








Leonardo Wanderley Lopes<sup>1</sup>   
 Estevão Silvestre da Silva Sousa<sup>2</sup>   
 Allan Carlos França da Silva<sup>2</sup>   
 Itacely Marinho da Silva<sup>2</sup>   
 Maxsuel Avelino Alves de Paiva<sup>2</sup>   
 Vinícius Jefferson Dias Vieira<sup>3</sup>   
 Anna Alice Almeida<sup>1</sup> 

### Descritores

Acústica  
 Qualidade Vocal  
 Distúrbio de Voz  
 Voz  
 Fonoaudiologia

### Keywords

Acoustics  
 Voice Quality  
 Voice Disorder  
 Voice  
 Speech-Language Pathology

### Endereço para correspondência:

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

Recebido em: Julho 24, 2018.

Aceito em: Novembro 24, 2018.

# Medidas cepstrais na avaliação da intensidade do desvio vocal

## *Cepstral measures in the assessment of severity of voice disorders*

### RESUMO

**Objetivo:** Analisar se existe associação entre a presença, a intensidade e o tipo de desvio vocal e as medidas cepstrais em amostras de indivíduos com queixa vocal. **Método:** Foram utilizadas 376 amostras da vogal /ε/ de indivíduos de ambos os gêneros, com queixa vocal. Utilizou-se uma escala analógico-visual para análise perceptivo-auditiva das vozes quanto à intensidade do desvio vocal (GG), graus de rugosidade (GR), soproidade (GS) e tensão (GT), incluindo-se a determinação da qualidade vocal predominante (rugosa, soprosa ou tensa). Foram extraídas as medidas relacionadas ao *Cepstral Peak Prominence-Smoothed (CPPS)* e o declínio espectral das amostras vocais. **Resultados:** Houve diferença dos valores do *CPPS* entre os grupos com e sem desvio vocal, assim como entre as diferentes intensidades e tipos de desvio vocal. Os valores do *CPPS* foram mais reduzidos em função da presença e intensidade do desvio vocal. Os valores do *CPPS* diferenciaram vozes rugosas x soprosas, rugosas x tensas e soprosas x tensas. O declínio espectral apenas diferenciou vozes soprosas x tensas. O *CPPS* se correlacionou de modo positivo e forte com os GG e GS, de modo negativo moderado com o GR, e de forma negativa fraca com o GT. O declínio espectral apresentou correlação positiva moderada com o GT e correlação negativa fraca com o GS. **Conclusão:** Existe associação entre a presença de desvio vocal, o GG, a qualidade vocal predominante e o *CPPS*. De modo especial, o GG é fortemente correlacionado ao *CPPS*. O declínio espectral está associado apenas aos parâmetros de soproidade e tensão.

### ABSTRACT

**Purpose:** To analyze whether there is an association between the presence, intensity and type of voice disorder and the cepstral measures in samples of individuals with voice complaints. **Methods:** We used 376 vowel /ε/ samples from individuals of both genders that had voice complaints. An analogue-visual scale was used for the perceptual-auditory analysis of voices regarding the degree of voice disorder (GD), roughness (RD), breathiness (BD), and strain (SD), including a determination of voice quality (rough, breathy or strained). Measures related to cepstral peak prominence smoothed (CPPS) and spectral decline of vocal samples were extracted. **Results:** There were differences in the CPPS values between the groups with or without voice disorders as well as between the different intensities and types of voice disorder. CPPS values were lower because of the presence and intensity of voice disorders. The CPPS values differentiated the following voices: rough x breathy, rough x strained, and breathy x strained. The spectral decline only differentiated breathy x strained voices. CPPS correlated positively and strongly with GD and BD; moderately and negatively with RD, and negatively and weakly with SD. The spectral decline had a moderate positive correlation with the SD and a weak negative correlation with BD. **Conclusion:** There is an association between voice disorder, GD, predominant voice quality and CPPS. In particular, GD is strongly correlated with CPPS. The spectral decline is associated only with the parameters of breathiness and strain.

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

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

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

<sup>3</sup> Programa de Pós-graduação em Linguística, Universidade Federal da Paraíba – UFPB - João Pessoa (PB), Brasil.

**Fonte de financiamento:** nada a declarar.

**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 avaliação dos distúrbios da voz deve levar em consideração uma abordagem multidisciplinar e multidimensional<sup>(1)</sup>, incluindo: uma anamnese detalhada, para identificação da queixa, dos fatores de risco e dos sintomas presentes; a análise perceptivo-auditiva, para identificação da presença, tipo e intensidade do desvio vocal; a avaliação acústica, que fornece uma análise quantitativa e qualitativa do sinal vocal; a avaliação aerodinâmica, que fornece dados sobre o controle do fluxo aéreo para a fonação; e o exame visual laríngeo, para análise estrutural e funcional da laringe<sup>(1,2)</sup>.

A análise perceptivo-auditiva e o exame visual laríngeo são os principais métodos utilizados para avaliação dos distúrbios de voz pelo fonoaudiólogo e pelo médico otorrinolaringologista, respectivamente. Vale enfatizar que ambos os métodos possuem fatores de confundimento relacionados à subjetividade do avaliador<sup>(3)</sup> que faz julgamento auditivo da qualidade vocal (perceptivo-auditiva) e julgamento visual do exame laríngeo.

A análise acústica é complementar à avaliação perceptivo-auditiva e laríngea propiciando dados quantitativos e qualitativos da função vocal e alta reprodutibilidade para monitoramento do paciente<sup>(4,5)</sup>. Um aspecto relevante é a possibilidade da quantificação do desvio presente no sinal e a sua comparação com dados normativos<sup>(6)</sup>. Em síntese, a validade das medidas acústicas depende da sua capacidade de representar o desvio da qualidade vocal percebido auditivamente e os mecanismos fisiológicos subjacentes à produção vocal. Desse modo, um dos desafios para os clínicos e pesquisadores é compreender o quanto cada medida está relacionada à avaliação perceptivo-auditiva e ao exame visual laríngeo.

De modo geral, a análise acústica pode envolver a extração de medidas que quantificam determinada propriedade do sinal da voz, assim como pode incluir a análise descritiva de padrões visuais destes mesmos sinais<sup>(7,8)</sup>. Na extração das medidas clássicas de perturbação (*jitter* e *shimmer*) e de ruído (proporção harmônico-ruído), é necessário que haja uma estimativa dos valores da frequência fundamental ( $F_0$ ), com determinação clara dos ciclos glóticos, o que é mais provável de ser realizado apenas nos sinais vocais com desvios leves<sup>(1,9)</sup>.

Nos indivíduos disfônicos, os sinais de voz podem oscilar de quase periódicos a completamente aperiódicos, de forma que a complexidade de um sinal com desvios moderados e intensos pode comprometer a confiabilidade das medidas tradicionais baseadas em modelos lineares, como *jitter* e *shimmer*<sup>(1)</sup>. Sendo assim, embora essas medidas tradicionais apresentem uma correlação de moderada a forte com a percepção auditiva do desvio vocal<sup>(10)</sup>, elas podem ter uma aplicação restrita na análise de vozes com desvios mais intensos.

Por sua vez, a análise cepstral demonstrou ser uma alternativa para avaliação de sinais mais desviados, uma vez que ela é capaz de determinar a  $F_0$  e produzir estimativas de aperiodicidade e/ou ruído aditivo sem a identificação de limites de ciclo individuais, como preconizado na extração das medidas de perturbação e ruído<sup>(11)</sup>. De forma geral, o cepstro evidencia em que medida os harmônicos advindos da  $F_0$  são individualizados e se destacam em relação ao nível de ruído presente no sinal. Sinais com maior regularidade e menor quantidade de ruído apresentam

maior definição e amplitude do pico cepstral dominante<sup>(11)</sup>. Desse modo, as medidas cepstrais são mais confiáveis que as medidas tradicionais de perturbação e ruído para avaliação de vozes com ampla faixa de desvio e, além disso, demonstraram ser fortes preditoras da presença de desvio vocal<sup>(11-13)</sup>.

Nesse contexto, o objetivo do presente estudo foi analisar se existe associação entre a presença, a intensidade e o tipo de desvio vocal e as medidas cepstrais em amostras de indivíduos com queixa vocal.

## MÉTODOS

### Desenho do estudo

Este é um estudo descritivo, transversal e observacional, avaliado e aprovado pelo comitê de Ética em Pesquisa da Instituição de origem, com o parecer de número 52492/12.

### Amostra

A amostra desta pesquisa foi constituída de 376 pacientes com queixa vocal, de ambos os gêneros, atendidos no laboratório de voz da Instituição de origem. Todos os participantes assinaram o Termo de Consentimento Livre e Esclarecido, autorizando a sua participação na pesquisa.

Para seleção dos indivíduos, foram considerados os seguintes critérios de elegibilidade: apresentar queixa vocal, respondendo positivamente à pergunta: “Você considera que tem um problema de voz atualmente?”; terem realizado o exame visual laríngeo para confirmação diagnóstica de distúrbio de voz, nas duas semanas prévias ou subsequentes à sessão de coleta de dados; não apresentar comprometimento cognitivo ou neurológico que impedisse a gravação da voz; não ter realizado terapia vocal ou tratamento cirúrgico na laringe previamente.

Desse modo, foram selecionados 376 indivíduos, com idade média de  $41,20 \pm 14,04$  anos, sendo 294 mulheres e 82 homens. Esses pacientes apresentaram os seguintes diagnósticos médicos: 99 (26,30%) sem alteração estrutural ou funcional na laringe, 90 (23,90%) pacientes com nódulos vocais, 42 (11,20%) com distúrbio de voz secundário a refluxo laringofaríngeo, 38 (10,10%) com cisto vocal, 25 (6,66%) com fenda triangular médio-posterior, 22 (5,85%) com paralisia unilateral de prega vocal, 21 (5,60%) com distúrbio de voz secundário a doença neuromuscular, 20 (5,30%) com pólipos de prega vocal, 11 (2,90%) com sulco vocal e 8 (2,10%) com edema de Reinke.

Todos esses pacientes procuraram o serviço por demanda espontânea ou encaminhados pelo otorrinolaringologista e foram avaliados antes de realização da terapia vocal. Os pacientes com distúrbio de voz secundário à doença neuromuscular também apresentaram laudo médico do neurologista. Desse modo, todos os indivíduos participantes apresentavam queixa vocal e receberam confirmação diagnóstica do distúrbio de voz por meio do exame visual laríngeo. Considerando-se o objetivo descrito para esta pesquisa, a avaliação perceptivo-auditiva foi eleita como padrão de referência para determinação do desfecho (presença/ausência de desvio vocal, intensidade do desvio e qualidade vocal predominante), independentemente do resultado do exame visual laríngeo.

## Procedimentos de coleta de dados

Toda a coleta de dados desta pesquisa foi realizada no Laboratório de Voz de uma Instituição de Ensino Superior. Inicialmente, os pacientes responderam a uma ficha contendo dados pessoais e informações referentes à queixa vocal. Na sequência, submeteram-se à gravação da vogal /É/ sustentada.

A coleta de voz ocorreu no referido laboratório, utilizando-se o *software* Fonoview, versão 4.5, da CTS Informática, *desktop Dell 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. 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 a boca do paciente.

Para a coleta das vozes, os pacientes estavam 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 ao registro. Durante a gravação, o paciente foi solicitado a emitir a vogal [É] sustentada, em frequência e intensidade autorreferida como habitual. A vogal [É] 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.

Posteriormente, as vozes foram editadas no *software SoundForge* versão 10.0, sendo eliminados os dois segundos iniciais e finais da emissão da vogal sustentada, devido a maior irregularidade nesses trechos, preservando-se o tempo mínimo de três segundos para cada emissão.

A extração das medidas acústicas foi realizada no *software* de acesso livre Praat (Paul Boersma e David Weenink, University of Amsterdam, Holanda), versão 5.3.84, extraindo-se o *Cepstral Peak Prominence-Smoothed (CPPS)* e o declínio espectral das amostras vocais. O CPPS é uma modificação no algoritmo do CPP (*Cepstral Peak Prominence*) que produz uma melhoria notável na precisão da análise das vozes desviadas. Essa modificação envolve a suavização do cepstro antes de se extrair o pico cepstral. Ao invés do cepstro ser calculado a cada 10 ms, com o CPPS, ele é calculado a cada 2 ms, favorecendo assim a precisão na identificação das irregularidades presentes no sinal<sup>(11)</sup>.

Os seguintes comandos e parâmetros foram aplicados para gerar o CPPS no Praat:

1. Clicou-se em “*Analyze Periodicity*” e, na sequência, em “*Fo Power Cepstrogram*”.
2. No “*menu*”, prosseguiu-se com “*Pitch floor (Hz) = 60*”, “*Time Step (s) = 0,002*”, “*Maximum Frequency (Hz) = 5000*” e “*Pre-emphasis from (Hz) = 50*”.
3. Clicou-se em “*Query*” e selecionou-se “*Get CPPS*” no “*menu*”, seguiu-se adiante com “*Subtract tilt before smoothing*” e com “*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. O resultado desse procedimento foram as medidas CPPS, conforme descrito em Maryn e Weenink<sup>(14)</sup>.

Os seguintes comandos e parâmetros foram aplicados para obter o declínio espectral no Praat:

1. Clicou-se em “*Analyse spectrum*” e escolheu-se “*To Ltas*”;
2. Seguiu-se adiante em “*Bandwidth*” com 100 Hz;
3. Foi selecionado o novo sinal do “*Ltas*” e clicado em “*Query*”;
4. Avançou-se com “*Get slope*”, em “*Low Band*”, os valores foram mudados para 0 e 1250 Hz, e em “*High Band*”, os valores foram alterados para 1250 e 4000 Hz;
5. Ainda, em “*Query*”, obtiveram-se os valores de Declínio Espectral “*Report spectral tilt*”;
6. O resultado desse procedimento são as medidas do Declínio Espectral conforme descrito em Maryn e Weenink<sup>(14)</sup>.

Todos os valores do CPPS e do declínio espectral foram conferidos manualmente para identificação de *outliers*, que correspondem a valores espúrios advindos de erros na extração da medida analisada. Ao final da inspeção, não foram identificados *outliers* nos sinais avaliados.

Para a análise das medidas perceptivo-auditivas, as vozes foram reeditadas no controle “*normalize*” do *SoundForge*, no modo *peak level*, a fim de obter uma padronização na saída de áudio entre -6 e 6dB para todos os sinais, de modo que a intensidade do sinal de áudio não influenciasse no julgamento dos avaliadores quanto à intensidade do desvio vocal.

A avaliação perceptivo-auditiva foi realizada por três fonoaudiólogos, de modo independente. Inicialmente, os juízes foram treinados com 16 estímulos-âncora (vogal sustentada /É/), contendo quatro amostras de indivíduos com variabilidade normal da qualidade vocal (VNQV), quatro amostras de indivíduos com desvio vocal de leve a moderado, quatro amostras de indivíduos com desvio vocal moderado e quatro amostras de indivíduos com desvio vocal intenso. 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 em que esta pesquisa foi realizada.

Para a análise perceptivo-auditiva, foi utilizada a Escala Analógica Visual (EAV) de 0 a 100 mm<sup>(15)</sup>, para a avaliação da intensidade do desvio vocal (GG) e dos graus de rugosidade (GR), de soproidade (GS) e de tensão (GT) na emissão da vogal sustentada. Os juízes foram orientados de que a marcação mais próxima do 0 representaria vozes mais socialmente aceitáveis, produzidas de forma mais natural, com menor esforço, ruído ou condição instável<sup>(15)</sup>. Ao contrário, marcações mais próximas do 100 representariam vozes menos aceitas socialmente e com maior percepção de esforço, ruído ou instabilidade. 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>(16)</sup> e por possuírem correlatos conhecidos no plano fisiológico e acústico.

Para avaliação, cada emissão do vogal sustentada foi apresentada por três vezes através de caixa de som, em intensidade confortável autorreferida pelo avaliador. Após cada apresentação, os juízes fizeram a avaliação da GG e do GR, GS e GT, seguida pela identificação qualidade vocal (tipo de desvio) predominante nas vozes desviadas (rugosa, soproosa ou tensa).

No final da sessão de avaliação perceptiva, 20% (76 sinais) das amostras foram repetidas aleatoriamente, para a análise da confiabilidade da avaliação do juiz através do Coeficiente Kappa de Cohen. Foi selecionado o juiz com maior coeficiente (0,80), indicando uma boa confiabilidade do avaliador<sup>(17)</sup>.

Os valores do ponto de corte da EAV<sup>(15)</sup> foram utilizados para classificar as vozes quanto à presença de desvio vocal e GG. Dessa forma, 97 vozes foram classificadas com VNQV ( $GG \leq 35,5$  mm) e 279 vozes foram categorizadas como desviadas ( $GG > 35,5$  mm). Todos os indivíduos com VNQV apresentaram ausência de alteração estrutural ou funcional na laringe. Dos pacientes com vozes desviadas, apenas dois receberam laudo de ausência de alteração estrutural ou funcional na laringe, enquanto os 277 restantes apresentaram os diagnósticos médicos citados anteriormente. Na sequência, os valores do GG na EAV foram utilizados para classificação dos sinais em quatro grupos, a partir dos valores de corte indicados na literatura<sup>(15)</sup>: 97 vozes com VNQV (0-35,5 mm); 239 vozes com grau 2 (35,6-50,5 mm), correspondendo ao desvio de leve a moderado; 165 vozes com grau 3 (50,6-90,5 mm), relacionado ao desvio moderado; e 27 vozes com grau 4 (90,6-100 mm), relacionado ao desvio intenso.

Deve-se destacar que o estudo de referência<sup>(15)</sup> utilizado na realidade brasileira para determinação dos valores de corte da EAV utilizou como tarefa de fala apenas a contagem de 1 a 10 (fala encadeada). Embora tal fato possa se constituir em uma limitação do presente estudo, optou-se pela utilização dos valores de corte propostos por Yamasaki et al. (2017)<sup>(15)</sup> por utilizar apenas os quatro graus de desvio considerados internacionalmente (saudável ou VNQV, leve a moderado, moderado e intenso) e ser a principal referência brasileira para os valores de corte utilizados nessa classificação.

## ANÁLISE DOS DADOS

Foi realizada análise estatística descritiva para todas as variáveis analisadas, incluindo-se os valores de média e desvio padrão. O teste não paramétrico de Mann-Whitney foi utilizado para comparação das médias das medidas cepstrais entre os grupos com e sem desvio. O teste de Kruskal-Wallis foi utilizado para comparar a média das medidas cepstrais em função da intensidade do desvio vocal, utilizando-se o teste *post hoc* de Nemenyi para comparação dos grupos par a par.

O teste de correlação de Spearman foi utilizado para verificar a correlação entre a intensidade do desvio vocal e as medidas cepstrais. Os coeficientes de correlação foram utilizados para avaliar e quantificar o grau de relacionamento linear entre as duas variáveis, observando se as variáveis se modificam conjuntamente e em que grau. Para a classificação dos coeficientes de correlação, adotou-se, nesta pesquisa, que valores de 0,1 a 0,3 representam uma correlação fraca; entre 0,4 e 0,6 indicam correlação moderada; e acima de 0,6 é possível afirmar que o grau de correlação entre as variáveis é forte<sup>(18)</sup>.

Todas as análises foram realizadas pelo *software Statistical Package for the Social Sciences* (SPSS) versão 2.0. O nível de significância adotado foi de 5%.

## RESULTADOS

Inicialmente foi realizado o teste não paramétrico de Mann-Whitney para comparar a média das medidas cepstrais entre os grupos com e sem desvio vocal (Tabela 1). Houve diferença dos valores do *CPPS* entre os grupos ( $p < 0,001$ ), com valores mais elevados nos pacientes sem desvio vocal.

O teste de Kruskal-Wallis foi utilizado para comparar a média das medidas cepstrais em função da intensidade do desvio vocal (Tabela 2). Constatou-se diferença entre os grupos para a medida *CPPS* ( $p < 0,001$ ). Na sequência, utilizou-se o teste *post hoc* de Nemenyi para comparação dos grupos par a par. Houve diferença entre indivíduos do grupo com VNQV e do grupo com grau leve a moderado ( $p = 0,001$ ), com maiores valores no grupo com VNQV. Da mesma forma, houve diferença entre os grupos com grau leve a moderado e moderado ( $p = 0,001$ ), com valores superiores no grupo com grau leve a moderado. Também houve diferença entre os grupos com grau moderado e intenso ( $p = 0,001$ ), com maiores valores no grupo com grau moderado.

Foi realizado o teste não paramétrico de Kruskal-Wallis para comparar as medidas cepstrais em função da qualidade vocal predominante. Observou-se diferença nos valores de *CPPS* ( $p < 0,001$ ) e declínio espectral ( $p < 0,001$ ) entre os diferentes tipos de vozes (Tabela 3). Na análise *post hoc*, os valores de *CPPS* separaram vozes rugosas de soproas ( $p = 0,001$ ), com maiores valores da média de *CPPS* em vozes rugosas. Houve diferença nas médias do *CPPS* ( $p = 0,001$ ) e declínio espectral ( $p < 0,001$ ) entre vozes rugosas e tensas. Vozes rugosas apresentaram menores valores do *CPPS* e maiores valores do declínio espectral em relação às vozes tensas. Os valores de *CPPS* ( $p < 0,001$ ) e declínio espectral ( $p < 0,001$ ) também diferenciaram vozes soproas de vozes tensas. As vozes tensas apresentaram maiores valores do *CPPS* e menores valores do declínio espectral.

Por fim, realizou-se o teste de correlação de Spearman entre as medidas perceptivo-auditivas e cepstrais (Tabela 4). O *CPPS* apresentou correlação negativa forte com o GG ( $p < 0,001$ ) e GS ( $p < 0,001$ ), correlação negativa moderada com o GR ( $p < 0,001$ ) e correlação negativa fraca com o GT ( $p = 0,001$ ). Quanto ao declínio espectral, observou-se correlação positiva moderada com o GT ( $p < 0,001$ ) e correlação negativa fraca com o GS ( $p = 0,001$ ).

**Tabela 1.** Comparação das médias das medidas cepstrais entre os grupos com e sem desvio vocal

VARIÁVEIS	VNQV		VOZES DESVIADAS		p-valor
	Média	DP	Média	DP	
<b>CPPS</b>	16,35	2,40	13,93	3,54	<0,001*
<b>Declínio Espectral</b>	-13,63	6,25	-14,74	6,51	0,139

\*Valores significativos ( $p < 0,05$ ) – Teste de Mann-Whitney.

**Legenda:** VNQV = variabilidade normal da qualidade vocal; DP = desvio padrão; CPPS = *cepstral peak prominence smoothed*

**Tabela 2.** Comparação das médias das medidas cepstrais em função da intensidade do desvio vocal

Variáveis	Intensidade do desvio vocal								p-valor
	VNQV		Leve a moderado		Moderado		Intenso		
	Média	DP	Média	DP	Média	DP	Média	DP	
<b>CPPS</b>	16,37	2,38	15,05	2,63	12,58	3,25	7,56	3,62	8,75
<b>Declínio Espectral</b>	-13,48	6,67	-15,10	7,16	-15,07	7,74	-12,78	<0,001*	0,479

\*Valores significativos ( $p < 0,05$ ) – Teste de Kruskal-Wallis.

**Legenda:** VNQV = variabilidade normal da qualidade vocal; CPPS = *cepstral peak prominence smoothed*; DP = desvio padrão

**Tabela 3.** Comparação das medidas cepstrais em função da qualidade vocal predominante

Variáveis	Qualidade Vocal Predominante						p-valor
	Rugosidade		Soprosidade		Tensão		
	Média	DP	Média	DP	Média	DP	
<b>CPPS</b>	13,75	3,62	12,5	3,15	15,5	3,11	<0,001*
<b>Declínio Espectral</b>	-14,94	5,81	-16,74	7,24	-10,58	5,79	<0,001*

\*Valores significativos ( $p < 0,05$ ) – Teste de Kruskal-Wallis.

**Legenda:** DP = desvio padrão; CPPS = *cepstral peak prominence smoothed*

**Tabela 4.** Correlação entre a intensidade do desvio vocal, os graus de rugosidade, soprosidade e tensão e as medidas cepstrais.

Variáveis	GG		GR		GS		GT	
	Correlação	Valor de p	Correlação	Valor de p	Correlação	Valor de p	Correlação	Valor de p
<b>CPPS</b>	-0,629	<0,001*	-0,536	<0,001*	-0,618	<0,001*	-0,150	0,001*
<b>Declínio Espectral</b>	0,050	0,272	-0,021	0,624	-0,145	0,001*	0,308	<0,001*

\*Valores significativos ( $p < 0,05$ ) – Teste de correlação de Spearman.

**Legenda:** CPPS = *cepstral peak prominence smoothed*; GG = grau geral; GR = grau de rugosidade; GS = grau de soprosidade; GT = grau de tensão

## DISCUSSÃO

No contexto da avaliação vocal, há um esforço de clínicos e pesquisadores para identificar medidas capazes de identificar e monitorar com confiabilidade os desvios da qualidade vocal<sup>(19)</sup>. Sendo assim, as medidas cepstrais têm demonstrado potencial para avaliar vozes com ampla faixa de desvio, o que justifica o aumento de estudos com essas medidas, favorecendo a compreensão do seu papel na clínica vocal<sup>(20)</sup>.

Na presente pesquisa, observou-se que o *CPPS* é capaz de diferenciar indivíduos com e sem desvio da qualidade vocal, com maiores valores no último grupo. Essa diferença pode ser justificada pelo fato de que os sinais de voz sem desvio apresentam maior periodicidade, com configuração harmônica bem definida e, portanto, maiores valores de *CPPS*. Ao contrário, as vozes mais desviadas apresentam menor proporção entre a energia dos harmônicos e os componentes de ruído e aperiodicidade, com menores valores de *CPPS*<sup>(11)</sup>.

Neste estudo, a maioria dos indivíduos com desvio da qualidade vocal apresentava alteração estrutural e/ou funcional na laringe. Fisiologicamente, a presença de tais alterações pode ocasionar mudança dos padrões vibratórios e no fechamento

glótico, ocasionando a aperiodicidade e ruído no sinal vocal, respectivamente<sup>(21)</sup>.

Alguns estudos<sup>(20,22,23)</sup> investigaram a capacidade de as medidas cepstrais discriminarem vozes saudáveis de vozes desviadas. Nesses estudos, foram encontradas taxas de acurácia entre 71% e 85% de classificação entre sinais saudáveis e desviados. Os autores<sup>(20,23)</sup> utilizaram como padrão de referência a avaliação perceptivo-auditiva (acurácia=85%), seguida pelo exame visual laríngeo (acurácia=73%)<sup>(20,24)</sup> e pela autoavaliação vocal (acurácia=75%). As taxas de classificação se comportaram de modo diferente em função do padrão de referência utilizado, com melhor desempenho relacionado à análise perceptivo-auditiva em detrimento do exame visual laríngeo e da autoavaliação vocal. No entanto, em todos os casos, as medidas cepstrais foram capazes de diferenciar vozes saudáveis e desviadas.

A interpretação das taxas de classificação deve estar pautada no objetivo de teste, privilegiando os valores de sensibilidade em caso de medidas de triagem, em detrimento dos valores de especificidade quando o objetivo do teste é confirmação diagnóstica. O estudo de Awan<sup>(20)</sup> tinha por objetivo propor a utilização das medidas cepstrais para triagem dos distúrbios da voz. Desse modo, eles utilizaram um valor de corte mais baixo (19,09 dB,

19,01 dB e 19,46 dB para a avaliação perceptivo-auditiva, exame laríngeo e autoavaliação, respectivamente) para classificação dos sinais em saudáveis ou desviados a partir dos padrões de referência preconizados. Valores abaixo desse ponto de corte indicariam a presença de alteração a partir dos padrões de referência citados.

Dessa forma, os valores do *CPPS* encontrados no presente estudo para ambos os grupos (VNQV e desviado) estão abaixo dos valores de corte preconizados na literatura<sup>(20)</sup>. A partir desse achado, pode-se discutir as diferenças metodológicas entre o estudo citado<sup>(20)</sup> e a presente pesquisa. Há três diferenças principais, relacionadas à forma de julgamento perceptivo-auditivo das vozes analisadas, à tarefa de fala para avaliação perceptivo-auditiva e aos critérios de alocação dos sujeitos.

Awan et al.<sup>(20)</sup> utilizaram uma avaliação binária/categórica, de modo que os avaliadores deveriam indicar apenas se as vozes escutadas eram saudáveis ou desviadas. Em contrapartida, o presente estudo utilizou o ponto de corte da EAV para categorização das vozes em saudáveis e desviadas.

Eles<sup>(20)</sup> utilizaram duas tarefas de fala, incluindo a fala encadeada (leitura do “*The Rainbow Passage*”) e a vogal /a/ sustentada, enquanto no presente estudo foi utilizada a emissão da vogal sustentada /ε/. Segundo o mesmo autor<sup>(4)</sup>, ainda existe uma incerteza sobre quais tarefas de fala devem ser incluídas nos modelos de previsão da presença ou ausência do distúrbio vocal, principalmente quando comparam-se vogais sustentadas e fala encadeada.

A fala encadeada é mais próxima da conversação diária, entretanto, quando há a classificação da qualidade vocal, ela mostra-se mais variável, pois o foco perceptual é divergente dos fenômenos não vocais, por exemplo, a prosódia, a articulação de palavras e todo o contexto fonético e fonológico. Por outro lado, vogais sustentadas, são menos propensas a essa variabilidade fonética<sup>(25)</sup>. Além disso, a utilização de vogais é um dos procedimentos mais citados na prática clínica para avaliação da qualidade vocal. Embora, se saiba que as diferenças culturais, principalmente no que diz respeito à língua, podem influenciar os resultados das avaliações da qualidade vocal. No que diz respeito às medidas cepstrais, ainda não há ponto de corte dessas medidas em indivíduos falantes do Português Brasileiro.

Além disso, no estudo citado<sup>(20)</sup>, todos os indivíduos sem desvio vocal não apresentavam queixa vocal nem alteração no exame visual laríngeo. Tal fato pode justificar a presença de um ponto de corte mais elevado, visto que se cria um grupo mais homogêneo quando se utiliza um padrão de referência combinado. Os critérios utilizados nesse estudo<sup>(20)</sup> justificam-se pelo fato de que os autores buscavam identificar um ponto de corte para triagem de distúrbios da voz. De modo diferente, a presente pesquisa busca investigar a relação dessas medidas em relação ao desvio da qualidade vocal, utilizando como padrão de referência apenas a avaliação perceptivo-auditiva.

Um outro estudo<sup>(3)</sup> utilizou a avaliação perceptivo-auditiva como padrão de referência e o ponto de corte da EAV para alocação dos indivíduos com e sem desvio da qualidade vocal. O ponto de corte do *CPPS* para discriminação dos indivíduos saudáveis e com desvio vocal foi de 17,68 dB, estando mais próximo dos resultados encontrados na presente pesquisa.

Na comparação da média das medidas acústicas em função da intensidade do desvio vocal, observou-se diferença entre os grupos com diferentes graus de desvio, com menores valores nas vozes mais desviadas em cada grupo (VNQV x leve a moderado, leve a moderado x moderado, moderado x intenso). Desse modo, quanto maior a intensidade do desvio vocal menor a energia acústica da  $F_0$  e sua definição em relação à energia total do sinal acústico<sup>(26)</sup>, o que ocasiona um decréscimo no pico espectral em função da intensidade do desvio vocal<sup>(3,11,12)</sup>.

Com relação à qualidade vocal predominante, houve diferença nos valores do *CPPS* entre os diferentes tipos de desvio vocal. As vozes com predomínio de tensão apresentaram maiores valores do *CPPS* em relação às vozes predominantemente rugosas e soprosas. Por sua vez, vozes rugosas apresentam maiores valores do *CPPS* em detrimento das vozes soprosas. Em relação ao declínio espectral, as vozes tensas apresentaram menores valores em relação às vozes rugosas e soprosas.

A tensão fonatória é comumente caracterizada por um aumento na contração da musculatura intrínseca e extrínseca da laringe, ocasionando maior rigidez no sistema e maior pressão longitudinal nas pregas vocais, com aumento da pressão subglótica e aumento do tempo da fase fechada do ciclo glótico<sup>(27)</sup>. De modo geral, tal ajuste produz um sinal com maior nível de energia e maior definição de  $F_0$ , o que justifica os maiores valores do *CPPS* e menor declínio espectral nas vozes tensas, em relação às vozes rugosas e soprosas<sup>(27)</sup>.

Quanto à diferenciação entre as vozes soprosas e as vozes rugosas e tensas, o padrão fisiológico tipicamente associado às vozes soprosas é caracterizado pelo maior grau de separação entre os processos vocais, menor convexidade da borda livre das pregas vocais e menor tempo da fase fechada do ciclo glótico. Esse padrão fisiológico gera diminuição da energia abaixo de 2500 Hz e aumento de energia nas faixas de frequência mais elevadas, o que justificaria menores valores do *CPPS* em vozes soprosas, visto que o aumento de ruído em frequências elevadas é um dos fatores que mais influencia na diminuição do *CPPS*<sup>(27)</sup>.

Vozes rugosas têm maior componente de ruído nas baixas frequências em detrimento das altas frequências, o que pode estar relacionado com os maiores valores do *CPPS* em vozes rugosas quando comparadas às vozes soprosas. Em estudos anteriores, o declínio espectral<sup>(24,27,28)</sup> e o *CPP* foram os principais parâmetros para diferenciar entre vozes soprosas e saudáveis, embora não tenha diferenciado entre vozes rugosas e soprosas, nem tenha sido selecionada com a principal medida para diferenciar entre vozes rugosas e saudáveis. Na análise acústica multivariada, apenas a combinação das medidas de *shimmer* e média de  $F_0$  foram capazes de diferenciar vozes rugosas de vozes soprosas<sup>(24,27,28)</sup>.

Observou-se uma correlação negativa forte entre o *CPPS* e o GG e o GS, de modo que vozes mais desviadas e com maior componente de sopro apresentaram maior decréscimo do pico cepstral. Outros estudos<sup>(3,10,11)</sup> demonstraram que existe uma forte correlação entre a intensidade do desvio vocal e as medidas cepstrais, assim como entre a percepção de sopro na emissão vocal e tais medidas. De modo geral, as medidas cepstrais demonstram ter uma maior força de correlação com a intensidade do desvio vocal em comparação com as medidas baseadas no domínio do tempo (*jitter* e *shimmer*). Em estudo

anterior<sup>(8)</sup>, verificou-se uma correlação moderada positiva entre as medidas de *jitter* e *shimmer* e o GG.

As características espectrais do sinal vocal são estreitamente relacionadas às mudanças na duração do contato das pregas vocais<sup>(28)</sup>. Há uma correlação positiva forte entre o quociente de abertura e grau de convexidade das pregas vocais e o aumento da energia na região de 4 KHz. Desse modo, pode-se compreender a forte correlação encontrada entre o GS e o *CPPS* na presente pesquisa, uma vez que a diminuição do tempo de fase fechada do ciclo glótico é o principal correlato fisiológico da presença de sopro na emissão vocal.

Quanto ao GR, observou-se uma correlação negativa moderada com o *CPPS*. A rugosidade corresponde à irregularidade vibratória da vibração das pregas vocais ocasionada por modificações na pressão subglótica ou alterações estruturais na borda livre das pregas vocais<sup>(24)</sup>, produzindo uma emissão com presença de sub-harmônicos, modulação de amplitude e aumento da perturbação no sinal. A rugosidade é caracterizada pelo componente de ruído em baixas frequências<sup>(10)</sup>, associadas a um decréscimo da média de  $F_0$  e aumento do seu desvio padrão<sup>(8)</sup>.

Dessa forma, o *CPPS*, cujo valor está diretamente relacionado à diferença entre energia nas frequências mais baixas e presença de ruído aditivo nas frequências mais elevadas, parece ser menos correlacionado ao componente de rugosidade em comparação ao componente de sopro<sup>(10)</sup>. Sendo assim, a presença e o GR são caracterizados de forma mais adequada por métodos de análise acústica que envolvam medidas baseadas na distribuição de energia e nos aspectos temporais da emissão, como as medidas cepstrais/espectrais e *jitter/shimmer*, respectivamente<sup>(24)</sup>. O desempenho das medidas cepstrais é inferior ao desempenho das medidas baseadas no domínio do tempo para avaliar o GR<sup>(8,21)</sup>.

Houve uma correlação negativa fraca entre o GT e o *CPPS*. Entre os parâmetros perceptivo-auditivos, a tensão tem sido referida como a característica mais controversa e difícil de ser avaliada acusticamente<sup>(8,29)</sup>. A presença de tensão na emissão vocal está associada fisiologicamente com tensão longitudinal das pregas vocais, aumento da pressão subglótica, maior contração da musculatura extrínseca e intrínseca da laringe, posição mais verticalizada da laringe e aumento do tempo da fase fechada dos ciclos glóticos<sup>(29)</sup>. Em termos acústicos, vozes tensas tendem a apresentar aumento de energia nas altas frequências<sup>(26)</sup>, o que também pode ocorrer em indivíduos vocalmente saudáveis que utilizam a voz com maior projeção.

Um estudo<sup>(26)</sup> com análise acústica multivariada baseada em medidas cepstrais e espectrais identificou valores de *CPP* mais baixados em indivíduos disfônicos com tensão vocal em relação a indivíduos vocalmente saudáveis, além de um deslocamento do pico cepstral dominante para as frequências mais elevadas. Os autores também observaram uma correlação negativa forte entre o GT e as medidas cepstrais na fala encadeada, em detrimento de uma correlação negativa fraca na vogal sustentada, semelhante ao que foi encontrado no presente estudo.

A partir dos dados observados na presente pesquisa, as medidas acústicas cepstrais foram fortes indicadores da presença e intensidade do desvio vocal, assim como do GS, além de contribuir na diferenciação de diferentes tipos de desvios vocais. Os achados com relação à avaliação do GR e GT reforçam a importância

tendência atual de utilização de análise acústica multivariada, visto que nenhuma medida isolada é capaz de promover uma análise confiável de sinais com diferentes componentes de irregularidade, ruído e tensão concomitantes. De modo geral, os dados deste estudo demonstraram que as medidas cepstrais constituem uma ferramenta confiável de quantificação do desvio vocal, produzindo estimativas de aperiodicidade e/ou ruído aditivo, sem a necessidade de identificação de limites dos ciclos individuais<sup>(11)</sup>.

Deve-se destacar que no presente estudo apenas a vogal sustentada foi utilizada para a avaliação da relação entre as medidas cepstrais e a análise perceptivo-auditiva. Desse modo, sugere-se a avaliação com outras tarefas, como a fala automática e frases do CAPE-V, identificando-se a melhor tarefa para análise cepstral no contexto do Português Brasileiro. Além disso, faz-se necessário estabelecer o ponto de corte e o poder discriminatório dessas medidas para as diferentes tarefas de fala no Português Brasileiro, assim como para diferentes padrões de referência (exame visual laríngeo, avaliação perceptivo-auditiva e autoavaliação vocal).

Uma das possíveis limitações deste estudo também pode estar relacionada aos valores de referência utilizados para classificação das vozes em diferentes graus de desvio, visto que o estudo original de validação dos pontos de corte<sup>(15)</sup> utilizou fala encadeada e a presente pesquisa utilizou vogal sustentada. Dessa forma, reforça-se a importância de novos estudos com o *CPPS* na mesma tarefa de fala encadeada utilizado no estudo de Yamasaki et al.<sup>(15)</sup>, observando-se se há uma correspondência entre os achados do *CPPS* em diferentes graus na vogal sustentada e fala encadeada, mesmo utilizando valores de corte não definidos previamente para a vogal sustentada.

## CONCLUSÃO

Existe associação entre a presença de desvio vocal, o GG, a qualidade vocal predominante e o *CPPS*. Vozes desviadas apresentam menores valores do *CPPS* em relação às vozes saudáveis. As vozes com domínio de tensão possuem maiores valores do *CPPS* em relação às vozes predominantemente rugosas e soprosas. Vozes rugosas mostram maiores valores do *CPPS* em detrimento das vozes soprosas. O GG e o GS apresentam correlação negativa forte com o *CPPS*, enquanto o GR e o GT demonstram correlação negativa moderada e forte com o *PPS*, respectivamente. O declínio espectral está associado apenas aos parâmetros de sopro e tensão.

## 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 new assessment techniques: Guideline elaborated by the Committee on Phoniatrics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol*. 2001;258(2):77-82. <http://dx.doi.org/10.1007/s004050000299>. PMID:11307610.
2. Hunter EJ, Titze IR. Quantifying vocal fatigue recovery: dynamic vocal recovery trajectories after a vocal loading exercise. *Ann Otol Rhinol Laryngol*. 2009;118(6):449-60. <http://dx.doi.org/10.1177/000348940911800608>. PMID:19663377.

3. Awan SN, Roy N, Jetté ME, Meltzner GS, Hillman RE. Quantifying dysphonia severity using a spectral/cepstral-based acoustic index: comparisons with auditory-perceptual judgements from the CAPE-V. *Clin Linguist Phon.* 2010;24(9):742-58. <http://dx.doi.org/10.3109/02699206.2010.492446>. PMID:20687828.
4. Awan SN, Helou LB, Stojadinovic A, Solomon NP. Tracking voice change after thyroidectomy: application of spectral/cepstral analyses. *Clin Linguist Phon.* 2011;25(4):302-20. <http://dx.doi.org/10.3109/02699206.2010.535646>. PMID:21158501.
5. Uloza V, Verikas A, Bacauskiene M, Gelzinis A, Pribuisiene R, Kaseta M, et al. Categorizing normal and pathological voices: automated and perceptual categorization. *J Voice.* 2010;25(6):700-8. <http://dx.doi.org/10.1016/j.jvoice.2010.04.009>. PMID:20579842.
6. Barsties B, De Bodt M. Assessment of voice quality: current state-of-the-art. *Auris Nasus Larynx.* 2015;42(3):183-8. <http://dx.doi.org/10.1016/j.anl.2014.11.001>. PMID:25440411.
7. Lopes LW, Alves GAS, Melo LM. Evidência de conteúdo de um protocolo de análise espectrográfica. *Rev CEFAC.* 2017;19(4):510-28. <http://dx.doi.org/10.1590/1982-021620171942917>.
8. Lopes LW, Cavalcante DP, Costa PO. Intensidade do desvio vocal: integração de dados perceptivo-auditivos e acústicos em pacientes disfônicos. *CoDAS.* 2014;26(5):382-8. <http://dx.doi.org/10.1590/2317-1782/20142013033>. PMID:25388071.
9. Godino-Llorente JI, Osma-Ruiz V, Sáenz-Lechón N, Gómez-Vilda P, Blanco-Velasco M, Cruz-Roldán F. The 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.
10. Awan SN, Roy N. Outcomes measurement in voice disorders: application of an acoustic index of dysphonia severity. *J Speech Lang Hear Res.* 2009;52(2):482-99. [http://dx.doi.org/10.1044/1092-4388\(2008/08-0034\)](http://dx.doi.org/10.1044/1092-4388(2008/08-0034)). PMID:19339702.
11. Dejonckere PH, Wieneke GH. Cepstral of normal and pathological voices: correlation with acoustic, aerodynamic and perceptual data. In: Ball MJ & Duckworth M. editors. *Advances in clinical phonetics.* Amsterdam: John Benjamins; 1996. p. 217-226. <http://dx.doi.org/10.1075/sspcl.6.13dej>.
12. Awan SN, Roy N, Dromey C. Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model. *Clin Linguist Phon.* 2009;23(11):825-41. <http://dx.doi.org/10.3109/02699200903242988>. PMID:19891523.
13. Wolfe VI, Martin DP, Palmer CI. Perception of dysphonic voice quality by naive listeners. *J Speech Hear Res.* 2000;43(3):697-705. <http://dx.doi.org/10.1044/jslhr.4303.697>. PMID:10877439.
14. Maryn Y, Weenink D. Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and acoustic voice quality index. *J Voice.* 2015;29(1):35-43. <http://dx.doi.org/10.1016/j.jvoice.2014.06.015>. PMID:25499526.
15. 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.
16. 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.
17. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005;85(3):257-68. PMID:15733050.
18. Dancy C, Reidy J. Estatística sem matemática para psicologia: usando SPSS para Windows. Porto Alegre: Artmed; 2006.
19. Brockmann-Bauser 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.0b013e32834575fe>. PMID:21483265.
20. Awan SN, Roy N, Zhang D, Cohen SM. Validation of the Cepstral Spectral Index of Dysphonia (CSID) as a screening tool for voice disorders: development of clinical cutoff scores. *J Voice.* 2015;30(2):1-15. PMID:26361215.
21. McAllister A, Sederholm E, Ternström S, Sundberg J. Perturbation and hoarseness: a pilot study of six children's voices. *J Voice.* 1996;10(3):252-61. [http://dx.doi.org/10.1016/S0892-1997\(96\)80006-3](http://dx.doi.org/10.1016/S0892-1997(96)80006-3). PMID:8865096.
22. Watts CR, Awan SN. An examination of variations in the cepstral spectral index of dysphonia across a single breath group in connected speech. *J Voice.* 2015;29(1):26-34. <http://dx.doi.org/10.1016/j.jvoice.2014.04.012>. PMID:25108589.
23. Awan SN, Solomon NP, Helou LB, Stojadinovic A. Spectral-Cepstral estimation of dysphonia severity: external validation. *Ann Otol Rhinol Laryngol.* 2013;122(1):40-8. <http://dx.doi.org/10.1177/000348941312200108>. PMID:23472315.
24. Awan SN, Roy N. Acoustic prediction of voice type in women with functional dysphonia. *J Voice.* 2005;19(2):268-82. <http://dx.doi.org/10.1016/j.jvoice.2004.03.005>. PMID:15907441.
25. Barsties B, Maryn Y. External validation of the acoustic voice quality index version 03.01 with extended representativity. *Ann Otol Rhinol Laryngol.* 2016;125(7):571-83. <http://dx.doi.org/10.1177/0003489416636131>. PMID:26951063.
26. Lowell SY, Kelley RT, Awan SN, Colton RH, Chan NH. Spectral- and cepstral-based acoustic features of dysphonic, strained voice quality. *Ann Otol Rhinol Laryngol.* 2012;121(8):539-48. <http://dx.doi.org/10.1177/000348941212100808>. PMID:22953661.
27. Watts CR, Awan SN. Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *J Speech Lang Hear Res.* 2011;54(6):1525-37. [http://dx.doi.org/10.1044/1092-4388\(2011/10-0209\)](http://dx.doi.org/10.1044/1092-4388(2011/10-0209)). PMID:22180020.
28. Awan SN, Krauss AR, Herbst CT. An examination of the relationship between electroglottographic contact quotient, electroglottographic decontacting phase profile, and acoustical spectral moments. *J Voice.* 2014;29(5):519-29. <http://dx.doi.org/10.1016/j.jvoice.2014.10.016>. PMID:25795367.
29. 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.

### Contribuição dos autores

*Todos os autores (ESSS, LWL, ACFS, IMS, MAAP, VJDV, AAA) contribuíram na construção e desenvolvimento deste trabalho. ESSS, LWL, ACFS, IMS, MAAP realizaram a coleta de dados e análise dos resultados; ESSS, LWL, VJDV, AAA realizaram a revisão do manuscrito.*