

Leonardo Wanderley Lopes<sup>1</sup>  
 Jônatas do Nascimento Alves<sup>2</sup>  
 Deyverson da Silva Evangelista<sup>2</sup>  
 Fernanda Pereira França<sup>2</sup>  
 Vinícius Jefferson Dias Vieira<sup>2</sup>  
 Maria Fabiana Bonfim de Lima-Silva<sup>1</sup>  
 Leandro de Araújo Pernambuco<sup>1</sup>

# Acurácia das medidas acústicas tradicionais e formânticas na avaliação da qualidade vocal

## *Accuracy of traditional and formant acoustic measurements in the evaluation of vocal quality*

### Descritores

Voz  
 Acurácia  
 Análise Acústica  
 Qualidade Vocal  
 Distúrbios da Voz

### Keywords

Voice  
 Accuracy  
 Acoustic Analysis  
 Vocal Quality  
 Voice Disorders

### Endereço para correspondência:

Leonardo Wanderley Lopes  
 Departamento de Fonoaudiologia,  
 Centro de Ciências da Saúde,  
 Universidade Federal da Paraíba –  
 UFPB  
 Cidade Universitária, Campus I,  
 Bairro Castelo Branco, João Pessoa (PB),  
 Brasil, CEP: 58051-900.  
 E-mail: lwlopes@hotmail.com

Recebido em: Janeiro 08, 2018

Aceito em: Abril 09, 2018

### RESUMO

**Objetivo:** Investigar a acurácia das medidas acústicas, isoladas e combinadas, na discriminação da intensidade do desvio vocal (GG) e da qualidade vocal predominante (QVP) em pacientes disfônicos. **Método:** Participaram 302 pacientes do gênero feminino, com queixa vocal. A partir da vogal /ε/ sustentada, foram extraídas as medidas acústicas de média e desvio padrão (DP) da frequência fundamental ( $F_0$ ), o *jitter*, o *shimmer* e o *Glottal to noise excitation (GNE)* e a média dos três primeiros formantes (F1, F2, F3). A avaliação perceptivo-auditiva do GG e QVP foi realizada por três fonoaudiólogos especialistas em voz. **Resultados:** Isoladamente, apenas o GNE obteve desempenho satisfatório na discriminação do GG e da QVP. Houve uma melhora na classificação do GG e QVP com a combinação das medidas acústicas. A média de  $F_0$ , F2 e GNE (saudável × desvio leve a moderado), DP de  $F_0$ , F1 e F3 (leve a moderado × desvio moderado), *Jitter* e GNE (moderado × desvio intenso) foram as melhores combinações para discriminar o GG. As melhores combinações para discriminação da QVP foram média de  $F_0$ , *Shimmer* e GNE (saudável × rugosa), F3 e GNE (saudável × sopro), média de  $F_0$ , F3 e GNE (rugosa × tensa), média de  $F_0$ , F1 e GNE (soprosa × tensa). **Conclusão:** De forma isolada, o GNE mostrou-se o único parâmetro acústico capaz de discriminar o GG e a QVP. Houve um ganho no desempenho da classificação com a combinação das medidas acústicas tradicionais e formânticas, tanto para a discriminação do GG como da QVP.

### ABSTRACT

**Purpose:** Investigate the accuracy of isolated and combined acoustic measurements in the discrimination of voice deviation intensity (GD) and predominant voice quality (PVQ) in patients with dysphonia. **Methods:** A total of 302 female patients with voice complaints participated in the study. The sustained /ε/ vowel was used to extract the following acoustic measures: mean and standard deviation (SD) of fundamental frequency ( $F_0$ ), jitter, shimmer, glottal to noise excitation (GNE) ratio and the mean of the first three formants (F1, F2, and F3). Auditory-perceptual evaluation of GD and PVQ was conducted by three speech-language pathologists who were voice specialists. **Results:** In isolation, only GNE provided satisfactory performance when discriminating between GD and PVQ. Improvement in the classification of GD and PVQ was observed when the acoustic measures were combined. Mean  $F_0$ , F2, and GNE (healthy × mild-to-moderate deviation), the SDs of  $F_0$ , F1, and F3 (mild-to-moderate × moderate deviation), and mean jitter and GNE (moderate × intense deviation) were the best combinations for discriminating GD. The best combinations for discriminating PVQ were mean  $F_0$ , shimmer, and GNE (healthy × rough), F3 and GNE (healthy × breathy), mean  $F_0$ , F3, and GNE (rough × tense), and mean  $F_0$ , F1, and GNE (breathy × tense). **Conclusion:** In isolation, GNE proved to be the only acoustic parameter capable of discriminating between GD and PVQ. There was a gain in classification performance for discrimination of both GD and PVQ when traditional and formant acoustic measurements were combined.

Trabalho realizado no Departamento de Fonoaudiologia, Universidade Federal da Paraíba – UFPB - João Pessoa (PB), Brasil.

<sup>1</sup> Departamento de Fonoaudiologia, Universidade Federal da Paraíba – UFPB - João Pessoa (PB), Brasil.

<sup>2</sup> Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba – UFPB - João Pessoa (PB), Brasil.

**Fonte de financiamento:** Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Processo nº 480168/2013-0.

**Conflito de interesses:** nada a declarar.



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

## INTRODUÇÃO

A voz é essencialmente um fenômeno multidimensional, que envolve aspectos fisiológicos, perceptuais, aerodinâmicos, acústicos e emocionais. Assim, é necessário que a sua avaliação também siga esse princípio, contemplando e integrando essas dimensões em seu processo, a fim de alcançar uma visão global da disfonia<sup>(1)</sup>.

O objetivo da avaliação vocal é analisar a qualidade vocal, identificar se a voz é saudável ou não, diagnosticar a presença de um distúrbio, determinar um prognóstico, além de acompanhar a evolução do paciente durante a terapia de voz<sup>(2)</sup>. O processo de avaliação vocal, de modo geral, inclui procedimentos relacionados ao exame visual laríngeo, avaliação perceptivo-auditiva da voz, análise acústica, avaliação aerodinâmica e autoavaliação vocal<sup>(1)</sup>.

A análise perceptivo-auditiva é considerada o principal padrão de referência utilizado pelo fonoaudiólogo na avaliação vocal<sup>(2)</sup>. Ela é considerada um método subjetivo, que depende do julgamento do avaliador e possui caráter exclusivamente impressionístico<sup>(2,3)</sup>. Esse tipo de avaliação fornece informações acerca da caracterização da intensidade do desvio vocal, bem como da qualidade vocal predominante<sup>(4)</sup>.

A análise acústica é um procedimento mais objetivo, não invasivo e vem se tornando um método cada vez mais comum na clínica vocal. Na análise acústica tradicional, são utilizados dois tipos de medida, as de perturbação (*jitter* e *shimmer*) e de ruído. O *jitter* indica a variabilidade da frequência fundamental a curto prazo, medida entre ciclos glóticos vizinhos. O *shimmer* corresponde à variabilidade da amplitude da onda sonora a curto prazo. O *Glottal to Noise Excitation (GNE)* mede o ruído adicional no sinal sonoro, independentemente do ruído modulado pelo mecanismo glótico, indicando a origem do sinal vocal, se ele é proveniente da vibração das pregas vocais ou da corrente de ar turbulenta gerada no trato vocal. Sendo assim, as medidas de perturbação e ruído possuem um enfoque voltado à fonte glótica<sup>(3-5)</sup>.

Além dessas medidas, existem aquelas que estão relacionadas à ressonância da onda sonora no trato vocal, modificada pelas diferentes configurações de posicionamento das estruturas do trato vocal e volume das cavidades de ressonância durante a produção vocal. Tais medidas são chamadas de formantes e correspondem a concentrações de energia ao longo do trato vocal<sup>(3-6)</sup>.

O trato vocal possui uma configuração tridimensional e o som que é produzido na glote é modificado através do posicionamento das estruturas como a laringe, o palato mole, a língua, os lábios e a mandíbula. As frequências do sinal glótico que são reforçadas pelo trato vocal supraglótico são designadas formantes, de modo que esse tipo de análise fornece informações acerca dos ajustes que estão sendo realizados pelo trato vocal supraglótico<sup>(6-10)</sup>.

As modificações no posicionamento dos articuladores e no volume das cavidades de ressonância determinam os valores dos formantes<sup>(6-8,11)</sup>. Sendo assim, o aumento do primeiro formante (F1), por exemplo, está relacionado com um ajuste de mandíbula abaixada, abaixamento anterior da língua e estreitamento da faringe. Um ajuste de língua anteriorizada e língua posteriormente abaixada geram um aumento do segundo

formante (F2). Quanto ao terceiro formante (F3), a formação de uma menor cavidade imediatamente atrás dos incisivos pode elevar seus valores<sup>(6-8,10,11)</sup>.

Nesse contexto, há uma forte interação entre a fonte produtora do som (glote) e o filtro. O *feedback* da pressão encontrada pela onda sonora no trato vocal modifica o fluxo de ar glótico e o modo de vibração das pregas vocais<sup>(12)</sup>.

Além disso, alguns estudos<sup>(8-10,13-15)</sup> observaram que pacientes com distúrbio de voz realizam ajustes não apenas na glote mas também na supraglote. Esses pacientes apresentam valores dos formantes (F1, F2, F3) rebaixados em comparação a indivíduos sem distúrbio de voz<sup>(10,13,15)</sup>.

Sendo assim, esses ajustes podem estar relacionados à gênese, à manutenção ou em ocorrência com os distúrbios de voz<sup>(11,13)</sup>. Tais ajustes não são necessariamente avaliados pelas medidas acústicas tradicionais, visto que estão focadas na fonte glótica<sup>(16)</sup>.

Salienta-se que a análise acústica não substitui a análise perceptivo-auditiva, mas atua como integradora entre os níveis auditivos e fisiológicos<sup>(6-8)</sup>. Por sua vez, a combinação das medidas acústicas e perceptivo-auditivas aumenta a precisão para determinar a presença ou ausência de um distúrbio de voz, bem como a intensidade do desvio presente<sup>(17,18)</sup>.

Por esse motivo, é importante investigar se a combinação de medidas relacionadas à fonte (perturbação e ruído) e ao filtro (medidas formânticas) possibilita uma melhor classificação dos sinais vocais quanto à intensidade do desvio e qualidade vocal predominante.

Sendo assim, esta pesquisa tem por objetivo investigar a acurácia das medidas acústicas tradicionais e formânticas, isoladas e combinadas, na discriminação da intensidade do desvio vocal e da qualidade vocal predominante em pacientes disfônicos. Para a realização deste estudo, parte-se das hipóteses de que a combinação das medidas acústicas tradicionais e formânticas melhora a discriminação quanto à intensidade do desvio vocal; e que a combinação das medidas acústicas tradicionais e formânticas pode melhorar a discriminação entre diferentes qualidades vocais predominantes.

## MÉTODO

### Desenho do estudo

Este é um estudo descritivo, transversal e observacional, avaliado e aprovado pelo comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal da Paraíba (UFPB), com o parecer de número 52492/12. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido, autorizando a pesquisa

### Amostra

Participaram desta pesquisa pacientes atendidos no Laboratório de Voz do Departamento de Fonoaudiologia (UFPB), no período compreendido entre abril de 2012 e julho de 2015. Foram considerados como critérios de elegibilidade para participação nesta pesquisa:

- Ser do gênero feminino, considerando a relação entre essa variável e a medida da média de  $F_0$ , que está associada às características anatômicas das pregas vocais, desiguais entre homens e mulheres adultos<sup>(16)</sup>. Além disso, há uma maior prevalência de distúrbios da voz nessa população<sup>(19)</sup>;
- Idade superior a 18 anos e inferior a 65 anos, evitando, assim, o período da muda vocal e da presbifonia, respectivamente;
- Apresentar queixa vocal, respondendo de forma positiva à seguinte pergunta: “Você considera que você tem um problema de voz atualmente ou nos últimos seis meses?”;
- Ter realizado exame visual laríngeo e possuir laudo otorrinolaringológico.

Assim, do total de 530 pacientes avaliados no referido laboratório, 96 eram do gênero masculino, 75 tinham menos de 18 anos ou mais de 65 anos, e 57 indivíduos não apresentavam queixa vocal. Logo, foram excluídos 228 indivíduos que não se ajustaram aos critérios de elegibilidade, obtendo-se uma amostra final de 302 pacientes com média de idade de  $39,25 \pm 12,63$  anos. Nenhum paciente apresentou comprometimento neurológico ou cognitivo que impediu a gravação da voz.

Todos os pacientes da amostra apresentaram laudo laríngeo no momento da coleta de dados, conforme descrito a seguir: 78 (25,85%) pacientes com nódulos vocais, 63 (20,86%) sem alteração estrutural ou funcional na laringe, 41 (13,57%) com cisto vocal, 35 (11,60%) com hiperemia secundária a refluxo laringofaríngeo, 24 (7,94%) com fenda triangular médio-posterior, 24 (7,94%) com pólipos de prega vocal, 18 (5,96%) com paralisia unilateral de prega vocal, 11 (3,64%) com sulco vocal e 8 (2,64%) com edema de Reinke.

## Procedimentos

Toda a coleta de dados desta pesquisa foi realizada no Laboratório de Voz do Departamento de Fonoaudiologia (UFPB), durante a sessão inicial de avaliação vocal. Durante essa sessão, os pacientes foram avaliados por meio de uma ficha contendo perguntas relacionadas às informações pessoais e queixa vocal, responderam questionários de autoavaliação vocal e submeteram-se à gravação de tarefas de fala.

Para esta pesquisa, foram utilizados apenas os dados de identificação pessoal, queixa vocal e a amostra da vogal sustentada, conforme descrito posteriormente.

As vozes foram coletadas em cabine de gravação com tratamento acústico e ruído inferior a 50 dB NPS, com taxa de amostragem de 44000 Hz, com 16 bits por amostra e distância de 10 cm entre o microfone e os lábios do paciente. Utilizou-se o *software Fonoview*, versão 4.5, da CTS Informática, *desktop Del all-in-one*, microfone cardioide unidirecional, da marca *Senheiser*, modelo E-835, localizado em um pedestal e acoplado a um pré-amplificador Behringer, modelo U-Phoria UMC 204.

Para a coleta das vozes, o paciente permaneceu em pé, situando-se o pedestal à sua frente, de acordo com a distância preconizada entre a boca e o microfone. O paciente recebeu as instruções sobre a coleta de voz e, logo em seguida, procedeu-se com o registro. Durante a gravação, o paciente foi solicitado

a emitir a vogal /E/ sustentada, em frequência e intensidade autorreferida como confortável e habitual. A vogal /E/ foi selecionada para este estudo, pois é uma vogal oral, aberta, não arredondada e é considerada a vogal com a posição mais média no Português Brasileiro, o que permite uma posição mais neutra e intermediária do trato vocal. Além disso, é a vogal mais comumente utilizada para avaliação da qualidade vocal na realidade brasileira<sup>(20)</sup>.

Posteriormente, as vozes foram editadas no *software SoundForge* versão 10.0, sendo eliminados os dois segundos iniciais e finais da emissão da vogal, devido a maior irregularidade nesses trechos, preservando-se o tempo mínimo de três segundos para cada emissão. Para a avaliação perceptivo-auditiva, os sinais foram normalizados no controle “normalize” do *SoundForge*, no modo *peaklevel*, a fim de obter uma padronização na saída de áudio entre -6 e 6dB.

A extração das medidas acústicas da frequência fundamental (média e desvio padrão), do *jitter*, do *shimmer* e do *glottal to noise excitation (GNE)* foi realizada manualmente no *software VoxMetria*, versão 4.7h, da CTS Informática (Pato Branco, Paraná, Brazil), no módulo análise de qualidade vocal. No *software* utilizado, os valores de referência para os parâmetros de *jitter*, *shimmer* e *GNE* são 0,6, 6,5 e 0,5%, respectivamente. Valores maiores que os citados para o *jitter* e *shimmer* são considerados desviados, enquanto valores menores que o citado para o *GNE* podem ser considerados desviados.

Utilizou-se o *software Praat*, versão 5.3.77h, para a extração das medidas formânticas, a partir da representação da vogal em um espectrograma de banda larga, contendo os três primeiros formantes (F1, F2 e F3). Devido ao grande número de estimações envolvidas, foi utilizado um *script* (ferramenta que extrai automaticamente, de forma padronizada, as medidas paramétricas investigadas), o que viabilizou a otimização do tempo de processamento e evitou possíveis erros de manuseio na sequência dos procedimentos de estimação. Foram extraídos as médias e desvio padrão das frequências dos formantes para cada amostra. Na sequência, todos os valores foram conferidos e não houve identificação de *outliers*.

A avaliação perceptivo-auditiva foi realizada de modo independente por três fonoaudiólogos especialistas em voz, com mais de 10 anos de experiência nesse tipo de análise. Utilizou-se a Escala Analógica Visual (EAV) de 0 a 100 mm<sup>(21)</sup>, para a avaliação da intensidade do desvio vocal (GG) da vogal sustentada. A marcação mais próxima do 0 representa menor desvio, e quanto mais próxima do 100, maior o desvio vocal.

Antes da avaliação perceptivo-auditiva, foram utilizados oito estímulos-âncora da vogal sustentada /E/ para treinamento dos juízes, contendo duas amostras de indivíduos com variabilidade normal da qualidade vocal (VNQV), duas amostras de indivíduos com desvio vocal de leve a moderado, duas amostras de indivíduos com desvio vocal moderado e duas amostras de indivíduos com desvio vocal intenso. Todos os arquivos apresentados eram de vozes femininas. Os juízes foram orientados a escutar os estímulos-âncora imediatamente antes da análise das vozes desta pesquisa. Todas as amostras selecionadas para esse treinamento foram previamente analisadas por fonoaudiólogos com experiência em análise vocal e rotineiramente utilizadas

para treinamento perceptivo-auditivo e como estímulo-âncora no Laboratório onde esta pesquisa foi realizada.

A sessão de avaliação perceptiva ocorreu em ambiente silencioso. Inicialmente, cada juiz foi orientado de que as vozes deveriam ser consideradas com VNQV quando fossem socialmente aceitáveis, produzidas de forma natural, sem esforço, ruído ou condição instável durante a emissão. Eles também foram instruídos de que a rugosidade corresponderia à presença de irregularidade vibratória, a soproidade estaria relacionada ao escape de ar audível na emissão e a tensão corresponderia à percepção de esforço vocal ao longo da emissão.

Os parâmetros perceptivo-auditivos de rugosidade, soproidade e tensão foram escolhidos para caracterização dos sinais neste estudo por serem utilizados universalmente para caracterizar o desvio de qualidade vocal<sup>(2)</sup> e por possuírem correlatos conhecidos no plano fisiológico e acústico.

Para avaliação, cada emissão da vogal sustentada foi apresentada por três vezes através de caixa de som, em intensidade confortável autorreferida pelos avaliadores. Em seguida, os juízes fizeram a identificação da presença ou ausência de desvio vocal, da qualidade vocal predominante nas vozes desviadas (rugosa, sopro ou tensa) e, por fim, o julgamento da intensidade do desvio vocal.

A EAV foi posteriormente convertida em uma escala numérica com valores de 1 a 4, sendo o grau 1 referente aos indivíduos com VNQV (0-35,5 mm), grau 2 para sujeitos com desvio de leve a moderado (35,6-50,5 mm), grau 3 para desvio moderado (50,6-90,5) e grau 4 para desvio intenso (>90,5 mm)<sup>(22)</sup>.

Ao final da avaliação perceptivo-auditiva, 10% das amostras foram repetidas aleatoriamente, para a análise da confiabilidade da avaliação dos juízes por meio do Coeficiente Kappa de Cohen. Os resultados da análise perceptivo-auditiva do juiz com maior confiabilidade (coeficiente Kappa de 0,79) foram selecionados para utilização nesta pesquisa. Os outros dois juízes apresentam valores de Kappa < 0,70.

A partir dos resultados da análise perceptivo-auditiva, os pacientes foram categorizados em dois grupos: 33 pacientes com VNQV ( $GG \leq 35,5$  mm) e 269 pacientes com desvio da qualidade vocal ( $GG \geq 35,6$  mm). Entre os pacientes com desvio da qualidade vocal, 150 foram classificados com grau leve a moderado ( $35,6 \leq GG \leq 50,5$  mm), 112 com grau moderado ( $50,6 \leq GG \leq 90,5$  mm) e 7 indivíduos com desvio intenso ( $GG > 90,5$  mm). Dos 269 pacientes com desvio da qualidade vocal, 135 (50,18%) apresentaram qualidade vocal predominantemente rugosa, 95 (35,31%) com predomínio de soproidade e 39 (14,49%) com predomínio de tensão.

Os 33 pacientes com VNQV apresentavam queixa vocal e ausência de alteração estrutural e funcional na laringe, conforme laudo otorrinolaringológico. Entre os 269 pacientes com desvio da qualidade vocal, todos apresentavam queixa vocal, sendo que 30 receberam diagnóstico médico de ausência de alteração estrutural e funcional na laringe, e 239 receberam diagnóstico de alteração laringea, conforme descrito anteriormente.

Tal caracterização da amostra está em consonância com a literatura, visto que não há uma relação direta entre a presença de queixa vocal, a presença de desvio de qualidade vocal e a presença de alteração laringea<sup>(5)</sup>. Sendo assim, considerando-se que o objetivo do presente não é avaliar os parâmetros acústicos em função da presença vs. ausência de um distúrbio de voz,

mas elucidar relações entre os parâmetros perceptivo-auditivos e as medidas acústicas na avaliação da intensidade e do tipo de desvio vocal, optou-se por não excluir os indivíduos com queixa vocal e ausência de alteração laringea. Esses critérios permitiriam reforçar a validade interna do estudo e garantiriam que a variável independente (avaliação auditiva-perceptiva) é a única ou mais provável explicação para os resultados encontrados para a variável dependente (parâmetros acústicos).

## Análise de dados

Realizou-se análise estatística descritiva para todas as variáveis analisadas, incluindo os valores de média e desvio padrão. Para a classificação dos sinais em função do GG e da qualidade vocal predominante, foi realizada análise discriminante quadrática (QDA), utilizando como método auxiliar a validação cruzada com  $K$  subconjuntos.

A QDA foi selecionada para este estudo porque permite encontrar as variáveis individuais e combinadas que melhor discriminam grupos pré-estabelecidos (GG e qualidade vocal predominante). Considerando-se que foram analisadas oito medidas acústicas para a análise da combinação de medidas, elas foram combinadas 2 a 2, 3 a 3, 4 a 4, até 8 a 8.

Na validação cruzada pelo método de  $K$  subconjuntos, a classificação foi realizada dez vezes, variando no conjunto de dados, o que é utilizado para treino e para teste, sem repetição, de modo que seja obtida mais precisão nos resultados<sup>(22)</sup>. Dessa forma, os sinais com diferentes GG e qualidades vocais predominantes foram divididos aleatoriamente em subconjuntos, com um mínimo de 10 sinais em cada subconjunto, considerando-se essa quantidade mínima de sinais, possibilita as melhores estimativas de erro. Os sinais com desvios intensos foram excluídos da análise, visto que não satisfaziam à condição de possuir um número mínimo de 10 sinais.

Esses subconjuntos foram confrontados por meio do procedimento de validação cruzada e, a cada uma das iterações entre os subconjuntos, foram obtidas medidas de desempenho (acurácia, sensibilidade e especificidade) do classificador em discriminar o GG ou a qualidade vocal predominante. Ao final de todas as iterações dos subconjuntos, foram extraídos os valores da média e do desvio padrão dos subconjuntos formados, utilizados para interpretação dos dados finais do classificador.

As medidas de acurácia, sensibilidade e especificidade foram utilizadas para avaliar o desempenho do classificador. De modo geral, a interpretação das medidas de sensibilidade e especificidade é mais evidente quando os grupos a serem comparados pertencem a uma classe saudável (sem alterações) ou patológica (com alterações)<sup>(23)</sup>. Sendo assim, quando se realiza análise discriminante entre classes com alteração, como o que foi realizado nesta pesquisa (quando comparadas diferentes intensidades do desvio e qualidade vocal predominante), faz-se necessário determinar, no classificador utilizado, o grupo de sinais que terá sua correta classificação medida pela sensibilidade e o grupo que terá sua correta classificação medida pela especificidade.

Dessa forma, adotou-se como procedimento padrão que a primeira condição apresentada em cada tabela corresponderia

ao sinal que seria classificado corretamente pela especificidade, enquanto a segunda condição seria classificada corretamente pela sensibilidade (Quadro 1).

O desempenho de classificação levou em conta os sinais com diferentes GG e com diferentes predomínios de qualidade vocal. Também se considerou o poder individual de cada uma das medidas acústicas consideradas e as combinações possíveis entre essas medidas, identificando-se aquelas que forneceram

as melhores taxas de classificação entre os sinais de voz nas condições preconizadas nesta pesquisa.

Considerando-se que a acurácia pode ser classificada em excelente (>90%), boa (80-90%), aceitável (70-80%), pobre (60-70%) e sem capacidade de discriminação aceitável (<60%)<sup>(23)</sup>, foram analisadas apenas as classificações com desempenho superior a 70%. A análise discriminante (acurácia, sensibilidade e especificidade) foi realizada utilizando-se o *software* Matlab®, versão 7.9.

**Quadro 1.** Casos de discriminação e suas respectivas medidas de sensibilidade e especificidade

Casos de discriminação	Sensibilidade	Especificidade
<b>VNQV × Leve a moderado</b>	Taxa de correta classificação dos sinais com desvio de leve a moderado	Taxa de correta classificação dos sinais com VNQV
<b>Leve × Moderado</b>	Taxa de correta classificação dos sinais com desvio moderado	Taxa de correta classificação dos sinais com desvio leve
<b>VNQV × Soprosa</b>	Taxa de correta classificação dos sinais com qualidade vocal soprosa	Taxa de correta classificação dos sinais com VNQV
<b>VNQV × Rugosa</b>	Taxa de correta classificação dos sinais com qualidade vocal rugosa	Taxa de correta classificação dos sinais com VNQV
<b>Soprosa × Tensa</b>	Taxa de correta classificação dos sinais com qualidade vocal tensa	Taxa de correta classificação dos sinais com qualidade vocal soprosa
<b>Rugosa × Soprosa</b>	Taxa de correta classificação dos sinais com qualidade vocal soprosa	Taxa de correta classificação dos sinais com qualidade vocal rugosa
<b>Rugosa × Tensa</b>	Taxa de correta classificação dos sinais com qualidade vocal tensa	Taxa de correta classificação dos sinais com qualidade vocal rugosa

**Legenda:** VNQV = variabilidade normal da qualidade vocal

## RESULTADOS

Nas Tabelas 1 e 2, são apresentados os dados relacionados à média e desvio padrão das medidas acústicas em função do GG e da qualidade vocal predominante, respectivamente. Tais dados

não serão analisados isoladamente, mas em conjunto com o desempenho das classificações utilizadas.

Inicialmente, realizou-se o teste de acurácia das medidas acústicas isoladas na discriminação do GG nos pacientes. Assim, encontrou-se melhor desempenho (70,95%, DP = 3,05),

**Tabela 1.** Média e desvio padrão das medidas acústicas em diferentes intensidades do desvio vocal

Medidas	INTENSIDADE DO DESVIO VOCAL					
	VNQV		Leve a moderado		Moderado	
	Média	DP	Média	DP	Média	DP
<b>Média F<sub>0</sub></b>	179,87	43,19	182,06	60,78	183,75	69,97
<b>DP F<sub>0</sub></b>	7,28	15,82	10,80	21,58	24,78	37,83
<b>F1</b>	599,35	143,10	592,95	127,39	585,63	145,74
<b>F2</b>	2014,08	232,43	2018,87	213,42	2033,47	231,94
<b>F3</b>	2812,05	216,47	2843,75	245,48	2888,86	219,74
<b>Jitter</b>	0,25	0,50	0,50	1,22	1,87	2,86
<b>Shimmer</b>	3,91	3,09	5,32	4,29	9,11	9,11
<b>GNE</b>	0,90	0,119	0,83	0,19	0,68	0,24

**Legenda:** F<sub>0</sub> = frequência fundamental; DP = desvio padrão; F1 = primeiro formante; F2 = segundo formante; F3 = terceiro formante; GNE = glottal to noise excitation; VNQV = variabilidade normal da qualidade vocal

**Tabela 2.** Média e desvio padrão das medidas acústicas em função da qualidade vocal predominante

Medidas	QUALIDADE VOCAL PREDOMINANTE							
	VNQV		Soprosa		Rugosa		Tensa	
	Média	DP	Média	DP	Média	DP	Média	DP
<b>Média F<sub>0</sub></b>	181,02	42,62	171,62	64,52	196,75	68,82	203,83	64,81
<b>DP F<sub>0</sub></b>	7,36	15,95	18,86	33,70	13,99	22,00	19,63	35,42
<b>F1</b>	597,42	143,545	581,58	108,75	586,78	155,31	672,93	188,06
<b>F2</b>	2011,29	233,41	2005,97	229,23	2046,27	207,30	2063,18	210,52
<b>F3</b>	2808,08	216,124	2837,71	254,40	2911,63	202,524	2910,97	230,60
<b>Jitter</b>	0,25	,509	1,44	2,46	1,02	2,75	1,42	3,27
<b>Shimmer</b>	3,91	3,09	8,71	7,45	6,33	5,76	8,827	11,46
<b>GNE</b>	0,90	0,11	0,76	0,22	0,69	0,27	0,83	0,20

**Legenda:** F<sub>0</sub> = frequência fundamental; DP = desvio padrão; F1 = primeiro formante; F2 = segundo formante; F3 = terceiro formante; GNE = glottal to noise excitation; VNQV = variabilidade normal da qualidade vocal

para a medida *GNE*, com a sensibilidade de 86,67%±5,44 e especificidade de 55,83%±5,13 (Tabela 3).

Ao investigar o poder discriminatório das medidas acústicas combinadas na classificação do GG na amostra investigada, encontrou-se maior acurácia nas seguintes combinações: média de  $F_0$ , F2 e *GNE* (75,24%±4,86) para distinção entre VNQV e desvio leve a moderado; DP de  $F_0$ , F1, F3, *jitter* e *GNE* (74,02%±3,26) para discriminação de desvio leve *versus* moderado (Tabela 3).

Na sequência, investigou-se a acurácia das medidas isoladas na discriminação da qualidade vocal predominante. O *GNE* apresentou o melhor desempenho na discriminação entre VNQV e rugosa (73,57%±5,56), entre VNQV e sopro (82,38%±3,73), além da discriminação entre sopro e tensa (71,43%±4,76) (Tabela 4).

Por fim, verificou-se o desempenho das medidas acústicas combinadas na discriminação da qualidade vocal. A média de  $F_0$ , *shimmer* e *GNE* (78,57%±4,21) constituíram a melhor combinação para discriminar VNQV e qualidade vocal rugosa. As medidas de F3 e *GNE* (84,05%±3,29) foram a melhor combinação para classificar VNQV e qualidade vocal sopro. A média de  $F_0$ , F3 e *GNE* (73,75%±3,75) foi selecionada como melhor combinação para discriminar entre vozes rugosas e tensas. A combinação da média de  $F_0$ , F1 e *GNE* (75,71%±6,41) apresentou a melhor *performance* para discriminar entre vozes sopro e tensas (Tabela 4).

## DISCUSSÃO

Este estudo investigou a acurácia das medidas acústicas tradicionais e formânticas, isoladas e combinadas, na discriminação do GG e qualidade vocal predominante em pacientes disfônicos. Foram levantadas duas hipóteses: 1) a combinação das medidas acústicas tradicionais e formânticas melhora a discriminação das vozes quanto ao GG; 2) a combinação das medidas acústicas tradicionais e formânticas pode melhorar a discriminação entre vozes com diferentes qualidades vocais predominantes. Assim, a sessão de discussão foi organizada para elucidar as conclusões encontradas no que diz respeito a essas hipóteses.

### Medidas acústicas tradicionais e formânticas na discriminação da intensidade do desvio vocal

Ao analisar as medidas acústicas isoladas, apenas o *GNE* apresentou desempenho aceitável (70,95%±3,05) na discriminação entre vozes com VNQV e vozes com desvio de leve a moderado, com maior sensibilidade (86,67%±5,44) para identificação correta dos sinais com desvio.

A medida de *GNE* apresentou-se mais rebaixada nos pacientes com desvio vocal de leve a moderado em relação aos indivíduos com VNQV. No entanto, em nenhum dos dois grupos essa medida apresentou valores inferiores a 0,5%, considerado ponto de corte para a presença de desvio nesse parâmetro. Por sua vez, na análise comparativa, pode-se inferir que os pacientes com desvio vocal de grau leve a moderado, possuem mais fluxo de

**Tabela 3.** Acurácia, sensibilidade e especificidade das melhores medidas acústicas isoladas e das melhores combinações das medidas acústicas na discriminação da intensidade do desvio vocal

Intensidade do desvio vocal	Medida isolada	Ac (%)	Sens (%)	Esp (%)
<b>VNQV x Leve a moderado</b>	<i>GNE</i>	70,95±3,05	86,67±5,44	55,83±5,13
Melhor Combinação				
<b>VNQV x Leve a moderado</b>	Média de $F_0$ , F2, <i>GNE</i>	75,24±4,86	84,17±5,34	67,50±7,90
<b>Leve a moderado x Moderado</b>	DP de $F_0$ , F1, F3, <i>Jitter</i> , <i>GNE</i>	74,02±3,26	87,62±2,51	56,14±6,28

**Legenda:** Ac = acurácia; Sens = sensibilidade; Esp = especificidade;  $F_0$  = frequência fundamental; DP = desvio padrão; F1 = primeiro formante; F3 = terceiro formante; *GNE* = *glottal to noise excitation*; VNQV = variabilidade normal da qualidade vocal

**Tabela 4.** Acurácia, sensibilidade e especificidade das melhores medidas acústicas isoladas e das melhores combinações das medidas acústicas na discriminação da qualidade vocal predominante

Qualidade Vocal Predominante	Medida isolada	Ac (%)	Sens (%)	Esp (%)
<b>VNQV x Rugosa</b>	<i>GNE</i>	73,57±5,56	88,33±4,84	59,17±11,00
<b>VNQV x Sopro</b>	<i>GNE</i>	82,38±3,73	87,50±5,16	78,33±7,88
<b>Sopro x Tensa</b>	<i>GNE</i>	71,43±4,76	57,50±8,82	81,67±4,08
Melhor Combinação				
<b>VNQV x Rugosa</b>	Média de $F_0$ , <i>Shimmer</i> , <i>GNE</i>	78,57±4,21	87,50±5,16	70,00±6,36
<b>VNQV x Sopro</b>	F3, <i>GNE</i>	84,05±3,29	90,00±5,09	77,50±7,03
<b>Rugosa x Tensa</b>	Média de $F_0$ , F3, <i>GNE</i>	73,75±3,75	60,83±6,34	84,17±5,75
<b>Sopro x Tensa</b>	Média de $F_0$ , F1, <i>GNE</i>	75,71±6,41	71,67±7,05	78,33±8,16

**Legenda:** Ac = acurácia; Sens = sensibilidade; Esp = especificidade;  $F_0$  = frequência fundamental; F1 = primeiro formante; F3 = terceiro formante; *GNE* = *glottal to noise excitation*; VNQV = variabilidade normal da qualidade vocal

ar não sonorizado entre as pregas vocais do que aqueles com VNQV<sup>(5,11)</sup>.

Uma pesquisa<sup>(4)</sup> realizada com 226 pacientes, sendo 53 indivíduos do grupo controle e 173 pacientes com desvio vocal, demonstrou que o GNE apresentou uma excelente acurácia (95%) para diferenciar vozes saudáveis e desviadas. Dessa forma, infere-se que o GNE poderia representar uma boa medida para avaliação vocal, justamente por mostrar uma maior discriminação entre vozes saudáveis e com desvio.

A partir da análise das medidas acústicas combinadas, a hipótese de que a combinação das medidas tradicionais e formânticas melhora o desempenho do classificador na discriminação do GG foi confirmada. Além do aumento dos valores de acurácia e especificidade, a combinação de medidas incluiu a discriminação entre desvio de grau leve a moderado e moderado, o que não ocorreu com as medidas isoladas. A combinação das medidas relacionadas à média de  $F_0$ , F2 e GNE obteve uma acurácia de 75,24%±4,86 na discriminação entre sinais com VNQV e desvio de leve a moderado. Pacientes com desvio leve a moderado apresentaram menores valores de GNE, e maiores valores de média de  $F_0$  e F2 em relação a pacientes com VNQV.

Valores menores de GNE podem indicar fechamento glótico ineficiente, maior ruído aditivo na voz e possível diminuição na intensidade<sup>(4,5,24)</sup>. Por sua vez, os dados do presente estudo quanto ao GNE estão sendo analisados comparativamente entre os grupos, visto que os valores não se encontraram abaixo do ponto de corte em nenhum dos grupos de sinais.

Os valores da média de  $F_0$  encontrados estariam ligados à presença de tensão longitudinal das pregas vocais, gerando um maior número de ciclos glóticos por segundo e, conseqüentemente, maior elevação de  $F_0$ <sup>(25)</sup>.

Valores aumentados de F2 estariam relacionados aos ajustes de anteriorização de língua<sup>(6-8)</sup>. Tal ajuste promove a elevação do complexo laríngeo e, por uma ação biomecânica, há maior tensão longitudinal nas pregas vocais, com conseqüente elevação da  $F_0$ , aumento do esforço fonatório e diminuição da projeção vocal<sup>(14,25)</sup>.

Uma pesquisa<sup>(26)</sup> analisou as medidas formânticas de vogais sustentadas, mostrando um aumento nos valores dessas medidas quando o complexo laríngeo estava elevado. Além disso, os valores de  $F_0$  diminuíam quando havia um aumento no comprimento do trato vocal (larínge mais baixa), da mesma forma, aumentavam quando havia diminuição do comprimento do trato vocal (larínge elevada).

A partir desses achados, infere-se que, em relação aos indivíduos sem desvio da qualidade vocal, pacientes com desvio de grau leve a moderado podem implementar ajustes supraglóticos a fim de compensar condições glóticas disfuncionais, com presença de maior fluxo de ar não sonorizado. Tais achados estão em consonância com outros estudos<sup>(8-10,13-15)</sup> que mostraram que pacientes disfônicos tendem a realizar ajustes no trato vocal como forma de compensar o seu problema vocal.

Por outro lado, pode-se questionar se o ajuste supraglótico pode estar relacionado à gênese do problema vocal nesses pacientes, uma vez que a elevação da laringe, com aumento da tensão longitudinal das pregas vocais, diminui a convexidade da curvatura da borda livre das pregas vocais, que é um dos

mecanismos responsáveis pelo aumento do fluxo aéreo translótico não sonorizado<sup>(27)</sup>.

De modo geral, a descrição e análise das medidas formânticas no grupo com desvio leve a moderado parece ser interessante para a compreensão dos ajustes supraglóticos realizados por esses pacientes, os quais podem ter implicações na evolução clínica do caso em terapia vocal.

A combinação das medidas de DP de  $F_0$ , F1, F3, *jitter* e GNE também obteve um desempenho aceitável (74,02%±3,26) na discriminação entre sinais com desvio de leve a moderado e com desvio moderado. As medidas de DP de  $F_0$ , F1, F3 e *jitter* foram maiores nos pacientes com desvio moderado, enquanto os valores de GNE foram menores nesses pacientes em relação aos indivíduos com desvio leve a moderado. Considerando-se os valores de referência para as medidas de GNE e *jitter*, apenas esse último apresentou valores acima do ponto de corte para ser considerado desviado.

Em termos fisiológicos, o desvio padrão de  $F_0$  está diretamente ligado à condição neuromuscular e à regularidade de vibração da mucosa das pregas vocais, assim, maiores valores do DP de  $F_0$  encontrados em pacientes com desvio moderado podem indicar instabilidade fonatória, maior irregularidade da vibração das pregas vocais, ocasionando desvio na produção vocal<sup>(24,25)</sup>.

O *jitter* avalia as perturbações na frequência entre ciclos vibratórios vizinhos<sup>(11,18)</sup>, sendo a medida reconhecida mais correlacionada ao GG<sup>(17)</sup> e sensível à presença de desvio vocal, o que justifica seu aumento nos indivíduos com desvio vocal moderado no presente estudo.

Assim, de acordo com esses dados, infere-se que pacientes com desvio vocal de grau moderado possuem uma maior irregularidade de vibração das pregas vocais e instabilidade fonatória (aumento do DP de  $F_0$ ), maior fluxo de ar não sonorizado, com mais ruído na voz (diminuição do GNE) e uma maior intensidade global do desvio vocal (aumento de *jitter*) em relação aos pacientes com desvio de leve a moderado.

A elevação dos valores de F1 está relacionada ao maior abaixamento do complexo oromandibular e ao estreitamento na orofaringe<sup>(6-8,10,11)</sup>. Esses ajustes supraglóticos citados podem ocorrer em compensação a condições glóticas disfuncionais, uma vez que uma maior abertura de mandíbula e o estreitamento da faringe podem ocasionar decréscimo na soprosonidade percebida auditivamente<sup>(27)</sup> e aumentar a intensidade vocal<sup>(8-10,17)</sup>. Por sua vez, a elevação de F1 também está associada ao esforço fonatório presente em pacientes com disfonia por tensão muscular<sup>(14)</sup>.

A hipótese de que as medidas acústicas tradicionais e formânticas combinadas podem melhorar a discriminação quanto ao GG foi confirmada. As informações parecem ter um caráter complementar, visto que as medidas formânticas, isoladamente, não apresentaram *performance* aceitável para discriminação dos casos estudados. Por sua vez, vale salientar que, na presente pesquisa, utilizou-se uma escala de avaliação perceptivo-auditiva focada na fonte glótica e, por isso, poder-se-ia esperar uma maior contribuição das medidas acústicas relacionadas à fonte glótica.

Entretanto, vozes mais desviadas parecem apresentar maiores ajustes supraglóticos, visto que os maiores valores encontrados na combinação das medidas estariam relacionados à sensibilidade, ou seja, indicando de forma correta os sinais mais desviados.

## Medidas acústicas tradicionais e formânticas na discriminação da qualidade vocal predominante

Ao analisar as medidas acústicas isoladas, apenas o *GNE* apresentou desempenho aceitável na discriminação entre vozes quanto à qualidade vocal predominante.

Quanto à discriminação entre vozes com VNQV e rugosas, encontrou-se uma acurácia de  $73,57 \pm 5,56\%$ , com maior sensibilidade ( $88,33\% \pm 4,84$ ) para identificação correta das vozes rugosas. Com relação à discriminação VNQV x sopro, encontrou-se acurácia de  $82,38\% \pm 3,73$ , com maior sensibilidade ( $87,50\% \pm 5,16$ ) para identificação correta de vozes soprosas. Já na discriminação sopro x tenso, encontrou-se uma acurácia de  $71,43 \pm 4,76\%$ , com maior especificidade ( $81,67\% \pm 4,08$ ) para identificação correta de vozes soprosas.

Mais uma vez, apenas a medida *GNE*, de forma isolada, mostrou valores aceitáveis na discriminação das diferentes qualidades vocais. Nesse contexto, o *GNE* mostrou-se importante principalmente para diferenciar vozes soprosas de outros tipos de vozes, provavelmente por estar relacionado diretamente com a origem do sinal vocal, se esse é proveniente da vibração das pregas vocais ou da corrente de ar turbulenta gerada no trato vocal<sup>(4,5)</sup>, fator esse que poderia explicar sua relação direta com esse parâmetro.

A hipótese de que as medidas acústicas tradicionais e formânticas combinadas podem melhorar a discriminação da qualidade vocal predominante foi confirmada, visto que a combinação dessas medidas melhorou o desempenho do classificador para a discriminação VNQV x rugosa, VNQV x sopro, sopro x tenso, além de proporcionar uma discriminação aceitável entre rugosa e tenso.

Para a discriminação VNQV x rugosa, a melhor combinação encontrada selecionou as medidas da média de  $F_0$ , *shimmer* e *GNE*, com acurácia de  $78,57\% \pm 4,21$  e maior sensibilidade ( $87,50 \pm 5,16\%$ ) para a identificação correta de vozes rugosas. Os valores da média de  $F_0$  e de *shimmer* mostram-se elevados na população com qualidade vocal rugosa, enquanto os valores de *GNE* encontraram-se diminuídos em relação às vozes com VNQN.

Em geral, é esperado que vozes rugosas apresentem valores de  $F_0$  rebaixados<sup>(18)</sup>. Porém, para justificar o aumento dessa medida na presente pesquisa, podemos inferir que os pacientes com vozes rugosas tiveram, possivelmente, tensão associada à emissão e que, por esse motivo, houve um aumento do  $F_0$ <sup>(2,14,28)</sup> em relação aos pacientes com VNQV.

O *shimmer* é uma medida relacionada à variabilidade de amplitude entre ciclos adjacentes, estando assim relacionada à irregularidade vibratória e à resistência glótica<sup>(4,29)</sup>. No plano perceptivo-auditivo, estudos anteriores mostram que o *shimmer* estaria relacionado à rugosidade<sup>(17,18)</sup>. Os valores de *shimmer*, na presente pesquisa, contribuíram na identificação correta das vozes rugosas. Pode-se destacar que, apesar de os valores de *shimmer* estarem mais desviados nas vozes com rugosidade, esses valores ainda estão dentro da normalidade considerando os valores de corte adotados.

Uma pesquisa<sup>(18)</sup> teve, entre os objetivos, analisar o poder discriminatório das medidas acústicas na classificação da intensidade

do desvio e na diferenciação do tipo de voz predominante. Participaram 186 pacientes disfônicos e as medidas utilizadas foram a frequência fundamental ( $F_0$ ), *jitter*, *shimmer* e *GNE*. Entre os resultados, o *shimmer* e o *GNE* mostram-se úteis na detecção de vozes rugosas e soprosas, respectivamente.

Os dados deste estudo<sup>(18)</sup>, mostram-se semelhantes aos achados na presente pesquisa, uma vez que o *shimmer* apresentou relação com o parâmetro de rugosidade e o *GNE*, apesar de aparecer em todas as combinações, pareceu ser mais sensível quando se trata de vozes com qualidade vocal sopro.

As medidas de  $F_3$  e *GNE* foram selecionadas como a melhor combinação para discriminar VNQV x Sopro ( $84,05\% \pm 3,29$ ), com uma elevada sensibilidade ( $90,00\% \pm 5,09$ ) para a identificação correta das vozes soprosas. Pacientes com vozes soprosas apresentaram maiores valores de  $F_3$  e valores mais rebaixados de *GNE*.

A frequência de  $F_3$  está relacionada com as duas cavidades estabelecidas pela posição da língua, ou seja, à cavidade atrás da constrição da língua e aquela à frente. A frequência de  $F_3$  pode ser influenciada também por ajuste de lábios, laringe e faringe, havendo uma tendência em diminuição com ajuste de labiodentalização, arredondamento labial e aumento com constrição próxima à faringe<sup>(3,10,11,20)</sup>. Dessa forma, pode-se inferir que pacientes com qualidade vocal predominantemente sopro apresentam maior constrição próxima à faringe e maior estiramento de lábios, provavelmente, como mecanismo compensatório para aumentar a intensidade vocal.

Os achados na pesquisa reforçam que a medida de *GNE* apresenta-se bastante relacionada com a qualidade vocal sopro<sup>(4,5,18,28)</sup>, sendo a única medida isolada com acurácia aceitável para discriminação entre sinais com VNQV e com sopro.

Para a discriminação entre vozes rugosas e tensas, a melhor combinação encontrada foi das medidas de média de  $F_0$ ,  $F_3$  e *GNE* ( $73,75 \pm 3,75\%$ ), com maior especificidade ( $84,17 \pm 5,75\%$ ) na identificação de vozes rugosas. A média de  $F_0$  foi menor em pacientes com rugosidade, do que aqueles com vozes tensas; encontrou-se  $F_3$  com maiores valores em pacientes com rugosidade; e o *GNE* mais elevado em pacientes com a qualidade vocal tenso.

Os achados sugerem que pacientes com qualidade vocal tenso podem ter uma maior tensão longitudinal nas pregas vocais, devido aos maiores valores na média de  $F_0$ , além disso parece que pacientes com rugosidade possuem uma menor cavidade no trato vocal, por conta do aumento de  $F_3$ <sup>(11,13)</sup> e que pacientes com qualidade vocal tenso parecem ter menos ruído na voz<sup>(4,5)</sup> do que pacientes com rugosidade, aspecto sugerido pelo valor de *GNE* menos desviado nas vozes tensas.

Essa categoria de discriminação, isto é, rugosa x tenso, apareceu apenas quando houve a combinação das medidas, não tendo nenhum valor individual aceitável. O que demonstra a importância de encontrar-se a melhor combinação das medidas formânticas na identificação da qualidade vocal<sup>(4,24)</sup>

As medidas relacionadas à média de  $F_0$ ,  $F_1$  e *GNE* foram selecionadas para discriminação de vozes soprosas e tensas, com acurácia de  $75,71 \pm 6,41\%$  e com maior especificidade ( $78,33 \pm 8,16$ ) na identificação correta de vozes soprosas.

Os valores de  $F_0$  e F1 mostraram-se maiores em pacientes com vozes tensas, já o *GNE* apresentou valores mais rebaixados em pacientes com vozes soprosas.

Em relação à média de  $F_0$  e à qualidade vocal tensa, salienta-se que a frequência fundamental é determinada, entre outros fatores, pela tensão das pregas vocais, que é controlada pelos músculos intrínsecos da laringe, especificamente, o cricotireoideo<sup>(2,11,15)</sup>. Assim, pacientes com tensão fonatória costumam realizar maior contração da musculatura extrínseca e intrínseca, inclusive com maior tensão longitudinal das pregas vocais, maior pressão subglótica e maior constrição no trato vocal, gerando um maior número de ciclos glóticos por segundo e, conseqüentemente, uma frequência fundamental mais elevada<sup>(25)</sup>.

O grau geral e a rugosidade parecem ser parâmetros mais relacionados com  $F_0$ <sup>(28,30)</sup>. Estando os valores da média de  $F_0$  mais elevados tanto no grau geral, como na tensão fonatória e os valores do desvio padrão de  $F_0$  também elevados em vozes rugosas. Os achados no presente estudo parecem concordar quanto ao aumento de  $F_0$  em pacientes com vozes tensas e a relação positiva de  $F_0$  quanto ao grau geral do desvio vocal.

Já em relação aos valores de F1 elevado, infere-se que os pacientes com qualidade vocal tensa poderiam realizar ajustes no trato vocal, tendo uma maior abertura vertical da boca e maior constrição faríngea<sup>(6-8,10,11)</sup> em relação aos pacientes com qualidade vocal soprosa.

Uma pesquisa<sup>(14)</sup> realizada com 111 mulheres com disфония por tensão muscular encontrou dados semelhantes. Os formantes F1 e F2 encontram-se elevados nessa população em relação a vozes saudáveis, o que sugere ajustes na supraglote relacionados a uma maior abertura vertical da boca, maior constrição faríngea e uma postura de língua mais abaixada e anteriorizada. Os ajustes encontrados, nessa pesquisa, assemelham-se ao presente estudo, quanto a uma maior abertura vertical da boca e maior constrição faríngea, indicados pelo aumento de F1 em pacientes com qualidade vocal tensa.

Ao analisar as medidas acústicas combinadas na discriminação da qualidade vocal predominante, mais uma vez, observou-se que a medida *GNE* apareceu em todas as combinações aceitáveis encontradas. A medida de  $F_0$  mostrou-se presente na maioria das combinações na discriminação da qualidade vocal predominante, o que atesta resultados encontrados em estudos anteriores<sup>(8,13,18,29,30)</sup>, em que a frequência fundamental mostra-se uma medida interessante na discriminação da qualidade vocal dos pacientes, provavelmente por estar relacionada, em termos fisiológicos, à condição neuromuscular e à regularidade de vibração da mucosa das pregas vocais, tendo, em termos acústicos e perceptivos, uma relação direta com a periodicidade do sinal sonoro<sup>(6,9,11,30)</sup>.

Em síntese, a combinação de medidas de perturbação/ruído com medidas formânticas promove uma discreta melhora (75,24%) na taxa de classificação entre vozes com VNQV e com desvio de leve a moderado em relação à medida *GNE* isolada (70,95%), além de possibilitar a discriminação entre vozes com desvio de leve a moderado *vs.* moderado, o que não foi observado com medidas isoladas. Tais achados põem em evidência que quanto maior a intensidade do desvio vocal, mais complexo é o sinal em termos de aperiodicidade e ruído,

o que demanda a necessidade de um conjunto de medidas para caracterizá-lo de forma adequada.

Além disso, a análise combinada de medidas relacionadas à fonte glótica (perturbação e ruído) e ao filtro (medidas formânticas) contribui para ampliar a compreensão dos mecanismos de interação fonte-filtro em vozes desviadas e pode se tornar útil enquanto medidas de resultado de tratamento e monitoramento ao longo da terapia vocal. O fato de mais medidas formânticas (F1 e F3) serem selecionadas pelo classificador para discriminação de vozes mais desviadas evidencia que indivíduos com desvios mais intensos implementam mais ajustes no trato vocal, provavelmente, como mecanismo compensatório à ineficiência funcional da fonte glótica.

Quanto à qualidade vocal predominante, as medidas formânticas demonstraram ser importantes para as classificações entre vozes VNQV *vs.* soprosa (F3), rugosa *vs.* tensa (F3), soprosa *vs.* tensa (F1). De modo específico, pode-se observar que as medidas formânticas parecem ter uma maior contribuição na discriminação do parâmetro perceptivo-auditivo de tensão. Provavelmente, indivíduos com vozes tensas implementam mais ajustes supraglóticos, seja de modo compensatório ou em coocorrência com as alterações em nível glótico.

A presença de um distúrbio de voz tende a alterar o sinal vocal de diferentes modos, combinando vários tipos de perturbação e ruído na emissão vocal, assim como possíveis ajustes supraglóticos. Desse modo, o uso combinado de medidas para avaliação, caracterização e classificação do sinal vocal pode representar melhor as características da produção vocal e evidenciar manifestações que não seriam detectadas com a utilização de medidas isoladas. Outros estudos<sup>(3,28,30)</sup> já demonstraram que a combinação de medidas de perturbação e ruído melhora a discriminação entre sinais com e sem desvio vocal. No entanto, com o presente estudo, pode-se concluir que incluir as medidas relacionadas aos ajustes de trato vocal às medidas tradicionais de perturbação e ruído pode melhorar o desempenho da classificação da intensidade e do tipo de desvio vocal, além de fornecer *insights* acerca da interação fonte e filtro nos pacientes com desvio vocal.

## CONCLUSÃO

A medida acústica *GNE*, de forma isolada, mostrou-se a única capaz de discriminar a intensidade do desvio vocal e a qualidade vocal predominante. Houve um ganho no desempenho da classificação com a combinação das medidas acústicas tradicionais e formânticas, tanto para a discriminação da intensidade do desvio vocal quanto da qualidade vocal predominante.

## REFERÊNCIAS

1. Dejonckere PH, Bradley P, Clemente P, Cornut G, Crevier-Buchman L, Friedrich G, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating assessment techniques. *Eur Arch Otorhinolaryngol.* 2001;258(2):77-82. <http://dx.doi.org/10.1007/s004050000299>. PMID:11307610.
2. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol.* 2009;18(2):124-32. [http://dx.doi.org/10.1044/1058-0360\(2008/08-0017\)](http://dx.doi.org/10.1044/1058-0360(2008/08-0017)). PMID:18930908.

3. Brockmann-Bausier M, Drinnan MJ. Routine acoustic voice analysis: time to think again? *Curr Opin Otolaryngol Head Neck Surg.* 2011;19(3):165-70. <http://dx.doi.org/10.1097/MOO.0b013e328345757fe>. PMID:21483265.
4. Godino-Llorente JI, Osma-Ruiz V, Sáenz-Lechón N, Gómez-Vilda P, Blanco-Velasco M, Cruz-Roldán F. Effectiveness of the glottal to noise excitation ratio for the screening of voice disorders. *J Voice.* 2010;24(1):47-56. <http://dx.doi.org/10.1016/j.jvoice.2008.04.006>. PMID:19135854.
5. Treole K, Trudeau MD. Changes in sustained production tasks among women with bilateral vocal nodules before and after voice therapy. *J Voice.* 1997;11(4):462-9. [http://dx.doi.org/10.1016/S0892-1997\(97\)80043-4](http://dx.doi.org/10.1016/S0892-1997(97)80043-4). PMID:9422281.
6. Fant G. *Acoustic theory of speech production.* 2nd ed. Paris: Mouton; 1970.
7. Ladefoged P. *Elements of acoustic phonetics.* Chicago: University of Chicago Press; 1996.
8. Camargo ZA. Análise da qualidade vocal de um grupo de indivíduos disfônicos: uma abordagem interpretativa e integrada de dados de natureza acústica, perceptiva e eletroglotográfica [tese]. São Paulo: Pontifícia Universidade Católica de São Paulo; 2002. 283 p.
9. Silva MFBL, Madureira S, Rusilo LC, Camargo Z. Vocal quality assessment: methodological approach for a perceptive data analysis. *Rev CEFAC.* 2017;19(6):831-41. <http://dx.doi.org/10.1590/1982-021620171961417>.
10. Magri A, Stamado T, Camargo ZA. Influência da largura de banda de formantes na qualidade vocal. *Rev CEFAC.* 2009;11(2):296-304. <http://dx.doi.org/10.1590/S1516-18462009005000010>.
11. Lee S-H, Yu J-F, Hsieh Y-H, Lee G-S. Relationships between formant frequencies of sustained vowels and tongue contours measured by ultrasonography. *Am J Speech Lang Pathol.* 2015;24(4):739-49. [http://dx.doi.org/10.1044/2015\\_AJSLP-14-0063](http://dx.doi.org/10.1044/2015_AJSLP-14-0063). PMID:26254465.
12. Titze I, Palaparthi A. Sensitivity of source-filter interaction to specific vocal tract shapes. *IEEE Trans Audio Speech Lang Process.* 2016;24(12):2507-15. <http://dx.doi.org/10.1109/TASLP.2016.2616543>.
13. Camargo ZA, Vilarim GS, Cukier S. Parâmetros perceptivo-auditivos e acústicos de longotermo da qualidade vocal de indivíduos disfônicos. *Rev CEFAC.* 2004;6(2):189-96.
14. Roy N, Nissen SL, Dromey C, Sapir S. Articulatory changes in muscle tension dysphonia: evidence of vowel space expansion following manual circumlaryngeal therapy. *J Commun Disord.* 2009;42(2):124-35. <http://dx.doi.org/10.1016/j.jcomdis.2008.10.001>. PMID:19054525.
15. Muhammad G, Mesallam TA, Malki KH, Farahat M, Alsulaiman M, Bukhari M. Formant analysis in dysphonic patients and automatic Arabic digit speech recognition. *Biomed Eng Online.* 2011;10:41. PMID:21624137.
16. Schwartz SR, Cohen SM, Dailey SH, Rosenfeld RM, Deutsch ES, Gillespie MB, et al. Clinical practice guideline: hoarseness (dysphonia). *Otolaryngol Head Neck Surg.* 2009;141(3, Supl 2):S1-S31. <http://dx.doi.org/10.1016/j.otohns.2009.06.744>. PMID:19729111.
17. Ma EP, Yiu EM. Multiparametric evaluation of dysphonic severity. *J Voice.* 2006;20(3):380-90. <http://dx.doi.org/10.1016/j.jvoice.2005.04.007>. PMID:16185841.
18. 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.
19. Cohen SM, Pitman MJ, Noordzij JP, Courey M. Management of dysphonic patients by otolaryngologists. *Otolaryngol Head Neck Surg.* 2012;147(2):289-94. <http://dx.doi.org/10.1177/0194599812440780>. PMID:22368039.
20. Gonçalves MIR, Pontes PAL, Vieira VP, Pontes AAL, Curcio D, Biase NG. Função de transferência das vogais orais do Português brasileiro: análise acústica comparativa. *Rev Bras Otorrinolaringol.* 2009;75(5):680-4.
21. Ozkan H. A Comparison of classification methods for telediagnosics of Parkinson's disease. *Entropy.* 2016;18(115):1-14.
22. Yamasaki R, Madazio G, Leão SHS, Padovani M, Azevedo R, Behlau M. Auditory-perceptual evaluation of normal and dysphonic voices using the voice deviation scale. *J Voice.* 2017;31(1):67-71. <http://dx.doi.org/10.1016/j.jvoice.2016.01.004>. PMID:26873420.
23. Hosmer DW, Lemeshow S. *Applied logistic regression.* New York: Wiley; 2000. <http://dx.doi.org/10.1002/0471722146>.
24. González CMT, Hernandez JBA, Orozco-Arroyave JR, Casals JS, Gallego-Jutgla E. Automatic detection of laryngeal pathologies in running speech based on the HMM transformation of the nonlinear dynamics. *Lect Notes Comput Sci.* 2013;1:136-43.
25. Van Houtte E, Van Lierde K, Claeys S. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *J Voice.* 2011;25(2):202-7. <http://dx.doi.org/10.1016/j.jvoice.2009.10.009>. PMID:20400263.
26. Macari AT, Ziade G, Turfe Z, Chidiac A, Alam E, Hamdan AL. Correlation between the position of the hyoid bone on lateral cephalographs and formant frequencies. *J Voice.* 2016;30(6):757.e21-6. <http://dx.doi.org/10.1016/j.jvoice.2015.08.020>. PMID:26604010.
27. Samlan RA, Story BH, Bunton K. Relation of perceived breathiness to laryngeal kinematics and acoustic measures based on computational modeling. *J Speech Lang Hear Res.* 2013;56(4):1209-23. [http://dx.doi.org/10.1044/1092-4388\(2012\)12-0194](http://dx.doi.org/10.1044/1092-4388(2012)12-0194). PMID:23785184.
28. Lopes LW, Costa SLNC, Costa WCA, Correia SEN, Vieira VJD. Acoustic assessment of the voices of children using nonlinear analysis: proposal for assessment and vocal monitoring. *J Voice.* 2014;28(5):565-73. <http://dx.doi.org/10.1016/j.jvoice.2014.02.013>. PMID:24836362.
29. Madazio G, Leão S, Behlau M. The phonatory deviation diagram: a novel objective measurement of vocal function. *Folia Phoniatr Logop.* 2011;63(6):305-11. <http://dx.doi.org/10.1159/000327027>. PMID:21625144.
30. Lopes LW, Simões LB, Silva JD, Evangelista DS, Ugulino ACN, Costa Silva PL, et al. Accuracy of acoustic analysis measurements in the evaluation of patients with different laryngeal diagnoses. *J Voice.* 2017;31(3):382.e15-26. <http://dx.doi.org/10.1016/j.jvoice.2016.08.015>. PMID:27742492.

### Contribuição dos autores

*Todos os autores (LWL, JNA, DSE, FPF, VJDV, MFBL e LAP) contribuíram na construção e desenvolvimento deste trabalho. JNA, DSE, FPF realizaram a coleta, tabulação e análise dos dados; LWL e VJDV supervisionaram a realização da pesquisa, executaram a análise estatística e revisão do manuscrito. MFBL e LAP participaram da análise estatística e revisão do manuscrito.*