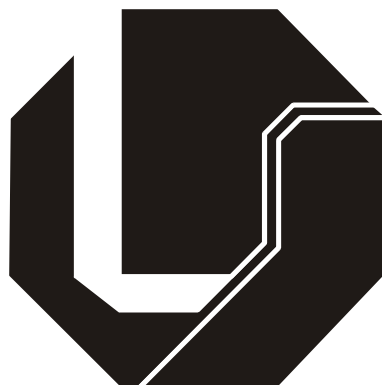


UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA



**Estudo da Separação entre Voz Patológica e
Normal por Meio da Avaliação da Energia
Global do Sinal de Voz**

Marlice Fernandes de Oliveira

JULHO
2007

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Estudo da Separação entre Voz Patológica e Normal por Meio da Avaliação da Energia Global do Sinal de Voz

Marlice Fernandes de Oliveira

Texto da dissertação apresentada à Universidade Federal de Uberlândia, perante a banca de examinadores abaixo, como parte dos requisitos necessários para a obtenção do título de Mestre em Ciências.

Banca Examinadora:

Prof. Adriano de Oliveira Andrade, PhD - Orientador (UFU)

Prof. Luciano Vieira Lima, Dr.

Prof. Adriano Alves Pereira, Dr.

Profa. Maria Inês Rebelo Gonçalves, PhD

Estudo da Separação entre Voz Patológica e Normal por Meio da Avaliação da Energia Global do Sinal de Voz

Marlice Fernandes de Oliveira

Texto da dissertação apresentada à Universidade Federal de Uberlândia, como parte dos requisitos para a obtenção do título de Mestre em Ciências.

Prof. Adriano de Oliveira Andrade, PhD
Orientador

Prof. Darizon Alves de Andrade, PhD
Coordenador do curso de Pós-Graduação

**“O valor das coisas não está no tempo que elas duram,
mas na intensidade com que acontecem.
Por isso, existem momentos inesquecíveis,
coisas inexplicáveis e pessoas incomparáveis.”
(Fernando Pessoa)**

Agradecimentos

Agradeço primeiro a Deus pelo dom da vida e por todas as oportunidades que Ele me concedeu: o lar maravilhoso no qual Ele me colocou, as pessoas boas que Ele pôs no meu caminho, e principalmente, por todos os ensinamentos que Ele deixou para que eu pudesse me guiar.

Agradeço ao meu pai, Marcílio, exemplo de esforço, trabalho, honestidade, humildade, amor e dedicação acima de tudo.

Agradeço a minha mãe, Clarice, que me acompanhou até o meio desta caminhada e que tanto me orientou e consolou nos momentos de desânimo e cansaço. Você foi meu descanso nos momentos mais difíceis. Como eu sinto sua falta, mãezinha.

Agradeço aos meus filhos, Pedro Henrique, Gabriel e Renato, que souberam ser sempre compreensivos nas minhas ausências, nos meus momentos de omissão, ou de irritação. Vocês são a coisa mais importante da minha vida.

Agradeço ao Luciano, tão amigo, que sempre fez o papel de protetor, de mediador nos momentos difíceis, orientador em tudo que eu precisei e acima de tudo, sempre disposto a me ajudar a transpor qualquer obstáculo que porventura aparecesse no meu caminho. Obrigada por tudo, Luciano.

Agradeço aos meus irmãos, Viviana, Eliane, Sandra e Sérgio, por este laço tão forte que nos une e que me proporcionou tanta segurança em todos os momentos em que eu precisei.

E também a minha nora Raquel e aos meus sobrinhos e sobrinhas, cada um por um motivo especial, mas todos participaram desta caminhada.

Agradeço ao João, meu amigo inseparável, meu mestre tantas vezes, e por que não dizer meu SOS, sempre pronto a me socorrer sempre que eu o solicitei. João, você foi muito importante para que eu chegasse até aqui, obrigada.

Agradeço a Kheline, amiga e companheira nesta trajetória, com quem eu dividi minhas dúvidas, minhas dificuldades e com quem eu sempre pude contar. Como foi bom ter você junto comigo, Khel.

Ao Fernando, meu parceiro neste trabalho, eu tenho que agradecer imensamente. Esse trabalho só pôde ser finalizado contando com o seu esforço, dedicação e competência. Eu vou te dever sempre.

Ao Adriano, meu orientador, eu agradeço de coração, por ter sido uma luz que me direcionou e que possibilitou que eu concluísse esse trabalho. Eu agradeço muito a Deus por ter colocado em meu caminho alguém tão centrado, firme e competente, que não permitiu que eu vacilasse mais: um verdadeiro orientador.

Agradeço também aos colegas de disciplina Alexandre e Marlene, com quem também eu pude contar.

Agradeço enfim, a todos os companheiros do laboratório que sempre demonstraram interesse em me ajudar.

Agradeço à Marly pela atenção e carinho sempre que eu precisei.

Resumo

OLIVEIRA, Marlice Fernandes de. **Estudo da separação entre voz patológica e normal por meio da avaliação da energia global do sinal de voz.** 2007. Dissertação (Mestrado em Ciências) – Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, 2007.

A análise do sinal de voz é uma ferramenta importante no diagnóstico dos distúrbios laríngeos. Dentre as diversas técnicas para o processamento da voz destaca-se o espectrograma por permitir uma visualização da variação da energia do sinal em função do tempo e frequência. Neste contexto, esta pesquisa investiga a energia global do sinal de voz, estimada a partir do espectrograma, como ferramenta capaz de discriminar esses sinais, obtidos de pacientes com diferentes doenças, daqueles coletados de sujeitos saudáveis. O estudo ainda verifica a possibilidade do uso da energia global na discriminação de distúrbios laríngeos. No total 94 indivíduos participaram desse estudo, sendo 46 disfônicos e 48 eufônicos. Inicialmente os sujeitos foram submetidos a um exame de videolaringoscopia para a determinação do diagnóstico. Posteriormente eles foram submetidos a um exame clínico de análise acústica vocal por meio da gravação da vogal sustentada /ε/. Utilizando a energia global foi possível discriminar as vozes normais das vozes disfônicas. Por meio da energia global foi possível ainda separar as vozes de pacientes portadores de paralisia da prega vocal esquerda de todas as outras doenças laríngeas investigadas. Podemos sugerir a energia global da voz como uma ferramenta auxiliar no diagnóstico diferencial entre vozes normais e disfônicas.

Palavras-chave: Distúrbios Vocais e Laríngeos; Análise Acústica; Espectrograma; Energia Global.

Abstract

OLIVEIRA, Marlice Fernandes de. **Investigation of discrimination between healthy and pathological voice through the analysis of the global energy of the voice signal.** 2007. Dissertation (Máster of Sciences) – Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, 2007.

Voice analysis is an important tool in the diagnosis of laryngeal disorders. Among distinct signal processing techniques employed for voice analysis, the spectrogram is commonly used, as it allows for a visualization of the variation of the energy of the signal as a function of the both time and frequency. In this context, this study investigates the use of the global energy of the voice signal, estimated through the spectrogram, as a tool for discrimination between signals obtained from healthy and pathological subjects. This research has also exploited the potential use of the global energy of the voice signal to discriminate distinct laryngeal disorders. In total, 94 subjects were involved in this study, from which 46 were dysphonic and 48 normal. The diagnosis of laryngeal disorders was confirmed by means of a videolaryngoscopic examination. Participants were also subjected to a clinical examination of vocal acoustic through the recording of the sustained vowel /ε/. The global energy allowed for the discrimination between normal and dysphonic voice. Furthermore, this technique could discriminate the voice signal of patients suffering from left vocal fold paralysis from those suffering from other investigated disorders. The results suggest the global energy of the signal as an auxiliary and alternative tool for the diagnosis between normal and dysphonic voice.

Keywords: Vocal and Laryngeal Disorders; Acoustic Analysis; Spectrogram; Global Energy.

Lista de Figuras

Figura 1- Nasofibrosópio flexível	2
Figura 2- Telefibrosópio flexível	2
Figura 3 - Pregas Vocais em adução e abdução	20
Figura 4- Estruturas Laríngeas	21
Figura 5- Pregas Vocais	21
Figura 6- Inervação Laríngea	23
Figura 7- Ciclos Glóticos	25
Figura 8- Modelo Fonte-Filtro	27
Figura 9- Prega Vocal- Respiração	32
Figura 10- Prega Vocal- Fonação	32
Figura 11-Paralisia de Prega Vocal	33
Figura 12- Pólipo	33
Figura 13- Fenda Fusiforme	33
Figura 14- Granuloma	33
Figura 15- Fenda Triangular Médio-Posterior	33
Figura 16- Sulco	33
Figura 17- Papiloma	34
Figura 18- Edema de REINKE	34
Figura 19 - Interface do Programa Voian (Matlab)	36
Figura 20- Diagrama de Blocos	38

Figura 21-Janelamento	40
Figura 22 - Espectrograma a) Voz Normal b) Voz Disfônica	45
Figura 23- Espectrograma Médio a) Vozes Normais b) Vozes Disfônicas	47
Figura 24- Espectro de Potência	48
Figura 25- Energia Global das Vozes Normais e Disfônicas	49
Figura 26- Energia Global dos Distúrbios Laríngeos	50

Sumário

Capítulo 1

Diagnóstico dos Distúrbios Laríngeos	1
1.2 Revisão do Estado da Arte	3
1.2. Análise Acústica Vocal	3
1.2.2 Processamento dos Sinais	8
1.2.2. Sinais	8
1.2.2.2 Histórico	9
1.3 Motivação	11
1.4 Justificativas	12
1.5 Objetivos	13
1.6 Descrição das Técnicas para Análise de Voz	13

Capítulo 2

Produção da Voz	19
2.1 Laringe	19
2.1.1 Estruturas da Laringe	21
2.1.2 Função da Musculatura Laríngea	22
2.1.3 Inervação da Laringe	22
2.1.4 Efeito de Bernoulli	24

2.1.5 Os Ciclos Glóticos	24
2.2 Trato Vocal	26
2.2.1 Modelo Fonte Filtro	27

Capítulo 3

Material e Método	30
3.1 Coleta de Dados	30
3.2 Estratégia para Análise de Dados	35
3.2.1 Diagrama de Blocos	37
3.2.2 Descrição da Normalização	38
3.2.3 Descrição do Janelamento	39
3.2.4 Periodograma	41
3.2.5 Espectrograma	42
3.2.6 Energia Global	42

Capítulo 4

Resultados	44
4.1 Análise Visual do Espectrograma	44
4.2 Análise do Espectrograma Médio	45
4.3 Análise do Periodograma	47

4.4 Análise da Energia Global	48
-------------------------------	----

Capítulo 5

Discussão	52
-----------	----

Capítulo 6

Conclusões	57
------------	----

Capítulo 7

Estudos Futuros	58
-----------------	----

Apêndice 1

Glossário	60
-----------	----

Apêndice 2

Termo de Consentimento Livre e Esclarecido	64
--	----

Anexo 1

Cálculo do Intervalo de Confiança	65
-----------------------------------	----

Referências

67

CAPÍTULO 1

DIAGNÓSTICO DOS DISTÚRBIOS LARÍNGEOS

A voz é produzida na laringe e expressa as condições físicas e emocionais de cada indivíduo. A sua produção normal se faz com o equilíbrio entre duas forças que atuam sobre as pregas vocais (a força aerodinâmica dos pulmões e a força mioelástica das pregas vocais) que produzem a abertura e o fechamento das mesmas de modo harmônico e sucessivo. Essa energia aerodinâmica gerada pelo fluxo expiratório é convertida em energia acústica pela vibração das pregas vocais [1].

Quando ocorrem alterações nas estruturas que compõem o aparelho fonador, a produção vocal se modifica e ocorrem disfonias ou distúrbios da voz. Essas alterações podem ser incapacitantes, porque além de prejudicarem a inteligibilidade da fala, provocam fadiga vocal, ardência, dor, sensação de tensão, estrangulamento, esforço, redução do tempo de fonação, extensão e projeção vocal, quebras de sonoridade e ainda falta de estabilidade [2].

O otorrinolaringologista diagnostica as desordens laríngicas e avalia as condições orgânicas da laringe e do trato vocal por meio do exame de videolaringoscopia. Neste procedimento utilizam-se aparelhos que podem ser introduzidos pela cavidade oral, o telefibróscópio rígido ou pela cavidade nasal, o nasofibróscópio flexível. Por meio de uma microcâmera acoplada na extremidade do aparelho, uma imagem dinâmica das pregas vocais, que pode ser gravada, é obtida. (Fig. 1 e 2).



Fig. 1: Nasofibrosópio flexível.



Fig. 2: Telefibrosópio rígido.

Outra forma de realizar o diagnóstico das doenças laríngeas é por meio de exames anátomo-patológicos quando há necessidade de remoção de parte do tecido para avaliação histológica em laboratórios de análises clínicas. Tais exames histopatológicos são contudo invasivos, razão pela qual não têm indicação de rotina, com utilização apenas em pós-operatórios e em casos específicos.

A realização do diagnóstico das doenças laríngeas depende de um conjunto de informações que devem ser cuidadosamente analisadas. A interpretação imprecisa dos dados fornecidos em qualquer dos tipos de avaliação pode gerar erro de diagnóstico e consequentemente desvio na conduta terapêutica prejudicial à evolução e cura do paciente.

A videolaringoscopia consiste no procedimento indicado à realização do diagnóstico das lesões das pregas vocais, entretanto não informa quanto à qualidade da voz e às medidas de perturbação e ruído vocal.

Várias formas de avaliação dos distúrbios vocais e laríngeos são utilizadas; porém, necessita-se realizar a somatória das informações para fornecer o máximo de subsídios compatíveis ao diagnóstico.

Para tanto o fonoaudiólogo ao realizar a análise da voz fornece informações, que associadas aos resultados da videolaringoscopia, contribuem para a constatação do diagnóstico dos distúrbios laríngeos.

1.2 REVISÃO DO ESTADO DA ARTE

1.2.1 ANÁLISE ACÚSTICA VOCAL

A voz tem sido objeto de estudo de vários profissionais, os quais têm utilizado envolvidos na avaliação da mesma diversas técnicas e ferramentas específicas em sua análise objetiva.

O levantamento de dados da literatura envolveu uma busca por publicações fundamentadas em métodos de análise acústica do sinal de voz e suas possibilidades de aplicação em discriminação entre vozes normais e disfônicas. Ainda nesta revisão realizou-se um breve histórico a respeito do processamento digital dos sinais.

Os estudos encontrados abordam diversos métodos e estratégias para a realização da análise acústica da voz, abaixo relacionados.

PARRAGA (2002) utilizou a Transformada WAVELET PACKET na análise e classificação de sinais de vozes patológicas e afirmou ser possível separar vozes normais de disfônicas com um erro de classificação da ordem de 23,07% para falsos positivos e de 14,58 % para falsos negativos [3].

HADDAD et al.(2006) avaliaram vozes disfônicas de sujeitos submetidos à cordectomia por meio da análise acústica vocal, e detectaram diferenças significativas na frequência fundamental do sinal de voz dos diversos tipos de pesquisados. Porém, não encontraram diferenças estatisticamente significativas nas medidas de ruído [4].

ANDRADE (2003) determinou os limiares de normalidade dos parâmetros acústicos da voz e validou algumas medidas de ruído como confiáveis na determinação dos mesmos. São eles: SFR-Suavidade Espectral do Resíduo-, SFF-Suavidade Espectral do Filtro-, e EX-Coefficiente de Excesso- [5].

ZITTA (2005) comparou a frequência fundamental e a medida de ruído GNE- *Glottal to Noise Excitation Ratio*- e encontrou diferenças significativas entre vozes normais e com nódulos vocais por meio da avaliação da frequência fundamental, contudo não as encontrou pela utilização do GNE [6].

NIETO (1996) avaliou a medida de ruído HNR-*Harmonic to Noise Ratio*- em sujeitos portadores de pólipos, nódulos e edema de REINKE, no pré e pós-operatório de microcirurgia de laringe, e encontrou diferenças significativas nos valores desta medida entre os dois respectivos grupos [7].

SANTOS (2005), em seu estudo com análise acústica afirma que a avaliação de diferentes parâmetros vocais como as medidas de perturbação ou as de ruído oferecem informações sobre importantes aspectos da função laríngea [8].

MOTA (2005) realizou a análise acústica em um sujeito portador de membrana laríngea com vasculodisgenesia por meio do espectrograma e percebeu presença de componente de ruído em altas e médias frequências [9].

FIGUEIREDO et al. (2003) avaliaram a medida de ruído PHR-Proporção Harmônico-Ruído- no sinal de voz de sujeitos fumantes e não fumantes e as contrastaram entre ambos os gêneros pelo relacionamento entre os componentes harmônico e de ruído da voz. Encontraram valores aumentados no grupo feminino em relação ao masculino entre os não fumantes, o que não ocorreu no grupo de fumantes [10].

SADER et al. (2004) utilizaram análise perceptiva e acústica vocais com o objetivo de comparar sinais de vozes de crianças sem queixas vocais e disfônicas. Consideraram mais fiéis os dados extraídos dos traçados espectrográficos do que os obtidos de medidas isoladas, mesmo ao utilizarem o mesmo programa de análise [11].

CORAZZA et al. (2004) avaliaram os parâmetros acústicos vocais, JITTER, SHIMMER, NNE-Energia do Ruído Glótico- e HNR-Proporção Harmônico-Ruído- em sujeitos adultos sem queixa vocal e perceberam alterações principalmente nas medidas de JITTER e SHIMMER [12].

VIEIRA et al. (2005) pesquisaram as alterações estruturais mínimas –AEMs- da cobertura das pregas vocais. Em tal estudo correlacionaram os diferentes tipos de fendas glóticas geradas pelas AEMs e os achados perceptivos e acústicos. Para a análise acústica considerou-se os harmônicos presentes no traçado espectrográfico, a quantidade de harmônicos e de ruído que ultrapassam os 3,0 kHz. Foi utilizada uma escala de ruído de 0 a 3, conforme a presença do sinal aperiódico no gráfico, em que o zero constitui-se o indicativo de ausência e o três de ruído excessivo. Com esta análise concluíram que os casos de fenda fusiforme ântero-posterior predominantes nos sujeitos portadores de sulco estria maior apresentaram alterações acentuadas nos parâmetros perceptivos e acústicos investigados [13].

MURPHY (2006) investigou o índice derivado do espectro de potência na discriminação de vozes normais e disfônicas e concluiu que essa medida possibilita apenas uma informação indireta da proporção harmônico-ruído do sinal glótico [14].

ESTELLA et al. (2006) utilizaram escalas perceptuais, medidas aerodinâmicas e análise acústica com o objetivo de discriminar padrões vocais normais de alterados. As estratégias de análise utilizadas foram: a escala GRBAS (Grau de Rouquidão,

Soprosidade, Astenia e Tensão), o TMF-Tempo Máximo de Fonação-, o Fonetograma e o Índice de Perturbação Vocal-JITTER-, cuja combinação possibilita níveis confiáveis de resultados [15].

JIANG et al. (2006) com o objetivo de estudar o sinal de voz de diversos distúrbios laríngeos como, por exemplo, pólipos, nódulos e paralisias, sugerem a Teoria do Caos e tecem comentários a respeito da aplicabilidade da mesma em diferentes sistemas físicos, como: turbulência, reações químicas, sistema solar, circuitos não lineares. Referenciam este método como parâmetro para o estudo da vibração das pregas vocais, porquanto apresentam-se falhas para tal finalidade JITTER e SHIMMER.

ZHANG et al. (2005) utilizaram as medidas de perturbação vocal-JITTER e SHIMMER- para discriminar vozes normais das de pacientes portadores de paralisia unilateral. Os resultados da pesquisa demonstram que a significância estatística é maior ao tratar-se de vozes de paralisia com sinais aperiódicos, ou seja, nos quadros mais severos [17].

SCALARRARA et al. (2005) utilizaram técnicas de processamento digital de sinais para discriminar vozes de sujeitos saudáveis das de portadores de nódulos vocais. No estudo realizado relacionaram os picos do espectro de potência do sinal e encontraram aumento na frequência fundamental nos casos de nódulos vocais [18]

CARDING et al. (2004) com o objetivo de discriminar vozes normais de disfônicas, avaliaram três medidas acústicas: JITTER, SHIMMER e HNR. Com base nos resultados, consideraram as medidas acústicas isoladas insuficientes para este tipo de discriminação e sugeriram cautela quanto a este tipo de procedimento. [19].

NEMR et al. (2005) realizaram um estudo com 29 indivíduos com queixas vocais. O objetivo foi relacionar os resultados da análise perceptivo-auditiva vocal aos

da análise acústica e ainda com as avaliações médicas, por meio da laringoscopia indireta e da videolaringoscopia. Concluíram que as análises perceptivo-auditivas e médicas são concordantes no diagnóstico de alterações vocais e/ou laríngeas, enquanto o percentual deconcordância entre as análises acústica e perceptivo-auditiva foi da ordem de 62% [20].

ARAÚJO (2002) objetivou em seu estudo a normatização das medidas acústicas da voz normal. Para isto avaliou a frequência fundamental, as medidas de perturbação vocal (*JITTER* e *SHIMMER*) e as medidas de ruído, tais como: *SNL-Spectral Noise Level*, *HNR-Harmonic-to-Noise Ratio*, *NNE-Normalized Noise Energy* e *BR-Breathiness Ratio*, pelo que encontraram grande variabilidade entre as vozes normais [21].

MASTER et al. (2006) avaliaram o sinal de voz com o objetivo de discriminar gênero, idade, vozes treinadas e disfônicas com o uso do *LTAS-Long-Term Average Spectrum*, método que mostra no eixo das abscissas o nível de pressão sonora em dB e no das ordenadas a frequência em Hz. Este tipo de análise utiliza amostras de fala de longa duração, ou seja, no intervalo entre 20 a 40 s, em amostragem suficientemente longa para que o espectro resultante não seja afetado por diferenças no material de fala. Nas amostras de longa duração, as frequências dos primeiros formantes, F1, F2 e F3, que sofrem variações de acordo com a vogal emitida, são, neste tipo de análise, representadas por uma média e os demais, F4 e F5, mais relacionados à qualidade da voz, mostram-se mais evidenciados. Outra peculiaridade consiste na necessidade da eliminação das pausas e os sons não vozeados, sem vibração laríngea, da amostra, por poderem mascarar a informação da fonte glótica. Isto deve ser feito, principalmente na análise de vozes disfônicas para que o ruído produzido na emissão dos sons não

vozeados, a exemplo de /s/, /f/, não seja confundido com o ruído aperiódico presente nas vozes disfônicas. No LTAS são avaliados: a indicação da inclinação da curva, a mensuração dos picos e a diferença entre a amplitude de F_0 e $F1$. Como resultado, verificou-se que este método não diagnóstica as alterações laríngicas, porém evidencia características espectrais distintas em alguns tipos de qualidade vocal [22].

Em suma, a supra revisão de literatura apontou diferentes estratégias utilizadas no processamento do sinal de voz com o objetivo de investigar os distúrbios laríngicos, com ênfase no processamento digital de sinais que tem sido atualmente utilizado em diferentes tipos de análises e aplicações em diversas áreas.

1.2.2 PROCESSAMENTO DOS SINAIS

1.2.2.1 SINAIS

Existem na natureza uma infinidade de sinais que podem ser medidos e analisados, como por exemplo, a temperatura, a luz ou o som. Os ouvidos convertem a emissão sonora em sinais elétricos que podem ser analisados em suas principais características: a amplitude, a frequência e a fase. Por meio da percepção tátil, sinais elétricos são transmitidos da pele ao cérebro que decodifica esta percepção como agradável ou desagradável e gera decisões a ela pertinentes. Da mesma forma, a visão também converte as imagens em sinais elétricos e os envia ao cérebro que os analisa em cor, forma, dentre outros.

Um sinal é definido como uma função de uma ou mais variáveis, que traz informações sobre a natureza de um fenômeno físico e a transporta a um receptor que vai analisá-la nos seus componentes principais. Sinais bioelétricos como o

eletrocardiograma e o eletroencefalograma têm sido muito estudados com o objetivo de extrair informações sobre os estados patológicos dos órgãos em questão.

Os sinais de voz codificam uma variedade de informações a respeito de diversos parâmetros distintos e são largamente utilizados em sistemas de telefonia, sinais de rádio, de televisão e, desde a década de 1980, também para fins clínicos.

Com a finalidade clínica, tal análise do sinal propicia dados referentes à diferentes realidades vocais, quer sejam profissionais ou de portadores de distúrbios. Deste modo auxilia na detecção precoce de distúrbios vocais e laríngeos. A análise do sinal com fins clínicos proporciona documentação com dados específicos respectivos que permitem ao profissional e ao paciente melhor entender os parâmetros vocais alterados bem como acompanhar o tratamento [6].

1.2.2.2 HISTÓRICO

A história do processamento digital de sinais teve início no final do século XIX, com as Transformadas de LAPLACE e FOURIER que tiveram grande aplicação para as Ciências, em especial, a Matemática, e a Engenharia. FOURIER, matemático francês, realizou a primeira análise espectral de um som em 1822. HELMOLTZ, em 1863, construiu caixas de ressonância que possibilitaram uma análise espectrográfica simples. Posteriormente foram desenvolvidos analisadores mecânicos. Porém, por serem processadores lentos, só possibilitaram a análise de curtas amostragens de fala [23].

A história da análise acústica iniciou-se de fato, por volta de 1920, com a criação do oscilograma, fornecedor do gráfico da amplitude por tempo, e avaliação do sinal contínuo. Mas o avanço significativo nesta área ocorreu durante a segunda guerra mundial, com o desenvolvimento do espectrógrafo de som. Este aparelho teve

implicação revolucionária por permitir registro tridimensional do sinal sonoro, com integração dos aspectos de tempo, frequência e intensidade num único gráfico de dois eixos. A sua utilização possibilitou a realização da análise espectrográfica de forma automática e com rapidez. O objetivo principal era o rastreamento das tropas militares, por meio do reconhecimento da voz. Após a guerra, o referido instrumento passou a ser usado nos laboratórios de fonética e análise vocal.

Em 1950, a partir da referida análise, tornou-se usual a utilização de filtros para analisar a onda complexa em seus componentes de frequência.

Em 1965, COOLEY e TUKEY criaram um algoritmo que reduziu consideravelmente o tempo de cálculo da FFT-*Fast Fourier Transform*-, enquanto na década de 1970, com a tecnologia digital ocorreu notória aceleração no processamento digital de sinais e na análise acústica vocal [23].

A partir de 1990, a avaliação da voz por meio da análise acústica foi introduzida no Brasil e acarretou a implementação de laboratórios de voz capacitados para tal finalidade com o intuito de obter dados mais confiáveis sobre o sinal de voz que possibilitassem a objetivação de análises perceptivo-auditivas, até então só explicáveis subjetivamente.

A introdução do processamento digital de sinais proporcionou confiabilidade e objetividade na obtenção de parâmetros acústicos complexos [24].

A análise de FOURIER consiste num processo matemático que divide a onda senoidal em um número finito de outras de mesma natureza.

Reitera-se que esta transformada é usada na engenharia de telecomunicações, radares, sonares, processamento de sinais, engenharia biomédica, simulações, síntese

musical, e outros. Por meio da utilização da FFT, é possível diminuir o número de elementos de processamento e também corrigir possíveis distorções.

A transformada de FOURIER é normalmente apresentada em sua forma contínua, porém para a sua implementação computacional, é necessário discretizá-la. Para tanto vale-se da taxa de amostragem que indica quantas vezes em um segundo se pode extrair amostras da voz, a qual é expressa em Hertz, ou seja, número de amostras por segundo. As mais comuns são: 5.500, 11.025, 22.050 e 44.100 Hz. A mesma necessita ainda ser dotada de uma frequência duas vezes mais alta do que o som que se pretende amostrar.

Nesta pesquisa, as amostras vocais foram gravadas numa taxa de amostragem de 44.100 Hz, pois quanto mais elevada melhor será a qualidade do som. Para o processamento dos sinais utilizou-se a STFT-*Short-Time Fourier Transform*.

Os vários passos empregados no processamento digital dos sinais de voz deste estudo são descritos no Capítulo 3.

1.3 MOTIVAÇÃO

A análise perceptivo-auditiva é muito utilizada para a avaliação da voz. Esse é um método que auxilia no diagnóstico dos distúrbios laríngeos quando realizado por profissionais experientes. Porém, uma das dificuldades por ele apresentada consiste na variabilidade da capacidade humana de julgamento do que se percebe auditivamente pela interferência de fatores desde julgamentos individuais a aspectos culturais, econômicos e sociais. Contudo, é inegável a contribuição da análise perceptivo-auditiva, que no entanto isoladamente não fornece resultados mensuráveis e precisos, pelo que optou-se pela análise acústica. Ressalte-se que apesar da mesma fornecer resultados

com maior objetividade quando comparada com a análise perceptivo-auditiva, deve ser minuciosamente monitorada para proporcionar efetiva confiabilidade [25].

Em vista disso, serviu de motivação para este estudo, a exploração da análise do sinal de voz, com o aprofundamento em técnicas de processamento digital de sinais e o propósito de levantar recursos possibilitadores da discriminação entre vozes normais e disfônicas, capazes de conferir mais precisão e objetividade à análise acústica vocal.

1.4 JUSTIFICATIVAS

A disfonia, que é a característica mais marcante das doenças laríngeas, provoca modificações no sinal de voz, as quais fornecem diferentes informações e têm sido objeto de estudo de vários pesquisadores. Diversas técnicas e múltiplas ferramentas e algoritmos matemáticos têm sido empregados como meta para melhor entender o comportamento desse sinal nas diferentes doenças da laringe. Neste contexto, justifica-se o interesse em realizar pesquisas laboratoriais com a utilização do sinal de voz, com a exploração de diversas ferramentas viabilizadoras de novos métodos de análise como recurso não invasivo e fornecedor de maiores subsídios ao diagnóstico dos distúrbios laríngeos.

Estudos como este, que envolvem análise de sinais, necessitam da contribuição de experiências de profissionais de várias áreas do conhecimento: médicos, fonoaudiólogos, engenheiros e matemáticos, uma vez que métodos isolados apenas trazem informações parciais sobre as desordens vocais.

A interdisciplinaridade profissional contribui cada vez mais para que as características relevantes de cada distúrbio vocal que se queira explorar dotem-se de mais adequado e amplo instrumental.

1.5 OBJETIVOS

O objetivo geral deste estudo consistiu na busca e desenvolvimento de métodos para a análise do sinal de voz.

Constituíram-se em objetivos específicos, de início, verificar se o espectro de potência constitui-se ferramenta capaz de discriminar vozes normais de vozes disfônicas. A seguir, avaliar as aplicações e limitações do espectrograma.

E como objetivo principal, optou-se por avaliar a energia global da voz para discriminação entre padrões normais e padrões alterados e ao término desta investigação verificar se a energia global da voz permite a separabilidade ou identificação de diversos distúrbios laríngeos.

1.6 DESCRIÇÃO DE TÉCNICAS PARA ANÁLISE DE VOZ

A avaliação da voz pode ser realizada por meio de uma análise perceptivo-auditiva pela extração dos diferentes graus de perturbação vocal de forma subjetiva, que conta para isto com o conhecimento prévio do especialista. Com esta finalidade as escalas GRBAS e RASAT [26], que avaliam diversos parâmetros da voz (e.g., rouquidão, aspereza, soprosidade, astenia e tensão), são comumente utilizadas. Cada parâmetro é avaliado por meio de uma escala com a graduação de 0 a 3, com o zero (0) indicativo de ausência, um (1) considerado grau leve, dois (2) grau moderado e três (3) grau extremo.

A escala RASAT é uma modificação da escala GRBAS que foi desenvolvida pelo Comitê para Testes de Função Fonatória da Sociedade Japonesa de Logopedia e

Foniatria-Committee for Phonatory Function Tests - Japan Society of Logopedics and Phoniatrics- em 1969 [26], [6]. A escala RASAT, em contraste com a GRBAS, leva em consideração a avaliação da aspereza da voz. Note que as mesmas relacionam a avaliação subjetiva da voz, em consideração ao que se percebe auditivamente e aos aspectos fisiológicos envolvidos na geração da voz.

Vários pesquisadores utilizam a escala GRBAS ou suas extensões para realizar a análise perceptivo-auditiva vocal [15], [27]. Porém, apesar destas escalas fornecerem resultados aceitáveis quando aplicadas por profissionais suficientemente especializados, necessitam de complementação por serem subjetivas [5].

A avaliação de voz por meio da análise perceptivo-auditiva tem caráter subjetivo e por esse motivo tornou-se necessário criar ferramentas mais objetivas que pudessem quantificar os distúrbios da voz. Atualmente, o fonoaudiólogo dispõe de uma diversidade de programas computacionais para a realização da análise acústica vocal. Por meio desses programas que utilizam processamento digital de sinais e de algoritmos específicos, obtém-se vários parâmetros que permitem descrever diversas características do sinal de voz [1]. Porém, geralmente, esses programas não descrevem, com clareza, os algoritmos utilizados no processamento do referido sinal.

Em presença de distúrbios vocais, observa-se que o sinal de voz sofre modificações da periodicidade e amplitude dos ciclos glóticos e aumento da componente de ruído. Essas modificações ocorrem devido às alterações nos padrões de vibração da mucosa das pregas vocais [16] em caso de lesões de massa, retrações cicatriciais, decorticação, ou remoção parcial ou total das pregas vocais na presença de câncer de laringe.

Qualquer tipo de lesão ou alteração nas pregas vocais provoca modificações na qualidade da voz. É por esse motivo que algumas pesquisas ressaltam a importância de se utilizar a análise acústica vocal como técnica não invasiva capaz de fornecer suporte ao diagnóstico das disfunções laríngeas [4], [17], [28], [29]. [19], [18], [30].

Na literatura encontram-se vários métodos para avaliar a voz de forma objetiva. Um deles é a medida dos índices de perturbação vocal, pelo cálculo dos valores de JITTER e SHIMMER. Chama-se de JITTER o quanto um ciclo glótico se diferencia em frequência do seu sucessor ou do seu antecessor, e de SHIMMER, a mesma variação, porém, em torno da amplitude. Para a extração do JITTER encontra-se o PFF-Fator de Perturbação de Frequência-, o DPF-Fator de Perturbação Direcional-, o RAP-Perturbação Média Relativa- e o JR-Razão de Jitter-.

O PFF representa a frequência relativa das perturbações com períodos maiores que 0,5 ms. Alguns estudos indicam que vozes disfônicas têm PFF maior do que aquele encontrado em vozes normais [31]. O DPF soma o número de vezes que a diferença entre períodos sucessivos muda de sinal algébrico e posteriormente divide este valor pelo número total de períodos [32]. O RAP consiste também numa medida de perturbação em torno da frequência e calcula a diferença entre os períodos pelo uso do período atual e uma média entre o anterior, o atual e o posterior. Usa-se, portanto, uma janela de três períodos. Outros pesquisadores usaram a mesma técnica com a utilização de uma janela de cinco períodos [33]. O JR fornece a relação entre a média das perturbações pelo período médio e apresenta semelhança com o parâmetro PFF [34].

Para a extração do SHIMMER encontra-se o APQ-Quociente de Perturbação de Amplitude- e o AVI-Índice de Variabilidade de Amplitude- [35].

O APQ foi uma adaptação do RAP para o SHIMMER com cálculos similares aos utilizados para o JITTER. O AVI que avalia a perturbação da amplitude foi definido por DEAL [35].

As medidas de perturbação da frequência- JITTER- e de perturbação da amplitude- SHIMMER- são muito utilizadas na análise acústica vocal. Porém, estudos recentes demonstram a ineficácia destes métodos de análise para vozes disfônicas severas. Diversas pesquisas atuais revelam que a laringe com alterações pela presença de lesões de massa ou alterações neuromusculares produz ciclos erráticos e por esse motivo, essas medidas têm sido sugeridas apenas em vozes normais e disfonias leves [6]. Portanto, nota-se a relevância de estudos com a utilização de outros métodos que possam ser aplicados a sinais de vozes disfônicas em graus extremos.

Encontrou-se também várias formas para quantificar as medidas do ruído provocado pela turbulência do ar através da glote, como: HNR- *Harmonic-to Noise Ratio*-; NNE ou ERG- *Normalized Noise Energy*-, PSR- *Signal to Noise Ratio*-; SNL- *Spectral Noise Level*-; e BR- *Breathiness Ratio*- [21]. Essas medidas de ruído fornecem dados importantes a respeito das características do sinal de voz, por verificar a proporção e a relação entre elas. A análise do componente aperiódico do sinal sonoro, ruído, tem sido apontada como excelente recurso na avaliação das disfonias [6].

A HNR expressa em decibel, contrasta a energia do componente harmônico ou componente regular do sinal com a energia do ruído presente ou componente irregular do sinal. Alguns autores comentam que o componente aperiódico da voz é diretamente proporcional ao grau de disfonia. [36].

A NNE- *Neutralized Noise Ratio*-, em decibel, mede o componente de ruído do sinal por meio da subtração do componente harmônico da energia total do sinal

acústico. Alguns pesquisadores consideram a NNE como uma medida mais sensível à diferenciação entre vozes eufônicas e disfônicas do que a HNR, principalmente quando vozes disfônicas severas integram a análise [37].

A PSR contrasta a energia total do sinal com o componente de ruído.

A medida de ruído SNL está relacionada, segundo alguns autores, ao aspecto perceptual de aspereza vocal [38].

A BR estima características perceptuais de soproidade vocal [21].

Para a avaliação do sinal de voz é comum utilizar a emissão sustentada das vogais /a/, /ε/ e /i/, por tempo de fonação de 3 s e remoção dos momentos instáveis iniciais e finais da emissão. O sinal de voz referente à emissão de um único fonema pode ser definido como quase-periódico, não estacionário e que possui características estatísticas variantes no tempo. Portanto, em processamento digital de sinais de voz, torna-se interessante o uso de técnicas que realcem o comportamento de determinada característica do sinal ao longo do tempo [39].

As principais ferramentas utilizadas no processamento do sinal de voz são a análise dos parâmetros temporais da forma de onda, como a variação da amplitude do sinal; o espectro de potência referente a distribuição de energia do sinal em função da frequência e o espectrograma que estima a variação local da energia do sinal em função do tempo e da frequência [40].

O espectrograma constitui-se num método de larga utilização para análise de voz, auxiliar no diagnóstico dos distúrbios laríngeos, por proporcionar uma visualização da energia contida no sinal de voz, por permitir a separação da contribuição de cada frequência [1].

Pela análise do espectrograma pode-se visualizar a variação da energia do sinal por meio de uma representação monocromática ou de escala de cores. Porém, esta análise é ainda pouco objetiva, por ser dependente de uma avaliação subjetiva do especialista.

Vários profissionais, por meio da extração das características do sinal se empenham em pesquisar diferenças nele presentes que possibilitem relação direta entre o sinal de voz e a fisiologia vocal, e, por conseguinte estabelecer relações entre o que pode ser considerado um padrão normal e alterado.

CAPÍTULO 2

PRODUÇÃO DA VOZ

A voz, do ponto de vista físico, é o som produzido pela vibração das pregas vocais e modificado nas cavidades de ressonância.

A produção da voz depende da associação de órgãos de dois sistemas do corpo humano, o respiratório e o digestório. Este conjunto de órgãos não existe como uma unidade física e devido a isso não existe um aparelho fonador, apesar de ser uma denominação normalmente utilizada. A voz é produzida por meio da utilização desses dois sistemas que devem funcionar em perfeita harmonia.

A produção da voz depende fundamentalmente da laringe que é um órgão esquelético-membranoso e que tem como limite inferior a traquéia e superior a faringe.

2.1 LARINGE

A laringe é um órgão fibromuscular que se assemelha a um arcabouço tubular constituído de cartilagens, músculos e ligamentos. As cartilagens da laringe são: tireóide, cricóide, aritenóides, corniculadas e cuneiformes.

A laringe tem aproximadamente 5 cm de comprimento no adulto e está situada na região anterior do pescoço. Estende-se da epiglote (C4) à borda inferior da cartilagem cricóide, onde se inicia a traquéia.

Durante a puberdade, no homem, a laringe cresce rapidamente em tamanho e as pregas vocais tornam-se cerca de 1 cm mais longas o que faz com que o limite

inferior da gama tonal masculina decresça em uma oitava. Nas mulheres, estas alterações são muito menos acentuadas.

A laringe é palpável anteriormente e constitui importante referencial em anatomia de superfície. Relaciona-se posteriormente com a laringo-faringe.

A cavidade da laringe é dividida em três porções: vestíbulo, ventrículo e cavidade infraglótica.

O vestíbulo vai do ádito da laringe às pregas vestibulares. Os ventrículos, um de cada lado, são limitados superiormente pelas pregas vestibulares e inferiormente pelas pregas vocais, (Fig.3, Fig. 4 e Fig. 5) as quais estendem-se da cartilagem tireóide ao processo vocal das cartilagens aritenóides com inclusão do ligamento e o do músculo vocal.

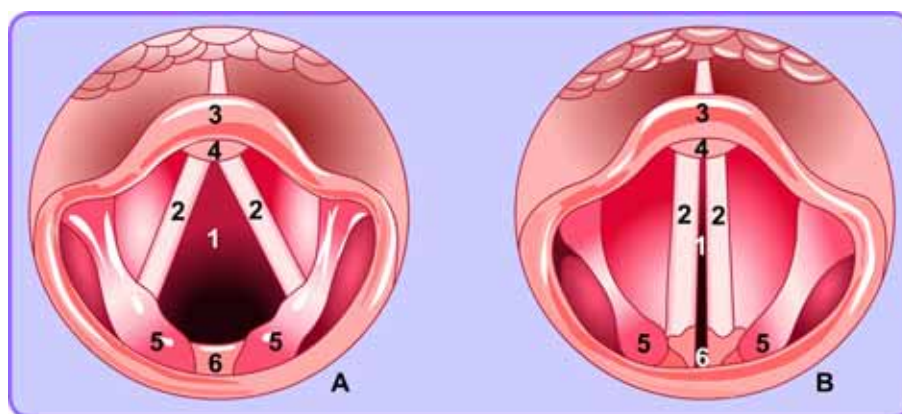


Fig. 3: A) Pregas vocais em abdução. B) Pregas vocais em adução. 1- Glote, 2-Pregas Vocais, 3 e 4- Epiglote, 5-Processo Vocal, 6-Músculo ariaritenóideo.

A parte mais anterior da glote denomina-se glote membranácea e a parte posterior, entre os processos vocais, constitui-se na glote cartilágnea.

O adito da laringe é considerado a sua entrada, e o fechamento do mesmo protege a via respiratória contra a penetração de partículas alimentares e corpos estranhos.

2.1.1 ESTRUTURAS DA LARINGE

As estruturas laringeas podem ser visualizadas nas Figs. 3 e 4.

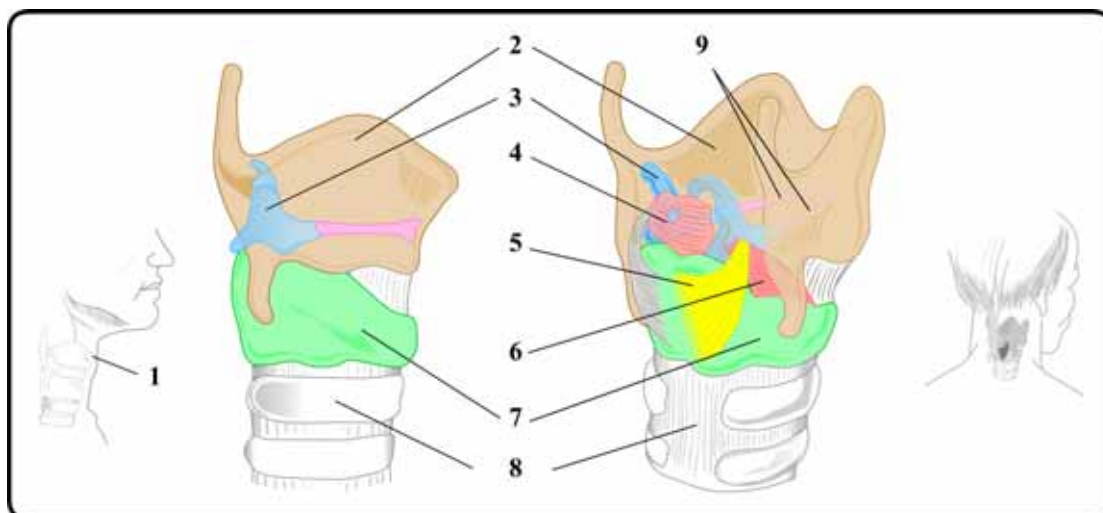


Fig. 4: 1- Laringe, 2- Cartilagem tireóide, 3- Cartilagens aritenóides, 4- Músculo interaritenideo, 5- Músculo cricoaritenóideo-posterior, 6- Músculo cricoaritenóideo-lateral, 7- Cartilagem cricóide, 8- Traquéia, 9- Pregas vocais.

Fonte: Voice Problem.org (imagem modificada)

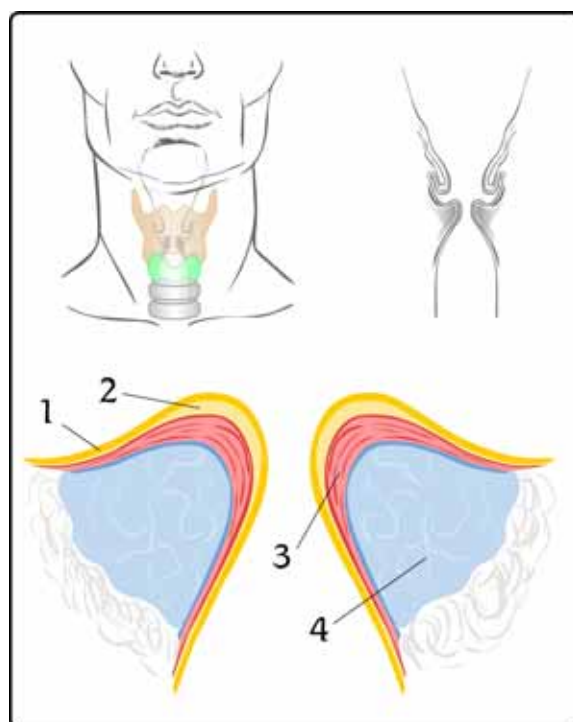


Fig. 5: Pregas vocais 1- Epitélio, 2- Lâmina-própria, 3- Ligamento vocal, 4- Músculo vocal.

Fonte: Voice problem.org (imagem modificada)

2.1.2 FUNÇÃO DA MUSCULATURA LARÍNGEA

A laringe tem em sua formação uma musculatura extrínseca e uma musculatura intrínseca. Os músculos extrínsecos são responsáveis pela sustentação e fixação da mesma, enquanto os músculos intrínsecos (Fig. 4) estão relacionados com as funções próprias da laringe, ou seja, proteção, respiração e fonação [41]. Estes músculos são sinérgicos e são classificados de acordo com o tipo de ação que executam:

- a) Abdutores: separam as cartilagens aritenóides (Fig.4-3) e as pregas vocais para o ato respiratório;
- b) Adutores: aproximam as cartilagens aritenóides e as pregas vocais para a fonação e proteção;
- c) Tensores: alongam e estiram as pregas vocais;
- d) Relaxadores: encurtam as pregas vocais.

A musculatura intrínseca da laringe é innervada pelo nervo Vago, X par craniano, e forma um sistema capaz de executar as ações necessárias para a produção vocal [6].

2.1.3 INERVAÇÃO DA LARINGE

Abaixo visualiza-se na Fig. 6 o nervo laríngeo superior e o nervo laríngeo recorrente, os quais são responsáveis pela innervação da musculatura de produção da voz. Ambos são ramificações do nervo vago-X par craniano- e estão localizados lateralmente no pescoço. O percurso do nervo laríngeo recorrente difere nos lados esquerdo e direito.

Pela relevância ao propósito do presente trabalho detalha-se especificamente o nervo larígeo recorrente esquerdo.

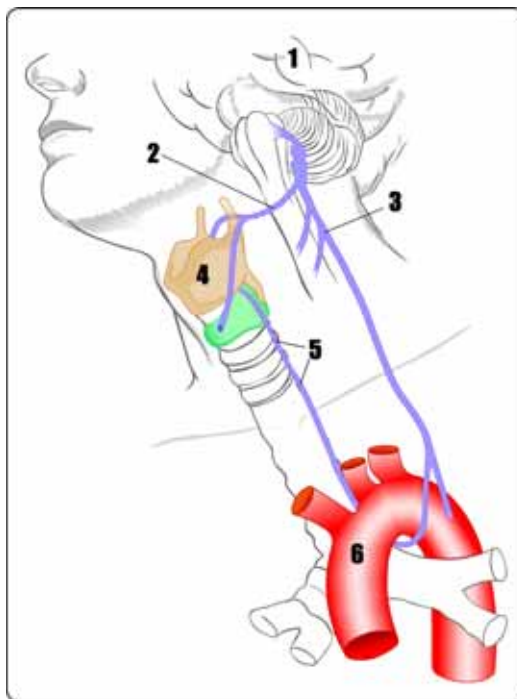


Fig. 6: 1- Córtex Cerebral, 2- Nervo larígeo superior, 3- Ramo descendente do Nervo larígeo recorrente, 4- Cartilagem tireóide, 5- Ramo ascendente do Nervo larígeo recorrente, 6- Artéria Aorta.

Fonte: Voice problem.org (imagem modificada)

Note-se que o nervo larígeo recorrente do lado esquerdo (Fig.6-3) tem um longo trajeto, pois após ramificar-se do nervo vago apresenta um ramo descendente que termina por abraçar a artéria aorta em nível torácico, para posteriormente ascender e inervar a laringe. O grande percurso desse nervo do lado esquerdo, comparado com o curto percurso do lado direito torna-se no principal agravante para a grande incidência de paralisias da prega vocal esquerda objeto desse estudo.

A glote é o espaço entre as pregas vocais e as estruturas que a sobrepõem formam a cavidade supraglótica. A infraglote é a região imediatamente inferior que se

estende até o primeiro anel da traquéia. A produção sonora é gerada na glote e amplificada na supraglote e cavidades de ressonância.

2.1.4 EFEITO DE BERNOULLI

Na inspiração a laringe se abaixa e as pregas vocais são abduzidas, ou seja, afastadas da linha média para permitir a passagem do ar aos pulmões. Quando inicia o processo da expiração, a laringe se eleva e as pregas vocais são aduzidas e se aproximam da linha média, o que ocasiona o início do processo de vibração, ou seja, a produção da onda sonora. Para isto ocorre um movimento de coordenação e de força contrátil da musculatura intrínseca associado ao fluxo aéreo. A aceleração do fluxo aéreo que advém dos pulmões e passa pela glote em fechamento associa-se à redução da pressão entre as pregas vocais e desta forma provoca o efeito de BERNOULLI. Este efeito ocorre pela pressão negativa que provoca a sucção da mucosa de ambas as pregas vocais para a linha média. O fenômeno de BERNOULLI afirma que a velocidade do fluxo de um gás ou fluido através de um tubo é inversamente proporcional à pressão nas suas paredes.

2.1.5 OS CICLOS GLÓTICOS

A vibração das pregas vocais consiste nas diversas aproximações e afastamentos das mesmas da linha média na produção dos ciclos glóticos (Fig. 7). Portanto, a laringe é um transdutor de energia aerodinâmica em acústica, por intermédio dos ciclos de abertura e fechamento das pregas vocais [24].

Os músculos e ligamentos que têm flexibilidade e que foram removidos da posição original tendem a retornar ao equilíbrio, ou seja, para a linha média, em oposição ao fluxo aéreo, o que provoca nova obstrução da passagem do ar.

A repetição destes ciclos de abertura e fechamento proporcionam a vibração para a produção de sons complexos, constituídos de uma frequência fundamental e uma série de harmônicos, ou seja, o sinal de voz [42].

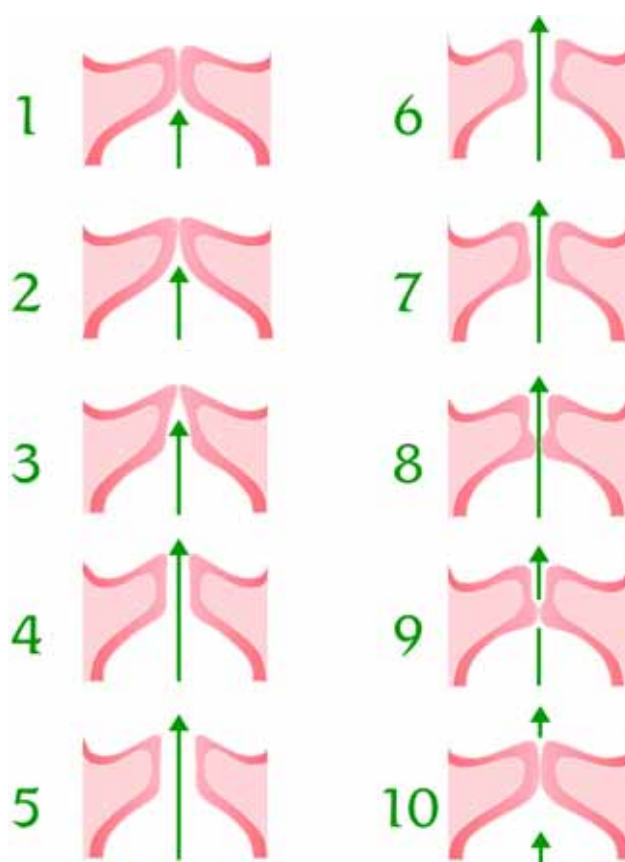


Fig. 7: As várias fases do ciclo glótico. 1- Fase fechada, 2- Início da fase de abertura, 3 e 4- Abertura, 5- Fase aberta, 6- Início da fase de fechamento, 7, 8 e 9- Fechamento, 10- Fase fechada novamente

Fonte: Voice problem.org (imagem modificada)

O ciclo glótico tem várias fases e se inicia quando a pressão subglótica torna-se superior à resistência glótica, o que desencadeia o processo de vibração. Podem acontecer diferenças na duração de cada fase do ciclo. Em casos de fadiga vocal ou em

algumas disfonias, observa-se a fase fechada mais curta e em presença de padrão vibratório mais estável, mais longa. Nos ciclos glóticos sucessivos pode-se observar pequenas perturbações no período ou na amplitude, denominados, respectivamente de JITTER e SHIMMER.

A frequência de vibração das pregas vocais é determinada pelo tamanho, comprimento e tensão das mesmas. Estes valores foram estudados por BHELAU (1993) [24], que encontrou na população de falantes de São Paulo, valores médios de 113Hz para os homens adultos, 204 Hz para mulheres adultas e 235 Hz para crianças.

Apesar da média apontar para os valores supra citados, as vozes masculinas podem variar de 80 a 150 Hz, as femininas de 150 a 250 Hz e as infantis podem estar acima de 250 Hz [43].

Reitera-se que as porções vibrantes das pregas vocais adotam diferentes configurações e a produção sonora ocorre pela interação dos três fatores físicos que são: massa, comprimento e tensão, determinados pela contração muscular e vibração da mucosa.

A estrutura harmônica da onda sonora é determinada pelo padrão vibratório das pregas vocais associado ao tamanho e forma das cavidades de ressonância , ou seja, do trato vocal [44], [22].

2.2 TRATO VOCAL

O trato vocal se assemelha a um tubo com dimensões aproximadas de 17 cm de comprimento por 4 cm de diâmetro correspondente ao espaço compreendido da glote aos lábios.

Na produção dos diferentes sons da fala, vogais ou consoantes, ocorrem modificações constantes na configuração do trato vocal, que resultam em combinações acústicas variadas. Porém, as considerações relacionadas às consoantes não são objetivo deste estudo.

2.2.1 MODELO FONTE-FILTRO

Para a produção das vogais, FANT [45], [46], descreveu o modelo Fonte-Filtro (Fig. 8), por meio do qual o explica os fenômenos acústicos da produção da voz.

O autor separa o que ocorre na fonte que se constitui na produção acústica decorrente da vibração das pregas vocais do que ocorre no filtro que é a produção acústica gerada pela amplificação sonora decorrente da impedância das cavidades supraglóticas, ou seja, do trato vocal [46].

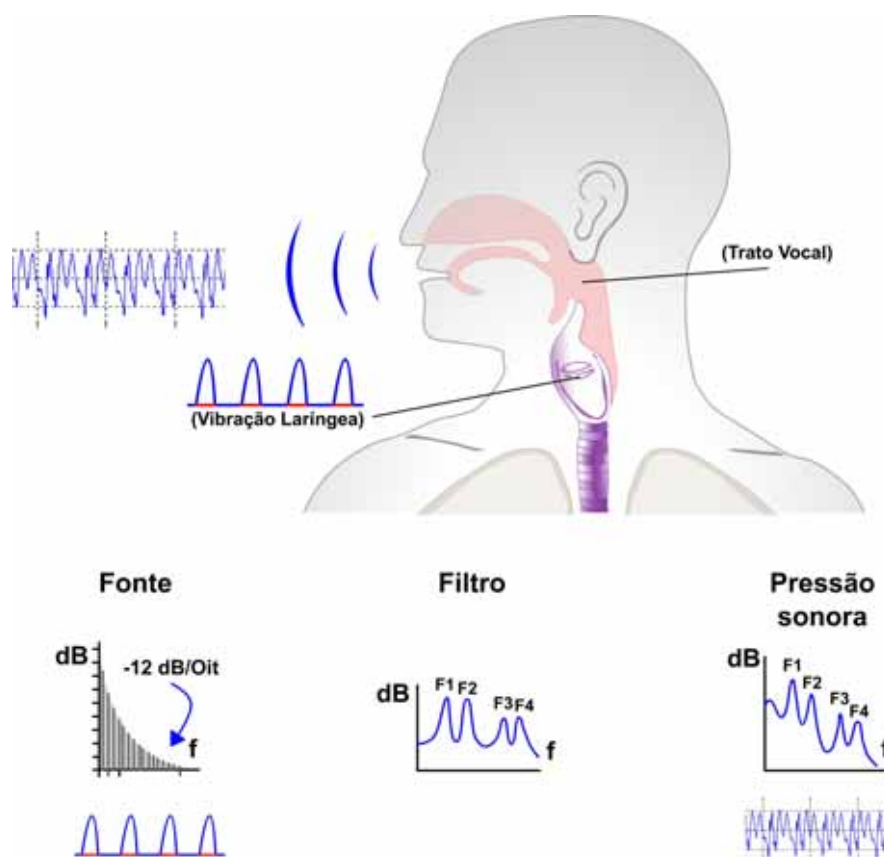


Fig. 8: Modelo Fonte-Filtro.

O modelo Fonte-Filtro (Fig. 8) descreve as características acústicas do sinal no domínio da frequência. Este espectro é caracterizado por picos que são chamados de formantes, que são as faixas de frequência que concentram maior energia acústica, decorrentes da ressonância do trato vocal. As frequências correspondentes a estes picos geralmente são designados por F1, F2, ..., Fn- primeiro formante, segundo formante...n-ésimo formante-.

Os três primeiros picos ou formantes fornecem informações quanto à identificação da vogal e os 4º e 5º têm menos conteúdo fonético, e trazem informação quanto à qualidade da voz.

As vogais são emissões sonoras, intensas e contínuas. Apresentam um trato vocal aberto e têm suas frequências de ressonância, ou seja, seus formantes bem definidos. Para a determinação de uma vogal, os formantes mais importantes são o primeiro e o segundo [24], [45]. O primeiro formante tem relação direta com o deslocamento da língua no plano vertical e o segundo com o grau de anteriorização da mesma, ou seja, o quanto a língua se deslocou no plano horizontal.

A vogal escolhida neste estudo foi a vogal /ε/, oral, anterior, média, aberta, não arredondada. A mesma consiste em uma vogal oral porque não tem ressonância nasal, ou seja, pela contração dos músculos palatofaríngeos, ocorre a elevação palatina que impede o escape de ar pela cavidade nasal. É uma vogal anterior, porque na produção da mesma há uma anteriorização da língua no plano horizontal, e pode ser considerada média em virtude da língua não estar muito elevada nem muito rebaixada no plano vertical, mas numa posição intermediária. Trata-se de uma vogal aberta, por não ocorrer nenhum tipo de estreitamento do trato vocal para a sua produção. E finalmente, é uma

vogal não arredondada, porque os lábios não estão protruídos e mantém-se numa posição neutra.

A escolha desta vogal foi motivada pelas características da sua estrutura formante, ou seja, menor amplificação das componentes não harmônicas- ruído- e pela configuração favorável da cavidade supraglótica.

Após a coleta da vogal /ε/ pelo microfone, iniciou-se o processamento deste sinal.

CAPÍTULO 3

MATERIAL E MÉTODO

No total, 94 sujeitos participaram desta pesquisa. Antes da coleta dos dados eles foram informados sobre o objetivo do estudo e assinaram o Termo de Consentimento Livre e Esclarecido, o qual foi aprovado pelo Comitê de Ética e Pesquisa da Universidade Federal de Uberlândia, sob o registro CEP 208/06.

3.1 COLETA DE DADOS

Nesse estudo, participaram 94 sujeitos adultos, do sexo masculino, com faixa etária entre 19 e 64 anos- média = 43 anos e desvio padrão = 16,97 anos-. Desses sujeitos, 48 eram portadores de voz normal e 46 de voz disfônica.

A definição da faixa etária desse estudo levou em consideração que a voz madura ocorre aos 18 anos, já com a frequência fundamental da voz nesta idade em nível em que posteriormente persiste por várias décadas [47], [48]. Observa-se também que alguns estudos [49], [50] consideram que a partir dos 65 anos diversas alterações fisiológicas provocam uma deteriorização da laringe, decorrentes de calcificações das cartilagens e modificações histológicas dos tecidos conjuntivos, que provocam variações acústicas no sinal de voz. Assim, os sujeitos na faixa etária investigada portaram características vocais comuns.

Note que crianças e sujeitos do sexo feminino foram excluídos do estudo, pois apresentam características vocais completamente distintas dos sujeitos do sexo masculino.

Inicialmente a totalidade da amostra foi submetida a exame de videolaringoscopia para a avaliação das respectivas condições orgânico-funcionais da laringe e do trato vocal. Este procedimento foi realizado com a utilização de telefibroscópio rígido 70° Endomed- FiberScope, endolight 1000, fonte de luz HTI- e nasofibroscópio flexível Machid- Ent- 30P III, câmera CCD Toshiba-, conforme a necessidade de cada caso. Por uma microcâmera acoplada na extremidade do aparelho obteve-se a gravação da imagem dinâmica das pregas vocais durante a produção da vogal /i/. A escolha dessa vogal baseou-se no fato da mesma facilitar a visualização das pregas vocais.

Os resultados da videolaringoscopia revelaram diferentes distúrbios e lesões laríngeas, os quais são relacionados a seguir: uma fenda fusiforme ântero-posterior grau leve pós-cirurgia de carcinoma in situ, dois cistos, duas constrições medianas, uma disфонia espasmódica com tremor vocal, quatro edemas de REINKE, cinco fendas fusiformes posteriores, uma fenda triangular ântero-posterior, quatro granulomas, um hemangioma, três leucoplasias, três papilomas, duas paralisias, quatorze pólipos, um sulco vocálic e duas vásculodisgenesias.

Os padrões normais e as alterações laríngeas encontradas nesta pesquisa estão classificados abaixo:

- Quanto ao comportamento do vestíbulo laríngeo: constrições ântero-posteriores e mediana;

- Quanto às fendas glóticas: fenda em ampulheta, paralela, irregular, triangular posterior, triangular médio-posterior, triangular ântero-posterior, fusiforme anterior, fusiforme posterior e fusiforme ântero-posterior [43].
- Quanto a alterações estruturais mínimas da mucosa de cobertura das pregas vocais: assimetria laríngea, sulco vocalis, cisto, ponte de mucosa, vásculodisgenesia e microdiafragma laríngeo [43];
- Quanto a lesões nas pregas vocais: nódulos vocais, pólipos, edema de REINKE, papiloma, sinéquia, granuloma, leucoplasia, neoplasmas malignos- câncer- ou benignos e
- Quanto a alterações neuro-motoras: paralisias das pregas vocais, Mal de Parkinson, Miastenias Gravis e outros. [51].

Alguns distúrbios laríngeos podem se visualizados nas Fig. de 9 a 18 extraídos de exames de videolaringoscopia. As imagens não são dos sujeitos dessa pesquisa, mas apenas ilustrativas.



Fig. 9: Prega Vocal Normal- Respiração



Fig.10: Prega Vocal Normal –Fonação.



Fig. 11: Paralisia de prega vocal esquerda. Observe a assimetria laríngea.



Fig. 12: Pólipo. Lesão unilateral no terço médio da prega vocal esquerda.



Fig. 13: Fenda fusiforme. Observe que a coaptação glótica é insuficiente, e mantém uma fenda em forma de fusão.



Fig. 14: Granuloma. Lesão no terço posterior da prega vocal direita.



Fig. 15: Fenda Triangular médio posterior.. O fechamento glótico não é completo. Mantém uma fenda triangular.



Fig. 16: Sulco. Observe a escara longitudinal na região médio posterior da prega vocal direita.

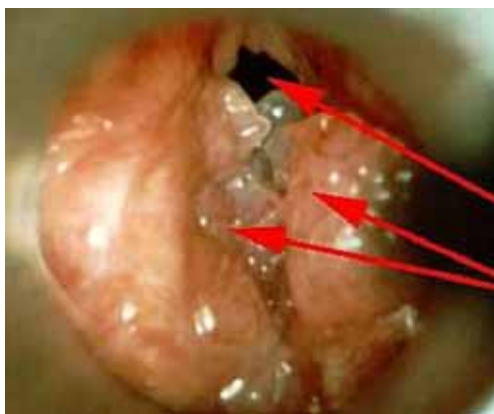


Fig. 17: Papiloma Bilateral. Observe o comprometimento de quase todo o espaço glótico.

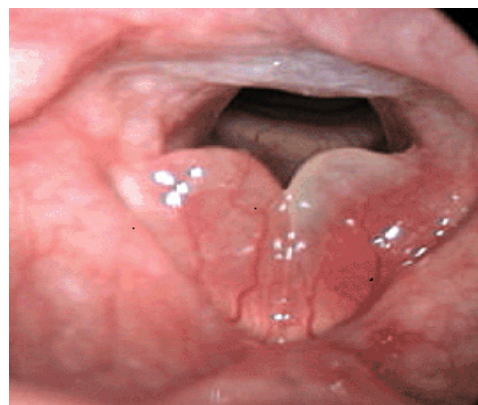


Fig. 18: Edema de REINKE bilateral. O edema das pregas vocais compromete quase todo o espaço glótico.

Fonte: International Archives of Otorhinolaryngology; Revista Brasileira de Otorrinolaringologia

Os portadores de voz normal, ou seja, sem queixa vocal foram avaliados por meio da análise perceptivo- auditiva e também submetidos à videolaringoscopia para confirmação das condições anátomo-fisiológicas da laringe.

A captação do sinal de voz para a realização da análise acústica foi feita por meio do microfone unidirecional Behringer- super cardioid, XM2000-, e durante a avaliação, os indivíduos permaneceram em pé com o microfone posicionado em ângulo de 90° à frente da boca mantida a distância de 10 cm da mesma. Conforme sugerido por MADAZIO [36], esta distância atenua interferências no sinal e mantém uma proporção sinal-ruído elevada. Os sujeitos foram orientados a respirar profundamente antes de emitir a vogal oral anterior média aberta não arredondada /ε/, de forma sustentada [52], [53], [36]. O sinal de voz, amostrado a uma taxa Fs da ordem de 44,1 kHz e com duração de três segundos, foi gravado no formato WAVE- .wav-.

Os dados técnicos do microfone utilizado nesse estudo são:

- Princípio de conversor: condensador, membrana de 16.

- Característica direcional- nodal-: microfone com membrana grande, dupla. Capta prioritariamente o som pela parte anterior e parcialmente pelas laterais. O som proveniente da região posterior é amortecido.
- Voltagem do circuito aberto: -40 dBV/pa (10 mV/pa)
- Resposta de frequência: 40 Hz-18 kHz
- Nível máx. de pressão acústica: 142 dB
- Alcance dinâmico: 119 dB
- Impedância nominal: 350 Ω

Após as gravações, estes arquivos de voz foram armazenados e um programa para realizar a análise acústica dos mesmos foi implementado.

Tal programa é apresentado na Fig. 19.

3.2 ESTRATÉGIA PARA ANÁLISE DE DADOS

A *toolbox* de Processamento de Sinais do Matlab foi utilizada para a análise dos dados. Foi desenvolvido um programa para a realização do processamento dos sinais de voz, cuja interface é visualizada na Fig. 19.

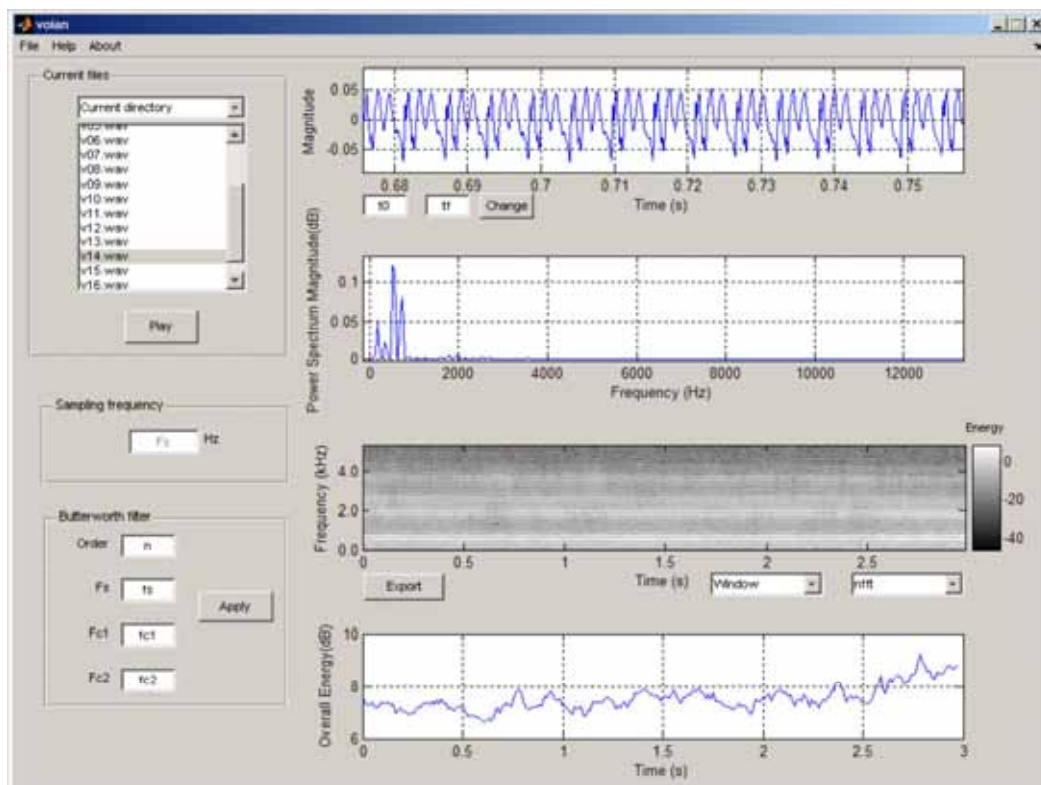


Fig. 19: Interface do programa desenvolvido neste trabalho para o processamento dos sinais das vozes normais e disfônicas.

O programa apresentado lê o arquivo .wav quando selecionado e processa o sinal.

Por meio deste programa, foi possível selecionar cada voz para análise individual. Ao selecionar uma voz, quatro gráficos são gerados simultaneamente. O primeiro gráfico é um gráfico de magnitude (%) por tempo (s). Esse valor de amplitude está normalizado entre -1 e 1, com -1 igual a 0%, 1 igual a 100%, e o 0 é valor de referência. Por meio desse gráfico é possível visualizar os ciclos glóticos.

O segundo consiste no espectro de potência do sinal. Trata-se de um gráfico de amplitude (em dB) por frequência (Hz) no qual visualiza-se a presença de picos que correspondem aos formantes da vogal /ε/, analisada neste estudo. Os primeiros picos do espectro de potência evidenciam as frequências de maior contribuição para o sinal.

O terceiro gráfico constituiu-se no espectrograma do sinal. O espectrograma, de faixa estreita, mostra no eixo vertical, a frequência (Hz) e no eixo horizontal, o tempo (s) e na escala em tons cinza pode-se visualizar a energia do sinal de voz. Por meio deste método pode-se analisar a presença dos componentes harmônico e de ruído do sinal.

O quarto gráfico integra os dados de energia do sinal, obtidos a partir do espectrograma, com a contribuição de somatória das frequências em cada instante de tempo, tanto do componente periódico- harmônico- como o componente aperiódico- ruído- do sinal. Revela no eixo vertical a energia global pelo tempo (s) no eixo horizontal.

3.2.1 DIAGRAMA DE BLOCOS

Para entender melhor o desenvolvimento desse trabalho, criou-se um diagrama de blocos que descreve a sequência dos passos empregados (Fig. 20).

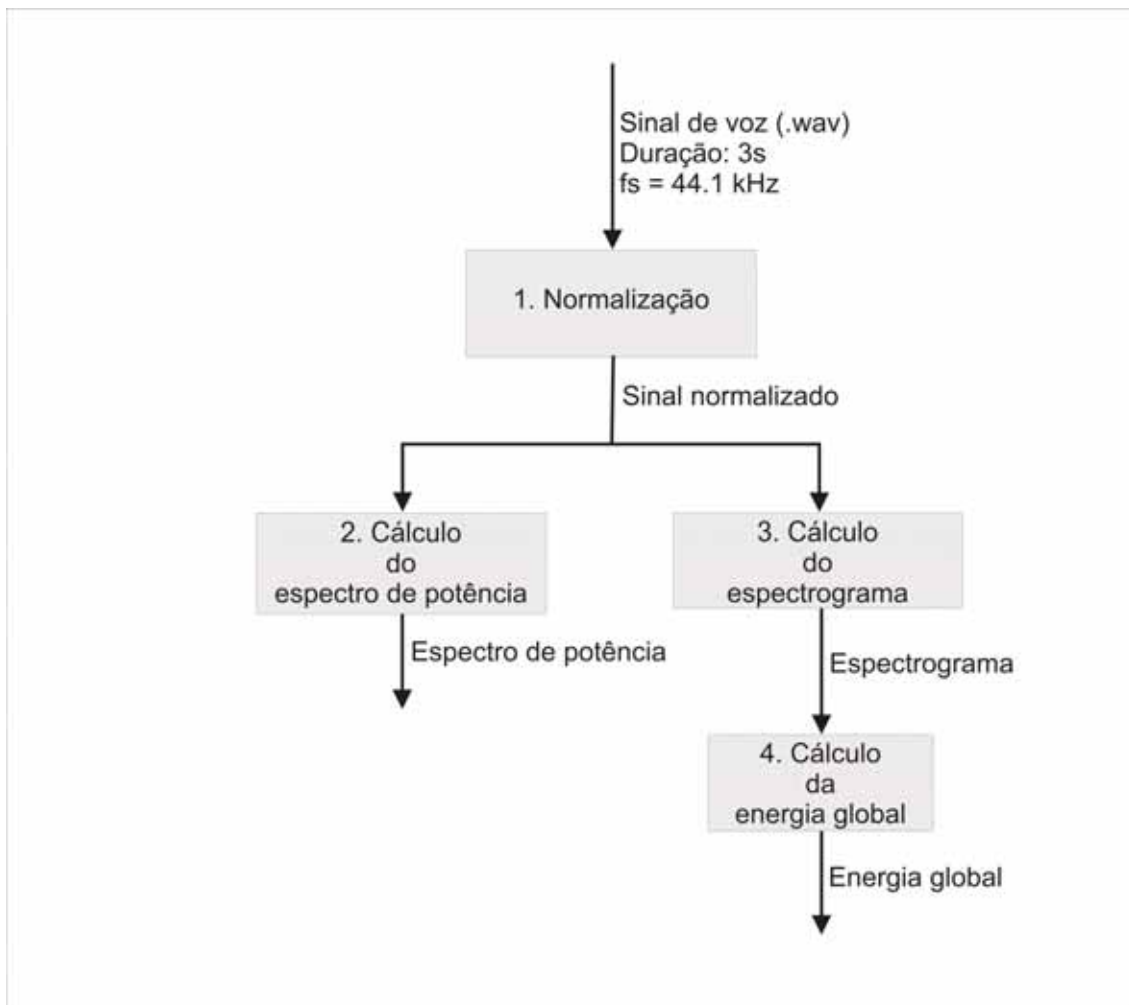


Fig. 20: Diagrama de blocos ilustrativo das etapas empregadas na análise dos sinais de voz.

3.2.2 DESCRIÇÃO DA NORMALIZAÇÃO

Inicialmente, os dados foram normalizados (Etapa 1) linearmente de modo que os valores de amplitude do sinal estivessem compreendidos no intervalo $[-1,1]$. Para isto, a Equação 1 foi utilizada, onde $x(n)$ é o sinal de voz no tempo discreto n , b é o número de bits do conversor A/D utilizado e $y(n)$ é o sinal de voz (no tempo discreto n) normalizado.

$$y(n) = \frac{x(n)}{2^b} \quad (1)$$

Note que esta normalização não altera as características do sinal, porém pode ser relevante no emprego de algumas ferramentas de Processamento Digital de Sinais [54], [39].

3.2.3 DESCRIÇÃO DO JANELAMENTO

Os dados normalizados foram janelados (Etapa 2), por meio de uma janela retangular de tamanho igual a 0,7 s, com o limite inferior de 0,5 s e superior de 1,2 s. Esse janelamento permitiu a seleção da região mais estável do sinal de voz visto que em algumas disfunções laríngeas os pacientes têm dificuldade em iniciar e finalizar a emissão do sinal. A estabilidade vocal foi identificada visualmente e auditivamente nesse intervalo. Este tipo de identificação de estabilidade, embora subjetiva, é uma prática comum na área que tem sido empregada em diversos estudos [36], [55]. Observe que o janelamento foi usado tanto para sinais de sujeitos saudáveis quanto para sujeitos disfônicos. Isto garantiu que o tamanho da série temporal analisada fosse padronizada no estudo.

A Fig. 21 ilustra esse janelamento para o caso de um sinal de voz obtido de um paciente portador de paralisia de prega vocal esquerda. Nos casos de paralisia de prega vocal não existe coaptação glótica ideal o que dificulta o padrão vibratório das pregas vocais. Na Fig. 19 verifica-se a presença de componente harmônico apenas neste intervalo selecionado, no qual ocorreu vibração laríngea. Nota-se que no restante do sinal existe apenas componente de ruído, no qual devido à alteração vocal, não foi possível manter a vibração da laringe.

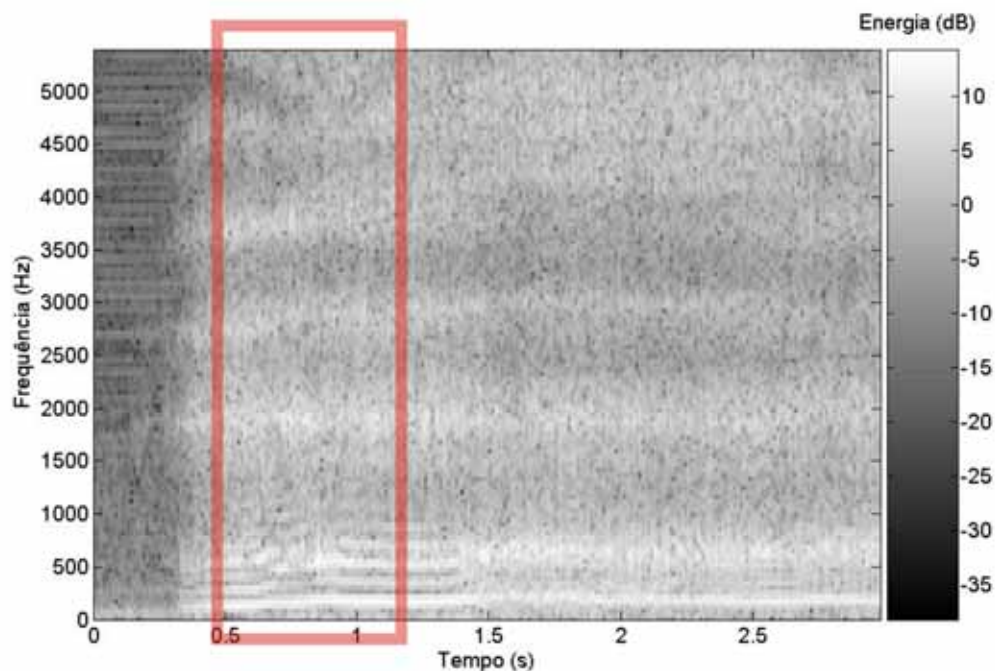


Fig. 21: Janela retangular de 0,7 s, com a seleção de região mais estável do sinal de voz, utilizada para análise neste estudo. Observa-se que essa região mais estável do sinal é identificada no espectrograma como faixas de energia contínuas e isoladas umas das outras.

Observa-se que o janelamento escolhido coincide com o intervalo de tempo mais estável da emissão, o qual Este foi necessário pelo fato de nos quadros de disfonias severas com decorticação da mucosa das pregas vocais ou coaptação glótica prejudicada, o tempo de fonação ser extremamente curto. Isto dificulta o início da emissão, que geralmente não se dá de forma isocrônica como nas vozes normais. Ocorre por meio de ataque brusco e dificulta a manutenção da emissão que muitas vezes não chega a atingir 1 s. Isto pode ser verificado pela presença de harmônicos no sinal apenas nesse intervalo de tempo selecionado e ocorre principalmente em casos mais severos, como o de câncer, paralisias ou papilomatoses. Para que este intervalo de tempo fosse padronizado em todas as amostras, optou-se por fazer um janelamento nessa região de maior estabilidade. Apesar de subjetivo, vários autores fazem a opção de realizar a

análise do sinal vocal com eliminação do início e o final da amostra por considerarem o meio da emissão o fornecedor de informações relevantes e confiáveis para a realização de análise adequada [36], [55]. E, neste estudo adotou-se o mesmo procedimento.

3.2.4 PERIODOGRAMA

Na Etapa 3, o Periodograma [54] baseado na Transformada de FOURIER foi utilizado para o estudo da distribuição de energia dos sinais em função da frequência. O Periodograma é uma estimativa mais precisa do espectro de potência [54] visto que este reduz o efeito do espalhamento de energia do sinal por meio do emprego de técnicas de janelamento. A Equação 2 apresenta a Transformada de FOURIER para uma janela de L amostras do sinal discreto $x[n]$, onde $w[n]$ representa a função janela utilizada e ω , a frequência em Hz..

Neste estudo, utilizou-se 8 janelas consecutivas de HANNING (sem sobreposição), para o cálculo do Periodograma. A Equação 3 apresenta o método utilizado para o cálculo do Periodograma a partir da Transformada de FOURIER, onde U representa uma constante de normalização. Note que o Periodograma é a média do espectro de potência de cada uma das 8 janelas utilizadas..

$$V(e^{j\omega}) = \sum_{n=0}^{L-1} w[n]x[n]e^{-j\omega n} \quad (2)$$

$$P_{ss}(\omega) = \frac{1}{LU} |V(e^{j\omega})|^2 \quad (3)$$

3.2.5 ESPECTROGRAMA

A Etapa 4 consistiu na avaliação da variação da energia instantânea dos sinais por meio do espectrograma. O mesmo permite a visualização dessa energia em função da frequência e do tempo. Esta ferramenta tem sido empregada em diversas pesquisas [1], [24], [42], [13], [56], e é comumente utilizada na prática clínica para a análise acústica vocal. Porém esta avaliação é frequentemente subjetiva e dependente da experiência do profissional. Para o cálculo do espectrograma foi utilizada a STFT- *Short-Time Fourier Transform*- e o tipo de janela empregada foi a de HUNNING com 1024 pontos. O espectrograma do sinal de voz pode ser obtido como o resultado da aplicação da STFT, conforme exposto na Equação 4.

$$X[n, \lambda] = \sum_{m=-\infty}^{\infty} x[n+m]w[m]e^{-j\lambda m} \quad (4)$$

onde $x[n+m]$ representa o sinal analisado, como uma função de uma variável de tempo discreta n , deslocando em m posições, λ representa o espectro (contínuo) de frequências e X representa a energia do sinal, em função de n e de λ . Novamente, $w[m]$ é a janela de HUNNING neste estudo.

3.2.6 ENERGIA GLOBAL

Na Etapa 5, avaliou-se o emprego da energia global do sinal, estimada por meio da integral do espectrograma (Equação 5) como uma técnica para a discriminação entre vozes normais e disfônicas e também para a discriminação dos distúrbios laríngeos.

O procedimento para a determinação da energia global a partir do espectrograma é apresentada na Equação 5:

$$U(n) = \int_0^{\lambda_m} X[n, \lambda] d\lambda \quad (5)$$

onde $U(n)$ representa a energia global do sinal de voz em função de uma variável de tempo discreta n , λ_m representa a maior frequência com contribuição significativa para a formação do sinal e $X[n, \lambda]$ representa o espectrograma do sinal em questão [57].

A determinação da energia global da voz foi realizada pela integração dos dados de energia obtidos no espectrograma, para um valor discretizado de tempo. Tal procedimento equivale ao somatório das contribuições de cada frequência na formação do sinal, para cada instante de tempo.

CAPÍTULO 4

RESULTADOS

4.1 ANÁLISE VISUAL DO ESPECTROGRAMA

Um método muito utilizado na análise acústica vocal e na classificação dos distúrbios laríngeos é o espectrograma [1], [45], [58], [28]. Com o objetivo de discriminar vozes normais e disfônicas, escolheu-se esta ferramenta. Ao analisar visualmente o espectrograma de vozes normais, percebe-se melhor definição dos harmônicos, que pode ser constatada por meio das raias horizontais na Fig. 22 a. Nesta mesma figura observa-se ainda que estes harmônicos aparecem em grande número, e podem ser percebidos inclusive, nas faixas de frequência acima de 3,0 kHz. Esta característica é compatível com as vozes normais segundo alguns pesquisadores [1]. A quantidade de ruído entre os harmônicos de vozes normais é praticamente nula, em contraste à quantidade de ruído presente nas vozes disfônicas. Na maior parte das vozes disfônicas, o componente harmônico, quando presente, ocorre apenas nas frequências mais baixas. Nas disfônias extremamente severas percebe-se grande quantidade de ruído difuso em todo ou praticamente todo o espectrograma (Fig. 22 b).

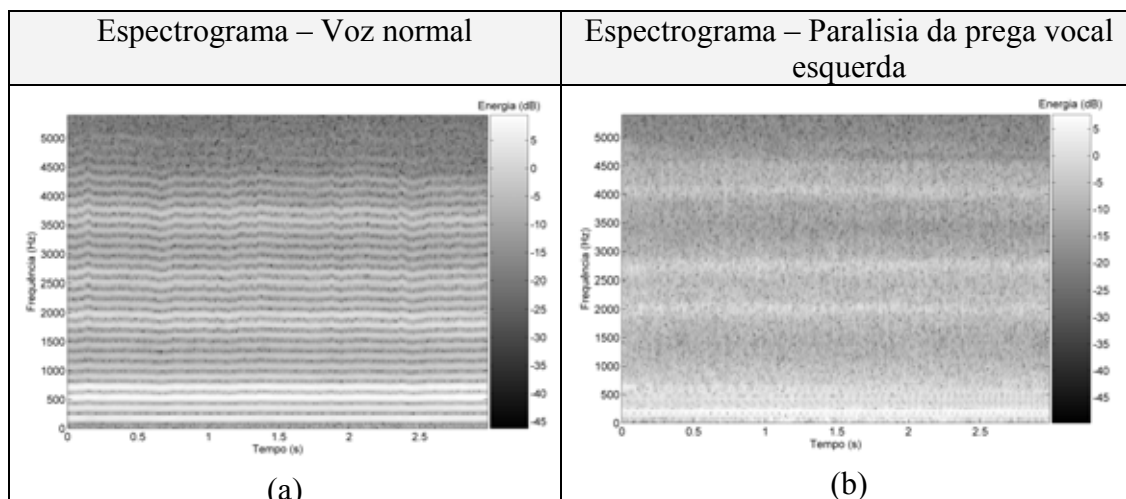


Fig. 22: a) Espectrograma da emissão da vogal oral anterior média aberta não arredondada /ε/ de um sujeito portador de voz normal.

b) Espectrograma da vogal oral anterior média aberta não arredondada /ε/ de um sujeito portador de paralisia da prega vocal esquerda.

4.2 ANÁLISE DO ESPECTROGRAMA MÉDIO

Foram analisados os espectrogramas médios dos grupos de vozes normais e disfônicas. Nessa análise destaca-se a contribuição da faixa de frequências de 500 a 550 Hz, na qual se concentra grande parte da energia dos sinais, tanto no grupo de vozes normais quanto no de vozes disfônicas. Essa faixa de frequências é compatível com o primeiro formante da vogal utilizada neste trabalho- vogal /ε/- para o sexo masculino [24]. Nota-se, portanto, que nas vozes normais existe apenas uma faixa de energia a qual é bem definida (Fig. 23 a). O mesmo não ocorre nas vozes disfônicas, nas quais é perceptível a presença de duas faixas de frequência em que se concentra a energia desses sinais (Fig.23 b). Alguns pesquisadores afirmam que a definição dos primeiros formantes das vogais anteriores /a/, /e/, /i/ e /ε/ é muito clara, tanto no português como no inglês [24], [59]. Neste contexto, acredita-se que a presença de ruído nas vozes

disfônicas gerado pela soproidade excessiva alterou a energia do sinal apenas nas vozes disfônicas e conferiu-lhe mais uma faixa de concentração de energia.

O componente de ruído presente nas disfonias é uma das características mais valorizadas na análise do sinal de voz, por se correlacionar a quantidade deste componente ao grau de rouquidão. Estudos anteriores, com o objetivo de interpretar melhor este componente de ruído adotaram uma escala com variação de 1 a 4 [60]. O grau 1 desta escala refere-se a sinais de voz com componente harmônico regular associado a um componente de ruído concentrado nas faixas de formantes. O grau 2 está relacionado a sinais de voz com predominância de componente de ruído na região do 2º formante das vogais /ε/ e /i/ e ainda ruído adicional acima de 3,0 kHz. O grau 3 é caracterizado por componente de ruído com energia intensa que substitui totalmente os harmônicos das vogais /ε/ e /i/. O grau 4 se relaciona-se com presença de ruído no 2º formante das vogais /a/, /ε/ e /i/.

Neste contexto, observa-se que ao avaliar os sujeitos disfônicos nesta pesquisa, encontrou-se resultados semelhantes aos dados da literatura, com a utilização da mesma vogal /ε/. Isto pode ser verificado no espectrograma médio, que evidencia mais uma faixa de energia nas vogais do grupo disfônico. Além disso, como o grupo disfônico tem diferentes graus de comprometimento vocal, o componente de ruído está presente em diferentes faixas de frequência.

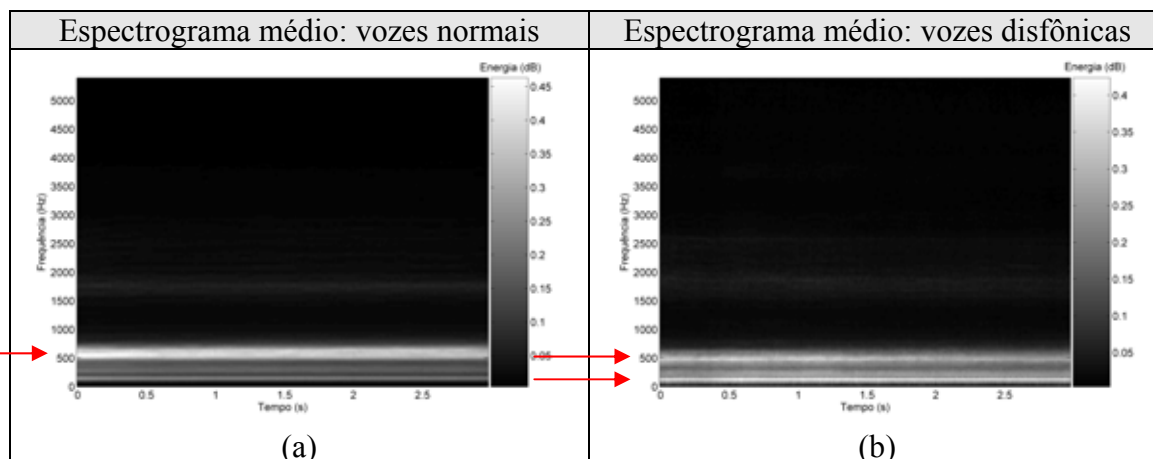


Fig. 23: a) Espectrograma médio das vozes normais. b) Espectrograma médio das vozes disfônicas. As setas indicam as faixas de frequência onde a energia está concentrada em ambos os casos.

4.3 ANÁLISE DO PERIODOGRAMA

Inicialmente utilizou-se o espectro de potência com o objetivo de discriminar vozes normais de vozes disfônicas. Para esta análise, escolheu-se a faixa de frequência entre 0 e 1,0 kHz. Esta escolha justifica-se pelo fato de que a maior parte da energia dos sinais está concentrada nessa faixa de frequência. Isto ocorre porque o espectro laríngeo é linear e a energia dos harmônicos, que são múltiplos inteiros da frequência fundamental, decresce em frequência na taxa de 12 dB por oitava. Por essa razão, apenas esta faixa de frequência foi relevante para esta análise.

Os resultados apresentados na Fig. 24 mostram uma separabilidade entre os espectros médios de potência de vozes normais e disfônicas em algumas faixas de frequência. Esta separabilidade evidencia-se nos intervalos de frequência de 0 – 150 Hz e 550 – 800 Hz. Em outras faixas de frequência, observa-se sobreposição de energia. Isso limita a aplicação do espectro de potência com a finalidade discriminatória entre vozes normais e disfônicas.

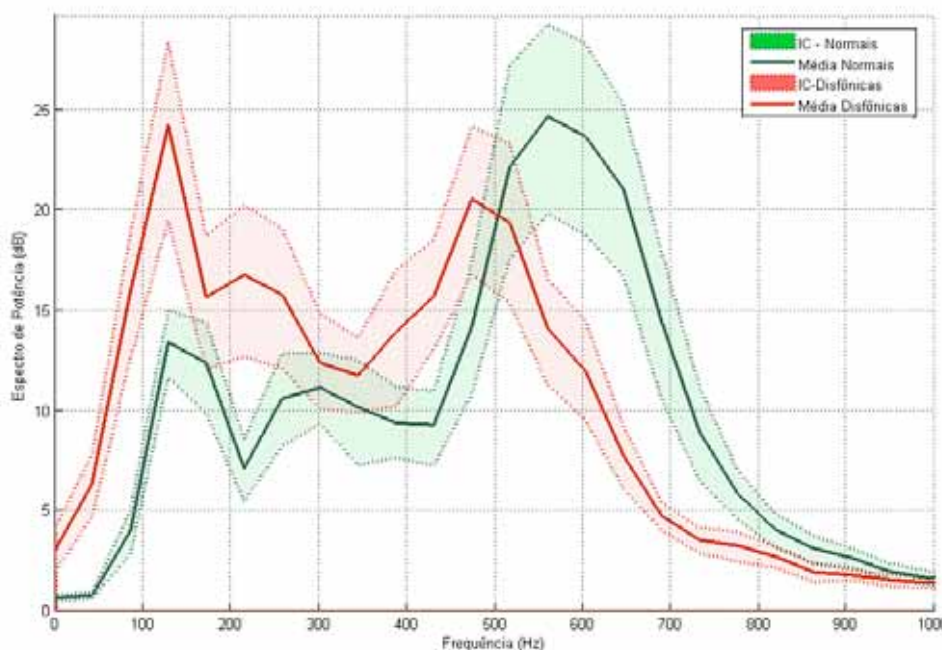


Fig. 24: Espectro de Potência médio e seus intervalos de confiança para vozes normais e disfônicas.

4.4. ANÁLISE DA ENERGIA GLOBAL

A energia global é calculada pela integração da energia contida em cada frequência que compõe o sinal em um dado instante de tempo. Esse é um método que fornece, de forma precisa, a variação da energia da voz ao longo do tempo, por considerar as contribuições de todas as frequências do sinal, isto é, tanto a contribuição da energia dos harmônicos quanto a da energia do componente de ruído. O componente de ruído é provocado pela turbulência do ar ao passar pela glote e o componente harmônico é produzido pela vibração das pregas vocais e alterado de acordo com a impedância do trato vocal.

As diversas alterações nas estruturas que compõem a laringe, seja por atrofias, retrações cicatriciais, lesões de massa ou alterações neuro-musculares, acabam por

mdificar o fluxo do ar ao passar pela glote. Essas alterações provocam maior turbulência e têm como consequência a amplificação do componente de ruído amplificado em relação ao sinal de voz. Esta componente de ruído é intensa e está presente nas doenças laringeas [1].

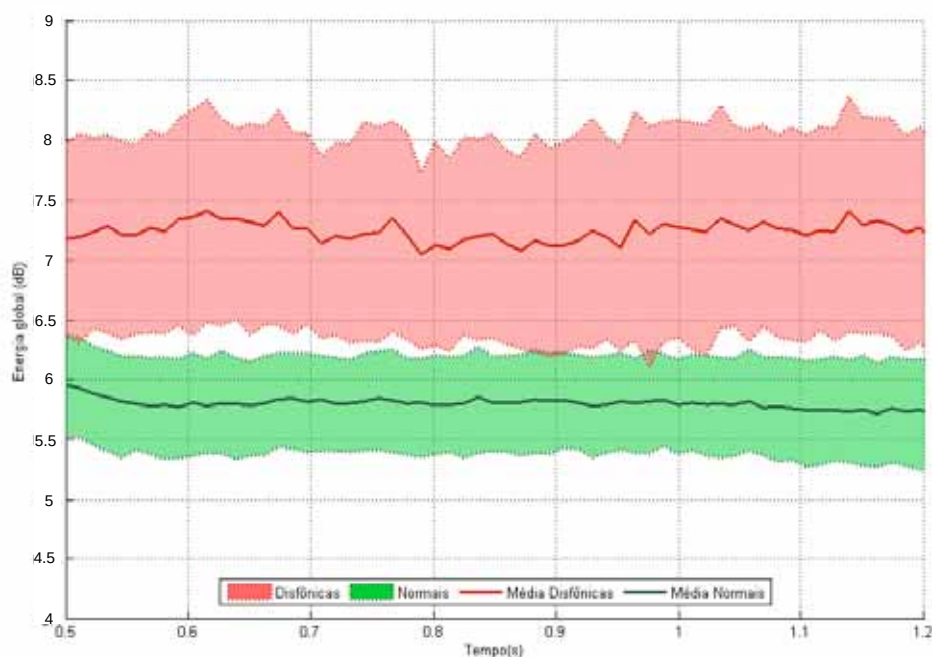


Fig. 25: Gráfico da média e intervalo de confiança da energia global de vozes normais e disfônicas.

Ao avaliar os dados encontrados por meio da análise da energia global, observa-se que existe uma separabilidade estatisticamente significativa entre as vozes normais e disfônicas. Esta separabilidade foi confirmada pelo intervalo de confiança da média da energia global estimado pela técnica BOOTSTRAP, o que é apresentado na Fig. 25.

Com os resultados obtidos e suportados por meio do intervalo de confiança, verifica-se que a energia global foi uma ferramenta eficaz na discriminação entre vozes normais e disfônicas.

A energia global média foi maior para as vozes disfônicas quando comparadas a vozes normais. Este fato pode ser justificado pela maior presença do componente de ruído em vozes disfônicas, em decorrência das disfunções biomecânicas das pregas vocais, da alteração da pressão subglótica e/ou dos distúrbios neuromotores.

Em uma avaliação complementar utilizou-se os sinais de voz do grupo disfônico com o objetivo de discriminar os diversos distúrbios laríngeos, por meio da energia global destes sinais.

Neste estudo, encontrou-se vozes com parâmetros vocais extremamente perturbados e outras bastante próximas das vozes normais. Isto deve-se ao fato de constatar-se diferentes tipos de doenças laríngicas, e em cada qual existirem diferentes graus de comprometimento.

A Fig. 26 apresenta a média da energia global das vozes das diversas doenças laríngicas.

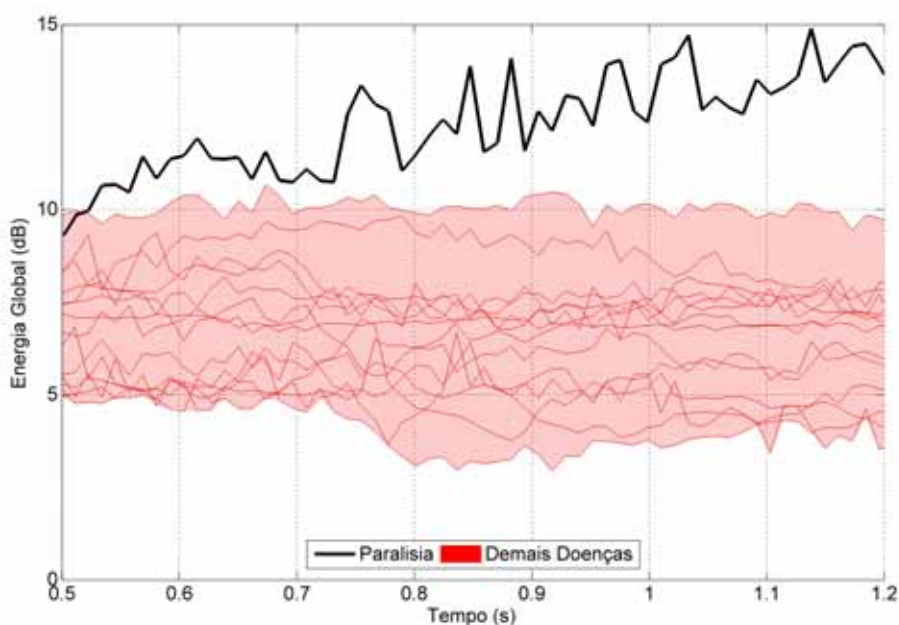


Fig. 26: Média da energia global dos diversos distúrbios laríngicos.

Não foi possível encontrar separabilidade entre todas as desordens laríngicas investigadas neste estudo. A maior parte das doenças investigadas apresentou energia global muito semelhante, o que pode ser constatado na Fig. 26. Porém observa-se separabilidade entre o grupo de vozes de paralisia e as demais doenças laríngicas investigadas.

Os dois casos de paralisia de prega vocal deste estudo se fixaram em posição intermediária por lesão do nervo laríngeo recorrente. A lesão foi unilateral esquerda. A análise perceptivo-auditiva em ambos os casos revelou qualidade vocal rouco-soprosa extrema.

Com base nestes resultados, pode-se sugerir a energia global como uma ferramenta auxiliar importante na discriminação entre vozes normais e disfônicas.

CAPÍTULO 5

DISCUSSÃO

Nesse estudo utilizou-se ferramentas de Processamento Digital de Sinais no domínio da frequência e do tempo com o propósito de discriminação entre vozes normais e disfônicas.

Para a análise no domínio da frequência torna-se necessário entender o modelo Fonte-Filtro descrito por FANT [45] [46]. O filtro oral é caracterizado por picos (F1, F2, F3, e outros). Estes picos são chamados de formantes e correspondem aos modos normais de ressonância dos tubos acústicos. Os três primeiros formantes trazem informações quanto à identificação da vogal, ou seja, têm maior conteúdo fonético, porém certos pesquisadores consideram o esforço vocal e as alterações da musculatura intrínseca da laringe como fatores contribuinte para o espectro resultante [61], [45]. Ao utilizar o espectro de potência observou-se diferenças ao avaliar os espectros médios das vozes normais e disfônicas. Pode-se inferir que o componente de ruído presente nas vozes disfônicas modificou o espectro resultante. Isto determinou a diferença entre os dois grupos. Porém, esta separabilidade não ocorreu em toda faixa de frequência do sinal, o que limita a aplicação desse método na discriminação das vozes normais e disfônicas. Estes resultados são compatíveis com pesquisas similares [14].

Apesar da diferença dos espectros de potência entre os dois grupos, esta separabilidade não ocorreu em toda a faixa de frequência, e por essa razão, não foi eficaz na discriminação entre padrões normais e disfônicos.

Neste estudo, o diagnóstico de voz normal foi atribuído aos indivíduos sem queixas vocais, que apresentaram na videolaringoscopia diagnóstico de pregas vocais

normais. Tal diagnóstico é compatível com coaptação glótica completa, ausência de alterações estruturais mínimas e de lesões nas pregas vocais e ainda produção fonatória equilibrada. Vários pesquisadores ao definirem voz normal, afirmam que não existem padrões definidos sobre suas características [1], mas aceitam a evidência de uma produção vocal equilibrada, considerada neutra, sem comportamento hipofuncional de adução glótica, nem hiperfuncional [62].

O diagnóstico de vozes disfônicas foi atribuído aos sujeitos portadores de qualidade vocal alterada nos vários graus de rouquidão, aspereza e soprosidade constatados pela análise perceptivo-auditiva vocal. Estes sujeitos portadores de disfonia apresentaram na videolaringoscopia presença de lesões nas pregas vocais, alterações estruturais mínimas ou quanto à coaptação glótica de caráter hipofuncional ou hiperfuncional.

Ao comparar os espectrogramas das vozes normais e disfônicas, nota-se que as vozes normais têm maior componente harmônico e menor componente de ruído, enquanto que nas vozes disfônicas têm acontece o oposto [14]. Portanto, nesse estudo observou-se que o componente de ruído gerado pela soprosidade vocal contribuiu com maior energia na formação do sinal.

Pesquisas anteriores fazem referência à presença de ruído intenso na região dos primeiros formantes da vogal /ε/ nas vozes disfônicas, e constata que este componente pode substituir totalmente os harmônicos [60]. Estes achados de ruído intenso nas disfonias em pesquisas anteriores explicam o fato das vozes disfônicas deste estudo terem energia global maior do que as vozes normais.

Os harmônicos são múltiplos inteiros da frequência fundamental que é a mais baixa da série, e são, portanto, periódicos. O ruído consiste num fenômeno acústico não

periódico, resultante da superposição desarmônica de sons provenientes de várias fontes que têm movimentos de vibração com diferentes frequências, sem apresentar relação entre si [42].

Na paralisia laríngea, doença que teve a energia global maior em comparação ao grupo de doenças analisadas neste trabalho, ocorre irregularidade no fechamento glótico. Isto contribui significativamente à presença de ruído no sinal de voz.

As disfonias paralíticas têm o componente de ruído aumentado em decorrência de múltiplas variáveis. A onda mucosa da prega vocal tem o seu movimento parcial ou totalmente afetado e a mesma pode estar fixada em posições mediana, paramediana, intermediária ou lateral. A prega vocal paralisada muitas vezes está desnivelada em relação à prega vocal sadia e a atrofia causada pela paralisia pode arquear a borda livre da mesma. A gravidade da alteração vocal está relacionada à denervação da musculatura e atenua-se em presença de mecanismo compensatório pela prega sadia. O quadro fonatório varia muito, porém predomina, na maioria das vezes, uma voz rouco-soprosa, emitida com esforço. A voz pode ser diplofônica- bitonal- ou entrar em falsete, e recebe a denominação de falsete paralítico [63]. Nas vozes de paralisia deste estudo o componente de ruído era muito intenso em toda a extensão de frequência e os harmônicos praticamente ausentes, o que justifica a energia global maior nessa doença quando comparada com os demais distúrbios laríngeos investigadas.

Como o sinal de voz é composto de harmônico e ruído, pode-se sugerir a energia global da voz como uma ferramenta auxiliar na discriminação entre vozes normais e disfônicas considerando por considerar a contribuição dos dois componentes.

Tanto o componente harmônico como o de ruído podem ser verificados ao analisar na formação do sinal, a contribuição de toda a faixa de frequência em cada instante de tempo, ou seja, pela energia global da voz.

Por meio da análise da energia global da voz, percebe-se que vozes disfônicas têm energia global maior do que as vozes normais. E, ao utilizar o mesmo parâmetro de comparação percebe-se que apenas as vozes de paralisia intermediária de prega vocal esquerda se destacaram e tiveram energia global maior do que as vozes de outras doenças laringeas.

O fato da energia global comportar-se de forma semelhante em diferentes alterações laringeas e diferenciar-se apenas nos quadros de paralisia sugere que a turbulência provocada pelo escape excessivo de ar causado pela má coaptação glótica das paralisias confere mais componente de ruído que ocasiona aumento de energia do sinal.

Uma mesma patologia laringea pode ter diferentes graus de comprometimento vocal, desde leve a severo. Este fato leva a acreditar-se que a energia global possa comportar-se de forma diferente e talvez ter caráter discriminativo nestes diferentes graus de disфонia. Neste contexto, a energia global deve ser testada com o objetivo de discriminar diferentes graus de comprometimento inerentes a um mesmo distúrbio laringeo e ainda, e testada em diferentes momentos do tratamento.

Com base nos resultados acima apresentados, ressalta-se que os métodos utilizados neste estudo contribuem para esclarecer certas características relevantes para o diagnóstico das desordens laringeas. Porém, constatou-se que só foi possível discriminar vozes normais de vozes disfônicas por meio da energia global, pois o espectro de potência foi limitado neste tipo de discriminação. Daí, sugere-se o uso da

energia global como ferramenta computacional com o fim de discriminar vozes normais de vozes disfônicas.

CAPÍTULO 6

CONCLUSÕES

Este estudo esclarece certos conceitos a respeito da energia vocal, pois evidencia a maior contribuição do componente de ruído quando comparado à contribuição dos harmônicos na formação do sinal.

A análise da energia global não possibilita uma classificação dos diversos distúrbios laríngeos, pois não há separabilidade estatisticamente significativa na energia global dos distúrbios investigados. Porém, esse método de análise mostrou ser importante ferramenta de processamento de sinais que propicia discriminação entre vozes normais e disfônicas.

Pode-se ainda concluir, com estes resultados, que vozes disfônicas têm energia global maior do que vozes normais.

Conclui-se ainda que o espectro de potência mostrou-se limitado quando utilizado para a discriminação entre vozes normais e disfônicas.

Este estudo proporcionou um acréscimo à avaliação fonoaudiológica por conferir recursos e fornecer uma estratégia de análise objetiva que proporciona a discriminação entre padrões vocais normais e disfônicos.

CAPÍTULO 7

ESTUDOS FUTUROS

Como trabalhos futuros sugere-se realizar a análise da energia global, com a utilização de um único distúrbio laríngeo com o objetivo de obter separabilidade nos diferentes graus de comprometimento.

Neste estudo realizou-se uma comparação por meio da energia global, entre vozes de paralisia esquerda por lesão do nervo laríngeo recorrente e as demais doenças laríngeas, em decorrência do que sugere-se também como trabalho futuro, utilizar a energia global para analisar o sinal de voz de paralisias em estudos longitudinais de um mesmo paciente, em diversas etapas da evolução da doença, ou seja, no início do processo de instalação da doença, após seis meses, no pré-cirúrgico e no pós-cirúrgico das tireodectomias.

Sugere-se também utilizar a energia global na discriminação entre os diversos tipos de leucoplasias.

Indica-se a realização de trabalhos que utilizem a energia global da voz como parâmetro de comparação entre os diferentes tipos de fendas glóticas.

Sugere-se a energia global para discriminar os diversos tipos de paralisias laríngeas do nervo laríngeo recorrente nos diversos graus, ou seja, nas paralisias medianas, paramedianas, intermediárias e laterais e também nas paralisias por lesão do nervo laríngeo superior.

A energia global pode ser uma ferramenta importante na discriminação dos sinais de voz dos vários tipos de sulco vocalis.

A energia global pode ser utilizada em estudos comparativos entre lesões de massa de pregas vocais e lesões atróficas com coaptação glótica insuficiente.

Sugere-se também a energia global para comparar os diversos tipos de corpectomia, de acordo com o grau de ressecção, ou seja, nos tipo I ou subepitelial, no tipo II ou subligamentar, no tipo III ou transmuscular, no tipo IV ou total ou no tipo V ou estendida.

Preconiza-se também a utilização da análise da energia global como um parâmetro de comparação na voz esofágica, por meio da comparação de pacientes que façam uso de prótese traqueo-esofágica com aqueles pacientes que não a utilizam.

Enfim, sugere-se também a realização de estudos com a utilização do espectro de potência na discriminação dos distúrbios laríngeos, pois esta ferramenta mostrou-se limitada na discriminação entre vozes normais e disfônicas, mas não foi testada na discriminação dos diversos distúrbios laríngeos ou nos seus respectivos diversos graus de comprometimento.

APÊNDICE 1

GLOSSÁRIO

Aspereza: irregularidade vibratória relacionada à rigidez da mucosa das pregas vocais. Este padrão encontra-se mais presente no sulco vocálico, nas paralisias de prega vocal em abdução, nas leucoplasias e em outras lesões com redução de massa. Presença de características vocais de pouca projeção e com componente de ruído em alta frequência na análise acústica [64].

Astenia: as pregas vocais estão em hipofunção, com projeção vocal reduzida. A voz é caracterizada por harmônicos pouco definidos na espectrografia acústica [26].

Câncer: é caracterizado pelo carcinoma que tem atividade destrutiva local. Este tipo de câncer representa 4% dos tumores malignos e afeta principalmente os fumantes do sexo masculino. A severidade da malignidade é avaliada pelo uso do sistema “TNM” ou suas variáveis- *American Joint Committee for Cancer Staging and Results Reporting, 1983*-. O **T** refere-se à localização do tumor primário, o **N** indica o envolvimento de nódulos linfáticos e o **M** significa propagação da lesão para outras partes do corpo (metástases). Quanto à localização encontra-se uma graduação de 0 a 4, na qual o **T0** (zero) corresponde à ausência de tumor; **Tis** carcinoma in situ, **T1** confinado às pregas vocais, **T2** extensão supraglótica ou subglótica, **T3** confinado à laringe mas com prega fixa e **T4** tumor massivo [70].

Cisto: é uma alteração estrutural mínima que cresce da camada superficial da lâmina própria para a superfície da mucosa da prega vocal, e pode também inserir-se parcialmente no ligamento vocal [71].

Edema de REINKE: consiste no inchaço das pregas vocais que pode se estender a partir da comissura anterior ao início do processo vocal das cartilagens aritenóides e aumentar a massa do revestimento vocal. Acomete mais às mulheres, mas pode ocorrer em ambos os sexos. Está associado ao tabagismo [69].

Fendas glóticas: são caracterizadas pela manutenção de parte do espaço glótico aberto na situação de fonação. Na fonação normal espera-se fechamento glótico completo. Essas fendas podem ocorrer por inaptações miodinâmicas, como por exemplo, nos estados de contração excessiva da musculatura abdução intrínseca da laringe, ou seja, dos músculos cricoaritenóideos posteriores ou também relacionar-se à flacidez (hipofunção) por diminuição da atividade da musculatura adutora. Podem ocorrer também por inaptações anatômicas [65]. Nesse estudo houve cinco casos de fenda glótica fusiforme posterior e um de fenda triangular ântero-posterior. A fenda fusiforme posterior é uma variação da fusiforme ântero-posterior, e apresenta menor rigidez na sua região anterior em comparação à fenda fusiforme ântero posterior. O outro tipo de fenda encontrado foi a fenda triangular ântero-posterior que tem como característica a hipotonia, ou seja, a diminuição da contração de toda a musculatura intrínseca da laringe [43].

Granuloma: consiste numa afecção caracterizada pelo trauma causado pelo abuso vocal nos indivíduos que usam a voz de forma intensa e que tem associado o refluxo gastro-esofágico (RGE). O refluxo de substâncias ácidas associado ao trauma vocal gera inflamação da mucosa e estimula a formação de tecido de granulação [68].

Hemangioma: consiste no tumor vascular mais comum que acomete a região da cabeça e pescoço em 60% dos casos. É raro na laringe. No adulto, os quadros são caracterizados por disfonia leve ou disfagia. As pregas vocais podem encontrar-se

atróficas, com fenda fusiforme e diminuição da vibração da onda mucosa, sinais estes indicativos de provável sulco vocálico. É comum a associação de hemangioma com AEMs- alterações estruturais mínimas das pregas vocais- [74].

Jitter: medida de perturbação ou variabilidade da frequência fundamental ciclo a ciclo. É um fenômeno que está presente em todo sinal vocal. Isto acontece porque a voz humana é quase periódica, ou seja, cada ciclo sofre pequenas variações no tempo. Em presença de disfonias acontece alteração nos padrões de vibração das pregas vocais, o que decorre em aumento do Jitter. O Jitter é uma medida em porcentagem (%) e pode ser considerado normal, valores de até 0,5 [64].

Leucoplasias: lesões hiperplásicas que enrijecem a túnica mucosa. Resultam de trauma vocal nas alterações estruturais mínimas ou associadas a cistos abertos, fechados e/ou a carcinomas laríngeos [65].

Nódulos: são as lesões mais superficiais da lâmina própria, constituídos por variada composição de edema e fibrose. Localizam-se no 1/3 médio das pregas vocais, e são provocados pelo atrito constante gerado pelos abusos vocais [65], [66].

Papiloma: tumor verrucoso, benigno, que afeta o epitélio da prega vocal e é causado pelo vírus HPV. Deve ser removido, pois prolifera rapidamente e obstrui a via aérea. Quando aparece na idade adulta pode evoluir para tumores malignos [70].

Paralisia da prega vocal: incapacidade de uma ou ambas as pregas vocais se moverem devido à falta de inervação da musculatura intrínseca da laringe -nervo laríngeo superior ou recorrente-, Pode ter origem central ou periférica, com etiologia traumática, operatória, compressiva ou neurítica [73].

Pólipo: é um aumento de massa da cobertura das pregas vocais que pode estar relacionado a um único incidente de abuso vocal traumático. Pode ser pediculado ou

séssil e a rigidez depende do tipo histológico- edema, hemorragia, trombose, degenerações e outros-. [67].

Rouquidão: presença de irregularidade vibratória da mucosa das pregas vocais durante a fonação, mais comum em lesões de massa, que gera nos traçados espectrográficos componente de ruído em baixa frequência entre os harmônicos inferiores [1].

Shimmer: medida de perturbação ou variabilidade da amplitude ciclo a ciclo. É uma medida em porcentagem (%), sendo considerados normais valores inferiores à 3,0 [64].

Soprosidade: corresponde à presença de componente de ruído semelhante a vazamento de ar pela análise perceptivo-auditiva vocal e o correlato fisiológico mais freqüente é a presença de fenda glótica. O correlato acústico é a presença de componente de ruído intenso entre as frequências acima de 4000 HZ [6].

Sulco Vocalis: faz parte das alterações estruturais mínimas das pregas vocais, de origem congênita, com a formação de escaras longitudinais nas mesmas. A massa e a cobertura das pregas vocais encontram-se diminuídas e provocam, provocando rigidez na mucosa [72].

Tensão: as pregas vocais estão em hiperfunção, com esforço vocal provocado pelo excesso de coaptação glótica. Presença de componente de ruído nas frequências altas do espectrograma e harmônicos altos marcados [6].

APÊNDICE 2

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Eu _____ RG _____
 assino esse Termo de Consentimento com a finalidade de autorizar a realização do exame de Análise Acústica Vocal Computadorizada e Videolaringoscopia da Pesquisa: *"ESTUDO DOS DISTÚRBIOS VOCAIS ELARÍNGEOS POR MEIO DE MODELOS PROBABILÍSTICOS"*, realizada pela Universidade Federal de Uberlândia-Faculdade de Engenharia Elétrica, sob a responsabilidade do Prof. Adriano O. Andrade, PhD (Orientador), Telefone: (34) 3214 1138, e da mestrande e fonoaudióloga Marlice Fernandes de Oliveira, Telefone: (34) 32291293 e afirmo que foram dadas todas as explicações necessárias para eu tomar essa decisão de livre e espontânea vontade.

Tenho conhecimento que:

- 1-A presente pesquisa objetiva avaliar a voz por meio da realização da análise acústica vocal computadorizada, possibilitando a obtenção de informações sobre as medidas de perturbação vocal, medidas de ruído e qualidade vocal. Este estudo objetiva também avaliar as condições funcionais e orgânicas das pregas vocais por meio da realização de uma videolaringoscopia;
- 2-Os exames serão feitos por profissionais experientes, capacitados e habilitados;
- 3-Serão realizados inicialmente os exames para a avaliação vocal e posteriormente será realizada a videolaringoscopia;
- 4-A avaliação da voz é um exame simples, obtido por meio da captação do sinal vocal pelo microfone;
- 5-A videolaringoscopia é um exame realizado por meio da introdução de um aparelho com micro câmera pela boca ou pelo nariz para fazer a filmagem das pregas vocais;
- 6-Para a realização da videolaringoscopia, injeta-se medicamento vasoconstritor nas narinas antes do procedimento;
- 7-Os procedimentos não causarão nenhum tipo de sangramento, dor ou choque; podendo em alguns casos provocar apenas reflexo nauseante e devido a isso é usado xilocaína spray para aliviar o reflexo de vômito;
- 8-A avaliação será realizada apenas uma vez, com duração de aproximadamente 20 minutos, podendo existir a possibilidade de sua repetição caso haja necessidade;
- 9-Este procedimento se mostra seguro e vem sendo realizado na prática clínica durante anos, onde não se encontra registro sobre presença de danos a saúde;
- 10-Os pesquisadores manterão sigilo sobre minha participação na pesquisa, não serão divulgados: nome, assim como os dados pessoais;
- 11-Minha participação é voluntária e tenho o direito de interrompê-la em qualquer momento;
- 12-É um trabalho voltado para fins científicos, sem fins lucrativos;
- 13-Autorizo a publicação dos resultados do exame da pesquisa em revistas científicas, livros, internet, congressos e etc., onde minha identidade será preservada, não sendo de nenhuma forma mencionada.
- 14-Poderei ter acesso às informações sobre o estudo (resultado da pesquisa, achados encontrados);
- 15-Terei uma cópia deste termo de consentimento antes do início das avaliações.

 (Assinatura do voluntário)

Data do exame: ____/____/____

Comitê de Ética em Pesquisa com Seres Humanos - CEP/UFU
 Universidade Federal de Uberlândia
 Av. João Naves de Ávila, 2121 – B. Sta Mônica, Uberlândia-MG,
 CEP: 38408-100, Fone: 34-3239 4531/4131

ANEXO 1

CÁLCULO DO INTERVALO DE CONFIANÇA

O cálculo do intervalo de confiança da média do sinal, foi realizado utilizando a técnica Bootstrap [75]. Esta técnica é baseada em um processo de re-amostragem que seleciona amostras, aleatoriamente, a partir do espaço amostral original, gerando novos conjuntos de amostras diferente do original, contudo, mantendo suas características estatísticas.

O presente estudo emprega esta técnica com o objetivo de calcular o intervalo de confiança para a média das amostras dos sujeitos disfônicos e eufônicos.

A aplicação do algoritmo, baseado no Bootstrap, para o cálculo do intervalo de confiança para a média é ilustrado pelos passos abaixo [76]:

1- Experimento: Conduza o experimento. Supondo que a amostra seja $X = \{-2.41, 4.86, 6.06, 9.11, 10.20, 12.81, 13.17, 14.10, 15.77, 15.79\}$ de tamanho 10, com média $\mu = 9,946$ estimada levando-se em conta todos os valores de X .

2- Re-amostragem: Utilizando-se um gerador numérico pseudo-aleatório, selecione uma amostra, com reposição, a partir dos 10 valores de X . Deste modo podemos obter a amostra Bootstrap $X^* = \{9.11, 9.11, 6.06, 13.17, 10.20, -2.41, 4.86, 12.81, -2.41, 4.86\}$. Observe que algumas amostras do valor original aparecem mais de uma vez, e outras nem sempre.

3- Cálculo da média para os valores de X^* : a média para todos os 10 valores em X^* é calculada ($\mu^* = 6,54$).

4 - Repetição: Repita o passo 2 N vezes, para obter o total de N médias, $\mu_1^*, \dots, \mu_N^*$. Por exemplo, $N=1000$.

5- Distribuição aproximada de μ^* : Ordene, em ordem ascendente, os valores das médias estimadas no passo 4 para obter: $\mu_{(1)}^* \leq \mu_{(2)}^* \leq \dots \leq \mu_{(1000)}^*$, onde $\mu_{(k)}$ é o k -ésimo menor valor de $\mu_1^*, \dots, \mu_N^*$.

6 – Intervalo de Confiança: O intervalo de confiança desejado, $(1-\alpha) 100\%$, é dado por: $\mu_{(q_1)}^*$, $\mu_{(q_2)}^*$, onde q_1 é a parte inteira de $(N\alpha/2)$ e $q_2 = N - q_1 + 1$. Para $\alpha=0,05$ e $N=1000$, $q_1=25$ e $q_2=976$.

REFERÊNCIAS

1. Pontes P.A.L., et al., *Characteristics of Hoarse, Rough and Normal Voices: Acoustic Spectrographic Comparative Analysis*, in *Revista Brasileira de Otorrinolaringologia*. 2002. p. 12.
2. Casper J., *Reabilitação Vocal para Disfonia por Tensão Muscular*, in *O Melhor que Vi o Ouvi II*, Revinter, Editor. 2000: Rio de Janeiro. p. 16-28.
3. Parraga A., *Aplicação da Transformada Wavelet Packet na Análise e Classificação de Sinais de Vozes Patológicas*, in *Escola de Engenharia Elétrica*. 2002, Universidade Federal do Rio Grande do Sul: Porto Alegre. p. 148.
4. Haddad L., et al., *Vocal Assessment in Patients Submitted to CO2 Laser Cordectomy*. *Revista Brasileira de Otorrinolaringologia*, 2006. **72**(3): p. 295-302.
5. Andrade L.M.O., *Determinação dos Limiares de Normalidade dos Parâmetros Acústicos da Voz*, in *Escola de Engenharia De São Carlos e Escola de Medicina de Ribeirão Preto*. 2003, Universidade de São Paulo: São Carlos. p. 48.
6. Zitta S.M., *Análise Perceptivo-Auditiva e Acústica em Mulheres com Nódulos Vocais*, in *Centro Educacional de Educação Tecnológica do Paraná*. 2005, Faculdade de Engenharia Elétrica e Informática Industrial: Curitiba. p. 111.
7. Nieto A., et al., *Harmonic/noise ratio and spectrographic analysis in vocal abuse pathology*. *Acta Otorrinolaringol Esp*, 1996. **47**(5): p. 370-6.
8. Santos I.R., *Análise Acústica da Voz de Indivíduos na Terceira Idade*, in *Escola de Engenharia de São Carlos e Faculdade de Medicina de Ribeirão Preto*. 2005, Universidade de São Paulo: São Carlos. p. 188.
9. Mota P.H.M., Mourão L.F., and Silvério K.C., *Membrana Laríngea Congênita associada a Vásculodisgenesia: relato de caso*. *Revista Brasileira de Otorrinolaringologia*, 2005. **71**(4): p. 20-24.
10. Figueiredo D.C., et al., *Auditory perceptual, acoustic, computerized and laryngological analysis of young smokers' and nonsmokers' voice*. *Revista Brasileira de Otorrinolaringologia*, 2003. **69**(6): p. 791-799.
11. Sader R.C.M. and Hanayama E.M., *Theoretical considerations on the acoustical approach of the infant voice*, in *Revista CEFAC*. 2004. p. 312-318.
12. Corazza V.R., et al., *Correlação entre os achados estroboscópicos, perceptivoauditivos e acústicos em adultos sem queixa vocal*, in *Revista Brasileira de Otorrinolaringologia*. 2004. p. 30-4.
13. Vieira V.P., Biase N.D., and Pontes P., *Acoustical and perceptive auditive analysis verses glottic coaptation in minimal structural alteration*, in *Acta ORL*. 2005. p. 1-14.
14. Murphy P.J., *Spectral noise estimation in the evaluation of pathological voice*. *Logopedics Phoniatrics Vocology*, 2006. **31**(4): p. 182 - 189.
15. Estella P.M., et al., *Multiparametric Evaluation of Dysphonic Severity*. *Journal of Voice*, 2006. **20**(3): p. 380-390.
16. Jiang J.J., Zhang Yu., and McGilligan C., *Chaos in Voice, From Modeling to Measurement*. *Journal of Voice*, 2006. **20**(1): p. 2-17.
17. Zhang Yu., et al., *Perturbation and Nonlinear Dynamic Analysis of Voices from Patients with Unilateral Laryngeal Paralysis*. *Journal of Voice*, 2005. **19**(4): p. 519-528.

18. Scalassara P.R., Pereira J.C., and Maciel C.D. *Análise do sinal de voz usando processamento de sinais*. in *5º EncoBio de São Carlos*. 2005. São Carlos -SP: Programa de Interunidades em Bioengenharia EESC-USP.
19. Carding P.N., et al., *The reability and Sensitivity to change of acoustic measures of voice quality*. *Clínical Otolaryngol.*, 2004. **29**: p. 538-544.
20. Nemr K., et al., *Comparative analysis of perceptual evaluation, acoustic analysis and indirect laryngoscopy for vocal assessment of a population with vocal complaint*. *Revista Brasileira de Otorrinolaringologia*, 2005. **vol.71**(1): p. 1-9.
21. Araújo S.A., et al., *Normatização de Medidas Acústicas da Voz Normal*, in *Revista Brasileira de Otorrinolaringologia*. 2002. p. 1-9.
22. Master S., et al., *O espectro médio de longo termo na pesquisa e na clínica fonoaudiológica*, in *Pró-Fono Revista de Atualização Científica*. 2006. p. 111-120.
23. Matuck G. R., *Processamento de Sinais de Voz - Padrões Comportamentais por Redes Neurais Artificiais*, in *Instituto Nacional de Pesquisas Espaciais*. 2005, Ministério da Ciência e Tecnologia: São José dos Campos. p. 56.
24. Russo I. and Behlau M., *Percepção da Fala: Análise Acústica do Português Brasileiro*. Vol. 1. 1993, São Paulo: Editora Lovise. 57.
25. Nordemberg M. and Sundberg J., *Effect on LTAS of vocal loudness variation*. *Logoped Phoniatric Vocology*, 2004. **29**(4): p. 183-91.
26. Pinho S. and Pontes P.A.L., *Escala de avaliação perceptiva da fonte glótica*. *Jornal do Conselho Federal de Fonoaudiologia*, 2002. **7**(15): p. 14-15.
27. Medrado R., Ferreira L.P., and Behlau M., *Voice-over: Perceptual and Acoustic Analysis of Vocal Features*. *Journal of Voice*, 2005. **19**(3): p. 340-349.
28. Ledda G.P., et al., *Functional Outcomes after CO2 Laser Treatment of Glottic Carcinoma*. *Laryngoscope*, 2006. **116**(6): p. 1007-1011.
29. Mitev P. and Hadjitodorov S., *Fundamental frequency estimation of voice of patients with laryngeal disorders*. *Information Sciences*, 2003. **156**(1-2): p. 3-19.
30. Rosa O., Pereira C. J. C., and Carvalho A. *Evaluation of Neural Classifiers using Statistic Methods for Identification of Laryngeal Pathologies*. in *Proceedings of the Vth Brazilian Symposium on Neural Networks*. 1998.
31. Lieberman P., *Some acoustics measures of the fundamental periodicity of normal and pathologic larynges*. *Journal of the Acoustics Society of America*, 1963. **35**: p. 344-53.
32. Hecker M.H.L. and Kreul E.J., *Description of the speech of patients with cancer of the vocal folds- Measures of fundamental frequency*. *Journal of the Acoustic Society of America*, 1970. **49**(4): p. 1275-1282.
33. Davis S.B., *Acoustic characteristics of normal and pathological voices*. In. *Speech and language: Advances in Basic and Praticce*. Vol. 1. 1979, New York: Academic Press. 271-335.
34. Horii Y., *Fundamental frequency perturbation observed in sustained phonation*. *Journal of Speech and Hearing Research*, 1979. **22**: p. 5-19.
35. Deal R.E., *Some waveform and spectral features of vowel roughness*. *Journal of Speech and Hearing Research*, 1978. **21**: p. 250-264.
36. Madazio G., Behlau M., and Pontes P., *Análise da proporção harmônico-ruído pré e pós-reabilitação vocal*, in *Tópicos em fonoaudiologia*, M.I.Q. (org.), Editor. 1998, Editora Lovise: São Paulo. p. 169-189.

37. Behlau M.S., *Considerações sobre a análise acústica em laboratórios computadorizados de voz*, in *Fonoaudiologia Atual*, E. Lovise, Editor. 1996, Bompét A. R.,: São Paulo. p. 93-115.
38. Sansone F.E. and Emanuel F.W., *Spectral Noise Levels and roughness severity ratings for normal and simulated rough vowels produced by adult males*. Journal of Speech and Hearing Research, 1970. **13**: p. 472-88.
39. Rocchesso D., *Introduction to Sound Processing*. Vol. 1. 2003, Verona-Italy: PHASAR Srl. 256.
40. Hayes, M.H., *Schaum's outline of Theory and Problems of Digital signal Processing*. Vol. 1. 1999, New York: McGraw-Hill. 436.
41. Douglas C.R., *Tratado de Fisiologia aplicada à Fonoaudiologia*, ed. 1ª. Vol. 1. 2002, São Paulo: Robe Editorial. 774.
42. Russo I.C.P., *Acústica e Psicoacústica aplicadas à Fonoaudiologia*. 2ª edição, ed. E. Lovise. Vol. 1. 1999, São Paulo: Editora Lovise. 262.
43. Pontes P. and Behlau M., *Avaliação e Tratamento das Disfonias*. Vol. 1. 1995, São Paulo: Lovise. 312.
44. Camargo Z., Vilarim G.S., and C. S., *Parâmetros Perceptivo-Auditivos e Acústicos de Longo Termo da Qualidade Vocal de Indivíduos Disfônicos*, in *Revista Cefac*. 2004. p. 189-96.
45. Camargo Z., *Da Fonação à Articulação: Princípios Fisiológicos e Acústicos*. Revista CFFa, 1999. **2**(2).
46. Vieira M.N. *Uma Introdução a Acústica da Voz Cantada*. in *I Seminário Música Ciência Tecnologia: Acústica Musical*. 2005. Belo Horizonte.
47. Filho L.P.B., *O Processo do Envelhecimento e o Comportamento Vocal*, in *CEFAC*. 1999, Centro de Especialização em Fonoaudiologia Clínica: Rio de Janeiro. p. 43.
48. Mac-Kay A.P.M.G., *Linguagem e Envelhecimento*, in *Tópicos em Fonoaudiologia*, E. Lovise, Editor. 1998, Marchesan I.Q.. org: São Paulo. p. 415-420.
49. Boone D.R. and McFarlane S.C., *A voz e a Terapia Vocal*. 5ª ed. Vol. 1. 1994, Porto Alegre: Artes Médicas. 300.
50. Minoru H. and Bless D., *Exame Videostroboscópico da Laringe*, ed. A. Médicas. Vol. 1. 1997, Porto Alegre. 267.
51. Greene M.C.L., *Distúrbios da Voz*. Vol. 1. 1989, São Paulo: Editora Manole Ltda. 489.
52. Dajher M.E., Pereira J.C., and Maciel C.D. *Padrões Visuais Comparativos de vozes normais e patológicas*. in *5º EncoBio de São Carlos*. 2005. São Carlos SP.
53. Yiu E., et al., *Analysing vocal quality of connected speech using Kay's computerized speech lab: a preliminary finding*. Clinical Linguistics & Phonetics, 2004. **14**(4): p. 295-305.
54. Oppenheim A.V., Schafer R.W., and Buck J.R., *Discret-Time signal Processing*, ed. S. Edition. Vol. 1. 1999, New Jersey: Prentice-Hall. 895.
55. Spinelli I.C.P. and Behlau M., *Estudo Comparativo das Medidas de Frequência Fundamental, Jitter e Shimmer em Diferentes sistemas de Análise Vocal*, in *A Voz do Especialista*, Revinter, Editor. 2001, Behlau M.- org: Rio de Janeiro. p. 265-271.
56. Figueiredo L.C., et al., *Vocal behavior during menstrual cycle: perceptual-auditory, acoustic and self-perception analysis*. Revista Brasileira de Otorrinolaringologia, 2004. **70**(3): p. 331-339.

57. Oppenheim A.V., *Signal and Systems*. Signal Processing Series, ed. Prentice-Hall. Vol. 1. 1997, Cambridge: Alan V. Oppenheim Editor. 402.
58. Carvalho M.B., *Tratado de Cirurgia de Cabeça e Pescoço e Otorrinolaringologia*. Tratamento Endoscópico dos Tumores da Laringe, ed. Pinto J.A. Vol. 2. 2001, São Paulo: Atheneu. 911-925.
59. Behlau M.S., et al., *Análise Perceptual Acústica das Vogais do Português Brasileiro Falado em São Paulo*. Acta Awho, 1988. 7: p. 67-73.
60. Ferreira L.P., Lopes D.M.B., and Limongi S.C.O., *Tratado de Fonoaudiologia*, ed. Roca. Vol. 1. 2005, São Paulo. p. 1076.
61. Laver J., *The Phonetic Description of Voice Quality*. Vol. 1. 1980, Cambridge: Cambridge University Press. 186.
62. Sundberg J., *Research on singing voice in retrospect*. Speech, Music and Hearing KTH, 2003. 45: p. 11-22.
63. Lazzer R., *The Vocal Paralysis in the Adult: Glottic Configuration in the Unilateral Paralysis with Involvement of the Recurrent Laryngeal Nerve and its Relationship with the Therapeutic Procedures*, in *Revista Cefac: Atualização Científica em Fonoaudiologia*. 1999. p. São Bernardo do Campo – SP.
64. Andrade L.M.O., et al., *Medidas de Perturbação da Voz: um novo enfoque*, in *Revista Fonoaudiologia Brasil*. 2002. p. 39-46.
65. Behlau M. and Pontes P., *Avaliação e Tratamento das Disfonias*, ed. 1ª. Vol. 1. 1995, São Paulo: Editora Lovise. 312.
66. Boone D.R., P.D., *Comunicação Humana e seus Distúrbios*. 2ª edição ed. 1994, Porto Alegre: Artes Médicas.
67. Paula M.A.P. and Rehder M.I., *Atuação Multidisciplinar em um Caso de Pólipo de Prega Vocal*, in *O melhor Que Vi e Ouvi em Laringe e Voz*, Revinter, Editor. 2001, Behlau M.-org: Rio de Janeiro. p. 243-251.
68. Hamam A.C.S. and Kyrillos L.C.R., *RGE e Trauma vocal: Uma associação que resulta em Granuloma de Prega Vocal*, in *O melhor que Vi e ouvi em Laringe e Voz*, o. Mara Behlau, Editor. 2000, Revinter: Rio de Janeiro. p. 208-217.
69. Lourinho L.A., T.M.S.T., Teixeira M.E.P.A., Barbosa M.T.S.O., *Edema de Reinke: Uma Abordagem Diferenciada*, in *O Melhor que Vi e Ouvi II*, B. M.-org., Editor. 2001, Revinter: Rio de Janeiro. p. 252-256.
70. Colton R.H. and Casper J.K., *Compreendendo os Problemas de Voz*. 1ª ed, ed. A. Médicas. Vol. 1. 1996, Porto Alegre. 386.
71. Ferreira A.E.M. and Fujita R.R., *Cisto em Voz Profissional*, in *O melhor que Vi e Ouvi II- Atualização em Laringe e Voz*, Revinter, Editor. 2000, Mara Behlau, org: Rio de Janeiro. p. 119-127.
72. Feijó A. and Steffen N., *Fonoterapia em um Caso de Sulco Estria e Escara*, in *O melhor que Vi e Ouvi II-Atualização em Laringe e Voz*, B. M.-org., Editor. 2000, Revinter: Rio de Janeiro. p. 285.
73. Brinckmann C.A., Krimberg C.F.D., and Sant'Anna G.D., *Paralisia Unilateral de Prega Vocal: Um Caso com Duas Compensações Atípicas*, in *O Melhor que Vi e Ouvi II- Atualização em Laringe e Voz*, B. M.-org., Editor. 2000, Revinter: Rio de Janeiro. p. 29-33.
74. Martins R.H.G., et al., *Hemangioma laríngeo*, in *Revista Brasileira de Otorrinolaringologia*. 2006.
75. Efron R.T.B., *An Introduction to the Bootstrap*. 1993.

76. Andrade, A.O., *Decomposition and Analysis of Electromyographic Signals*, in *School of Systems Engineering*. 2005, The University Of Reading: Reading. p. 268.