

# MEDIDAS CEPSTRAIS NA AVALIAÇÃO DA INTENSIDADE DO DESVIO VOCAL EM DIFERENTES TAREFAS DE FALA

## *Cepstral measures in the assessment of severity of voice disorders in different speech tasks*

Emanuel Gustavo Rodrigues Diniz <sup>(1)</sup>, Leonardo Wanderley Lopes <sup>(1)</sup>

(1) Departamento de Fonoaudiologia da Universidade Federal da Paraíba – UFPB, João Pessoa, Paraíba, Brasil.

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

### RESUMO

**Objetivo:** verificar se existe correlação entre as medidas cepstrais e a intensidade do desvio vocal em diferentes tarefas de fala em indivíduos com queixa vocal. **Método:** Foram utilizadas 873 amostras da vogal /ε/, contagem e frases do CAPE-V 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). O *Cepstral Peak Prominence-Smoothed (CPPS)* foi extraído de cada uma das tarefas. **Resultados:** Houve diferença dos valores do *CPPS* entre os grupos com e sem desvio vocal e entre as diferentes intensidades e tipos de desvio vocal em todas as tarefas. Na comparação com e sem lesão laríngea houve diferença para *CPPS* vogal e frases do CAPE-V. Apenas *CPPS* vogal diferencia vozes rugosas x soprosas, rugosas x tensas e soprosas x tensas. *CPPS* vogal se correlacionou de modo negativo e forte com GG e negativo moderado com GR e GS. *CPPS* contagem se correlacionou negativamente moderado com GG e GR. *CPPS* frases do CAPE-V se correlacionou de modo negativo moderado com o GG, GR e GS. **Conclusão:** Existe associação entre a presença de desvio vocal, a qualidade vocal predominante e o CPSS em quaisquer das tarefas estudadas. Há associação entre presença de lesão laríngea e o *CPPS* nas tarefas de vogal e frases do CAPE-V. Os GG, GR e GS apresentam correlações de moderada a forte com *CPPS* nas tarefas utilizadas. De todas as tarefas apenas *CPPS* vogal apresentou correlação forte com GG.

**Descritores:** Acústica; Qualidade Vocal; Distúrbio da voz; Voz; Fonoaudiologia.

## ABSTRACT

**Porpouse:** To verify whether there is a correlation between cepstral measures and the intensity of vocal deviation in different speech tasks in individuals with voice complaints. **Methods:** We used 376 samples of the vowel /  $\epsilon$  /, count and CAPE-V sentences of individuals of both genders, with vocal complaint. An analogue-visual scale was used for auditory-perceptual analysis of voices regarding vocal deviation intensity (GG), degree of roughness (GR), breathiness (GS) and tension (GT), including the determination of vocal quality. predominant (rough, breathy or tense). Cepstral Peak Prominence-Smoothed (CPPS) was extracted from each task. **Results:** There was a difference in CPPS values between the groups with and without vocal deviation and between the different intensities and types of vocal deviation in all tasks. Comparing with and without laryngeal lesion, there was a difference for CPPS vowel and CAPE-V sentences. Only CPPS vowel differentiates rough x breathy, rough x strained, and breathy x strained voices. CPPS vowel correlated negatively and strongly with GG and moderate negative with GR and GS. CPPS count correlated negatively moderately with GG and GR. CPPS CAPE-V sentences correlated moderately negatively with GG, GR, and GS. **Conclusion:** There is an association between the presence of vocal deviation, the presence of laryngeal injury, the predominant vocal quality and the CPSS in any of the tasks studied. There is an association between presence of laryngeal lesion and CPPS in vowel tasks and CAPE-V sentences. GG, GR and GS present moderate to strong correlations with CPPS in the tasks used. Of all tasks, only vowel CPPS showed a strong correlation with GG.

**Keywords:** Acoustics; Voice Quality; Voice Disorder; Voice; Speech-Language Pathology.

## INTRODUÇÃO

A produção vocal pode ser definida como resultado de um processo integrativo que reúne diversos fatores anatomofisiológicos, comportamentais, emocionais, orgânicos e ambientais referentes ao indivíduo. Para realizar uma boa avaliação vocal é necessário considerar todos esses aspectos no intuito de se obter uma visão global das disfonias<sup>(1)</sup>. Para isso, as avaliações multidimensionais da voz geralmente incluem questionários de autoavaliação, avaliações perceptivo auditivas, acústicas e visuais<sup>(2)</sup>.

A avaliação perceptivo auditiva é o principal instrumento utilizado na clínica fonoaudiológica e permite classificar subjetivamente as vozes quanto à qualidade e intensidade dos desvios. No entanto, pode sofrer interferências em situações como: desvio vocal leve, disфонia inconsistente, pouca experiência do avaliador, atenção ou fadiga do ouvinte<sup>(1,3)</sup>. Por sua vez, a análise acústica, de caráter objetivo, permite extração de medidas que possibilitam a quantificação ou qualificação de aspectos do sinal vocal, também referência na clínica e pesquisa fonoaudiológicas<sup>(4)</sup>.

Em relação à análise acústica, existe um problema no uso de medidas mais tradicionais, como *jitter* e *shimmer*. Estas medidas são mais utilizadas para análise de vogais sustentadas, pois consideram a individualização dos ciclos glóticos e análise de periodicidade no tempo. Ao se estender estes métodos para a análise de amostras de fala, os dados podem ser alterados em decorrência dos efeitos de entonação e segmentos desvozeados como também do tempo curto de duração das vogais dentro de um contexto de fala<sup>(5)</sup>.

O uso de medidas cepstrais pode ser uma alternativa razoável para contornar tais problemas. Como, por exemplo, o uso do *Cepstral Peak Proeminence Smoother* (CPPS), medida de maior sucesso para mensuração objetiva da qualidade vocal e recomendada pela *American Speech-Language-Hearing Association* (ASHA) como medida acústica eficaz para avaliar tanto a quantidade de ruído, quanto a qualidade global da voz. Válida e confiável tanto em contextos de vogais sustentadas quanto de fala<sup>(5,6,7)</sup>.

O CPPS consiste basicamente na extração da frequência fundamental ( $F_0$ ) de um sinal vocal que é transformado em um cepstro. É gerado a partir da aplicação da Transformada de Fourier, fazendo com que a energia espectral de  $F_0$  seja separada do ruído de fundo, gerando picos cepstrais. Depois, é calculada uma regressão linear entre frequência e magnitude cepstral para normalizar a amplitude total do sinal. Antes de se extrair os picos, o sinal é suavizado, através do alisamento do sinal. Dessa forma, o CPPS é exatamente o resultado da diferença entre a amplitude do pico cepstral e seu valor correspondente na linha de regressão linear. Assim, quanto maior o pico, mais distante do ruído está a  $F_0$ <sup>(7,8)</sup>.

No entanto, também há dúvidas relacionadas com o melhor tipo de tarefa a ser utilizada como base da avaliação. A escolha entre o uso de vogais sustentadas ou de fala encadeada na perceptivo auditiva pode influenciar no julgamento do avaliador, tornando o resultado questionável<sup>(9)</sup>. Na análise acústica, os métodos de análise mais tradicionalmente empregados como *jitter* e *shimmer*, só demonstram confiabilidade

em amostras de vogais sustentadas e em vozes normais ou levemente desviadas. Dessa forma, existe uma dificuldade em se utilizar tarefas de fala encadeada nas avaliações<sup>(10)</sup>. Porém, isso não ocorre para o *CPPS*, uma vez que ele é robusto tanto para amostras de fala quanto de vogal devido ao seu princípio de extração que não leva em consideração ciclos glóticos, mas sim relação sinal/ruído a partir da  $F_0$ <sup>(11)</sup>.

Mesmo assim, tanto em tarefas de voz como de fala, o grau da disфонia pode ser percebido distintamente nestes diferentes contextos, desempenhando funções essenciais na avaliação vocal. Por exemplo, a fala contínua resulta em diferentes padrões vocais e sonoros que são relativamente representativos do uso diário da voz. Já as vogais sustentadas, evocam uma fonação mais controlada, razoavelmente estável e um pouco menos natural<sup>(12)</sup>.

Observa-se que estes dois modelos avaliativos associados às diferentes tarefas de fala se complementam em suas potencialidades e falhas, e integrados permitiriam a obtenção de dados robustos que corroborariam com avaliações e diagnósticos ecologicamente válidos para a gravidade das disfonias<sup>(2,12)</sup>.

Desse modo, considera-se que o *CPPS* possibilita analisar sinais com ampla faixa de desvio, possui correlação com o desvio percebido auditivamente, possibilita análise de segmentos vocálicos e, principalmente, de fala encadeada e é uma medida acessível em *softwares* livres de análise vocal. A partir disso, esse estudo buscou verificar se existe correlação entre as medidas cepstrais e a intensidade do desvio vocal em diferentes tarefas de fala em indivíduos com queixa vocal.

## **MÉTODO**

### **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 873 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.

Assim, foram selecionados 873 indivíduos, com idade média de  $41,20 \pm 14,04$  anos. 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.

Todos os indivíduos participantes apresentavam queixa vocal e receberam confirmação diagnóstica do distúrbio de voz por meio do exame visual laríngeo. 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 / $\text{E}$ / sustentada e contagem de 1 a 10 e frases do Protocolo CAPE-V<sup>(13)</sup>, a saber: “Olha lá o avião azul”, “Sônia sabe sambar sozinha”, “Érica tomou suco de pera e amora”, “Minha mãe namorou um anjo”, “Papai trouxe pipoca quente”, “Agora é hora de acabar”.

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 [ $\text{E}$ ] sustentada, contagem de 1 a 10 e as frases do Protocolo CAPE-V<sup>(13)</sup> em frequência e intensidade autorreferida como habitual. A vogal [ $\text{E}$ ] 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. As tarefas duas tarefas de fala contínua foram escolhidas por apresentarem diferentes padrões de voz e sons mais próximos do uso diário da voz, uma de forma automática (contagem) e outra por repetição ou leitura (frases do CAPE-V).

Posteriormente, as vozes foram editadas no software *Audacity 2.3.0*, sendo eliminados os dois segundos iniciais e finais da emissão das tarefas de vogal e contagem. Quanto às frases do CAPE-V, foram editadas excluindo-se os intervalos de silêncio pré e pós emissão para cada frase dentro do mesmo arquivo, 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)*.

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>(14)</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 PowerCepstrogram*”.
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, Weenink (2015)<sup>(15)</sup>.

Todos os valores do *CPPS* 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>(16)</sup>, para a avaliação da intensidade do desvio vocal (GG) e dos graus de rugosidade (GR), de sopro (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>(16)</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 sopro 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, sopro e tensão foram escolhidos para caracterização dos sinais neste estudo por serem utilizados universalmente para caracterizar o desvio de qualidade vocal<sup>(13)</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 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, sopro 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.

Os valores do ponto de corte da EAV<sup>(16)</sup> foram utilizados para classificar as vozes quanto à presença de desvio vocal e GG. 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>(16)</sup>: VNQV (0-35,5 mm); grau 2 (35,6-50,5 mm), desvio de leve a moderado; grau 3 (50,6-90,5 mm), desvio moderado; grau 4 (90,6-100 mm), desvio intenso.

Deve-se destacar que o estudo de referência<sup>(16)</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>(16)</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 e com e sem lesão laríngea. O teste de *Kruskal-Wallis* foi utilizado para comparar a média das medidas cepstrais em função qualidade vocal predominante, utilizando-se o teste *post hoc* da Anova 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>(17)</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

Para comparar a média das medidas cepstrais nas tarefas de vogal sustentada, contagem e frases do CAPE-V entre os grupos com e sem desvio vocal foi realizado o teste não paramétrico de *Mann-Whitney* (Tabela 1). Houve diferença dos valores de *CPPS* entre os grupos com e sem desvio ( $p < 0,001$ ), com valores maiores nos indivíduos sem desvio vocal em todas as tarefas.

Na Comparação das médias das medidas cepstrais entre os grupos com e sem lesão laríngea (Tabela 2), realizada pelo teste não paramétrico de *Mann-Whitney*, apenas *CPPS* em vogal e frases do CAPE-V apresentou diferença ( $p < 0,001$ ).

Foi realizado o teste não paramétrico de *Kruskal-Wallis* para comparar as medidas cepstrais em função da qualidade vocal predominante (Tabela 3). *CPPS* vogal, contagem e frases do CAPE-V apresentaram diferença ( $p < 0,001$ ) em relação a qualidade vocal predominante seja de rugosidade, sopro ou tensão. Na análise *post hoc* (Tabela 4) *CPPS* vogal diferenciou vozes rugosas de soprosas ( $p = 0,029$ ), rugosas de tensas ( $p = 0,027$ ) e soprosas de tensas ( $p < 0,001$ ).

Por fim, realizou-se o teste de correlação de *Spearman* entre as medidas perceptivo-auditivas e cepstrais (Tabela 5). *CPPS* vogal apresentou correlação negativa forte com GG ( $p < 0,001$ ), negativa moderada com GR e GS ( $p < 0,001$ ) e negativa fraca com tensão ( $p < 0,001$ ). *CPPS* contagem apresentou correlação negativa moderada com GG e GR ( $p < 0,001$ ) e negativa fraca com GS e GT ( $p < 0,001$ ). No *CPPS* frases do CAPE-V houve correlação negativa moderada com GG, GR, GS e negativa fraca com GT ( $p < 0,001$ ).

## DISCUSSÃO

Os resultados deste estudo confirmam a robustez atribuída ao *CPPS* como medida preditora de distúrbios vocais<sup>(5,18,19,20,21,22)</sup>. Na tabela 1, ao comparar a média das medidas cepstrais entre os grupos com e sem desvio vocal se observa que o *CPPS* foi capaz de predizer, em qualquer uma das tarefas utilizadas (vogal, contagem e frases), quando a voz do indivíduo estava dentro da normalidade ou desviada. Isso significa que independentemente da tarefa, o *CPPS* consegue distinguir o quanto uma voz é saudável ou não a partir da separação de  $F_0$  do restante do ruído do sinal, sem apresentar problemas no que se refere às dificuldades encontradas nas extrações de medidas de fala por medidas tradicionais<sup>(5,11)</sup>.

Deste modo, os resultados encontrados corroboram com alguns estudos<sup>(5,11,18,22,23)</sup> que usaram *CPP* e *CPPS*, buscando identificar grupos com e sem desvio vocal a partir da comparação entre amostras de vogais e frases. Todos eles assumiram essas duas medidas como capazes de diferenciar vozes normais de disfônicas em quaisquer das tarefas utilizadas.

Na comparação das medidas cepstrais entre os grupos com e sem lesão laríngea, apenas *CPPS* de vogal e frases do CAPE-V apresentaram significância (Tabela 2). Dessa forma, *CPPS* conseguiu diferenciar quando há ou não lesão laríngea mesmo que ela não necessariamente altere a qualidade vocal.

Ao comparar grupos disfônicos (com lesão laríngea) e não disfônicos (sem lesão laríngea) em tarefas com frases, algumas pesquisas<sup>(11,23)</sup> também apresentaram resultados com forte correlação para diferenciar os grupos. Na correlação *CPPS* x vogal também foram encontradas diferenças nos grupos com e sem lesão<sup>(23,24)</sup>. No entanto, vale ressaltar que eles não utilizaram as mesmas tarefas do presente estudo. Consideraram para as suas amostras a extração da vogal /a/ e de outros tipos de sentenças que não pertenciam ao CAPE-V.

O porquê da tarefa de contagem não se mostrar significativa nessa comparação também pode estar influenciado pela quantidade do material linguístico da amostra. A curta duração da emissão da contagem pode não ser suficiente para o julgamento do desvio vocal, influenciada por questões suprasegmentais da fala. Já nas vogais, isso não ocorre, uma vez que a captação da amostra é diretamente de fonte glótica (região da lesão) sem interferência do trato vocal. Nas frases do CAPE-V, embora haja também interferências de fala, existe um maior tempo de emissão que pode favorecer um melhor julgamento da alteração vocal, somada a presença de maior diversidade linguística e de sinal vocal se comparada com a tarefa de contagem<sup>(20)</sup>.

Ainda, o fato de a tarefa de contagem não apresentar diferença nessa comparação pode ser encontrado no estudo que buscou encontrar qual a melhor tarefa de fala para validação do *Acoustic Voice Quality Index (AVQI)* para o português brasileiro. Foi verificado que as frases do CAPE-V produziram uma melhor acurácia diagnóstica que a contagem de 1 a 10. Evidenciando forte correlação perceptivo-acústica<sup>(25)</sup> que pode ser atribuída ao resultado encontrado neste estudo entre o *CPPS* e as frases do CAPE-V.

Quando se tratou elencar a qualidade vocal predominante das amostras, as três tarefas foram capazes de diferenciar quais vozes eram rugosas, soprosas e tensas (Tabela 3). Estes resultados corroboram com os achados de Lopes et. al (2019)<sup>(22)</sup>, no qual as vozes com predomínio de tensão apresentaram maiores valores do *CPPS* em relação às vozes predominantemente rugosas e soprosas. E as vozes rugosas apresentam maiores valores do *CPPS* em detrimento das vozes soprosas.

Um estudo de meta-análise<sup>(26)</sup> verificou a correlação do julgamento perceptivo-auditivo com várias medidas acústicas em diferentes tarefas e mostrou que: *CPPS* e *CPP* foram as medidas acústicas que mais fortemente se correlacionaram com rugosidade em tarefas de fala contínua ( $r = 0.6$ ); e para soprosidade, na fala contínua, das 21 medidas investigadas, *CPPS* e *CPP* foram as melhores preditoras desse parâmetro.

A justificativa para os maiores valores de *CPPS* nas vozes tensas pode estar relacionada com a sua produção. Na maioria das vezes, a tensão fonatória é gerada pelo aumento na contração da musculatura intrínseca e extrínseca da laringe, somada ao aumento da pressão subglótica e tempo da fase fechada do ciclo glótico. Assim, produz um sinal com maior nível de energia e maior definição de  $F_0$  que se evidencia na amplitude do pico cepstral<sup>(11,22)</sup>.

Nas vozes soprosas há diminuição da energia abaixo de 2500 Hz e aumento de energia nas faixas de frequência mais elevadas geradas pela maior separação e concavidade da borda livre das pregas vocais e menor tempo da fase fechada do ciclo glótico. Dessa forma, os menores valores do *CPPS* em vozes soprosas são justificáveis, uma vez que o aumento de ruído em frequências elevadas influencia no rebaixamento dos valores de *CPPS*<sup>(11,22)</sup>.

Já as vozes rugosas possuem quantidade de ruído maior nas baixas frequências do que nas altas, gerando menos ruído e, por fim, se relacionando com os maiores valores do *CPPS* encontrados em vozes rugosas quando comparadas às soprosas<sup>(22)</sup>.

Vale destacar que os maiores valores de média foram encontrados na tarefa de vogal. Isso pode estar correlacionado com a menor quantidade de ruído adicional causado pelo trato vocal, uma vez que é uma amostra direta de fonte glótica<sup>(20)</sup>.

Na comparação pós-teste (Tabela 4) apenas o *CPPS* vogal foi capaz de diferenciar os tipos de vozes entre si (rugosas de soprosas, rugosas de tensas e soprosas de tensas). Mais uma vez se destaca a diferença do material linguístico observado na contagem e frases do CAPE-V e a interferência do trato vocal. Esses parâmetros de qualidade vocal são melhor detectados em vogais, uma vez que são inicialmente gerados em fonte glótica e só depois sofrem acréscimos de filtro<sup>(20)</sup>.

Um dos critérios para a excelência de uma medida acústica é que ela represente o desvio percebido auditivamente com fidedignidade, quantifique a função vocal e tenha uma boa correlação com o julgamento perceptivo-auditivo<sup>(22)</sup>. Por isso, verificamos essa correlação do *CPPS* nas tarefas estudadas (vogal, contagem, frases do CAPE-V) com a intensidade do desvio vocal e os GG, GR, GR, GT (Tabela 5).

No *CPPS* vogal a correlação negativa forte com GG ( $p < 0,001$ ) indica a sua predisposição para diferenciação entre vozes normais e desviadas através desse parâmetro. Isso significa que quanto menor o valor do *CPPS* maior é o grau de severidade da disfonia<sup>(11,22,27)</sup>. A correlação negativa moderada com GR e GS ( $p < 0,001$ ) confirma também alguns achados que apontam *CPPS* como bom preditor desses parâmetros<sup>(28,29)</sup>. No entanto, há um estudo que encontrou correlação negativa forte para GS e negativa moderada para GR<sup>(22)</sup>.

Já GT, com correlação negativa fraca ( $p < 0,001$ ) indica a dificuldade do *CPPS* vogal em discriminar esse parâmetro, corroborando com o estudo de Lopes et. al (2019)<sup>(22)</sup>. Isso acontece devido ao fenômeno fisiológico na produção da tensão. Ela provoca uma contração excessiva da musculatura vocal que gera maior quantidade de energia concentrada em  $F_0$  ou em altas frequências, como no caso de vozes saudáveis que produzem tensão. Além disso, o GT pode gerar confusão na sua identificação em amostras de vogais se comparado a de fala, principalmente a depender da experiência do juiz<sup>(22)</sup>.

Para a tarefa de contagem, *CPSS* apresentou correlação negativa moderada com GG e GR ( $p < 0,001$ ) e negativa fraca com GS e GT ( $p < 0,001$ ). O que indica a dificuldade dessa tarefa em diferenciar o ruído nas frequências altas. Infere-se que o ruído gerado pode ser confundido com alguns fonemas do português brasileiro produzidos durante a fala<sup>(20,22)</sup>.

Já no *CPPS* frases do CAPE-V houveram melhores correlações do que no *CPPS* contagem. Apresentou correlação negativa moderada com GG, GR, GS e somente negativa fraca com GT ( $p < 0,001$ ). Isso implica que a quantidade do material linguístico apresentado nas frases do CAPE-V, bem como seu balanceamento fonético, pode ser mais bem utilizada para correlações com a intensidade e grau de desvio do que o material de fala oferecido pela contagem<sup>(20,22)</sup>.

Dessa forma, entre as três tarefas utilizadas, *CPPS* vogal se mostrou mais fortemente correlacionada com a maioria dos graus (GG, GR, GS) do que as tarefas de fala encadeada (contagem e frases do CAPE-V).

## CONCLUSÃO

Assim, existe associação entre a presença de desvio vocal, a qualidade vocal predominante e o *CPPS* nas tarefas de vogal, contagem e frases do CAPE-V. Há associação entre a presença de lesão laríngea e o *CPPS* nas tarefas de vogal e frases do CAPE-V. Vozes desviadas e com lesão laríngea apresentam menores valores do *CPPS*, em todas as tarefas, em relação às vozes sem desvio e lesão. As vozes com domínio de tensão possuem maiores valores do *CPPS*, em todas as tarefas, em relação às vozes predominantemente rugosas e soprosas. Vozes rugosas mostram maiores valores do *CPPS* em quaisquer tarefas em detrimento das vozes soprosas. Para *CPPS* vogal, houve correlação negativa forte para GG e negativa moderada para GR, GS. Para *CPPS* contagem houve correlação negativa moderada com GG e GR. *CPPS* frases do CAPE-V apresenta correlação negativa moderada com GG, GR, GS.

## REFERÊNCIAS

1. Lopes LW, Cavalcante DP, Costa PO. Intensidade do desvio vocal: integração de dados perceptivo-auditivos e acústicos em pacientes disfônicos. *CoDAS*. 2014; 26 (5): 382-88.
2. Kim GH, Lee JS, Lee CY, Lee YW, Bae IH, Park HJ, Lee BJ, Kwon SB. Effects of Injection Laryngoplasty with Hyaluronic Acid in Patients with Vocal Fold Paralysis. *Osong Public Health Res Perspect*. 2018; 9 (6): 354-61.
3. Awan SN, Roy N. Acoustic Prediction of Voice Type in Women with Functional Dysphonia. *Journal of Voice*. 2005; 2 (19): 268-82.
4. Lopes LW, Alves GAS, Melo LM. Evidência de conteúdo de um protocolo de análise espectral. *Rev. CEFAC*. 2017; 19 (4): 510-28.
5. 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. *Clinical Linguistics & Phonetics*. 2010; 24 (9): 742-758.
6. Riesgo CAF, Nöth E. What Makes the Cepstral Peak Prominence Different to Other Acoustic Correlates of Vocal Quality. *Journal of Voice*. 2019.
7. Awan SN, Solomoon NP, Helou LB, Stojadinovic A. Spectral-Cepstral Estimation of Dysphonia Severity: External Validation, *Annals of Otology, Rhinology and Laryngology*. 2013; 122 (1): 40-8.
8. Reddy MSB, Pebilli GK, Soonan VV. Smoothened cepstral peak prominence in choral singers, trained singers and non-singers. *JAIISH*. 2014. 33: 18-23.
9. Baravieira PB, Brasolotto AG, Montagnoli AN, Silvério KCA, Yamasaki R, Behlau M. Análise perceptivo-auditiva de vozes rugosas e soprosas: correspondência entre a escala visual analógica e a escala numérica. *CoDAS*. 2016; 28 (2): 163-7.
10. Maryn Y, Roy N, De Bodt M, Cauwenberg PV, Corthals LSP. Acoustic Measurement of overall voice quality: a meta-analysis. *J. Acoust. Soc. Am*. 2009; 126 (50).
11. 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; 54:1525-37.
12. Maryn Y, Roy N. Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. *J Soc Bras Fonoaudiol*. 2012; 24 (2): 107-12.
13. Kempster GB, Gerratt BR, Verdolini Abbott K, Barkmeier Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Americ J of Speech-Lang Pathol*. 2009; 18 (2): 124-32.

14. 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. 217-226
15. Maryn Y, Weenink D. Objective Dysphonia Measures in the Program Praat: Smoothed Cepstral Peak Prominence and Acoustic Voice Quality Index. *Journal of voice*. 2015; 29 (1): 35-43.
16. 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*. 2016; 31 (1): 67-71.
17. Dancey C, Reidy J. *Estatística sem matemática para psicologia: usando SPSS para Windows*. Porto Alegre: Artmed; 2006.
18. Awan SN, Helou LB, Stojadinovic A, Solomon Np. Tracking voice change after thyroidectomy: application of spectral/cepstral analyses. *Clinical Linguistics & Phonetics*. 2011; 25 (4): 302–320.
19. Heman-Ackah DY, Sataloff RT, Laureyns G, Lurie D, Michael DD, Heuer R, Rubin A, Eller R, Chandran S, Abaza M, Lyons K, Divi V, Lott J, Johnson J, Hillenbrand J. Quantifying the Cepstral Peak Prominence, a Measure of Dysphonia. *Journal of Voice*. 2014; 28 (6): 783-8.
20. Gerratt BR, Kreiman J, Garellek M. Comparing measures of voice quality from sustained phonation and continuous speech. *Journal of Speech, Language, and Hearing Research*. 2016; 59: 994–1001.
21. Lee YK, Kim GH, Bae IH, Park HJ, Wang SG, Kwon SB. The cut-off analysis using visual analogue scale and cepstral assessments on severity of voice disorder. *Logopedics Phoniatrics Vocology*. 2018; 43 (4): 1-6.
22. Lopes LW, Sousa ESS, Silva ACF, Silva IM, Paiva MAA, Vieira VFD, Almeida AA. Medidas cepstrais na avaliação da intensidade do desvio vocal. *CoDAS*. 2019; 31(4): 1-8.
23. Brinca LF, Batista APF, Tavares AI, Gonçalves ICE, Moreno ML. Uso de Análises Cepstrais para Diferenciar Normal de Vozes Disfônicas: Um Estudo Comparativo de Fala Conectada Versus Vogal Sustentada em Falantes Femininas de Português Europeu. *Journal of Voice*. 2014; 28 (3): 282-6.
24. Delgado-Hernández J, León-Gómez NM, Izquierdo-Arteaga LM, Llanos-Fumero Y. Cepstral Analysis of Normal and Pathological Voice in Spanish Adults. Smoothed Cepstral Peak Prominence in Sustained Vowels Versus Connected Speech. *Acta Otorrinolaringol Esp*. 2018; 69:134-140.
25. Englert M, Lima L, Constantini AC, Latoszek BB, Maryn Y, Behlau M. Acoustic Voice Quality Index - AVQI para o português brasileiro: análise de diferentes materiais

de fala. CoDAS. 2019; 31 (1): 1-7.

26. Latoszek BB, Maryn Y, Gerrits E, De Bodt MA. Meta-Analysis- Acoustic Measurement of Roughness and Breathiness. *Journal of Speech, Language, and Hearing Research*. 2018; 61: 298–323.

27. Awan SN, Roy N. Outcomes measurement in voice disorders: application of an acoustic index of dysphonia severity. *J Speech Lang Hear Res*. 2009; 52: 482–499.

28. Heman-Ackah YD, Michael DD, Goding GSJR. The Relationship Between Cepstral Peak Prominence and Selected Parameters of Dysphonia. *Journal of voice*. 2002.6 (1): 20-7

29. Halberstam B. Acoustic and Perceptual Parameters Relating to Connected Speech Are More Reliable Measures of Hoarseness than Parameters Relating to Sustained Vowels. *ORL*. 2004. 66 (2): 70-3.

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

	VNQV		VOZES DESVIADAS		p-valor
	Média	DP	Média	DP	
<b>CPPS Vogal</b>	16,26	2,31	13,65	3,42	<0,001*
<b>CPPS CTG</b>	9,543	1,98	8,169	2,0	<0,001*
<b>CPPS CAPE-V</b>	10,63	1,67	9,3	1,86	<0,001*

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

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

**Tabela 2.** Comparação das médias das medidas cepstrais entre os grupos com e sem lesão laríngea

	SEM LESÃO LARÍNGEA		COM LESÃO LARÍNGEA		p-valor
	Média	DP	Média	DP	
<i>CPPS Vogal</i>	15,21	2,68	13,19	3,44	<0,001*
<i>CPPS CAPE-V</i>	9,9	1,59	9,08	1,89	<0,001*

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

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

**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	
<i>CPPS Vogal</i>	13,86	3,52	12,75	3,07	15,43	3,15	<0,001*
<i>CPPS CTG</i>	8,19	2,06	8,01	1,97	8,86	1,76	<0,001*
<i>CPPS CAPE-V</i>	9,29	1,86	9,22	1,09	9,82	1,67	<0,001*

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

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

**Tabela 4.** Comparação pós-teste das medidas cepstrais em função da qualidade vocal predominante

Variáveis	Tipo de Voz	Diferença Média	Erro Padrão	p-valor
<b>CPPS</b> <b>Vogal</b>	Rug x Sop	1,111	0,368	0,029*
	Rug xTen	-1,592	0,524	0,027*
	Sop x Ten	-2,702	0,562	<0,001*

\*Valores significativos ( $p < 0,05$ ) – Teste *post hoc* ANOVA

**Legenda:** *CPPS* = *cepstral peak prominence smoothed*; Rug = rugosidade; Sop = sopro; Ten = tensão

**Tabela 5.** Correlação entre a intensidade do desvio vocal, os graus de rugosidade, sopro 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 Vogal</b>	-0,612	<0,001*	-0,513	<0,001*	-0,587	<0,001*	-0,192	<0,001*
<b>CPPS CTG</b>	-0,463	<0,001*	-0,447	<0,001*	-0,381	<0,001*	-0,210	<0,001*
<b>CPPS CAPE-V</b>	-0,535	<0,001*	-0,501	<0,001*	-0,456	<0,001*	-0,273	<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 sopro; GT = grau de tensão