



**UNIVERSIDADE FEDERAL DA PARAÍBA - UFPB**  
**CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA - CCEN**  
**PROGRAMA DE PÓS GRADUAÇÃO EM MODELOS DE DECISÃO E SAÚDE**

**VALIDAÇÃO DAS MEDIDAS CEPSTRAIS PARA A AVALIAÇÃO DE DISTÚRBIO  
DE VOZ EM FALANTES DO PORTUGUÊS BRASILEIRO**

**ESTEVÃO SILVESTRE DA SILVA SOUSA**

**JOÃO PESSOA-PB**

**2021**

**ESTEVÃO SILVESTRE DA SILVA SOUSA**

**VALIDAÇÃO DAS MEDIDAS CEPSTRAIS PARA A AVALIAÇÃO DE DISTÚRBO  
DE VOZ EM FALANTES DO PORTUGUÊS BRASILEIRO**

Dissertação apresentada ao Programa de Pós-Graduação em Modelos de Decisão e Saúde – Nível Mestrado do Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba.

**Linha de Pesquisa:** Modelos em Saúde

**Orientadores:**

Prof. Dr. Leonardo Wanderley Lopes

Prof. Dr. Ronei Marcos de Moraes

**JOÃO PESSOA-PB**

**2021**

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

S725v Sousa, Estevão Silvestre da Silva.

Validação das medidas cepstrais para a avaliação de  
distúrbio de voz em falantes do português brasileiro /  
Estevão Silvestre da Silva Sousa. - João Pessoa, 2021.  
84 f. : il.

Orientação: Leonardo Wanderley Lopes, Ronei Marcos de  
Moraes.

Dissertação (Mestrado) - UFPE/CCEN.

1. Acústica da fala. 2. Modelos logísticos. 3.  
Qualidade vocal. 4. Voz. 5. Distúrbios de voz. 6.  
Ponoaudiologia. I. Lopes, Leonardo Wanderley. II.  
Moraes, Ronei Marcos de. III. Título.

UFPE/BC

CDU 612.78(043)

**ESTEVÃO SILVESTRE DA SILVA SOUSA**

**VALIDAÇÃO DAS MEDIDAS CEPSTRAIS PARA A AVALIAÇÃO DE DISTÚRPIO  
DE VOZ EM FALANTES DO PORTUGUÊS BRASILEIRO**

João Pessoa, 24 de Fevereiro de 2021

**BANCA EXAMINADORA**

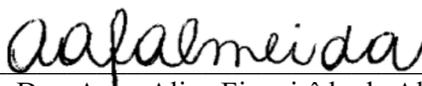


---

Prof. Dr. Leonardo Wanderley Lopes  
Orientador (UFPB)

---

Prof. Dr. Ronei Marcos de Moraes  
Orientador (UFPB)



---

Profa. Dra. Anna Alice Figueirêdo de Almeida  
Membro Interno (UFPB)

---

Profa. Dra. Liliane dos Santos Machado  
Membro Interno (UFPB)

---

Profa. Dra. Ana Cristina Côrtes Gama  
Membro Externa (UFMG)

*Para onde poderia eu escapar do teu Espírito?  
Para onde poderia fugir da tua presença?  
Se eu subir aos céus, lá estás;  
Se eu fizer a minha cama na sepultura, também lá estás.  
Se eu subir com as asas da alvorada e morar na extremidade do mar,  
Mesmo ali a tua mão direita me guiará e me susterá.*

## AGRADECIMENTOS

Primeiramente agradeço à Deus, minha base, rocha, alicerce que por sua graça me concedeu o dom da vida e que por Ele e somente por Ele tenho a capacidade de finalizar esse capítulo da minha vida.

Aos meus pais e avós que com tanta sabedoria me ensinaram a importância da educação e de ser “gente” na vida. Obrigado Welma, Wellington, Ubirajara, Valdicélia (*in memoriam*), Lúcia e Levi. À toda minha família, que através de gestos, palavras e orações me incentivaram a continuar e não desistir de alcançar meus objetivos. Em especial a minha irmã Ester Silvestre.

À minha namorada Samara Ruth que com sua sabedoria e amor esteve do meu lado, que suportou minhas ausências e mesmo de longe por muitos momentos esteve comigo.

Aos meus queridos e preciosos amigos que por muitas vezes em meio ao mar de ansiedade e medo ofereceram um bote de amor e amizade para me resgatar.

À minha segunda casa, a Igreja Cristã Nova Vida onde eu aprendo todos os dias a viver de acordo com a vontade do Senhor, agradecimento em especial na pessoa do Pr. Augusto.

Dedico essa dissertação aos meus eternos amigos de pesquisa Maxsuel Alves, Itacely Marinho e Allan Carlos. Vocês sempre farão parte da minha história. Assim como às minhas amigas de mestrado e agora de vida: Brenda Lopes, Maria Letícia e Manuela Leitão.

À Samuel Abreu que com tanta paciência me ajudou na construção desse trabalho, com sua vasta sabedoria e paciência.

Agradeço ao meu querido orientador Leonardo Wanderley Lopes que está comigo desde a iniciação científica onde me ensina a ser o melhor profissional e pesquisador que eu puder, me orgulho todos os dias em ser seu orientando.

Ao meu orientador Ronei Marcos de Moraes que com sua extrema inteligência e tranquilidade me ensinou a melhorar meu olhar científico e que por muitas vezes sem dizer nada me ensinou a questionar tudo o que me viera.

A minha banca examinadora que carregou extrema admiração e respeito. Obrigado por todo o ensino e confiança depositados em mim.

À CAPES pelo apoio financeiro.

Meu agradecimento a todos os pesquisadores negros que me permitiram o sentimento de pertencimento e representatividade.

Foram anos desafiadores, em muitos momentos cansativos, entretanto a vontade de fazer acontecer sempre foi maior. Trago comigo um amadurecimento científico e a certeza de que a ciência no Brasil continua viva.

## RESUMO

A análise acústica é um procedimento com maior objetividade, não-invasivo e que atualmente vem sendo usado comumente na clínica vocal. A análise acústica envolve medidas capazes de quantificar determinada propriedade do sinal vocal. Nos sinais vocais com desvio leve há a possibilidade de estimar valores da frequência fundamental ( $f_0$ ) através da determinação do número de ondas sonoras. Entretanto, nos sinais de voz com média ou alta irregularidade, as medidas tradicionais, que hoje são as mais utilizadas, podem ter sua confiabilidade comprometida. Diante disso, as medidas cepstrais possuem a perspectiva de determinar a  $f_0$  e produzir uma estimativa de aperiodicidade e/ou ruído aditivo sem a necessidade de identificar limites de ciclo individuais. Na análise acústica as medidas cepstrais atualmente são as medidas consideradas mais robustas para quantificar desvios de voz, entretanto ainda não se encontraram valores de cortes para falantes do português brasileiro. Sendo assim, o objetivo do presente estudo foi construir um modelo de decisão baseado nas medidas cepstrais que predizem indivíduos falantes do português brasileiro com e sem desvio da qualidade vocal. Este é um estudo descritivo, documental, transversal e observacional. A amostra desta pesquisa foi constituída por 305 indivíduos de ambos os sexos, com laudo do otorrinolaringologista (78 sem distúrbio de voz e 227 com queixa vocal e com patologia laríngea). Foram extraídas as medidas: *Cepstral Peak Prominence* (CPP), *Cepstral Peak Prominence-Smoothed* (CPPS) e declínio espectral. O poder discriminativo das medidas cepstrais se deu por intermédio de modelos de Regressão Logística com a variável: distúrbio de voz e se utilizou o modelo de regressão linear múltiplo para análise das medidas cepstrais quanto ao desvio da qualidade vocal. Além disso, também foi obtido o ponto de corte das medidas cepstrais para o português brasileiro em diferentes tarefas de fala. Se resultou que o modelo reduzido de regressão logística para identificação da presença e ausência de distúrbio de voz, inclui as variáveis: CPP CAPE-V, CPPS VGE e declínio CAPE-V, apresentando um desempenho de 70,7%, com sensibilidade de 85% e especificidade de 53%. Isoladamente, as variáveis CPPS VGE e CPPS VGA apresentaram desempenho satisfatório para discriminar indivíduos com e sem distúrbio de voz, com altos valores de especificidade e baixos valores de sensibilidade. As medidas: CPP CAPE-V, CPPS VGE, CPPS VGA e CPPS CAPE-V são capazes de prever a variabilidade da intensidade do desvio vocal em 64,32%.

Palavras-chave: Acústica da fala, Modelos Logísticos, Qualidade vocal, Voz, Distúrbios de voz, Fonoaudiologia.

## ABSTRACT

The acoustic analysis is a procedure with greater objectivity, non-invasive and that is currently being used commonly in the vocal clinic. The acoustic analysis involves measures capable of quantifying a certain property of the vocal signal. In vocal signals with mild deviation, it is possible to estimate values of the fundamental frequency ( $f_0$ ) by determining the number of sound waves. However, in voice signals with medium or high irregularity, traditional measures, which today are the most used, may have their reliability compromised. In view of this, cepstral measurements have the perspective of determining  $f_0$  and producing an estimate of aperiodicity and / or additive noise without the need to identify individual cycle limits. In the acoustic analysis, cepstral measurements are currently considered the most robust measures to quantify voice deviations, however, cut values have not yet been found for Brazilian Portuguese speakers. Thus, the objective of the present study was to build a decision model based on cepstral measures that predict individuals who speak Brazilian Portuguese with and without deviation in vocal quality. This is a descriptive, documentary, cross-sectional and observational study. The sample of this research was constituted by 305 individuals of both sexes, with the report of the otorhinolaryngologist (78 without voice disorder and 227 with vocal complaint and with laryngeal pathology). The measurements were extracted: Cepstral Peak Prominence (CPP), Cepstral Peak Prominence-Smoothed (CPPS) and spectral decline. The discriminative power of cepstral measures was given by means of Logistic Regression models with the variable: voice disorder and the multiple linear regression model was used to analyze the cepstral measures regarding the deviation of vocal quality. In addition, the cutoff point for cepstral measurements for Brazilian Portuguese in different speech tasks was also obtained. It turned out that the reduced logistic regression model for identifying the presence and absence of voice disorder, includes the variables: CPP CAPE-V, CPPS VGE and CAPE-V decline, presenting a performance of 70.7%, with a sensitivity of 85 % and specificity of 53%. In isolation, the variables CPPS VGE and CPPS VGA showed satisfactory performance to discriminate individuals with and without voice disorder, with high specificity values and low sensitivity values. The measures: CPP CAPE-V, CPPS VGE, CPPS VGA and CPPS CAPE-V are able to predict the variability in the intensity of vocal deviation by 64.32%.

Keywords: Speech acoustics, Logistic models, Vocal quality, Voice, Voice disorders, Speech therapy.

## LISTA DE QUADROS E FIGURAS

<b>Quadro 1</b> – Instrumentos de análise vocal e suas funções na avaliação multidimensional da voz.....	19
<b>Quadro 2</b> – Estudos que utilizaram análise cepstral para identificação de desvios Vocais.....	27
<b>Quadro 3</b> – Modelos Multiparamétricos e suas funções na avaliação multidimensional da voz.....	35
<b>Quadro 4</b> - Estudos utilizados como referência para a média populacional da dissertação...	38
<b>Figura 1</b> – Modelo simplificado de produção de fala (COSTA, 2008) .....	24
<b>Figura 2</b> – Cepstro de um segmento de fala (ZWETSCH, 2006) .....	25
<b>Figura 3</b> - Cepstro para uma voz normal (AGUIAR NETO et al, 2007) .....	26
<b>Figura 4</b> – Cepstro para uma voz patológica (AGUIAR NETO et al, 2007) .....	26
<b>Figura 5</b> - Fluxograma de vozes no banco de dados da atual.....	39
<b>Figura 6</b> - Exemplo de um modelo de regressão múltipla com duas variáveis independentes.....	49
<b>Figura 7</b> - Curva ROC de um teste de diagnóstico hipotético.....	52
<b>Figura 8</b> - Curva ROC para a variável distúrbio de voz.....	55
<b>Figura 9</b> – Curva ROC associada a medida CPP CAPE-V.....	59
<b>Figura 10</b> – Curva ROC associada a medida DECL CAPE-V.....	60
<b>Figura 11</b> – Curva ROC associada a medida CPP VGE.....	61
<b>Figura 12</b> – Curva ROC associada a medida CPP VGA.....	62
<b>Figura 13</b> – Curva ROC associada a medida CPP TODAS.....	63
<b>Figura 14</b> – Curva ROC associada a medida CPPS CAPE-V.....	64
<b>Figura 15</b> – Curva ROC associada a medida CPPS TODAS.....	65
<b>Figura 16</b> – Curva ROC associada a medida Decl. VGE.....	66
<b>Figura 17</b> – Curva ROC associada a medida Decl. VGA.....	67
<b>Figura 18</b> – Curva ROC associada a medida Decl. TODAS.....	68
<b>Figura 19</b> – Curva ROC associada a medida Decl. CONTAGEM.....	69
<b>Figura 20</b> – Curva ROC associada a medida CPPS VGE.....	70
<b>Figura 21</b> – Curva ROC associada a medida CPPS VGA.....	71

## LISTA DE TABELAS

<b>Tabela 1.</b> Classificação do teste coeficiente de Kappa, segundo Landis et al., (1977) .....	42
<b>Tabela 2.</b> Tabela descritiva para os grupos sem distúrbio e com distúrbio de voz.....	45
<b>Tabela 3.</b> Modelo de regressão logística para a variável: distúrbio de voz.....	53
<b>Tabela 4.</b> Modelo de regressão logística reduzido com base no método de <i>Stepwise</i> para a variável distúrbio de voz.....	54
<b>Tabela 5.</b> Valores associados à Variável CPP CAPE-V.....	59
<b>Tabela 6.</b> Valores associados à Variável DECL CAPE-V.....	60
<b>Tabela 7.</b> Valores associados à Variável CPP VGE.....	61
<b>Tabela 8.</b> Valores associados à Variável CPP VGA.....	62
<b>Tabela 9.</b> Valores associados à Variável CPP TODAS.....	63
<b>Tabela 10.</b> Valores associados à Variável CPPS CAPE-V.....	64
<b>Tabela 11.</b> Valores associados à Variável CPPS TODAS.....	65
<b>Tabela 12.</b> Valores associados à Variável DECL VGE.....	66
<b>Tabela 13.</b> Valores associados à Variável DECL VGA.....	67
<b>Tabela 14.</b> Valores associados à Variável DECL TODAS.....	68
<b>Tabela 15.</b> Valores associados à Variável DECL CONTAGEM.....	69
<b>Tabela 16.</b> Valores associados à Variável CPPS VGE.....	70
<b>Tabela 17.</b> Valores associados à Variável CPPS VGA.....	71
<b>Tabela 18.</b> Modelo de regressão linear múltipla para a variável EAV-G.....	72
<b>Tabela 19.</b> Modelo de regressão linear múltipla reduzido com base no método de <i>Stepwise</i> para a variável dependente EAV-G.....	72

## LISTA DE ABREVIATURAS E SIGLAS

<b>f0-</b>	Frequência Fundamental
<b>CSID-</b>	<i>Cepstral Spectral Index of Dysphonia</i>
<b>ADSV-</b>	<i>Analysis of Dysphonia in Speech and Voice program</i>
<b>ROC-</b>	<i>Receiver Operating Characteristic</i>
<b>LIEV-</b>	Laboratório Integrado de Estudos da Voz
<b>CPP-</b>	<i>Cepstral Peak Prominence</i>
<b>ABI-</b>	<i>Acoustic Breathiness Index</i>
<b>MFCC-</b>	<i>Mel frequency Cepstral coefficients</i>
<b>PLP-</b>	<i>Perceptual Linear Prediction</i>
<b>RASTA – PLP</b>	<i>ReAlitive SpecTrAl PLP</i>
<b>SVM-</b>	<i>Support Vector Machines</i>
<b>KNN-</b>	<i>K-Nearest Neighbors</i>
<b>DA-</b>	<i>Discriminant Analysis</i>
<b>NB-</b>	<i>Native Bayes</i>
<b>CT-</b>	<i>Classification Tree</i>
<b>AVQI-</b>	<i>Acoustic Voice Quality Index</i>
<b>CPPS-</b>	<i>Smoothed ceptral peak prominence</i>
<b>HNR-</b>	<i>Harmonic-to-noise ratio</i>
<b>DNN-</b>	<i>Deep Neural Network</i>
<b>UFPB-</b>	Universidade Federal da Paraíba
<b>CAPE-V-</b>	<i>Consensus Auditory-Perceptual Evaluation of Voice</i>
<b>EAV-</b>	Escala Analógica Visual
<b>EN-</b>	Escala Numérica
<b>VNQV-</b>	Variabilidade Normal da Qualidade Vocal
<b>VPP-</b>	Valor Preditivo Positivo
<b>VPN-</b>	Valor Preditivo Negativo
<b>RVN-</b>	Razão de Verossimilhança Negativo
<b>RVP-</b>	Razão de Verossimilhança Positivo
<b>RV-</b>	Razão de Verossimilhança
<b>GG-</b>	Grau Geral

<b>AC-</b>	Acurácia
<b>SEN-</b>	Sensibilidade
<b>ESP-</b>	Especificidade
<b>OR-</b>	<i>OddsRatio</i>
<b>D-</b>	<i>Devience</i>
<b>AUC-</b>	<i>Area Underthe Curve</i>
<b>ABI-</b>	<i>Acoustic Breathiness Index</i>
<b>DSI-</b>	<i>Dysphonia Severity Index</i>
<b>TMF-</b>	Tempo Máximo de Fonação
<b>IDV-</b>	Índice de Desvantagem Vocal
<b>QVV-</b>	Qualidade de Vida em Voz
<b>PPAV-</b>	Perfil de Participação e Atividades Vocais
<b>LTAS-</b>	<i>Long Term Avarage Spectrum</i>
<b>AIC</b>	Critério de Akaike
<b>DeCS</b>	Descritores em Ciências da Saúde
<b>GT</b>	Grupo Trabalho
<b>GC</b>	Grupo Controle

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	<b>14</b>
<b>2. OBJETIVOS.....</b>	<b>18</b>
2.1 GERAL.....	18
2.2 ESPECÍFICOS.....	18
<b>3. REFERENCIAL TEÓRICO.....</b>	<b>19</b>
3.1 ANÁLISE ACÚSTICA NO CONTEXTO DA AVALIAÇÃO MULTIDIMENSIONAL DA VOZ.....	19
3.2 ANÁLISE CEPSTRAL.....	21
3.3 ANÁLISE ACÚSTICA MULTIPARAMÉTRICA.....	33
<b>4. METODOLOGIA.....</b>	<b>37</b>
4.1 DESENHO DO ESTUDO.....	37
4.2 POPULAÇÃO DO ESTUDO.....	37
4.3 AMOSTRA DE FALA.....	40
4.4 MÉTODO DE ANÁLISE DOS DADOS.....	41
4.4.1 EXTRAÇÃO DAS MEDIDAS ACÚSTICAS.....	41
4.4.2 ANÁLISE DAS GRAVAÇÕES VOCAIS.....	42
4.4.3 ANÁLISE ESTATÍSTICA DOS DADOS .....	43
4.4.3.1 Análise Estatística Descritiva.....	44
4.4.3.2 Regressão Logística.....	45
4.4.3.3 Regressão Linear Múltipla.....	48
4.5 RAZÃO DE CHANCES ( <i>ODDS RATIO</i> ) .....	50
4.6 VERIFICAÇÃO DA QUALIDADE DO MODELOS AJUSTADO.....	50
4.7 CURVA ROC.....	51
4.8 CONSIDERAÇÕES ÉTICAS.....	53
<b>5 RESULTADOS E DISCUSSÕES.....</b>	<b>53</b>
5.1 REGRESSÃO LOGÍSTICA.....	53
5.1.1 MODELO DE REGRESSÃO LOGÍSTICA PARA VARIÁVEL DISTÚRBO DE VOZ.....	53
5.1.2 CURVA ROC PARA AS MEDIDAS CEPSTRAIS.....	57
5.2 REGRESSÃO LINEAR MÚLTIPLA.....	72

<b>6 ÍNDICE CEPSTRAL.....</b>	<b>74</b>
<b>7 CONCLUSÃO.....</b>	<b>74</b>
<b>REFERÊNCIAS.....</b>	<b>76</b>
<b>ANEXOS.....</b>	<b>83</b>

## 1 INTRODUÇÃO

A produção vocal se forma como uma parte muito importante no processo de comunicação interpessoal e profissional (BEHLAU et al., 2015). A voz envolve mecanismos fisiológicos, biomecânicos e aerodinâmicos que produzem uma saída acústica que é percebido pelo sistema auditivo. Além disso, depende da integridade e do funcionamento harmônico dos sistemas e organismos envolvidos na produção vocal. Por esse motivo, a avaliação multidimensional da voz compreende um conjunto de análises que se complementam (PATEL et al., 2018).

Basicamente, os procedimentos recomendados para a avaliação clínica da voz englobam: anamnese, exame visual laríngeo, avaliação da qualidade vocal (percepção auditiva, acústica e aerodinâmica) e a autoavaliação. Esses mecanismos se complementam e devem ser analisados de forma conjunta, com a finalidade de fornecer uma visão global da voz, estabelecer o diagnóstico, definir a conduta e, nos casos em que for necessário terapia ou treinamento, também acompanhar a evolução clínica do quadro (DEJONCKERE et al., 2001).

De acordo com os Descritores em Ciências da Saúde (DeCS), distúrbios da voz pode ser definido como sendo processos patológicos que afetam a produção da voz, geralmente envolvendo as pregas vocais e a mucosa laríngea. Os distúrbios da voz podem ser causados por fatores orgânicos (anatômicos) ou funcionais (emocionais ou psicológicos), levando à disfonia, afonia e defeitos na qualidade da voz, na intensidade e frequência.

Presentemente há uma grande variabilidade de métodos usados para avaliação do distúrbio da voz, incluindo diferenças na coleta de dados, medidas, tarefas de fala e assim por diante. Tais diferenças nos procedimentos de avaliação também são refletidas na literatura de pesquisa, o que dificulta a comparação dos resultados e sua interpretação entre os estudos, contribuindo assim para as dificuldades em recomendar diretrizes baseadas em evidências para avaliação vocal (PATEL et al., 2018).

A análise acústica tem por objetivo detalhar de forma qualitativa e quantitativa os atributos físicos de uma onda sonora no domínio do tempo. Trata-se de um método computadorizado, objetivo, não invasivo, com baixo custo, rápido, de fácil realização e interpretação. Além disso, apresenta a possibilidade de alta reprodutibilidade (EADIE et al., 2005). Um aspecto relevante é a possibilidade da quantificação do distúrbio presente no sinal e a sua comparação com dados normativos (BARSTIE & DE BODT, 2015). É importante ressaltar que a análise acústica serve como complementação das análises de percepção auditiva, autoavaliação e do exame visual

laríngeo, assim como, irá auxiliar junto com as outras análises o fechamento do diagnóstico laríngeo e no delineamento de análise e intervenção pré, durante e pós intervenção (CASMERIDES & COSTA, 2001).

São diversos os pontos positivos da análise acústica, dentre os quais podemos citar: facilita a compreensão da produção da voz, gera dados normativos, produz uma documentação vocal, monitora o resultado do tratamento, acompanha o desenvolvimento da voz e detecta de forma precoce os possíveis problemas vocais (CARRASCO et al., 2010). A diversidade de aspectos etiológicos e de sinais de um distúrbio de voz, nem sempre são possíveis de serem visualizadas através de imagens laríngeas, conseqüentemente não podendo ser utilizada como referência isolada para comprovação da ausência ou a presença de um distúrbio de voz, no que diz respeito ao diagnóstico inicial ou na avaliação pré e pós-intervenção, levando também em consideração que um distúrbio de voz pode ser qualificado por diferentes ajustes do trato vocal, ajustes esses que não são visualizados na laringoscopia (CHRISTMANN et al., 2015).

A análise acústica pode envolver a extração de medidas que quantificam determinada propriedade do sinal vocal, da mesma maneira pode incluir a análise descritiva de padrões visuais dos mesmos sinais (LOPES et. al., 2017). A análise acústica descritiva diz respeito a descrições de padrões visuais da forma da onda do sinal de voz, do espectro de energia, da espectrografia de faixa larga (formantes) ou de faixa estreita, entre outros (LOPES et. al., 2020). A análise descritiva de padrões visuais assim como a avaliação da percepção auditiva da voz, são métodos subjetivos que dependem do julgamento de um avaliador, conseqüentemente sofrem influência de variáveis como a experiência do avaliador (SOUZA & GAMA, 2015).

Na extração das medidas clássicas de perturbação (*jitter* e *shimmer*) e de ruído (proporção harmônico-ruído), precisa haver uma estimativa dos valores da frequência fundamental ( $f_0$ ), com estabelecimento claro das ondas sonoras, o que é mais provável de ser realizado apenas nos sinais vocais com distúrbios leves (LOPES et. al., 2020). Nos indivíduos com distúrbio de voz, os sinais de voz podem variar de quase periódicos a totalmente aperiódicos, de modo que a complexidade de um sinal com distúrbios moderados e intensos pode interferir na confiabilidade das medidas tradicionais baseadas em modelos lineares (DEJONCKERE et al., 2001). Portanto, embora essas medidas tradicionais apresentem uma correlação de moderada a forte com a percepção auditiva do desvio vocal (GODINO-LLORENTE et al., 2010), elas podem ter uma aplicação limitada na análise de vozes com distúrbios mais intensos.

A análise cepstral foi originalmente descrita por Noll (1964) como um procedimento para extrair a frequência fundamental do espectro de uma onda sonora. O cepstro, uma transformada de Fourier do espectro de potência do sinal de voz, mostra em forma de gráfico à medida que os harmônicos espectrais e, em particular, a frequência fundamental vocal, são individualizados e emergem do nível de ruído de fundo. Quanto mais regulares os picos harmônicos medidos em uma forma de onda, maior a amplitude do pico cepstral dominante, uma vez que a amplitude do pico cepstral é afetada pelo grau de periodicidade e energia geral do sinal de voz (HILLENBRAND et al., 1994; DEJONCKERE & WIENEKE, 1996).

Com o objetivo de apresentar uma meta-análise dos coeficientes de correlação entre o julgamento da percepção auditiva da rugosidade e soproidade e várias medidas acústicas nas vogais sustentadas e fala encadeada, Latoszek et al., (2018) pôde concluir que a medida mais investigada nos casos de rugosidade que utilizaram a fala encadeada foram as medidas cepstrais, apenas o *CPPS* conseguiu prever a rugosidade com uma correlação moderada. Já nos casos de soproidade com fala encadeada, de 21 medidas mais estudadas apenas no *CPP* e *CPPS* foram encontrados os melhores preditores acústico de soproidade.

A *American Speech-Language-Hearing Association (ASHA)* indicou o *CPP* para medição do nível geral de ruído de um sinal de voz. As maiores discussões no campo clínico e de pesquisa foram relacionadas à escolha de uma medida acústica que se correlacionasse com a percepção da qualidade vocal. Após muito debate, o *CPP* foi escolhido como uma medida geral de distúrbio de voz que reflete a relação global de energia periódica *versus* aperiódica em um sinal.

Um dos desafios da análise acústica é a escolha da melhor tarefa de fala para tal investigação, pelo fato de que a qualidade vocal pode ser diferente entre as tarefas de fala; por exemplo, a fala contínua incorpora atributos vocais importantes, como padrões de entonação (variações no tom, frequência e intensidade) que podem ser relevantes para a caracterização do distúrbio de voz em situações cotidianas e que estão ausentes nas produções de vogais sustentadas. (AWAN, 2010). Além disso, ainda não foram identificados estudos que ofereçam pontos de corte das medidas cepstrais para o português brasileiro. Visto que a maioria dos estudos que utilizam estas medidas são baseados na língua inglesa e conseqüentemente não são possíveis de aplicar para o português brasileiro, dado que características fonéticas e linguísticas podem alterar os resultados dessa avaliação (AWAN et al., 2011).

A voz é um fenômeno acústico, sendo assim, deve ser avaliada de maneira multidimensional e a extração de fatores isolados parece ser incapaz de realmente caracterizá-la. Por esse motivo,

vem aumentando o interesse por modelos acústicos multiparamétricos, tais como o *Cepstral Spectral Index of Dysphonia (CSID)*, o *Acoustic Vocal Quality Index (AVQI)*, o *Dysphonia Severity Index (DSI)*, e o *Acoustic Breathiness Index (ABI)*. (BARSTIES B & DE BODT, 2014), (AWAN et al., 2015).

Esta pesquisa justifica-se por haver a necessidade de aprofundamento acerca das medidas cepstrais no âmbito clínico, contribuindo de forma efetiva e confiável para investigação de distúrbios da voz. Além de auxiliar profissionais e pesquisadores na área de voz e gestores de saúde, a analisarem os resultados que porventura podem contribuir no processo de análise, classificação e acompanhamento clínico dos pacientes com distúrbio de voz.

Optou-se pela utilização da fala encadeada pelo fato de se assemelhar à conversação diária, mesmo essa tarefa de fala mostrando-se variável na classificação da qualidade vocal, em virtude dos fenômenos não vocais (prosódia, articulação de palavras e toda conjuntura fonética e fonológica). As vogais sustentadas serão utilizadas por ser uma das tarefas mais utilizadas na clínica para avaliação da qualidade vocal (AWAN et al., 2015). As divergências culturais, principalmente relacionadas a questão linguística, podem interferir os resultados da análise de qualidade vocal (BARSTIES & MARYN, 2016). Essa informação nos remete a importância de se obter o ponto de corte no contexto linguístico brasileiro, assim como em diferentes tarefas de fala, visto que são medidas utilizadas tanto para fala encadeada quanto para vogal sustentada.

Quando se considera a adoção de qualquer ferramenta de avaliação, é essencial que o novo método de teste objetivo seja avaliado em termos de precisão diagnóstica. Um teste impreciso pode resultar na classificação incorreta de indivíduos com distúrbio de voz como vocalmente saudáveis e em indivíduos vocalmente saudáveis com presença de distúrbio de voz (AWAN et al., 2015).

Diante do exposto no parágrafo anterior, o índice cepstral permite o desenvolvimento de uma ferramenta automatizada de avaliação objetiva com precisão conhecida, o que possibilita ao examinador um avanço potencialmente significativo na identificação de distúrbio de voz. Visto que, dependendo da natureza do distúrbio, algumas doenças estão associadas a altas taxas de morbidade, como nódulos vocais, esse quadro pode ser mudado com a identificação precoce do problema o que resulta em um diagnóstico confiável e tratamento efetivo (AWAN et al., 2015).

Desta forma, o presente trabalho se propôs a responder aos seguintes questionamentos: Qual seria o potencial discriminatório das medidas cepstrais para diferenciar indivíduos com e

sem distúrbio de voz? Assim como, qual o ponto de corte para discriminar indivíduos com e sem distúrbio de voz em diferentes tarefas de fala? A partir disso, ter como resultado de qual seria a melhor tarefa de fala a se utilizar na análise acústica da voz com a utilização de medidas cepstrais para o português brasileiro? E por fim, será que a combinação de medidas cepstrais tem uma maior potencial discriminatório?

Inicialmente, a ideia do estudo seria desenvolver valores de corte em duas categorias, para presença e ausência de distúrbio de voz e para presença e ausência de desvio vocal. No entanto, o que será apresentado nesta dissertação, terá relação com os valores de corte ao desempenho dessas medidas para identificação do distúrbio de voz.

Com a chegada ao Brasil da pandemia do Coronavírus (COVID-19), a população passou a enfrentar maiores desafios relacionados a uma reorganização de suas vidas diárias. Com o distanciamento social as relações pessoais ficaram restritas, o que culminou na interrupção dos atendimentos presenciais na relação de prestadores e usuários de serviços de saúde, tanto no âmbito clínico quanto no acadêmico. Sendo assim, não foi possível ampliar um grupo controle exclusivamente em relação à qualidade vocal, dada as questões de distanciamento social.

## **2 OBJETIVOS**

### **2.1 GERAL**

- Construir um modelo de decisão baseado nas medidas cepstrais que predizem indivíduos falantes do português brasileiro com e sem distúrbio de voz.

### **2.2 OBJETIVOS ESPECÍFICOS**

- Estabelecer o ponto de corte das medidas cepstrais para identificar indivíduos falantes do português brasileiro com e sem distúrbio de voz em diferentes tarefas de fala;
- Verificar o quanto as medidas cepstrais são explicadas pela variabilidade do desvio vocal medido pela EAV.

### 3 REFERENCIAL TEÓRICO

#### 3.1 ANÁLISE ACÚSTICA NO CONTEXTO DA AVALIAÇÃO MULTIDIMENSIONAL DA VOZ

No campo dos distúrbios vocais, não existe um padrão de referência perfeito ou teste definitivo e, como em muitos outros campos de estudo, diversos padrões de referência estão disponíveis para se definir uma avaliação vocal. Ao avaliar indivíduos com distúrbio de voz, é preferível, sempre que possível, caracterizar o impacto dos distúrbios em todos os mecanismos e dimensões pertinentes, seguindo uma bateria de avaliações que inclua: avaliação da percepção auditiva, exame visual laríngeo, avaliação acústica, avaliação aerodinâmica e a autopercepção dos pacientes sobre o impacto dos distúrbios vocais em sua vida diária (HIRANO, 1989; ROY et al., 2013).

No Quadro 1 é possível observar os itens de avaliação vocal e suas respectivas funções no contexto da avaliação multidimensional da voz:

**Quadro 1 - Instrumentos de análise vocal e suas funções na avaliação multidimensional da voz (continua)**

Item de Avaliação	Objetivo	Obtenção
Percepção Auditiva	Identificar a presença, tipo e intensidade do distúrbio de voz.	Julgamento do avaliador e impressão auditiva acerca do sinal vocal através de escalas de avaliação, as mais usadas são: <i>Consensus Auditory-Perceptual Evaluation of Voice</i> (CAPE-V), EAV e GRBAS.
Exame Visual Laríngeo	Analisar a estrutura e função da laringe através do julgamento e experiência do otorrinolaringologista.	Laringoscopia, Estroboscopia da laringe.
Análise Acústica	Fornecer uma análise quantitativa e qualitativa do sinal vocal oferecendo a possibilidade de detalhamento do processo de geração do sinal sonoro.	Medidas de perturbação e ruído; Medidas Formânticas; Medidas não lineares; Medidas Cepstrais.

**Quadro 2 - Instrumentos de análise vocal e suas funções na avaliação multidimensional da voz (conclusão)**

Item de Avaliação	Objetivo	Obtenção
Análise Aerodinâmica	Propiciar dados sobre o controle do fluxo aéreo para a fonação.	Fluxo de ar; Tempo Máximo de Fonação (TMF); Pressão Subglótica; Resistência Laríngea; Pressão Limiar de Fonação; Fluxo do Limiar de Fonação.
Autoavaliação	Avaliar os efeitos psicossociais das desvantagens e vantagens da condição vocal dos indivíduos sobre o impacto relativo de sua condição vocal nas atividades diárias.	Protocolo IDV (Índice de Desvantagem Vocal); Protocolo IDV-10; Protocolo QVV (Qualidade de Vida em Voz); Protocolo PPAV (Perfil de Participação e Atividades Vocais). Protocolo URICA-VOZ (Verificar o estágio de prontidão e a motivação do paciente disfônico à terapia vocal); Protocolo ESV (Escala de Sintomas Vocais).

FONTE: ASHA, 1998; ASHA, 2004; Patel et al., 2018.

A análise acústica serve como complemento da avaliação laríngea e percepção auditiva, o que resulta em dados quantitativos e qualitativos da função vocal e alta reprodutibilidade para monitoramento do paciente (AWAN et al., 2011). Com esses dados quantitativos é possível mensurar o distúrbio presente no sinal e compará-lo com dados normativos (BARSTIE B; DE BODT M, 2015). A validade da análise acústica depende da capacidade de representação do distúrbio de voz que é auditivamente percebido assim como, os mecanismos fisiológicos contidos na produção vocal. Diante disso, um desafio para clínicos e pesquisadores é o entendimento do quanto cada medida se relaciona com o exame visual laríngeo e a análise da percepção auditiva.

A avaliação acústica pode envolver a análise descritiva de padrões visuais, como o espectrograma de faixa larga e faixa estreita, o Diagrama de Desvio Fonatório (DDF) e o espectro de longo termo e a extração de medidas de energia, de perturbação e do ruído presente no sinal vocal (LOPES et al., 2020). O espectrograma é um gráfico tridimensional, com o domínio das frequências no eixo vertical, o domínio do tempo no plano horizontal e a amplitude dos componentes da onda sonora representada pelo contraste de cores no traçado (LOPES et al., 2017). O DDF, oferece a possibilidade de análise combinada das medidas de perturbação (*jitter*, *shimmer* e correlação) e ruído (GNE) (MADAZIO et al., 2011).

Segundo Nordemberg e Sundberg (2003) o espectro médio de longo termo (*Long Term Average Spectrum – LTAS*) reflete a contribuição tanto da fonte glótica quanto do trato vocal na qualidade de uma voz. Na análise acústica tradicional, se utiliza duas medidas, as de perturbação (*jitter e shimmer*) e de ruído. O *jitter* indica o quanto a frequência fundamental varia a curto prazo entre ciclos glóticos vizinhos. O *shimmer* diz respeito à variabilidade da amplitude da onda sonora a curto prazo. O *Glottal to Noise Excitation (GNE)* mensura o ruído adicional no sinal sonoro, de maneira independente do ruído modulado pelo mecanismo glótico, indicando a origem do sinal vocal, se ele é advém da vibração das pregas vocais ou da corrente de ar turbulenta gerada no trato vocal. (BROCKMANN-BAUSER et al., 2011).

Os métodos tradicionais dependem da identificação precisa dos limites do ciclo (ou seja, os limites nos quais um ciclo de vibração começa e termina), esses métodos se tornam inválidos quando aplicados a sinais que são altamente perturbados como resultado de vibração aperiódica das pregas vocais ou ruído aditivo. A validade questionável dos métodos de perturbação tradicionais no que diz respeito à quantificação da gravidade do distúrbio de voz é problemática ao tentar avaliar a mudança pré e pós-tratamento, uma vez que medidas válidas de gravidade do distúrbio de voz podem não ser adequadas para pacientes que inicialmente apresentam distúrbios vocais de moderados a graves (AWAN et al., 2009).

### 3.2 ANÁLISE CEPSTRAL

Por sua vez, a análise cepstral demonstrou ser uma possibilidade para análise de sinais com maior irregularidade, visto que as medidas cepstrais são capazes de determinar a  $f_0$  e produzir estimativas de aperiodicidade e/ou ruído aditivo sem a necessidade de identificar limites de ondas sonoras individuais. De forma geral, o cepstro evidencia em que medida que os harmônicos advindos da  $f_0$  são individualizados e se destacam em relação ao nível de ruído presente no sinal. Sinais com maior irregularidade e maior quantidade de ruído apresentam menor definição e amplitude do pico cepstral dominante. Sendo assim, as medidas cepstrais são mais confiáveis que as medidas tradicionais de perturbação e ruído para avaliação de vozes com ampla faixa de desvio, demonstrando também ser fortes preditoras da presença de desvio vocal (DEJONCKERE et al., 1996).

Um aspecto importante na avaliação dos distúrbios de voz é a escolha da tarefa de fala. A maioria das pesquisas relacionadas à análise dos distúrbios vocais usam como amostra as

vogais sustentadas. Um ponto negativo desse tipo de amostragem é que o ajuste feito durante a emissão da vogal sustentada é diferente do ajuste de fala encadeada. O que compromete a confiabilidade da análise de percepção auditiva, visto que, o mais comumente escutado é a fala encadeada (HILLENBRAND & HOUDE, 1996).

Será descrito a seguir alguns estudos que utilizaram as medidas *CPP* em conjunto com modelos preditivos para análise vocal.

Latoszek et al. (2018) desenvolveram um estudo exploratório onde foi analisado um conjunto de medidas para avaliação da rugosidade. A amostra foi constituída por 970 indivíduos com presença de distúrbio de voz e 88 sujeitos classificados com ausência de distúrbio de voz. A tarefa de fala utilizada foram amostras de fala contínua e vogal sustentada /a/ e essas amostras foram julgadas perceptivamente quanto à gravidade do parâmetro rugosidade. Foram extraídas 24 medidas acústicas e dentre essas medidas se utilizou SNL nas faixas de frequências entre 100-5.100 Hz, 100-8.000 Hz, 100-2.600 Hz, 100-3.000 Hz, 2.600-5.100 Hz e 5.100-8.000 Hz; H2A; CPPS; APQ5; GNE; etc. Em seguida foi aplicada uma análise de regressão linear múltipla por etapas para construir um modelo acústico dos melhores preditores, obtendo como resultado uma combinação de 12 medidas acústicas. Esse estudo apresentou como melhor resultado sensibilidade de 51,9% e especificidade de 94,9%.

Com o objetivo de realizar a discriminação entre dois grupos de indivíduos, (pessoas que sofrem da doença de Parkinson e sujeitos que sofrem de outros distúrbios neurológicos), Benda et. al (2016) coletaram amostras de vogal sustentada /a/ de 50 indivíduos. Foram utilizadas três diferentes técnicas para extração das medidas cepstrais (*MFCC - Mel frequency Cepstral Coefficients; PLP - Perceptual Linear Prediction* e *RASTA - PLP - ReAlitive SpecTrAl PLP*). Utilizou-se cinco classificadores diferentes de aprendizado não supervisionado (*SVM- Support Vector Machines; KNN - K-Nearest Neighbors; DA - Discriminant Analysis; NB - Native Bayes* e *CT - Classification Tree*) para construção de um modelo com melhor predição. O melhor resultado se deu com o *SVM* isolado com sensibilidade e especificidade de 90%.

Soren et al. (2013) objetivaram prever a gravidade do distúrbio de voz e diferenciar os tipos de qualidade da voz (soprosas e rugosas) através de medidas baseadas em análise cepstral com tarefa de fala encadeada. A amostra foi constituída por 28 indivíduos com distúrbio de voz e 14 com ausência de distúrbio de voz. As duas medidas utilizadas foram o *CPP* e o desvio padrão do *CPP*. Foi utilizada Regressão linear para determinar a capacidade preditiva de cada

variável para a gravidade do distúrbio de voz. Com o modelo de regressão as medidas utilizadas apresentaram acurácia de 79% na discriminação da qualidade vocal.

Em um estudo de Awan e Watts (2014), se utilizou o modelo de regressão linear com o objetivo de avaliar o valor preditivo do *CPP* e desvio padrão do *CPP* na diferenciação entre indivíduos com distúrbio de voz de indivíduos com ausência de distúrbio de voz usando vogal sustentada /a/ e fala contínua. A amostra totalizou 32 indivíduos, 16 com distúrbio de voz e 16 saudáveis. O melhor resultado foi uma sensibilidade de 93.75% e especificidade de 87.5%.

Devido ao número de pesquisas e experimentos para estudar a tensão percebida da voz serem menores quando comparadas as pesquisas que envolvem as qualidades vocais de rugosidade e soproidade, Anand et al. (2018) através de uma pesquisa exploratória, apontaram medidas acústicas com maior correlação com a qualidade vocal de tensão. A amostra foi constituída por 28 sujeitos com distúrbio de voz usando a vogal sustentada /a/ como tarefa de fala. Foram extraídas medidas acústicas relacionadas ao cepstro, momentos espectrais e *sharpness*. Além disso, uma série de análises de regressão linear foram calculadas para determinar qual dos índices melhor previu a percepção de tensão. Como resultado, *CPP* foi a medida acústica que teve o maior valor de correlação dentre as medidas analisadas, obtendo uma forte correlação com a tensão vocal, com acurácia maior que 70%.

Fant et al. (2018) utilizaram *Deep Neural Network (DNN)*, *support vector machine*, e *Gaussian mixture model* na detecção de distúrbio de voz através da medida *Mel frequency Cepstral coefficients*. O estudo obteve uma amostra de 462 indivíduos, sendo 60 saudáveis e 402 com desvio vocal. A tarefa de fala utilizada foi a vogal sustentada /a/. Os resultados demonstraram que o *DNN* supera o modelo de *support vector machine*, e *Gaussian mixture model* com melhor acurácia de 94,26%.

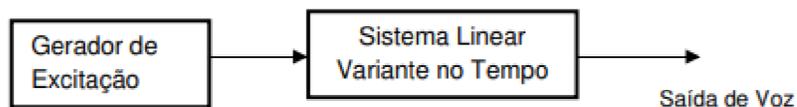
Diante de todo o exposto, é possível observar que a maioria dos estudos realizados com as medidas cepstrais e o uso de algum modelo de decisão é feito com a língua inglesa, sendo assim as medidas de corte para o *CPP* encontradas na literatura não podem ser usadas no português brasileiro.

Lopes et. al (2018) trabalharam com vozes sintetizadas e o objetivo do estudo era analisar quais seriam os parâmetros, usados na produção de vozes sintetizadas, seriam preditores da gravidade do distúrbio de voz. Foram utilizadas 871 vozes sintetizadas, com a vogal /E/ sustentada de ambos os gêneros e com desvios da qualidade vocal (rugosidade, soproidade e tensão). Fora utilizada a análise através da regressão linear beta com o objetivo de estimar um preditor que avaliasse os parâmetros de intensidade do distúrbio de voz. O

modelo conseguiu explicar 75% (pseudo  $R^2 = 0,75$ ) da variabilidade da gravidade geral do distúrbio de voz.

Mesquita (2006) teve como objetivo descrever a saúde vocal em professores da rede municipal de ensino de Belo Horizonte e investigar a associação dos distúrbio de voz com fatores relacionados ao trabalho de 1996 à 2004. Foi utilizado um total de 697 pessoas que receberam um questionário auto aplicado contendo questões sociodemográficas, de saúde geral e mental (GHQ-12), ambiente e organização do trabalho e aspectos relacionados à voz. A análise dos fatores associados à variável dependente (distúrbio de voz ausente, possível e provável) foi realizada em três etapas utilizando-se a regressão logística multinomial, tendo como categoria de referência a ausência de distúrbio de voz. O modelo final testou a associação independente de todas as variáveis que permaneceram estatisticamente associadas ao distúrbio de voz possível ( $p < 0,05$ ) em cada subconjunto de variáveis.

O uso da análise cepstral é direcionado para problemas centrados em voz sonora. A voz obtida a partir de um som sonoro, por exemplo, a produção da vogal sustentada /a/ permite avaliar o comportamento das pregas vocais, com vibração obtida na produção desse tipo de som. O modelo básico para produção da fala é constituído por um gerador de excitação e um sistema linear variante no tempo (**Figura 1**). O gerador de excitação deve fornecer dois tipos de saída: um trem de pulsos (glotais) para sinais sonoros e ruído aleatório para sinais não-sonoros (COSTA, 2008).



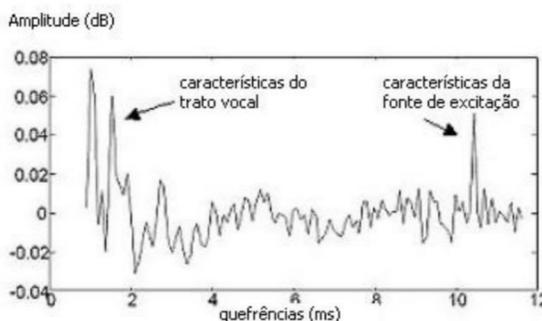
**Figura 1 – Modelo simplificado de produção de fala (COSTA, 2008).**

Visto que a saída de voz é o resultado da convolução do gerador de excitação (glote) com o sistema linear variante no tempo (trato vocal), seria útil separar ou “deconvoluir” as duas componentes. A “deconvolução” cepstral converte um produto de dois espectros na soma de dois sinais, que podem ser separados por um processo de filtragem linear, intitulado “lifteragem” ou *liftering*, facilitando o estudo individualizado das modificações ocorridas na excitação e da parte ressonantal. Dentre as propriedades matemáticas envolvidas no processo, destacam-se, principalmente, as transformadas de Fourier e funções logarítmicas que resultam

em uma função chamada cepstral ou cepstro, responsável pela dissociação do sinal de voz. A transformação desejada é logarítmica, dada pela Equação (1), sendo  $Ex(\omega)$  a transformada de Fourier da forma de onda da excitação e  $V(\omega)$  a transformada de Fourier do modelo do trato vocal. Assim, a análise cepstral do sinal de voz permite trabalhar com as componentes do sinal correspondentes à excitação e as respectivas modificações introduzidas pelo efeito de ressonância no trato vocal, separadamente. O estudo das alterações na voz produzidas pelas pregas vocais é, então, facilitada devido às suas propriedades homomórficas, que permitem a separação das características do filtro do trato vocal da sequência de excitação. (COSTA, 2008).

$$\log(Ex(\omega).V(\omega)) = \log(Ex(\omega)) + \log(V(\omega)) \quad (1)$$

Na **Figura 2**, vê-se o cepstro de um segmento de voz, em que o pico correspondente ao período fundamental (excitação) está próximo da quefrência de 10 ms, separado das componentes do trato vocal, que são as de baixas quefrências.



**Figura 2 – Cepstro de um segmento de fala (ZWETSCH, 2006).**

Na **Figura 3** é apresentada a representação do cepstro para um sinal de voz normal e, na **Figura 4**, o cepstro para um sinal de voz patológico com edemas nas pregas vocais (AGUIAR NETO et al, 2007).

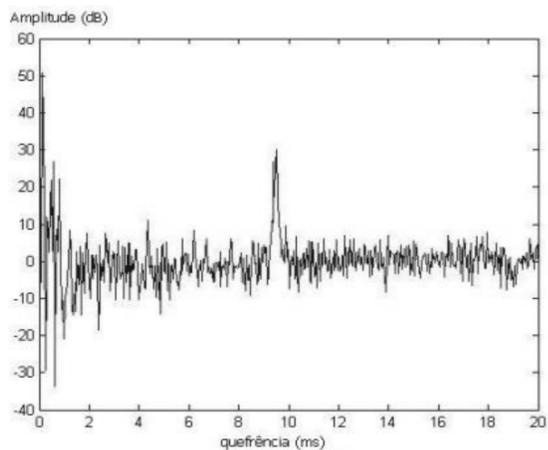


Figura 3 - Cepstro para uma voz normal (AGUIAR NETO et al, 2007).

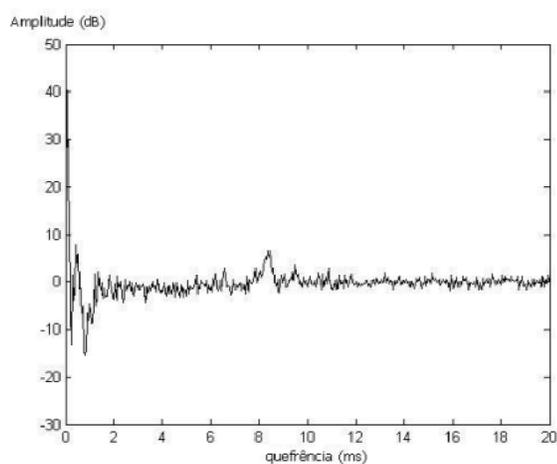


Figura 4 – Cepstro para uma voz patológica (AGUIAR NETO et al, 2007).

As medidas cepstrais vêm sendo cada vez mais usadas como forma de análise dos distúrbios vocais. Um crescente corpo de trabalhos (**Quadro 2**) demonstraram que as medidas cepstrais têm a capacidade de diferenciar perceptivamente indivíduos com e sem distúrbio de voz, além de vozes em diferentes idiomas e tarefas de fala.

**Quadro 2 - Estudos que utilizaram análise cepstral para identificação de desvios vocais (continua)**

<b>Autor (Ano)</b>	<b>Língua</b>	<b>Objetivo</b>	<b>Medida Cepstral utilizada</b>	<b>População</b>	<b>Tarefa de Fala</b>	<b>Resultados</b>
Hillenbrand et al. (1994)	Inglês	Avaliar a eficácia de várias medidas acústicas na previsão de classificações de soproidade.	<i>Spectral Tilt</i>	15 indivíduos vocalmente saudáveis	Vogal sustentada	O <i>spectral tilt</i> apresentou uma fraca correlação com a soproidade percebida
Hillenbrand et al. (1996)	Inglês	Avaliar a eficácia de várias medidas acústicas na previsão de classificações de soproidade.	<i>Spectral Tilt</i>	5 indivíduos vocalmente saudáveis e 20 com distúrbio de voz	Vogal sustentada e fala encadeada	Para a vogal sustentada O <i>Spectral Tilt</i> se correlacionou moderadamente com as classificações de soproidade. Já com a fala encadeada, forneceu previsões precisas de classificações de soproidade, respondendo por 70% -85% da variância.
Heman-Ackah et al. (2002)	Inglês	Avaliar a capacidade do CPP e outras medidas acústicas para prever o grau geral, grau de soproidade e grau de rugosidade em amostras de vozes com distúrbio de voz.	<i>CPPS</i>	19 indivíduos com distúrbio de voz	Vogal sustentada e fala encadeada	O <i>CPPS</i> apresentou uma forte correlação com o grau geral e grau de soproidade nas amostras de vogal sustentada. Também apresentou forte correlação com o grau geral na fala encadeada.
Heman-Ackah et al. (2003)	Inglês	Verificar se as medidas cepstrais são bons preditores de vozes com distúrbio de voz.	<i>CPPS</i>	281 indivíduos com distúrbio de voz	Vogal sustentada e fala encadeada	O <i>CPPS</i> se mostrou como sendo um bom preditor da disфонia com sensibilidade de 0.89 e 0.87 e especificidade de 0.77 e 0.90 na fala encadeada e vogal sustentada, respectivamente.

**Quadro 2 - Estudos que utilizaram análise cepstral para identificação de desvios vocais (continuação)**

Autor (Ano)	Língua	Objetivo	Medida Cepstral utilizada	População	Tarefa de Fala	Resultados
Halberstam (2004)	Inglês	Investigar as correlações entre os parâmetros acústicos e de percepção auditiva do grau de rugosidade.	<i>CPPS</i> e <i>CPP</i>	40 indivíduos com distúrbio de voz	Vogal sustentada e fala encadeada	O <i>CPPS</i> e <i>CPP</i> apresentaram uma correlação forte com a fala encadeada, com a vogal sustentada o <i>CPP</i> apresentou correlação forte e o <i>CPPS</i> uma correlação moderada.
Awan & Roy (2005)	Inglês	Avaliar as contribuições de várias medidas acústicas baseadas no tempo e espectrais para a categorização do tipo de voz para uma amostra diversa de vozes coletadas tanto com distúrbio de voz quanto normais.	<i>CPP</i>	51 mulheres vocalmente saudáveis e 83 com distúrbio de voz	Vogal sustentada	O <i>CPP</i> foi significante na discriminação de vozes normais e desviadas, porém não foi significante nos grupos específicos (soprosidade, rugosidade e tensão).
Eadier & Baylor (2006)	Inglês	Determinar se as mudanças na confiabilidade intra e interexaminador ocorrem para julgamentos de ouvintes inexperientes da gravidade geral, rugosidade e soprosidade em Falantes com distúrbio de voz e normais após 2 horas de treinamento de ouvinte e determinar as bases acústicas dos julgamentos de ouvintes inexperientes antes e depois do treinamento.	<i>CPP</i> e <i>CPPS</i>	30 indivíduos com distúrbio de voz e 6 vocalmente saudáveis	Vogal sustentada e fala encadeada	Medidas relacionadas ao <i>CPP</i> previram significativamente todos os julgamentos de qualidade da voz, exceto rugosidade nas vogais, que foi predita melhor pelo <i>shimmer</i> .

**Quadro 2 - Estudos que utilizaram análise cepstral para identificação de desvios vocais (continuação)**

<b>Autor (Ano)</b>	<b>Língua</b>	<b>Objetivo</b>	<b>Medida Cepstral utilizada</b>	<b>População</b>	<b>Tarefa de Fala</b>	<b>Resultados</b>
Awan et al. (2009)	Inglês	Identificar um subconjunto de métodos de análise com base nas medidas cepstrais que predissessem a gravidade do distúrbio de voz (conforme estimado por meio da análise da percepção auditiva) em amostras de fala contínua.	<i>CPP</i>	104 mulheres com distúrbio de voz	Fala encadeada	As medidas cepstrais se correlacionaram fortemente com a avaliação da percepção auditiva.
Radish Kumar et al. (2010)	Inglês	Observar se as medidas cepstrais possuem boa predição em indivíduos com nódulos vocais	<i>CPP</i>	50 indivíduos com distúrbio de voz	Vogal sustentada	Houve uma redução dos picos cepstrais em indivíduos com presença de nódulo vocal.
Awan et al. (2010)	Inglês	Avaliar a utilidade da análise acústica através das medidas cepstrais no monitoramento da alteração vocal pré e pós tireoidectomia.	<i>CPP</i>	70 indivíduos com distúrbio de voz	Vogal sustentada e fala encadeada	Houve uma moderada correlação com o grau geral de 0.68 na fala encadeada. Na tarefa de vogal sustentada houve uma correlação fraca com o grau geral.
Lowell et al. (2012)	Inglês	Determinar se as medidas acústicas cepstrais eram eficazes na distinção de indivíduos com distúrbio da qualidade vocal de indivíduos vocalmente saudáveis e se essas medidas estavam relacionadas às avaliações da percepção auditiva da gravidade da tensão.	<i>CPP</i>	23 indivíduos vocalmente saudáveis e 23 com distúrbio de voz	Vogal sustentada e fala encadeada	O <i>CPP</i> apresentou forte discriminação entre o grupo saudável e com distúrbio de voz e possui uma forte correlação com a avaliação da percepção auditiva (0.8).

**Quadro 2 - Estudos que utilizaram análise cepstral para identificação de desvios vocais (continuação)**

Autor (Ano)	Língua	Objetivo	Medida Cepstral utilizada	População	Tarefa de Fala	Resultados
Lowell et al. (2013)	Inglês	Analisar as medidas cepstrais quanto à previsão da gravidade do distúrbio de voz e na diferenciação dos tipos de qualidade vocal.	CPP e Desvio padrão do CPP (CPP SD)	28 indivíduos com distúrbio de voz e 14 vocalmente saudáveis	Fala contínua	Um modelo incorporou as medidas cepstrais e produziu um valor que explica 81% da variância na gravidade do distúrbio de voz quanto à percepção auditiva. O CPP apresentou a maior contribuição preditiva para a gravidade do distúrbio de voz. A análise discriminante produziu duas funções discriminantes que incluíram tanto CPP quanto seu desvio padrão como contribuintes significativos com uma precisão de classificação geral para as funções combinadas de 79%.
rinca et al. (2014)	Português-pt	Investigar o uso do CPP e CPPS na diferenciação de vozes com distúrbio de voz e sem de distúrbio de voz.	CPP e CPPS	30 indivíduos com distúrbio de voz e 30 vocalmente saudáveis	Vogal sustentada e fala encadeada	Para a vogal sustentada, ambas as medidas (CPP CPPS) foram significativas na diferenciação entre os grupos com distúrbio de voz e saudáveis. Para a fala encadeada, apenas os valores do CPP revelaram diferenças significativas entre os dois grupos. Houve correlação com a análise da percepção auditiva tanto na vogal sustentada quanto na fala encadeada, a correlação mais forte (0.66) foi entre o CPP e o grau de sopro.

**Quadro 2 - Estudos que utilizaram análise cepstral para identificação de desvios vocais (continuação)**

Autor (Ano)	Língua	Objetivo	Medida Cepstral utilizada	População	Tarefa de Fala	Resultados
Watts (2015)	Inglês	Investigar o efeito da fala encadeada (frases CAPE-V) sobre as medidas acústicas cepstrais.	CPP, CPP SD	20 indivíduos com distúrbio de voz	Fala encadeada	Os resultados revelaram que o CPP foi significativamente afetado pelas sentenças. Essa descoberta confirma a base teórica original para incluir estímulos de fala encadeada com diferentes contextos fonéticos.
Shim et al. (2016)	Coreano	Examinar as características acústicas cepstrais de pacientes com disfonia por tensão muscular.	CPP	30 indivíduos vocalmente saudáveis e 30 com distúrbio de voz	Vogal sustentada	Os coeficientes de correlação entre CPP e grau geral, rugosidade, sopro e tensão foram $-0,657$ , $-0,555$ , $-0,492$ , $-0,428$ , respectivamente; demonstrando uma alta correlação com a percepção auditiva.
Watts et al. (2017)	Inglês e Flamengo	Investigar a relação e a confiabilidade das medidas cepstrais de dois softwares acústicos, ADSV e Praat.	CPP	22 indivíduos vocalmente saudáveis e 22 com distúrbio de voz	Vogal sustentada e fala encadeada	As análises revelaram correlações muito fortes (0,88) entre as medidas CPP derivadas de ADSV e aquelas derivadas de Praat, independentemente do contexto (vogal e fala encadeada) ou idioma falado.
Delgado-Hernández (2018)	Espanhol	Estabelecer a utilidade do CPPS na avaliação objetiva de alterações da voz em espanhol e determinar que tipo de amostra de voz é a mais sensível na avaliação da gravidade do distúrbio de voz.	CPPS	20 indivíduos vocalmente saudáveis e 20 com distúrbio de voz	Vogal sustentada e fala encadeada	Houve significância para vogais e para a fala encadeada. Quanto à amostra de voz mais adequada para avaliação da gravidade das alterações vocais, houve forte correlação com a fala encadeada ( $-0,73$ ) e correlação moderada com a vogal ( $-0,56$ ).

**Quadro 2 - Estudos que utilizaram análise cepstral para identificação de desvios vocais (conclusão)**

<b>Autor (Ano)</b>	<b>Língua</b>	<b>Objetivo</b>	<b>Medida Cepstral utilizada</b>	<b>População</b>	<b>Tarefa de Fala</b>	<b>Resultados</b>
Lopes et al. (2019)	Português-br	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.	<i>CPPS, CPPS SD</i>	99 indivíduos vocalmente saudáveis e 277 com distúrbio de voz	Vogal sustentada	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 sopro e tensão.
Murton et al. (2020)	Inglês	Empregar tarefas e métodos de análise frequentemente usados para identificar valores de <i>CPP</i> que podem auxiliar na avaliação clínica da voz. O experimento 1 objetivou identificar valores de <i>CPP</i> para distinguir falantes com e sem distúrbio de voz. O experimento 2 estimou as avaliações da percepção auditiva da gravidade geral da disфонia usando valores de <i>CPP</i> .	<i>CPP e CPPS</i>	50 indivíduos sem alteração vocal e 295 com distúrbio de voz	Vogal sustentada e fala encadeada	Na diferenciação de pacientes com e sem desvio vocal o <i>CPP</i> apresentou acurácia de 79.4 e 87.7 para vogal e fala encadeada, respectivamente. O <i>CPPS</i> apresentou acurácia de 77.4 e 94.5 para vogal e fala encadeada, respectivamente. Os resultados sugerem fortes relações entre os valores de <i>CPP</i> e a percepção auditiva da gravidade geral do distúrbio de voz.

FONTE: Hillenbrand et al. (1994), Hillenbrand et al. (1996), Heman-Ackah et al. (2002), Heman-Ackah et al. (2003), Halberstam (2004), Awan & Roy (2005), Eadier & Baylor (2006), Awan et al. (2009), Radish Kumar et al. (2010), Awan et al. (2010), Lowell et al. (2012), Lowell et al. (2013), Brinca et al. (2014), Watts (2015), Shim et al. (2016), Watts et al. (2017), Delgado-Hernández (2018), Lopes et al. (2019), Murton et al. (2020).

### 3.3 ANÁLISE ACÚSTICA MULTIPARAMÉTRICA

As medidas acústicas oferecem informações mais objetivas acerca dos diversos parâmetros vocais, tais como: frequência fundamental, *jitter*, *shimmer*, proporção harmônico-ruído, porém são medidas acústicas isoladas. A voz por ser um fenômeno acústico, deve ser avaliada de maneira multidimensional e a extração de parâmetros isolados parece ser insuficiente para caracterizá-la. Diante disso, vem ampliando a procura por modelos acústicos multiparamétricos, como por exemplo o *Cepstral Spectral Index of Dysphonia (CSID)*, o *Acoustic Vocal Quality Index (AVQI)*, o *Acoustic Breathiness Index (ABI)* e o *Dysphonia Severity Index (DSI)*. Esses modelos consideram diversos parâmetros acústicos para fornecer um único escore para a qualidade vocal (BARSTIES & DE BODT, 2014).

*DSI* é um índice multivariado, que inclui quatro parâmetros que são ponderados e depois computados para quantificar a qualidade da voz. O *DSI* foi avaliado como uma medida útil e viável, com variabilidade aceitável entre examinadores. Esse índice tem sido utilizado clinicamente como uma medida para avaliar os efeitos da terapia vocal, treinamento vocal e em casos de cirurgias laríngeas (WUYTS et al., 2000).

O *AVQI* é um modelo multivariado, que inclui seis parâmetros acústicos e que abrange a fala contínua e a fonação sustentada para a análise. É recomendável usar as duas tarefas de fala, que adicionam representação de contextos variados de fala para avaliação da qualidade da voz. O tipo e a gravidade da qualidade da voz diferem substancialmente entre as duas tarefas de fala. A fala contínua representa o uso diário da voz, mas seu julgamento da qualidade da voz mostra maior variabilidade em relação aos fenômenos não vocais (por exemplo, flutuação prosódica, contexto fonético e uso fonológico da disfonia). As vogais sustentadas, por outro lado, são menos sensíveis a essa variabilidade fonética, mas também implicam menos fonação natural. Esse índice acústico se correlaciona significativamente com o julgamento da percepção auditiva da qualidade geral da voz (MARYN et al., 2010).

Em um estudo de Maryn et. al (2010), se teve como objetivo analisar o *Acoustic Voice Quality Index (AVQI)*. Esse índice combina seis medidas acústicas, a saber: *Smoothed cepstral peak prominence CPPS*; *Harmonic-to-noise ratio HNR*; *Shimmer local Shim*; *Shimmer local dB ShdB*; *General slope of the spectrum Slope* e *Tilt of the regression line through the spectrum Tilt*. O objetivo foi a avaliação do *AVQI* para análise da gravidade do distúrbio de voz de pessoas submetidas a procedimentos cirúrgicos laríngeos. As tarefas de fala escolhidas foram a vogal sustentada /a/ e fala encadeada (com texto balanceado foneticamente) de 33 pacientes

classificados com distúrbio de voz. Foi aplicada uma análise de regressão linear múltipla por etapas para construir um modelo acústico dos melhores preditores, obtendo como resultado uma combinação de 6 medidas acústicas. O *CPP* apresentou uma melhor contribuição para a medição objetiva da qualidade vocal quando as medidas foram combinadas, Esse estudo apresentou como melhor resultado sensibilidade e especificidade acima de 80%. Vale salientar que pelo tamanho da amostra não se pode representar uma população com presença de distúrbio de voz, entretanto os resultados reforçam a confiabilidade no uso do *CPP*.

Outro modelo utilizado na análise acústica da voz é o *Acoustic Breathiness Index (ABI)*, que também analisa amostras de vogal sustentada e fala contínua e, assim, avalia a gravidade da soproiedade, parâmetro caracterizado por ruído audível. A soproiedade é causada pelo fechamento incompleto das pregas vocais durante a fonação e é um sinal proeminente de doenças laringeas associadas, como nódulos nas pregas vocais, pólipos e paralisia das pregas vocais (BARSTIES et al., 2017).

No estudo de Latoszek et al. (2017) foi desenvolvido um modelo acústico multivariado para avaliação de vozes soprosas. A amostra foi constituída por 970 indivíduos com presença de distúrbio de voz e 88 sujeitos classificados sem distúrbio de voz. A tarefa de fala utilizada foram amostras de fala contínua e vogal sustentada /a/ e essas amostras foram julgadas perceptivamente quanto à gravidade do parâmetro soproiedade. Se extraiu 28 medidas acústicas, sendo elas: Hfno, HNR, H1-H2, CPPS, LNPSD, APQ5, GNE, entre outras. Posteriormente foi aplicada uma análise de regressão linear múltipla por etapas para construção de um índice acústico (*ABI, Acoustic Breathiness Index*) dos melhores preditores e como resultado uma combinação de 9 medidas acústicas. Com o índice foi possível realizar a classificação das vozes e o melhor resultado se apresentou com sensibilidade de 82,4% e especificidade de 92,9%.

Na literatura já é observado um índice cepstral para a língua inglesa intitulado *Cepstral Spectral Index of Dysphonia (CSID)*, também é uma ferramenta multivariada utilizada como método de análise acústica que tem por objetivo a estimativa da gravidade do distúrbio de voz produzida pelo programa *ADSV* e incorpora medidas espectrais e cepstrais extraídas de tarefas de fala contínua ou amostras de vogal sustentada. A gravidade estimada do distúrbio de voz pelo *CSID* é teoricamente um número entre 0 e 100, sendo 100 classificado como o mais grave. No entanto, em vezes, o *CSID* pode gerar um número abaixo de 0 ou acima de 100; essa estimativa representa uma voz extremamente normal e periódica ou uma voz desviada e aperiódica, respectivamente (PETERSON et al., 2013).

Em um estudo, Awan et al. (2015) tiveram como objetivo avaliar o desempenho do *Cepstral Spectral Index of Dysphonia (CSID* - uma estimativa multivariada da gravidade do distúrbio de voz) como uma potencial ferramenta de triagem para identificação de distúrbio de voz e identificar pontos de corte clínicos de classificação entre casos com distúrbios vocais e grupo controle. Os grupos foram definidos utilizando análise da percepção auditiva, auto avaliação através do *Voice Handicap Index* e o exame visual laríngeo. As amostras de voz foram analisadas através do programa *Analysis of Dysphonia in Speech and Voice program (ADSV)* utilizando-se as medidas cepstrais para análise da gravidade do distúrbio de voz. A capacidade do *CSID* de classificar com precisão as amostras quanto ao padrão de referência foi feita através de análises de regressão logística e por meio da curva ROC. Como conclusão foi visto que o *CSID* apresentou forte discriminação na classificação entre indivíduos com e sem distúrbio de voz.

No Quadro 3 é representado a função dos modelos multiparamétricos citados anteriormente e como eles auxiliam na avaliação multidimensional da voz.

**Quadro 3 - Modelos Multiparamétricos e suas funções na avaliação multidimensional da voz (continua)**

Modelo	Método	Medidas que compõem o modelo	Desempenho	Tarefas de fala	Língua de validação
<i>CSID</i>	O <i>software</i> gera dois escores para a qualidade vocal que varia de 0 a 100 pontos. Esse índice é comercializado com o programa <i>Analysis of Dysphonia in Speech and Voice</i> da KayPENTAX	<i>CPP</i> - <i>Cepstral Peak Prominence</i> , <i>SR</i> - <i>Spectral Ratio</i> , e <i>the standard deviation</i> do <i>Spectral Ratio</i> .	Forte correlação com a percepção auditiva, que varia entre 0,75 e 0,85.	Avalia separadamente a vogal sustentada e a fala encadeada.	Desenvolvido para o inglês.

**Quadro 3 - Modelos Multiparamétricos e suas funções na avaliação multidimensional da voz (conclusão)**

Modelo	Método	Medidas que compõem o modelo	Desempenho	Tarefas de fala	Língua de validação
<i>AVQI</i>	Utiliza um script no programa Praat – um <i>software</i> livre que pode ser baixado pela internet para quantificar a intensidade de desvio da qualidade vogal. Fornece um único escore de 0 a 10 pontos.	<i>Smoothed cepstral peak prominence</i> CPPS; <i>Harmonic-to-noise ratio</i> HNR; <i>Shimmer local</i> Shim; <i>Shimmer local</i> dB ShdB; <i>General slope of the spectrum</i> <i>Slope</i> e <i>Tilt of the regression line through the spectrum</i> <i>Tilt</i> .	Forte correlação com a análise da percepção auditiva, que varia entre 0,794 e 0,929, e precisão diagnóstica consistente.	Avalia concomitantemente a vogal sustentada e fala encadeada.	Originalmente desenvolvido em holandês, já foi validado para outros idiomas (alemão, inglês, francês, finlandês, coreano e lituano).
<i>ABI</i>	O modelo foi desenvolvido para a quantificação da soproidade e utiliza um script no programa Praat – um <i>software</i> livre que pode ser baixado pela internet para quantificar a intensidade de desvio da qualidade vocal. Fornece um único escore de 0 a 10 pontos.	Hfno-6000 Hz, H1-H2, CPPS, HNR, PSD, <i>Jitter</i> , <i>Shimmer</i> , Shim-dB, GNEmax-4500 Hz.	Forte correlação com a análise da percepção auditiva com acurácia de 0.84.	Avalia concomitantemente vogal sustentada e fala encadeada.	Desenvolvido para o holandês e já foi validado para o alemão, coreano, japonês e espanhol.
<i>DSI</i>	Analisa a voz por meio de um conjunto de parâmetros acústicos e compara-os com uma escala da percepção auditiva. Os resultados variam de +5 (vozes normais) a -5 (distúrbios intensos).	<i>Jitter</i> , Frequência Fundamental mais alta ( $f_0$ high), menor intensidade (1-low em dB SPL) e Tempo Máximo de Fonação.	Forte correlação com a análise da percepção auditiva com acurácia de 0.79.	Vogal sustentada.	Originalmente desenvolvida com um banco de dados da Bélgica.

FONTE: Englert et al. 2019; Barsties et al. 2017; Awan et al. 2015.

## 4 METODOLOGIA

### 4.1 DESENHO DO ESTUDO

Trata-se de um estudo quantitativo, visto que os resultados desta pesquisa podem ser quantificados; Documental e retrospectivo pois foi utilizado um banco de vozes já disponível no Laboratório Integrado de Estudos da Voz (LIEV). A pesquisa é do tipo descritiva, pois foi feita uma observação, registro, análise, e em seguida correlacionados fatos ou variáveis (MANZATO; SANTOS, 2012). A pesquisa também é do tipo transversal, pois os dados são coletados em um ponto do tempo, utilizando como base uma amostra, selecionada com finalidade de descrever uma população em um determinado momento (RICHARDSON, 1999). Quanto à validação das hipóteses, foi observacional, pois não houve influência do pesquisador sobre as variáveis.

### 4.2 POPULAÇÃO DO ESTUDO

A determinação do tamanho da amostra nesta pesquisa baseou-se em uma média dos indivíduos com e sem distúrbio de três outros estudos (AWAN et al., 2015; AWAN et al., 2014 e MARYN et al., 2014) realizados com o objetivo de validação dos valores de corte das medidas cepstrais (Quadro 4). Estes estudos foram escolhidos por terem relevância na área da análise acústica principalmente no que diz respeito às medidas cepstrais, além disso todos esses estudos usaram como desfecho o laudo do otorrinolaringologista quanto à classificação de indivíduos normais e com patologia laríngea. Foi encontrada a média de 259 participantes nos grupos com distúrbios de voz e 59 no grupo sem distúrbio de voz. No presente estudo 78 indivíduos foram classificados como sem distúrbio de voz e 227 com presença de distúrbio de voz.

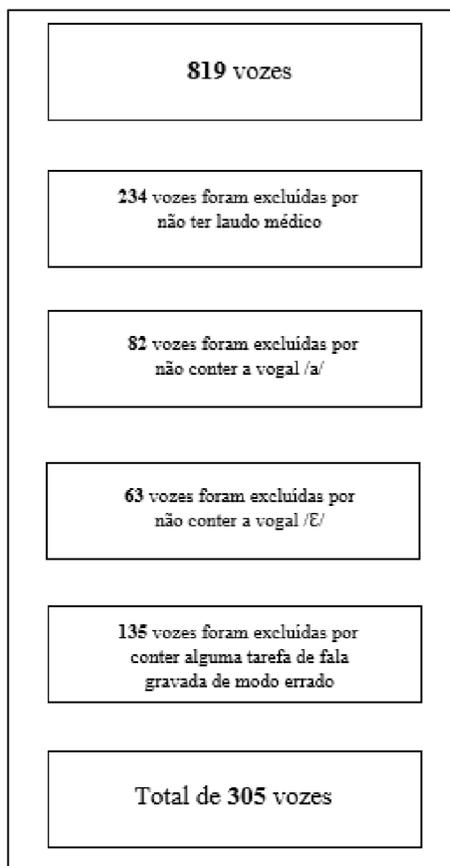
Historicamente, as mulheres apresentam uma maior preocupação com a saúde em relação aos homens, sendo elas as responsáveis pela maior parte dos registros de consultas clínicas (PINHEIRO et al., 2002). Em relação aos problemas relacionados à voz, as mulheres são mais afetadas (BEHLAU et al. 2001), fato que pode indicar a prevalência do gênero feminino na amostra do estudo.

**Quadro 4.** Estudos utilizados como referência para o tamanho amostral da dissertação.

<b>Autores/Ano</b>	<b>Objetivo</b>	<b>População</b>	<b>Padrão de Referência</b>
Maryn e Weenink (2015)	Estimar a relação entre o CPPS no programa <i>SpeechTool</i> e no <i>Praat</i> . Além disso, investigar a proporção entre o AVQI em versão original e a segunda versão com o CPPS.	261 indivíduos com distúrbio de voz e 28 sem distúrbio de voz sendo 193 mulheres e 96 homens.	Presença e ausência de distúrbio da voz.
Awan et al., (2015)	Avaliar o desempenho do CSID como uma ferramenta de triagem potencial para identificação de distúrbio de voz e identificar pontuações de corte clínicas potenciais para classificar casos com distúrbio de voz versus controles.	258 indivíduos com distúrbio de voz e 74 sem distúrbio de voz sendo 216 mulheres e 116 homens.	Presença e ausência de distúrbio da voz.
Awan et al., (2014)	Avaliar a relação entre as medidas cepstrais e a medida do VHI em um grupo grande e diversificado de indivíduos com distúrbios vocais e controles. Além disso, examinar a relação entre o VHI total e o CSID.	258 indivíduos com distúrbio de voz e 74 sem distúrbio de voz sendo 216 mulheres e 116 homens.	Presença e ausência de distúrbio da voz.

Fonte: Awan et al., 2015; Awan et al., 2014 e Maryn et al., 2014.

Dessa forma, a amostra foi composta por 305 indivíduos adultos, com média de idade de 36,00 (DP = 12,13) sendo 65 homens e 240 mulheres. Os participantes que fizeram parte do estudo selecionados a partir da base de dados do Laboratório Integrado de Estudos da Voz (LIEV) do Departamento de Fonoaudiologia da Universidade Federal da Paraíba – UFPB. Na figura 5 é possível observar o fluxograma com o processo de escolha das vozes do banco de dados:



**Figura 5: Fluxograma de vozes no banco de dados da atual pesquisa**

A partir da consulta à base de dados do referido laboratório, os participantes foram alocados em dois grupos: grupo de trabalho (GT - com distúrbio de voz) e grupo- controle (GC - sem distúrbio de voz). Ambos os grupos deveriam cumprir os seguintes critérios de elegibilidade:

- Possuir idade entre 18 e 60 anos, pois são limites que diminuem a probabilidade da presença de alterações vocais decorrentes de mudanças hormonais da adolescência ou senescência (BEHLAU, 2001);
- Não ter realizado terapia vocal ou tratamento cirúrgico na laringe previamente.

Os participantes do GT deveriam estar de acordo com 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, com laudo de lesão estrutural ou funcional na laringe.

Os indivíduos do GC deveriam apresentar os seguintes critérios de elegibilidade:

- Não apresentar queixa vocal, respondendo negativamente à pergunta: “Você considera que tem um problema de voz atualmente?”;
- Terem realizado o exame visual laríngeo, com laudo de laringe normal;

Os critérios de exclusão foram:

- Apresentar comprometimento cognitivo ou neurológico que impedisse a gravação da voz.

A variável dependente do estudo é a presença ou ausência de distúrbio de voz e dentre as variáveis independentes se têm as medidas cepstrais de CPPS, CPP e declínio espectral extraídas das tarefas de fala de vogal sustentada /a/ e /ɛ/, contagem de 1 a 10, frase do CAPE-V e de todas essas tarefas de fala juntas em um mesmo arquivo de áudio.

#### 4.3 AMOSTRA DE FALA

Esse estudo foi documental e retrospectivo. Dessa forma, passaremos a descrever as condições nas quais as tarefas de fala foram coletadas e armazenadas na base de dados utilizada para esta pesquisa.

Os procedimentos de coleta de dados para as amostras de voz foram conduzidos de acordo com a rotina do protocolo de avaliação vocal do laboratório (LIEV). Todos os dados foram registrados digitalmente (sinal de voz) ou de forma escrita nos prontuários clínicos nas bases de dados do laboratório (anamnese ou laudo do exame visual laríngeo). Utilizou-se o *software* Fonoview, versão 4.5, microfone cardioide unidirecional, localizado em um pedestal e acoplado a um pré-amplificador Behringer. 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 a 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 /ɛ/ e /a/ sustentada, em frequência e intensidade autorreferida como habitual. Foram selecionadas por serem as vogais mais comumente utilizadas para avaliação da qualidade vocal na realidade brasileira (GONÇALVES et al., 2009). Na sequência, submeteram-se à gravação de fala automática, correspondendo à contagem de 1 a 10.

Em seguida, os informantes procederam com a gravação de uma frase do Protocolo CAPE-V traduzida e adaptada para o português brasileiro (BEHLAU et al., 2020), a saber: “Olha lá o avião azul”. Optou-se por essa frase por ela ter um segmento totalmente vozeado, pois é possível que o ruído das consoantes não vozeadas seja adicionado de forma indevida às análises e, assim, influenciar na estimativa de gravidade computada, sendo confundido com ruído proveniente do distúrbio de voz (AWAN et al., 2011). As 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 (frase do CAPE-V).

Após o processo de coleta, as vozes foram editadas no *software SoundForge* versão 10.0, sendo eliminados os dois segundos iniciais e finais da emissão da vogal sustentada, contagem e frase do CAPE-V devido a maior irregularidade nesses trechos, preservando-se o tempo de três segundos para cada emissão, no caso das vogais.

#### 4.4 MÉTODO DE ANÁLISE DOS DADOS

##### 4.4.1 Extração das medidas acústicas

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* (CPP), *Cepstral Peak Prominence-Smoothed* (CPPS) e o declínio espectral das amostras vocais. O CPPS é uma modificação no algoritmo do CPP (*Cepstral Peak Proeminence*) 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. Os comandos e parâmetros serão aplicados através de uma extração automática, utilizando um *script* (ANEXO A) para gerar o CPP, CPPS e declínio espectral no aplicativo Praat.

#### 4.4.2 Análise das gravações vocais

A avaliação da percepção auditiva foi realizada por um fonoaudiólogo especialista em voz, com mais de 10 anos de experiência nesse tipo de análise. O juiz foi treinado com o estímulos-âncora (vogal sustentada /É/) de vozes saudáveis e com distúrbios vocais leves, moderados e intensos. O juiz foi instruído de que as vozes deveriam ser consideradas normais quando fossem socialmente aceitáveis, produzidas de forma natural, sem irregularidade, ruído ou esforço observável durante a emissão. Para avaliação, cada tarefa de fala foi apresentada por três vezes através de caixa de som, em intensidade confortável autorreferida pelo avaliador.

Para esta avaliação de percepção auditiva foi utilizada a Escala Analógica Visual (EAV) (ANEXO B) que consiste em uma linha horizontal de 0 a 100 mm, ao longo dessa linha, o juiz realizará uma marcação, representando a quantidade de uma característica julgada, e cada milímetro representará um grau de desvio. A marcação mais próxima do 0 representa menor desvio, e quanto mais próxima do 100, maior o desvio vocal. Esse instrumento possibilita a avaliação de diversos parâmetros vocais, porém, para a presente pesquisa, foi utilizado apenas o grau geral (GG), que consiste na impressão global da voz (KEMPSTER et al., 2009).

No final da sessão de avaliação perceptiva, 20% das amostras foram repetidas aleatoriamente, para a análise da confiabilidade da avaliação do juiz através do Coeficiente Kappa de Cohen. (SIM J, WRIGHT CC, 2005). Para a definição dos resultados do teste, será utilizado, neste trabalho, a classificação proposta por Landis et al., (1977), que serão descritos a seguir na tabela 1:

**Tabela 1** – Classificação do teste coeficiente de Kappa, segundo Landis et al., (1977)

Valores	Classificação
Entre 0,80 e 1,00	Quase perfeita
Entre 0,60 e 0,80	Boa
Entre 0,40 e 0,60	Moderada
Entre 0,20 e 0,40	Regular
Entre 0 e 0,20	Discreta
Entre -1 e 0	Pobre

O coeficiente de Kappa pode ser calculado da seguinte forma:

$$K = \frac{P_0 - P_c}{1 - P_c} \quad (2)$$

onde,  $P_0 = \sum_{i=1}^M (n_{ii}) / N$ ,  $P_c = \sum_{i=1}^M (n_{i+} + n_{+i}) / N^2$ , onde  $n_{ii}$  é o total da diagonal principal da Matriz de Confusão;  $n_{i+}$  é o total da linha  $i$  na Matriz de Confusão;  $n_{+i}$  é o total da coluna da mesma matriz;  $M$  é o total de classes de desempenho na matriz e  $N$  é o número total de decisões presentes na matriz.

A variância do Kappa ( $\sigma_{kappa}^2$ ) pode ser estimada por (MORAES; MACHADO, 2014):

$$\sigma_{kappa}^2 = \frac{P_0(1-P_0)}{N(1-P_c)^2} + \frac{2(1-P_0)+2P_0P_c-\theta}{N(1-P_c)^3} \quad (3)$$

onde:

$$\theta = \frac{\sum_{i=1}^M n_{ii}(n_{i+} + n_{+i})}{N^2} \quad (4)$$

O valor resultante do coeficiente Kappa varia entre -1 e +1.

Os valores do julgamento perceptivo-auditivo foram utilizados para caracterização da qualidade vocal do GT e GC, assim como para a análise de regressão

#### 4.4.3 Análise estatística dos dados

Os dados foram tabulados em planilha eletrônica no Microsoft Office Excel. A análise estatística foi realizada por meio do *software* estatístico livre R, versão 3.6.1.

Foi realizada a estatística descritiva para a obtenção de informações sintéticas sobre os dados. A distribuição normal, tanto para o caso univariado quanto para o multivariado, possui um papel muito importante na inferência estatística. A não verificação dessa pressuposição influencia diretamente na qualidade e confiabilidade da inferência, podendo conduzir em resultados e conclusões incorretos.

Destaca-se, assim que foi realizado o teste de Shapiro-Wilk para testar normalidade das variáveis independentes, o p-valor do teste para o modelo foi menor de (p-valor= <0,05) o que significa não normalidade dos dados, também foi realizada a análise de Regressão Logística Binomial com a variável dependente: distúrbio de voz, se utilizou a regressão linear múltipla onde o objetivo foi utilizar a variável: grau geral do desvio vocal segundo a EAV e identificar quais medidas cepstrais em diferentes tarefas de fala conseguiram prever os diferentes graus de

desvio vocal. A Regressão Logística analisa a relação estatística entre variáveis independentes e uma variável dependente assim podendo determinar se as medidas cepstrais e quais tarefas de fala mais influenciam na identificação de indivíduos com e sem desvio vocal (FÁVERO et al, 2009). A regressão linear múltipla ocorre quando se admite que o valor da variável dependente é função linear de duas ou mais variáveis independentes (FÁVERO et al, 2009).

#### 4.4.3.1 Análise estatística descritiva

A análise estatística descritiva foi realizada para obtenção de informações sintéticas sobre os dados. Nela foram obtidos os valores de média e desvio padrão de todas as variáveis. A amostra foi constituída por 305 indivíduos de ambos os sexos com média de idade de 36 anos (DP = 12,13).

Foi realizada uma análise descritiva (**Tabela 2**) entre o GT e GC com as medidas: CPPS, CPP e Declínio espectral com todas as tarefas de fala, afim de se obter dados como a média, desvio padrão, mínimo e máximo das variáveis. A análise descritiva mostrou que, em média, o grupo sem distúrbio apresentam valores maiores das medidas cepstrais.

**Tabela 2.** Tabela descritiva para os grupos sem distúrbio e com distúrbio de voz

VARIÁVEIS	SEM DISTÚRBO DE VOZ				COM DISTÚRBO DE VOZ			
	MÉDIA	DP	MÍN	MÁX	MÉDIA	DP	MÍN	MÁX
CPP VGE	28,20	3,06	21,77	3,06	24,57	5,38	10,52	35,81
CPP VGA	27,44	3,52	16,13	3,52	24,28	5,75	9,56	38,15
CPP CAPE-V	23,01	5,72	7,60	5,72	19,66	6,12	9,51	33,67
CPP TODAS	27,37	3,52	16,33	3,52	24,16	5,68	9,01	37,43
CPP	12,84	3,04	6,67	3,04	13,12	3,55	7,04	30,91
CONTAGEM								
CPPS VGE	16,53	2,30	10,95	2,30	13,89	3,78	2,90	23,52
CPPS VGA	16,40	2,39	11,98	2,39	13,78	4,01	2,95	23,93
CPPS CAPE-V	13,14	1,50	9,82	1,50	11,78	2,75	3,06	17,39
CPPS TODAS	10,92	1,28	8,32	1,28	9,75	2,15	3,00	14,91
CPPS	7,60	1,50	4,52	1,50	7,07	1,64	2,93	12,07
CONTAGEM								
DECL. VGE	-13,96	5,69	-29,78	5,69	-16,63	6,65	-35,47	-2,12
DECL. VGA	-12,42	4,67	-28,23	4,67	-13,79	6,15	-34,54	5,49
DECL. CAPE-V	-13,38	6,01	-27,94	6,01	-15,89	5,14	-31,94	14,02
V								
DECL. TODAS	-12,97	4,20	-26,45	4,20	-15,29	5,25	-29,97	-2,79
DECL.	-14,64	6,35	-28,57	6,35	-17,59	6,54	-32,50	15,31
CONTAGEM								

**Legenda:** CPP: Cepstral Peak Prominence; CPPS: Cepstral Peak Prominence Smoothed; DECL: Declínio espectral; VGE: Vogal /E/ sustentada; VGA: Vogal /a/ sustentada, DP: Desvio Padrão. CPP TODAS: Tarefas de fala em um único arquivo de áudio.

#### 4.4.3.2 Regressão Logística

A tomada de decisão científica é diferenciada de acordo com as particularidades do tema proposto, as informações acessíveis, as possíveis de serem utilizadas, a qualidade dessas informações e do viés profissional do indivíduo tomador de decisão. Sendo assim, este processo é representado em alto nível, objetivando a diminuição da perda de informações na sua representação e por consequência, utilizar ao máximo a informação disponibilizada. A maneira de modelar esse processo irá variar conforme as fontes, suas possíveis representações estatísticas, matemáticas e do contexto a qual está inserido. Daí surgem as formas interdisciplinares de modelagem, suas representações e suas diversas utilizações em cenários complexos de tomada de decisão (MORAES et al., 2016).

Se destaca na presente dissertação o modelo de regressão logística. O modelo de regressão logística pode ser definido como uma técnica estatística que, a partir de uma série de variáveis explicativas ou independentes, produz um modelo que permite a predição de valores assumidos por uma variável dependente dicotômica quantitativa ou categórica, e torna ainda possível aferir a probabilidade de ocorrência direta desse evento ocorrer (FÁVERO et al., 2009).

O modelo de regressão logística pode ser expresso da seguinte forma, sabendo que  $\pi$  (X), consiste na probabilidade de sucesso dado o valor x de uma variável explicativa qualquer

$$\text{logit} \left( \frac{\pi_i}{1-\pi_i} \right) = \beta_0 + \beta_1 X_1 + \dots + \beta_q X_q + \varepsilon, \quad i = 1, \dots, n, \quad (5)$$

em que,

$\beta_0, \beta_1, \dots, \beta_q$  são parâmetros do modelo;

$x_1, x_2, \dots, x_q$ : variáveis explicativas (independentes) explicativas;

$\varepsilon$  é o erro aleatório;

$\text{logit} \left( \frac{\pi_i}{1-\pi_i} \right)$  – logaritmo de razão de chance. Esta função normalmente é indicada pelo fato de sua melhor aplicabilidade a variáveis dicotômicas e aos parâmetros de regressão que podem ser interpretados em termo de razão de chances (*odds ratio*) (HOSMER E LEMESHOW, 2000).

Devido a sua utilização para variáveis do tipo binária, se faz necessário uma transformação para tornar a resposta da variável a ser estimada pelo modelo em uma função

não linear, para isto utiliza-se a função exponencial, resultando na função de resposta logística, uma vez que  $\pi_i$  trata-se de uma probabilidade (BARRETO, 2011). Assim:

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_q X_q)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_q X_q)}, \quad (6)$$

Na presente pesquisa, será considerado o fato de ocorrência ou não do distúrbio de voz, desta forma, a probabilidade de presença ou não deste distúrbio de acordo com as variáveis explicativas, pode ser expressa da seguinte forma:

$$\text{Pr}(Y|X_q) = \beta_0 + \beta_1 X_1 + \dots + \beta_q X_q + \varepsilon, \quad (7)$$

em que, se  $Y = \begin{cases} 1, & \text{ocorre distúrbio} \\ 0, & \text{não ocorre distúrbio} \end{cases}$

Para o ajuste do modelo de regressão logística a um conjunto de dados, ou seja, a estimação dos parâmetros  $\beta_0$  e  $\beta_i$ , em geral, se utiliza o Método de Máxima Verossimilhança, o qual busca maximizar a probabilidade de que o evento ocorra (FÁVERO, 2009).

A estimação dos parâmetros  $\beta_0$  e  $\beta_i$  tem como objetivo encontrar um modelo logístico de forma que as ponderações sofridas pelas variáveis explicativas estabeleçam a importância de cada variável para ocorrência do evento de interesse, bem como possibilitem o cálculo da probabilidade de ocorrência desse evento (FÁVERO, 2009).

Neste estudo a regressão logística foi utilizada como ferramenta de modelagem estatística, no intuito de estimar a capacidade preditiva das medidas cepstrais em relação à presença/ausência do distúrbio de voz.

Para o ajuste deste modelo, foi realizada inicialmente, uma triagem para a pré-seleção de variáveis mais significativas. Todas elas foram submetidas individualmente ao teste de associação estatística Qui-quadrado e o p-valor menor ou igual a 0,05 ( $\alpha=0,05$ ; nível de significância estatística = 5%) foi utilizado como referência para a escolha dos itens, de modo que a permitir que todas as variáveis conceitualmente relevantes para a ocorrência de distúrbio de voz permanecessem para a análise posterior do modelo de regressão.

Após esta etapa, o ajuste do modelo de regressão foi realizado, sucedendo-se com a investigação do melhor ajuste, aquele mais reduzido, que inclui apenas as covariáveis e

interações mais importantes para explicar a probabilidade da ocorrência do distúrbio de voz. Tal procedimento ocorreu por meio da seleção das variáveis explicativas ( $\alpha = 0,10$ ), passo a passo. Aquelas que não contribuíram ou contribuíram pouco na variabilidade da variável dependente ( $p\text{-valor} \geq 0,10$ ) foram excluídas do modelo uma a uma.

Vale ressaltar que o p-valor adotado nesse momento ( $\alpha=0,5$ ) foi selecionado de acordo com a necessidade da análise estatística desenvolvida. O método de inclusão e eliminação de covariáveis segundo a importância das mesmas, de acordo com algum critério estatístico utilizado para o modelo de regressão logística e linear, denomina-se Método *Stepwise* (HOSMER e LEMESHOW, 2000), e trata-se de uma das ferramentas mais aplicadas em regressão logística para a seleção de modelos. As variáveis de interesse nem sempre apresentam a significância estatística esperada, quando isso acontece se deve impor um limite mais elevado para estes níveis discriminativos, com o intuito de atrair variáveis importantes, em princípio, para participarem do modelo. Nesta pesquisa o p-valor adotado como critério para permanência das covariáveis no modelo foi de 0,5, o método (PAULA, 2010).

Com a obtenção do modelo final, foi necessário verificar sua qualidade, testar sua validade e interpretar seus resultados. O método utilizado para selecionar o modelo final, verificando qual deles apresenta o melhor ajuste, foi o Critério de Akaike (AIC), que corresponde a um processo de minimização do conjunto de variáveis explicativas, não envolvendo testes estatísticos. Segundo Akaike (1974) esse critério auxilia na escolha do modelo mais parcimonioso, com um número reduzido de parâmetros. De acordo com Paula (2010), como o logaritmo da função de verossimilhança  $L(\_)$  cresce com o aumento do número de parâmetros do modelo, a melhor proposta é encontrar o modelo com menor valor para a função:  $AIC = -L(\_) + p$ , em que  $p$  denota o número de parâmetros do modelo ajustado. O modelo que apresentou menor valor de AIC foi considerado como de melhor ajuste para os parâmetros selecionados.

#### 4.4.3.3 Regressão Linear Múltipla

Neste estudo também foi utilizada a regressão linear múltipla, que pode ser utilizada para analisar a relação entre uma única variável dependente e várias variáveis independentes. O objetivo da análise de regressão múltipla é usar as variáveis independentes, cujos valores são conhecidos, para prever os valores da variável dependente selecionada pelo pesquisador. Os coeficientes da regressão denotam a contribuição relativa das variáveis independentes para a

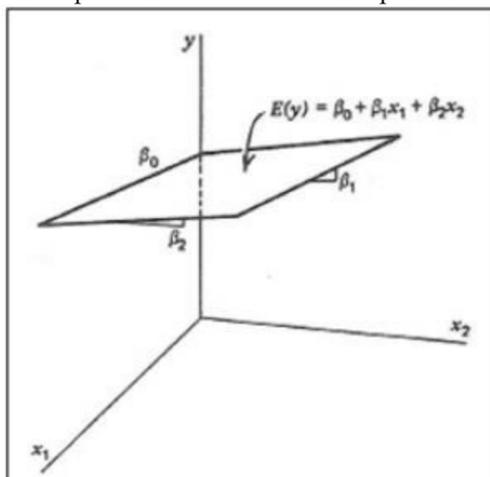
previsão geral da variável dependente e assim facilitam a interpretação sobre a influência de cada variável no modelo (HAIR JR. Et al, 2009).

O conjunto das variáveis independentes ponderadas formam o modelo estatístico de regressão, uma combinação linear das variáveis independentes que melhor prevê a variável dependente (SPIEGEL et al, 2013). No caso em que temos duas variáveis independentes sua forma geral é expressa pela relação:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \quad (8)$$

O termo “linear” é usado porque: (8) é uma função linear dos parâmetros desconhecidos. O modelo descreve um plano no espaço bidimensional das variáveis independentes  $x_1$  e  $x_2$ , tal como mostra na figura 6.

**Figura 6.** Exemplo de um modelo de regressão múltipla com duas variáveis independentes



Fonte: (MONTGOMERY e PECK, 1982)

O parâmetro  $\beta_0$  também chamado de intercepto, é a média de  $y$  quando  $x_1 = x_2 = 0$ . O parâmetro  $\beta_1$  indica a mudança esperada na resposta ( $y$ ) por unidade de mudança em  $x_1$  quando  $x_2$  é mantida constante. Da mesma forma, a medida  $\beta_2$  detém a variação esperada do  $y$  por unidade de mudança em  $x_2$  quando  $x_1$  é mantida constante. Em geral a resposta  $y$  pode estar relacionada com  $k$  variáveis independentes. Onde o modelo pode ser expresso por:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i \quad (9)$$

Em que,

$y_i$  é a variável dependente;

$\beta_0$  é o intercepto ou termo independente da variável;

$\beta_1, \beta_2, \beta_3, \dots, \beta_k$  são as inclinações de  $y$  em relação as variáveis  $x_1, x_2, x_3, \dots, x_k$  respectivamente, mantendo constante as demais covariáveis.

$\mathcal{E}_i$  é o erro aleatório em  $y_i$ , para a observação  $i, i=1, 2, \dots, n$ .

Dessa forma temos um modelo de regressão linear múltipla (9) com  $k$  variáveis independentes.

No que se refere ao modelo de ajuste, o  $R^2$  ajustado exibe o coeficiente de determinação múltipla, que é uma medida do grau de ajustamento da equação de regressão múltipla aos dados amostrais. Um ajuste perfeito resulta em  $R^2 = 1$ , um ajuste muito bom acarreta um valor próximo de 1 e um ajuste fraco ocasiona um valor de  $R^2$  próximo de zero. O coeficiente múltiplo de determinação de  $R^2$  é uma medida de aderência da equação de regressão aos dados amostrais. O coeficiente de determinação ajustado é o coeficiente múltiplo de determinação  $R^2$  modificado de modo a levar em conta o número de variáveis e o tamanho da amostra. O coeficiente de determinação ou de explicação  $R^2$ , mede a parcela da variação de  $Y$  explicada pela variação dos  $X$  (FÁVERO et al., 2009).

#### 4.5 RAZÃO DE CHANCES (*ODDS RATIO*)

A razão de chances ou *odds ratio* (OR) é uma medida da intensidade do efeito de  $x_i$  (por exemplo, mudança de  $x_i = 0$  para  $x_i = 1$ ) sobre a probabilidade de obter um resultado favorável. A razão de chances é planejada para determinar como as chances de sucesso, aumentam conforme certas mudanças que ocorram nos valores do regressor (variável explicativa) (HOFFMANN, 2012). A *odds ratio* é definida por:

$$OR = \exp(\beta_1 X_1 + \dots + \beta_q X_q), \quad (10)$$

em que,  $X_q$  são variáveis preditoras e  $\beta_q$  são parâmetros a serem estimado. Através da OR é possível mensurar o acréscimo ou decréscimo na  $P(Y=1)$  em relação a  $P(Y=2)$ , baseado na oscilação da variável  $X_q$ .

De maneira geral, para sua análise, aquelas variáveis que apresentarem valores acima de 1 são classificadas como fatores de risco para o problema em questão e aquelas que apresentarem valores abaixo de 1 classificam-se como fatores de proteção.

#### 4.6 VERIFICAÇÃO DA QUALIDADE DO MODELO AJUSTADO

Após realizar os ajustes necessários da análise de regressão logística, se faz necessário verificar a qualidade do modelo, ou seja, analisar se o modelo final obtido explica a relação entre as variáveis selecionadas com o menor erro possível, pois na ausência dessa verificação há a possibilidade de uma inferência equivocada dos dados. Sendo assim, são descritos na literatura alguns critérios, como a função de desvio (*Deviance*), o teste de *Hosmer e Lemeshow* e análise gráfica por meio da curva ROC (*Receiver Operating Characteristic*) (MARTINEZ et al, 2003; CORRAR et al, 2007; PAULA, 2010).

A função de desvio *Deviance* (D), que é uma distância entre o logaritmo da função de verossimilhança do modelo saturado, que corresponde ao melhor modelo possível (com  $n$  parâmetros) e do modelo sob investigação (com  $p$  parâmetros) avaliado na estimativa de máxima verossimilhança, é definido como:

$$D = -2 \ln \frac{LLO}{LL\beta} , \quad (11)$$

em que  $LL\beta$  trata-se do logaritmo da função de verossimilhança do modelo saturado e o  $LLO$  é o logaritmo da função de verossimilhança do modelo em investigação (PAULA, 2013). Na sua interpretação, um valor pequeno da função desvio indica que, para um número de parâmetros, se obtém um ajuste tão bom quanto o ajuste com um modelo saturado. Portanto, com um valor do desvio menor que a estatística qui-quadrado, com  $n - p$  graus de liberdade, não rejeitamos a hipótese nula que o modelo está bem ajustado.

Através do teste de *Hosmer e Lemeshow*, é possível verificar se as classificações previstas são iguais às observadas para cada grupo, aplicando o teste qui-quadrado é avaliado se existem diferenças estatisticamente significativas entre as frequências observadas e esperadas (CORRAR et al., 2007).

## 4.7 CURVA ROC

A curva ROC (*Receiver Operating Characteristic*) trata-se de uma ferramenta utilizada para a visualização do problema em avaliação. Através da análise é possível estabelecer a taxa de acertos e erros do modelo. A sensibilidade irá se referir a taxa de verdadeiros positivos, ou seja, quando o modelo discrimina os doentes que realmente estão doentes. Já a especificidade é a taxa que verifica os falsos positivos que é quando o modelo classifica os doentes que na verdade são saudáveis. A curva ROC precisa se apresentar numa forma convexa e crescente (FÁVERO et al., 2009).

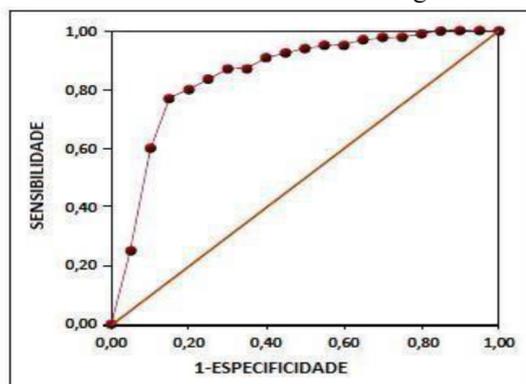
Existe uma área sobre a curva ROC denominada de AUC (*Area Under the Curve*) que é uma medida usada para comparar modelos. A melhor acurácia é aquela mais próxima de 1, mostrando assim ter uma melhor taxa de acerto (FÁVERO et al., 2009). A área sob a curva ROC fornece a medida de precisão do modelo, ou seja, a AUC fornece uma medida da capacidade modelo de discriminar entre aqueles indivíduos que vivenciam o desfecho de interesse daqueles que não o fazem. Quanto maior a AUC, maior é a capacidade do modelo em discriminar a ocorrência ou não ocorrência do evento de interesse, e, quanto mais próxima a curva ROC estiver da reta diagonal, pior é o poder discriminatório do modelo (FÁVERO, 2009). Dessa forma, pode-se observar as seguintes situações de acordo com a AUC:

- AUC = 1 → discriminação perfeita
- AUC = 0,90–0,99 → discriminação excelente
- AUC = 0,80–0,89 → discriminação boa
- AUC = 0,70–0,89 → discriminação moderada
- AUC = 0,60–0,69 → discriminação pobre
- AUC = 0,50–0,59 → discriminação falha

A classificação dos indivíduos como “com distúrbio de voz” e “sem distúrbio de voz” é obtida através do ponto de corte da curva ROC. É necessário selecionar um ponto que tenha uma ótima combinação entre a sensibilidade e especificidade, além de se aproxima do canto superior esquerdo do gráfico. Acima do ponto de corte os indivíduos são classificados como “com distúrbio de voz” e abaixo do ponto de corte como “sem distúrbio de voz” (SILVA, 2006).

Na figura 7, podemos observar que a cada ponto da curva ROC, círculos vermelhos ilustram a possibilidade de um possível ponto de corte do teste.

**Figura 7** - Curva ROC de um teste de diagnóstico hipotético



Após estabelecer o melhor ponto de corte para o modelo, pode-se seguir com a tabela de contingência (ou matriz de confusão) com o intuito de estabelecer as taxas de acerto e de erro do modelo final. Através desta tabela, também é possível observar o valor da sensibilidade, especificidade, acurácia, valor preditivo positivo, valor preditivo negativo, razão de verossimilhança positivo e negativo e o ponto de corte (SILVA, 2006).

#### 4.8 CONSIDERAÇÕES ÉTICAS

A presente pesquisa faz parte de um projeto maior intitulado “Acurácia das medidas de análise acústica linear e não linear na avaliação dos distúrbios da voz”, que foi apreciado e aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde/UFPB, com o parecer de nº 52492/12. Todos os participantes receberam explicação sobre a pesquisa e assinaram o termo de consentimento livre e esclarecido.

## 5 RESULTADOS E DISCUSSÕES

Os resultados obtidos nesse estudo serão descritos em duas subseções. A primeira subseção apresentará os resultados da análise de regressão logística e a segunda subseção apresentará os resultados da análise de regressão linear múltipla.

### 5.1 REGRESSÃO LOGÍSTICA

#### 5.1.1 Modelo de Regressão Logística para variável distúrbio de voz

O modelo de regressão logística foi aplicado para a variável distúrbio de voz, realizando-se um processo de triagem com todos os itens (**Tabela 3**).

**Tabela 3. Modelo de regressão logística para a variável: distúrbio de voz.**

Variável	Estimativa	Erro padrão	Valor Z	p-valor	OR
Intercept	3.8009	140.426	2.707	0.0068*	44.740
CPP_VGE	-0.03764	0.06636	-0.567	0.5706	0.9630588
CPP_VGA	0.05500	0.16191	0.340	0.7341	10.565.443
CPP_CAPE-V	-0.04635	0.02902	-1.597	0.1103	0.9547057
CPP_TODAS	-0.05122	0.16937	-0.302	0.7623	0.9500680
CPP_CONTAGEM	0.02184	0.04561	0.479	0.6321	10.220.771
CPPS_VGE	-0.08532	0.13010	-0.656	0.5119	0.9182177
CPPS_VGA	-0.10016	0.11963	-0.837	0.4024	0.9046887
CPPS_CAPE-V	-0.06353	0.16960	-0.375	0.7080	0.9384443
CPPS_TODAS	-0.15699	0.48090	-0.326	0.7441	0.8547129
CPPS_CONTAGEM	0.29912	0.27291	1.096	0.2731	13.486.759
DECL_VGE	-0.02228	0.05398	-0.413	0.6798	0.9779665
DECL_VGA	0.03348	0.04950	0.676	0.4987	10.340.494
DECL_CAPE-V	-0.07343	0.03755	-1.955	0.0500*	0.9292046
DECL_TODAS	-0.04111	0.09317	-0.441	0.6591	0.9597279
DECL_CONTAGEM	-0.02243	0.03931	-0.571	0.5683	0.9778231

**Legenda:** CPP: *Cepstral Peak Prominence*; CPPS: *Cepstral Peak Prominence Smoothed*; DECL: Declínio espectral; VGE: Vogal /E/ sustentada; VGA: Vogal /a/ sustentada, OR: *odds ratio* \*Significativo ao nível de 5% de probabilidade, Critério de Informação de Akaike (AIC) = 315.824.

Com base nas observações da Tabela 3, pode-se afirmar que apenas os parâmetros Declínio espectral da frase do CAPE-V foi estatisticamente significativo ao nível de 5% de probabilidade. Diante disto, mantendo-se as demais variáveis, quando eleva-se o Declínio espectral da frase do CAPE-V em uma unidade espera-se que o log das chances (desfecho= 1 = Ter distúrbio) se altere em -0,073, em termos de Razão de Chances quer dizer que, com o aumento em uma unidade no declínio espectral da frase do CAPE-V espera-se uma diminuição 7,1% (1-0.929) da probabilidade de apresentar um distúrbio de voz (desfecho=1).

No entanto como em sua maioria as variáveis explicativas são não significativas faz-se necessário investigar um modelo reduzido. Em geral, os valores estimados possuem efeitos negativos para determinação das chances do preditor ser igual a 1, caso contrário constariam com sinal positivo. A fim de se obter um modelo que melhor descreva a variável distúrbio de voz, foi utilizado o método de *Stepwise* para se encontrar o modelo mais adequado, como se pode visualizar na tabela 4.

**Tabela 4. Modelo de regressão logística reduzido com base no método de *Stepwise* para a variável distúrbio de voz.**

Variável	Estimativa	Erro padrão	Valor Z	p-valor	OR	IC (OR; 95%)
Intercept	3.6314	1.0171	3.570	0.00023*	37.7658	5.43-296.86
CPP_CAPE-V	-0.05817	0.02646	-2.198	0.02109*	0.9434929	0.89-0.99
CPPS_VGE	-0.27914	0.05868	-4.757	0.0000*	0.7564370	0.67-0.84
DECL_CAPE-V	-0.10679	0.03104	-3.441	0.00058*	0.8987108	0.84-0.95

**Legenda:** CPP: *Cepstral Peak Prominence*; CPPS: *Cepstral Peak Prominence Smoothed*; DECL: Declínio espectral; VGE: Vogal /E/ sustentada. \*Significativo ao nível de 5% de probabilidade, IC: Intervalo de Confiança, OR: *odds ratio*, Critério de Informação de Akaike (AIC) =300.175.

Com base nas observações da Tabela 4, pode-se afirmar que em geral os parâmetros do modelo foram significativos ao nível de 5% de probabilidade, e ao se comparar os valores de AIC do modelo completo e reduzido, é possível inferir que o modelo reduzido com as variáveis CPP CAPE-V, CPPS VGE e declínio CAPE-V, é o mais indicado para descrever a variável distúrbio de voz. Os valores estimados possuem efeitos negativos para determinação das chances do preditor ser igual a 1, caso contrário constariam com sinal positivo.

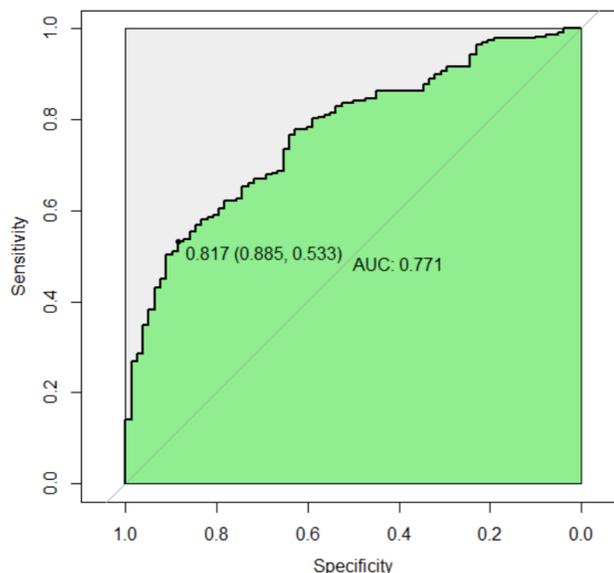
Avaliando-se as variáveis do modelo, uma a uma, é possível inferir com base no modelo que:

- Mantendo-se as demais variáveis constantes, elevando-se a variável CPPS VGE em uma unidade espera-se que o log das chances (desfecho=1) se altere em -0.279. Dessa forma, com o aumento em uma unidade do CPPS VGE espera-se uma diminuição de 24,36% (1-0.7564) na probabilidade de o indivíduo apresentar distúrbio de voz.
- Mantendo-se as demais variáveis constantes, elevando-se a variável DECL-CAPE-V em uma unidade espera-se que o log das chances (desfecho=1) se altere em -0.107, em termos de Razão de Chances. Dessa forma, com o aumento em uma unidade da mesma variável espera-se uma diminuição 10,13% (1-0.8987) de se ter distúrbio de voz.
- Mantendo-se as demais variáveis constantes, elevando-se a variável CPP CAPE-V em uma unidade espera-se que o log das chances (desfecho=1) se altere em -0,058, em termos de Razão de Chances. Dessa forma, que com o aumento em uma unidade da mesma variável espera-se uma diminuição de 5,46% (1-0.9434) na probabilidade de ocorrência de um distúrbio de voz.

Os limites do intervalo de confiança foram calculados para todas as variáveis contidas no modelo final. Quando avaliados os limites superior e inferior de cada variável, foi possível observar que estão abaixo de 1, permitindo assim classificar que o resultado obtido através da OR tem as variáveis CPPS VGE, DECL CAPE-V e CPP CAPE-V como fatores de proteção, ou seja, diminuem a chance de ocorrência do distúrbio de voz.

Na figura 8, é exposto o comportamento gráfico do modelo logístico, por meio da curva ROC.

**Figura 8.** Curva ROC para a variável distúrbio de voz



Com base na Figura 8, estima-se uma acurácia moderada de 70,7%, sensibilidade de 85%, especificidade de 53%, RVN de 0,215, RVP de 1,895, sendo o ponto de corte padrão, o valor de 0,79: acima deste valor, a predição é classificada como presença de distúrbio de voz e, abaixo dele com ausência de distúrbio de voz.

Através das informações obtidas, é possível afirmar que este modelo é sensível para detectar indivíduos com distúrbio de voz. Diante disso, pode ser utilizado para redução de possibilidades diagnósticas, sendo indicado por exemplo, para se utilizar em triagens vocais. O que possibilita um diagnóstico precoce de uma doença, interferindo desse modo no prognóstico (LIU et al., 2005).

A presença de queixa vocal e visualização direta das estruturas da laringe e do padrão vibratório das pregas vocais são importantes tanto para o diagnóstico como para o tratamento apropriado do distúrbio de voz. O processo de decisão clínica e a interpretação da imagem laríngea envolvem julgamentos perceptivos da imagem das estruturas da laringe e dos seus movimentos vibratórios (OMORI et al., 1998).

Ao final desta análise, é possível observar que as variáveis que melhor descrevem o modelo com base nos parâmetros do distúrbio de voz foram CPPS VGE, CPP CAPE-V e Declínio espectral CAPE-V. Nos estudos de Awan et al., (2013) e Awan et al., (2015) foi investigado a capacidade das medidas cepstrais discriminarem vozes saudáveis de vozes com

distúrbio de voz, eles utilizaram um modelo de regressão binomial e diversos parâmetros e um deles foi a presença/ausência de distúrbio de voz onde obtiveram também com a curva ROC uma acurácia de 73%. No presente estudo o valor de acurácia foi próximo a dos estudos dispostos na literatura.

O CPPS deriva de dois processos de “suavização” antes do cálculo da proeminência do pico cepstral. Maryn et al., (2014) destaca a relevância dessa medida: eles realizaram uma análise dos coeficientes de correlação entre as medidas acústicas e a avaliação da percepção auditiva da qualidade vocal, e afirmam que o CPPS é uma medida robusta tanto em amostras de vogais sustentada quanto em fala encadeada. Igualmente no modelo final, foi possível observar que a variável CPPS em contexto de vogal foi significativa no modelo final.

Com esse modelo é possível notar que todas as medias acústicas que foram extraídas estão presentes, tanto o CPP quanto o CPPS e o declínio espectral, sendo assim é possível afirmar que as três medidas são necessárias no que diz respeito à identificação de indivíduos com e sem distúrbio de voz.

Na produção da vogal sustentada há uma maior estabilidade do filtro do trato vocal. Sendo assim, nessa tarefa se tem uma maior chance em identificar indivíduos sem distúrbio de voz, visto que, as vogais sustentadas são menos propensas à variabilidade fonética (LOPES et al., 2019). Halberstam (2004) também analisou as medidas cepstrais de 60 indivíduos classificados como sem distúrbio e com distúrbio de voz e relatou uma maior correlação da fala encadeada para identificar o distúrbio de voz, quando comparada com a vogal sustentada.

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 (LOPES et al., 2019). A triagem de indivíduos saudáveis ou em risco para sinais precoces de anormalidade vocal se faz relevante porque algumas formas de distúrbio de voz podem ser tratadas com maior efetividade se detectadas precocemente. Por exemplo, ser professor é uma profissão de alto risco para o desenvolvimento de distúrbios da voz. Dadas as intensas demandas vocais da profissão, não é surpresa que os distúrbios vocais sejam um risco ocupacional relativamente comum, com 11% dos professores relatando um distúrbio vocal atual e 58% relatando um histórico de distúrbio vocal durante a vida. O distúrbio vocal interfere na satisfação no trabalho, no desempenho e na frequência, fazendo com que 18% dos professores relatem faltas ao trabalho anualmente (AWAN et al., 2015). Sendo assim, o modelo final de decisão possui potencial para ser usado em casos de triagens vocais.

### 5.1.2 Curva ROC para as Medidas Cepstrais

Sensibilidade e especificidade são qualidades intrínsecas no processo de um diagnóstico. Quando o resultado desse procedimento oferece duas categorias que definem a presença (diagnóstico positivo) ou a ausência (diagnóstico negativo), então esse procedimento apresenta apenas um único par de valores referentes à sensibilidade e à especificidade. Porém, na maior parte dos procedimentos diagnósticos, se produz um resultado sob a forma de variável contínua. Sendo assim, os valores de sensibilidade e especificidade irão variar de acordo com o ponto de corte empregado para que se possa chegar a um diagnóstico. Sensibilidade e especificidade possuem uma relação inversamente proporcional. Sendo assim, quando mudado o ponto de corte de um procedimento diagnóstico, é possível aumentar a sensibilidade desse procedimento em detrimento de sua especificidade e vice-versa, pelo fato desses dois parâmetros variarem de acordo com diferentes pontos de corte. Visto isso, o ponto de corte ideal é aquele que consegue o equilíbrio entre sensibilidade e especificidade de um procedimento diagnóstico (LIU et al., 2005).

A melhor maneira de representar o comportamento de um pontos de corte sobre a sensibilidade e a especificidade de um procedimento diagnóstico é através da curva ROC. O valor da área sob a curva pode ser utilizado para realizar inferências estatísticas acerca do desempenho de um procedimento diagnóstico, testando sua significância estatística em relação a outros procedimentos ou a uma situação aleatória. Sendo assim, pode-se derivar um valor de  $p$  para toda a área sob a curva ROC associada a um procedimento diagnóstico. Esse valor testa a hipótese nula de que o procedimento diagnóstico não seria capaz de diferenciar as pessoas entre o grupo saudável e o clínico, ou seja, os resultados desse procedimento clínico não são aleatórios (LIU et al., 2005).

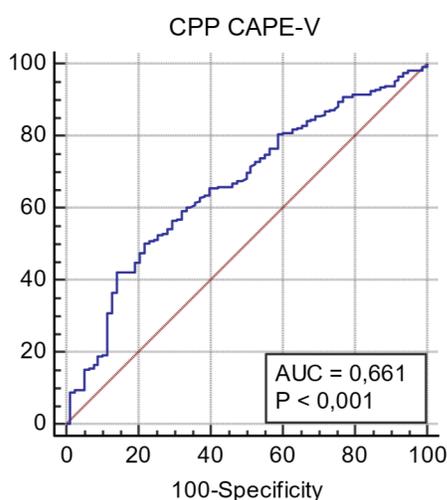
Sendo assim, se objetivou aplicar a curva ROC para as medidas cepstrais de maneira isolada. Tendo como objetivo a extração dos valores de: Acurácia, sensibilidade, especificidade, valores preditivos positivos e negativos, razão de verossimilhança negativo (RVN) e positivo (RVP) e o ponto de corte. As únicas variáveis que não foram significativas ao nível de 5% com valor de ( $p > 0,005$ ) foram: CPP e CPPS da contagem, não sendo assim capazes de diferenciar indivíduos com e sem distúrbio de voz.

Sobre a Razão de Verossimilhança (RV), quanto maior a taxa de probabilidade positiva, maior a probabilidade de um distúrbio de voz; quanto menor a razão de probabilidade negativa, menor a probabilidade de um desvio. Razões de probabilidade sempre se referem à

probabilidade de se ter um distúrbio de voz; a designação positiva e negativa simplesmente se refere ao resultado do teste. Assim, RV + indica quanto a probabilidade de um distúrbio aumentar se o teste for positivo. A RV - indica quanto a probabilidade de um distúrbio diminuir se o teste for negativo (AWAN et. al, 2013). A curva ROC das medidas estudadas na presente pesquisa foram extraídas do *software* MedCalc - versão 19.6.4.

As medidas com menor valor de acurácia (<70%) foram as seguintes:

**Figura 9** – Curva ROC associada a medida CPP CAPE-V



**Tabela 5- Valores associados à Variável CPP CAPE-V.**

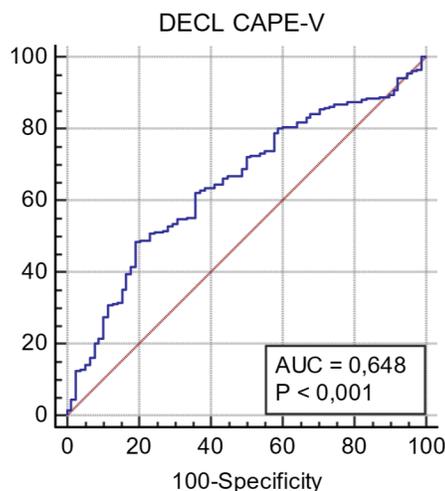
Variável	CPP CAPE-V
Acurácia	0,661
Sensibilidade	50,22
Especificidade	78,21
Valor preditivo Positivo	2,18
Valor Preditivo Negativo	0,65
Ponto de Corte	18,84
RVN	0,636
RVP	2,304

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPP: *Cepstral Peak Prominence*.

Com base na figura 9, estima-se uma acuraria aceitável de 66,1%, Intervalo de confiança para a sensibilidade [43,5;56], intervalo de confiança para especificidade [66,0;85,7]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPP na tarefa de fala “frase CAPE-V” acima de

18,84 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 10** – Curva ROC associada a medida DECL CAPE-V

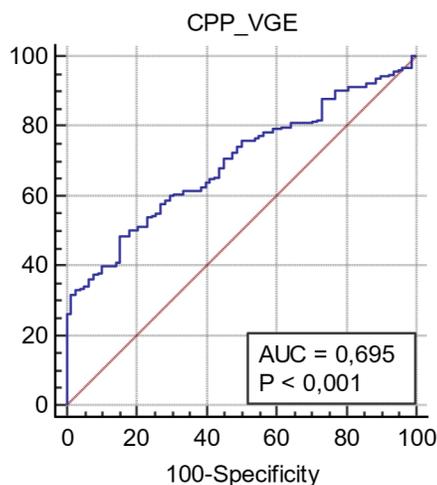


**Tabela 6. Valores associados à Variável DECL CAPE-V**

Variável	DECL CAPE-V
Acurácia	0,648
Sensibilidade	48,44
Especificidade	80,77
Valor preditivo Positivo	0,192
Valor Preditivo Negativo	0,515
Ponto de Corte	-16,20
RVN	0,638
RVP	2,518

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, DECL: Declínio espectral.

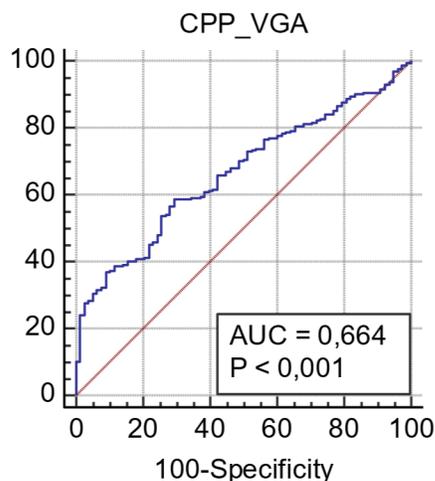
Com base na figura 10, estima-se uma acuraria aceitável de 64,8%, Intervalo de confiança para a sensibilidade [41,8;55,2], intervalo de confiança para especificidade [70,3;88,8]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de declínio espectral na tarefa de fala “frase CAPE-V” acima de -16,10 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 11** – Curva ROC associada a medida CPP VGE**Tabela 7. Valores associados à Variável CPP VGE**

Variável	CPP VGE
Acurácia	0,695
Sensibilidade	48,44
Especificidade	84,62
Valor preditivo Positivo	3,15
Valor Preditivo Negativo	0,61
Ponto de Corte	24,86
RVN	0,609
RVP	3,149

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPP: *Cepstral Peak Prominence*, VGE: Vogal /É/ sustentada.

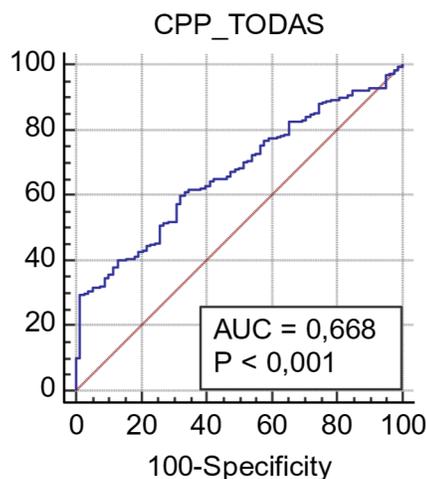
Com base na figura 11, estima-se uma acuraria aceitável de 69,5%, Intervalo de confiança para a sensibilidade [41,8;55,2], intervalo de confiança para especificidade [74,7;91,8]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPP na tarefa de fala “VGE” acima de 24,86 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 12** – Curva ROC associada a medida CPP VGA**Tabela 8. Valores associados à Variável CPP VGA**

Variável	CPP VGA
Acurácia	0,664
Sensibilidade	58,67
Especificidade	70,51
Valor preditivo Positivo	1,99
Valor Preditivo Negativo	0,59
Ponto de Corte	25,92
RVN	0,586
RVP	1,989

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPP: *Cepstral Peak Prominence*, VGA: Vogal /a/ sustentada.

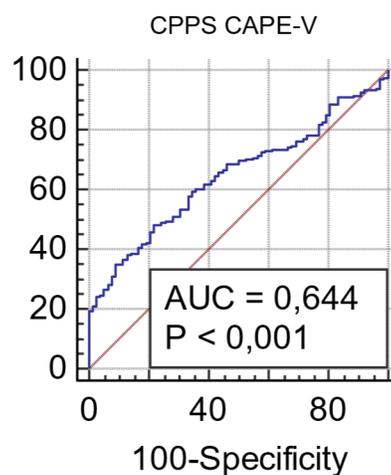
Com base na figura 12, estima-se uma acuraria aceitável de 66,4%, Intervalo de confiança para a sensibilidade [51,9;65,2], intervalo de confiança para especificidade [59,180,3]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPP na tarefa de fala “VGA” acima de 25,92 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 13** – Curva ROC associada a medida CPP Todas**Tabela 9. Valores associados à Variável CPP TODAS**

Variável	CPP TODAS
Acurácia	0,668
Sensibilidade	29,33
Especificidade	98,72
Valor preditivo Positivo	22,88
Valor Preditivo Negativo	0,72
Ponto de Corte	21,69
RVN	0,715
RVP	22,914

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPP: *Cepstral Peak Prominence*.

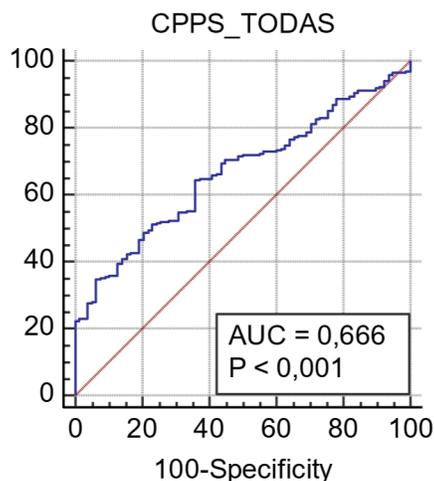
Com base na figura 13, estima-se uma acuraria aceitável de 66,8%, Intervalo de confiança para a sensibilidade [23,5;35,8], intervalo de confiança para especificidade [93,1;100,00]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPP em todas as tarefas de fala, acima de 21,69 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 14** – Curva ROC associada a medida CPPS CAPE-V**Tabela 10. Valores associados à Variável CPPS CAPE-V**

Variável	CPPS CAPE-V
Acurácia	0,644
Sensibilidade	48,00
Especificidade	78,21
Valor preditivo Positivo	2,20
Valor Preditivo Negativo	0,66
Ponto de Corte	11,98
RVN	0,664
RVP	2,202

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*.

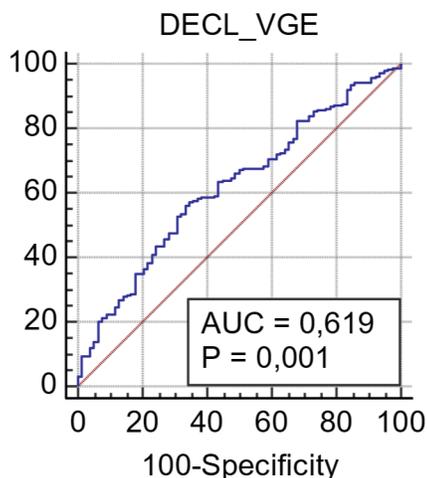
Com base na figura 14, estima-se uma acuraria aceitável de 64,4%, Intervalo de confiança para a sensibilidade [41,3;54,7], intervalo de confiança para especificidade [67,4;86,8]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPPS “frase CAPE-V”, acima de 11,98 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 15** – Curva ROC associada a medida CPPS TODAS**Tabela 11. Valores associados à Variável CPPS TODAS**

Variável	CPPS TODAS
Acurácia	0,666
Sensibilidade	64,44
Especificidade	64,10
Valor preditivo Positivo	1,80
Valor Preditivo Negativo	0,55
Ponto de Corte	10,5975
RVN	0,554
RVP	1,794

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*.

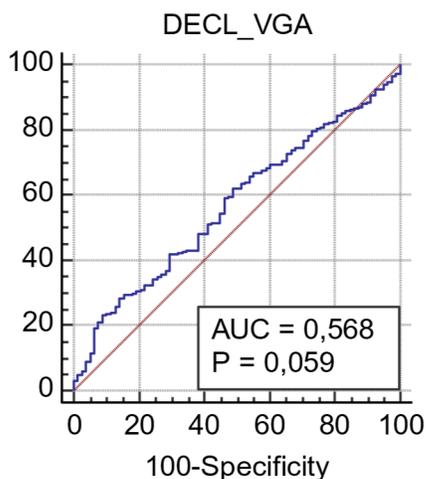
Com base na figura 15, estima-se uma acuraria aceitável de 66,6%, Intervalo de confiança para a sensibilidade [57,8;70,7], intervalo de confiança para especificidade [52,4;74,7]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPPS com todas as tarefas de fala, acima de 10,5975 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 16** – Curva ROC associada a medida Decl. VGE**Tabela 12. Valores associados à Variável DECL VGE**

Variável	DECL VGE
Acurácia	0,619
Sensibilidade	57,33
Especificidade	65,38
Valor preditivo Positivo	1,66
Valor Preditivo Negativo	0,65
Ponto de Corte	-15,327
RVN	0,652
RVP	1,655

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, DECL: Declínio espectral, VGE: Vogal /E/ sustentada.

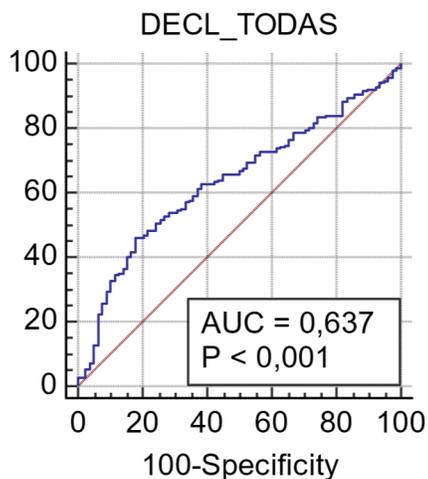
Com base na figura 16, estima-se uma acuraria aceitável de 61,9%, Intervalo de confiança para a sensibilidade [50,6;63,9], intervalo de confiança para especificidade [52,4;74,7]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de declínio espectral “VGE”, acima de -15,327 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 17** – Curva ROC associada a medida Decl. VGA**Tabela 13. Valores associados à Variável DECL VGA**

Variável	DECL VGA
Acurácia	0,568
Sensibilidade	28,44
Especificidade	85,90
Valor preditivo Positivo	2,02
Valor Preditivo Negativo	0,83
Ponto de Corte	-16,50
RVN	0,833
RVP	2,017

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, DECL: Declínio espectral, VGA: Vogal /a/ sustentada.

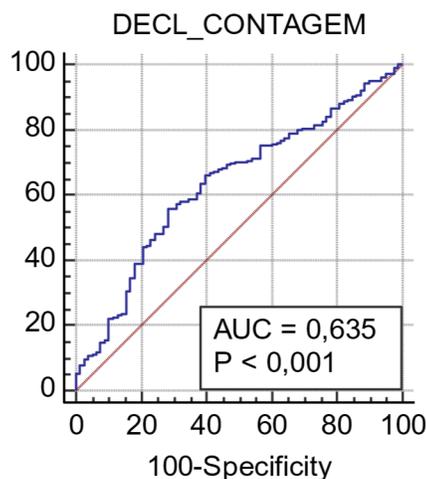
Com base na figura 17, estima-se uma acuraria aceitável de 56,8%, Intervalo de confiança para a sensibilidade [22,6;34,8], intervalo de confiança para especificidade [76,2;92,7]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de declínio espectral “VGA”, acima de -16,50 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 18** – Curva ROC associada a medida Decl. TODAS**Tabela 14. Valores associados à Variável DECL TODAS**

Variável	DECL TODAS
Acurácia	0,637
Sensibilidade	46,22
Especificidade	82,05
Valor preditivo Positivo	2,58
Valor Preditivo Negativo	0,66
Ponto de Corte	-17,71
RVN	0,655
RVP	2,574

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, DECL: Declínio espectral.

Com base na figuras 18, estima-se uma acuraria aceitável de 63,7%, Intervalo de confiança para a sensibilidade [39,6;53,0], intervalo de confiança para especificidade [71,7;89,8]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de declínio espectral em todas as tarefas de fala, acima de -17,71 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 19** – Curva ROC associada a medida Decl. CONTAGEM**Tabela 15. Valores associados à Variável DECL CONTAGEM**

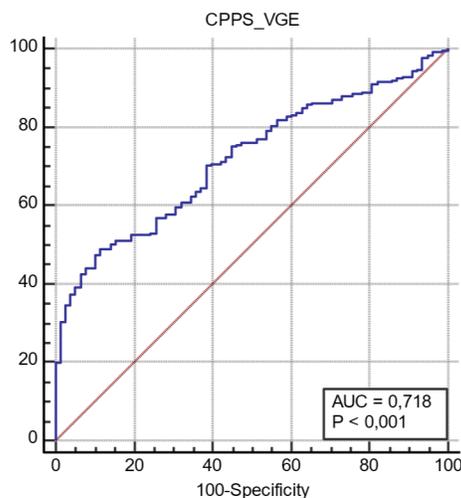
Variável	DECL CONTAGEM
Acurácia	0,635
Sensibilidade	56,00
Especificidade	71,79
Valor preditivo Positivo	1,99
Valor Preditivo Negativo	0,61
Ponto de Corte	-16,43
RVN	0,612
RVP	1,985

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, DECL: Declínio espectral.

Com base na figura 19, estima-se uma acuraria aceitável de 63,5%, Intervalo de confiança para a sensibilidade [49,2;62,6], intervalo de confiança para especificidade [60,5;81,4]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de declínio espectral em “contagem”, acima de -16,43 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

As medidas com maior valor de acurácia (>70%) foram as seguintes:

**Figura 20** – Curva ROC associada a medida CPPS VGE

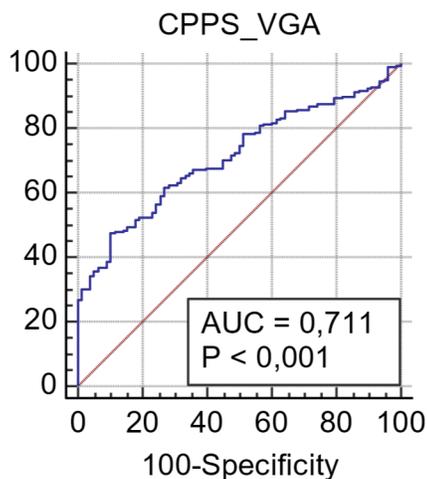


**Tabela 16. Valores associados à Variável CPPS VGE**

Variável	CPPS VGE
Acurácia	0,718
Sensibilidade	48,89
Especificidade	88,46
Valor preditivo Positivo	0,115
Valor Preditivo Negativo	0,515
Ponto de Corte	14,01
RVN	0,577
RVP	4,236

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, VGE: Vogal /É/ sustentada.

Com base na figura 20, estima-se uma acuraria aceitável de 71,8%. Intervalo de confiança para a sensibilidade [42,2;55,6], intervalo de confiança para especificidade [76,2;92,7]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPPS na tarefa de fala “VGE” acima de 14,01 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

**Figura 21** – Curva ROC associada a medida CPPS VGA**Tabela 17. Valores associados à Variável CPPS VGA**

Variável	CPPS VGA
Acurácia	0,711
Sensibilidade	47,56
Especificidade	89,74
Valor preditivo Positivo	4,64
Valor Preditivo Negativo	0,58
Ponto de Corte	13,60
RVN	0,587
RVP	4,635

**Legenda:** RVN: Razão de verossimilhança negativo, RVP: Razão de verossimilhança positivo, CPPS: *Cepstral Peak Prominence.Smoothed*, VGA: Vogal /a/ sustentada.

Com base na figura 21, estima-se uma acuraria aceitável de 71,1%, Intervalo de confiança para a sensibilidade [40,9;54,3], intervalo de confiança para especificidade [80,8;95,5]. Dessa forma, tomando-se como base o valor do ponto de corte, espera-se que indivíduos que apresentem um valor de CPP “VGA”, acima de 13,60 classifiquem-se sem distúrbio de voz e, de modo contrário, um valor abaixo do ponto de corte indicaria, possivelmente, que o indivíduo apresenta algum grau de distúrbio de voz.

As medidas com maior acurácia foram: CPPS VGA e CPPS VGE. Nessas duas medidas os valores de especificidade foram maiores do que os de sensibilidade. Sendo assim, essas medidas de maneira isolada são relevantes no processo de monitoramento e confirmação diagnóstica de indivíduos com distúrbio de voz. Desse modo, é possível reabilitar um indivíduo com distúrbio de voz e monitorá-lo de acordo com esses valores, indicando assim se a condição laríngea está melhorando. Vale salientar, que esse desfecho foi dado de acordo com a presença e ausência de distúrbio, os resultados podem ser diferentes com outros desfechos.

## 5.2 REGRESSÃO LINEAR MÚLTIPLA

Considerando o desvio padrão que foi encontrado na análise descritiva no GT e GC, tivemos a intenção com a regressão linear múltipla de observar o quanto o desvio da qualidade vocal explicaria a variabilidade nas medidas cepstrais.

**Tabela 18. Modelo de regressão linear múltipla para a variável EAV-G.**

Variável	Estimativa	Erro padrão	Valor Z	p-valor
CPP_VGE	0,059871	0,218208	0,274	0,783992
CPP_VGA	-0,4086	0,568602	-0,719	0,472969
CPP_CAPE-V	-0,24037	0,103433	-2,324	0,020826
CPP_TODAS	0,192724	0,592927	0,325	0,745387
CPP_CONTAGEM	0,005697	0,158363	0,036	0,971328
CPPS_VGE	-0,55083	0,421795	-1,306	0,192617
CPPS_VGA	-1,39069	0,398711	-3,488	0,000562*
CPPS_CAPE-V	-1,0342	0,590818	-1,75	0,081101
CPPS_TODAS	-1,11105	1,730109	-0,642	0,521262
CPPS_CONTAGEM	0,54177	0,972238	0,557	0,577795
DECL_VGE	-0,00529	0,177399	-0,03	0,976224
DECL_VGA	0,244654	0,158623	1,542	0,12408
DECL_CAPE-V	-0,08834	0,142522	-0,62	0,535869
DECL_TODAS	-0,05994	0,317751	-0,189	0,850514
DECL_CONTAGEM	-0,06299	0,150807	-0,418	0,676483

**Legenda:** CPP: *Cepstral Peak Prominence*; CPPS: *Cepstral Peak Prominence Smoothed*; DECL: Declínio espectral; VGE: Vogal /É/ sustentada; VGA: Vogal /a/ sustentada. \*Significativo ao nível de 5% de probabilidade,  $R^2_{aj}=63,6\%$ .

Com base nas observações da Tabela 18, pode-se afirmar que o parâmetro CPPS VGA foi significativo ao nível de 5% de probabilidade. A fim de se obter um modelo que melhor descreva a variável EAV-G, utilizou-se o método de *Stepwise* para se encontrar o modelo mais adequado, como pode-se visualizar na tabela 19.

**Tabela 19. Modelo de regressão linear múltipla reduzido com base no método de *Stepwise* para a variável dependente EAV-G.**

Variável	Estimativa	Erro padrão	Valor Z	p-valor	OR	IC (OR; 95%)
CPP_CAPE-V	-0,24345	0,09811	-2,481	0,0136*	-0,24345	-0,436;-0,050
CPPS_VGE	-0,60314	0,28047	-2,15	0,0323*	-0,60314	-1,155;-0,051
CPPS_VGA	-1,7195	0,25881	-6,644	1,44E-10*	-1,7195	-2,228;-1,210
CPPS CAPE-V	-1,42596	0,32198	-4,429	1,33E-05*	-1,42596	-2,059;-0,792

**Legenda:** CPP: *Cepstral Peak Prominence*; CPPS: *Cepstral Peak Prominence Smoothed*; DECL: Declínio espectral; VGE: Vogal /É/ sustentada; VGA: Vogal /a/ sustentada, IC: Intervalo de Confiança, OR: *odds ratio* \*Significativo ao nível de 5% de probabilidade,  $R^2_{aj} = 64,32\%$ .

Com base nas observações da Tabela 19, pode-se afirmar que em geral os parâmetros do modelo foram significativos ao nível de 5% de probabilidade, e ao se comparar os valores de  $R^2_{aj}$  é maior que o do modelo completo, indicando maior qualidade no ajuste, bem como menor erro padrão. Os limites do intervalo de confiança foram calculados para todas as variáveis contidas no modelo final. Quando avaliados os limites superior e inferior de cada variável, foi possível observar que em geral os intervalos foram enxutos, não assumindo valor zero o que confere qualidade às estimativas.

- Mantendo-se as demais variáveis constantes, com o aumento médio em uma unidade da variável CPP CAPE-V espera-se em média uma diminuição de 0,24 na EAV-G;
- Mantendo-se as demais variáveis constantes, com o aumento médio em uma unidade da variável CPPS VGE espera-se em média a uma diminuição de 0,64 na EAV-G;
- Mantendo-se as demais variáveis constantes, com o aumento médio em uma unidade da variável CPPS VGA espera-se em média uma diminuição de 1,71 na EAV-G;
- Mantendo-se as demais variáveis constantes, com o aumento médio em uma unidade da variável CPPS CAPE-V espera-se em média uma diminuição de 1,42 na EAV-G.

Com o modelo final foi observado que, com o aumento médio do CPP CAPE-V, CPPS VGE, CPPS VGA e CPPS CAPE-V há a diminuição na EAV-G, ou seja, a voz é marcada como tendo menor desvio. Esses dados corroboram para o que há na literatura sobre as medidas cepstrais, onde sinais com maior regularidade e menor quantidade de ruído apresentam maior definição e amplitude do pico cepstral dominante, ou seja, quanto menor o distúrbio maior o valor das medidas cepstrais (DEJONCKERE et al., 1996).

Awan et al., (2011) observaram no estudo que a previsão mais forte de gravidade do distúrbio de voz quanto ao grau geral foi para a frase CAPE-V totalmente vozeada ("*We were away a year ago*"). Os resultados do presente estudo corroboram para os achados na literatura onde a frase vozeada “olha lá o avião azul” obteve significância na predição de presença de desvio vocal.

Além disso, a variável CPPS VGA apresenta maior valor na OR. Portanto, se pode afirmar que essa variável é a que exerce maior influência no que diz respeito a diminuição dos valores na escala EAV-G. De acordo com a literatura (LAW et al., 2012) a vogal sustentada é mais fácil de ser obtida e de padronizar, pois não é afetada pela articulação dos sons da fala e é caracterizada por uma configuração relativamente estática da laringe e dos músculos supralaríngeos com o mínimo de variação durante sua produção. Esse fato colabora para uma

alta concordância interavaliadores. Lu e Matteson (2014) correlacionaram a tarefa de fala à confiabilidade encontrada na percepção auditiva da voz, e a tarefa que apresentou maior concordância interavaliadores foi a emissão da vogal /a/ sustentada de forma habitual, para o grau geral.

De acordo com o modelo ajustado foi possível observar, que com as variáveis presentes ele foi capaz de explicar em 64,32% a variabilidade da intensidade do desvio vocal. Correlacionando este modelo com o de regressão logística que não considerou a presença de desvio da qualidade vocal se nota que as medidas cepstrais (CPP, CPPS e declínio espectral) foram significativas em ambos os modelos, além disso, ambas as tarefas de fala estiveram presentes nos modelos, tanto a fala encadeada quanto a vogal sustentada. O que nos faz refletir, que para análise de presença e ausência de distúrbio de voz, assim como a presença de desvio da qualidade vocal a presença dessas tarefas de fala e das medidas cepstrais se faz relevante visto que apresentam uma boa adequação ao problema trabalhado e desta maneira podem ser uma ferramenta de apoio na tomada de decisão para a conduta fonoaudiológica como também para construção de estratégias terapêuticas.

Se destaca a importância da realização de um estudo (que era objetivo inicial desta pesquisa), para verificar o quanto o grau de desvio vocal pode justificar o desfecho: presença e ausência de distúrbio de voz.

## 6. ÍNDICE CEPSTRAL

O índice cepstral para o português brasileiro ( $IC_{Br}$ ) foi calculado a partir de duas medidas da frase do CAPE-V “Olha lá o avião azul”, as medidas foram: declínio espectral e *cepstral peak prominence* e uma medida da vogal sustentada /E/ que foi o *cepstral peak prominence smoothed*. A vogal e frase foram inseridas no mesmo arquivo de áudio.

$$IC_{Br} = \frac{e^{3,6314-0,27914CPPSvge-0,10679DECLfrase-0,05817CPPfrase}}{1 + e^{3,6314-0,27914CPPSvge-0,10679DECLfrase-0,05817CPPfrase}}$$

## 7. CONCLUSÃO

O modelo reduzido de regressão logística para identificação da presença e ausência de distúrbio de voz, inclui as variáveis: CPP CAPE-V, CPPS VGE e declínio CAPE-V. O modelo

apresentou um desempenho de 70,7%, com sensibilidade de 85% e especificidade de 53%, o que pode corroborar com a sua recomendação para procedimentos de triagem vocal.

Isoladamente, apenas as variáveis CPPS VGE e CPPS VGA apresentaram desempenho satisfatório para discriminar indivíduos com e sem distúrbio de voz, com altos valores de especificidade e baixos valores de sensibilidade. Dessa forma, os valores de corte na situação de validação desta pesquisa podem ser indicados para procedimentos de confirmação diagnóstica e monitoramento da condição laríngea do paciente durante o tratamento, seja ele terapêutico ou cirúrgico. Valores abaixo de 14,01 no CPPS VGE indicam a presença de um distúrbio de voz e valores abaixo de 13,60 no CPPS VGA indicam presença de um distúrbio de voz.

As medidas: CPP CAPE-V, CPPS VGE, CPPS VGA e CPPS CAPE-V são capazes de prever a variabilidade da intensidade do desvio vocal em 64,32%, quanto maiores os valores nessas medidas menor a probabilidade de desvio vocal percebido auditivamente.

As medidas cepstrais incluídas nesse estudo fornecem informações relativamente únicas, significativas e complementares, mostrando ter relações com os diferentes graus de desvio da qualidade vocal em falantes do português brasileiro. Essas medidas informam o clínico sobre os aspectos de um distúrbio da voz que refletem na sua dimensão multidimensional. Além disso, nesses modelos tem-se uma estratégia mais rápida quanto a tomada de decisão no que diz respeito a identificação de desvios vocais. Os resultados também contribuem para a literatura na área da análise acústica e modelos de decisão em saúde, servindo de subsídio para futuros estudos.

Uma das possíveis limitações deste estudo pode estar relacionada ao uso apenas de um desfecho: presença e ausência de distúrbio de voz. Considerando que as medidas acústica possuem grande relação com a qualidade vocal, deve-se explorar valores de corte para outros desfecho como: presença e ausência de desvio da qualidade vocal, independente da presença de distúrbio laríngeo e o padrão de referência combinado, incluindo a presença do desvio da qualidade vocal com a presença do distúrbio de voz. Além disso, se sugere como perspectivas de estudos futuros a aplicação de modelos de decisão com as medidas cepstrais levando em consideração a intervenção terapêutica, pré, durante e pós fonoterapia.

## REFERÊNCIAS

AGUIAR NETO, B. G., FECHINE, J. M., COSTA, S. CUNHA, MUPPA, M., Feature Estimation for Vocal Fold Edema Detection Using Short-Term Cepstral Analysis. Proceedings of the 7th International Conference on Bioinformatics and Bioengineering, 14-17 Oct., page(s) 1158-1162, 2007.

AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION. The roles of otolaryngologists and speech-language pathologists in the performance and interpretation of stroboscovideolaryngoscopy [Relevant paper]. Retrieved from <http://www.asha.org/policy>, 1998.

AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION. Evidence based practice in communication disorders: An introduction [Technical report]. Retrieved from <http://www.asha.org/policy>, 2004.

ANAND, S., KOPF, L. M., SHRIVASTAV, R., & EDDINS, D. A. Objective Indices of Perceived Vocal Strain. *Journal of Voice*. doi:10.1016/j.jvoice..06.005, 2018.

AWAN, S. N., & ROY, N. Acoustic Prediction of Voice Type in Women with Functional Dysphonia. *Journal of Voice*, 19(2), 268–282, 2005.

AWAN SN, ROY N. Outcomes measurement in voice disorders: application of an acoustic index of dysphonia severity. *J Speech Lang Hear Res*.52(2):482-99, 2009.

AWAN, S. N., ROY, N., & DROMEY, C. Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectral/cepstral model. *Clinical Linguistics & Phonetics*, 23(11), 825–841. doi:10.3109/02699200903242988, 2009.

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*. 24(9):742-58, 2011.

AWAN, S. N., HELOU, L. B., STOJADINOVIC, A., & SOLOMON, N. P. Tracking voice change after thyroidectomy: application of spectral/cepstral analyses. *Clinical Linguistics & Phonetics*, 25(4), 2010.

AWAN SN, SOLOMON NP, HELOU LB, STOJADINOVIC A. Spectral-Cepstral estimation of dysphonia severity: external validation. *Ann Otol Rhinol Laryngol*. 122(1):40-8, 2013.

AWAN, S. N.; ROY, N.; COHEN, S. M. Exploring the Relationship Between Spectral and Cepstral Measures of Voice and the Voice Handicap Index (VHI). *Journal of Voice*, 28(4), 430–439, 2014.

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*. Mar;30(2):130-44, 2015.

BARRETO, A. S.; Modelos de Regressão: Teorias e Aplicações com o Programa Estatístico R. 1ª Ed. Brasília: Ed. do Autor, 2011.

BARSTIES B, DE BODT M. Assessment of voice quality: current state-of-the-art. *Auris Nasus Larynx*. 42:183–188, 2014.

BARSTIES B, MARYN Y. External validation of the acoustic voice quality index version 03.01 with extended representativity. *Ann Otol Rhinol Laryngol*. 125(7):571-83. <http://dx.doi.org/10.1177/0003489416636131>. PMID:26951063, 2016.

BARSTIES V, LATOSZEK B, MARYN Y, GERRITS E, DE BODT M. The Acoustic Breathiness Index (ABI): a multivariate acoustic model for breathiness. *J Voice*. 31:511–e511, 2017.

BEHLAU, MARA; ROCHA, BRUNA; ENGLERT, MARINA; MADAZIO, GLAUCYA. Validation of the Brazilian Portuguese CAPE-V Instrument-Br CAPE-V for Auditory-Perceptual Analysis. *Journal of Voice*, S0892199720302575–. doi:10.1016/j.jvoice.2020.07.007, 2020.

BEHLAU M, MADAZIO G, OLIVEIRA G. Functional dysphonia: strategies to improve patient outcomes. *Patient Relat Outcome Meas*. 6:243- 53. <http://dx.doi.org/10.2147/PROM.S68631>. PMID:26664248, 2015.

BEHLAU M.; MADAZIO G.; FEIJÓ D.; PONTES P. Avaliação de voz. In: BEHLAU, M. *Voz: o livro do especialista*, 1. Rio de Janeiro: Revinter, cap. 3, p. 85-180, 2001.

BRINCA, L. F., BATISTA, A. P. F., TAVARES, A. I., GONÇALVES, I. C., & MORENO, M. L. Use of Cepstral Analyses for Differentiating Normal From Dysphonic Voices: A Comparative Study of Connected Speech Versus Sustained Vowel in European Portuguese Female Speakers. *Journal of Voice*, 28(3), 282–312, 2014.

BROCKMANN-BAUSER M, DRINNAN MJ. Routine acoustic voice analysis: time to think again? *Curr Opin Otolaryngol Head Neck Surg*. 19(3):165-70, 2011.

CARRASCO, E. R., OLIVEIRA, G., & BEHLAU, M. Análise perceptivo-auditiva e acústica da voz de indivíduos gagos. *Revista CEFAC*, 12(6), 925–935. doi:10.1590/s1516-18462010005000058, 2010.

CASMERIDES MCB, COSTA HO. Laboratório computadorizado de voz: caracterização de um grupo de usuários. In: Ferreira LP, Costa HO. *Voz ativa falando sobre a clínica fonoaudiológica*. Rocca; 263-79, 2011.

CORRAR, L. J.; PAULO, E.; DIAS FILHO, J. M. Análise multivariada: para os cursos de administração, ciências contábeis e economia. São Paulo: Atlas, 2007.

CHRISTMANN MK, BRANCALIONI AR, FREITAS CR, VARGAS DZ, KESKE-SOARES M, MEZZOMO CL, MOTA HB. Uso do programa MDVP em diferentes contextos: revisão de literatura. *Rev CEFAC*. 17(4):1341-9, 2015.

DEJONCKERE PH, WIENEKE GH. Cepstral of normal and pathological voices: Correlation with acoustic, aerodynamic and perceptual data. In M. J. Ball, & M. Duckworth (Eds.), *Advances in Clinical Phonetics*. (pp. 217–226), 1996.

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 Phoniatics of the European Laryngological Society (ELS). *Eur Arch Otorhinolaryngol*. 258(2):77-82, 2001.

DELGADO-HERNÁNDEZ, J., LEÓN-GÓMEZ, N. M., IZQUIERDO-ARTEAGA, L. M., & LLANOS-FUMERO, Y. Cepstral Analysis of Normal and Pathological Voice in Spanish Adults. Smoothed Cepstral Peak Prominence in Sustained Vowels Versus Connected Speech. *Acta Otorrinolaringologica (English Edition)*, 69(3), 134–140, 2018.

EADIE TL, DOYLE PC. Classification of dysphonic voice: acoustic and auditory-perceptual measures. *J Voice*. 19(1):1-14. <http://dx.doi.org/10.1016/j.jvoice.2004.02.002>. PMID:15766846, 2005.

EADIE, T. L., & BAYLOR, C. R. The Effect of Perceptual Training on Inexperienced Listeners' Judgments of Dysphonic Voice. *Journal of Voice*, 20(4), 527–544. doi:10.1016/j.jvoice.2005.08.007, 2006.

ENGLERT M, LIMA L, BEHLAU M. Acoustic Voice Quality Index and Acoustic Breathiness Index: Analysis With Different Speech Material in the Brazilian Portuguese. *J Voice*, 34(2), 163–304. doi: <https://doi.org/10.1016/j.jvoice.2019.03.015>, 2019.

FÁVERO, L. P.; BELFIORE, P.; SILVA, F. L.; CHAN, B. Análise de Dados: Modelagem Multivariada para Tomada de Decisões. CAMPUS, 2009.

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*. 24(1):47- 56. <http://dx.doi.org/10.1016/j.jvoice.2008.04.006>. PMID:19135854, 2010.

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*. 75(5):680-4, 2009.

HAIR, JR, J.F., BLACK, W.C., BABIN, B.J., ANDERSON, R.E., TATHAM, R.L. Análise multivariada de dados. Porto Alegre: Bookman. 6ª edição, 2009.

HALBERSTAM, B. Acoustic and Perceptual Parameters Relating to Connected Speech Are More Reliable Measures of Hoarseness than Parameters Relating to Sustained Vowels. *ORL*, 66(2), 70–73. doi:10.1159/000077798, 2004.

HEMAN-ACKAH, Y. D., MICHAEL, D. D., & GODING, G. S. The Relationship Between Cepstral Peak Prominence and Selected Parameters of Dysphonia. *Journal of Voice*, 16(1), 20–27, 2002.

HEMAN-ACKAH, Y. D., MICHAEL, D. D., BAROODY, M. M., OSTROWSKI, R., HILLENBRAND, J., HEUER, R. J., ... SATALOFF, R. T. Cepstral Peak Prominence: A More Reliable Measure of Dysphonia. *Annals of Otology, Rhinology & Laryngology*, 112(4), 324–333, 2003.

HILLENBRAND, J., CLEVELAND, R. A., & ERICKSON, R. L. Acoustic Correlates of Breathly Vocal Quality. *Journal of Speech Language and Hearing Research*, 37(4), 769, 1994.

HIRANO M. Objective evaluation of the human voice: Clinical aspects. *Folia Phoniatica*, 41, 89–144, 1989.

HOFFMANN, R. *Análise estatística de relações lineares e não-lineares*. LIVROPRONTO, 2012.

HOSMER, D. W. & LEMESHOW, S. *Applied Logistic Regression*. Edition Wiley Series in Probability and Statistics texts and references section (2<sup>a</sup> Edition), ISSN 0094306, doi: 10.2307/2074954, 1-397, 2000.

KEMPSTER, G. B.; GERRATT, B. R.; VERDOLINI-ABBOTT, K, BARKMEIER-KRAEMER, J.; HILLMAN, R. E. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol*. v. 18, p. 124- 132, 2009.

LANDIS JR, KOCH GG. The measurement of observer agreement for categorical data. *Biometrics*. 33:159-74, 1977.

LATOSZEK, B. B. V. et al. The Acoustic Breathiness Index (ABI): a multivariate acoustic model for breathiness. *Journal of Voice*, Elsevier, v. 31, n. 4, p. 511–e11, 2017.

LATOSZEK, B. B. V. et al. A meta-analysis: Acoustic measurement of roughness and breathiness. *Journal of Speech, Language, and Hearing Research*, ASHA, v. 61, n. 2, p. 298–323, 2018.

LAVÉRE J. The concept of articulatory settings: an historical survey. *Historiogr Linguist* 5:1-14, 1978.

LAVÉRE J. *The phonetic description of voice quality*. Cambridge: Cambridge University Press; 1980.

LAW, T., KIM, JH, LEE, KY, TANG, EC, LAM, JH, VAN, HASSELT AC, et al. Comparison of rater's reliability on perceptual evaluation of different types of voice sample. *J Voice*. 26(5):666.e13-21, 2012.

LIU, H., LI, G., CUMBERLAND WG, WU, T. Testing statistical. significance of the area under a receiving operating. Characteristics curve for repeated measures design with. bootstrapping. *J Data Science*. 3:257, 2005.

LOPES LW, ALVES GAS, MELO LM. Evidência de conteúdo de um protocolo de análise espectrográfica. *Rev CEFAC*. 19(4):510-28, 2017.

LOPES LW, SILVA JD, SIMÕES LB, EVANGELISTA DS, SILVA POC, ALMEIDA AA,

LIMA-SILVA MFB. Relationship between acoustic measurements and self-evaluation in patients with voice disorders. *J Voice*. 31(1):119. e1-10. <http://dx.doi.org/10.1016/j.jvoice.2016.02.021>. PMID:27049448, 2018.

LOPES LW, ALVES GAS, MELO ML. Content evidence of a spectrographic analysis protocol. *Rev CEFAC*. 19(4):510-28. <http://dx.doi.org/10.1590/1982-021620171942917>, 2017.

LOPES, LW et al . Desempenho do diagrama de desvio fonatório na avaliação de vozes sintetizadas rugosas e soproas,. *Braz. j. otorhinolaryngol.*, São Paulo , v. 84, n. 4, p. 460-472, Ago. 2018 .

LOPES, LW, SILVESTRE, ESS, FRANÇA, ACS, MARINHO IS, AVELINO, MAP, VIEIRA, VJD, ALMEIDA, AA. Medidas cepstrais na avaliação da intensidade do desvio vocal. *CoDAS* vol.31 no.4 São Paulo Epub Aug 15, 2019.

LOPES, LW et al. Classificação espectrográfica do sinal vocal: relação com o diagnóstico laríngeo e a análise perceptivo-auditiva. *Audiol., Commun. Res.*, São Paulo , v. 25, e2194, 2020.

LOWELL, S. Y., COLTON, R. H., KELLEY, R. T., & MIZIA, S. A. Predictive Value and Discriminant Capacity of Cepstral- and Spectral-Based Measures During Continuous Speech. *Journal of Voice*, 27(4), 393–400, 2013.

LOWELL, S. Y., KELLEY, R. T., AWAN, S. N., COLTON, R. H., & CHAN, N. H. Spectral- and Cepstral-Based Acoustic Features of Dysphonic, Strained Voice Quality. *Annals of Otology, Rhinology & Laryngology*, 121(8), 539–548. doi:10.1177/000348941212100808, 2012.

LU FL, MATTESON S. Speech tasks and inter-rater reliability in perceptual voice evaluation. *J Voice*. 28(6):725-32, 2014.

MADAZIO G, LEÃO S, BEHLAU M. The phonatory deviation diagram: a novel objective measurement of vocal function. *Folia Phoniatr Logop*. 63:305-11, 2011.

MARTINEZ, EZ, LOUZADA-NETO F, PEREIRA, BB. A curva ROC para testes diagnósticos. *Cad Saude Colet*. 11(1):7-31, 2003.

MARYN, Y., CORTHALS, P., VAN CAUWENBERGE, P. et al. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice* 24:540–555, 2010.

MARYN, Y.; WEENINK, D. Objective Dysphonia Measures in the Program Praat: Smoothed Cepstral Peak Prominence and Acoustic Voice Quality Index. *Journal of Voice*, 29(1), 35–43, 2014.

MORAES, R. M., & MACHADO, L. S. Psychomotor skills assessment in medical training based on virtual reality using a Weighted Possibilistic approach. *Knowledge-Based Systems*, 70, 97–102, 2014.

MORAES, R. M., ALVES, R. S. S. Modelos de Decisão aplicados à Saúde: teoria e prática. *Tempus, actas de saúde colet*, Brasília, 10(2), 07-10, jun, 2016.

MURTON, O., HILLMAN, R., MEHTA, D. Cepstral Peak Prominence Values for Clinical Voice Evaluation. *American Journal of Speech-Language Pathology* Vol. 29 1596–1607, 2020.

NORDEMBERG, M.; SUNDBERG, J. Effect on LTAS of vocal loudness variation. *TMH-QPSR, KTH*, v. 45, p. 87-91, 2003.

OMORI K, SLAVIT DH, KACKER A, BLAUGRUND SM. Influence of size and etiology of glottal gap in glottic incompetence dysphonia. *Laryngoscope*. 108(4 Pt 1):514-8, 1998.

PATEL, R. R., AWAN, S. N., BARKMEIER-KRAEMER, J., COUREY, M., DELIYSKI, D., EADIE, T., ... HILLMAN, R. Recommended Protocols for Instrumental Assessment of Voice: American Speech-Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function. *American Journal of Speech-Language Pathology*, 27(3), 887, 2018.

PAULA, G. A. Modelos de Regressão: com apoio computacional. São Paulo: IME-USP São Paulo, 2010.

PETERSON, E. A., ROY, N., AWAN, S. N., MERRILL, R. M., BANKS, R., & TANNER, K. Toward Validation of the Cepstral Spectral Index of Dysphonia (CSID) as an Objective Treatment Outcomes Measure. *Journal of Voice*, 27(4), 401–410, 2013.

RADISH KUMAR, B., BHAT, J. S., & PRASAD, N. Cepstral Analysis of Voice in Persons With Vocal Nodules. *Journal of Voice*, 24(6), 651–653. doi:10.1016/j.jvoice.2009.07.008, 2010.

ROY, N., BARKMEIER-KRAEMER, J., EADIE, T., SIVASANKAR, M. P., MEHTA, D., PAUL, D., & HILLMAN, R. Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*, 22, 212–226., 2013.

S. L. DO N. C. COSTA, Análise Acústica, Baseada no Modelo Linear de Produção da Fala, para Discriminação de Vozes Patológicas. 161 f. Tese de Doutorado, Universidade Federal de Campina Grande, PB, 2008.

SHIM HJ, et al. Voice Quality of Muscle Tension Dysphonia. *Clinical Archives of Communication Disorders*. Vol. 1, No. 1: 42-47, 2016.

SILVA, C. A. M. Exploração de métodos de seleção de variáveis pela técnica de regressão logística para análise de dados epidemiológicos. Dissertação de Mestrado. Campinas, SP: Universidade Estadual de Campinas–Programa de Pós Graduação em Ciências Médicas, 2006.

SOUZA BO, GAMA ACC. Apoio visual do traçado espectrográfico: impacto na confiabilidade

da análise perceptivo-auditiva da voz por avaliadores inexperientes. *Disturb. Comun.* 27(3):479-86, 2015.

SPIEGEL, M. R., SCHILLER, J., SRINIVASAN, A. *Probabilidade e estatística*. Porto Alegre, Bookman 3ª edição, 2013.

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.* 29(1):26-34, 2014.

WATTS, C. R. The Effect of CAPE-V Sentences on Cepstral/Spectral Acoustic Measures in Dysphonic Speakers. *Folia Phoniatica et Logopaedica*, 67(1), 15–20. doi:10.1159/000371656, 2015.

WATTS, C. R., AWAN, S. N., & MARYN, Y. A Comparison of Cepstral Peak Prominence Measures From Two Acoustic Analysis Programs. *Journal of Voice*, 31(3), 387.e1–387.e10. doi:10.1016/j.jvoice.2016.09.012, 2017.

WUYTS, FL., DE BODT, MS., MOLENBERGHS, G., et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *J Speech, Lang Hear Res.* 43:796–809, 2000.

ZWETSCH, I. C., RIBEIRO, R. D., FAGUNDES, T. R., SCOLARI, D. *Processamento Digital de Sinais no Diagnóstico Diferencial de Doenças Laríngeas Benignas*. Scientia Medica, Porto Alegre: PUCRS, Vol. 16, n. 3, jul./set. 2006.

## ANEXO A

### *Scripts para extração das medidas acústicas: Declínio espectral, CPP, CPPS*

#### - **DECLÍNIO ESPECTRAL**

```
clearinfo
obj_audio=Read from file: "Escreva o diretório do arquivo . wav"
select obj_audio
obj_ltas=To Ltas: 100
select obj_ltas
declinio=Get slope: 0, 1250, 1250, 4000, "energy"
printline Declínio 'declinio'
select obj_audio
plus obj_ltas
Remove
```

#### - **CPP**

```
clearinfo
obj_audio=Read from file: "Escreva o diretório do arquivo . wav"
select obj_audio
obj_cepstogram=To PowerCepstrogram: 60, 0.002, 5000, 50
select obj_cepstogram
obj_powercepstrum=To PowerCepstrum (slice): 0.1
select obj_powercepstrum
cpp=Get peak prominence: 60, 333.3, "Cubic", 0.001, 0, "Straight", "Robust"
printline CPP 'cpp'
select obj_audio
plus obj_cepstogram
plus obj_powercepstrum
Remove
```

#### - **CPPS**

```
clearinfo
obj_audio=Read from file: "Escreva o diretório do arquivo . wav"
select obj_audio
obj_cepstogram=To PowerCepstrogram: 60, 0.002, 5000, 50
select obj_cepstogram
cpps=Get CPPS: "yes", 0.01, 0.001, 60, 330, 0.05, "Parabolic", 0.001, 0, "Straight",
printline CPPS 'cpps'
select obj_audio
plus obj_cepstogram
Remove
```

## ANEXO B

## ESCALA ANALÓGICA VISUAL

## AVALIAÇÃO PERCEPTIVO-AUDITIVA POR CONSENSO

Pesquisa: \_\_\_\_\_

Número do sujeito: \_\_\_\_\_. Amostra: ( )Vg "é" | ( )Vg "a" | ( )Contagem | ( )Frases | ( )FSE

Predomínio: ( )VNQV | ( )Rugosa | ( )Soprosa | ( )Tensa | ( )Instável | ( )Outro: \_\_\_\_\_

**Qualidade vocal (EAV)**

Grau geral	_____	C	I	_____	/100 EN=
Rugosidade	_____	C	I	_____	/100 EN=
Soprosidade	_____	C	I	_____	/100 EN=
Tensão	_____	C	I	_____	/100 EN=
Instabilidade	_____	C	I	_____	/100 EN=
_____	_____	C	I	_____	/100 EN=
_____	_____	C	I	_____	/100 EN=

**Natureza da alteração:** ( ) Fonte | ( ) Filtro | ( ) Fonte e filtro

## ANEXO C

## PARECER CONSUBSTANCIADO DO CEP

UFPB - CENTRO DE CIÊNCIAS  
DA SAÚDE DA UNIVERSIDADE  
FEDERAL DA PARAÍBA



## PARECER CONSUBSTANCIADO DO CEP

## DADOS DA EMENDA

**Título da Pesquisa:** Acurácia das medidas de análise acústica linear e não linear na avaliação dos distúrbios da voz

**Pesquisador:** Leonardo Wanderley Lopes

**Área Temática:**

**Versão:** 4

**CAAE:** 19237213.2.0000.5188

**Instituição Proponente:** Universidade Federal da Paraíba

**Patrocinador Principal:** Financiamento Próprio

## DADOS DO PARECER

**Número do Parecer:** 2.677.777

**Apresentação do Projeto:**

Trata-se da apresentação de Emenda versão 4 ao projeto de pesquisa vinculado ao curso de Fonoaudiologia do Centro de Ciências da Saúde da Universidade Federal da Paraíba, sob orientação do Prof. Leonardo Wanderley Lopes

**Justificativa da Emenda:**

Continuidade da pesquisa durante Projeto PIBIC/PIVIC 2018-2019, ampliando os dados da pesquisa.

**Objetivo da Pesquisa:****Objetivo Primário:**

Analisar a acurácia das medidas de análise acústica linear e não linear na avaliação dos distúrbios da voz.

**Objetivo Secundário:**

Verificar a acurácia das medidas acústicas tradicionais e não lineares combinadas e isoladas, na identificação de vozes saudáveis e alteradas;

Analisar o poder discriminatório das medidas acústicas tradicionais e não lineares combinadas e isoladas, na classificação de vozes com diferentes graus de alteração;

Avaliar o poder discriminatório das medidas acústicas tradicionais e não lineares combinadas e isoladas, na categorização do tipo de voz predominante;

Endereço: UNIVERSITARIO S/N  
Bairro: CASTELO BRANCO CEP: 58.051-900  
UF: PB Município: JOÃO PESSOA  
Telefone: (83)3216-7791 Fax: (83)3216-7791 E-mail: comitecetica@ccs.ufpb.br