



Giulliana Karla Lacerda Pereira de Queiroz

**ANÁLISE DINÂMICA NÃO LINEAR E ANÁLISE DE
QUANTIFICAÇÃO DE RECORRÊNCIA APLICADAS NA
CLASSIFICAÇÃO DE DESVIOS VOCAIS**

João Pessoa/PB
Junho, 2017

Instituto Federal da Paraíba
Programa de Pós-Graduação em Engenharia Elétrica

Giulliana Karla Lacerda Pereira de Queiroz

**ANÁLISE DINÂMICA NÃO LINEAR E ANÁLISE DE
QUANTIFICAÇÃO DE RECORRÊNCIA APLICADAS NA
CLASSIFICAÇÃO DE DESVIOS VOCAIS**

Dissertação de Mestrado submetida à
Banca Examinadora do Programa de
Pós-Graduação em Engenharia Elétrica
do Instituto Federal da Paraíba como
requisito necessário à obtenção do grau
de Mestre em Engenharia Elétrica.

Orientadora: Prof^ª Dr^ª Silvana Luciene do Nascimento Cunha Costa

Coorientadora: Prof^ª Dr^ª Suzete Élide Nóbrega Correia

João Pessoa/PB
Junho, 2017

Dados Internacionais de Catalogação na Publicação – CIP
Biblioteca Nilo Peçanha – IFPB, *campus* João Pessoa

Q3a Queiroz, Giulliana Karla Lacerda Pereira de.
Análise dinâmica não linear e análise de quantificação de recorrência aplicadas na classificação de desvios vocais / Giulliana Karla Lacerda Pereira de Queiroz. – 2017.
104 f.: il.

Dissertação (Mestrado em Engenharia Elétrica) – Instituto Federal de Educação, Ciência e Tecnologia da Paraíba – IFPB / Coordenação de Pós-Graduação em Engenharia Elétrica, 2017.

Orientadora: Prof^a. Dra. Silvana Luciene do N. Cunha Costa.
Coorientadora: Prof^a. Dra. Suzete Élide Nóbrega Correia.

1. Processamento digital de sinais de voz. 2. Análise dinâmica não linear. 3. Análise de quantificação de recorrência. 4. Avaliação da qualidade vocal. I. Título.

CDU 621.391

GIULLIANA KARLA LACERDA PEREIRA DE QUEIROZ

**ANÁLISE DINÂMICA NÃO LINEAR E ANÁLISE DE
QUANTIFICAÇÃO DE RECORRÊNCIA APLICADAS NA
CLASSIFICAÇÃO DE DESVIOS VOCAIS**

Dissertação aprovada pela Banca Examinadora em: 16/06/2017.

BANCA EXAMINADORA

Silvana Luciene do N. Cunha Costa

Silvana Luciene do Nascimento Cunha Costa, Dra. – IFPB
Orientadora

Suzete Élide Nóbrega Correia

Suzete Élide Nóbrega Correia, Dra. – IFPB
Coorientadora

Joseana Macêdo Fechine Régis de Araújo

Joseana Macêdo Fechine Régis de Araújo, Dra. – UFCG
Membro da Banca

Carlos Danilo Miranda Regis

Carlos Danilo Miranda Regis, Dr. – IFPB
Membro da Banca

Leonardo W. Lopes

Leonardo Wanderley Lopes, Dr. – UFPB
Membro da Banca

João Pessoa-PB
Junho de 2017

Agradecimentos

- » A Deus, a minha fortaleza de todos os dias, aquele que me sustenta fazendo-me não desistir e enviando instrumentos de paz e sabedoria;
- » À minha família, aos meus filhos (Alice, Isaac e Rebeca) e ao meu esposo (Hugo) por todo apoio, estímulo e paciência;
- » Aos meus pais, Marcus e Valteíze, que estiveram presentes todo instante, participando deste sonho desde o princípio e aos meus irmãos, Kelly e Júnior, pelos conselhos e apoio;
- » À minha Orientadora, professora Silvana Costa, que esteve todo instante compartilhando seus conhecimentos com tanta dedicação e paciência, preparando-me não só para seguir uma área acadêmica, mas também para vida com suas imensas palavras de sabedoria;
- » À minha Coorientadora, professora Suzete Correia, por todas as orientações, dedicação e ensinamentos, desde o início do mestrado, enriquecendo os meus conhecimentos;
- » Aos professores examinadores Leonardo Lopes, Joseana Fachine e Carlos Danilo, por avaliar este trabalho de forma tão cuidadosa contribuindo ainda mais para o enriquecimento da pesquisa com suas valiosas colocações;
- » A todos os professores da Pós-Graduação em Engenharia Elétrica (IFPB), pela dedicação ao programa;
- » A todos os meus amigos, em especial Kallynna, Pablo, Evandson, Everaldo, Vinícius, Taciana, Ana Lúcia, Francisco e Andrécia, pelos conselhos e orientações;
- » Ao PRPIG, pelo apoio financeiro ofertado durante o mestrado.

“O sucesso não é medido pelos patamares alcançados, mas pelos obstáculos superados para chegar lá.”

(Booker T. Washington)

Resumo

Disfonia representa qualquer dificuldade na emissão vocal que prejudique a produção natural da voz. Técnicas de processamento digital de sinais vêm sendo empregadas como ferramenta auxiliar na avaliação de desvios vocais, trazendo maior conforto ao paciente. Algumas medidas não lineares, baseadas na teoria do caos, foram empregadas, neste trabalho, em conjunto com medidas de quantificação de recorrência para a análise discriminativa destes desvios. Dois estudos de caso foram realizados nesta pesquisa. No caso 1 foi feita a discriminação de vozes adultas saudáveis e desviadas (rugosidade, soprosidade e tensão) e no caso 2 foi avaliada a discriminação da intensidade dos graus dos desvios vocais de vozes adultas (Grau 1-voz normal, Grau 2 - voz considerada com desvio leve e Grau 3 - voz considerada com desvio moderado). As características da análise dinâmica não linear empregada no processo de classificação foram a Dimensão de Correlação e o Primeiro Mínimo da Função de Informação Mútua. As medidas de quantificação empregadas foram o Determinismo, a Entropia de Shannon, o Comprimento Médio das Linhas Diagonais, o Comprimento Máximo das Linhas Verticais e a Transitividade. O Passo de Reconstrução também foi utilizado no processo de classificação. Por meio dos testes estatísticos, foi avaliado o potencial de cada característica em discriminar os tipos de sinais de voz considerados. Foi utilizada a rede neural MLP (*Multilayer Perceptron*), com o algoritmo de aprendizado supervisionado Gradiente Conjugado Escalonado (SCG), no processo de classificação. Avaliando o desempenho do classificador utilizando as medidas, de forma individual e combinada, foram obtidos, como melhores resultados, uma acurácia média de 91,17% na distinção entre as vozes saudáveis e soprosas com as medidas Transitividade e Passo de reconstrução. Com relação à discriminação entre a intensidade dos graus dos desvios, obteve-se uma acurácia média de 94,5% entre os Graus 1 e 3, com a combinação das medidas Determinismo, Entropia, Transitividade, Primeiro Mínimo da Função de Informação Mútua e o Comprimento máximo das linhas verticais. Os resultados encontrados, nesta pesquisa, indicam que as medidas não lineares, baseadas na teoria do caos, com as medidas de quantificação de recorrência foram eficientes para detectar a presença e o grau dos desvios vocais, podendo ser empregada em métodos de avaliação, triagem e monitoramento vocal.

Palavras-chaves: Análise Não Linear, Análise de Quantificação de Recorrência, Análise acústica, Processamento Digital de Sinais de Voz, Avaliação da qualidade vocal.

Abstract

Dysphonia represents any difficulty in vocal emission that impairs the natural production of voice. Digital signal processing techniques have been used as an auxiliary tool in the evaluation of vocal deviations, bringing greater comfort to the patient. Some nonlinear measures, based on chaos theory, were used in this work, together with recurrence quantification measures for the discriminative analysis of these deviations. Two case studies were conducted in this research. In case 1, the discrimination of adult healthy and deviant voices (roughness, breathiness and strain) was done, and in case 2, the intensity degrees of deviation in adult voices was evaluated (Grade 1-normal voice, Grade 2 - voice considered with slight deviation and Grade 3 - voice considered with moderate deviation). The characteristics of the nonlinear dynamic analysis used in the classification process were the Correlation Dimension and the First Minimum of the Mutual Information Function. The quantification measures employed were Determinism, Shannon's Entropy, the Mean Length of the Diagonal Lines, the Maximum Length of Vertical Lines, and Transitivity. In the characterization process, the potential of each characteristic in discriminating the types of speech signals considered was investigated through statistical tests. The MLP (Multilayer Perceptron) neural network was used, with the Scheduling Conjugate Graded (SCG) supervised learning algorithm, in the classification process. In order to evaluate the performance of the classifier using the measures, individually and in combination, a mean accuracy of 91.17% was obtained in the distinction between healthy and breathy voices with Transitivity and Reconstruction Step Measurements. With respect to the discrimination between the intensity of the degrees of deviations, an average accuracy of 94.5% between Degrees 1 and 3 was obtained, with the combination of the Determinism, Entropy, Transitivity, First Minimum of Mutual Information Functions, Maximum Length of Vertical Lines. The results found in this study indicate that the non-linear measures based on chaos theory with recurrence quantification measures were efficient to detect the presence and degree of vocal deviations and could be used in evaluation, screening and vocal monitoring.

Key-Words: Nonlinear analysis, Recurrence Quantification Analysis, Acoustic Analysis, Digital Processing of Speech Signals.

SUMÁRIO

1. INTRODUÇÃO	19
1.1 Motivação	19
1.2 Justificativa	19
1.3 Objetivos da Pesquisa	22
1.2.1 Objetivo Geral	22
1.2.2 Objetivos Específicos	22
2. PESQUISAS CORRELATAS	24
2.1 Considerações Finais do Capítulo	34
3. FUNDAMENTAÇÃO TEÓRICA	35
3.1 O Sistema de Produção Vocal e Distúrbios Vocais	35
3.1.1 Fisiologia da Produção Vocal	35
3.1.2 Distúrbios Vocais	37
3.1.3 Qualidade Vocal	39
3.1.4 Análise Perceptivo-Auditiva	40
3.2 Sistemas Dinâmicos	41
3.2.1 Métodos das Coordenadas Defasadas	43
3.2.2 Medida de Análise Dinâmica Não Linear	46
3.2.3 Gráficos de Recorrência	47
3.2.4 Medidas de Quantificação de Recorrência	53
3.3 Considerações Finais do Capítulo	55
4. MATERIAL E MÉTODOS	56
4.1 Base de Dados	57
4.1.1 Análise Perceptivo- Auditiva	57
4.2 Metodologia	58
4.2.1 Critérios para Escolha dos Sinais	58
4.2.2 Extração das Características	61
4.2.3 Classificação Rede Neural MLP (<i>Multilayer Perceptron</i>)	62
4.2.4 Medidas de Avaliação de Desempenho (Acurácia, Sensibilidade e Especificidade)	64
4.3 Considerações Finais do Capítulo	65
5. RESULTADOS E DISCUSSÃO	66
5.1 Introdução	66
5.2 Análise Estatística no Estudo de Caso 1	66
5.3 Análise Estatística no Estudo de Caso 2	69
5.4 Classificação	72

5.4.1 Classificação do Estudo de Caso 1	72
5.4.2 Classificação do Estudo de Caso 2	78
5.5 Estudos comparativos	84
5.5.1 Estudo de Caso 1	84
5.5.2 Estudo de Caso 2	87
5.6 Considerações Finais do Capítulo.....	90
6. CONSIDERAÇÕES FINAIS	92
6.1 Contribuições da pesquisa	93
6.2 Sugestões para Trabalhos Futuros	93
REFERÊNCIAS BIBLIOGRÁFICAS	95
APÊNDICE- Classificações segundo a Escala Analógica Visual (EAV)	100

LISTA DE FIGURAS

Figura 1: Esquema da estrutura do aparelho fonador.....	36
Figura 2: Vista posterior dos principais componentes da laringe.....	37
Figura 3: Escala Analógico-Visual (EAV), dada em milímetros.....	41
Figura 4: Função de Informação Mútua e determinação do passo de reconstrução ótimo.	45
Figura 5: Comportamento dos valores da dimensão de correlação para um segmento de 32 ms para um sinal de voz saudável em função da dimensão de imersão e do valor de ϵ , em que as curvas correspondem as diferentes dimensões de imersão ($m = 4, \dots, 10$) de baixo para cima, respectivamente.....	47
Figura 6: Trajetória no espaço de fase bidimensional de um sistema dinâmico arbitrário.	48
Figura 7: Trajetória no espaço de fase de um sistema dinâmico arbitrário. O estado ξ_j passa na vizinhança ϵ de ξ_i , sendo assim um ponto recorrente.....	49
Figura 8: Ruído branco: (a) Retrato de fase e (b) Gráfico de Recorrência.....	50
Figura 9: Sinal senoidal (a) Retrato de fase e (b) Gráfico de recorrência.	50
Figura 10: Gráficos de recorrência de sinais de voz: (a) saudável, (b) rugosa, (c) soprosa e (d) tensa.....	51
Figura 11: Gráficos de recorrência de sinais de voz: (a) saudável, (b) Grau leve (G2) e (c) Grau moderado (G3).	52
Figura 12: Modelo Geral da Metodologia Empregada.....	56
Figura 13: Quantidade de sinais utilizados para cada tipo de desvio.	59
Figura 14: Modelo de um neurônio artificial.....	62
Figura 15: Arquitetura FeedForward de Múltiplas Camadas.	63
Figura 16: Distribuição dos valores médios das medidas associadas à teoria não linear, caso 1.	67
Figura 17: Distribuição dos valores médios das medidas de quantificação de recorrência, caso 1.	68
Figura 18: Distribuição dos valores médios das medidas associadas à teoria não linear, caso 2.	70
Figura 19: Distribuição dos valores médios das medidas de quantificação de recorrência, caso 2.	71

Figura 20: Comparação das médias das acurácias dos resultados individuais e combinados.....	77
Figura 21: Comparação das médias das acurácias dos resultados individuais e combinados.....	83
Figura 22: Comparativo entre os resultados obtidos por Couras e Lopes et al.com os resultados obtidos nesta pesquisa, na classificação individual e combinada entre as discriminações do Estudo de Caso 1.	85
Figura 23: Comparativo entre os resultados obtidos por Vieira e os resultados obtidos nesta pesquisa, na classificação individual e combinadas entre as discriminações do Estudo de Caso 2.	88

LISTA DE QUADROS E TABELAS

Quadro 1: Pesquisas realizadas correlacionadas à avaliação de distúrbios da voz por meio de análise acústica.....	29
Tabela 1: Medidas de quantificação de recorrência em relação as suas estruturas.	53
Tabela 2: Quantidade de sinais utilizados na análise discriminativa entre os sinais de vozes saudáveis e desviados.	59
Tabela 3: Valores médios e desvios padrões de cada classe analisada de acordo com a análise perceptivo-auditiva realizada pela escala EAV.	60
Tabela 4: Quantidade de sinais utilizados para a discriminação com relação às intensidades dos desvios.	60
Tabela 5: Valores médios e desvios padrões de cada classe analisada, com relação as intensidades dos desvios, de acordo com a análise perceptivo-auditiva realizada pela escala EAV.	61
Tabela 6: Potencial discriminativo das medidas da análise dinâmica não linear e de quantificação de recorrência com relação ao Estudo de Caso 1.	66
Tabela 7: Potencial discriminativo das medidas da análise dinâmica não linear e de quantificação de recorrência com relação ao Estudo de Caso 2.	69
Tabela 8: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e desviadas com as medidas empregadas no estudo.	72
Tabela 9: Melhor desempenho na classificação dos sinais de vozes saudáveis e desviadas com as medidas combinadas, empregadas no estudo.	73
Tabela 10: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e soprosas com as medidas empregadas no estudo.	73
Tabela 11: Melhor desempenho na classificação dos sinais de vozes saudáveis e soprosas com as medidas combinadas, empregadas no estudo.	74
Tabela 12: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e rugosas com as medidas empregadas no estudo.	75
Tabela 13: Melhor desempenho na classificação dos sinais de vozes saudáveis e rugosas com as medidas combinadas, empregadas no estudo.	75
Tabela 14: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e tensas com as medidas empregadas no estudo.	76

Tabela 15: Melhor desempenho na classificação dos sinais de vozes saudáveis e tensas com as medidas combinadas, empregadas no estudo.	77
Tabela 16: Melhor desempenho individual na classificação entre as intensidades de Grau 1 (G1) e Grau 2 (G2), com as medidas empregadas no estudo.	78
Tabela 17: Melhor desempenho na classificação das medidas combinadas entre as intensidades do Grau 1 (G1) e Grau 2(G2), com as medidas empregadas no estudo. ...	79
Tabela 18: Melhor desempenho individual na classificação entre as intensidades de Grau 1 (G1) e Grau 3 (G3), com as medidas empregadas no estudo.	79
Tabela 19: Melhor desempenho na classificação das medidas combinadas entre as intensidades do Grau 1 (G1) e Grau 3 (G3), com as medidas empregadas no estudo. ...	80
Tabela 20: Melhor desempenho individual na classificação entre as intensidades de Grau 2 (G2) e Grau 3 (G3), com as medidas empregadas no estudo.	81
Tabela 21: Melhor desempenho na classificação das medidas combinadas entre as intensidades do Grau 2 (G2) e Grau 3(G3), com as medidas empregadas no estudo. ...	81
Tabela 22: Melhor desempenho individual na classificação entre as intensidades de Grau 1 (G1) e Graus 2 e 3 (G2+G3), com as medidas empregadas no estudo.....	82
Tabela 23: Melhor desempenho na classificação entre as intensidades de Grau 1 (G1) e Graus 2 e 3 (G2+G3), com as medidas combinadas empregadas no estudo.....	83
Tabela 24: Resultado das acurácias e das medidas que mais se destacaram na pesquisa de Couras, 2017.	85
Tabela 25: Resultado das acurácias e das medidas que mais se destacaram na pesquisa de Lopes et al., 2016.	86
Tabela 26: Resultado das acurácias e das medidas que mais se destacaram nesta pesquisa.	86
Tabela 27: Resultado das acurácias e das medidas que mais se destacaram na pesquisa de Vieira, 2014.	88
Tabela 28: Resultado das acurácias e das medidas que mais se destacaram nesta pesquisa.	89
Tabela 29: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes soprosas.	100
Tabela 30: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes rugosas.	101
Tabela 31: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes tensas.	102

Tabela 32: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes com predominância de grau 2.....	103
Tabela 33: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes com predominância de grau 3.....	104
Tabela 34: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes da junção dos graus (G2 + G3).....	105

LISTA DE ABREVIATURAS E SIGLAS

ACUR	Acurácia
APQ	<i>Amplitude Perturbation Quotient</i>
CAPE-V	<i>Consensus Auditory Perceptual Evaluation of Voice</i>
CRP	<i>Cross Recurrence Plot</i>
DESV	Desviada
Det	Determinismo
DIV	Medida divergência
EAV	Escala Analógica-Visual
<i>Entr</i>	Entropia de Shannon
ESP	Especificidade
ESV	Escala de Sintomas Vocais
FMMI	<i>First Minimum of Mutual Information</i>
FN	Falso negativo
FP	Falso positivo
GG	Grau geral
G1	Grau 1 (Grupo de sinais de voz com grau normal)
G2	Grau 2 (Grupo de sinais de voz com grau leve)
G3	Grau 3 (Grupo de sinais de voz com grau moderado)
G4	Grau 4 (Grupo de sinais de voz com grau intenso)
G2 e G3	Grau 2 e Grau 3 (Grupo de sinais de vozes com desvios)
GRBA	<i>Grade, Roughness, Breathiness, Asthenia</i>
GRBAS	<i>Grade, Roughness, Breathiness, Asthenia, Strain</i>
GRBASI	<i>Grade, Roughness, Breathiness, Asthenia, Strain, Instability</i>
Hz	Unidade de medida Hertz
IDV	Índice de Desvantagem Vocal
LIEV	Integrado de Estudos da Voz
LAM	Medida laminaridade
LDA	Linear Discriminant Analysis
<i>Lmed</i>	Comprimento médio das estruturas diagonais no gráfico de recorrência
LPC	<i>Linear Predictive Coding</i> - Codificação por predição linear
MEEI	Massachusetts Eye and Ear Infirmary

MLP	<i>Multi-Layer Perceptron</i>
MNL	Medidas Não Lineares
MQR	Medidas de Quantificação de Recorrência
MDVP	<i>Multi-Dimensional Voice Program</i>
MQR's	Medidas de quantificação de recorrência
PM	Primeiro Mínimo da Função de Informação Mútua
PPQ	<i>Pitch Perturbation Quotient</i>
PSO	<i>Particle Swarm Optimization</i>
PVDV	Padrões Visuais da Dinâmica Vocal
QDA	<i>Quadratic Discriminant Analysis</i> - Análise discriminante quadrática
<i>RATIO</i>	Medida da relação entre determinismo e taxa de recorrência
RNAs	Redes Neurais Artificiais
RPDE	Medida da entropia do tempo de recorrência do tipo 1 (Recurrence Probability Density Entropy)
RQA	<i>Recurrence Quantification Analysis</i>
RUG	Rugosa
SDL	Saudável
SENS	Sensibilidade
SNR	Sinal-ruído
SOP	Soprosa
SPSS	<i>Statistical Package for the Social Sciences</i>
SVM	<i>Support Vector Machine</i>
<i>T1</i>	Medida tempo de recorrência do tipo 1
<i>T2</i>	Medida tempo de recorrência do tipo 2
TEN	Tensa
TISEAN	<i>Time Series Analysis</i>
<i>Trans</i>	Transitividade
<i>TREND</i>	Medida tendência
<i>TT</i>	Medida tempo de permanência (Trapping Time)
<i>Vmax</i>	Comprimento máximo das estruturas verticais
VN	Verdadeiro negativo
VP	Verdadeiro positivo

LISTA DE SÍMBOLOS

p	Nível de significância em um teste de hipóteses
N	Números de estados de um sistema
$\{x_i\}$	Série temporal
$\vec{\xi}_l$	Vetores m -dimensionais da série temporal
ε	Raio vizinhança (<i>threshold</i>) no ponto $\vec{\xi}_i$
τ	Passo de Reconstrução
$I(\tau)$	Função de Informação Mútua
t	Tempo
$x(t)$	Série temporal
$x(t + \tau)$	Versão de $x(t)$ defasada de τ
T	Comprimento da série temporal $x(t)$
d	Dimensão fractal do atrator
m	Dimensão de imersão (graus de liberdade).
D_2	Dimensão de Correlação
$Entr$	Entropia de Shannon
Det	Determinismo em um RP
L_{med}	Comprimento Médio das Linhas Diagonais
PM	Primeiro Mínimo da Função de Informação Mútua
V_{max}	Comprimento Máximo das estruturas verticais presentes no gráfico de recorrência
$\mathcal{R}_{i,j}^{m,\varepsilon}$	Definição para gráfico de recorrência
\mathcal{R}^m	Espaço m - dimensional
p	Nível de significância em um teste de hipóteses
N	Número de estados $\vec{\xi}_i$ considerados
N_l	Número total de linhas diagonais no gráfico de recorrência
N_v	Número total de linhas verticais
V_l	Comprimento da linha vertical
$Trans$	Transitividade
$\ \cdot \ $	Norma da vizinhança, comumente a norma euclidiana
$\Theta(\cdot)$	Função de degrau unitário

1. INTRODUÇÃO

1.1 Motivação

A voz é uma ferramenta de grande complexidade e possui inúmeros fatores envolvidos na sua produção. As estruturas anatômicas e os processos fisiológicos envolvidos no sistema de produção da fala são responsáveis pelos diversos fonemas emitidos e, conseqüentemente, carregam informações sobre suas formas, fisiologia e dinâmica (BEHLAU, 2001).

A produção da voz pode ser relacionada aos aspectos anatomofisiológicos, emocionais, orgânicos, ambientais e comportamentais. Portanto, é necessário que a avaliação vocal também siga esse princípio, mapeando os aspectos da produção da voz e correlacionando-os, permitindo, assim, uma visão global de disfonia (LOPES, 2014).

A partir da análise da voz, é possível extrair informações sobre a saúde geral e bem-estar de uma pessoa. Além de revelar quem somos e como nos sentimos, a voz fornece percepções consideráveis sobre a estrutura e função de certas partes do corpo. A presença de algum tipo de disfonia causa mudanças significativas em seus padrões vibratórios, afetando assim a qualidade da produção vocal.

A qualidade vocal é uma interação entre um sinal acústico de voz e um ouvinte. O sinal de voz não possui em si mesmo uma qualidade, mas evoca no ouvinte uma determinada qualidade (KREIMAN *et al.*, 1998 apud DAJER, 2010). Logo, a qualidade vocal é a impressão global que se relaciona com o timbre, a frequência, a amplitude e a composição dos harmônicos da onda sonora produzida, bem como com o modo de vibração das pregas vocais, o tamanho e o formato do trato vocal, a tensão e os tons das paredes faríngeas, podendo ser modificada por fatores anatômicos ou patológicos (FANT, 2004).

1.2 Justificativa

No início do século XIX surgiu a avaliação perceptivo-auditiva, trazendo uma análise subjetiva da voz, exigindo somente o ouvido humano como instrumento de avaliação. Esta análise é uma das formas utilizadas na avaliação da voz, que detecta os distúrbios e caracteriza o tipo de voz, tais como: ressonância, tessitura vocal, qualidade

vocal, velocidade de fala, respiração, *pitch* e tensão geral do aparelho fonador. No entanto, deve ser ressaltada a importância da utilização de instrumentos complementares para a avaliação fonoaudiológica, possibilitando uma avaliação mais precisa da qualidade vocal (CRUZ *et al.*, 2014).

A análise acústica é um método que vem sendo utilizado para avaliar a voz de forma objetiva. Este método pode ser considerado um recurso complementar não invasivo, de baixo custo, quando comparado aos exames realizados por aparelhos invasivos, a exemplo de videoestroboscopia, videolaringoscopia, entre outros.

A análise acústica permite um registro e oferece a possibilidade de detalhamento do processo de geração do sinal sonoro, os quais correspondem a eventos das porções glóticas e supraglótica do trato vocal (CAMPISI, 2000). Além disso, possibilita ainda um tratamento à distância (a voz do paciente pode ser gravada e enviada para um médico em outra localidade em forma de arquivo, tendo assim outra opinião do diagnóstico), auxiliar no tratamento pré-diagnóstico de alterações na laringe, bem como na evolução de um tratamento medicamentoso ou pós-cirúrgico (COSTA, 2012).

A hipótese de que a voz é produzida por um sistema de fonte-filtro (modelo linear de produção da voz), no entanto, carrega intrinsecamente suas limitações. A presença de patologias na laringe acarreta, dependendo da severidade, alterações no padrão anatômico ou no padrão de vibração das pregas vocais (BEHLAU, 2001), impossibilitando a identificação do tipo da patologia/alteração.

Existem várias não linearidades envolvidas na vibração das pregas vocais e na geração da onda glotal (COSTA, 2012 apud ROSA, 2002). Estas incluem efeitos devidos às fortes forças de restauração na colisão das pregas vocais, turbulência no fluxo de ar devido a uma constrição do trato vocal durante a produção de sons surdos, e ao acoplamento não linear entre a fonte e o trato vocal durante a pronúncia de alguns sons sonoros (KUMAR *et al.*, 1996). Devido a tais fatores, os métodos clássicos de análise de dados baseados em um modelo linear têm sido enriquecidos com novos métodos que são derivados da teoria dos sistemas dinâmicos não lineares (JIANG *et al.*, 2006).

As teorias não lineares têm sido utilizadas cada vez mais para a interpretação, explicação e previsão do comportamento de fenômenos físicos e biológicos. Os modelos não lineares ainda não são capazes de explicar toda complexidade presente nos sistemas humanos e é fundamental que estudos surjam de modo a buscar o refinamento

desta técnica. Estes modelos estão ajudando a representar certos tipos de sistemas que os modelos tradicionais lineares não explicam (COSTA, 2012; VIEIRA, 2014).

Do ponto de vista dos sistemas dinâmicos, um sinal de voz pode ser considerado como uma observação do sistema de produção da fala, que pode ser usada para revelar e modelar sua dinâmica usando as técnicas não lineares (COSTA, 2012).

Nesse contexto, o interesse pela análise acústica como uma ferramenta para o auxílio ao diagnóstico e tratamento de doenças na laringe, vem aumentando cada vez mais nos últimos anos (COSTA, 2012; VIEIRA, 2014). A escolha de características relevantes, que possam representar bem as variações existentes no sinal de voz introduzidas pela presença de disfonias e/ou de alterações na laringe pode aumentar a eficiência no processo de detecção de desvios vocais, como também na classificação do tipo e do grau do desvio vocal.

As medidas não lineares, baseada na teoria do caos, e as medidas de quantificação de recorrência vêm ganhando destaque na discriminação entre as vozes saudáveis e alteradas/desviadas, considerando que o sistema de produção vocal seja não linear (JIANG *et al.*, 2006, COSTA, 2012; VIEIRA, 2014; SANTOS, 2015; LOPES, 2016).

A análise de quantificação de recorrência surgiu como forma de potencializar a avaliação do comportamento de séries temporais, proporcionando análise de sinais tanto na forma visual (gráficos de recorrência) quanto paramétrica, a partir das medidas de quantificação de recorrência (VIEIRA, 2014 apud WEBBER *et al.*, 1994; MARWAN, 2003). Embora pequenas perturbações em tais sistemas causem divergência exponencialmente crescente entre os estados, após algum tempo o sistema retorna a um estado que é arbitrariamente próximo a um estado precedente e passa por uma evolução similar. Os gráficos de recorrência permitem visualizar o comportamento recorrente dos sistemas dinâmicos de baixa ou de alta dimensionalidade.

Neste trabalho, o comportamento do sinal de voz será avaliado utilizando a medida não linear de Dimensão de Correlação, as medidas de quantificação de recorrência (Determinismo, Entropia de Shannon, Comprimento Médio das Linhas Diagonais, Comprimento Máximo das Linhas Verticais e Transitividade) e os parâmetros de Passo de Reconstrução e Primeiro Mínimo da Função de Informação Mútua, para a classificação de desvios (rugosidade, soprosidade e tensão), em vozes adultas masculinas e femininas, assim como também a análise da intensidade dos graus dos sinais em vozes adultas. As medidas serão avaliadas de forma individual e

combinadas. A escolha da maioria destas medidas se deu pelo fato de terem apresentado os melhores resultados em trabalhos já existentes (COSTA, 2012; VIEIRA, 2014; LOPES, 2016). Os sinais serão avaliados de forma a indicar a presença ou não do desvio (SDL x DESV) e qual o desvio presente (SDL x SOP, SDL x RUG, SDL x TEN), além de classificar os sinais com relação ao grau do desvio vocal (G1- sinal de voz com grau normal, G2- sinal de voz com grau leve e G3- sinal de voz com grau moderado e intenso). O processo de classificação é realizado empregando redes neurais artificiais, com validação cruzada.

1.3 Objetivos da Pesquisa

1.2.1 Objetivo Geral

Avaliar o desempenho da classificação de sinais de vozes saudáveis e com desvios vocais por meio de análise dinâmica não linear e medidas de quantificação de recorrência.

1.2.2 Objetivos Específicos

- Caracterizar os sinais de voz de pacientes saudáveis e com desvios vocais, assim como o grau do desvio vocal em adultos, por meio das medidas empregadas;
- Avaliar o potencial discriminativo das medidas na classificação entre vozes saudáveis e desviadas (rugosa, soprosa e tensa) e entre os graus de desvio fonatório dos sinais de vozes adultas;
- Analisar os gráficos de recorrência com relação aos desvios e ao grau dos sinais de voz;
- Utilizar técnicas de classificação de padrões como rede neural artificial, para discriminar entre vozes saudáveis e desviadas e entre os graus de desvio vocal;
- Avaliar uma medida ou um conjunto de medidas que melhor representem o desvio vocal dos sinais analisados.

1.3 Organização do Texto

O Capítulo 1 apresenta a parte introdutória do trabalho, abordando o motivo da pesquisa, o estado da arte os objetivos do trabalho.

No capítulo 2 é apresentado o estado da arte, em que são apresentados os principais resultados de algumas pesquisas encontradas na literatura, as quais tratam da investigação de métodos de avaliação dos distúrbios da voz.

Os Capítulos 3 e 4 apresentam a fundamentação teórica necessária para compreensão do método aplicado, abordando conceitos básicos como: sistema de produção vocal e distúrbios vocais, qualidade vocal, análise perceptivo-auditiva e análise dinâmica não linear e a análise de quantificação de recorrência.

No Capítulo 4 é apresentada a metodologia empregada, tais como informações sobre a base de dados utilizada nesta pesquisa, assim como todo o desenvolvimento metodológico.

No Capítulo 5 são apresentados e discutidos os resultados. E, por fim, no Capítulo 6, são apresentadas as considerações finais, destacando as contribuições desta pesquisa e as sugestões para trabalhos futuros.

2. PESQUISAS CORRELATAS

Neste capítulo são apresentados, os resultados de algumas pesquisas encontradas na literatura que buscam desenvolver novos métodos para dar mais precisão à análise de distúrbios da voz.

Jiang *et al* (2006) fizeram um estudo comparativo entre a análise dinâmica não linear (dimensão de correlação) e análise de perturbação (*jitter e shimmer*), em termos dos efeitos do comprimento do sinal analisado, taxa de amostragem e nível de ruído. Nos experimentos foram usadas vogais sustentadas /e/ de vozes saudáveis e de vozes afetadas por patologias na laringe. Segundo os autores, as grandes variações nos valores das medidas de perturbação para os casos dos sinais aperiódicos e caóticos tornam inapropriado o uso dessas medidas para análise de tais sinais. Em contrapartida, a análise dinâmica não linear pode quantificar sinais caóticos. Além disso, segundo os autores, a dimensão de correlação oferece uma medida de análise mais confiável para sinais quase periódicos de menor comprimento, taxa de amostragem mais baixa e níveis maiores de ruído.

Zhang e Jiang (2008) pesquisaram as características de vogais faladas de forma contínua e de forma sustentada de locutores com vozes saudáveis e pacientes com patologias da laringe. Foram usados métodos de perturbação (incluindo *jitter e shimmer*), relação sinal-ruído (SNR), e métodos de dinâmica não linear (dimensão de correlação e entropia de segunda ordem). Os resultados obtidos mostraram que em todos os casos estudados, ou seja, fala contínua, sustentada, saudável e/ou patológica, os sinais de voz têm características dinâmicas de baixa dimensão. Além disso, para as vogais sustentadas, a análise de *jitter, shimmer*, dimensão de correlação, e entropia de segunda ordem revelaram diferenças significativas entre vozes saudáveis e patológicas. Por outro lado, para voz contínua, *jitter e shimmer* não discriminaram, estatisticamente, vozes saudáveis de patológicas, porém uma diferença significativa foi encontrada para a SNR (Relação sinal-ruído), dimensão de correlação, e entropia de segunda-ordem.

Henríquez *et al.* (2009) pesquisaram a possibilidade de se quantificar a qualidade da voz (saudável ou patológica) por meio do uso de seis medidas não lineares, são elas: as entropias de Rényi de primeira e segunda ordem, a entropia de correlação (ou entropia de segunda ordem), a dimensão de correlação, a entropia de Shannon e o

valor do primeiro mínimo da função de informação mútua. A fim de verificar a utilidade dessas medidas são usados dois bancos de dados: um banco de dados não comercial, composto por quatro diferentes tipos de qualidade da voz (voz saudável e três níveis de vozes patológicas - alta, média e baixa severidade), de acordo com o parâmetro do grau de rouquidão (*Grade of hoarseness*), o (G) da escala GRBAS; e um banco de dados comercial, "MEEI Voice Disorders", composto de dois níveis de qualidade vocal (vozes saudáveis e vozes patológicas). Para um classificador baseado em redes neurais MLP (*Multi-Layer Perceptron*) e usando todas as características combinadas, a taxa global de sucesso de um sistema de classificação foi de 82,47% (banco de dados não comercial) e a do outro 99,69% (banco de dados MEEI). Segundo os autores, a diferença entre as taxas de classificação obtidas a partir dos dois bancos de dados deve-se ao fato da não existência, no banco de dados MEEI, de vozes patológicas de baixa severidade.

Dajer (2010), em sua Tese, avalia vozes saudáveis e com alterações patológicas aplicando análise de Padrões Visuais da Dinâmica Vocal (PVDV) em conjunto com a análise acústica e análise perceptivo-auditiva. A autora obteve valores de *jitter*, *shimmer* e Frequência Fundamental. Na análise perceptivo-auditiva foram avaliadas rugosidade, soproidade, tensão e instabilidade. Foi utilizada a técnica de reconstrução de espaço de fase e qualitativamente parâmetros de laços, regularidade e convergência de traçados para analisar a dinâmica dos sinais de voz dos PVDV. A análise qualitativa descrita dos PVDV mostrou resultados diferenciais para cada grupo. Mostrando que configuração, regularidade e convergência dos traçados das vozes normais apresentam uma dinâmica diferencial em relação as vozes disfônicas.

Vaziri *et al.* (2010) utilizaram as medidas de dinâmica não linear: dimensão de correlação, maior expoente de Lyapunov, entropia aproximada, dimensão fractal, medida de complexidade de Ziv-Lempel, bem como duas medidas clássicas de perturbação: o quociente de perturbação de amplitude (APQ - *Amplitude Perturbation Quotient*) e o quociente de perturbação de *pitch* (PPQ - *Pitch Perturbation Quotient*), na tarefa de diagnóstico da qualidade da voz. Nos experimentos foram analisados 51 sinais de vozes saudáveis e 112 sinais de vozes afetadas por patologias nas pregas vocais. Utilizando um classificador baseado em SVM (*Support Vector Machine*), os resultados encontrados revelam as seguintes taxas de classificação: dimensão de correlação, 94,44%; maior expoente de Lyapunov, 88,89%; Ziv-Lempel, 88,80%; dimensão fractal, 87,85%; entropia aproximada, 79,78%, PPQ, 73,33% e APQ, 71,12%. Segundo os autores, isto significa que a dimensão de correlação e o maior expoente de Lyapunov

podem ser usados como um método confiável no diagnóstico clínico de patologias na laringe e em outras aplicações na área.

Sobrinho (2011) analisou sinais de vozes normais e patológicas (com nódulos e edemas), usando seção de Poincaré de vários trechos do espaço de fase reconstruído e calculou a dispersão em relação ao ponto médio da seção em relação à distribuição dos pontos sobre os eixos coordenados. Essa dispersão foi calculada utilizando o conceito estatístico de desvio padrão. A avaliação da ferramenta desenvolvida para esse estudo foi realizada utilizando um sinal senoidal inserindo gradativamente *jitter* e *shimmer*. Os resultados obtidos para esse conjunto de vozes mostraram que o desvio padrão da dispersão e o desvio padrão em relação aos eixos coordenados dos pontos de vozes normais são menores do que os encontrados para vozes com edema e com nódulo.

Costa (2012), em sua Tese trata da caracterização e da classificação de sinais de vozes saudáveis e vozes afetadas por diferentes patologias laríngeas (edema, paralisia e nódulos nas pregas vocais), por meio da análise dinâmica não linear (e teoria do caos), como também por meio da análise de quantificação de recorrência. O autor investigou, por meio de testes estatísticos, o potencial de cada característica em discriminar os tipos de vozes considerados. Na classificação ele empregou a técnica de análise discriminante com função linear ou quadrática, com validação cruzada, sendo considerado um intervalo de confiança de 95% para as médias das taxas de acurácia do classificador. A partir da combinação de características dos conjuntos das medidas de análise não linear (MNL) e das medidas de quantificação de recorrência (MQR), as médias da taxa de acurácia obtidas variaram nos intervalos de confiança: [95,44%; 100%] para a classificação entre vozes saudáveis e patológicas; [94,75%; 100%] entre vozes saudáveis e afetadas por edema, e entre saudáveis e nódulos. Para a classificação entre saudável e paralisia, obteve-se acurácia de 100%. Também foram avaliados os efeitos do uso de vetores híbridos formados por características MNL, MQR e coeficientes extraídos da análise preditiva linear (LPC). Neste caso, as taxas de acurácia variaram nos intervalos de confiança: [95,02%; 97,62%] na discriminação entre vozes afetadas por paralisia e edema; [98,29%; 99,93%] para paralisia *versus* nódulos e [97,98%; 99,84%] para edema *versus* nódulos. Os resultados encontrados indicam que o método utilizado é promissor, podendo ser empregado no desenvolvimento de uma ferramenta computacional para apoio ao diagnóstico de patologias laríngeas.

Galdino (2012), em sua dissertação de mestrado, analisou a dinâmica não linear de vozes de homens adultos, sem queixas vocais, por meio de padrões visuais de

dinâmica vocal (PVDV). Para a análise não linear a autora adotou o método de reconstrução do espaço de fase e gerou um gráfico bidimensional. A mesma analisou o gráfico qualitativamente considerando três aspectos para a análise: número de laços das órbitas, regularidade e espaçamento dos traçados. Cada aspecto foi classificado em uma escala de 4 a 0, da melhor para a pior classificação. Os resultados mostraram que todos os indivíduos avaliados apresentaram uma qualidade de voz adaptada na variabilidade normal de acordo com os resultados da avaliação perceptivo-auditiva do protocolo CAPE-V (ASHA, 2002).

Vieira (2014), em seu trabalho de dissertação, empregou a análise de quantificação de recorrência para a avaliação dos distúrbios da voz. Em uma das suas abordagens foi analisado sinais de vozes infantis diagnosticadas de acordo com o grau de intensidade do desvio vocal. As medidas empregadas foram: Determinismo, Comprimento Máximo das Linhas Diagonais, Entropia de Shannon, Tendência, Laminaridade, Tempo de Permanência, Comprimento Máximo das Linhas Verticais, Transitividade, Relação entre Determinismo e Taxa de Recorrência, Divergência, Comprimento Médio das Linhas Diagonais, Tempo de Recorrência do tipo 1, Tempo de Recorrência do tipo 2, Entropia do Tempo de Recorrência do tipo 1 e o Raio de Vizinhaça. Na avaliação dos desvios vocais das vozes infantis, com segmentação, na discriminação entre vozes saudáveis e vozes consideradas com desvio moderado, a maior acurácia foi de $84,04\% \pm 1,24\%$.

Santos (2015), em sua dissertação avaliou as medidas de energia e entropia dos coeficientes de detalhe da transformada wavelet para a avaliação da qualidade vocal em crianças. Dois estudos de caso foram abordados em sua pesquisa: 1) Análise acústica do grau da intensidade do desvio vocal e 2) Análise acústica da qualidade vocal predominante (rugosidade e soproidade). Para o primeiro caso foi obtida uma acurácia acima de 95% e para o segundo caso as medidas de acurácia foram superiores a 90%, com a utilização de um vetor híbrido de medidas combinadas para ambos os casos. Os resultados da classificação mostraram que a técnica é promissora e pode ser considerada para ser empregada como uma ferramenta para a análise acústica.

Souza (2015) em sua dissertação analisou a discriminação entre sinais de vozes saudáveis e patológicos por meio de duas abordagens, as quais se diferenciam pela forma de extração dos padrões representativos dos sinais: extração dos descritores de textura diretamente das sub-bandas da transformada wavelet; e extração dos descritores de Haralick, a partir da matriz de co-ocorrência. Paralisia, edema de Reinke e nódulos

nas pregas vocais foram às patologias laríngeas consideradas na pesquisa. Os melhores resultados foram obtidos com os descritores de Haralick, empregando redes neurais MLP (Multilayer Perceptron) na classificação, em conjunto com o algoritmo de otimização por enxame de partículas PSO (*Particle Swarm Optimization*) empregado na seleção das características mais representativas, atingindo uma acurácia de 91,79 % entre as vozes saudáveis e vozes afetadas por nódulos nas pregas vocais.

Baravieira (2016), em sua tese, utilizou uma rede neural artificial para a avaliação de vozes rugosas e soprosas. A avaliação perceptivo-auditiva foi feita pela escala visual analógica de 100 mm e pela escala numérica de quatro pontos. A extração de características do sinal de voz foi realizada por meio da Transformada Wavelet Packet e dos parâmetros acústicos: *jitter*, *shimmer*, amplitude da derivada e amplitude do *pitch*; e validação do classificador por meio da parametrização, treino, teste e avaliação das redes neurais artificiais. O melhor desempenho, com a utilização da rede neural, encontrou-se na discriminação da soprosidade e rugosidade e dos seus respectivos graus no qual obteve uma taxa de acerto de 74%.

Lopes *et al.* (2016), em sua pesquisa, analisaram a precisão das medidas de quantificação de recorrência na discriminação de pacientes com e sem problema na voz. Doze medidas de quantificação de recorrência foram extraídas usando 1% de taxa de recorrência: Determinismo, Comprimento Médio e Máximo das Linhas Diagonais, Entropia de Shannon, Laminaridade, Tempo de Permanência, Comprimento Máximo das Linhas Verticais, Tempo de Recorrência do tipo 1 e do tipo 2, Entropia do Tempo de Recorrência do tipo 1, Transitividade e Divergência. Parâmetros relacionados com a topologia recorrente do sistema vocal também foram analisados (dimensão de imersão, passo de reconstrução e raio vizinhança) totalizando um conjunto de 15 medidas. Os valores médios das medidas foram empregados no processo de classificação por meio de análise discriminante quadrática. Os resultados mostram que as medidas de recorrência, isoladas ou combinadas, apresentaram um bom desempenho na discriminação de indivíduos com e sem problema na voz. As combinações das medidas melhoraram na eficácia da classificação, com melhor equilíbrio entre os valores de sensibilidade e especificidade, atingindo uma acurácia de 83,27%.

Couras (2017), em sua dissertação, empregou técnicas de processamento digital de sinais, baseadas no modelo linear de produção da fala, para analisar a qualidade vocal. Foi avaliado o desempenho de medidas tradicionalmente empregadas na análise acústica, tais como frequência fundamental, medidas de perturbação (*jitter* e *shimmer*),

GNE (*Glottal to Noise Excitation Ratio*) e frequências formânticas. Também é avaliado o potencial discriminativo dos coeficientes da análise de predição linear (*Linear Predictive Coding- LPC*), coeficientes cepstrais e mel cepstrais na classificação de desvios vocais (rugosidade, sopro e tensão). Com o auxílio de um classificador baseado em redes neurais artificiais MLP (*Multilayer Perceptron*), é realizada a classificação dos sinais utilizando as medidas extraídas individualmente e de forma combinada. Foram obtidas taxas de classificação de 86% na discriminação entre as vozes soprosas e saudáveis.

No Quadro 1 será apresentada um resumo descritivo relacionados as informações das pesquisas acima.

Quadro 1: Pesquisas realizadas correlacionadas à avaliação de distúrbios da voz por meio de análise acústica.

Autoria	Características Extraídas	Metodologia	Resultados Alcançados
Jiang <i>et al</i> , 2006	1- <i>Jitter</i> ; 2- <i>Shimmer</i> ; 3- Dimensão de correlação; 4- Entropia de segunda ordem.	Base de dados: 53 pacientes saudáveis; 173 pacientes com patologias na laringe. Base da Kay Elemetrics (ELEMETRICS, 1994). Classificador baseado em modelos de misturas gaussianas.	Acurácia da classificação: Saudável x Patologias – 94,07% ± 3,28%.
Zhang e Jiang, 2008	1- <i>Jitter</i> ; 2- <i>Shimmer</i> ; 3- Dimensão de correlação; 4- Relação sinal-ruído; 5- Entropia de segunda ordem.	Base e dados: 23 pacientes saudáveis; 23 pacientes com patologias laringeas. Base construída na pesquisa. Análise com vogal sustentada e fala contínua. Teste estatístico de Mann-Whitney com nível de significância de 5%.	Potencial discriminativo: saudável x patologias (vogal sustentada) – todas as medidas ($p < 0,001$); saudável x patologias (fala contínua) – medidas 3, 4 e 5 ($p < 0,001$).

Continuação do Quadro 1.

Autoria	Características Extraídas	Metodologia	Resultados Alcançados
Henríquez <i>et al.</i> , 2009	1- Entropias de Rényi de primeira e segunda ordem; 2- Entropia de correlação (ou entropia de segunda ordem); 3-Dimensão de correlação; 4- Entropia de Shannon; 5-Primeiro mínimo da função de informação mútua média.	Base de dados: 85 pacientes saudáveis; 57 pacientes com patologias na laringe. Base construída no hospital espanhol “Doctor Negrin” (ALONSO, 1900) Análise perceptivo-auditiva realizada com a escala GRBAS. Classificação com redes <i>multilayer perceptron</i> .	Acurácia da classificação: saudável x patologias – 82,47% _ 3,1%.
Dajer, 2010	1- <i>Jitter</i> ; 2- <i>Shimmer</i> ; 3-Frequência Fundamental	Base de dados: 46 sinais de vozes saudáveis; 45 sinais de voz com disfonias. Base do Comitê para Testes de Função Fonatoria da Sociedade Japonesa de Logopedia e Foniatria (SJLF, 1969). Análise perceptivo-auditiva com a escala GRBAS. Padrões Visuais de Dinâmica Vocal. Testes estatísticos com 5% de significância;	PVDV mostrou melhores resultados na discriminação entre vozes saudáveis x disfônicas.
Vaziri <i>et al.</i> , 2010	1-Quociente de perturbação de amplitude; 2-Quociente de perturbação de frequência; 3- Dimensão de correlação; 4-Maior expoente de Lyapunov; 5- Entropia aproximada; 6- Dimensão fractal; 7- Medida de Ziv-Lempel.	Base de dados: 51 pacientes com vozes normais; 112 pacientes com vozes com patologias laríngeas; Base construída na pesquisa. Classificação com máquina de vetor suporte.	Acurácia da classificação: saudável x patológica: 94,44% (melhor desempenho com a medida de dimensão de correlação).
Sobrinho, 2011	Dispersão da trajetória do espaço de fase em relação a distribuição do ponto médio dos eixos das coordenadas, inserindo <i>Jitter</i> e <i>Shimmer</i> .	Base de dados: 16 vozes normais; 16 vozes com nódulos; 16 vozes com edema. Base do Laboratório de Processamento de Sinais-LPS, da Escola de Engenharia de São Carlos. Vogal sustentada /a/.	Os resultados mostraram que o desvio padrão de dispersão e o desvio padrão dos pontos dos eixos das coordenadas são menores para as vozes normais.

Continuação do Quadro 1.

Autoria	Características Extraídas	Metodologia	Resultados Alcançados
Costa, 2012	1- Dimensão de correlação; 2-Entropia de correlação 3-Entropia de Shannon; 4-Entropia aproximada; 5-Entropia de Tsallis; 6-Maior expoente de Lyapunov; 7-Primeiro mínimo da função de informação mútua; 8-Expoente de Hurst; 9-Determinismo; 10-Comprimento máximo das linhas diagonais; 11-Entropia de Shannon da distribuição das linhas diagonais; 12-Tendência; 13-Laminaridade; 14-Tempo de permanência; 15-Comprimento máximo das linhas verticais; 16-Coefficientes LPC.	Base de dados: 53 pacientes saudáveis; 55 casos com paralisia; 44 casos com edema; 18 casos de nódulos. Base da Kay Elemetrics (ELEMETRICS, 1994). Classificação com análise e discriminante linear (LDA) Análise discriminante quadrática (LQDA).	Acurácia da classificação. saudável x patologias- 98,11% (LDA); saudável x paralisia- 97,74% (LDA); saudável x edema- 99,72% (QDA); saudável x nódulo- 99,95% (QDA); paralisia x edema- 99,88% (QDA); paralisia x nódulos- 99,92% (QDA); edema x nódulos – 99,86 % (QDA). OBS: Principais resultados.
Galdino, 2012	Número de laços das órbitas; Regularidade e espaçamento dos traçados.	Base de dados: 77 sinais de vozes adultas (homens), sem queixa vocal. Método de reconstrução do espaço de fase; Base de dados desenvolvida na pesquisa aprovada pelo Comitê de Ética em Pesquisas da Faculdade de Odontologia de Bauru; Análise perceptivo-auditiva.	Os resultados mostraram que todos os indivíduos avaliados apresentaram uma qualidade de voz adaptada na variabilidade normal, segundo a escala CAPE-V (ASHA, 2002).

Continuação do Quadro 1.

Autoria	Características Extraídas	Metodologia	Resultados Alcançados
Vieira, 2014	1-Determinismo; 2-Comprimento médio e máximo das linhas diagonais; 3-Entropia de Shannon; 4-Laminariedade; 5-Tendência; 6-Tempo de permanência; 7-Comprimento máximo das linhas verticais; 8-Tempo de recorrência tipo 1 e 2; 9-Entropia do tempo de recorrência tipo 1; 10-Transitividade; 11-Divergência; 12-Raio vizinhança; 13-Relação entre determinismo e taxa de recorrência (<i>RATIO</i>).	Base de dados: Laboratório de Voz e Deglutição do Departamento de Fonoaudiologia- UFPB Intensidade dos desvios 10 vozes normais; 70 vozes com desvio leve; 13 vozes com desvio moderado. Análise perceptivo-auditiva realizada com a escala EVA. Vozes infantis.	Acurácia de classificação: Melhor resultado: 84,04% entre as vozes normais e de grau moderado, com a combinação de 8 medidas
Santos, 2015	1-Energia e entropia dos coeficientes de detalhe da transformada wavelet.	Base de dados: 10 vozes normais; 83 vozes com desvios; Base do Laboratório Integrado do Estudo da Voz da Universidade Federal da Paraíba. Análise perceptivo-auditiva utilizando a EAV. Vozes infantis.	Acurácia da classificação: Acurácia acima de 95%.
Souza, 2015	Descritores de texturas dos coeficientes de cada sub-banda obtida pela decomposição wavelet bidimensional.	Base de dados: 51 sinais saudáveis; 43 Edema de Reinke; 18 Nódulos Vocais. Base de dados Kay Elemetrics (ELEMETRICS, 1994). Classificação com redes neurais MLP (<i>Multilayer Perceptron</i>), em conjunto com o algoritmo de otimização por enxames de partículas PSO (<i>Particle Swarm Optimization</i>).	Acurácia da classificação: O melhor resultado foi de 91,79% entre as vozes saudáveis x nódulos.

Continuação do Quadro 1.

Autoria	Características Extraídas	Metodologia	Resultados Alcançados
Baravieira, 2016	1-Característica Transformada <i>Wavelet Packet</i> ;; 2- <i>Jitter</i> ; 3- <i>Shimmer</i> ; 4- Amplitude do <i>pitch</i> .	Base de dados: 123 sinais de vozes saudáveis e vozes com presença de desvio vocal. Avaliação perceptivo-auditiva realizada através da escala EAV e pela escala numérica de 4 pontos. Utilizou um classificador de RNA (Rede Neural Artificial) validado por meio de parametrização.	Na caracterização de sinais de voz com sopro, as melhores características foram <i>jitter</i> , amplitude de <i>pitch</i> e frequência fundamental com taxa de acerto de 74%.
Lopes, 2016	1-Determinismo; 2-Comprimento médio e máximo das linhas diagonais; 3-Entropia de Shannon; 4-Laminariedade; 5-Tempo de permanência; 6-Comprimento máximo das linhas verticais; 7-Tempo de recorrência tipo 1 e 2; 8-Entropia do tempo de recorrência tipo 1; 9-Transitividade; 10-Divergência; 11-Dimensão de imersão; 12-Passo de reconstrução; 13-Raio vizinhança;	Base de dados: 53 pacientes sem problema na voz; 437 pacientes com problema na voz. Base do LIEV (Laboratório Integrado de Estudos da Voz)- UFPB Discriminação entre vozes saudáveis e desviadas. Análise perceptivo-auditiva realizada com a escala EVA.	Acurácia da classificação: Melhor resultado foi de 83,27% na classificação entre as vozes saudáveis e desviadas, com a combinação de 8 medidas.
Couras, 2017	1-Frequência fundamental; 2-Medidas de perturbação (<i>jitter</i> e <i>shimmer</i>); 3-GNE (<i>Glottal to Noise Excitation Ratio</i>) ; 4-Frequências formânticas; Coeficientes da análise de predição linear; 5-Coeficientes cepstrais e mel-cepstrais	Base de dados: 30 pacientes sem problema na voz; 90 pacientes com desvios vocais. LIEV (Laboratório Integrado de Estudos da Voz)- UFPB Discriminação entre vozes saudáveis com relação aos tipos de desvios (soprosa, rugosa e tensa). Análise perceptivo-auditiva realizada com a escala EVA Utilizou vozes femininas	Acurácia de classificação: Melhor resultado: 86% na classificação entre as vozes saudáveis e soprosas, com três medidas.

2.1 Considerações Finais do Capítulo

Nas pesquisas correlacionadas encontradas na literatura, algumas realizam apenas análise estatística com as medidas de análise acústica, umas baseadas no modelo linear da produção de voz e outras baseadas na análise dinâmica não linear. Há, também, técnicas híbridas, contendo medidas de ambos os métodos, com a utilização de medidas tradicionais como: frequência fundamental, *shimmer* e *jitter*. Além disso, não foi encontrado nenhum trabalho na literatura que utilizam as medidas não lineares com as medidas de quantificação de recorrência, com vozes adultas, para a discriminação de vozes saudáveis e com desvios (rugosa, soprosa e tensa) e entre as suas intensidades vocais, como foi empregado nesta pesquisa.

No próximo capítulo será abordada a fundamentação teórica relacionado ao estudo.

3. FUNDAMENTAÇÃO TEÓRICA

A voz é um importante instrumento para o meio da comunicação. Voz, articulação e linguagem são elementos mais importantes da produção da fala. Quando um desses elementos apresenta alguma desordem, a habilidade para se comunicar pode se tornar comprometida (COSTA, 2012). Logo, distúrbios de voz podem produzir um impacto na qualidade de vida de um indivíduo, limitando a comunicação no trabalho, assim como em outros aspectos da vida social (DAJER, 2010).

Neste capítulo, serão abordados os aspectos teóricos relacionados, tais como: a descrição sobre a fisiologia da produção vocal, bem como aspectos relacionados aos distúrbios da voz, à qualidade vocal, à avaliação perceptivo-auditiva e a descrição das medidas utilizadas nesta pesquisa.

3.1 O Sistema de Produção Vocal e Distúrbios Vocais

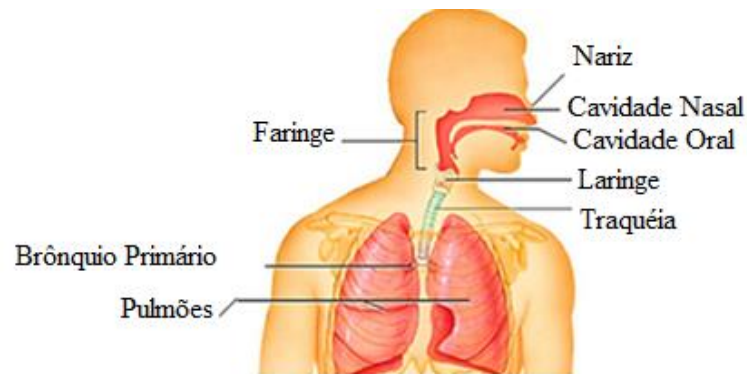
Nesta seção serão abordados aspectos tais como: a descrição sobre a Fisiologia da Produção Vocal, Distúrbios Vocais, Qualidade Vocal e Análise Perceptivo-Auditiva.

3.1.1 Fisiologia da Produção Vocal

O processo inicial da produção vocal é a respiração. Na respiração, nosso diafragma (músculo largo e horizontal localizado logo abaixo dos pulmões) abaixa. Logo, o pulmão, ao inalar o ar, aumenta seu volume. Na expiração, há um relaxamento da musculatura da caixa torácica. A voz é produzida pela vibração das pregas vocais durante a passagem de ar fornecida pelos pulmões. Quando a corrente de ar passa pela região de estreitamento criada pelas pregas vocais em posição fonatória, há vibração das mesmas e produção de ondas sonoras. A produção da voz é decorrente da ação conjunta e harmoniosa de várias estruturas anatômicas. Esse conjunto de estruturas é denominado aparelho fonador ou trato vocal, constituído pela laringe e pelas cavidades de ressonância (BEHLAU, 2001; SCLASSARA, 2009, apud SOUZA, 2015). Na Figura 1 está representada a estrutura do aparelho fonador.

Os pulmões e os músculos associados atuam como uma fonte de ar para excitar a ação conjunta da laringe e das cavidades oral e nasal, a fim de tornar o fluxo de ar em sinal de voz, por meio da atuação dos articuladores (VIEIRA, 2014 apud BEHLAU, 2001; COLTON *et al*, 2006).

Figura 1: Esquema da estrutura do aparelho fonador.

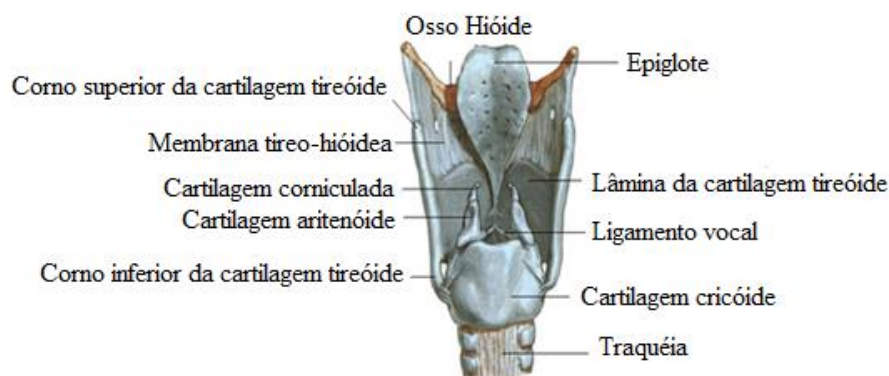


Fonte: SCALASSARA, 2009 apud SOUZA, 2015.

O ar inspirado pelo nariz ou pela boca atravessa a laringe, ao longo do seu percurso, para os pulmões e o mesmo processo acontece durante a expiração com o ar proveniente dos pulmões no seu caminho para o exterior. A respiração só é efetuada com normalidade se o órgão estiver funcionando na sua perfeita normalidade e não oferecer qualquer tipo de resistência à passagem do ar. Todavia, a laringe tem outra missão, igualmente importante, pois é nela que são produzidos os sons que caracterizam a voz.

A laringe pode ser dividida em três espaços: supraglote, glote e infraglote. As pregas vocais estão situadas na laringe e o espaço entre as pregas vocais é chamado de glote. O som da voz é produzido na glote, sendo imediatamente acrescido de ressonância na própria supraglote (SOUZA, 2015 apud COSTA, 2012). Na Figura 2 está representada a vista posterior dos principais componentes da laringe.

Figura 2: Vista posterior dos principais componentes da laringe.



Fonte: Adaptado NETTER, 1997, apud DAJER, 2010.

Entre a cartilagem existente a maior e mais importante cartilagem da laringe, em respeito à fonação, é a cartilagem tireóidea. Essa cartilagem possui o formato de um escudo, sendo composta de duas lâminas laterais, de forma quadrangular, e dois pares de cornos posteriores. O ângulo de união entre as lâminas direita e esquerda, chamado de proeminência laríngea, varia de acordo com o sexo. Na população masculina, observa-se um ângulo de aproximadamente 90° e na população feminina o ângulo é mais aberto, com cerca de 120°. Esta variação é responsável pela definição do tamanho das pregas vocais e contribui na definição da frequência vocal emitida (COSTA, 2012 apud BEHLAU, 2001).

3.1.2 Distúrbios Vocais

A preocupação na área de voz, desde os primórdios da Fonoaudiologia, tem sido a de levantar dados referentes às suas alterações (SANTOS *et al*, 2009). A voz de uma pessoa pode ser modificada em decorrência de algumas doenças que atingem as pregas vocais ou do seu mau uso. Distúrbios da voz podem ser entendidos como sendo processos patológicos que afetam a produção da voz, geralmente causados por fatores anatômicos ou funcionais (emocionais ou psicológicos), levando ao que é comumente chamado, na literatura, de disfonia (emissão da voz sem harmonia ou com dificuldade) e afonia (perda parcial ou total da voz), e ainda resultar em alterações na qualidade da voz, na altura e no tom (VIEIRA, 2014 apud BIREME, 2014).

O sintoma vocal é uma queixa em que o indivíduo relata o que sente de diversos modos, abordando sensações relacionadas à fonação, como dor de pescoço ou de

garganta após conversações prolongadas. Algumas outras queixas podem estar relacionadas às características perceptuais da voz, como garganta arranhada, rouquidão, entre outras, e a prevalência de sintomas vocais pode ser indicativa de suspeita de uma alteração vocal (OLIVEIRA *et al.*, 2004, SAPIR *et al.*, 1993). Os hábitos inadequados com o uso da voz referem-se a comportamentos de produção vocal que não atingem um desempenho vocal eficiente. Alguns fatores estão relacionados ao uso inadequado da voz, são eles (VIEIRA, 2014 apud BEHLAU, 2001; COLTON *et al.*, 2006; HIRSCHBERG *et al.*, 1995):

- Aumento de tensão ou esforço: para atingir certo resultado vocal (por exemplo, aumentar a intensidade da voz), é possível que haja tensão nos músculos ligados ao sistema de produção vocal. Em alguns casos, este comportamento pode ser refletido em dores ao falar. Contudo, um paciente pode ser considerado disfônico quando a sua produção vocal requer mais esforço que o suficiente;
- Uso inadequado de frequência: a frequência fundamental é um parâmetro que difere para vozes masculinas, femininas e infantis. A média da frequência fundamental encontrada no português brasileiro varia de 80 a 150 Hz para homens, 150 a 250 Hz para mulheres e acima de 250 Hz para crianças. Uma alteração vocal pode ser percebida quando a frequência fundamental do indivíduo difere da média do gênero/idade;
- Distúrbios psicogênicos: os aspectos psicológicos que podem contribuir para o surgimento de afonia ou disфонia geralmente surgem na infância. Distúrbios emocionais causados por traumas físicos ou por conflitos familiares, e até mesmo a relação com outras crianças podem ser fatores de incidência de distúrbios da voz que podem acompanhar o indivíduo ao longo dos anos.

Os distúrbios de voz podem produzir um impacto na qualidade de vida de uma pessoa, limitando a comunicação no meio social. Logo, a manifestação de uma alteração vocal deve ser rapidamente diagnosticada. O diagnóstico de uma alteração vocal é um processo extenso que precisa combinar diversas técnicas de avaliação e análise para contemplar a complexidade da dinâmica vocal (DAJER, 2010 apud BEHLAU, 2008).

Nas últimas décadas, com o objetivo de aperfeiçoar o processo de avaliação na clínica vocal, diversas áreas da ciência, medicina e engenharia principalmente, vem desenvolvendo pesquisas relacionadas à fisiologia normal da voz humana e suas alterações. Desde então, surgiram diversas medidas, técnicas e ferramentas de análise como videolaringoscopia, análise perceptivo-auditiva e análise acústica (DAJER, 2010). Estes estudos têm por objetivo fazer uma avaliação mais objetiva e aprofundada do estudo vocal.

3.1.3 Qualidade Vocal

A qualidade vocal está relacionada à frequência, à amplitude e à composição dos harmônicos da onda sonora produzida, bem como ao modo de vibração das pregas vocais, ao tamanho e ao formato do trato vocal, à tensão e ao tônus das paredes faríngeas, podendo ser modificada por fatores anatômicos ou patológicos (VIEIRA, 2014 apud FANT, 2004). Segundo Behlau e Ziemer (1987), a qualidade vocal é o índice mais completo dos atributos da emissão de um indivíduo.

Na clínica vocal as qualidades vocais mais comuns são: as vozes rugosas, quando a impressão auditiva é de irregularidade fonatória; soprosas, na presença de ar não sonorizado ao longo da emissão; e tensas, quando o esforço à fonação é perceptível. A identificação destas qualidades vocais é uma tarefa que se torna complexa em determinadas situações, principalmente pelo fato delas aparecerem, com frequência, de forma associada no mesmo paciente (MADAZIO, 2000).

Os desvios vocais analisados nesta pesquisa podem ser definidos como: (SANTOS, 2015 apud PINHO *et al.*, 2002):

- Rugosidade: irregularidade de vibração das pregas vocais. Engloba o conceito de rouquidão, crepitação, bitonalidade e também aspereza. Assim, a voz é percebida com ruídos presentes em baixa frequência, com característica rugosa e ruidosa. Este parâmetro verifica-se em casos de: fenda glótica, presença isolada de uma alteração orgânica ou fenda de qualquer dimensão com alterações da mucosa das pregas vocais (exemplo: nódulos, pólipos ou edemas).
- Soprosidade: presença de ruído de fundo, audível, que corresponde fisiologicamente à fenda glótica (abertura entre as pregas vocais).
- Tensão: associada a esforço vocal por aumento da adução glótica (hiperfunção),

geralmente inerente ao aumento da atividade da musculatura extrínseca da laringe, com elevação desta. Exemplo: disfonia espasmódica e síndromes de abuso vocal com conseqüente alteração da mucosa (nódulos ou pólipos).

3.1.4 Análise Perceptivo-Auditiva

Os profissionais que trabalham com pacientes que apresentam vozes com distúrbios sabem a dificuldade de documentar com precisão a qualidade de uma voz normal. Apesar de não existir um consenso quanto ao conceito de voz normal e voz alterada (BEHLAU, 2001; YAMASAKI, 2008 apud MARTINS, 2015), uma emissão considerada de boa qualidade pelos ouvintes e produzida sem desconforto pelo falante caracteriza uma voz normal, enquanto uma emissão considerada como ruidosa pelos ouvintes e produzida com certo desconforto pelo falante caracteriza uma disfonia (BEHLAU, 2001 apud MARTINS 2015).

A quantificação da avaliação por meio de escalas é um procedimento comum da análise perceptivo-auditiva, com as quais é possível a classificação do grau de severidade dos distúrbios da voz. Além dessa aplicação, escalas são comumente empregadas em análises clínicas para mensurar subjetivamente fenômenos como dor, ansiedade, náusea, fadiga e dispneia (VIEIRA apud YAMASAKI, 2008).

O procedimento da avaliação perceptivo-auditiva, apesar da subjetividade, é tradicional na rotina clínica fonoaudiológica, considerada padrão ouro para a qualidade vocal. Com o objetivo de reduzir potencialmente a variabilidade e as inconsistências da análise perceptivo-auditiva, foram desenvolvidas diversas escalas de avaliação (MARTINS, 2015).

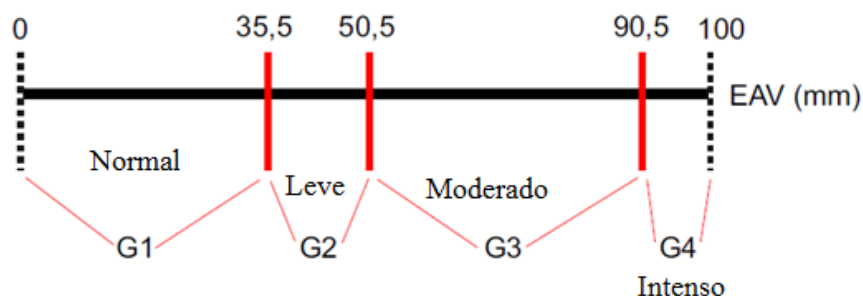
As escalas mais utilizadas, na literatura, com relação à avaliação de distúrbios da voz por meio de análise perceptivo-auditiva são: CAPE-V (*Consensus Auditory-Perceptual Evaluation of Voice*) (KEMPSTER, 2009), GRBAS (*Grade, Roughness, Breathiness, Asthenia, Strain*) (HIRANO, 1981), GRBASI (*Grade, Roughness, Breathiness, Asthenia, Strain, Instability*) (BELE-2005, GAMA-2009), e EAV (Escala Analógico-Visual) (YAMASAKI, 2008; LOPES, 2012).

Nesta pesquisa a análise acústica é relacionada com a análise perceptivo-auditiva, disponibilizada pelo Laboratório do Estudo Integrado da Voz da Universidade Federal da Paraíba, onde o estudo foi realizado a partir da EAV. Essa escala constitui de um intervalo de 100 milímetros, em que os ouvintes devem marcar o ponto que

corresponde à extensão da variação de uma dada característica, como o grau de desvio da voz, por exemplo, (MARTINS, 2015). Estes pontos de corte estão marcados em 35,5 mm, 50,5 mm e 90,5 mm. Os registros da autopercepção vocal e da avaliação do fonoaudiólogo, para fins de análise do resultado da EAV, foram classificados de acordo com os pontos de corte definidos na escala numérica proposta no estudo de Yamasaki *et al.*, 2008.

Na Figura 3 estão definidas as escalas, suas classificações e o grau correspondente.

Figura 3: Escala Analógico-Visual (EAV), dada em milímetros.



Fonte: Adaptado VIEIRA, 2014.

Caso o profissional, treinado, marque na escala algum valor entre 0 e 35,5 mm, a voz é considerada normal. Valores entre 35,5 e 50,5 mm indicam uma voz com desvio leve. Valores entre 50,5 e 90,5 mm indicam uma voz com desvio moderado. Se o profissional marcar algum valor entre 90,5 e 100 mm, o desvio vocal é considerado intenso (YAMASAKI, 2008). Esses quatro intervalos são denotados como grau 1, grau 2, grau 3 e grau 4, respectivamente.

3.2 Sistemas Dinâmicos

Nesta seção serão abordados aspectos relacionados ao sistema dinâmico, assim como os métodos não lineares utilizados na pesquisa, a fim de aplica-los na avaliação dinâmica não linear de um sinal de voz. Entre eles: medidas não lineares, gráficos de recorrências e suas medidas de quantificação.

Para fazer uma análise dinâmica de sinais pode-se utilizar uma modelagem matemática ou fazer uma análise de séries temporais. A primeira técnica implica em se

construir um modelo que possa ser usado para obter informações úteis associadas à produção do fenômeno em estudo. Por outro lado, a análise por séries temporais considera simplesmente uma série temporal escalar, em geral, associada com uma aquisição experimental para entender o comportamento dinâmico de um sistema. O ponto essencial dessa análise é que uma série temporal contém informações sobre variáveis não observáveis do sistema, o que permite a reconstrução do espaço de estados (COSTA, 2012).

Quando se pretende estudar a avaliação vocal é importante fazer uma observação por vários ângulos para poder interpretar as informações que fornece uma voz. Existem diversas técnicas que são utilizadas para o diagnóstico de distúrbios vocais, mas na maioria das vezes estas ferramentas são baseadas em modelos que consideram um sinal de voz um fenômeno linear. Logo, vozes com distúrbios podem apresentar resultados questionáveis (DAJER, 2013). Por este motivo, técnicas não lineares se apresentam como ferramentas promissoras para analisar sinais aperiódicos como é um sinal de voz que apresentam desvios vocais.

A principal característica ligada a um sistema que possui dinâmica caótica é a sua sensibilidade e variação das condições iniciais, ou seja, uma pequena variação nas condições iniciais faz com o que sistema evolua de forma imprevisível (LORENZ, 1963).

Segundo Kantz e Schreiber (2004), a ligação mais direta entre a teoria do caos e o mundo real é a análise de dados de séries temporais, nos termos da dinâmica não linear. Devido, principalmente, ao desenvolvimento de uma base matemática coerente para a descrição do caos em dinâmicas determinísticas não lineares, se tem dado uma atenção significativa à análise dinâmica não linear e ao modelamento determinístico dos dados de uma série temporal (COSTA, 2012; KUMAR *et al.*, 1996).

Uma série temporal é um conjunto de observações ordenadas no tempo, não necessariamente igualmente espaçadas, que apresentam dependência serial, isto é, dependência entre instantes de tempo. A notação usada aqui para denotar uma série temporal é $S_1, S_2, S_3, \dots, S_T$ que indica uma série de tamanho T .

Quando se deseja determinar se a dinâmica que deu origem a uma série temporal é determinística (ou não), há de se lançar mão de outros métodos. Se for determinística, é possível ter condições de desenvolver um modelo capaz de descrever a dinâmica. No caso de comportamento estocástico, ao contrário, trata-se um processo a muitos graus de

liberdade, que só se pode esperar uma descrição estatística do sistema (FIEDLER-FERRARA *et al.*, 1994).

Enquanto a definição das invariantes dinâmicas refere-se ao sistema dinâmico em questão, suas estimativas a partir da série temporal tornam-se possíveis por meio de um conjunto de teoremas formulados por Takens, 1981. Efetivamente, esses teoremas estabelecem que as invariantes dinâmicas estimadas das séries temporais observadas a partir da reconstrução de sua trajetória no espaço de estados usando coordenadas defasadas serão as mesmas do sistema dinâmico em análise sob certas condições.

Na seção seguinte está descrito o método utilizado para a reconstrução do espaço de estados (TAKENS, 1981). Este método foi escolhido por ser o método mais citado na literatura.

3.2.1 Métodos das Coordenadas Defasadas

O método das coordenadas defasadas é baseado no teorema de imersão de Takens, 1981. Ele demonstrou que, com o uso da técnica dos tempos de retardo ou método das coordenadas defasadas, é possível reconstruir certas propriedades topológicas do espaço de estados (atrator) a partir da série temporal, $\{x_i\}$, em que vetores $\vec{\xi}_i$ m -dimensionais são reconstruídos, de acordo com a Equação 1 :

$$\vec{\xi}_i = \{x(t_i), x(t_i + \tau), \dots, x(t_i + (m - 1)\tau)\} \quad (1)$$

em que, $i = 1, 2, \dots, T - (m - 1)\tau$, T é o número total de amostras, m é dimensão de imersão e τ é o passo de reconstrução ou tempo de defasagem (VIEIRA, 2014 *apud* TAKENS, 1981).

Embora o atrator reconstruído não seja idêntico ao original, pode-se demonstrar que as propriedades topológicas são preservadas (difeomorfismo). Além disso, a série temporal de uma única variável é suficiente para a reconstrução desde que a dimensão de imersão seja suficientemente grande (COSTA, 2012 *apud* TAKENS, 1981).

- **Determinação do Passo de Reconstrução (τ)**

A escolha eficiente do Passo de Reconstrução, τ , é bastante importante para a identificação de invariantes geométricos. Caso o τ seja muito pequeno $x(t)$ e $x(t + \tau)$ terão praticamente o mesmo valor. Logo, um atrator reconstruído em um espaço bidimensional com um passo de reconstrução menor do que o adequado fica comprimido em torno da diagonal, ou seja, esse atrator apresentará uma dependência linear entre as variáveis do sistema. Por outro lado, como a trajetória real está restrita a um volume finito do espaço de fases, o passo não pode ser muito grande, sob pena dos vetores reconstruídos serem completamente não-correlacionados cobrindo todo o espaço de fases (FIEDLER-FERRARA, 1994).

Neste trabalho para determinação do passo de reconstrução foi utilizado o método da função de informação mútua média, pois é o mais indicado na literatura, em que o valor do tempo de defasagem ideal corresponde ao primeiro mínimo local (COSTA, 2012).

- **Método da Função de Informação Mútua Média**

Segundo esse método, pode-se garantir a reconstrução de vetores com o menor nível de informação redundante (linearmente independentes), mas, ainda, correlacionados. A teoria da informação procura identificar o quanto de informação se pode ter de uma medida realizada em um determinado instante de tempo t , quando se observa outra medida, do mesmo sinal, em um tempo posterior $t + \tau$ (SOUZA, 2008).

A informação mútua média entre $x(t)$ e sua versão defasada $x(t + \tau)$ é obtida a partir de um histograma de b intervalos de classe (bins), criado para estimar a distribuição de probabilidade dos dados do sinal $x(t)$, dada pela Equação 2:

$$I(\tau) = \sum_{i=1}^b \sum_{j=1}^b P_{i,j}(x(t), x(t + \tau)) \log_2 \left[\frac{P_{i,j}(x(t), x(t + \tau))}{P_i(x(t)) P_j(x(t + \tau))} \right], \quad (2)$$

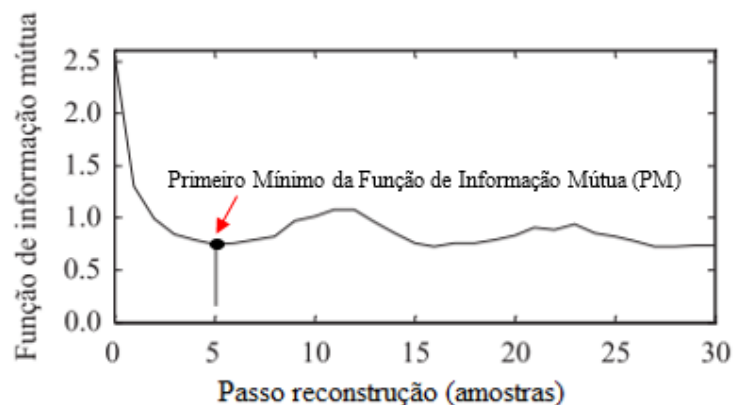
em que $P_i(x(t))$ é a probabilidade de que o sinal $x(t)$ assume um valor dentro do i -ésimo intervalo do histograma, $P_j(x(t + \tau))$ é a probabilidade de que $x(t + \tau)$ esteja no j -ésimo intervalo e $P_{i,j}(x(t), x(t + \tau))$ é a probabilidade de que $x(t)$ esteja, simultaneamente, no i -ésimo intervalo e $x(t + \tau)$ no j -ésimo intervalo. Sendo que $I(\tau) \geq 0$. Se $x(t)$ e $x(t + \tau)$ são

iguais, então $I(\tau)$ é máximo. Por outro lado, se $x(t)$ e $x(t + \tau)$ são completamente independentes, então o argumento do termo logarítmico é a unidade e $I(\tau) = 0$ (VIEIRA, 2014 apud FRASER *et al.*, 1986).

Segundo Fraser *et al.*, 1986, o valor do tempo de defasagem ideal corresponde ao primeiro mínimo local, quando este existir, da função de informação mútua média $I(\tau)$. Desta forma, ao traçar $I(\tau)$ versus, a melhor defasagem corresponde ao valor de τ no primeiro mínimo local. Com isso, garante-se que o espaço de estados reconstruído seja topologicamente equivalente ao espaço de estados do sistema dinâmico original (SAVI, 2006).

Na Figura 4 está ilustrado o comportamento da $I(\tau)$ em função de τ . O ponto preto indica onde ocorre o Primeiro Mínimo da Função de Informação Mútua (PM) e assim a determinação do passo de reconstrução "ótimo" para este segmento ($\tau = 5$).

Figura 4: Função de Informação Mútua e determinação do passo de reconstrução ótimo.



Fonte: Adaptado de COSTA, 2012.

Costa *et al.*, empregaram o Primeiro Mínimo da Função de Informação Mútua, juntamente com outras medidas não lineares e de codificação preditiva linear, na discriminação entre vozes saudáveis e patológicas (Edema, Paralisia e Nódulo), e esta proporcionou uma acurácia de 62,55% na classificação individual entre vozes saudáveis x patológicas. Porém, quando combinada com outras medidas (Dimensão de Correlação, Expoente de Hurst, Entropia de Shannon e Entropia de Tsallis), atingiu uma acurácia de 97,50 % na discriminação de vozes saudáveis x paralisia.

3.2.2 Medida de Análise Dinâmica Não Linear

A medida da análise dinâmica não linear utilizada neste trabalho foi a Dimensão de Correlação, além do valor do Primeiro Mínimo da Função de Informação Mútua e do Passo de Reconstrução.

- **Dimensão de Correlação (D_2)**

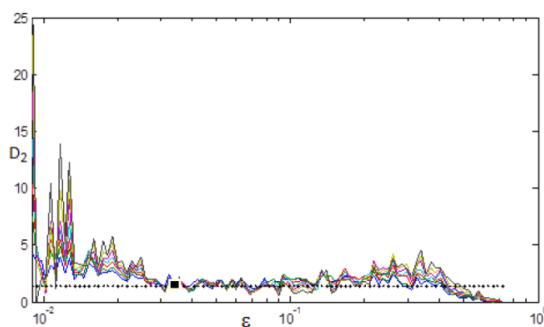
A Dimensão de Correlação (D_2) é uma medida geométrica de uma trajetória no espaço de fase que descreve o quanto dois pontos dessa trajetória são correlacionados. Esta medida representa a complexidade dos sistemas. Assim, um sistema mais complexo apresenta uma maior dimensão, o que significa que um maior número de variáveis de estado é necessário para descrever sua dinâmica. O valor de D_2 de uma fonte de ruído aleatório, por exemplo, não converge com aumento da dimensão de imersão m . Por outro lado, um sistema determinístico apresenta um valor finito para D_2 (JIANG *et al.*, 2006). Logo, a Dimensão de Correlação mede a probabilidade de se encontrar um par aleatório de pontos dentro de um determinado volume elementar e é definida da seguinte forma, dada pela Equação 3.

$$D_2 = \lim_{\varepsilon \rightarrow 0} \frac{\log \sum_{i=1}^{N(\varepsilon)} (f_i)^2}{\log(\varepsilon)} \quad (3)$$

em que o termo $(f_i)^2$ representa a frequência relativa com que dois pontos de um atrator estejam dentro da i -ésima caixa de tamanho ε . Onde ε é o raio em torno de ξ_i e a função $\theta(x)$ satisfaz à condição: $\theta(x) = 1$ para $x > 0$ e $\theta(x) = 0$ para $x \leq 0$. ξ_i e ξ_j , que representam vetores reconstruídos a partir de defasagens na série temporal associada ao sinal de voz.

Na Figura 5 é apresentado o comportamento da dimensão de correlação, D_2 (definida na seção 3.2.4 na página 28), em função de m e de ε , para um segmento de 32 ms de um sinal de voz saudável.

Figura 5: Comportamento dos valores da dimensão de correlação para um segmento de 32 ms para um sinal de voz saudável em função da dimensão de imersão e do valor de ε , em que as curvas correspondem as diferentes dimensões de imersão ($m = 4, \dots, 10$) de baixo para cima,



Fonte: Costa, 2012.

A partir da inspeção visual do comportamento dos valores da dimensão de correlação em função de m (variando de 4 a 10) e de ε é levada a efeito a determinação do valor de ε que oferece o menor coeficiente de variação entre as curvas dentro de um intervalo de convergência. O valor de $D2$ para o segmento é estimado como sendo a média dos valores obtidos de cada curva para este valor de ε . A dimensão de correlação do sinal é tomada como sendo a média calculada sob todos os 30 quadros considerados. No caso do segmento ilustrado na Figura 4.6, o quadrado sobreposto ao gráfico indica os valores encontrados ($\varepsilon = 0,034$ e $D2 = 1,47$) (COSTA, 2012).

3.2.3 Gráficos de Recorrência

Métodos clássicos de análise de dados baseados em modelo linear têm sido enriquecidos com novos métodos que são derivados da teoria do caos, nos últimos anos. Vários métodos de análise têm tentado estimar medidas e propriedades não lineares de processos naturais. Entretanto, a maioria dos métodos da análise não linear exigem séries longas ou estacionárias (MARWAN, 2003).

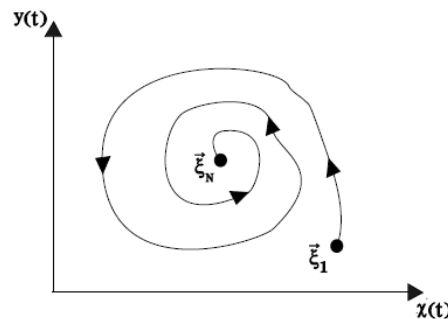
Diante de algumas limitações dos métodos não lineares já existentes, estudos apontam uma nova forma de se avaliar os sistemas caóticos com o uso de gráficos de recorrência e suas medidas de quantificação.

A dinâmica dos sistemas é caracterizada por estados que mudam com o tempo. A partir das variáveis de estado, que correspondem à dimensão do sistema dinâmico, pode-se investigar o comportamento da dinâmica do sistema, por meio do chamado espaço de fase (VIEIRA, 2014).

O Espaço de fase representa a evolução de um sistema dinâmico, cujas dimensões são as variáveis de estado. Cada estado do sistema representa um ponto nesse espaço. A sequência dos estados constitui a trajetória do espaço de fase (VIEIRA, 2014 apud MONTEIRO, 2006). Na Figura 6 está representado um exemplo da trajetória no espaço de fase bidimensional de um sistema dinâmico arbitrário.

Ao longo da trajetória, no espaço bidimensional, das grandezas $y(t)$ e $x(t)$ em que t representa o tempo, estão distribuídos os N estados do sistema. A evolução temporal dessas grandezas é retratada com o estado inicial $\vec{\xi}_1$ e estado final $\vec{\xi}_N$.

Figura 6: Trajetória no espaço de fase bidimensional de um sistema dinâmico arbitrário.



Fonte: VIEIRA, 2014.

Os Gráficos de Recorrência (ECKMANN *et al.*, 1987) têm surgido como mais uma técnica de análise não-linear de sistemas dinâmicos. A análise de quantificação de recorrência surgiu como forma de potencializar as avaliações, a partir do desenvolvimento das medidas de quantificação de recorrência (WEBBER *et al.*, 1994; MARWAN, 2003).

A principal vantagem dos Gráficos de Recorrência é a possibilidade do uso dessa técnica para a análise de séries curtas e não estacionárias (COSTA, 2012).

Geralmente, um experimento não mede todas as variáveis de estado do sistema e tem-se disponível a evolução no tempo de apenas uma variável de estado, representado por uma série temporal, $x(t)$, que representa a trajetória de um sistema dinâmico.

O gráfico de recorrência é definido pela seguinte expressão matemática (MARWAN, 2003), dada pela Equação 4.

$$\mathcal{R}_{i,j}^{m,\varepsilon} = \Theta (\varepsilon - \|\vec{\xi}_i - \vec{\xi}_j\|), \vec{\xi}_i \in \mathcal{R}^m, i, j = 1 \dots N, \quad (4)$$

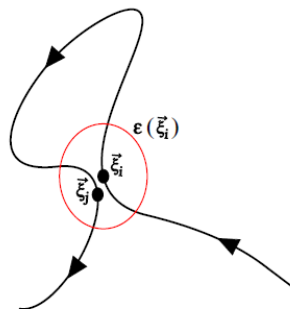
em que:

- » N é o número de estados $\vec{\xi}_i$ considerados;
- » ε é o raio vizinhança (*threshold*) no ponto $\vec{\xi}_i$;
- » $\| \cdot \|$ é a norma da vizinhança, comumente a norma euclidiana;
- » $\Theta (\cdot)$ é a função de degrau unitário;
- » m é a dimensão de imersão (graus de liberdade).

Os gráficos de recorrência são formados por uma matriz quadrada de ordem N , em que N é o número de vetores (estados do sistema) de dimensão m , preenchida por pontos brancos e pretos. O ponto preto, chamado de ponto recorrente, é colocado na matriz de recorrência com coordenadas i e j somente se o estado $\vec{\xi}_i$ for suficientemente próximo ao estado $\vec{\xi}_j$, ou seja, se a distância entre o estado corrente do sistema e o estado a ser comparado for menor que certa distância (raio) ε , fixada no centro do estado corrente (VEIRA, 2014 apud ECKMANN *et al*, 1987). Logo, o raio de vizinhança, ε , é um parâmetro fundamental na construção do gráfico de recorrência, pois se for escolhido um valor muito alto, falsas recorrências podem ser analisadas, caso contrário, pode deixar de ser observados pontos recorrentes no sistema.

A Figura 7 mostra o estado $\vec{\xi}_j$ passando na vizinhança ε de $\vec{\xi}_i$, sendo assim um ponto recorrente (VIEIRA, 2014).

Figura 7: Trajetória no espaço de fase de um sistema dinâmico arbitrário. O estado $\vec{\xi}_j$ passa na vizinhança ε de $\vec{\xi}_i$, sendo assim um ponto recorrente.



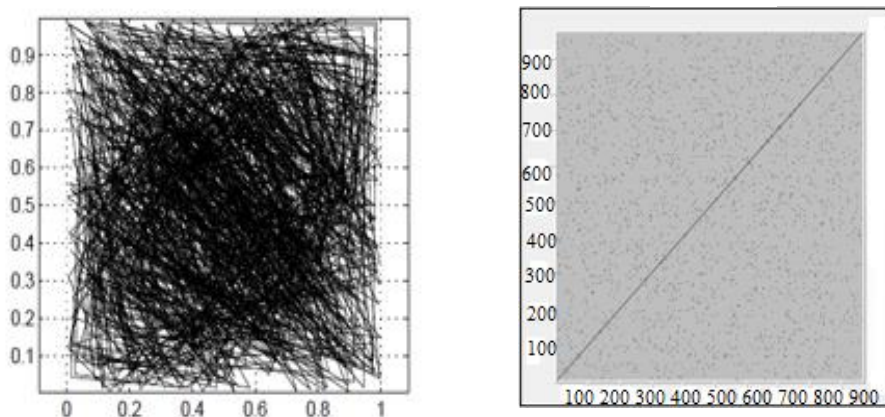
Fonte: VIEIRA, 2014.

Os gráficos de recorrência são compostos por estruturas de pequena escala como pontos e linhas (diagonais, verticais ou horizontais) e estruturas de grande escala ,

também chamadas de texturas. A combinação destas podem formar estruturas variadas (MARWAN *et al.*, 2007).

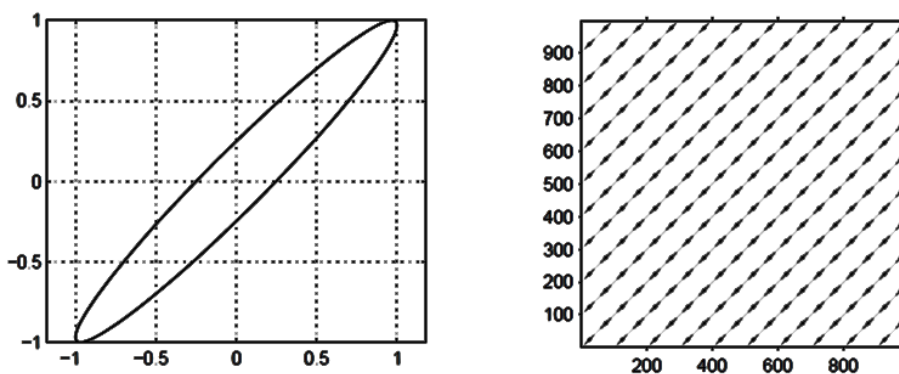
A análise visual dos gráficos de recorrência fornece resultados de forma qualitativa. Por exemplo, processos com comportamentos estocásticos tendem a não apresentar estruturas diagonais (Figura 8). Por outro lado, processos determinísticos causam diagonais mais longas e menos pontos de recorrência isolados (Figura 9) (COSTA,2012).

Figura 8: Ruído branco: (a) Retrato de fase e (b) Gráfico de Recorrência.



Fonte: COSTA, 2012 (a) and VIEIRA, 2014 (b).

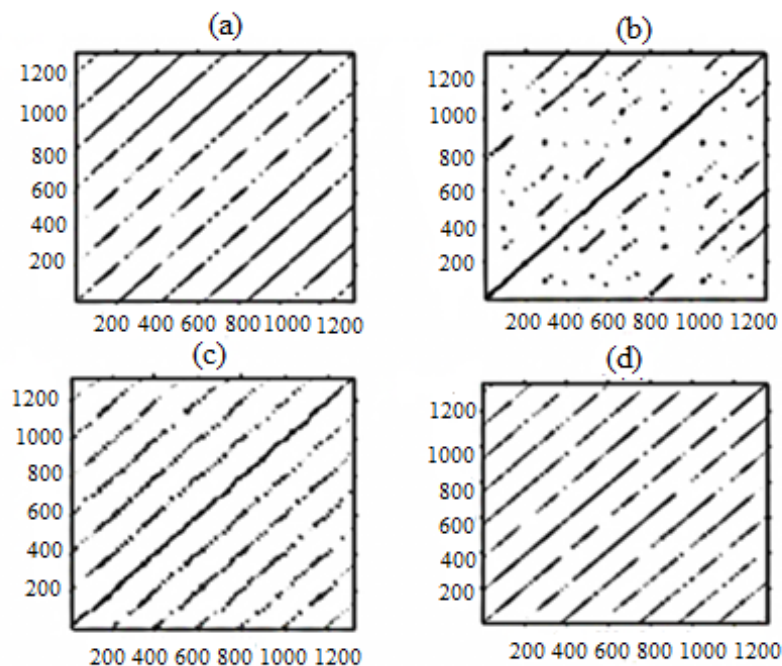
Figura 9: Sinal senoidal (a) Retrato de fase e (b) Gráfico de recorrência.



Fonte: COSTA, 2012.

Nas Figuras 10, cada gráfico de recorrência possui 32 ms do sinal, onde são apresentados exemplos de quatro gráficos de recorrência obtidos a partir de sinais de vozes saudáveis e com desviadas (rugosa, soprosa e tensa).

Figura 10: Gráficos de recorrência de sinais de voz: (a) saudável, (b) rugosa, (c) soprosa e (d) tensa.



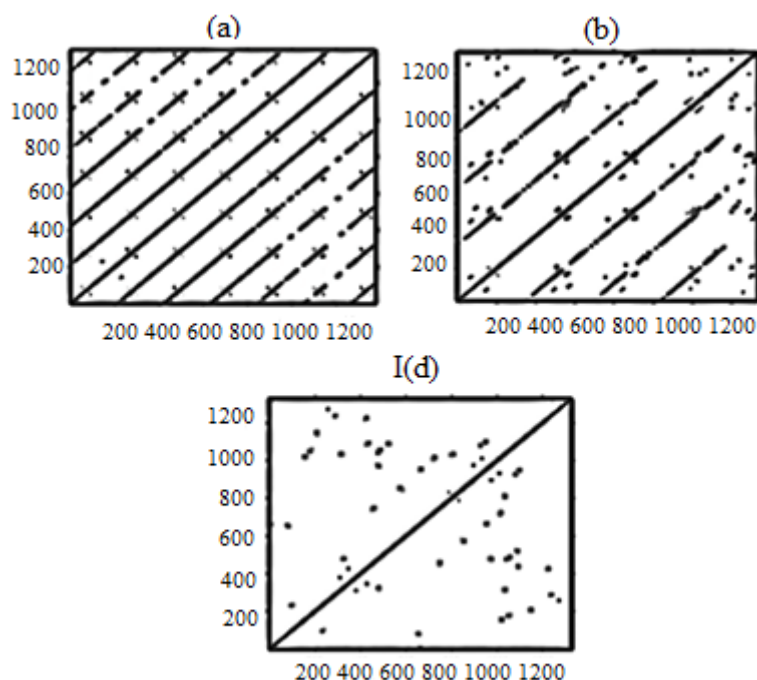
Fonte: Autoria própria.

Pode-se observar que o gráfico de recorrência do sinal de voz saudável, (Fig. 10(a)), apresenta uma formação periódica nas linhas diagonais. No sinal vocal desviado e com predomínio de rugosidade, (Fig.10(b)), pode-se perceber uma menor formação das linhas diagonais (o que parece ter uma relação com a presença de maior aperiodicidade no sinal). A qualidade vocal desviada e predominantemente soprosa, (Fig. 10(c)), apresenta menor número de pontos de recorrência (no gráfico, pode observar praticamente cada ponto individualmente, embora dê a impressão que as linhas diagonais têm maior comprimento). A configuração da qualidade vocal tensa, (Fig. 10(d)), é semelhante à configuração da voz considerada saudável, similar ao que foi encontrado em outros tipos de análise como o “diagrama de desvio fonatório”, em que

as características das vozes tensas são muito próximas às das vozes saudáveis (LOPES, 2016).

Na Figura 11 são apresentados exemplos dos gráficos de recorrência com relação à voz saudável (Grau 1- G1) e às intensidades dos desvios leves (Grau 2 - G2) e moderado (Grau 3 - G3).

Figura 11: Gráficos de recorrência de sinais de voz: (a) saudável, (b) Grau leve (G2) e (c) Grau moderado (G3).



Fonte: Autoria própria.

O gráfico de recorrência do sinal de voz saudável, (Fig. 11(a)), apresenta uma formação periódica nas linhas diagonais, com maior preenchimento que as vozes com o Grau 2 (Fig. 11(b)) e o Grau 3 (Fig. 11(c)). O sinal com Grau 3 possui uma quantidade maior de pontos isolados, onde as formações das linhas diagonais tornam-se inexistentes. Porém, a análise visual dos gráficos de recorrência fornece um resultado subjetivo. Para trazer objetividade na quantificação da dinâmica de um sistema não linear, Webber Jr. e Zbilut (1994) desenvolveram medidas de complexidade, as chamadas Medidas de Quantificação de Recorrência. As medidas inicialmente propostas: taxa de recorrência, determinismo, comprimento máximo das linhas diagonais, Entropia de Shannon da distribuição de frequências dos comprimentos das linhas diagonais e tendência, estão relacionados à densidade dos pontos de recorrência e

as estruturas diagonais presentes nos gráficos de recorrência. Posteriormente, Marwan (2003) propôs outras três medidas de recorrência, desta vez baseadas nas estruturas verticais (ou horizontais), são elas: a Laminaridade, o Tempo de Permanência e o Comprimento Máximo das Estruturas Verticais (COSTA, 2012).

3.2.4 Medidas de Quantificação de Recorrência

Esta análise é um método que tem por objetivo extrair medidas quantitativas dos gráficos de recorrência. De modo geral, a distribuição dos pontos recorrentes pode ser analisada por meio de inspeção visual a partir de três elementos: as linhas diagonais, as estruturas verticais/horizontais e os pontos recorrentes (MARWAN *et al.*, 2005; MARWAN *et al.*, 2007 apud VEIRA, 2014). A Tabela 1 mostra as medidas que foram utilizadas na pesquisa em relação as suas estruturas.

Tabela 1: Medidas de quantificação de recorrência em relação as suas estruturas.

Características	Medidas
Medidas baseadas na formação das estruturas diagonais.	<ul style="list-style-type: none"> • Entropia de <i>Shannon</i> (<i>Entr</i>); • Determinismo (<i>Det</i>); • Comprimento Médio das Linhas Diagonais (<i>Lmed</i>).
Medidas baseadas na formação das estruturas verticais.	<ul style="list-style-type: none"> • Comprimento Máximo das linhas verticais (<i>Vmax</i>).
Medida relacionada à formação e evolução dos pontos recorrentes.	<ul style="list-style-type: none"> • Transitividade (<i>Trans</i>).

A seguir serão descritas a fundamentação teórica das medidas de quantificação de recorrência utilizadas na pesquisa.

- **Transitividade (*Trans*) e Determinismo (*Det*)**

A medida Transitividade (*Trans*) é uma espécie de taxa de recorrência local, onde pode ser utilizada em conjunto com o raio vizinhança para a construção do gráfico de recorrência, como mostra a Equação 5.

$$Trans = \frac{\sum_{i,j,k=1}^N R_{i,j}^{m,\varepsilon} R_{i,k}^{m,\varepsilon} R_{k,i}^{m,\varepsilon}}{\sum_{i,j,k=1}^N R_{i,j}^{m,\varepsilon} R_{k,i}^{m,\varepsilon}} \quad (5)$$

Dessa forma, o raio é escolhido de modo que a Taxa de Recorrência não ultrapasse determinado valor. Essa metodologia é empregada quando se deseja que outras medidas não tenham valores saturados (WEBBER, 2005), como é o caso da medida Determinismo (*Det*), definida pela Equação 6.

$$Det = \frac{\sum_{l=lmin}^N l \cdot P^\varepsilon(l)}{\sum_{ij} R_{i,j}^{m,\varepsilon}} \quad (6)$$

em que $P^\varepsilon(l) = \{l_i, i = 1...Nl\}$, representa a distribuição de frequência dos comprimentos l das Nl estruturas diagonais e $lmin$ o número mínimo de pontos para formar uma linha diagonal dentro do gráfico de recorrência (MARWAN,2003). O Determinismo mede a quantidade dos pontos de recorrência presentes na formação das linhas diagonais em relação a todo o conjunto dos pontos de recorrência.

- **Comprimento médio das linhas diagonais (*Lmed*)**

Está relacionado ao tempo médio em que dois segmentos de uma trajetória estão próximos um do outro e pode ser interpretado como o tempo médio de predição. É calculado da acordo com a Equação 7:

$$Lmed = \frac{\sum_{l=lmin}^N l \times P^\varepsilon(l)}{\sum_{ij} P^\varepsilon(l)} \quad (7)$$

- **Entropia de Shannon (*Entr*)**

Refere-se à entropia de Shannon da distribuição de frequência dos comprimentos das linhas diagonais. Esta medida reflete a complexidade da estrutura determinística no sistema, dada pela Equação 8.

$$Entr = - \sum_{l=lmin}^N p(l) \cdot \ln p(l), \quad (8)$$

em que $p(l) = \frac{P^\varepsilon(l)}{\sum_{l=lmin}^N P^\varepsilon(l)}$.

- **Comprimento máximo das linhas verticais (*Vmax*)**

Esta medida mede o tempo máximo em que o sistema permanece em um estado específico, estando relacionada à duração máxima de um comportamento caótico, dada pela Equação 9.

$$Vmax = \text{máx}(\{Vl: i = 1 \dots Nv\}), \quad (9)$$

em que Nv é o número total de linhas verticais e Vl o comprimento da linha vertical.

3.3 Considerações Finais do Capítulo

Os gráficos de recorrência permitem uma análise da dinâmica não linear dos sistemas. Constituem numa representação bidimensional do comportamento dinâmico que pode até ser multidimensional. A disposição dos pontos no gráfico de recorrência pode formar linhas diagonais ou verticais e/ou horizontais, bem como combinar esses dois tipos de estruturas e formar outras (VIEIRA, 2014).

A evolução da dinâmica de um sistema de produção vocal saudável é mais regular do que a evolução de um sistema de produção vocal com desvio. A formação de estruturas verticais maiores e em maior número é uma característica mais presente nos sinais de voz desviados. Porém, a formação de estruturas no gráfico de recorrência depende do sistema que esteja sendo analisado, e de acordo com a aplicação do método, não é possível uma completa investigação da dinâmica presente na série temporal submetida à avaliação. Além disso, a sua análise é subjetiva e depende da interpretação de quem está avaliando. Logo, as medidas de quantificação de recorrência trazem objetividade, uma vez que quantifica as estruturas dos gráficos de recorrência que podem não ser observadas visualmente.

No capítulo seguinte, será apresentada a metodologia empregada nesta pesquisa, bem como os materiais utilizados no desenvolvimento da mesma.

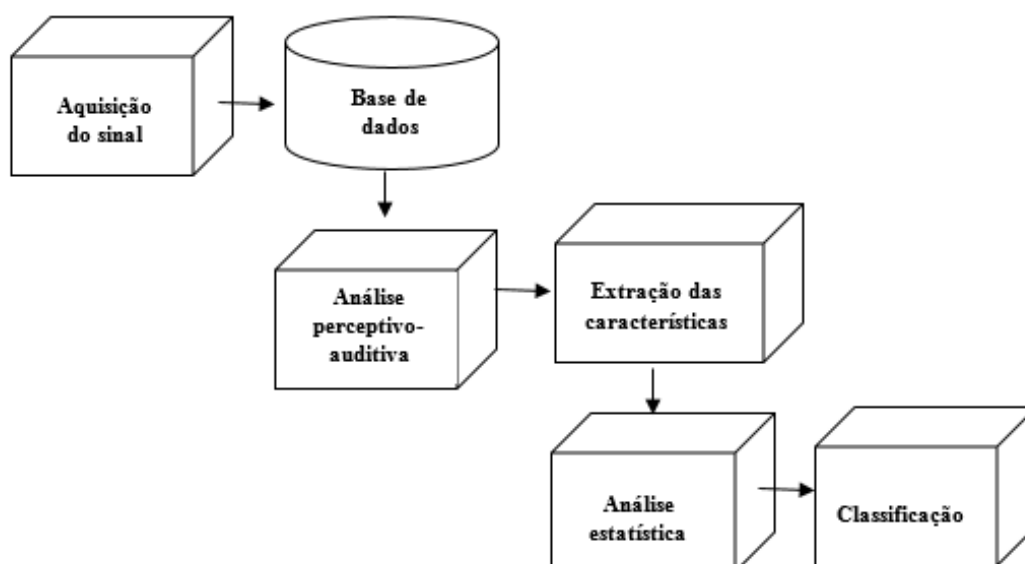
4. MATERIAL E MÉTODOS

Nesta pesquisa, dois casos de estudos foram considerados: 1) Análise acústica de sinais de voz saudáveis e desviadas (rugosidade, soprosidade e tensão), em vozes masculinas e femininas e 2) Análise acústica do grau de intensidade de distúrbios, em vozes adultas.

Os sinais foram avaliados a partir da medida de análise dinâmica não linear (Dimensão de Correlação) e cinco medidas de quantificação obtidas dos gráficos de recorrência (Determinismo, Entropia de Shannon, Comprimento médio das linhas diagonais, Comprimento máximo das linhas verticais e Transitividade). Além de serem utilizados também parâmetros do Primeiro Mínimo da Função de Informação Mútua e o Passo de reconstrução.

Na Figura 12 está apresentado o diagrama, em blocos, da metodologia aplicada..

Figura 12: Modelo Geral da Metodologia Empregada.



Fonte: Autoria própria.

4.1 Base de Dados

O banco de dados dos sinais de voz, utilizados na pesquisa, é proveniente do Laboratório Integrado de Estudos da Voz (LIEV) da Universidade Federal da Paraíba (UFPB), o qual faz parte de um projeto intitulado "Integração das medidas acústicas, perceptivas, fisiológicas e de autoavaliação em pacientes com e sem alteração laríngea", avaliado e aprovado pelo Comitê de Ética em Pesquisa do Centro de Ciências da Saúde/UFPB, com o parecer número 52492/12. Todos os participantes assinaram o termo de consentimento livre e esclarecido.

Foram incluídos no estudo, pacientes com idade superior a 18 anos e inferior a 65 anos, do sexo feminino e masculino, que apresentassem queixa vocal e que já tivessem realizado avaliação laringológica, com laudo otorrinolaringológico por escrito, entregue no momento da coleta.

Esta base, até a data da pesquisa, incluía 800 amostras de sinais de pacientes que responderam aos protocolos de Índice de Desvantagem Vocal (IDV) e a Escala de Sintomas Vocais (ESV) e gravaram a vogal /ε/ sustentada. As vozes foram coletadas em cabine de gravação com tratamento acústico e ruído inferior a 50 dB NPS, com taxa de amostragem de 44000 amostras/segundo, com 16 bits por amostra.

4.1.1 Análise Perceptivo- Auditiva

A análise acústica desta pesquisa foi realizada a partir da análise perceptivo-auditiva, feita pelos profissionais do LIEV, responsáveis pela criação da base de dados, da seguinte forma: A sessão de avaliação perceptivo-auditiva ocorreu em ambiente silencioso e foi realizada por um fonoaudiólogo especialista em voz e com mais de 10 anos de experiência nesta tarefa.

O avaliador foi instruído de que as vozes deveriam ser consideradas normais quando fossem produzidas sem irregularidade, ruído ou esforço observável durante a emissão. Ele também foi instruído de que a rugosidade corresponderia à presença de irregularidade vibratória, a sopro estaria relacionada ao escape de ar audível na emissão e a tensão corresponderia à percepção de esforço vocal. Além disso, o juiz foi treinado com estímulos-âncora, contendo emissões normais e com desvio em diferentes graus, assim como vozes predominantemente rugosas, soprosas e tensas.

Para avaliar a intensidade do desvio vocal (GG-grau geral), o avaliador utilizou uma escala analógica visual (EAV), com uma métrica de 0 a 100 mm. A marcação mais próxima do zero representa menor desvio vocal, e quanto mais próximo dos 100, maiores são os desvios.

Cada emissão da vogal sustentada foi apresentada através de caixa de som, em intensidade confortável autorreferida pelo avaliador. Em seguida, fez-se a identificação da intensidade do desvio vocal e, caso a emissão fosse considerada desviada, o juiz deveria identificar a qualidade vocal predominante (rugosa, soprosa ou tensa).

No final da sessão de avaliação perceptiva, 10% das amostras foram repetidas aleatoriamente, para a análise da confiabilidade do avaliador, utilizando-se Coeficiente Kappa de Cohen. O valor de Kappa foi de 0,79, o que indica boa confiabilidade do avaliador (LOPES *et al.*, 2015).

Ao final, foi realizada uma correspondência entre a EAV e a escala numérica (YAMASAKI *et al.*, 2008), sendo o grau 1 (0-35,5 mm) relacionado às vezes com variabilidade normal da qualidade vocal (VNQV), grau 2 (35,6-50,5 mm) desvio de leve a moderado, grau 3 (50,6-90,5 mm) desvio moderado e grau 4 (90,6-100 mm) desvio intenso.

4.2 Metodologia

Nesta pesquisa, para a análise acústica, foram selecionados trechos de 2 segundos dos sinais, aproximadamente, divididos em 62 segmentos de 32 ms.

4.2.1 Critérios para Escolha dos Sinais

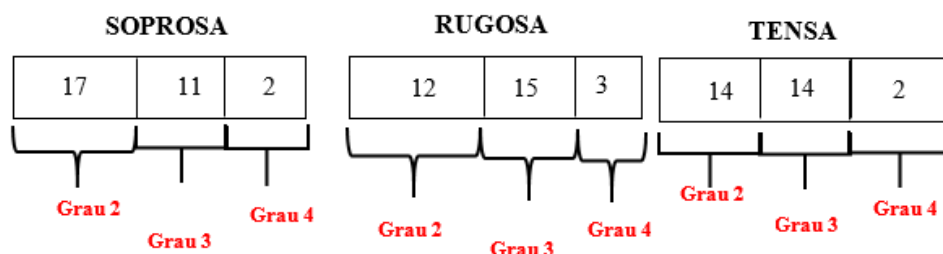
➤ Estudo de Caso 1

Neste Estudo de Caso foram avaliados, do banco de dados, 120 sinais, sendo 30 sinais para cada tipo de voz (saudável, rugosa, soprosa e tensa).

Na escolha destes sinais foram utilizados critérios relacionados aos graus. Ou seja, caso fosse escolhido mais sinais com grau 3, por exemplo, os resultados poderiam ser tendenciosos. Logo, fazendo-se uma proporcionalidade com relação aos graus dos sinais avaliados, pressupõe-se que haja maior confiabilidade nos seus resultados do classificador.

A Figura 13 mostra a quantidade de sinais utilizados referentes a cada tipo de desvio (soprosa, rugosa e tensa), relacionando-se a quantidade de sinais ao grau correspondente de cada desvio.

Figura 13: Quantidade de sinais utilizados para cada tipo de desvio.



Fonte: Autoria própria.

Na Tabela 2 é mostrada a quantidade de sinais utilizados referentes as vozes saudáveis (Grau 1) e desviadas, em que as vozes desviadas contém sinais com desvios de graus leve até intenso (G2 , G3 e G4) .

Tabela 2: Quantidade de sinais utilizados na análise discriminativa entre os sinais de vozes saudáveis e desviados.

Descrição	Grau	Quantidade de sinais utilizados
Saudáveis	G1	30
	G2	15
Desviadas	G3	10
	G4	5

G1:Grau 1- grau normal; G2:Grau 2 - grau leve; G3:Grau 3 - grau moderado; G4: Grupo Grau 4 - grau intenso.

Na Tabela 3 são apresentados os valores médios e seus desvios padrões de acordo com a avaliação perceptivo-auditiva realizada pela escala EAV, para as vozes selecionadas, de acordo com as tabelas apresentadas no Apêndice.

Tabela 3: Valores médios e desvios padrões de cada classe analisada de acordo com a análise perceptivo-auditiva realizada pela escala EAV.

Classes	Médias e desvios padrões
Rugosa	56,90±16,21
Soprosa	53,83±14,39
Tensa	55,5±13,78

➤ Estudo de Caso 2

Para o Estudo de Caso 2 foram avaliados, do banco de dados, 90 sinais, sendo 30 sinais de cada tipo de desvio.

A escolha dos sinais foi feita de forma que não apresentasse uma tendência nos resultados. Portanto, para as discriminações entre os graus dos desvios foram escolhidos os sinais rugosos, soprosos e tensos com a mesma proporcionalidade, pois como os sinais rugosos e soprosos apresentam mais irregularidades em suas linhas diagonais, por exemplo, a escolha predominante destes levaria ao aumento de sua acurácia.

Na Tabela 4, é apresentada a quantidade de sinais utilizados referentes a cada grau de intensidade, relacionando-os aos desvios correspondentes.

Tabela 4: Quantidade de sinais utilizados para a discriminação com relação às intensidades dos desvios.

Grau	Quantidade de sinais utilizados relacionando-os aos desvios		
	Rugosos	Soprosos	Tensos
G2	10	10	10
G3	12	8	10
G2+G3	10	10	10

G2:Grau 1- grau normal; G2:Grau 2 - grau leve; G3:Grau 3 - grau moderado.

Na Tabela 5 são apresentados os valores médios e seus desvios padrões de acordo com a avaliação perceptivo-auditiva realizada pela escala EAV, para as vozes selecionadas, de acordo com as tabelas apresentadas no Apêndice.

Tabela 5: Valores médios e desvios padrões de cada classe analisada, com relação as intensidades dos desvios, de acordo com a análise perceptivo-auditiva realizada pela escala EAV.

Classes	Médias e desvios padrões
G2	42,88±4,31
G3	59,40±8,09
(G2+G3)	51,16±10,96

G2:Grau 1- grau normal; G2:Grau 2 - grau leve; G3:Grau 3 - grau moderado.

4.2.2 Extração das Características

Para a extração das medidas foram empregados os seguintes *softwares*: TISEAN - *Time Series Analysis* v.3.01, onde foram extraídas as medidas de Passo de reconstrução (τ), Primeiro mínimo da função de informação mútua (*PM*) e Dimensão de correlação (*D2*), e o CRP- *Cross Recurrence Plot*, onde foram extraídas as medidas de Determinismo (*Det*), Comprimento médio das linhas diagonais (*Lmed*), Entropia de Shannon (*Entr*), Comprimento máximo das linhas verticais (*Vmax*) e Transitividade (*Trans*). Além de elaboração de rotinas e uso de funções pré-definidas no ambiente Matlab®.

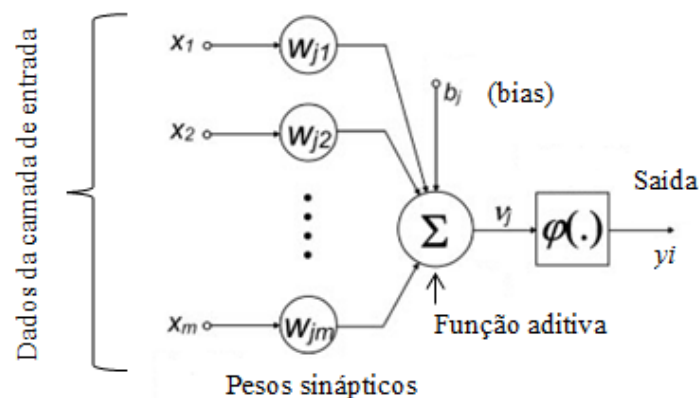
Após a extração das medidas consideradas nesta pesquisa, foram realizados os testes de hipóteses com a utilização do *software Statistical Package for the Social Sciences*- SPSS Statistics®, versão 19.0, com o objetivo de validar as conclusões a respeito do potencial discriminativo de cada medida. O nível de significância considerado em todos os testes estatísticos realizados foi de 5%. Para o Teste de Normalidade, foi feito o Teste de Shapiro-Wilk, pois o número de amostras analisadas foi menor que 50 ($N < 50$). Em seguida, foram aplicados os Testes paramétricos (Teste de Levene), quando os testes de normalidade sugeriram que as distribuições dos resultados fossem normais e Testes não-paramétricos (Teste de Mann-Whitney por se comparar dois grupos), quando a variável de observação não apresenta distribuição normal. Após a etapa da análise estatística, foi realizada a classificação dos sinais (STATISTICS, 2010).

4.2.3 Classificação Rede Neural MLP (*Multilayer Perceptron*)

Uma rede neural consiste em uma máquina que busca modelar a maneira como o cérebro humano realiza uma determinada tarefa (SOUZA,2015). Segundo Haykin (2001), as Redes Neurais Artificiais (RNA) podem ser definidas como sistemas paralelos distribuídos e constituídos de unidades de processamento simples, chamadas "neurônios", que têm capacidade computacional relacionada à aprendizagem e à generalização. Nesse sistema, o conhecimento é adquirido por um processo chamado "treinamento" ou "aprendizagem", que é realizado pelo algoritmo de aprendizagem.

Os neurônios, que são unidades de processamento fundamentais de uma rede neural. A Figura 14 mostra um modelo de neurônio, o qual são constituídos por três elementos básicos: sinapses ou elos de conexão, cada um caracterizado por um peso sináptico que é multiplicado por um sinal de entrada $x_{1...}$, um somador, que realiza a soma dos sinais de entrada, ponderados pelas respectivas sinapses do neurônio; e uma função de ativação, para restringir a amplitude da saída de um neurônio (SOUZA, 2015 apud HAYKIN, 1999).

Figura 14: Modelo de um neurônio artificial.



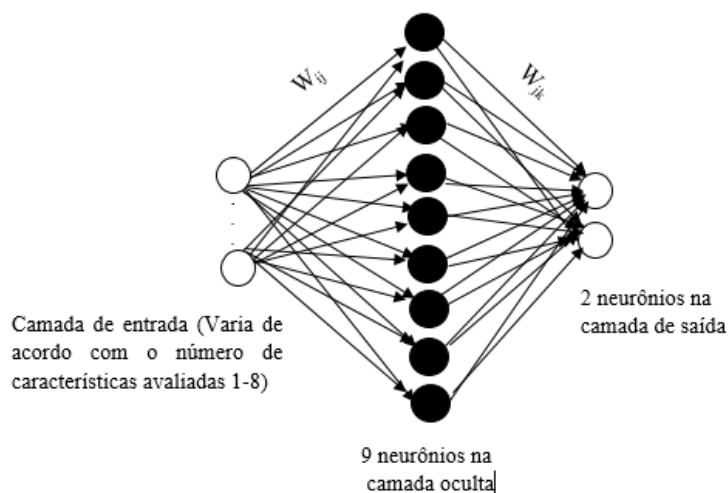
Fonte: Adaptada HAYKIN, 2001.

As redes neurais MLP (*Multilayer Perceptron*), são redes sem realimentação capazes de classificar padrões que não são linearmente separáveis (ZHANG, 2000). Esse tipo de rede neural conhece o ambiente por meio de aprendizado supervisionado, realizado por correção do erro, em que os pesos sinápticos das conexões entre as camadas são ajustados (HAYKIN, 1999).

Dentre os diversos tipos de RNAs, as redes perceptrons de múltiplas camadas (Multilayer Perceptron - MLP) vêm obtendo sucesso em problemas de reconhecimento de padrões. As redes MLPs são ditas totalmente conectadas, pois possuem um conjunto de unidades sensoriais que formam a camada de entrada, uma ou mais camadas ocultas ou intermediárias de neurônios computacionais e uma camada de saída. O treinamento de uma MLP é do tipo supervisionado por meio de algoritmo de retropropagação do erro (error-backpropagation), detalhes sobre o algoritmo são apresentados em (SIMON, 2001).

Nesta pesquisa foi utilizado um classificador MLP (*Multilayer Perceptron*) com o algoritmo de aprendizado supervisionado do Gradiente Conjugado Escalonado (SCG) (COURAS, 2017 apud SIMSEK *et al.*, 2016). Este algoritmo, definido por Moller (MOLLER, 1993), inicia a discussão sobre algoritmos de ordem mais elevada e utilizam informações mais complexas com o objetivo de melhorar o desempenho da rede. A arquitetura utilizada nesta pesquisa foi a rede *FeedForward* (FF) de Múltiplas Camadas por possuir uma camada oculta ou intermediária. Adicionando-se esta camada é possível aumentar a capacidade de processamento de uma rede *FeedForward*. Na Figura 15 é representado o modelo da arquitetura desta rede, utilizada nesta pesquisa.

Figura 15: Arquitetura FeedForward de Múltiplas Camadas.



Fonte: Autoria própria

No processo de classificação, foram realizadas duas fases: treinamento e teste. Na fase de treinamento foram utilizados 09 neurônios na camada oculta, pois foi

observado que seria a quantidade mais eficiente para esta classificação, e 02 na camada de saída. Na camada de entrada, a quantidade de neurônios varia de acordo com o número de características empregadas, podendo variar de 01 a 08. Os pesos W_{ij} e W_{jk} são ajustados automaticamente.

Os sinais foram separados, aleatoriamente, em dois conjuntos distintos, dos quais 2/3 foram utilizados para o treinamento da rede e 1/3 para o teste de eficiência. Em seguida, foram analisados utilizando as características de forma individual e combinada.

A classificação dos sinais foi feita por validação cruzada, em que a rede é executada 10 vezes, a fim de dar maior confiabilidade aos resultados. Esse processo é repetido 10 vezes, para que todos os subconjuntos sejam empregados na fase de teste. Logo, 90% dos dados foram empregados na fase de treino e 10% na fase de teste.

4.2.4 Medidas de Avaliação de Desempenho (Acurácia, Sensibilidade e Especificidade)

Quando se faz um teste em um laboratório para a detecção da condição de saúde, com os possíveis resultados sendo positivo ou negativo, pode haver erros de precisão ou até mesmo de interpretação. Caso o resultado seja positivo, o indivíduo pode ter a doença (Verdadeiro Positivo - VP) ou pode não tê-la (Falso Positivo - FP). Da mesma forma, quando o resultado for negativo, o indivíduo pode não ter a doença (Verdadeiro Negativo - VN) como pode tê-la (Falso Negativo - FN), (COSTA, 2012).

No processo de classificação foram utilizadas três medidas de avaliação de desempenho: acurácia, sensibilidade e a especificidade (SAVI, 2006).

A Acurácia (ACUR) mede a taxa de classificação correta global, refletindo a capacidade do classificador de identificar corretamente quando há e quando não há a presença do desvio vocal.

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN} . \quad (10)$$

A Sensibilidade (SENS) mede a capacidade do classificador em identificar a presença do distúrbio quando ele de fato existe.

$$Sensibilidade = \frac{VP}{VP+FN} . \quad (11)$$

A Especificidade (ESP) mede a capacidade do classificador em identificar corretamente a ausência do distúrbio quando de fato ele não existe.

$$\textit{Especificidade} = \frac{VN}{VN+FP} . \quad (12)$$

O classificador apresenta bom desempenho caso seja capaz de obter altos valores para acurácia, sensibilidade e especificidade.

4.3 Considerações Finais do Capítulo

Este capítulo apresentou a metodologia aplicada na pesquisa. Foram analisados dois casos de estudo um em relação à discriminação das vozes saudáveis e com desvios vocais e o outro com relação a discriminação da intensidade destes desvios. Para ambos os casos foi utilizada a rede neural MLP, utilizando três medidas de avaliação do desempenho: acurácia, sensibilidade e especificidade.

5. RESULTADOS E DISCUSSÃO

5.1 Introdução

Neste Capítulo, são apresentados os resultados experimentais, obtidos para as discriminações referentes aos dois casos de estudo, por meio das seguintes medidas: Dimensão de Correlação, Primeiro Mínimo da Função de Informação Mútua e Passo de reconstrução. Além disso, são apresentados os resultados obtidos considerando as seguintes medidas de quantificação de recorrência: Determinismo, Entropia de Shannon, Transitividade, Comprimento Médio das Linhas Diagonais e Comprimento Máximo das Linhas Verticais. São apresentados, também, os resultados dos testes estatísticos com o objetivo de evidenciar o potencial discriminativo de cada uma dessas características. Em ambos os casos de estudo a análise foi realizada com as medidas extraídas com os sinais segmentados. Foi avaliado se cada medida empregada apresentava potencial estatístico discriminativo entre as classes envolvidas.

5.2 Análise Estatística no Estudo de Caso 1

Os resultados do potencial discriminativo das medidas de quantificação de recorrência, juntamente com as outras medidas não lineares tradicionais, estão apresentados na Tabela 6.

Tabela 6: Potencial discriminativo das medidas da análise dinâmica não linear e de quantificação de recorrência com relação ao Estudo de Caso 1.

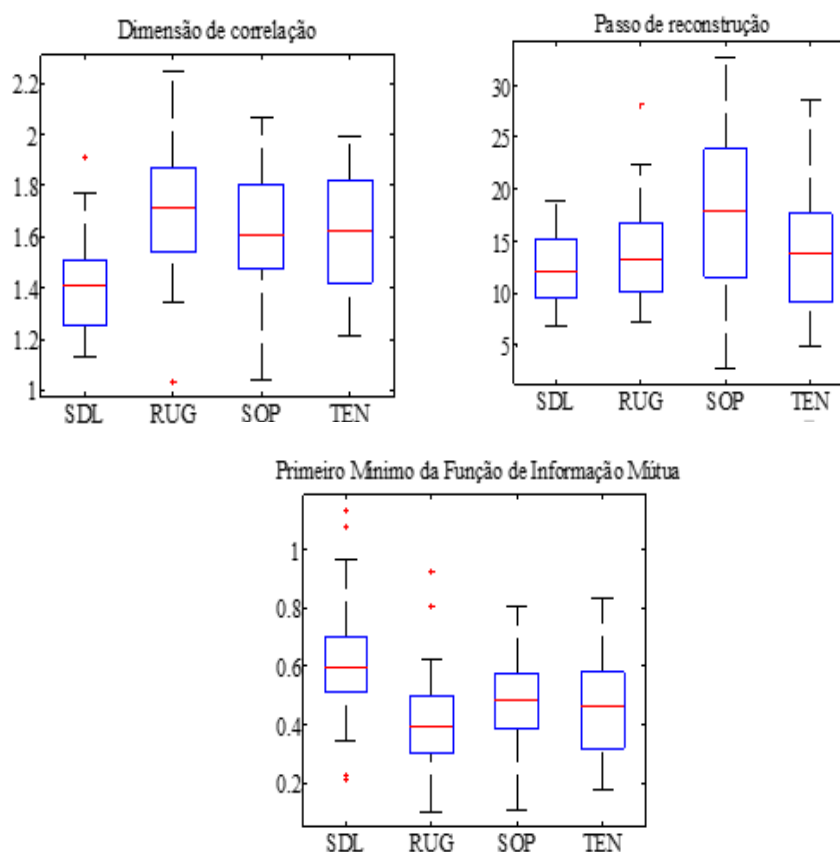
Casos de Discriminação	Medidas							
	τ	PM	D_2	Det	$Trans$	$Lmed$	$Entr$	$Vmax$
SDLxDESV	$p=0,20$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$
SDLxSOP	$p<0,001$	$p=0,005$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,05$	$p<0,001$	$p=0,003$
SDLxRUG	$p=0,23$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$
SDLxTEN	$p=0,21$	$p=0,005$	$p<0,001$	$p=0,002$	$p<0,001$	$p=0,048$	$p<0,017$	$p=0,16$

Na discriminação entre os sinais de vozes saudáveis e o conjunto de vozes tensas, apenas as medidas de Passo de Reconstrução (τ) e Comprimento Máximo das Linhas Verticais ($Vmax$), não apresentaram diferença estatística.

A medida de Passo de Reconstrução (τ), só proporcionou diferença estatística entre os sinais de vozes saudáveis e soprosas. Assim, é possível concluir que esta medida pode ser usada para diferenciar vozes saudáveis e soprosas. Entretanto, ela não se mostrou capaz de evidenciar diferenças estatísticas nas outras classes analisadas.

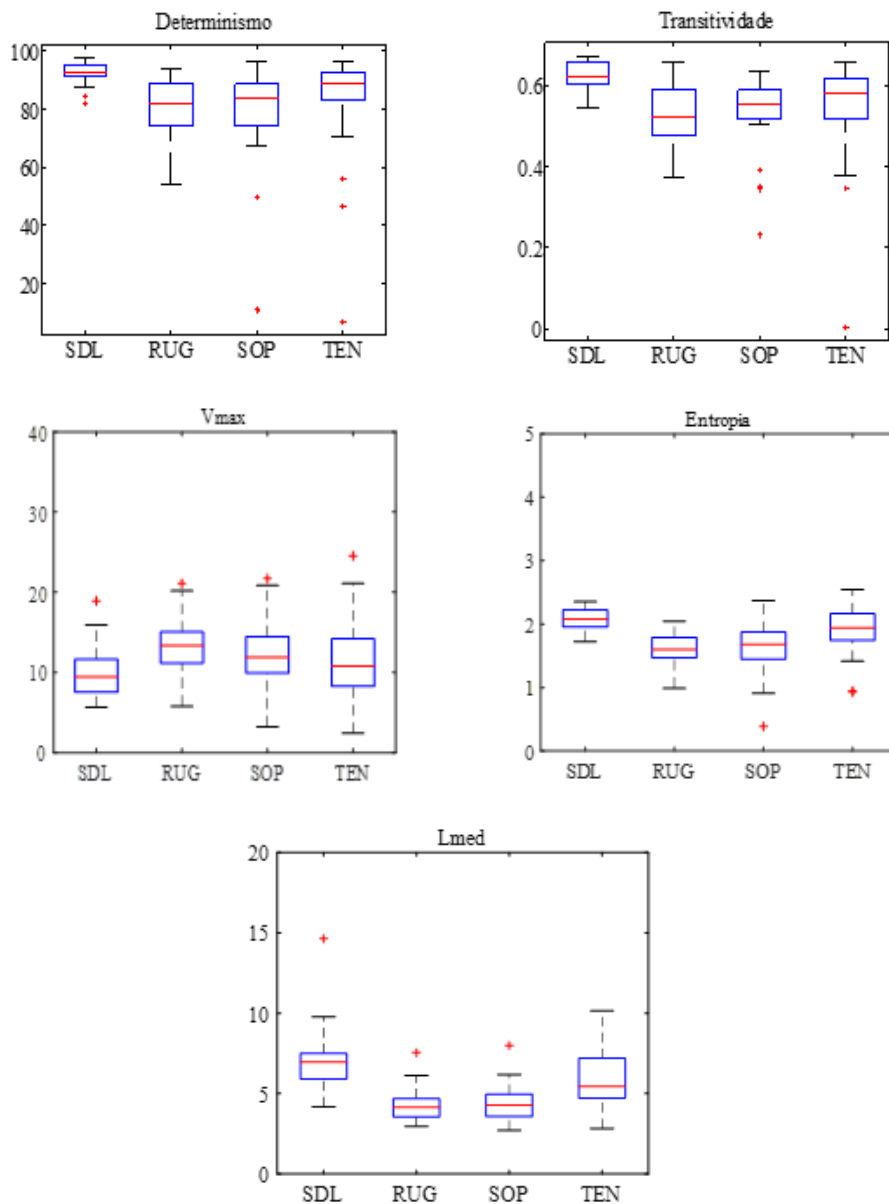
Nas Figuras 16 e 17, respectivamente, estão ilustradas as distribuições dos valores médios das medidas não lineares e os valores médios das medidas de quantificação de recorrência analisadas, sendo 30 sinais de vozes para cada tipo (Saudável, Soprosa, Rugosa e Tensa).

Figura 16: Distribuição dos valores médios das medidas associadas à teoria não linear, caso 1.



Fonte: Autoria própria.

Figura 17: Distribuição dos valores médios das medidas de quantificação de recorrência, caso 1.



Fonte: Autoria própria

Nos gráficos de caixas (*boxplots*), apresentados nas Figuras 16 e 17, observa-se que a medida de Passo de reconstrução para as vozes soprosas, que são considerados sinais mais complexos, apresenta mediana superior às demais classes, assim como a dispersão. No caso da Dimensão de correlação, as vozes saudáveis apresentam menor mediana, sendo que as vozes alteradas apresentam valores dentro do mesmo patamar. As vozes saudáveis e tensas relacionadas as medidas de *Lmed* e *Entr*, encontram-se no mesmo patamar de variabilidade. Isso pode ser explicado pelo fato delas apresetarem

semelhanças em suas características. Para o grupo de vozes saudáveis, a medida de Determinismo proporcionou valores concentrados entre 80% e 100%. Nos sinais com desvios vocais os valores são mais dispersos. Isso pode ser explicado pelo fato desses sinais estarem compreendidos entre diferentes tipos de graus de intensidade do desvio em um mesmo desvio de um mesmo grupo. Além disso, nesta pesquisa não houve separação dos sinais por gênero.

5.3 Análise Estatística no Estudo de Caso 2

Na Tabela 7 estão representados os resultados dos testes estatísticos do Estudo de Caso 2, em que são discriminados os valores de significância obtidos em cada caso de discriminação.

Tabela 7: Potencial discriminativo das medidas da análise dinâmica não linear e de quantificação de recorrência com relação ao Estudo de Caso 2.

Casos de Discriminação	Medidas							
	τ	PM	D_2	Det	$Trans$	$Lmed$	$Entr$	$Vmax$
G1xG2	$p=0,09$	$p=0,07$	$p=0,003$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p=0,4$
G1xG3	$p=0,16$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$
G2xG3	$p=0,95$	$p<0,001$	$p=0,014$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,015$
G1x(G2+G3)	$p=0,17$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$	$p<0,001$

Na discriminação entre os sinais de vozes G1 e G2, apenas as medidas de Passo de Reconstrução (τ) e Comprimento Máximo das Linhas Verticais ($Vmax$), não apresentaram diferença estatística.

A medida de Passo de Reconstrução (τ), não proporcionou diferença estatística em nenhum dos casos discriminativos. Logo, pode-se dizer que esta medida não se mostrou capaz de evidenciar diferenças estatísticas nas classes analisadas.

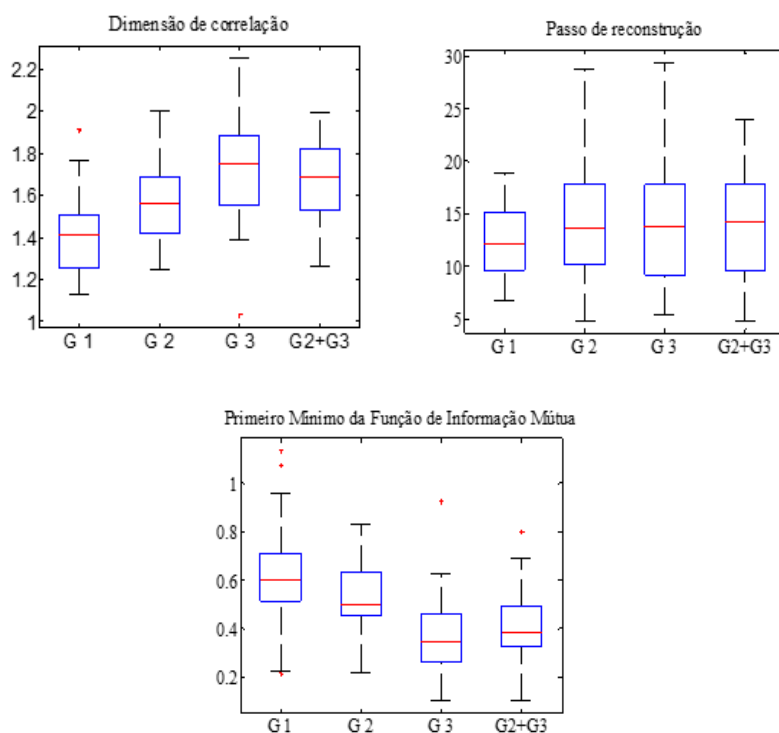
Na maioria dos casos houve diferença estatística entre os grupos considerados. A discriminação entre G1xG2 foi que apresentou menos medidas com diferenças estatísticas.

Nos dois casos de estudo, a medida de Passo de Reconstrução só obteve diferença estatística significativa entre a discriminação de vozes saudáveis e soprosas (SDL x SOP). Pode-se observar também que na classificação entre Grau1 (G1) e Grau 2 (G2), as medidas de Passo de reconstrução (τ), Primeiro Mínimo da Função de Informação Mútua (PM) e Comprimento máximo das linhas verticais ($Vmax$), não proporcionaram diferenças estatísticas. Isso significa dizer que estas medidas não são

tão relevantes no caráter discriminativo de grupos considerados entre sinais de vozes com características próximas.

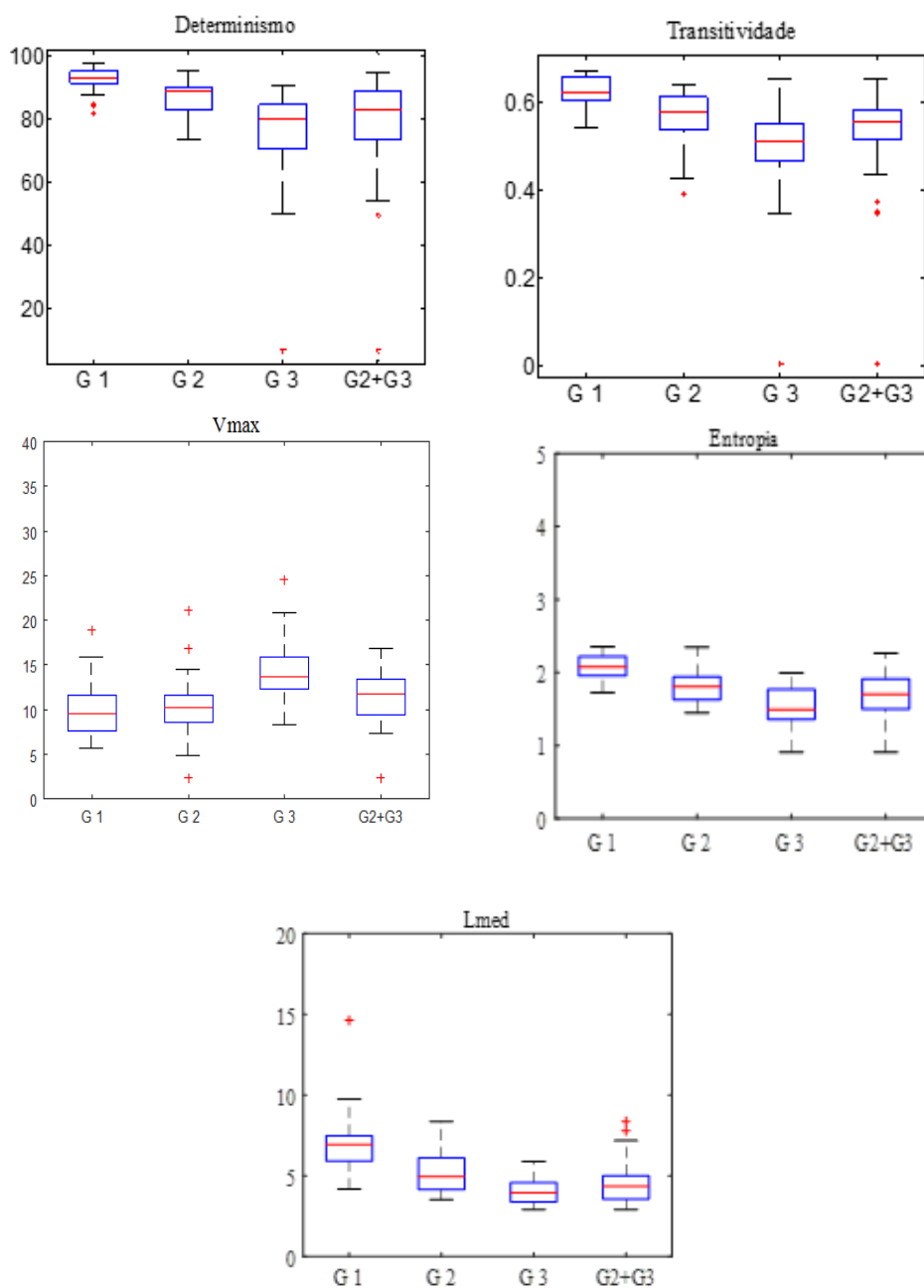
Nas Figuras 18 e 19, respectivamente, estão ilustradas as distribuições dos valores médios das medidas não lineares e os valores médios das medidas de quantificação de recorrência analisadas, sendo 30 sinais de vozes para cada tipo (G1-Grau 1, G2, Grau 2 e G3-Grau3).

Figura 18: Distribuição dos valores médios das medidas associadas à teoria não linear, caso 2.



Fonte: Autoria própria.

Figura 19: Distribuição dos valores médios das medidas de quantificação de recorrência, caso 2.



Fonte: Autoria própria.

Neste caso de estudo a Dimensão de correlação e o Passo de reconstrução da classe de vozes saudáveis (G1), apresenta uma mediana inferior às demais classes. Com relação às medidas de Primeiro Mínimo da Função de Informação Mútua e Transitividade, as vozes saudáveis apresentaram uma mediana superior às demais. A medida de Determinismo proporcionou valores concentrados entre 80% e 100%, com relação ao G1 e G2. Para os demais graus os valores são mais dispersos.

5.4 Classificação

5.4.1 Classificação do Estudo de Caso 1

Nesta Seção estão apresentados os resultados referentes à classificação do Estudo de Caso 1. A metodologia empregada no processo de classificação está descrita no Capítulo 4.

Classificação das Vozes Saudáveis (SDL) x Desviadas (DESV)

Na Tabela 8, estão apresentados os melhores valores obtidos das medidas individuais, em ordem de classificação, da Acurácia, Sensibilidade e Especificidade, no processo discriminativo entre as vozes saudáveis e desviadas.

Tabela 8: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e desviadas com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Entr</i>	84,67±0,65	79,00±2,05	90,33±0,92
<i>Trans</i>	83,00±0,85	82,33±3,72	83,67±3,80
<i>Det</i>	83,50±0,68	83,00±2,25	84,00±1,39
<i>Lmed</i>	82,83±1,45	83,00±3,77	82,67±1,71
<i>D₂</i>	73,17±1,10	67,67±4,69	78,67±2,64
<i>PM</i>	71,83±2,52	73,67±3,46	70,00±6,94
<i>Vmax</i>	63,50±2,98	66,00±2,57	61,00±5,24
τ	60,33±2,33	57,67±5,35	63,00±8,07

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade

Entr: Entropia de Shannon; *Trans*: Transitividade; *Det*: Determinismo; *Lmed*: Comprimento médio das linhas diagonais; *D₂*: Dimensão de correlação; *PM*: Primeiro Mínimo da Função de Informação Mútua, *Vmax*: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

As medidas de Entropia de Shannon, Transitividade, Determinismo e Comprimento Médio das Linhas Diagonais proporcionaram os melhores resultados, atingindo uma acurácia acima de 80%, indicando que estas medidas, relacionadas às linhas diagonais (*Lmed*, *Det* e *Entr*) e a medida relacionada aos pontos de recorrência (*Trans*), são fatores principais na discriminação entre as classes consideradas.

Na Tabela 9, estão representados os melhores desempenhos da classificação com as medidas combinadas.

Tabela 9: Melhor desempenho na classificação dos sinais de vozes saudáveis e desviadas com as medidas combinadas, empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR(%)	SENS (%)	ESP (%)
<i>PM+Lmed</i>	2x2	88,33±0,56	91,33±1,74	85,33±1,13
<i>PM+τ + Trans</i>	3x3	89,00±0,51	86,33±1,26	91,67±1,24
<i>PM+Lmed+Trans+Det</i>	4x4	90,00±0,50	92,33±1,12	87,67±1,00
<i>Trans + τ+ Lmed+Ent+ D₂</i>	5x5	90,17±0,80	89,00±1,65	91,33±0,74
<i>PM+Lmed+Det+Entr+Trans+Vmax</i>	6x6	90,33±0,60	91,33±1,02	89,33±1,20
<i>τ +Lmed+Det+Entr+Trans+Vmax+ D₂</i>	7x7	89,17±1,00	88,33±1,67	90,00±1,11
Todas	8x8	86,50±1,87	81,00±4,95	92,00±1,74

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ: Passo de reconstrução.

O melhor desempenho de classificação foi obtido a partir da classificação das medidas *PM*, *Lmed*, *Det*, *Entr*, *Trans* e *Vmax*. Porém, levando em consideração as diferenças estatísticas, os resultados ficam praticamente em um mesmo patamar, com relação aos resultados das medidas combinadas.

Classificação das Vozes Saudáveis (SDL) x Soprosas (SOP)

Na Tabela 10 estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis e soprosas, obtidos com as medidas individuais.

Tabela 10: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e soprosas com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Entr</i>	80,33±1,61	82,33±5,44	78,33±2,64
<i>Trans</i>	78,67±0,65	80,67±2,71	76,67±1,70
<i>Lmed</i>	76,83±2,38	78,00±2,78	75,67±5,38
<i>Det</i>	76,83±2,91	86,67±1,79	67,00±7,29
<i>D₂</i>	73,33±1,08	73,33±2,17	73,33±4,07
<i>Vmax</i>	70,50±1,89	69,00±5,58	72,00±2,64
<i>τ</i>	67,00±4,50	73,33±7,87	60,67±3,10
<i>PM</i>	62,50±3,55	60,00±3,72	65,00±5,07

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ: Passo de reconstrução.

Nos resultados das medidas individuais, as medidas Entropia de Shannon (*Entr*) e Transitividade (*Trans*), vem se destacando com os melhores resultados como na classificação entre as vozes saudáveis e soprosas.

Na Tabela 11 estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis e soprosas, obtidos com as medidas combinadas.

Tabela 11: Melhor desempenho na classificação dos sinais de vozes saudáveis e soprosas com as medidas combinadas, empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
$\tau + Trans$	2x2	91,17±0,43	91,67±1,20	86,67±2,11
$\tau + Trans+Det$	3x3	89,67±0,74	92,67±1,20	86,67±2,11
$\tau + Trans+Lmed+D_2$	4x4	91,00±0,75	92,00±1,74	90,00±1,22
$\tau + Trans+Det+Entr+ D_2$	5x5	90,33±0,89	94,00±0,83	86,67±2,49
$\tau+Trans+Lmed+Entr+Vmax+PM$	6x6	90,67±0,75	91,67±1,02	89,67±1,26
$\tau+Trans+Lmed+Entr+Vmax+PM+Det$	7x7	90,33±0,60	90,33±1,16	90,33±0,92
Todas	8x8	90,17±1,23	90,00±2,28	90,33±1,05

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D_2 : Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

A medida de Transitividade (*Trans*) e o Passo de Reconstrução (τ) foram as medidas que mais se destacaram, quando as medidas foram combinadas, apresentando assim a maior acurácia, entre todas as outras combinações.

Classificação das Vozes Saudáveis (SDL) x Rugosas (RUG)

Na Tabela 12, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis e rugosas, obtidos com as medidas individuais.

Na classificação das medidas individuais mais uma vez as medidas de Entropia (*Entr*), Transitividade (*Trans*) e Determinismo (*Det*) proporcionaram os melhores resultados.

Tabela 12: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e rugosas com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Entr</i>	84,50±0,50	89,33±1,63	79,67±1,75
<i>Det</i>	84,67±1,02	82,33±1,65	87,00±3,04
<i>Trans</i>	81,17±0,26	92,33±1,00	70,00±1,11
<i>Lmed</i>	79,00±1,52	71,67±5,00	86,33±2,87
<i>PM</i>	74,17±3,03	65,00±8,30	83,33±3,10
<i>D₂</i>	71,50±2,49	67,00±7,34	76,00±4,36
<i>Vmax</i>	71,33±3,29	63,67±6,69	79,00±1,86
τ	53,50±0,46	69,00±4,69	38,00±4,67

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; *Trans*: Transitividade; *Det*: Determinismo; *Lmed*: Comprimento médio das linhas diagonais; *D₂*: Dimensão de correlação; *PM*: Primeiro Mínimo da Função de Informação Mútua, *Vmax*: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

Na Tabela 13, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis e rugosas, obtidos com as medidas combinadas.

Tabela 13: Melhor desempenho na classificação dos sinais de vozes saudáveis e rugosas com as medidas combinadas, empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
<i>D₂ + Lmed</i>	2x2	86,33±0,42	84,33±1,49	88,33±1,24
<i>D₂ + Entr + τ</i>	3x3	87,33±0,90	88,33±2,18	86,33±1,36
<i>Lmed+Entr+Trans+D₂</i>	4x4	89,50±1,32	91,00±1,22	88,00±2,40
$\tau + Det + Entr + Vmax + PM$	5x5	88,50±0,87	93,67±0,92	83,33±1,22
<i>Det+Lmed+Entr+Trans+Vmax+PM</i>	6x6	88,33±0,75	91,33±1,24	85,33±1,59
$\tau + Det + Lmed + Entr + D_2 + Vmax + PM$	7x7	87,33±1,22	89,00±1,93	85,67±2,33
Todas	8x8	85,83±1,81	87,00±3,56	84,67±2,44

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; *Trans*: Transitividade; *Det*: Determinismo; *Lmed*: Comprimento médio das linhas diagonais; *D₂*: Dimensão de correlação; *PM*: Primeiro Mínimo da Função de Informação Mútua, *Vmax*: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

As medidas de Comprimento médio das linhas diagonais, Transitividade, Entropia e Dimensão de correlação, quando combinadas, obtiveram a melhor acurácia na classificação entre vozes saudáveis e rugosas.

Classificação das Vozes Saudáveis (SDL) x Tensas (TEN)

Nas Tabelas 14, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis e tensas, obtidos com as medidas individuais.

Tabela 14: Melhor desempenho individual na classificação dos sinais de vozes saudáveis e tensas com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Trans</i>	68,00±1,66	77,67±4,10	58,33±5,65
<i>D₂</i>	64,50±1,55	68,33±3,23	60,67±4,27
<i>PM</i>	63,83±2,08	71,33±7,67	56,33±8,13
<i>Det</i>	62,17±2,83	79,33±7,58	45,00±5,95
<i>Lmed</i>	57,50±1,88	67,33±6,23	47,67±9,56
τ	57,33±2,10	70,00±5,31	44,67±6,41
<i>Vmax</i>	54,17±1,52	59,00±13,44	49,33±11,56
<i>Entr</i>	53,00±2,70	82,00±9,84	24,00±8,65

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

A medida de Transitividade (*Trans*) mais uma vez se destacou nos resultados da acurácia tanto das medidas individuais, atingindo uma acurácia de 68 %.

Na Tabela 15, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis e tensas, obtidos com as medidas combinadas.

A medida de Transitividade (*Trans*) mais uma vez se destacou nos resultados da acurácia tanto das medidas combinadas, estando presente em todas as classificações. Quando esta medida foi combinada com as medidas de Passo de Reconstrução, Comprimento Máximo das Linhas Verticais, Comprimento Médio das Linhas Diagonais, Determinismo e Dimensão de Correlação, houve um aumento significativo na acurácia atingindo 76,5%. Pode-se concluir que esta medida relacionada a taxa de recorrência local é eficiente para a discriminação de vozes que apresentam características semelhantes como é o caso das vozes saudáveis e tensas.

Tabela 15: Melhor desempenho na classificação dos sinais de vozes saudáveis e tensas com as medidas combinadas, empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
<i>Trans+PM</i>	2x2	71,00±1,03	80,67±4,58	61,33±3,92
<i>Trans + τ + D₂</i>	3x3	73,17±3,07	84,00±4,12	62,33±7,59
<i>Trans + τ + Vmax+ PM</i>	4x4	74,33±2,84	79,33±3,51	69,33±3,88
<i>Trans + τ + Lmed+Entr+PM</i>	5x5	76,33±1,11	82,67±2,62	70,00±3,14
<i>Trans+τ+Vmax+Lmed+Det+ D₂</i>	6x6	76,50±1,77	83,33±2,28	69,67±2,83
<i>Trans+τ+Lmed+PM+Entr+Vmax+ Det</i>	7x7	73,30±1,44	80,00±1,93	67,00±2,56
Todas	8x8	69,83±3,52	76,67±8,81	63,00±7,88

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

Na Figura 20, é apresentado um resumo com os principais resultados obtidos nas classificações realizadas com as medidas extraídas, tanto individuais como combinadas, indicando a acurácia máxima, para cada caso discriminativo desta análise.

Figura 20: Comparação das médias das acurácias dos resultados individuais e combinados.



Observa-se que houve um aumento nos resultados das classificações quando as medidas foram combinadas, destacando-se com um aumento de cerca de 10% na discriminação entre SDLxSOP.

5.4.2 Classificação do Estudo de Caso 2

Os resultados, referentes à classificação realizada no Estudo de Caso 2, são apresentados nesta Seção. A classificação, assim como no Estudo de Caso 1, foi realizada com as medidas extraídas com segmentação. A metodologia aplicada está apresentada no Capítulo 4.

Classificação entre as intensidades de Grau 1 e Grau 2 (G1xG2) dos desvios vocais

Na Tabela 16, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis (Grau 1 - G1) e de grau leve (Grau 2 - G2), obtidos com as medidas individuais.

Tabela 16: Melhor desempenho individual na classificação entre as intensidades de Grau 1 (G1) e Grau 2 (G2), com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Det</i>	72,17±2,07	72,67±4,24	71,67±5,61
<i>Entr</i>	69,83±4,77	61,67±5,19	78,00±5,12
<i>Trans</i>	68,67±0,69	68,33±4,64	69,00±4,58
<i>Lmed</i>	67,00±1,79	75,33±4,04	58,67±6,80
<i>D₂</i>	65,00±1,32	64,33±3,59	65,67±3,59
<i>PM</i>	62,17±2,88	66,00±9,92	58,33±7,97
τ	58,67±1,02	40,67±4,27	76,67±5,53
<i>Vmax</i>	55,83±2,04	52,67±9,60	59,00±8,70

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

Na classificação das medidas individuais, mais uma vez, as medidas de Entropia, Transitividade e Determinismo obtiveram os melhores resultados.

Na Tabela 17, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis (Grau 1 - G1) e de grau leve (Grau 2 - G2), obtidos com as medidas combinadas.

Tabela 17: Melhor desempenho na classificação das medidas combinadas entre as intensidades do Grau 1 (G1) e Grau 2(G2), com as medidas empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
<i>Lmed + D₂</i>	2x2	76,17±0,78	72,00±3,07	80,33±3,20
<i>Lmed + D₂ +Entr</i>	3x3	81,50±1,07	79,33±1,85	83,67±2,02
<i>Lmed + Det +Entr+PM</i>	4x4	82,33±2,18	82,00±2,69	82,67±3,29
<i>Trans + Entr+ PM+Det+ D₂</i>	5x5	80,00±0,93	79,33±1,09	80,67±1,78
<i>Lmed +τ+Vmax+PM+Entr+D₂</i>	6x6	79,83±0,76	78,33±1,67	81,33±2,59
<i>Trans+τ+Lmed+Vmax+Entr+D₂+Det</i>	7x7	77,83±1,78	74,33±3,025	81,33±3,59
Todas	8x8	74,17±2,72	72,67±5,62	75,67±7,01

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ: Passo de reconstrução

As medidas de Comprimento médio das linhas diagonais, Determinismo, Entropia e Primeiro Mínimo da Função de Informação Mútua, combinadas, apresentaram a maior acurácia na discriminação entre os graus 1 e 2.

Classificação entre as intensidades Grau 1 e Grau 3 (G1xG3) dos desvios vocais

Na Tabela 18 estão apresentados os melhores valores obtidos com as medidas individuais, da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis (Grau 1- G1) e de grau moderado (Grau 3- G3).

Tabela 18: Melhor desempenho individual na classificação entre as intensidades de Grau 1 (G1) e Grau 3 (G3), com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Det</i>	86,83±0,58	85,67±1,8	88,00±0,74
<i>Trans</i>	85,83±0,28	85,00±0,75	86,67±0,86
<i>D₂</i>	77,00±0,48	73,33±2,38	80,67±1,71
<i>Vmax</i>	77,5±1,06	78,33±3,95	76,67±1,93
<i>PM</i>	73,33±2,71	77,00±5,85	70,67±7,02
<i>Lmed</i>	72,00±5,55	53,67±13,75	90,33±2,74
<i>Entr</i>	66,67±6,29	56,00±14,03	77,33±12,93
<i>T</i>	59,83±2,27	57,33±7,11	62,33±5,14

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ: Passo de reconstrução.

Na Tabela 19 estão apresentados os melhores valores obtidos com as medidas combinadas, da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis (Grau 1- G1) e de grau moderado (Grau 3- G3).

Tabela 19: Melhor desempenho na classificação das medidas combinadas entre as intensidades do Grau 1 (G1) e Grau 3 (G3), com as medidas empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
<i>Det + Vmax</i>	2x2	89,00±0,79	89,00±2,11	89,00±1,22
<i>τ + Trans + PM</i>	3x3	92,67±0,67	93,33±0,99	92±0,89
<i>Det + Entr + Vmax + D₂</i>	4x4	93,83±1,52	93,33±1,49	94,33±1,86
<i>Det + Entr + Vmax + Trans + PM</i>	5x5	94,5±1,11	92,67±1,63	96,33±0,92
<i>Det + Lmed + τ + Vmax + PM + Trans</i>	6x6	92,5±0,37	92,00±0,74	93,00±0,78
<i>Det + Trans + τ + Lmed + Vmax + Entr + D₂</i>	7x7	93,00±0,65	92,33±1,12	93,67±0,6
Todas	8x8	89,33±0,94	87,33±2,48	91,33±1,24

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ: Passo de reconstrução.

Os melhores resultados individuais estão representados pelas medidas de Transitividade e Determinismo. Quando estas são combinadas com as medidas de Entropia, Comprimento Máximo das Linhas Verticais e Primeiro Mínimo da Função de Informação Mútua obtêm-se um aumento de, aproximadamente, 8% no valor da acurácia.

Classificação entre as intensidades Grau 2 e Grau 3 (G2xG3) dos desvios vocais

Na Tabela 20, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes de grau leve (Grau 2- G2) e de grau moderado (Grau 3 -G3), obtidos com as medidas individuais.

Tabela 20: Melhor desempenho individual na classificação entre as intensidades de Grau 2 (G2) e Grau 3 (G3), com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Trans</i>	71,00±0,75	73,00±2,30	69,00±3,44
<i>PM</i>	71,00±1,45	67,33±5,09	74,67±3,92
<i>Vmax</i>	67,67±3,51	60,33±9,36	75,00±9,13
<i>Det</i>	67,33±0,71	66,00±5,66	68,67±5,4
<i>D₂</i>	65,00±2,59	57,33±7,74	72,67±3,94
<i>Entr</i>	61,00±3,29	44,67±11,16	77,33±9,84
<i>Lmed</i>	56,67±2,64	42,33±13,99	71,00±11,15
<i>T</i>	54,33±1,74	56,33±5,57	52,33±3,75

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

Neste caso discriminativo as medidas individuais que mais se destacaram em sua acurácia foram Transitividade e Primeiro Mínimo da Função de Informação Mútua.

Na Tabela 21, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes de grau leve (Grau 2- G2) e de grau moderado (Grau 3 -G3), obtidos com as medidas combinadas.

Tabela 21: Melhor desempenho na classificação das medidas combinadas entre as intensidades do Grau 2 (G2) e Grau 3(G3), com as medidas empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
$\tau+PM$	2x2	74,50±1,51	70,67±3,33	78,33±3,11
<i>Det + Vmax+PM</i>	3x3	79,67±2,53	79,67±2,13	79,67±6,09
<i>Det + Vmax+PM + D₂</i>	4x4	81,17±1,29	76,67±1,93	85,67±1,41
<i>Det+ Ent+Trans +Vmax+PM</i>	5x5	79,50±0,71	75,33±1,02	83,67±1,44
<i>Det+Lmed+Entr+Vmax+ D₂+PM</i>	6x6	77,33±3,04	80,00±1,11	74,67±6,43
<i>Trans+τ+Lmed+PM+Entr+Vmax+Det</i>	7x7	80,50±1,11	77,67±2,11	83,33±1,79
Todas	8x8	73,83±3,04	65,00±7,67	82,67±2,37

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

A maior acurácia obtida na combinação das medidas foi de 81,17%, com as medidas Determinismo, Comprimento máximo das linhas verticais, Primeiro Mínimo da Função de Informação Mútua e Dimensão de correção.

Classificação entre as intensidades Grau 1 e Graus 2 e 3 (G1x (G2+G3)) dos desvios vocais

Na Tabela 22, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis (Grau 1 - G1) e da junção dos graus leve e moderado (Graus 2 e 3) obtidos com as medidas individuais.

Tabela 22: Melhor desempenho individual na classificação entre as intensidades de Grau 1 (G1) e Graus 2 e 3 (G2+G3), com as medidas empregadas no estudo.

MEDIDAS	ACUR (%)	SENS (%)	ESP (%)
<i>Det</i>	80,67±1,22	74,00±3,25	87,33±0,83
<i>Trans</i>	79,00±1,93	80,00±2,43	78,00±5,63
<i>D₂</i>	76,17±0,56	76,00±2,27	76,33±2,08
<i>PM</i>	74,17±2,13	75,67±5,84	72,67±3,91
<i>Lmed</i>	74,17±3,91	75,00±8,21	73,33±8,42
<i>Entr</i>	69,33±5,13	60,00±10,18	78,67±9,47
τ	60,00±1,79	52,33±5,02	67,67±2,54
<i>Vmax</i>	59,67±2,67	47,67±9,70	71,67±8,72

ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

Na Tabela 23, estão apresentados os melhores resultados da Acurácia, Sensibilidade e Especificidade de cada medida empregada, no processo discriminativo entre as vozes saudáveis (Grau 1 - G1) e da junção dos graus leve e moderado (Graus 2 e 3) obtidos com as medidas combinadas.

Tabela 23: Melhor desempenho na classificação entre as intensidades de Grau 1 (G1) e Graus 2 e 3 (G2+G3), com as medidas combinadas empregadas no estudo.

MEDIDAS	COMBINAÇÕES	ACUR (%)	SENS (%)	ESP (%)
<i>Det + D₂</i>	2x2	83,83±1,27	86,33±1,36	81,33±3,27
<i>Trans + τ+PM</i>	3x3	85,67±1,34	83,33±1,86	88,00±1,59
<i>Trans + τ+ Vmax+ D₂</i>	4x4	85,67±2,11	82,00±4,01	89,33±1,30
<i>Trans + τ+ Lmed+ PM+Det</i>	5x5	86,17±0,50	84,33±1,00	88,00±1,130
<i>Trans + τ+Vmax+PM+Entr+D₂</i>	6x6	84,67±0,78	83,00±1,53	86,33±1,89
<i>Trans + τ+ Lmed+ PM+Entr+ D₂+Det</i>	7x7	86,50±0,76	85,67±0,87	87,33±1,47
Todas	8x8	84,67±1,38	82,67±1,39	86,67±1,93

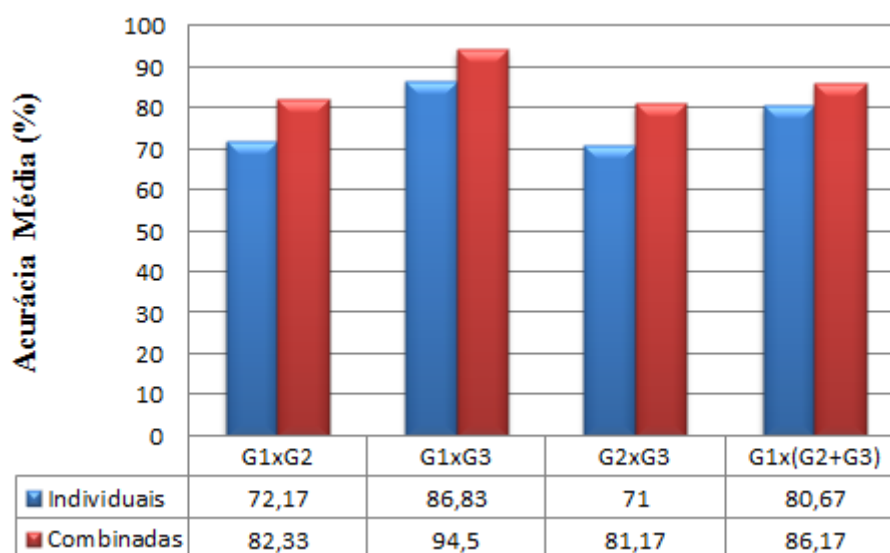
ACUR: Acurácia; SENS: Sensibilidade e ESP: Especificidade.

Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ: Passo de reconstrução.

A medida de Determinismo foi a que mais se destacou em sua acurácia na classificação individual, estando presente também na melhor combinação (7x7), com acurácia acima de 86%.

Na Figura 21 é apresentado um resumo com os principais resultados obtidos nas classificações realizadas com as medidas extraídas, tanto individuais como combinadas, indicando a acurácia máxima, para cada caso discriminativo desta análise (Caso 2).

Figura 21: Comparação das médias das acurácias dos resultados individuais e combinados.



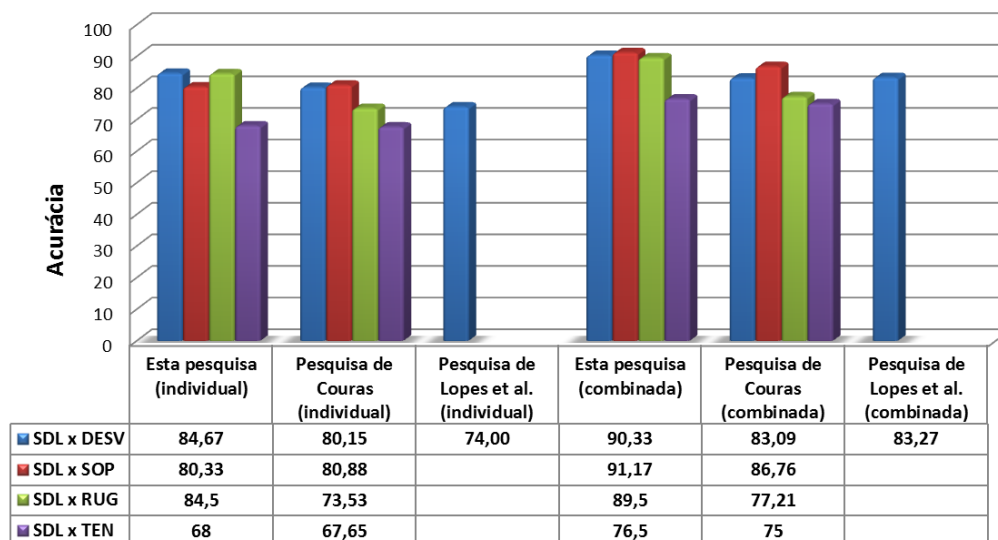
Observa-se que houve um aumento nas classificações quando as medidas foram combinadas, destacando-se com um aumento médio de 10% na maioria dos casos.

5.5 Estudos comparativos

5.5.1 Estudo de Caso 1

Na Figura 22 estão representados os resultados comparativos do estudo do caso 1 com os resultados obtidos na pesquisa de Couras (2017), relacionadas às medidas individuais e combinadas. Couras empregou técnicas de processamento digital de sinais, baseadas no modelo linear de produção da fala. Foram utilizadas medidas tradicionalmente empregadas como frequência fundamental, medidas de perturbação (*jitter e shimmer*), GNE (*Glottal to Noise Excitation Ratio*) e frequências formânticas. Também foi avaliado o potencial discriminativo dos coeficientes da análise de predição linear (*Linear Predictive Coding- LPC*), coeficientes cepstrais e mel-cepstrais na classificação de desvios vocais (rugosidade, sopro e tensão). Lopes *et al.*, analisaram a discriminação entre vozes saudáveis e desviadas com medidas de quantificação de recorrência (Determinismo, Comprimento Médio e Máximo das Linhas Diagonais, Entropia de Shannon, Laminaridade, Tempo de Permanência, Comprimento Máximo das Linhas Verticais, Tempo de Recorrência do tipo 1 e do tipo 2, Entropia do Tempo de Recorrência do tipo 1, Transitividade e Divergência), além dos parâmetros relacionados à topologia recorrente do sistema de produção vocal (Dimensão de Imersão, Passo de Reconstrução e Raio de Vizinhança).

Figura 22: Comparativo entre os resultados obtidos por Couras e Lopes *et al.* com os resultados obtidos nesta pesquisa, na classificação individual e combinada entre as discriminações do Estudo de Caso 1.



Fonte: Autoria própria.

Nas Tabelas 24, 25 e 26 estão representados um resumo das medidas que proporcionaram os melhores resultados individuais e combinadas, respectivamente, com relação a Figura 22, da pesquisa de Couras, Lopes *et al.* e desta pesquisa.

Tabela 24: Resultado das acurácias e das medidas que mais se destacaram na pesquisa de Couras, 2017.

Classificação	Resultados	Individual	Combinada
SDL x DESV	ACUR (%)	80,15± 0,98	83,09±0,98
	Medida (s)	<i>Shimmer</i>	$F_{1med}, F_{2med}, Jitter, Shimmer, GNE$
SDL x SOP	ACUR (%)	80,88±0,85	86,76±0,72
	Medida (s)	GNE	F_{1med}, F_{2med}, GNE
SDL x RUG	ACUR (%)	73,53±2,71	77,21±1,05
	Medida (s)	<i>Shimmer</i>	$F_{2med}, Shimmer, GNE$
SDL x TEM	ACUR (%)	67,65±2,76	75,00±2,31
	Medida (s)	<i>Shimmer</i>	$F_{0med}, F_{1med}, F_{2med}, F_{3med}, GNE$

ACUR: Acurácia; GNE: Glottal to Noise Excitation Ratio; F_0 : Frequência fundamental; F_{1med} : 1º Formante médio; F_{2med} : 2º Formante médio; Frequência fundamental; F_{3med} : 3º Formante médio.

Tabela 25: Resultado das acurácias e das medidas que mais se destacaram na pesquisa de Lopes *et al.*, 2016.

Classificação	Resultados	Individual	Combinada
	ACUR (%)	74,00±4,06	83,27±3,76
SDL x DESV	Medida (s)	<i>Entr</i>	$\tau, m, Lmed, Entr, TT, Vmax, T1, Trans$

ACUR: Acurácia; Entr: Entropia; Lmed: Comprimento médio das linhas diagonais, TT: Tempo de permanência; T1: Tempo de recorrência tipo 1, *m*: Dimensão de imersão; τ : Passo de reconstrução.

Tabela 26: Resultado das acurácias e das medidas que mais se destacaram nesta pesquisa.

Classificação	Resultados	Individual	Combinada
	ACUR (%)	84,67±0,65	90,33±0,60
SDL x DESV	Medida (s)	<i>Entr</i>	$PM, Lmed, Det, Entr, Trans, Vmax$
	ACUR (%)	80,33±1,61	91,17±0,43
SDL x SOP	Medida (s)	<i>Entr</i>	$\tau, Trans$
	ACUR (%)	84,50±0,50	89,50±1,32
SDL x RUG	Medida (s)	<i>Entr</i>	$Lmed, Entr, Trans, D_2$
	ACUR (%)	68,00±1,66	76,50±1,11
SDL x TEN	Medida (s)	<i>Trans</i>	$\tau, Trans, Lmed, Entr, PM$

ACUR: Acurácia; Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D_2 : Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

De um modo geral, as medidas combinadas demonstraram um potencial discriminativo melhor do que as medidas individuais.

Quando comparados os resultados das medidas individuais, desta pesquisa, com os resultados obtidos por Couras, observa-se que na discriminação das vozes saudáveis x soprosas e saudáveis x tensas, foram semelhantes os resultados da classificação. Já quando comparados os resultados relacionados as medidas combinadas, foi observado que nesta pesquisa o modelo empregado baseado em medidas não lineares, proporcionaram melhores resultados na maioria dos casos, exceto para a discriminação entre as vozes saudáveis e tensas que ficaram no mesmo patamar.

Para ambos os casos de estudo foram utilizadas a mesma base de dados, proveniente do LIEV/UFPB. Porém, Couras, utilizou vozes adultas feminina e nesta pesquisa foram utilizadas vozes adultas masculinas e femininas. O classificador utilizado nas duas pesquisas foi uma rede neural MLP.

Lopes *et al.*, conseguiram atingir uma acurácia máxima de 83,27%, com oito medidas. Vale ressaltar que com apenas as medidas τ e *Trans*, já conseguiram atingir

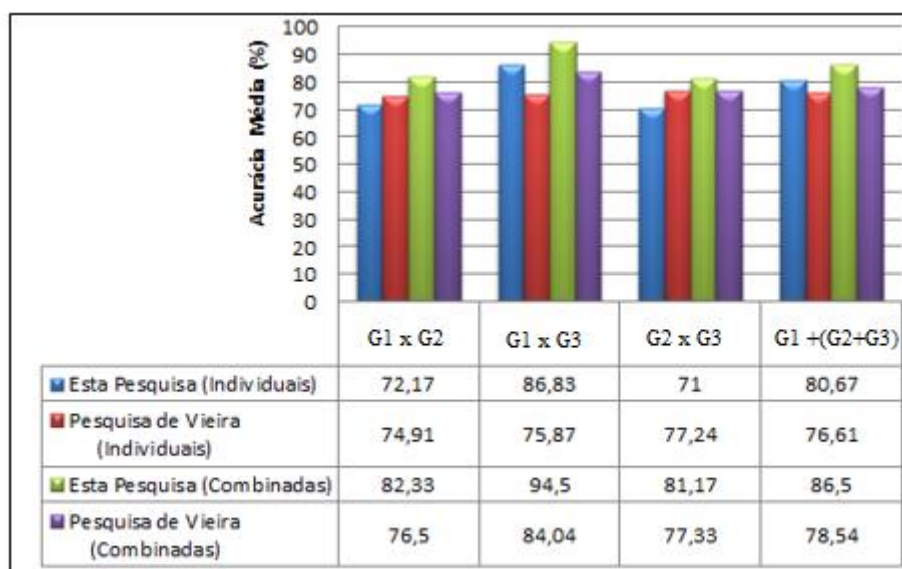
uma acurácia de 80,36%. No processo de classificação foi utilizada a análise discriminante quadrática.

Os resultados indicam o modelo não linear para a análise do estudo dos sinais de voz proporcionou um melhor resultado quando comparados ao estudo baseado no modelo linear, nestes casos de estudo especificamente, mostrando-se que o método foi mais eficiente. Isso pode estar relacionado as limitações que há no modelo linear no que se refere a dificuldade de se analisar um sistema dinâmico muito caótico, ou seja, caso um sinal de voz apresente uma severidade muito acentuada fica difícil analisar medidas relacionadas a frequência fundamental (*jitter*, *shimmer*), por exemplo, que foram uma das medidas utilizadas por Couras.

5.5.2 Estudo de Caso 2

Na Figura 23, estão representados os resultados comparativos do estudo do caso 2 com os resultados obtidos na pesquisa de Vieira (2014), com os sinais segmentados, relacionadas às medidas individuais e combinadas. Em sua pesquisa, Vieira (2014) analisou os sinais de vozes infantis diagnosticadas de acordo com o grau de intensidade do desvio vocal. As medidas de quantificação de recorrência foram extraídas com e sem segmentação, de acordo com as taxas de recorrência de 1% a 5%. As medidas empregadas na análise sem segmentação foram: Determinismo (*Det*), Comprimento Máximo das Linhas Diagonais (*Lmax*), Entropia de Shannon (*Entr*), Tendência (*TREND*), Laminaridade (*LAM*), Tempo de Permanência (*TT*) e Comprimento Máximo das Linhas Verticais (*Vmax*). Na análise dos sinais segmentados foram acrescentadas as medidas: Transitividade (*Trans*), relação entre Determinismo e Taxa de Recorrência (*RATIO*), Divergência (*DIV*), Comprimento Médio das Linhas Diagonais (*Lmed*), Tempo de Recorrência do Tipo 1 (*T1*), Tempo de Recorrência do tipo 2 (*T2*), Entropia do Tempo de Recorrência do tipo 1 (*RPDE*), Raio de Vizinhança (*RAIO*) e Medida da Taxa de Recorrência (*REC*).

Figura 23: Comparativo entre os resultados obtidos por Vieira e os resultados obtidos nesta pesquisa, na classificação individual e combinadas entre as discriminações do Estudo de Caso 2.



Fonte: Autoria própria.

Nas Tabelas 27 e 28, estão representados um resumo das medidas que proporcionaram os melhores resultados individuais e combinadas, respectivamente, com relação a Figura 23, da pesquisa de Vieira e desta pesquisa.

Tabela 27: Resultado das acurácias e das medidas que mais se destacaram na pesquisa de Vieira, 2014.

Classificação	Resultados	Individual	Combinada
G1 x G2	ACUR (%)	74,91±0,64	76,50±0,74
	Medida (s)	<i>Lmax</i>	<i>Lmed,TT,Vmax</i>
G1 x G3	ACUR (%)	75,87±0,68	84,04±1,24
	Medida (s)	<i>Lmed</i>	<i>RAIO,REC,Lmed,Entr,LAM,TT,Vmax,T²</i>
G2 x G3	ACUR (%)	77,24±0,17	77,33±0,46
	Medida (s)	<i>Det</i>	<i>Det, RATIO</i>
G1 x (G2 + G3)	ACUR (%)	76,61±0,56	78,54±0,31
	Medida (s)	<i>Lmax</i>	<i>Lmed, LAM,Vmax</i>

ACUR: Acurácia; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; Lmax: Comprimento máximo das linhas diagonais Vmax: Comprimento máximo das linhas verticais; TT: Tempo de permanência; LAM: Laminaridade; RATIO: Taxa de Recorrência; Entr: Entropia de Shannon; REC: Medida da Taxa de Recorrência; T2: Tempo de Recorrência do tipo 2.

Tabela 28: Resultado das acurácias e das medidas que mais se destacaram nesta pesquisa.

Classificação	Resultados	Individual	Combinada
G1 x G2	ACUR (%)	72,17±2,07	82,33±2,18
	Medida (s)	<i>Det</i>	<i>Lmed, Det, Entr, PM</i>
G1 x G3	ACUR (%)	86,83±0,58	94,50±1,11
	Medida (s)	<i>Det</i>	<i>Det, Entr, Vmax, Trans, PM</i>
G2 x G3	ACUR (%)	71,00±0,75	81,17±1,29
	Medida (s)	<i>Trans</i>	<i>Det, Vmax, PM, D₂</i>
G1 x (G2 + G3)	ACUR (%)	80,67±1,22	86,50±0,76
	Medida (s)	<i>Det</i>	<i>Trans, τ, Lmed, PM, Entr, D₂, Det</i>

ACUR: Acurácia; Entr: Entropia de Shannon; Trans: Transitividade; Det: Determinismo; Lmed: Comprimento médio das linhas diagonais; D₂: Dimensão de correlação; PM: Primeiro Mínimo da Função de Informação Mútua, Vmax: Comprimento máximo das linhas verticais e τ : Passo de reconstrução.

Nesta pesquisa os resultados com as medidas combinadas proporcionaram um melhor desempenho. Já na pesquisa de Vieira, os resultados das classificações individuais e combinadas ficaram em um mesmo patamar, exceto na classificação entre sinais de vozes saudáveis (G1) e considerados com desvio moderado (G3).

Em ambas as pesquisas as medidas relacionadas as linhas diagonais e verticais se destacaram em quase todos os casos, mostrando que a formação das linhas diagonais e verticais são fatores importantes para a separação destas discriminações.

A base de dados utilizada por Vieira é proveniente do Laboratório de Voz e Deglutição do Departamento de Fonoaudiologia/UFPB e nesta pesquisa a base é proveniente do LIEV/UFPB.

Os resultados desta pesquisa, com relação à intensidade dos desvios vocais proporcionaram um melhor resultado em comparação com a pesquisa de Vieira, 2014. Porém, é importante ressaltar que na pesquisa de Vieira foram utilizadas vozes infantis, que possuem particularidades anatômicas, as quais além de apresentar dimensões menores que as dos adultos, possuem epiglote em posição mais posterior, impedindo a exposição adequada das pregas vocais.

Nos resultados obtidos por Vieira, na discriminação da intensidade dos desvios vocais com segmentação, a maior acurácia foi de 84,04% com as medidas combinadas. Nesta pesquisa obteve-se uma acurácia máxima de 94,5%, também com as medidas combinadas. Em ambos os casos estas acurácias foram à discriminação entre as vozes saudáveis (Grau1- G1) e com desvio moderado (Grau 3- G3).

5.6 Considerações Finais do Capítulo

Nesta Seção, serão discutidos os resultados obtidos nesta pesquisa. Para o Estudo de Caso 1, as medidas de Entropia de Shannon (*Entr*), Determinismo (*Det*), Comprimento Médio das Linhas Diagonais (*Lmed*) e Transitividade (*Trans*), proporcionaram um melhor desempenho na maioria dos casos de classificação individual. Foi identificado que estas medidas, conseguem separar com mais precisão a discriminação de vozes saudáveis e com desvios (soprosas e rugosas). Isto pode ser um reflexo do fato das vozes saudáveis apresentarem mais regularidades em suas linhas diagonais, assim como as vozes rugosas, soprosas e tensas possuem um número maior de pontos de recorrência isolados. A medida de Transitividade (*Trans*), que está relacionada à taxa de recorrência local, destacou-se na discriminação entre as vozes saudáveis e tensas. O que significa que esta medida consegue separar sinais que possuem características semelhantes, como é o caso deste tipo de discriminação. Porém, quando combinadas com as medidas, entre elas as medidas não lineares baseadas na teoria do caos, há um aumento relevante em seus resultados da acurácia, como mostra a Figura 20.

No Estudo de Caso 2, as medidas que mais se destacaram individualmente foram a Transitividade (*Trans*) e o Determinismo (*Det*). Estas medidas relacionadas à taxa de recorrência local e à razão do número de pontos recorrente das linhas diagonais em relação a todo o conjunto de pontos recorrentes são fatores predominantes para a discriminação entre as vozes saudáveis, consideradas de Grau 1, e com desvios, considerados de Graus 2 (Leve) e Grau 3 (Moderado). Porém, quando combinadas com outras medidas obtêm-se um aumento médio da acurácia de 10%. Para este caso de estudo, estas medidas estavam presentes, na maioria das vezes, nos maiores resultados das acurácias. Pode-se concluir que medidas relacionadas aos pontos de recorrências são de bastante relevância para o fator discriminativo para este caso de estudo.

Para ambos os Casos de Estudo, a medida de Passo de Reconstrução só proporcionou diferença estatística na discriminação entre as vozes saudáveis e soprosas. Mas, ainda sim foi utilizada com as medidas combinadas a fim de observar o seu comportamento. Esta por sua vez, quando combinada com a medida de Transitividade na discriminação entre as vozes saudáveis e soprosas, proporcionou um aumento significativo em sua acurácia atingindo um aumento de 11%. Isso significa dizer que esta medida é eficiente quando combinada com outras medidas.

Em todos os casos de discriminação com as medidas combinadas, tanto no Estudo de Caso 1 como no Estudo de Caso 2, havia pelo menos uma medida relacionada as medidas não lineares (Passo de Reconstrução, Primeiro Mínimo da Função de Informação Mútua e Dimensão de Correlação). Pode-se concluir que quando combinadas com as medidas de quantificação de recorrência, proporcionam um aumento em suas acurácias.

Nesta pesquisa, atingiu uma acurácia máxima de 94,5%, na discriminação entre as vozes consideradas saudáveis (G1) de grau moderado (G3).

Os resultados desta pesquisa são considerados excelentes, segundo o critério de Hosmer-Lemeshow (2000).

No próximo capítulo serão abordadas as considerações finais desta pesquisa.

6. CONSIDERAÇÕES FINAIS

O uso da análise acústica, na fonoaudiologia, como ferramenta adicional para diagnóstico de sinais de voz na discriminação entre sinais saudáveis e sinais com alguma presença de patologia ou desvio vocal, vem sendo cada vez mais estudado. O grande desafio é encontrar características que melhor representem o comportamento caótico de um sinal de voz, pois cada patologia ou desvio vocal possuem características particulares que melhor os definem.

Nesta pesquisa foi investigada a aplicabilidade de algumas das medidas não lineares tradicionais na literatura, com as medidas de quantificação de recorrência. Para o processo de classificação foram empregadas oito medidas com o objetivo de identificar o comportamento não linear do sistema de produção vocal. As medidas relacionadas aos pontos recorrentes e as linhas diagonais foram as que mais se destacaram no processo de classificação. Entre elas estão as medidas de Determinismo (*Det*), Transitividade (*Trans*), Comprimento médio das linhas diagonais (*Lmed*) e Entropia de Shannon (*Entr*). O fato da formação das linhas diagonais no gráfico de recorrência refletir a regularidade de um sinal e os pontos recorrentes apresentarem-se mais isolados em sinais com menos regularidade é um dos fatores que mais foram observados nas discriminações das análises.

Nos resultados, percebe-se que as medidas do Passo de reconstrução (τ), a Dimensão de correlação (D_2) e o Primeiro mínimo da função de informação mútua (*PM*) estão presentes em quase todos os casos das medidas combinadas. Isso comprova que estas medidas tradicionais combinadas com as medidas de quantificação de recorrência contribuem para melhorar os resultados.

Nesta pesquisa, foram considerados dois casos de estudo usando a junção das medidas não lineares, baseada na teoria do caos, com as medidas de quantificação de recorrência.

No Estudo de Caso 1 atingiu-se uma acurácia máxima de 91,17%, tendo como melhores resultados as classificações das medidas combinadas de Passo de reconstrução (τ) e Transitividade (*Trans*) na discriminação entre as vozes soprosas e as vozes saudáveis. Com apenas duas medidas já se obteve um bom resultado, o que representa menor custo computacional.

No Estudo de Caso 2 obteve-se uma acurácia máxima de 94,5%, entre as intensidades dos desvios de Grau 1 (G1) e Grau 3 (G3), com a combinação das medidas de Determinismo (*Det*), Transitividade (*Trans*), Entropia de Shannon (*Entr*), Comprimento máximo das linhas verticais (*Vmax*) e Primeiro mínimo da função de informação mútua (*PM*).

Uma das principais motivações para estas pesquisas está no fato de não ter sido encontrado nenhum trabalho que relacionasse a análise de medidas não lineares com as medidas de quantificação de recorrência à avaliação de distúrbios da voz, assim como também a intensidade destes desvios. Além disso, também não foi encontrado nenhum trabalho relacionado à análise visual dos gráficos de recorrência com relação aos desvios vocais (rugosidade, soproidade e tensão) e nem com relação a intensidade destes desvios.

Diante das análises realizadas nos dois casos de estudo, conclui-se que as medidas não lineares combinadas com as medidas de quantificação de recorrência, podem ser consideradas promissoras para serem aplicadas em ferramentas de apoio aos diagnósticos de distúrbios vocais.

6.1 Contribuições da pesquisa

As principais contribuições desta pesquisa são as seguintes:

- » Utilização de medidas não lineares tradicionais combinadas às medidas de quantificação de recorrência para a discriminação entre vozes normais e com desvios vocais (rugosidade, soproidade, tensão), para vozes femininas e masculinas, em relação às intensidades dos desvios vocais (Grau 2 - G2 e Grau 3 - G3), para vozes adultas;
- » Investigação do uso de redes neurais MLP, com as medidas não lineares e de quantificação de recorrência;
- » Estudo sobre a formação de estruturas nos gráficos de recorrência com relação aos desvios vocais (rugosidade, soproidade e tensão) e a intensidade destes desvios (Grau 2- G2 e Grau 3 - G3);

6.2 Sugestões para Trabalhos Futuros

As sugestões para trabalhos futuros estão elencadas a seguir.

- » Avaliar a influência da presença de lesão laríngea nos desvios vocais e em seu processo de classificação;
- » Utilizar outros classificadores, além da rede neural MLP, bem como a investigação do potencial das medidas de quantificação de recorrência com a combinação desses classificadores;
- » Realizar a discriminação entre os desvios vocais (RUGxSOP, RUGxTEN, SOPxTEN), utilizando medidas não lineares e medidas de quantificação de recorrência.

REFERÊNCIAS BIBLIOGRÁFICAS

- ASHA. **Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V)**. Sponsored by American Speech- Language- Hearing Associations division 3: Voice and Voice disorders. Department of communication Science and Disorders. University of Pittsburgh, June 10-11, 2002. Disponível em: www.asha.org.
- ALONSO, J. B., ALONSO, I., FERRER et al., M. A. Automatic detection of pathologies in the voice by hos based parameters. **EURASIP Journal on Advances in Signal Processing**, vol. 2001, nº. 4, pp. 275–284, 1900.
- BARAVIEIRA, P. B. **Aplicação de uma rede neural artificial para a avaliação da rugosidade e soproidade vocal**. Universidade de São Paulo. Tese de Doutorado, 2016.
- BEHLAU, M., ZIEMER, R. **Psicodinâmica vocal**. In: FERREIRA, L.P. *Trabalhando Voz*. São Paulo: Summus, 1987.
- BEHLAU, M. **Voz, O Livro do Especialista**. E. Revinter. Ed., vol. 1, 2001.
- BEHLAU, M. Oficina de técnicas de avaliação da voz. In: **XV Congresso Brasileiro de Fonoaudiologia**. Campos de Jordão, São Paulo, 2008.
- BELE, I. V. Reliability in perceptual analysis of voice quality. **Journal of Voice**, vol. 19, nº 4, pp. 555-573, 2005.
- BIREME. Decs - descritores em ciências da saúde. Online: <http://decs.bvs.br/>, Acesso em 14 de janeiro de 2014.
- CAMPISI, P., TEWFIK, T. L., PELLAND, B. E., HUSEIN, M, SADEGHI N. MultiDimensional Voice Program analysis in children with vocal cord nodules. **J. Otolaryngol.** vol 29, nº 5, pp. 302-8, 2000.
- COLTON, R. H., CASPER, J. K. and LEONARD, R. **Understanding voice problems: A physiological perspective for diagnosis and treatment**. Wolters Kluwer Health, 2006.
- COSTA, W. C. de A. **Análise dinâmica não linear de sinais de voz para detecção de patologias laríngeas**. Universidade Federal de Campina Grande. Tese de Doutorado, 176 p., 2012.
- COSTA, W.C.de A., COSTA, S.L.N.C., ASSIS, F. M., NETO, B.G.A. Classificação de sinais de vozes saudáveis e patológicas por meio da combinação entre medidas da análise dinâmica não linear e codificação preditiva linear. **Revista Brasileira de Engenharia Biomédica**, vol 29, nº 1, pp. 3-14, 2013.
- COURAS, M. F.K. B. **Classificação de desvios vocais utilizando características baseadas no modelo linear de produção da fala**. Instituto Federal da Paraíba. Dissertação de Mestrado, 115 p, 2017.
- CRUZ, B. S. C., OLIVEIRA, I. B. Disfonias nas Laringetomias Parciais: revisão sistemática de literatura em procedimentos avaliativos. **Anais do XIX Encontro de Iniciação Científica**, Pontifícia Universidade Católica, Campinas, São Paulo, 2014.

- DAJER, M. E. **Padrões visuais de dinâmica vocal e imagens laríngeas de alta velocidade para análise de voz**, Universidade de São Paulo, 2013.
- DAJER, M. E. **Análise de sinais de voz por padrões visuais de dinâmica vocal**. Universidade de São Paulo. Tese de Doutorado, 154 p, 2010.
- ECKMANN, J. P.; KAMPHORST S. O.; RUELLE, D. Recurrence plots of dynamical systems. **Europhys, Letters**, 56(5):973–977, 1987.
- ELEMETRICS, K. **Kay elemetrics corp. disordered voice database**. Model 4337, 03 Ed., 1994.
- FANT, G. **Speech acoustics and phonetics**. Springer, vol. 24, 2004.
- FIEDLER-FERRARA N., PRADO, C. P. C., **Caos - uma Introdução**. Editora Edgard Blücher, 1994.
- FRASER, A. M. and SWINNEY, H. L., Independent coordinates for strange attractors from mutual information. **Physical Review A**, vol. 33, nº 2, p. 1134, 1986.
- GALDINO, D. G. **Análise acústica não linear dos padrões visuais da dinâmica vocal (PVDV) de homens e adultos**. Faculdade de Medicina de Ribeirão Preto. Dissertação de mestrado, 149 p, 2012.
- GAMA, A. C. C., ALVES, C. F. T., CERCEAU, J. S. B. and TEIXEIRA, L. C. Correlation between acoustic-perceptual data and voice-related quality of life in elderly women. **Pró-Fono Revista de Atualização Científica**, vol. 21, no. 2, pp. 125–130, 2009.
- HAYKIN, S. S **Neural networks: a comprehensive foundation**. Prentice Hall, Inc., 1999.
- HAYKIN, S. S. **Redes Neurais: Princípios e Prática**. Bookman Companhia Editora, Porto Alegre, 2001, 2ª ed., p. 900.
- HENRIQUEZ, P., ALONSO, J. B., FERRER, M. A., TRAVIESO, C. M., CODINO-LLORENTE and MARIA, F. D., Characterization of healthy ant pathological voice through measures based on nonlinear dynamics. **IEEE Transactions on Audio, Speech and Language Processing**, vol. 17, nº. 6, pp. 1186-1195, 2009.
- HIRANO M., **Clinical Examination of Voice**. Wein Austria: Springer-Verlag, 1981.
- HIRSCHBERG, J., DEJONCKERE, P., HIRANO, M., MORI, SCHULTZ-COULON, J. Voice disorders in children. **International journal of pediatric otorhinolaryngology**, vol. 32, pp. S109–S125, 1995.
- HOSMER, D. W. JR., LEMESHOW, S. **Applied logistic regression**. 2º ed. New York: John Wiley & Sons; 2000.
- JIANG, J., ZHANG, Y., MCGILLIGAN, C. Chaos in voice, from modeling to measurement. **Journal of Voice**, vol. 20, nº 1, pp. 2–17, January 2006.
- KANTZ, H. and SCHREIBER T. **Nonlinear time series analysis**. Cambridge University Press, 2004.
- KEMPSTER, G. B., GERRATT, B. R., VERDOLINI, A. K., BARKMEIER-KRAEMER, J. and HILLMAN R. E. Consensus auditory-perceptual evaluation of

- voice: development of a standardized clinical protocol. **American Journal of Speech-Language Pathology**, vol. 18, n° 2, p. 124, 2009.
- KREIMAN, J., GERRATT, B. R. Validity of rating scale measures of voice quality. **J. of the Acoustical Society of America**. 104, pp. 1598-1608, 1998.
- KUMAR, A., MULLICK, S. K. Mullick, Nonlinear dynamical analysis of speech, **Journal of Acoustical Society of America**, vol. 100, n°. 1, pp. 615–629, 1996.
- LOPES, L. W., LIMA, I. L. B., ALMEIDA, L. N. A., CAVALCANTE, D. P. and ALMEIDA, A.A.F. Severity of voice disorders in children: Correlations between perceptual and acoustic data. **Journal of Voice**, vol. 26, n° 6, pp. 819–825, 2012.
- LOPES, L. W. , CAVALCANTE, D. P., COSTA, P. O. **Intensidade do desvio vocal: integração de dados perceptivo-auditivos e acústicos em pacientes disfônicos**, vol 26, n° 5, pp.382-8, 2014.
- LOPES, L. W., SILVA, H. F., EVANGELISTA, D. S., SILVA, J. D., SIMÕES, L. B., SILVA, P. O. C., SILVA, M. F. B. L., ALMEIDA, A. A. F. Relação entre os sintomas vocais, intensidade do desvio vocal e diagnóstico laríngeo em pacientes com distúrbios de voz. **Revista CODAS**, vol 10, pp.1782-2317, 2015.
- LOPES, L. W., VIEIRA, V. J. D., COSTA, S. L. N. C., CORREIA, S. E. N., BEHLAU, M. Effectiveness of recurrence quantification measures in discriminating patients with and without voice disorders. **In: 10th International Conference on Voice Physiology and Biomechanics**; 2016; Viña del Mar, Chile, p. 14-17, 2016.
- LORENZ, E. N. Deterministic non-periodic flow. **Journal of Atmospheric Science**, vol 26, pp. 130–136, 1963.
- MADAZIO, G. **Diagrama de desvio fonatório na clínica vocal**. Universidade Federal de São Paulo, Tese de Doutorado, 2000.
- MARTINS, P. C., COUTO, T. E. , GAMA, A. C. C. Avaliação perceptiva auditiva do grau do desvio vocal: correlação entre escala visual analógica e escala numérica. **Pró-Fono Revista de Atualização Científica**, vol 27, n° 3, pp.279-84, 2015.
- MARWAN, N. **Encounters With Neighbours - Current Developments of Concepts Based on Recurrence Plots and Their Applications**. Ph.D. thesis, University of Potsdam, 159 p 2003.
- MARWAN, N., ROMANO, M. C., THIEL M. and KURTHS, J. Recurrence plots for the analysis of complex systems. **Physics Reports**, vol. 438, no. 5, pp. 237–329, 2007.
- MARWAN, N. and KURTHS, J. Line structures in recurrence plots. **Physics Letters A**, vol 336, n° 4, pp. 349-357, 2005.
- MOLLER, M. F. **A scaled conjugate gradient algorithm for fast supervised learning**. Neural Networks, vol. 6, n° 4, p. 525–533, 1993.
- MONTEIRO, L. H. A. **Sistemas dinâmicos**. Editora Livraria da Física, vol. 3, 2006.
- OLIVEIRA, I. B. **Avaliação fonoaudiológica da voz: reflexões sobre conduta, com enfoque à voz profissional**. In: Ferreira LP, Befi-Lopes DM, Limongi SCD. Tratado de Fonoaudiologia. São Paulo: Roca; 2004b.

- PINHO, S.M. R, PONTES, P., PINHO, S. M. and PONTES, P. Escala de avaliação perceptiva da fonte glótica: Rasat. **Vox Brasilis**, vol. 8, nº 3, pp. 11–3, 2002.
- ROSA, M. de O. **Laringe Digital**. Escola de Engenharia de São Carlos da Universidade de São Paulo. Tese de Doutorado, 2002.
- SANTOS, M. O. **Análise Acústica de Desvios Vocais Infantis utilizando a Transformada Wavelet**. Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, Dissertação de mestrado, 2015.
- SANTOS, J. G. and LIMA, M. F. B. Sintoma vocal e sua provável causa: levantamento de dados em uma população vocal. **Revista CEFAC**, vol 11, nº 1, pp.110-118, 2009.
- SAPIR, S., KEIDAS, A., MATHERS-SCHMIDT B. Vocal attrition in teachers: survey findings. **Eur J Disord Commun**, vol 28, nº 2, pp. 177-85,1993. 1993; 28(2):177-85.
- SAVI, M. A. **Dinâmica não-linear e Caos**. Rio de Janeiro: E-papers, 2006.
- SCALASSARA, P. R. **Utilização de medidas de previsibilidade em sinais de voz para discriminação de patologias de laringe**. Universidade de São Paulo. Tese de doutorado, 2009.
- SIMON, H. **Redes Neurais- Princípios e Prática**. Bookman, 2 ed., 2001.
- SIMSEK, B.; TANSEL, Y.; HÜSEYİN, E. Hybridizing a fuzzy multi-response Taguchi optimization algorithm with artificial neural networks to solve standard ready-mixed concrete optimization problems. In: **Int. J. of Computational Intelligence Systems**, Vol. 9, No. 3; pp.525-543, 2016.
- SOBRINHO, F. A. A. **Medida da dispersão da periodicidade de um sinal de voz normal e voz patológica através da seção de Poincaré**. Universidade de São Paulo, Dissertação de mestrado, 553 p, 2011.
- SOUZA, T.A. **Aplicação da Transformada Wavelet em Análise de Texturas de Gráficos de Recorrência para Detecção de Patologias Laríngeas**. Instituto Federal de Educação, Ciência e Tecnologia da Paraíba. Dissertação de mestrado, 2015.
- SOUZA, E.G. **Caracterização de sistemas dinâmicos de gráficos de recorrência**. Universidade Federal do Paraná. Dissertação de Mestrado, 2008.
- STATISTICS, SPSS. IBM. Version 19.2010.
- TAKENS, F. **Detecting strange attractors in turbulence, in Dynamical systems and turbulence**. Warwick 1980, pp. 366–381, Springer, 1981.
- VAZIRI, G., ALMASGANJ, F. and BEHROOZMAND, R. **Pathological assessment of patients' speech signals using nonlinear dynamical analysis**. Computers in Biology and Medicine, vol. 40, nº. 1, pp. 54–63, 2010.
- VIEIRA, V. J. D. **Avaliação de Distúrbios da Voz por meio de Análise de Quantificação de Recorrência**. Instituto Federal de Educação, Ciência e Tecnologia da Paraíba. Dissertação de mestrado, 217 p, 2014.

- WEBBER, C. Jr. and ZBILUT, J. **Recurrence quantification analysis of nonlinear dynamical systems**. Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences (M. A. Riley and G. C. V. Orden, eds.), 2005.
- WEBBER, C. Jr. and ZBILUT, J. Recurrence quantifications: Feature extractions from recurrence plots. **Int. J. Bifurcation and Chaos**, vol. 17, pp. 3467–3475, 2007.
- WEBBER, C. L. and ZBILUT, J. P. D. Dynamical assessment of physiological systems and states using recurrence plot strategies. **J. Appl. Physiol**, vol. 76, pp. 965–973, 1994.
- YAMASAKI, R., LEÃO, S., MADAZIO, G., PADOVANI, M., AZEVEDO, R. and BEHLAU, M. **Correspondência entre escala analógico-visual e a escala numérica na avaliação perceptivo-auditiva de vozes**. XVI Congresso Brasileiro de Fonoaudiologia, pp.24-27, 2008.
- ZHANG, Y. and JIANG. Acoustic analyses of sustained and running voices from patients with laryngeal pathologies. **Journal of Voice**, vol. 22, no. 1, pp. 1–9, January 2008.
- ZHANG, G. P. Neural networks for classification: a survey. **IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews**, vol. 30, n° 4, pp. 451–462, 2000.

APÊNDICE- Classificações segundo a Escala Analógica Visual (EAV)

Nas Tabelas 29 a 34, são apresentados os valores obtidos pela avaliação perceptivo-auditiva, realizada pelo LIEV/UFPB, referente aos sinais escolhidos para esta pesquisa nos dois casos de estudo.

Tabela 29: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes soprosas.

PREDOMINANTE SOPROSO	EAV (0-100 mm)			EAV(Grau)		
	SOP	RUG	TEN	SOP	RUG	TEN
VGE 10	46	31	14	2	1	1
VGE 12	64	50	57	3	2	3
VGE 15	50	42,5	10	2	2	1
VGE 19	41,5	36	4	2	2	1
VGE 22	45,5	40,5	37	2	2	2
VGE 36	57	47	31	3	2	1
VGE 45	72	62	57,5	3	3	3
VGE 49	60	52	22	3	3	1
VGE 54	46,5	41	38	2	2	2
VGE 58	39,5	18	33	2	1	1
VGE 65	46	39	20,5	2	2	1
VGE 68	49,5	44	10	2	2	1
VGE 78	47	36	7,5	2	2	1
VGE 84	38	18	4	2	1	1
VGE 94	40	28	6	2	1	1
VGE 109	41	35,5	19	2	1	1
VGE 110	93	72	61	4	3	3
VGE 114	50,5	38,5	11,5	2	2	1
VGE 118	51	43	36	3	2	2
VGE 119	48,5	41	35	2	2	1
VGE 121	76	68	43	3	3	2
VGE 125	44	31	3	2	1	1
VGE 159	43	32,5	15	2	1	1
VGE 166	52	36,5	5	3	2	1
VGE 172	55	45	19	3	2	1
VGE 178	49	38	8	2	2	1
VGE 227	56	47,5	34	3	2	1
VGE 304	52	37	5	3	2	1
VGE 319	96	70	83,5	4	3	3
VGE 327	65,5	55,5	13	3	3	1

Tabela 30: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes rugosas.

PREDOMINANTE RUGOSO	EAV (0-100 mm)			EAV (Grau)		
	SOP	RUG	TEN	SOP	RUG	TEN
VGE1	48	79	63	2	3	3
VGE 2	58	66	42	3	3	2
VGE 3	38	47	12	2	2	1
VGE 5	39	42	33	2	2	1
VGE 9	41	51,5	37	2	3	2
VGE 11	68	94	54	3	4	3
VGE 13	48	54	44	2	3	2
VGE 16	36,5	41	10,5	2	2	1
VGE 17	30	40	37	1	2	2
VGE 18	45	51	39	2	3	2
VGE 23	26	38,5	35	1	2	1
VGE 25	47	52	33	2	3	1
VGE 28	50	66	61	2	3	3
VGE 29	66	92	78	3	4	3
VGE 34	49	53,5	43	2	3	2
VGE 39	37	42	20	2	2	1
VGE 41	45	50	38	2	2	2
VGE 43	32,5	52,5	46	1	3	2
VGE 50	3,5	36	32,5	1	2	1
VGE 53	53	59,5	38,5	3	3	2
VGE 56	52	58	33	3	2	1
VGE 75	52,5	62	49	3	3	2
VGE 76	49,5	55,5	30,5	2	3	1
VGE 80	52	61,5	46	3	3	2
VGE 96	40	58,5	37	2	3	2
VGE 97	69	78	69,5	3	3	3
VGE 98	67,5	93	60	3	4	3
VGE 106	6	39,5	35	1	2	1
VGE 107	38	45	33,5	2	2	1
VGE 108	7	48,5	19,5	1	2	1

Tabela 31: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes tensas.

PREDOMINANTE TENSA	EAV (0-100 mm)			EAV (Grau)		
	SOP	RUG	TEN	SOP	RUG	TEN
VGE 4	11	36	41	1	2	2
VGE 14	10	57	64	1	3	3
VGE 21	7	57	64	1	3	3
VGE 30	61,5	55,5	68	3	3	3
VGE 31	28,5	40	46,5	1	2	2
VGE 37	5	40	46,5	1	2	2
VGE 38	16	38	43	1	2	2
VGE 42	3,5	61	93	1	3	4
VGE 60	3	45,5	52	1	2	3
VGE 70	18,5	46	54	1	2	3
VGE 81	5	43	55,5	1	2	3
VGE 100	9	36	40	1	2	2
VGE 148	27	45	50	1	2	2
VGE 156	3	50	61	1	2	3
VGE 180	5	66	80	1	3	3
VGE 200	3	56	64	1	3	3
VGE 248	6	44	55	1	2	3
VGE 284	8	44,5	52	1	2	3
VGE 299	8	39,5	48,5	1	2	2
VGE 312	3	39	47	1	2	2
VGE 530	5	43,5	54,5	1	2	3
VGE 561	9	41	54	1	2	3
VGE 583	6,5	36	42,5	1	2	2
VGE 607	37	44	51,5	2	2	3
VGE 654	4	37	45	1	2	2
VGE 711	4	37	44	1	2	2
VGE 719	8	37	50	1	2	2
VGE 722	4,5	39,5	54	1	2	3
VGE 736	69,5	85	96	3	3	4
VGE 763	6	48	48,5	1	2	2

Tabela 32: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes com predominância de grau 2.

PREDOMINANTE GRAU 2 (G2)	EAV (0-100 mm)			TIPO DE DESVIO
VGE 3	47	38	12	RUG
VGE 5	42	39	33	RUG
VGE 16	41	36,5	10,5	RUG
VGE 17	40	30	37	RUG
VGE 23	38,5	26	35	RUG
VGE 39	42	37	20	RUG
VGE 41	50	45	38	RUG
VGE 50	36	3,5	32,5	RUG
VGE 106	39,5	6	35	RUG
VGE 107	45	38	33,5	RUG
VGE 10	31	46	14	SOP
VGE 15	42,5	50	10	SOP
VGE 19	36	41,5	4	SOP
VGE 22	40,5	45	37	SOP
VGE 54	41	46	38	SOP
VGE 78	36	47	7,5	SOP
VGE 58	18	39,5	33	SOP
VGE 65	39	46	20,5	SOP
VGE 68	44	49,5	10	SOP
VGE 94	28	40	6	SOP
VGE 4	36	11	41	TEN
VGE 31	40	28,5	46,5	TEN
VGE 37	40	5	46,5	TEN
VGE 38	38	16	43	TEN
VGE 100	36	9	40	TEN
VGE 148	45	27	50	TEN
VGE 299	39,5	8	48,5	TEN
VGE 711	37	4	44	TEN
VGE 719	37	8	50	TEN
VGE 763	48	6	48,5	TEN

Tabela 33: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes com predominância de grau 3.

PREDOMINANTE GRAU 3 (G3)	EAV (0-100 mm)			TIPO DE DESVIO
VGE 1	79	48	63	RUG
VGE 2	66	58	42	RUG
VGE 9	51,5	41	37	RUG
VGE 13	54	48	44	RUG
VGE 18	51	45	39	RUG
VGE 25	52	47	33	RUG
VGE 28	66	50	61	RUG
VGE 53	59,5	53	38,5	RUG
VGE 75	62	52,5	49	RUG
VGE 76	55,5	49,5	30,5	RUG
VGE 80	61,5	52	46	RUG
VGE 96	58,5	40	37	RUG
VGE 12	50	64	57	SOP
VGE 36	47	57	31	SOP
VGE 45	62	72	57,5	SOP
VGE 322	48	59	12	SOP
VGE 327	55,5	65,5	13	SOP
VGE 121	68	76	43	SOP
VGE 172	45	55	19	SOP
VGE 118	43	51	36	SOP
VGE 14	57	10	64	TEN
VGE 21	57	7	64	TEN
VGE 30	55,5	61,5	68	TEN
VGE 60	45,5	3	52	TEN
VGE 70	46	18,5	54	TEN
VGE 81	43	5	55,5	TEN
VGE 156	50	3	61	TEN
VGE 180	66	5	80	TEN
VGE 200	56	3	64	TEN
VGE 248	44	6	55	TEN

Tabela 34: Classificação segundo a Escala Analógica Visual (EAV) para sinais de vozes da junção dos graus (G2 + G3).

G2+G3	EAV (0-100 mm)			TIPO DE DESVIO
VGE 1	79	48	63	RUG
VGE 2	66	58	42	RUG
VGE 9	51,5	41	37	RUG
VGE 13	54	48	44	RUG
VGE 18	51	45	39	RUG
VGE 12	50	64	57	SOP
VGE 36	47	57	31	SOP
VGE 45	62	72	57,5	SOP
VGE 322	48	59	12	SOP
VGE 327	55,5	65,5	13	SOP
VGE 14	57	10	64	TEN
VGE 21	57	7	64	TEN
VGE 30	55,5	61,5	68	TEN
VGE 60	45,5	3	52	TEN
VGE 70	46	18,5	54	TEN
VGE 3	47	38	12	RUG
VGE 5	42	39	33	RUG
VGE 16	41	36,5	10,5	RUG
VGE 17	40	30	37	RUG
VGE 23	38,5	26	35	RUG
VGE 10	31	46	14	SOP
VGE 15	42,5	50	10	SOP
VGE 19	36	41,5	4	SOP
VGE 22	40,5	45	37	SOP
VGE 54	41	46	38	SOP
VGE 4	36	11	41	TEN
VGE 31	40	28,5	46,5	TEN
VGE 37	40	5	46,5	TEN
VGE 38	38	16	43	TEN
VGE 100	36	9	40	TEN