic sound signal (voice output)[6,7].

The acoustic analysis provides visual and normative data on the glottal source and voice filter[6,8]. Some acoustic measures furnish quantitative data on the glottal source – e.g., the fundamental frequency (f0), defined by the number of vibrations per second produced by the vocal folds; jitter, the short-term variation in the fundamental frequency; shimmer, which indicates the short-term amplitude variability of the wave sound; glottal-to-noise excitation (GNE), which verifies the additional noise in the voice signal; and cepstral peak prominence-smoothed (CPPS), which determines the f0 and estimates the aperiodicity and/or additional noise in the voice signal[4,5,7].

Studies demonstrate that jitter is the best predictor during sustained phonation in the assessment of individuals with and without dysphonia[9]. The degree of vocal deviation is associated with CPPS[5,9], and GNE discriminates the overall degree of vocal deviation and quality predominantly in dysphonic patients[4].

The literature has demonstrated that voice disorders may manifest in various ways in the glottal source and filter (vocal tract), due to the constant interaction between the respiratory, phonatory, and resonant subsystems[7,8,10,11]. Thus, knowing investigative measures regarding the vocal tract can indicate new paths in voice diagnosis and rehabilitation.

The formants are the main investigative acoustic measures, as they provide information on the resonance produced in the vocal tract. They are influenced by the articulatory organs, according to the configuration and position of the articulators and the volume of the resonance cavities in the vocal tract as sounds are produced[12].

The first and second formants are the main ones to distinguish vowels. The first formant (F1) is related to the oromandibular complex; as the oral cavity opens, the F1 frequency decreases. The second formant (F2) is related to the degree of vowel anteriorization – i.e., the extent to which the pharynx is free based on tongue position; the more the tongue is anteriorized, the greater the F2 frequency[8,13,14].

The relationship between the harmonics H1 and H2 (i.e., the amplitude difference between the first two harmonics in the voice spectrum) is another acoustic measure that complements voice analysis, as it presents the direct influence of the velocity of vocal fold aerodynamic forces[15]. A study[15] investigated the voice behavior in 90 female patients diagnosed with voice nodules or polyps and 90 without lesions. They found that the weekly use of the voice in patients with phonotraumatic voice hyperfunction causes them to have more abrupt glottal closure, reducing the H1-H2 difference.

Other studies[6,16] identified that individuals with an irregular glottal source (e.g., incomplete glottal closure and irregular vocal fold vibration) can have compensatory vocal tract adjustments. Hence, glottal source measures have been generally considered important parameters to characterize the glottal closure and vibration mechanism in individuals with either healthy or deviated voices[5,17] and characterize the severity of dysphonia[15]. On the other hand, formant measures provide essential information on the source/filter integration process[10,11].

The cooperation between the subsystems involved in sound production can make adjustments in the vocal tract, as they adapt to the inefficiency of the phonatory mechanism[18]. Nevertheless, specific measures that shed light on source/filter coupling must be investigated to understand supraglottal adjustments and their influence on voice disorders.

Thus, the objective of this study was to compare the acoustic and auditory-perceptual measures in women with and without vocal nodules.

METHODS

Study design

This descriptive, analytical study was evaluated and approved by the Research Ethics Committee of the Federal University of Paraíba (UFPB) under evaluation report no. 2.158.960. The research was explained to all participants, who signed an informed consent form (ICF).

Sample

Two groups were formed for this research: a group of adult women with vocal nodules (VNG) followed up at the Integrated Laboratory for Voice Studies at UFPB and a group of adult women with no vocal nodules or voice disorders (NVNG). The eligibility criteria for VNG were as follows:

- Being otorhinolaryngologically diagnosed with vocal nodules.
- Being female, as voice disorders prevail in this sex. Also, due to anatomical vocal fold and vocal tract characteristics, differences between men and women influence the mean f0 values and formant measures.
- Being 19 to 59 years old, to avoid glottal source and vocal tract changes related to childhood, adolescence, and senescence.

NVNG was formed by actively recruiting voluntary employees and undergraduate students at UFPB who met the same eligibility criteria for VNG, except for the vocal nodule diagnosis, as follows:

- Not having voice complaints at the moment of the collection or in the 6 months before that, negatively answering the question: "Do you currently have any voice problem, or have you had any in the last 6 months?".
- Being otorhinolaryngologically diagnosed with a normal larynx.

The following exclusion criteria were used for both groups:

- o Having upper airway infections at the time of recording, which would change the resonance cavity and consequently the formant measures.
- o Having a short lingual frenulum, temporomandibular disorder, and/or articulator structural or functional changes, which would change supraglottal vocal tract adjustments.
- o Having cognitive or neurological changes that hindered the collection procedures.

o Having previously had speech-language-hearing therapy.

VNG and NVNG were matched for age, admitting a variation of 5 years in either direction, with the proportion of one control for every case (1:1).

All women in VNG were recruited in the initial consultation at the laboratory where the research was conducted, before beginning the voice therapy. Researchers initially surveyed the patients' assessment records to verify their overall sociodemographic data (date of birth, age, educational attainment, and occupation), voice complaint, and laryngeal diagnosis. The larynx was visually examined with video laryngoscopy by only one otorhinolaryngologist, who actively participated in the cooperative activities of the research and public outreach projects developed by the Integrated Laboratory for Voice Studies of the Department of Speech-Language-Hearing Sciences. These examinations followed the initial assessment from the Laboratory of Voice at UFPB. After voice assessment, those who met the eligibility criteria were invited to participate in the research upon reading and agreeing with the ICF. Hence, VNG had 12 women with a mean age of 36.47 ± 12.22 years.

NVNG women were then recruited based on the VNG participants' ages. The researcher contacted and informed them about the research, and those who agreed with and signed the ICF were invited to participate. Then, personal data were collected (date of birth, age, educational attainment, and occupation), and they were asked about voice complaints (to which they should give a negative answer), as defined in the eligibility criteria. Those who meth the NVNG criteria were referred for visual laryngeal examination (video laryngoscopy) by an otorhinolaryngologist from outside the institution to rule out any functional or structural changes in the larynx. Based on examination results, women with a normal larynx were recruited to continue participating in the study. Thus, the NVNG had 12 women with a mean age of 33.86 ± 11.59 years.

Sample collection procedure

VNG and NVNG participants selected based on the eligibility criteria were invited to have their voices collected. At the beginning of the collection session, the objectives of the research were once again read from the ICF, and their personal data were confirmed.

Firstly, the stomatognathic system structures were assessed, observing the morphology and mobility of the lips, tongue, cheeks, and soft palate; the tonus of the lips, tongue, and cheeks; mouth opening amplitude; and self-reported temporomandibular joint complaints and upper airway infections. This assessment aimed to rule out any temporomandibular disorder, lingual frenulum changes, or any structural or functional changes that might influence the study results by interfering with articulatory adjustments.

Then, the speech tasks were recorded with the Fonoview software, version 4.5, manufactured by CTS Informática, in an all-in-one Dell desktop computer, using a Sennheiser cardioid unidirectional microphone, model E-835, placed on a stand 10 cm away from the participant's mouth and connected to a Behringer preamplifier, model U-Phoria UMC 204. The voices were collected in an acoustically treated recording room in a laboratory for voice studies. The noise in the room was below 50 dBSPL, verified with an R8050 sound pressure level meter, by Reed Instruments, with a 44000 Hz sample rate and 16 bits per sample.

During voice collection, the women were standing in front of the microphone stand, keeping the said distance between their mouths and the microphone. They were instructed to breathe in naturally and then emit a sustained vowel /a/ for at least 5 seconds and repeat three times each of the following carrier phrases separately: "*Digo p**a**pa baixinho*"; "*Digo p**i**pa baixinho*", and "*Digo p**u**pa baixinho*" ("I say *papa*/*pipa*/*pupa* softly").

The sustained vowel /a/ was chosen because it is a low vowel, which is emitted with a stable phonatory system, ensuring a reliable laryngeal stability assessment(12).

The phrases used in this research had the [a, i, u] vowel segments in a consonant-vowel context, in an unstressed syllable at the beginning of the word, in which a voiceless bilabial plosive phoneme was used before and after the vowel. These phrases were used due to the little influence these consonants have on the formants of the adjacent vowels(12) and the need to homogenize the phonetic context of the collected vowels to investigate the speech samples. Thus, prosodic aspects were minimally controlled, without interfering with speech sound production to investigate acoustic distinctions between vowels.

The vowels [a, i, u] were used for the recognized acoustic distinctiveness between them, forming a vowel articulation triangle at their ends(12). Moreover, they follow a formant pattern about which researchers have reached a consensus, corresponding to the typical characteristics of vowels that have maximum and minimum vowel opening and tongue forward, backward, up, and down movements.

Auditory-perceptual analysis

For the perceptual measure analysis and then acoustic analysis, the voices were normalized with the "normalize" function in the Audacity program, version 2.3.1. Thus, all audio output signals were standardized at -6 to 6 dB to prevent audio signal intensity from interfering with the rater's judgment.

A speech-language-hearing therapist with more than 10 years of experience in the area conducted the auditory-perceptual analysis of the voices to describe and characterize the overall severity of the samples(7,19). He was instructed to consider a sample with normal variability of voice quality, produced with no irregularities, noises, or observable effects during vowel emission. The rater was also trained with 16 anchor stimuli with normal vowel emissions and different degrees of deviation.

The Visual Analog Scale (VAS) was used to classify vocal deviation, considering voice quality deviations greater than 35.5 mm on the overall level of the scale(9). Hence, it assessed the overall degree of vocal deviation (OD) and the degrees of roughness (DR), breathiness (DB), and strain (DS) in the sustained emission of vowel /a/. Speakers presented each sustained vowel emission to the speech-language-hearing therapist at a self-reported comfortable intensity. A total of 20% of the samples were repeated to calculate intra-rater reliability and Cohen's kappa coefficient. The kappa value was 0.89, which indicates excellent rater reliability.

Acoustic measure extraction

Mean f0, standard deviation, jitter, shimmer, and GNE were extracted from the sustained vowel /a/ sample, using the voice quality analysis module in the VoxMetria software, version 4.7, by CTS Informática. The following commands were applied to obtain mean f0 and measures of disturbance and noise:

1. Selecting the option "Voice quality".
2. Opening the file to be analyzed.
3. Selecting and eliminating the initial and final 2 seconds of the vowel emission due to their greater irregularity and maintaining at least 3 seconds of each emission.
4. Selecting the option "Voice analysis data", making available data on mean f0, standard deviation, jitter, shimmer, and GNE.

The other acoustic measures were extracted with the free Praat software (version 5.3.77h), developed by Paul Boersma and David Weenink of the University of Amsterdam, in the Netherlands[20].

CPPS was obtained following the parameters suggested by Maryn and Weenink[3], with the following procedures:

1. Choosing a file: open, open long sound file, extract part, ok. After choosing the file to be analyzed, the option "Analyze Periodicity" was chosen, and then "To Power Cepstrogram".
2. In the option "To Power Cepstrogram", it proceeded with the following parameters: Pitch floor (Hz) = 60, Time Step (s) = 0.002, Maximum Frequency (Hz) = 5000, and Pre-emphasis from (Hz) = 50.
3. The option "Query" was selected, and then "Get CPPS"; in the menu, it proceeded with the following parameters: Subtract tilt before smoothing selected, Time averaging window (s) = 0.01, Quefrency-averaging window (s) = 0.001, Peak search pitch range (Hz) = 60-330, Tolerance (0-1) = 0.05, Interpolation = Parabolic, Tilt line quefrency range (s) = 0.001-0.0 (= end), Line type = Straight, and Fit method = Robust.
4. The results of this procedure were the CPPS measures, expressed in decibels (dB).

Intensity (dB) and H1 and H2 measures were manually extracted using the following steps:

1. Selecting the option "Read from file"; choosing the file to be analyzed.
2. Selecting the option "View & edit"; the central portion of the voice signal was selected.
3. Selecting the option "sel", then "spectrum", and lastly "view spectral slice", resulting in a graph image with the peaks. The cursor is placed at the tip of the first peak (H1) to obtain its intensity (dB). Then, the values of the second peak (H2) were verified.

The mean F1 and F2 of vowels /a/, /i/, and /u/ were extracted from selected and segmented vowel sounds in consonant-vowel contexts. The means of the formants were extracted in Praat by selecting the option "Formant", obtaining F1, F2, and F3 values in Hertz (Hz).

The segments and duration of the vowel sounds in consonant-vowel contexts were obtained considering the first regular peak after the consonant and consonant-vowel transition as the initial limit of the vowel. The final limit was the last regular peak before the consonant and consonant-vowel transition[12]. The mean analysis duration was estimated at 0.13 seconds.

Data analysis

The statistical analysis was based on the mean and standard deviation of the investigated measures. Intergroup comparisons (VNG and NVNG) were made with Student's paired t-test and the Wilcoxon nonparametric paired test. The level of significance in the analyses was set at 5%.

RESULTS

Table 1 shows the means, standard deviation, and comparison of auditory-perceptual and specific glottal source acoustic measures between NVG and NVNG.

The comparative intergroup analysis revealed differences in the auditory-perceptual assessment regarding OD ($p = 0.0020$), DR ($p = 0.0105$), and DB ($p = 0.0059$) in VAS. VNG had higher values in these parameters than NVNG.

The only one of the investigated acoustic measures related to the glottal source with a difference between the groups was shimmer ($p = 0.0449$), whose mean values were higher in VNG.

Table 2 describes the characterization of means, standard deviation, and comparison of the specific vocal filter acoustic measures between VNG and NVNG.

Concerning supraglottal adjustment measures, there was a difference in F1 values for vowels /a/ ($p = 0.0145$) and /u/ ($p = 0.0007$) and F2 values for vowel /a/ ($p = 0.0284$) between the groups, with higher values in VNG.

DISCUSSION

The dissemination of studies on voice assessment parameters led to the investigation of different voice signal pick-up measures[7-9,21]. Insights based on the source-filter theory (in combination with nonlinear theories) and objective and subjective human measures lead to and assure the diagnosis and consequently the clinical voice intervention. Hence, clinicians and researchers must pay full attention to the physiological, acoustic, and perceptual mechanisms in the particularities of sound production.

It is thus understood that the acoustic speech production mechanism involves adjustments made by the vocal tract in both the glottal source and filter. Hence, this study aimed to investigate the source-filter relationship process in the development

Table 1. Comparison of auditory-perceptual and specific glottal source acoustic measures between the groups of women with and without vocal nodules

Auditory-perceptual/ acoustic measures	Group		p-value
	VNG	NVNG	
	Mean ± SD	Mean ± SD	
VAS OD	61.79 ± 8.06	47.83 ± 11.79	**0.0020***
VAS DR	55.13 ± 8.45	43.88 ± 10.96	**0.0105***
VAS DB	58.29 ± 14.25	38.13 ± 17.78	**0.0059***
VAS DS	30.92 ± 14.91	30.33 ± 11.48	0.9155
Mean f0	205.19 ± 20.44	202.37 ± 16.37	0.7130
SD f0	3.69 ± 3.14	2.55 ± 1.29	0.4188
Jitter	0.87 ± 1.73	0.28 ± 0.31	0.2142
Shimmer	6.91 ± 8.07	3.31 ± 1.01	**0.0449***
GNE	0.63 ± 0.23	0.82 ± 0.17	0.1484
CPPS	13.10 ± 1.57	13.15 ± 1.78	0.9097

Student's t-test/Wilcoxon test; *Significant values ($p < 0.05$)
Subtitle: VNG = group of women with vocal nodules; NVNG = group of women with no vocal nodules; VAS = visual analog scale; OD = overall degree; DR = degree of roughness; DB = degree of breathiness; DS = degree of strain; f0 = fundamental frequency; SD = standard deviation; GNE = glottal-to-noise excitation ratio; CPPS = cepstral peak prominence-smoothed

Table 2. Comparison of specific vocal filter acoustic measures between the groups of women with and without vocal nodules

Acoustic measures	Group		p-value
	VNG	NVNG	
	Mean ± SD	Mean ± SD	
H1-H2 (dB)	7.42 ± 5.91	4.94 ± 4.51	0.4697
F1 vowel /a/	804.40 ± 137.39	931.28 ± 83.31	**0.0145***
F1 vowel /i/	353.08 ± 47.66	377.32 ± 41.67	0.3474
F1 vowel /u/	369.01 ± 57.37	465.06 ± 50.97	**0.0007***
F2 vowel /a/	1377.63 ± 120.70	1471.79 ± 83.21	**0.0284***
F2 vowel /i/	2491.65 ± 170.93	2578.58 ± 137.92	0.3777
F2 vowel /u/	725.32 ± 225.65	726.55 ± 78.93	0.4776

Student's t-test/Wilcoxon test; *Significant values ($p < 0.05$)
Subtitle: VNG = group of women with vocal nodules; NVNG = group of women with no vocal nodules; H1-H2 (dB) = amplitude difference between the two first harmonics, in decibels; F1 = first formant; F2 = second formant

of various types of auditory-perceptual and acoustic analysis measures, as voice changes usually disturb sound in different ways[13,22,23].

It was found that women with vocal nodules had a greater overall degree of vocal deviation; consequently, they had a greater degree of roughness and breathiness than those without vocal nodules (Table 1). In general terms, it is known that mass lesions in the vocal folds can cause voice quality deviations, which may be explained by irregular vibration and/or glottal gaps, causing incomplete glottal closure during phonation, and resulting in roughness and breathiness in sound production[1].

With an intervention approach using resonant voice therapy, a study[21] assessed objective and subjective measures in 26 women with vocal nodules and 30 vocally healthy women. There were significant intergroup differences in f0, jitter, shimmer, maximum phonation time, and noise-to-harmonic ratio before the intervention (with high values in disturbance and noise measures in women with vocal nodules) and after the intervention (only in women with vocal nodules, with a decrease in these values).

The auditory-perceptual assessment of NVNG women showed that the overall degree, roughness, and breathiness were mildly deviated, which can be physiologically explained by the fact that they were females. Women have lower glottal proportion values (which reflect a greater tendency toward incomplete glottal closure) and are influenced by hormonal factors[24,25] that might influence voice quality results in women with no complaints or laryngeal changes. Moreover, the examination used in this research (video laryngoscopy) is not the most recommended one to assess the flexibility of vocal fold mucosa, mucosal tissue, and deeper layers – videostroboscopy should be used instead[26].

Videostroboscopy is the clinical gold standard. In the analysis, a light source is synchronized with the speaker's f0. The resulting image gives an illusion of slow-motion vocal fold vibration. This mechanism provides essential data for a laryngeal diagnosis regarding glottal closure, vibration amplitude, and mucosal wave[26].

On the other hand, the absence of voice quality deviation in NVNG women was not part of the inclusion criteria in the present study – which only stated that they could not have voice complaints or laryngeal lesions, as certain levels of instability and absence of periodicity can be normal. Voice quality instabilities can occur even when the subjects are careful to sustain a stable emission in terms of level and intensity[27]. Hence, these findings demonstrated that, even when there are no changes or complaints and acoustic measures are normal, the individual may have voice quality deviation. In this perspective, multidimensional voice assessment increasingly favors effective diagnoses of voice disorders[7].

Concerning disturbance and noise measure analysis, women with vocal nodules differed from women with no laryngeal changes only in shimmer, as VNG had higher values than NVNG. Shimmer is known as a measure related to glottal resistance and the maintenance of the closed phase of the glottal cycles. Thus, higher values are expected in this measure in women with vocal nodules because of incomplete glottal closure, commonly found in this type of laryngeal lesion[28].

Formant measure analysis showed that F1 values for vowels /a/ and /u/ and F2 values for vowel /a/ were lower in VNG than in NVNG – which can respectively indicate elevated oromandibular complex and posteriorized tongue during phonation, since F1 values are directly related to the posture of this complex, and F2 values are related to tongue position in the anteroposterior direction[12].

Differences found in formant measure values between the groups in this study may reinforce a bidirectional relationship between supraglottal and glottal adjustments. The inferred difference in tongue position between the groups suggests that the position and amplitude of movement of vocal tract articulators may be changed during speech in individuals with voice disorders[6,8].

Phonation associated with an elevated body of the tongue results in breathy voice quality. This statement is related to the results; however, the great issue for future directions is trying to understand whether these adjustments take place together with or as a consequence of voice disorders[8].

In this perspective, a study[29] investigated the voice quality and dynamics in 25 teachers with voice disorder complaints or manifestations and laryngeal changes. Vocal tract overload mechanisms led to voice quality adjustments (e.g., laryngeal hyperfunction and raspy voice) and adjustment changes throughout the vocal tract (elevated larynx, vocal tract hyperfunction, closed mandible, pharyngeal constriction, elevated body of the tongue, and air escape).

According to another study[10], habitual vocal tract adjustments in dysphonic and non-dysphonic women are different at rest and during phonation, as vocal effort can change the position of vocal tract articulators in patients with vocal nodules. The authors investigated with magnetic resonance the vocal tract adjustments in this population before and after resonance exercise with flexible tubes in water. Closed mandible, rounded lips, and retracted body of the tongue were supraglottal adjustments found in the population with dysphonia[10].

The analysis in this study reinforces that vocal fold functioning is not independent of the vocal tract, and that habitual vocal tract adjustments in women with and without vocal nodules are different during phonation[10,11]. Hence, they provide a worthy methodological/evaluative referential for clinical voice rehabilitation and monitoring.

Thus, the present study showed that understanding voice production physiology is correlated with effective voice disorder assessment since the acoustic signal results from the interaction between somatosensory and motor information[30]. Therefore, glottal source and vocal tract auditory and acoustic data are indispensable.

LIMITATIONS AND FUTURE DIRECTIONS

Some limiting factors in this study are the sample size in each subgroup, the few judges in the voice quality assessment process, and the lack of a more specific diagnosis examination – i.e., videostroboscopy, which is the gold standard for laryngeal assessment.

Future studies should use videostroboscopy to reach an effective laryngeal diagnosis, obtaining precise vocal fold physiology images. They should also consider the glottal cycles and, above all, have larger samples to make the study more robust.

CONCLUSION

The voices of women with nodules had greater deviation, roughness, breathiness, and higher shimmer values than those of women without vocal nodules. As for formant measures, women with vocal nodules had lower F1 values for vowels /a/ and /u/ and F2 values for vowel /a/ than women without vocal nodules. Hence, the analysis of women with vocal nodules can infer a possible decreased articulator amplitude, with elevated mandible, posteriorized tongue, and pharyngeal readjustments.

REFERENCES

1. Oliveira RCCD, Gama ACC, Genilhú PFL, Santos MAR. Videolaringoscopia digital de alta velocidade: avaliação de nódulos vocais e cistos em mulheres. CoDAS. 2021;33(3):e20200095. http://dx.doi.org/10.1590/2317-1782/20202020095. PMid:34008770.

2. Lopes LW, Cavalcante DP, Costa PO. Severity of voice disorders: integration of perceptual and acoustic data in dysphonic patients. CoDAS. 2014;26(5):382-8. http://dx.doi.org/10.1590/2317-1782/20142013033. PMid:25388071.

3. Maryn Y, Weenink D. Objective dysphonia measures in the program *Praat*: smoothed cepstral peak prominence and acoustic voice quality index. J Voice. 2015 Jan;29(1):35-43. http://dx.doi.org/10.1016/j.jvoice.2014.06.015. PMid:25499526.

4. Lopes LW, Alves JN, Evangelista DS, França FP, Vieira VJD, Lima-Silva MFB, et al. Acurácia das medidas acústicas tradicionais e formânticas na avaliação da qualidade vocal. CoDAS. 2018;30(5):e20170282. http://dx.doi.org/10.1590/2317-1782/20182017282. PMid:30365651.

5. Lopes LW, Sousa ESS, Silva ACF, Silva IM, Paiva MAA, Vieira VJD, et al. Medidas cepstrais na avaliação da intensidade do desvio vocal. CoDAS. 2019;31(4):e20180175. http://dx.doi.org/10.1590/2317-1782/20182018175. PMid:31433040.

6. Ishikawa K, Nudelman C, Park S, Ketring C. Perception and acoustic studies of vowel intelligibility in dysphonic speech. J Voice. 2021 Jul;35(4):659.e11-24. http://dx.doi.org/10.1016/j.jvoice.2019.12.022. PMid:31952898.

7. Lopes LW, França FP, Evangelista DS, Alves JN, Vieira VJD, Lima-Silva FB, et al. Does the combination of glottal and supraglottic acoustic measures improve discrimination between women with and without voice disorders? J Voice. 2022;36(4):583.e17-29. http://dx.doi.org/10.1016/j.jvoice.2020.08.006. PMid:32917459.

8. França FP, Almeida AA, Lopes LW. Configuração acústico-articulatória das vogais de mulheres com nódulos vocais e vocalmente saudáveis.

CoDAS. 2019;31(6):e20180241. http://dx.doi.org/10.1590/2317-1782/20192018241. PMid:31751443.

9. Hassan EM, Abdel Hady AF, Shohdi SS, Eldessouky HM, Din MHB. Assessment of dysphonia: cepstral analysis versus conventional acoustic analysis. Logoped Phoniatr Vocol. 2021 Out;46(3):99-109. http://dx.doi.org/10.1080/14015439.2020.1767202. PMid:32436465.

10. Yamasaki R, Murano EZ, Gebrim E, Hachiya A, Montagnoli A, Behlau M, et al. Vocal tract adjustments of dysphonic and non-dysphonic women pre- and post-flexible resonance tube in water exercise: a quantitative MRI study. J Voice. 2017 Jul;31(4):442-54. http://dx.doi.org/10.1016/j.jvoice.2016.10.015. PMid:28017460.

11. Cielo CA, Christmann MK. Finger kazoo: modificações vocais acústicas espectrográficas e autoavaliação vocal. Rev CEFAC. 2014;16(4):1239-54. http://dx.doi.org/10.1590/1982-021620145513.

12. Barbosa PA, Madureira S. Manual de fonética acústica experimental: aplicações a dados do português. São Paulo: Cortez; 2015. 591 p.

13. Lee Y, Park HJ, Bae IH, Kim G. Resonance characteristics in epiglottic cyst: formant frequency, vowel space area, vowel articulatory index, and formant centralization ratio. J Voice. 2021. http://dx.doi.org/10.1016/j.jvoice.2021.09.008. PMid:34642071.

14. Sundberg J, Lindblom B, Hefele AM. Voice source, formant frequencies and vocal tract shape in overtone singing: a case study. Logoped Phoniatr Vocol. 2021;1-13. http://dx.doi.org/10.1080/14015439.2021.1998607. PMid:34860148.

15. van Stan JH, Mehta DD, Ortiz AJ, Burns JA, Toles LE, Marks KL, et al. Differences in weeklong ambulatory vocal behavior between female patients with phonotraumatic lesions and matched controls. J Speech Lang Hear Res. 2020;63(2):372-84. http://dx.doi.org/10.1044/2019_JSLHR-19-00065. PMid:31995428.

16. Evitts PM, Starmer H, Teets K, Montgomery C, Calhoun L, Schulze A, et al. The impact of dysphonic voices on healthy listeners: listener reaction times, speech intelligibility, and listener comprehension. Am J Speech Lang Pathol. 2016 Nov 1;25(4):561-75. http://dx.doi.org/10.1044/2016_AJSLP-14-0183. PMid:27784031.

17. Bekerman AL. Uso del análisis acústico en el seguimiento de pacientes con patología vocal: estudio preliminary. Rev. Fed. Argent. Soc. Otorrinolaringol. 2017;24(3):28-35.

18. França FP, Evangelista DS, Lopes LW. Revisão sistemática sobre os formantes e a produção da voz e fala. Prolíngua. 2017;12(1):2-16.

19. Paz KEDS, Almeida AAF, Almeida LNA, Sousa ESDS, Lopes LW. Auditory perception of roughness and breathiness by dysphonic women. J Voice. 2022. http://dx.doi.org/10.1016/j.jvoice.2022.01.005. PMid:35082050.

20. Boersma P, Weenick D. Praat manual [Internet]. Amsterdam: Phonetic Sciences Department, University of Amsterdam; 2006 [citado em 2022 Jul 19]. Disponível em: http://www.fon.hum.uva.nl/praat/

21. Saltürk Z, Özdemir E, Sari H, Keten S, Kumral TL, Berkiten G, et al. Assessment of resonant voice therapy in the treatment of vocal fold nodules. J Voice. 2019;33(5):810.e1-4. http://dx.doi.org/10.1016/j.jvoice.2018.04.012. PMid:30017429.

22. Costa SC, Costa WCA, Correia SEN, Araújo JMFR, Vieira VJD. Análise de sinais de voz para caracterização de patologias na laringe. Rev Tecnol Inf Comun. 2014;4(2):63-70. http://dx.doi.org/10.12721/2237-5112.v04n02a09.

23. Maryn Y, Roy N, De Bodt M, van Cauwenberge P, Corthals P. Acoustic measurement of overall voice quality: a meta-analysis. J Acoust Soc Am. 2009;126(5):2619-34. http://dx.doi.org/10.1121/1.3224706. PMid:19894840.

24. Figueiredo LC, Gonçalves MIR, Pontes A, Pontes P. Estudo do comportamento vocal no ciclo menstrual: avaliação perceptive-auditiva, acústica e autoperceptiva. Rev Bras Otorrinolaringol. 2004;70(3):331-9. http://dx.doi.org/10.1590/S0034-72992004000300008.

25. Arruda P, Diniz da Rosa MR, Almeida LNA, de Araujo Pernambuco L, Almeida AA. Vocal acoustic and auditory-perceptual characteristics during fluctuations in estradiol levels during the menstrual cycle: a longitudinal study. J Voice. 2019 Jul;33(4):536-44. http://dx.doi.org/10.1016/j.jvoice.2018.01.024. PMid:29525075.

26. Powell ME, Deliyski DD, Zeitels SM, Burns JA, Hillman RE, Gerlach TT, et al. Efficacy of videostroboscopy and high-speed videoendoscopy to obtain functional results from perioperative classifications in patients with mass injuries in the vocal folds. J Voice. 2020 Set;34(5):769-82. http://dx.doi.org/10.1016/j.jvoice.2019.03.012. PMid:31005449.

27. Gama ACC, Behlau M. Estudo da constância de medidas acústicas em mulheres sem queixa de voz e em mulheres com disfonia. Rev Soc Bras Fonoaudiol. 2009;14(1):8-14. http://dx.doi.org/10.1590/S1516-80342009000100004.

28. Dajer ME, Andrade FAS, Montagnoli AN, Pereira JC, Tsuji DH. Vocal Dynamic Visual Pattern for voice characterization. J Phys Conf Ser. 2011;332(1):012026. http://dx.doi.org/10.1088/1742-6596/332/1/012026.

29. Lima-Silva MFB, Madureira S, Rusilo LC, Camargo Z. Perfil vocal de professores: análise integrada de dados de percepção e acústica. In: Camargo Z, editor. Fonética clínica. São Paulo: Pulso; 2016.

30. Raharjo I, Kothare H, Nagarajan SS, Houde JF. Speech compensation responses and sensorimotor adaptation to formant feedback perturbations. J Acoust Soc Am. 2021 Fev;149(2):1147-61. http://dx.doi.org/10.1121/10.0003440. PMid:33639824.