

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

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

Késsia Cecília Fernandes Conserva<sup>1</sup> , Fernanda Pereira França<sup>1</sup> , Leonardo Wanderley Lopes<sup>2</sup> 

### 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 soproidade 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 soproidade 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

### 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

Trabalho realizado no Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

<sup>1</sup>Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

<sup>2</sup>Universidade Federal da Paraíba – UFPB – João Pessoa (PB), Brasil.

**Conflito de interesses:** Não.

**Contribuição dos autores:** KCFC, FPF e LWL participaram da construção e desenvolvimento do trabalho; KCFC e FPF envolveram-se particularmente na coleta e tabulação dos dados; KCFC participou da interpretação dos dados e redação do artigo; LWL participou do delineamento do estudo, orientação e redação final do artigo.

**Financiamento:** Nada a declarar.

**Autor correspondente:** Fernanda Pereira França. E-mail: fgafernandap@hotmail.com

**Recebido:** Março 28, 2022; **Aceito:** Outubro 24, 2022

## INTRODUÇÃO

A integração da análise perceptivo-auditiva, análise otorrinolaringológica, autoavaliação e análise acústica tem a finalidade de alcançar o diagnóstico efetivo e abrir caminhos para reabilitação, na clínica vocal<sup>(1-3)</sup>.

A avaliação perceptivo-auditiva considera a descrição do sinal vocal por meio da audição humana. É uma análise subjetiva, porém, destacada como principal método avaliativo dos distúrbios de voz<sup>(4,5)</sup>. Nesse sentido, a busca por uma avaliação vocal objetiva propiciou o progresso da análise vocal, favorecendo, assim, a compreensão da relação entre fisiologia (*input*) e sinal sonoro acústico auditivo (*output* vocal)<sup>(6,7)</sup>.

A análise acústica fornece subsídios visuais e dados normativos sobre a fonte glótica e o filtro vocal<sup>(6,8)</sup>. Algumas medidas acústicas trazem dados quantitativos sobre a fonte glótica, tais como frequência fundamental ( $f_0$ ), definida pelo número de vibrações produzidas pelas pregas vocais por segundo; *jitter*, a variação da frequência fundamental em curto prazo; o *shimmer*, que indica a variabilidade da amplitude da onda sonora em curto prazo; o *Glottal to Noise Excitation* (GNE), que verifica o ruído adicional no sinal vocal e o CPPS (*Cepstral Peak Prominence-Smoothed*), que é capaz de determinar a  $f_0$  e produzir estimativas de aperiodicidade e/ou ruído aditivo do sinal vocal<sup>(4,5,7)</sup>.

Nessa perspectiva, estudos evidenciam que *jitter* tem a melhor capacidade preditiva durante a fonação sustentada na avaliação de indivíduos com e sem disfonia<sup>(9)</sup>; existe associação entre o grau do desvio vocal e o CPPS<sup>(5,9)</sup>; há discriminação do grau geral do desvio e da qualidade vocal predominante em disfônicos por meio do GNE<sup>(4)</sup>.

A literatura tem evidenciado que os distúrbios de voz podem apresentar diversas manifestações na fonte glótica e no filtro (trato vocal), devido à interação constante entre os subsistemas respiratório, fonatório e ressonante<sup>(7,8,10,11)</sup>. Dessa forma, o conhecimento de medidas investigativas a respeito do trato vocal pode elucidar novos caminhos para o diagnóstico e reabilitação vocal.

Nesse sentido, uma das principais medidas acústicas investigativas são os formantes, que trazem informações das ressonâncias geradas pelo trato vocal. Eles sofrem influência dos órgãos fonoarticulatórios, de acordo com a configuração e o posicionamento dos articuladores, assim como no volume das cavidades de ressonância no trato vocal, de acordo com a produção dos sons<sup>(12)</sup>.

Os principais formantes que distinguem as vogais são o primeiro e o segundo. O primeiro formante (F1) tem relação com o complexo oromandibular e, à medida que ocorre abertura da cavidade oral, a frequência desse formante reduz. Já o segundo formante (F2), está relacionado com o grau de anteriorização da vogal, ou seja, o quanto a faringe está livre pelo posicionamento da língua. Quanto maior o grau de anteriorização da língua, maior será a frequência de F2<sup>(8,13,14)</sup>.

Outra medida acústica que permite complementaridade à análise vocal é a relação entre os harmônicos H1–H2, que se refere à diferença de amplitude entre os dois primeiros harmônicos do espectro da voz. Apresenta influência direta da velocidade das forças aerodinâmicas das pregas vocais<sup>(15)</sup>. Um estudo<sup>(15)</sup> investigou o comportamento vocal entre 90 pacientes do gênero feminino com diagnóstico de nódulos ou pólipos vocais e 90 sem lesão e encontrou que o uso semanal da voz

de pacientes com hiperfunção vocal fonotraumática faz com que eles apresentem fechamento glótico mais abrupto, com redução da diferença H1-H2.

Outros estudos<sup>(6,16)</sup> identificaram que indivíduos que apresentam irregularidades na fonte glótica, como, por exemplo, fechamento glótico incompleto e irregularidade vibratória das pregas vocais, podem realizar ajustes compensatórios no trato vocal.

Dessa forma, de modo geral, as medidas de fonte glótica têm sido consideradas importantes parâmetros para caracterização do mecanismo de fechamento e vibração glótica, seja em indivíduos vocalmente saudáveis ou nos casos de produção vocal desviada<sup>(5,17)</sup>, assim como para a caracterização da gravidade da disfonia<sup>(15)</sup>. Por outro lado, as medidas formânticas trazem informações imprescindíveis sobre o processo de integração fonte-filtro<sup>(10,11)</sup>.

As cooperações dos subsistemas envolvidos na produção do som podem desenvolver ajustes no trato vocal, na medida em que se adaptam à ineficiência do mecanismo fonatório<sup>(18)</sup>. No entanto, cabe a investigação de medidas específicas que tragam luz ao acoplamento fonte-filtro, uma vez que é necessário entender os ajustes supraglóticos e sua influência nos distúrbios vocais.

Dessa forma, o objetivo desse estudo foi comparar as medidas acústicas e perceptivo-auditivas de mulheres com e sem nódulos vocais.

## MÉTODOS

### Desenho do estudo

Trata-se de um estudo analítico e descritivo. Foi avaliado e aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal da Paraíba – UFPB, com o parecer número 2.158.960. Todos os participantes receberam explicação sobre a pesquisa e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE).

### Amostra

Para realização desta pesquisa, foram constituídos dois grupos: grupo de mulheres com nódulos vocais (MNV), com a participação de mulheres adultas atendidas no Laboratório Integrado de Estudos da Voz (LIEV) da Universidade Federal da Paraíba, e o de grupo mulheres sem nódulos vocais (MSNV), formado por mulheres adultas sem distúrbio de voz. Para o MNV, foram considerados os seguintes critérios de elegibilidade:

- o Apresentar laudo otorrinolaringológico de nódulos vocais;
- o Ser do gênero feminino, devido à maior prevalência de distúrbios da voz nesse gênero e à influência dessa variável nos valores da média da frequência fundamental ( $f_0$ ) e das medidas formânticas, que apresentam diferenças entre homens e mulheres adultos, em razão das características anatómicas das pregas vocais e do trato vocal;
- o Ter idade entre 19 e 59 anos, para evitar as modificações na fonte glótica e no trato vocal relativas à infância, adolescência e senescência;

Para a composição do grupo de mulheres sem nódulos vocais (MSNV), foram recrutadas ativamente mulheres funcionárias e alunas de um curso de graduação da instituição de ensino superior onde foi realizada a pesquisa, que se disponibilizassem para tanto e que se enquadrassem nos mesmos critérios de elegibilidade citados para o MNV, com exceção do diagnóstico de nódulos vocais;

- o Não apresentar queixa vocal no momento da coleta ou nos seis meses anteriores, respondendo negativamente à pergunta: “Você apresenta um problema de voz atualmente ou nos últimos seis meses?”;
- o Apresentar laudo otorrinolaringológico de laringe normal.

Em ambos os grupos foram considerados os seguintes critérios de exclusão:

- o Apresentar infecções de vias áreas superiores no momento da gravação, o que geraria modificação nas cavidades de ressonância e, conseqüentemente, nas medidas formânticas;
- o Possuir frênulo lingual encurtado, disfunção temporomandibular e/ou alterações estruturais e funcionais dos articuladores, o que modificaria os ajustes do trato vocal supraglótico;
- o Ter alterações cognitivas ou neurológicas que impedissem a realização dos procedimentos de coleta;
- o Ter realizado terapia fonoaudiológica anteriormente.

A variável faixa etária foi utilizada para o pareamento do MSNV com o MNV, considerando uma variação da idade de cinco anos para mais ou para menos, seguindo uma proporção de um controle para cada caso (1:1).

Todas as mulheres do MNV foram recrutadas no momento da consulta inicial, no laboratório onde foi realizada a pesquisa, antes de iniciar o processo de terapia vocal propriamente dita. Inicialmente, os pesquisadores tiveram acesso à ficha de avaliação das pacientes para verificação dos dados sociodemográficos gerais (data de nascimento, idade, grau de instrução e profissão), queixa vocal e diagnóstico laríngeo. O exame visual laríngeo (videolaringoscopia) foi feito por apenas um médico otorrinolaringologista que participa ativamente de atividades colaborativas nos projetos de pesquisa e extensão desenvolvidas pelo LIEV, no Departamento de Fonoaudiologia, a partir de encaminhamento solicitado na avaliação inicial no Laboratório de Voz da Universidade Federal da Paraíba. Ao final da sessão de avaliação vocal, aquelas que cumpriam os critérios de elegibilidade definidos foram convidadas a participar da pesquisa, após leitura, concordância e assinatura do TCLE. Dessa forma, a amostra do MNV foi composta por 12 mulheres, com média de idade de  $36,47 \pm 12,22$  anos.

Para recrutamento das mulheres do MSNV, considerou-se, então, a observação da faixa etária das participantes do MNV. Sendo assim, elas foram contatadas pelo pesquisador, que fornecia informações sobre a pesquisa e, em caso de anuência e assinatura do TCLE, foram convidadas a participar. Em seguida, iniciou-se a coleta de informações pessoais (data de nascimento, idade, grau de instrução e profissão) e uma pergunta sobre a presença de queixa vocal, que deveria ser respondida negativamente, conforme critérios de elegibilidade. Aquelas que correspondiam aos critérios de elegibilidade para o MSNV eram encaminhadas para realização do exame visual laríngeo (videolaringoscopia) com médico otorrinolaringologista externo à instituição, para

descartar a presença de alteração funcional ou estrutural na laringe. Após os resultados dos exames, as mulheres com laudo de “laringe normal” foram recrutadas para dar continuidade na participação do estudo. Dessa forma, a amostra do MSNV foi composta por 12 mulheres, com média de idade de  $33,86 \pm 11,59$  anos.

## Procedimento da coleta das amostras

As participantes do MNV e MSNV selecionadas a partir dos critérios de elegibilidade foram convidadas a participar de uma sessão para coleta de voz. Ao iniciar a sessão, foram retomados os objetivos da pesquisa, a partir da leitura do TCLE, e confirmados os dados pessoais.

Inicialmente, foi realizada avaliação das estruturas do sistema estomatognático, observando-se: morfologia e mobilidade de lábios, língua, bochechas e palato mole; tonicidade de lábios, língua e bochechas; amplitude de abertura da boca e presença de queixas relacionadas à articulação temporomandibular e de queixas de infecções de vias aéreas superiores (de acordo com o autorrelato). O objetivo dessa avaliação foi descartar a presença de desordem temporomandibular, alteração no frênulo lingual, ou qualquer alteração estrutural e funcional que pudesse influenciar os resultados deste estudo, em função da interferência dessas alterações sobre os ajustes articulatorios.

Posteriormente, realizou-se a gravação das tarefas de fala. Para tanto, utilizou-se o *software* Fonoview, versão 4.5, da CTS Informática, *desktop* Dell *all-in-one*, microfone cardioide unidirecional, da marca Sennheiser, modelo E-835, localizado em um pedestal e acoplado a um pré-amplificador Behringer, modelo U-Phoria UMC 204. As vozes foram coletadas em cabine de gravação, em um laboratório de estudos da voz, com tratamento acústico e ruído inferior a 50 dBNPS, verificado com medidor de nível de pressão sonora R8050 da Reed Instruments, com taxa de amostragem de 44000 Hz, 16 bits por amostra e distância de 10 cm entre o microfone e a boca do falante.

Para a coleta das vozes, as mulheres ficaram em pé, situando o pedestal a sua frente, de acordo com a distância preconizada entre a boca e o microfone. Elas foram instruídas a realizar uma inspiração de modo habitual e, na sequência, emitir a vogal /a/ sustentada com duração mínima de cinco segundos, e repetir por três vezes cada uma das seguintes frases-veículo, separadamente: “Digo papa baixinho”; “Digo pipa baixinho” e “Digo pupa baixinho”.

A escolha da vogal /a/ sustentada é justificada pelo fato de ser uma vogal baixa, cuja emissão fornece uma condição estável do sistema fonatório, resultando em uma avaliação confiável da estabilidade da laringe<sup>(12)</sup>.

As frases utilizadas nesta pesquisa continham os segmentos vocálicos [a, i, u] em contexto CV (consoante-vogal), em sílaba inicial de palavra, não acentuada, com vogal antecedida e sucedida do fonema oclusivo bilabial desvozeado. A escolha dessas frases justifica-se pela pouca influência que essas consoantes têm sobre os formantes das vogais adjacentes<sup>(12)</sup> e pela necessidade de homogeneizar o contexto fonético para obtenção das vogais investigadas nas amostras de fala. Dessa forma, houve o mínimo controle dos aspectos prosódicos, sem interferências na realização dos sons vocálicos na investigação da distintividade acústica das vogais.

As vogais [a, i, u] foram escolhidas devido à reconhecida distintividade acústica entre elas, formando um triângulo

articulatório vocálico em suas extremidades<sup>(12)</sup>. Além disso, elas obedecem a um padrão formântico de consenso entre os pesquisadores, que corresponde às características típicas de vogais que apresentam o máximo e o mínimo de abertura vocálica e de movimento de recuo e de avanço, de abaixamento e levantamento da língua.

## Análise perceptivo-auditiva

Para a análise das medidas de percepção, assim como para a posterior análise acústica, as vozes passaram por processo de normalização no controle *normalize* do programa Audacity, versão 2.3.1, a fim de obter uma padronização na saída de áudio entre -6 e 6 dB para todos os sinais, de maneira que a intensidade do sinal de áudio não interferisse no julgamento do avaliador.

Para descrever e caracterizar a gravidade geral das amostras, um fonoaudiólogo com mais de dez anos de experiência na área realizou a análise perceptivo-auditiva das vozes<sup>(7,19)</sup>. Ele foi orientado a considerar uma amostra com variabilidade normal na qualidade da voz (NVVQ) quando fosse produzida sem irregularidades, ruídos ou efeitos observáveis durante a emissão vocal. Além disso, o avaliador foi treinado com 16 estímulos âncora contendo emissões vocais normais e com diferentes graus de desvio.

Para classificar o desvio vocal, utilizou-se a Escala Visual Analógica (EVA), considerando o desvio na qualidade vocal acima de 35,5 mm no nível geral da escala<sup>(9)</sup>. Por meio desta, avaliaram-se os parâmetros de grau geral do desvio vocal (GG) e os graus de rugosidade (GR), de soproidade (GS) e de tensão (GT) na emissão da vogal sustentada /a/. Cada emissão de vogal sustentada foi apresentada ao fonoaudiólogo por meio de um falante, em uma intensidade confortável autorreferida. Um total de 20% das amostras foi repetido para calcular a confiabilidade intra-avaliador e o coeficiente Kappa de Cohen. O valor de Kappa foi de 0,89, o que indica excelente confiabilidade do avaliador.

## Extração das medidas acústicas

Para extração das medidas acústicas de média f0, desvio padrão, *jitter*, *shimmer* e *GNE*, utilizou-se a amostra da vogal /a/ sustentada. Em seguida, foi utilizado o *software* VoxMetria, versão 4.7, da CTS Informática, no módulo análise da qualidade vocal. Os seguintes comandos foram aplicados para obtenção da média de f0 e medidas de perturbação e ruído:

1. Seleção da opção “Qualidade vocal”;
2. Escolha do arquivo a ser analisado;
3. Seleção e eliminação dos dois segundos iniciais e finais da emissão da vogal, devido à ocorrência de maior irregularidade em tais trechos, preservando-se o tempo mínimo de três segundos para cada emissão.
4. Seleção da opção “Dados da análise vocal”: são disponibilizados os dados de média f0, desvio padrão, *jitter*, *shimmer* e *GNE*.

Para extração das demais medidas acústicas, utilizou-se o *software* de livre acesso Praat (versão 5.3.77h), desenvolvido por Paul Boersma e David Weenink, University of Amsterdam, Holanda<sup>(20)</sup>.

A aquisição do CPPS seguiu os parâmetros sugeridos por Maryn e Weenink<sup>(3)</sup>. Os procedimentos foram os seguintes:

1. Escolha do arquivo: *open*, *open long sound file*, *extract part*, *ok*. Após escolha do arquivo a ser analisado, selecionou-se a opção *Analyze Periodicity* e, em seguida, *To Power Cepstrogram*.
2. Na opção *To Power Cepstrogram*, deu-se continuidade com os seguintes parâmetros: *Pitch floor (Hz) = 60*, *Time Step (s) = 0,002*, *Maximum Frequency (Hz) = 5000* e *Pre-emphasis from (Hz) = 50*.
3. Foi selecionada a opção *Query*, em seguida *Get CPPS*; no *menu*, deu-se continuidade com os seguintes parâmetros: *Subtract tilt before smoothing* selecionado, *Time averaging window (s) = 0.01*, *Quefrequency-averaging window (s) = 0.001*, *Peak search pitch range (Hz) = 60-330*, *Tolerance (0-1) = 0.05*, *Interpolation = Parabolic*, *Tilt line quefrequency range (s) = 0.001-0.0 (=end)*, *Line type = Straight* e *Fit method = Robust*.
4. Os resultados desse procedimento foram as medidas CPPS, expressas em decibéis (dB).

Para extração manual das medidas de intensidade (dB) dos harmônicos H1 e H2, as seguintes etapas foram utilizadas:

1. Seleção da opção *Read from file*; escolha do arquivo a ser analisado;
2. Seleção da opção *View & edit*; selecionou-se o trecho central do sinal de voz;
3. Seleção da opção *sel*, em seguida *spectrum* e, por último, *view spectral slice*, resultando em uma imagem gráfica dos picos, devendo-se posicionar o cursor na ponta do primeiro pico (H1), obtendo-se os valores de intensidade (dB) do primeiro harmônico. Em seguida, foram verificados os valores do segundo pico (H2).

A extração da média do primeiro e do segundo formante das vogais /a/, /i/ e /u/ ocorreu a partir da seleção e da segmentação dos sons vocálicos em contextos CV. Para a extração da média dos formantes no Praat, selecionou-se a opção denominada de *Formant*, obtendo-se o valor numérico de F1, F2 e F3 expresso em Hertz (Hz).

A segmentação e a duração dos sons vocálicos em contextos CV (consoante-vogal) foram realizadas considerando como limite inicial da vogal o primeiro pico regular após a consoante e ao período de transição entre a consoante-vogal. Como limite final, foi considerado o último pico regular antes da consoante e do período de transição entre a vogal-consoante<sup>(12)</sup> estimando-se uma média de análise de duração de 0,13 segundos.

## Análise dos dados

A análise estatística foi realizada considerando a média e desvio padrão das medidas investigadas. Para as comparações intergrupo (MNV e MSNV), utilizou-se o teste t de Student para dados pareados e o teste não paramétrico de Wilcoxon para dados pareados. O nível de significância considerado durante a análise foi de 5%.

## RESULTADOS

Na Tabela 1, podem ser observadas as médias, o desvio padrão e a comparação das medidas perceptivo-auditivas e acústicas específicas da fonte glótica, entre o MNV e MSNV.

A análise comparativa entre os grupos evidenciou que existiram diferenças na avaliação perceptivo-auditiva nos parâmetros: EVAGG ( $p=0,0020$ ), EVAGR ( $p=0,0105$ ) e EVAGS ( $p=0,0059$ ). O MNV apresentou valores mais elevados desses parâmetros, em relação ao MSNV.

Com relação às medidas acústicas investigadas, observou-se que a única relacionada à fonte glótica que apresentou diferença entre os grupos foi o *shimmer* ( $p=0,0449$ ), com presença de médias maiores no MNV.

A Tabela 2 descreve a caracterização das médias, do desvio padrão e da comparação das medidas acústicas específicas do filtro vocal entre o MNV e MSNV.

Quanto às medidas referentes aos ajustes supraglóticos, os valores de F1 para as vogais /a/ ( $p=0,0145$ ) e /u/ ( $p=0,0007$ ) e os valores de F2 para a vogal /a/ ( $p=0,0284$ ) apresentaram diferença entre os grupos, com valores mais elevados no MNV.

## DISCUSSÃO

A disseminação de estudos dos parâmetros de avaliação vocal resultou na investigação de diferentes medidas de captação do sinal de voz<sup>(7-9,21)</sup>. *Insights* baseados na teoria fonte-filtro,

associada a teorias não lineares, assim como medidas objetivas e subjetivas humanas efetivam e asseguram o diagnóstico e, conseqüentemente, a intervenção, na clínica vocal. Dessa forma, o olhar holístico do clínico e pesquisador deve estar voltado para os mecanismos fisiológicos, acústicos e perceptivos das particularidades da produção do som.

Assim, entende-se que o mecanismo acústico de produção da fala envolve ajustes, tanto na fonte glótica, quanto no filtro (ajustes realizados pelo trato vocal). Desse modo, buscou-se investigar o processo da relação fonte-filtro a partir do desenvolvimento de diversos tipos de medidas de análise perceptivo-auditivas e acústicas, visto que as alterações da voz geralmente perturbam o som de diferentes maneiras<sup>(13,22,23)</sup>.

Observou-se que mulheres com nódulos vocais apresentaram maior grau geral de desvio vocal e, conseqüentemente, maior grau de rugosidade e sopro, em relação a mulheres sem nódulos vocais (Tabela 1). De modo geral, sabe-se que a presença de lesão de massa nas pregas vocais pode ocasionar desvio da qualidade vocal, podendo ser justificado pela presença de irregularidade vibratória, e/ou aparecimento de fendas glóticas em detrimento ao fechamento glótico incompleto durante a fonação, resultando na presença de rugosidade e sopro durante a produção do som<sup>(1)</sup>.

Em uma perspectiva de intervenção com uso de terapia vocal ressoante, um estudo<sup>(21)</sup> avaliou medidas objetivas e subjetivas em 26 mulheres com nódulos vocais e 30 mulheres vocalmente saudáveis. Elas apresentaram diferenças significativas com relação às medidas de *f0*, *jitter*, *shimmer*, TMF, NHR (*noise-to-harmonic ratio*), antes da intervenção entre os grupos, com

**Tabela 1.** Comparação das medidas perceptivo-auditivas e acústicas específicas da fonte glótica entre o grupo de mulheres com nódulos vocais e o grupo de mulheres sem nódulos vocais

Medidas perceptivo-auditivas/ acústicas	Grupo		valor de p
	MNV	MSNV	
	Média ± DP	Média ± DP	
EVAGG	61,79 ± 8,06	47,83 ± 11,79	0,0020*
EVAGR	55,13 ± 8,45	43,88 ± 10,96	0,0105*
EVAGS	58,29 ± 14,25	38,13 ± 17,78	0,0059*
EVAGT	30,92 ± 14,91	30,33 ± 11,48	0,9155
Média f0	205,19 ± 20,44	202,37 ± 16,37	0,7130
DP 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

Testes t de Student/Wilcoxon; \*Valores significativos ( $p<0,05$ )

**Legenda:** MNV = mulheres com nódulos vocais; MSNV = mulheres sem nódulos vocais; EVA = escala visual analógica; GG = grau geral; GR = grau rugosidade; GS = grau sopro; GT = grau tensão; f0 = frequência fundamental; DP = desvio padrão; GNE = *glottal-to-noise excitation ratio*; CPPS = *cepstral peak prominence-smoothed*

**Tabela 2.** Comparação das medidas acústicas específicas do filtro vocal entre o grupo de mulheres com nódulos vocais e o grupo de mulheres sem nódulos vocais

Medidas acústicas	Grupo		valor de p
	MNV	MSNV	
	Média ± DP	Média ± DP	
H1-H2(dB)	7,42 ± 5,91	4,94 ± 4,51	0,4697
F1 VG /a/	804,40±137,39	931,28±83,31	0,0145*
F1 VG /i/	353,08±47,66	377,32±41,67	0,3474
F1 VG /u/	369,01±57,37	465,06±50,97	0,0007*
F2 VG /a/	1377,63±120,70	1471,79±83,21	0,0284*
F2 VG /i/	2491,65±170,93	2578,58±137,92	0,3777
F2 VG /u/	725,32±225,65	726,55±78,93	0,4776

Teste t de Student /Wilcoxon; \*Valores significativos ( $p<0,05$ )

**Legenda:** MNV = mulheres com nódulos vocais; MSNV = mulheres sem nódulos vocais; H1-H2(dB) = diferença de amplitude dos dois primeiros harmônicos em decibéis; F1 = primeiro formante; F2 = segundo formante; VG = vogal

valores elevados das medidas de perturbação e ruído no grupo de mulheres com nódulos vocais e, após intervenção, apenas no grupo de mulheres com nódulos vocais, apresentando redução dos valores dessas medidas.

Notou-se, na avaliação perceptivo-auditiva das mulheres do MSNV, que os parâmetros grau geral, rugosidade e soproidade apresentaram-se com desvio de grau leve, o que pode ser justificado fisiologicamente pelo gênero feminino, uma vez que mulheres apresentam menores valores da proporção glótica, o que reflete maior tendência para o fechamento glótico incompleto, e também por fatores hormonais<sup>(24,25)</sup>, que influenciariam o resultado da qualidade vocal de mulheres sem queixa e sem alteração laríngea. Além disso, o exame utilizado na presente pesquisa, a videolaringoscopia, não é o mais recomendado para avaliação da flexibilidade da mucosa das pregas vocais, do tecido mucoso e das camadas mais profundas, que seria a videoestroboscopia<sup>(26)</sup>.

A videoestroboscopia é o padrão-ouro clínico; na análise, há sincronia entre uma fonte de luz com a frequência fundamental do falante. A imagem resultante fornece uma ilusão de vibração das pregas vocais em câmera lenta e esse mecanismo traz dados fundamentais para o diagnóstico laríngeo, no que diz respeito ao fechamento glótico, amplitude de vibração e onda mucosa<sup>(26)</sup>.

Por outro lado, vale salientar que o presente estudo não utilizou como critério de inclusão que MSNV tivesse ausência de desvio na qualidade vocal, porém, não deveria apresentar queixa vocal e nem lesão laríngea, visto que certos níveis de instabilidade e ausência de periodicidade podem ser considerados normais. Mesmo com a máxima atenção do indivíduo para sustentar uma emissão estável em altura e intensidade, instabilidades na qualidade vocal podem ocorrer<sup>(27)</sup>. Desse modo, tais achados demonstraram que, mesmo na ausência de uma alteração e queixa e com medidas acústicas dentro dos valores de normalidade, o indivíduo pode ter um desvio na qualidade vocal. Nessa perspectiva, a avaliação multidimensional da voz impulsiona cada vez mais o diagnóstico dos distúrbios da voz de maneira efetiva<sup>(7)</sup>.

Com relação à análise das medidas acústicas de perturbação e ruído, as mulheres com nódulos vocais diferiram das mulheres sem alteração laríngea apenas quanto à medida de *shimmer*. O MNV apresentou maior valor quando comparado com o MSNV. Sabe-se que o *shimmer* é uma medida relacionada à resistência glótica e à manutenção da fase fechada dos ciclos glóticos. Sendo assim, são esperados maiores valores dessa medida em mulheres com nódulos vocais, em decorrência do fechamento glótico incompleto, comumente encontrado nesse tipo de lesão laríngea<sup>(28)</sup>.

A partir da análise das medidas formânticas, observou-se que os valores de F1 da vogal /a/ e /u/ e F2 da vogal /a/ apresentaram-se reduzidos no MNV, em relação ao MSNV. A diminuição nos valores de F1 para as vogais /a/ e /u/ e de F2 da vogal /a/ no MNV podem indicar posição do complexo oromandibular mais elevado e posteriorização da língua durante a fonação, respectivamente, uma vez que o valor de F1 está diretamente relacionado à postura desse complexo e F2 está relacionado à posição da língua no sentido anteroposterior<sup>(12)</sup>.

As diferenças encontradas nos valores das medidas formânticas entre os grupos deste estudo podem reforçar que há uma relação bidirecional entre ajustes supraglóticos e glóticos. Essa diferença inferida da posição da língua entre os grupos sugere que a posição e amplitude de movimentação dos articuladores do trato vocal podem estar modificadas durante a fala, em indivíduos com distúrbio de voz<sup>(6,8)</sup>.

A fonação associada ao corpo da língua elevado resulta em uma qualidade vocal soproosa. Essa afirmação se relaciona

com os resultados encontrados, porém, a grande questão para direcionamentos futuros é tentar entender se esses ajustes acontecem por efeito da coocorrência ou como consequência de um distúrbio de voz<sup>(8)</sup>.

Nessa perspectiva, um estudo<sup>(29)</sup> investigou a qualidade e a dinâmica vocal de 25 professoras com queixas ou manifestações de distúrbios da voz e alteração fonológica. Os mecanismos de sobrecarga do aparelho fonador provocaram ajustes na qualidade vocal, como hiperfunção laríngea e voz áspera, além de modificações nos ajustes de todo trato vocal: laringe elevada, hiperfunção do trato vocal, mandíbula fechada, constrição faríngea, corpo de língua elevado e escape de ar.

Segundo outro estudo<sup>(10)</sup>, ajustes habituais do trato vocal de mulheres disfônicas e não disfônicas são diferentes em posição de repouso e durante a fonação, uma vez que o esforço vocal pode mudar a posição dos articuladores do trato vocal de pacientes com nódulos vocais. Os autores investigaram os ajustes do trato vocal nessa população antes e após o exercício de ressonância com tubo flexível na água, através da ressonância magnética. Os ajustes supraglóticos encontrados na população com disфонia foram mandíbula fechada, lábios arredondados e corpo de língua retraído<sup>(10)</sup>.

A análise deste estudo reforça que o funcionamento das pregas vocais não é independente do trato vocal e que os ajustes habituais do trato vocal em mulheres com e sem nódulos vocais são diferentes na fonação<sup>(10,11)</sup>. Dessa forma, têm-se nortes metodológicos-avaliativos de grande valia para reabilitação e monitoramento na clínica vocal.

Dessa forma, o presente estudo mostrou que o entendimento da fisiologia da produção da voz está correlacionado com a avaliação eficaz dos distúrbios de voz, uma vez que o sinal acústico é resultado da interação entre informações somatossensoriais e motoras<sup>(30)</sup>. Assim, são imprescindíveis informações auditivas e acústicas da fonte glótica e do trato vocal.

## LIMITAÇÕES E DIREÇÕES FUTURAS

Alguns fatores limitantes deste estudo se referem ao tamanho da amostra presente em cada subgrupo, o número reduzido de juízes no processo de avaliação da qualidade vocal, assim como a ausência de realização de um exame mais específico de diagnóstico, considerado padrão-ouro na avaliação laríngea, a videoestroboscopia.

Para estudos posteriores, seria interessante realizar a videoestroboscopia, para efetividade do diagnóstico laríngeo, adquirindo imagens precisas da fisiologia das pregas vocais, levando em consideração os ciclos glóticos e, sobretudo, aumentar o tamanho da amostra para trazer robustez ao estudo.

## CONCLUSÃO

Mulheres com nódulos apresentaram vozes mais desviadas e com maior grau de rugosidade e soproidade, além de maiores valores de *shimmer* em relação às mulheres sem nódulos vocais. Quanto às medidas formânticas, mulheres com nódulos vocais demonstraram valores mais reduzidos de F1 da vogal /a/ e /u/ e F2 da vogal /a/ em relação às mulheres sem nódulos vocais. Nesse sentido, a análise das mulheres com nódulos vocais pode inferir uma possível redução na amplitude dos articuladores, com postura de mandíbula elevada e língua posteriorizada com reajustes na faringe.

## REFERÊNCIAS

1. Oliveira RCCD, Gama ACC, Genilhu 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 MFB, 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 Oct;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](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](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 preliminar. *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 Bodd 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 perceptivo-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, Deliyiski 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 disфония. *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 Feb;149(2):1147-61. <http://dx.doi.org/10.1121/1.0003440>. PMID:33639824.