María Eugenia Dajer

**PADRÕES VISUAIS DE SINAIS DE VOZ ATRAVÉS
DE TÉCNICA DE ANÁLISE NÃO LINEAR**

Dissertação apresentada ao Programa de Pós Graduação Interunidades em Bioengenharia/ Escola de Engenharia de São Carlos/ Faculdade de Medicina de Ribeirão Preto/ Instituto de Química de São Carlos da Universidade de São Paulo, como parte dos requisitos para a obtenção do Título de Mestre em Bioengenharia.

Orientador: Prof. Dr. José Carlos Pereira

São Carlos

2006

Dedico mi trabajo a Moni, Martha, Abudi, Tami y Nati, por el apoyo incondicional, el cariño y fundamentalmente por darme una FAMILIA.

AGRADECIMENTOS

A mis padres, porque la certeza de su presencia es mi fortaleza, mi luz y mi alegría.

A mis hermanos cordobeses y jujeños, Marisa, Pichi y Tutis, Cecil, Fede, Cesar, Viky, Robert, Virginia, Tomas, por demostrarme que existen vínculos que trascienden al tiempo y a la distancia.

Ao Prof. José Carlos Pereira, pela oportunidade e confiança de chefe, a paciência e a dedicação de pai, a guia de mestre, e o exemplo de vida. “Muchas Gracias, Pai”!!!

Ao Prof. Carlos Maciel, por a atenção e apoio no trabalho, pela amizade e a compreensão no dia a dia e as inúmeras lições de vida .

A os Prof. da Bioengenharia, Orivaldo e José Marcos, e ao Prof. Cintra, pela enorme contribuição para meu crescimento acadêmico e pessoal. E a Janete, pela paciência e a constante disposição para com todos nós.

Ao Prof. NeRson, pela disposição e ajuda desinteressada, a amizade e a terapia (de graça!!!).

Aos colegas da Bioengenharia, por os muitos bons momentos nas aulas e no “trabalho e o estudo”.

A Dona Judith, por tanto carinho e compreensão.

À Fga. Renata Fúria, por ser a peça fundamental na “mudança de planos”, e uma valiosa amiga e mestra, na profissão e na vida. E a Fga. Thaís Vanzela, colega na profissão e o trabalho e amiga extraordinária na vida.

Aos companheiros do LIM (Lindos, Inteligentes e *Maravishosos*), Paré, Mori, Paulo, Selma, Julián, Karen, Everthon, Leandro, Paulinha, e os “meninos

da Iniciação” Zé, Buda, Feliz, Sarubbi, Fabio, pelo apoio, ajuda e a disposição nas intermináveis horas de trabalho.

Aos funcionários da “Elétrica”, que fazem que todo “funcione”. Especialmente a Anderson, Alex e Roseli, por tantos pepinos resolvidos.

Ao pessoal da “Elétrica” (labs vizinhos e redondezas), pelos churras, o café da 4, as saídas, as festas, e as importantíssimas contribuições no meu aprendizado da língua portuguesa. Valeu galera!!!

A Aline e Cleber, que em pouco tempo de convivência tornaram-se pessoas e amigos valiosíssimos.

A ERtinho “Gaúcho” Valentini e Fernando “Brother” Castro, meus companheiros, meus amigos, meus irmãos. Muito obrigada pela alegria, a compreensão, o apoio, as idéias, o aprendizado, o carinho, e principalmente pelo valioso tempo compartilhado.

E por sobre todo e todos, a DEUS que é Pai, e me abençoa a cada passo com pessoas maravilhosas que são meus companheiros, mestres e meus irmãos.

RESUMO

DAJER, M.E. (2006). Padrões visuais de sinais de voz através de técnica de análise não linear. Dissertação (Mestrado)- Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006.

A voz tem sido objeto de estudos em diferentes áreas da ciência. Nas últimas duas décadas os pesquisadores demonstraram a presença do caos na produção de voz. O objetivo deste trabalho é estabelecer padrões visuais de sinais de voz humana através da técnica não linear de reconstrução de espaço de fase e associá-los com suas correspondentes análises auditivo-perceptiva e acústica. Foram analisados sinais de voz humana de sujeitos de ambos os gêneros, na faixa etária de 19 a 39 anos. Foram gravadas as vogais sustentadas /a/, /e/ e /i/ do português do Brasil, com uma taxa de amostragem de 22.050 Hz. Os sinais foram analisados a fim de obter medidas acústicas (Jitter, Shimmer e Coeficiente de Excesso). Foi utilizada a técnica de reconstrução de espaço de fase para descrever as características de dinâmica não linear dos sinais de voz, e para análise perceptivo auditiva foi utilizada a escala RASAT. Os resultados demonstram que métodos de dinâmica não linear como a reconstrução do espaço de fase, parece ser uma técnica apropriada para análise de sinais de voz, incorporando os componentes caótico e determinístico da voz humana. É sugerido que análise da dinâmica não linear não substitui as técnicas existentes, embora possa aperfeiçoar e complementar os métodos de análise disponíveis para os profissionais da saúde, como fonoaudiólogos e otorrinolaringólogos.

Palavras Chaves: análise não linear, análise de voz, caos determinístico, séries temporais.

ABSTRACT

Dajer, M.E. (2006). Voice signal discrimination with nonlinear analysis techniques. M.Sc. Dissertation. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006.

Human voice has been the focus of study for different areas of science. Researches in the last two decades have demonstrated the existence of chaos in human voice production. The purpose of this work is to use nonlinear dynamics techniques in the analysis of normal voices from healthy subjects and correlate them to traditional acoustic parameters as well as perceptual analysis. Human voice signals from healthy subjects, both male and female, ranging in age from 19 to 39 years old were analyzed. Sustained vowel sounds /a/, /e/ and /i/, from Brazilian Portuguese were recorded at a sampling rate of 22,050 Hz and analyzed in order to obtain acoustic measures (Jitter, Shimmer and coefficient of excess – EX). The phase space reconstruction technique was used to describe the nonlinear dynamic characteristics of voice signal samples. The results show, that non-linear dynamical method as phase space reconstruction seems to be a suitable technique for voice signals analysis, due to the chaotic component of the human voice. It is suggested, that non-linear dynamic analysis does not replace existing techniques instead, it may improve and complement the recent voice analysis methods available for health professionals, speech therapist and clinician.

Keywords: nonlinear analysis, voice analysis, deterministic chaos, temporal series.

LISTA DE FIGURAS

Figura 2.1:	Desenho esquemático dos sistemas envolvidos na produção dos sinais de voz.	16
Figura 2.2:	Desenho esquemático do sistema fonte- filtro da produção de voz.	17
Figura 2.3:	Representação esquemática da localização da laringe	18
Figura 2.4:	Vista anterior (esquerda) das cartilagens laríngicas Vista posterior (direita) das cartilagens laríngicas	19
Figura 2.5:	Desenho de um corte transversal da laringe. Músculos intrínsecos.	20
Figura 2.6:	Desenho de um corte sagital da laringe. Músculos intrínsecos.	20
Figura 2.7:	Esquema de um corte longitudinal das pregas vocais. modelo de "corpo e cobertura" de Hirano.	22
Figura 2.8:	Imagem de um corte histológico das pregas vocais.	23
Figura 2.9:	Imagem eletroestroboscópica das pregas vocais normais em fonação.	24
Figura 2.10:	Desenho esquemático do mecanismo das pregas vocais para a produção do pulso glotal.	25
Figura 2.11:	Imagem das pregas vocais em posição respiratória. Abertura glotal máxima.	27
Figura 2.12:	Imagem das pregas vocais em posição fonatória. Fechamento glotal máximo.	27
Figura 2.13A:	Desenho esquemático do espectro de frequências do som fundamental sem amplificação.	28
Figura 2.13B:	Curvas de ressonância do trato vocal.	28

Figura 2.13C:	Espectro resultante da frequência fundamental mais as ressonâncias do trato vocal.	28
Figura 2.14:	Exemplo de onda complexa de voz de uma vogal sustentada /e/.	34
Figura 2.15:	Exemplo de onda complexa de voz de uma vogal sustentada /e/.	34
Figura 2.16:	Exemplo de um espectrograma de uma vogal sustentada.	35
Figura 2.17:	Exemplo de um espectro de frequência de um sinal de voz, da vogal sustentada /a/	36
Figura 2.18:	Desenho da reconstrução do espaço de fase do comportamento de um pendulo	38
Figura 2.19:	Desenho esquemático de o ponto atrator.	39
Figura 2.20:	Desenho de um atrator periódico (circular).	40
Figura 2.21:	Desenho de um atrator estranho ou atrator de Lorentz.	40
Figura 2.22:	Desenho de um diagrama de bifurcação.	42
Figura 2.23:	Esquema das trajetórias de duas órbitas de um Sistema dinâmico	45
Figura 2.24A:	Exemplo de órbitas de um sistema dissipado com um ponto fixo atrator.	47
Figura 2.24B:	Exemplo de órbitas de um sistema dissipado com uma órbita atratora.	47
Figura 2.25:	Exemplo de duas órbitas de um sistema conservativo com um ponto fixo neutral e órbitas fixas neutrais.	47
Figura 2.26:	Órbitas num sistema discreto com $MEL > 0$.	48
Figura 3.1:	Sinal de voz completo da vogal sustentada /e/.	56
Figura 3.2:	Gráfico do trecho selecionado do sinal de voz da vogal sustentada /e/.	57

Figura 3.3:	Gráfico do espectro de freqüências do sinal de voz.	58
Figura 3.4:	Gráfico da curva de informação mutua por atrasos de amostras	59
Figura 3.5:	Gráfico da reconstrução do espaço de fase do trecho selecionado da vogal sustentada /e/.	60
Figura 3.6:	Gráfico da reconstrução do espaço de fase do da vogal sustentada /e/ em função do tempo.	61
Figura 3.7:	Gráfico da reconstrução do espaço de fase do da vogal sustentada /e/ em 3 dimensões .	62
Figura 3.8:	Curva de divergência das órbitas. Curva do Máximo expoente de Lyapunov	63
Figura 4.1:	Reconstrução do espaço de fase de um sinal artificial com uma freqüência pura.	66
Figura 4.2:	Reconstrução do espaço de fase de um sinal artificial no tempo de uma freqüência pura.	66
Figura 4.3:	Reconstrução do espaço de fase de um sinal artificial com 15 % de shimmer.	67
Figura 4.4:	Imagem aumentada do traçado das órbitas do espaço de fase (figura 4.3).	68
Figura 4.5:	Espaço de fase reconstruído de um sinal artificial com 3 % de jitter.	69
Figura 4.6:	Imagem aumentada do traçado das órbitas dos atratores na reconstrução do espaço de fase	70
Figura 4.7:	Exemplo da reconstrução do espaço de fase em 3 D	70
Figura 4.8:	(superior) Exemplo de um sinal periódico (inferior) reconstrução do espaço de fase do sinal.	71
Figura 4.9:	(superior) Exemplo de um sinal F0 e uma harmônica. (inferior) reconstrução do espaço de fase	71
Figura 4.10:	(superior) Sinal periódico F0 e dois harmônicos. (inferior) reconstrução do espaço de fase	72

Figura 4.11:	(superior) Sinal periódico F0 e três harmônicas. (inferior) reconstrução do espaço de fase .	72
Figura 4.12:	Reconstrução do espaço de fase de sinais de voz humana da vogal sustentada /a/.	73
Figura 4.13:	Reconstrução do espaço de fase de sinais de voz humana da vogal sustentada /e/.	74
Figura 4.14:	Reconstrução do espaço de fase de sinais de voz humana da vogal sustentada /i/.	75
Figura 4.15A:	Reconstrução do espaço de fase da vogal sustentada /e/	78
Figura 4.15B:	Reconstrução do espaço de fase da vogal sustentada /e/	78
Figura 4.16:	Reconstrução do espaço de fase da vogal sustentada /a/	80
Figura 4.17:	Reconstrução do espaço de fase da vogal sustentada /a/	81
Figura 4.18:	Imagem aumentada das órbitas dos atratores do sinal de voz da figura 4.16	82
Figura 4.19:	Imagem aumentada das órbitas dos atratores do sinal de voz da figura 4.17.	82
Figura 4.20:	Gráfico do MEL do sinal de voz da figura 4.16.	83
Figura 4.21:	Gráfico do MEL do sinal de voz da figura 4.18.	84
Figura 4.22:	Espaço de fase de voz com 0.4 % de jitter 0.98 % de shimmer e 139.13 de Ex	85
Figura 4.23:	Espaço de fase de voz com 0.46 % de jitter, 2.91 % de shimmer e 129.73 de Ex	86
Figura 4.24:	Imagem aumentada dos atratores da figura 4.22.	87
Figura 4.25:	Imagem aumentada dos atratores da figura 4.23.	87
Figura 4.26:	Gráfico do MEL da figura 4.22.	88

Figura 4.27:	Gráfico do MEL da figura 4.23.	89
Figura 4.28:	Espaço de fase de voz com 0.63 % de jitter, 2.4 % de shimmer e 110.64 de Ex,	90
Figura 4.29:	Espaço de fase de voz com 2.11% de jitter, 3.38 % de shimmer e 103.7 de Ex	91
Figura 4.30:	Imagem aumentada dos atratores da figura 4.28.	92
Figura 4.31:	Imagem aumentada dos atratores da figura 4.29.	93
Figura 4.32:	Gráfico do MEL da figura 4.28.	94
Figura 4.33:	Gráfico do MEL da figura 4.29.	94

ÍNDICE

RESUMO		v
ABSTRACT		
LISTA DE FIGURAS		vii
1	INTRODUÇÃO	14
1.1	Objetivo	15
2	REVISÃO BIBLIOGRÁFICA	16
2.1	VOZ HUMANA	16
2.1.1	Produção De Voz Humana	16
2.1.2	Estrutura e funções da laringe	18
2.1.2.1	Estrutura da Laringe	18
2.1.2.2	Funções da Laringe	25
2.1.2.2.1	Função de Proteção	25
2.1.2.2.2	Função respiratória	26
2.1.2.2.3	Função fonatória	27
2.1.3	Avaliação acústica e perceptiva auditiva de voz humana	29
2.1.3.1	Avaliação perceptiva auditiva da voz.	30
2.1.3.2	Avaliação acústica dos sinais de voz	32
2.1.3.2.1	Parâmetros acústicos no domínio do tempo	33
2.1.3.2.2	Parâmetros acústicos no domínio da frequência	34
2.2	CAOS DETERMINÍSTICO E ANÁLISE NÃO-LINEAR	37
2.2.1	Introdução:	37
2.2.2	Sistemas não lineares e Dinâmica não linear	38
2.2.2.1	Propriedade de bifurcação	41
2.2.2.2	Propriedade Fractal	42
2.2.2.3	Reconstrução do espaço de fase	43

2.2.2.4	Dimensão de correlação	44
2.2.2.5	Expoente de Lyapunov	45
2.2.2.6	Entropia de Kolmogorov	48
2.2.3	Caos na produção dos sinais de voz humana.	49
3	MATERIAIS E MÉTODOS	52
3.1	Materiais	52
3.1.1	Base de dados:	52
3.2	Métodos	53
3.2.1	Análise tradicional de parâmetros de perturbação acústica	53
3.2.2	Avaliação perceptivo auditiva	53
3.2.2.1	Procedimento	54
3.2.3	Análise de dinâmica não linear	54
3.2.3.1	Procedimento	55
4	RESULTADOS E DISCUSSÃO	64
4.1	Introdução:	64
4.2	Análise tradicional de parâmetros de perturbação acústica.	64
4.3	Avaliação Perceptiva Auditiva. "RASAT".	64
4.4	Análise Não Linear "ANL"	65
4.4.1	Validação da ferramenta.	65
4.4.1.1	Reconstrução do espaço de fase e parâmetros acústicos	65
4.4.1.2	Efeitos da complexidade do sinal de voz na reconstrução	71
4.4.2	Implementação da ferramenta "ANL" em de sinais de voz.	72
4.4.2.1	Estabelecimento de padrões das vogais de voz	72
4.4.2.2	Reconstrução do espaço de fase, cálculo da convergência	76
4.4.2.2.1	Reconstrução do espaço de fase e (Ex)	77
4.4.2.2.2	Reconstrução do espaço de fase e Shimmer	85
4.4.2.2.3	Reconstrução do espaço de fase e Jitter .	90
5	CONCLUSÃO:	96
6	REFERENCIAS BIBLIOGRAFICAS	98

1- INTRODUÇÃO:

A voz humana, principal responsável pela comunicação, é um sinal acústico formado por diversos sistemas cuja fonte é a laringe, local onde se encontram as pregas vocais.

A conjunção das dinâmicas de diversos sistemas de órgãos (sistema respiratório, sistema fonatório ou glotal, sistema de ressonância, sistema articulatorio, sistema nervoso, sistema ósseo-cartilaginoso e muscular) é a responsável da produção de voz humana, absolutamente individual e única para cada sujeito.

Para considerar um sinal vocal como estável e adequado, é necessário equilíbrio, coordenação e sinergia no acoplamento dos diferentes sistemas envolvidos na sua produção.

Devido à complexidade, aos inúmeros fatores envolvidos neste processo, e à importância da fala como ferramenta da comunicação, a voz tem sido objeto de estudos nos últimos anos visando tanto seu entendimento, através da análise acústica, quanto o desenvolvimento de técnicas de processamento e síntese de voz.

Diversos estudos observaram que as pequenas flutuações na frequência, amplitude e formato de onda se encontram sempre presentes nos sinais vocais (Banbrook, M; McLaughlin, S; Mann, I., 1999), (Goldberger, A L., 1996), refletindo os acoplamentos entre os diversos sistemas envolvidos na produção da voz.

Todo intento de um sujeito produzir sons perfeitamente regulares termina em uma forma de onda não periódica, característica fundamental da voz. Apesar de parecer movimentos estáveis numa escala macroscópica eles são sempre quase periódicos vistos em escala microscópica (Herzel, H., Berry,

D., Titze, I.R., & Saleh, M. , 1994) (Kumar, A.; Mullick, K., 1996). Esses micro-movimentos provocam flutuações em o que a simples vista parece uma atividade estável e uniforme.

As estruturas anatômicas e os processos fisiológicos envolvidos no complexo sistema de produção dos sinais de voz humana são responsáveis pelos diversos fonemas emitidos e conseqüentemente carregam informações sobre suas formas, fisiologia e dinâmica. A produção dos sinais de voz exige um adequado funcionamento e condições das estruturas do trato vocal. Qualquer interferência na dinâmica do processo implica na geração de um sinal acústico irregular, geralmente exigindo algum esforço adicional em sua produção.

Várias ferramentas para análise não linear têm sido propostas na literatura. Particularmente, os sinais sonoros apresentam uma componente determinística associada a uma componente aleatória. A análise através de *caos determinístico* parece ser a técnica mais adequada para os sinais de voz.

Diante da complexidade do fenômeno de produção da voz humana e da necessidade de uma compreensão mais aprofundada da dinâmica da produção dos sinais vocais, o objetivo deste trabalho é utilizar técnicas não lineares para análise do sinal de voz e associá-los com sua correspondente análise auditivo-perceptiva e acústica da voz.

1.1 Objetivo:

Estabelecer padrões visuais de sinais de voz humana normal através da técnica não linear de reconstrução de espaço de fase e associá-los com suas correspondentes análises auditivo-perceptiva e acústica.

2- REVISÃO BIBLIOGRÁFICA

2.1 VOZ HUMANA.

2.1.1 Produção dos sinais de voz

A voz humana é uma emissão acústica voluntária, considerada principal ferramenta de comunicação, de alta complexidade e inúmeros fatores envolvidos na sua produção. Para produzir um sinal de voz, é necessária a sincronia de um complexo sistema, que subdividimos em diferentes sub-sistemas (Behlau, 2001 p.26), esquematizados na figura 2.1:

- Sistema Respiratório. Fonte de Energia (pulmões)
- Sistema Fonatório. Fonte Vibratória (pregas vocais)
- Sistema de Ressonância. Fonte de Ressonância (cavidade oral e nasal)
- Sistema Articulatório. Articuladores (língua, lábios, mandíbula, palato, dentes)
- Sistema Nervoso Central e Periférico. Coordenação. (Córtex, estruturas sub-corticais e nervos).

Figura 2.1: Desenho esquemático dos sistemas envolvidos na produção dos sinais de voz.

Podemos representar o sistema de produção de voz de forma esquemática como um sistema de fonte-filtro, onde as pregas vocais serão consideradas a fonte sonora e o trato vocal, o filtro. O sinal resultante será o sinal acústico de voz humana. (Rabiner; Juang, 1993). (figura 2.2).

Figura 2.2: Desenho esquemático do sistema fonte- filtro da produção de voz.
<http://ispl.korea.ac.kr/~wikim/research/speech.html>

De forma resumida a voz é produzida a partir de um som básico gerado pela vibração das pregas vocais, o fonte glótica, localizadas na laringe. Este som é denominado *fonte sonora* ou *pulso glotal* – responsável pela frequência fundamental.

2.1.2 Estrutura e funções da laringe

2.1.2.1 Estrutura da Laringe

A *laringe* se localiza na parte anterior do pescoço, tem formato de tubo alongado, e comunica a faringe com a traquéia (figura 2.3). A fins didáticos pode ser dividida em: um espaço constituído por todas as estruturas acima da glote ou espaço supraglótico; um espaço que se estende da glote até o primeiro anel traqueal ou espaço infraglote; e o espaço entre as pregas vocais ou glote.

Figura 2.3: Representação esquemática da localização da laringe
<http://medimages.healthopedia.com/large/larynx.jpg>

A laringe é um órgão que constituído por: cartilagens, músculos, membranas e mucosa. As cartilagens laríngeas podem ser classificadas em pares e ímpares. Sendo três ímpares: a tireóidea, a cricóidea e a epiglote; e três cartilagens pares, a aritenóidea, considerada principal, as corniculadas e as cuneiformes, consideradas acessórias ou secundárias. Estas estruturas cartilagueas são sustentadas principalmente pelo osso hióide, também ponto de inserção para os músculos e ligamentos da laringe a língua e outros órgãos. Na figura 2.4 pode-se observar um desenho esquemático das cartilagens laríngeas em visão anterior e posterior.

Figura 2.4: Vista anterior (esquerda) das cartilagens laringeas e o osso Hióide. Vista posterior (direita) das cartilagens laringeas e da traquéia.
webanatomy.net/.../respiratory/larynx_images.jpgfv

A musculatura laringea pode ser classificada segundo o tipo de inserção. Os músculos esqueléticos com origem e inserção na laringe são denominados músculos intrínsecos; e os músculos com inserção na laringe e em estruturas externas são denominados extrínsecos. (Tucker, 1993)

Os músculos intrínsecos têm maior proporção de fibras musculares tipo IIA que outros músculos do corpo (Henick, Sataloff, 1997). E o tiroaritenóideo (TA) e o cricoaritenóideo lateral (CAP) são músculos altamente especializados em contração rápida. A musculatura intrínseca é responsável pela adução, a abdução e a tensão das pregas vocais, e participa ativamente nas funções de respiração, esfíncter e fonação. São seis pares: Tiroaritenóideo (TA), Cricoaritenóideo Posterior (CAP), Cricoaritenóideo Lateral (CAL), Cricotiróideo (CT), Ariepliglótico (AE) e Tireoepiglótico (TE); sendo o aritenóideo (A) ou interaritenóideo o único músculo ímpar. As figuras 2.5 e 2.6 ilustram de forma esquemática a musculatura intrínseca antes mencionada. (Hirano, 1981, p.69).

Figura 2.5: Desenho esquemático de um corte transversal da laringe. Músculos intrínsecos. Tiroaritenóideo (TA) e Aritenóideo (A) ou interaritenóideo (IA)
[escuela.med.puc.cl/.../ AnatLaringea09.gif](http://escuela.med.puc.cl/.../AnatLaringea09.gif)

Figura 2.6: Desenho esquemático de um corte sagital da laringe. Músculos intrínsecos. Ariepliglótico (AE), Tiroepiglótico (TE), Tiroaritenóideo (TA), Cricoaritenóideo Lateral (CAL) e Cricoaritenóideo Posterior (CAP).
[escuela.med.puc.cl/.../ AnatLaringea07.gif](http://escuela.med.puc.cl/.../AnatLaringea07.gif)

Os músculos extrínsecos têm como função principal manter e proporcionar estabilidade à laringe nas estruturas do pescoço e participam de forma indireta na fonação, mas são considerados importantes por serem os responsáveis do mecanismo secundário de controle de frequência (Le Huch, F. 1993). Estes grupos musculares podem ser classificados segundo sua localização em relação ao osso hióide em músculos supra-hióideos, que

elevam a laringe no pescoço; e em músculos infra-hióideos, que abaixam a laringe.

As cartilagens e músculos laríngeos encontram-se interconectados por ligamentos e membranas intrínsecos e extrínsecos, estes tecidos tem como função básica a interligação das estruturas laríngeas.

No interior da laringe encontram-se as pregas vestibulares e as *pregas vocais*. As pregas vestibulares ou bandas ventriculares, erroneamente chamadas de falsas pregas vocais, são duas dobras de tecido espesso e mole com glândulas internas que se localizam por cima das pregas vocais. (Hirano, 1993, 1996).

Segundo o modelo de “corpo e cobertura” proposto por Hirano no começo dos anos 70, as pregas vocais são duas estruturas multilaminadas, formadas basicamente por músculo e mucosa, constituídas por camadas com propriedades mecânicas diferentes e que se encontram localizadas em posição horizontal dentro da laringe e em sentido antero-posterior. Este modelo divide a prega vocal em mucosa, por sua vez, subdividida em duas camadas, o epitélio e lamina própria; e o músculo constituído pelo músculo vocal. (Hirano; Kakita, 1985, pp. 1-46).

O epitélio da mucosa das pregas vocais está formado por tecido escamoso e estratificado e sua função é cobrir e manter a forma da prega vocal. A lamina própria subdivide-se em três camadas: superficial, intermediária e profunda, onde cada uma delas apresenta características diferentes. Na figura 2.7 pode-se observar um esquema das diferentes camadas das pregas vocais segundo o modelo de Hirano. (Hirano, 1989, pp. 89-144).

Figura 2.7: Esquema de um corte longitudinal das pregas vocais. Pode-se visualizar as diferentes camadas segundo o modelo de “corno e cobertura” de Hirano.

A camada superficial da lamina própria (CSLP), também chamada de espaço de Reike, se apresenta como um tecido gelatinoso, solto e flexível; esta característica permite um descolamento acentuado, sendo a camada que vibra mais intensamente durante a fonação. A camada intermediária da lamina própria (CILP) está composta por um grande número de fibras elásticas e caracteriza-se por ser uma camada com maior densidade. A camada profunda da lamina própria (CPLP), é constituída por maior quantidade de fibras colágenas que proporcionam maior rigidez. Não existe um limite bem definido entre as camadas intermediária (CILP) e profunda (CPLP), mas pode-se observar a diminuição de fibras elásticas e o aumento de fibras colágenas à medida que os tecidos se aproximam do corpo da prega vocal. (Hirano; Kakita, 1985, pp. 1-46). (Hirano, 1996).

O corpo da prega vocal está constituído pelo músculo vocal ou músculo tiroaritenóideo, e funciona como um feixe elástico rígido no momento da contração muscular. É composto por fibras elásticas, fibras colágenas e fibras musculares distribuídas longitudinal e paralelamente à borda livre da prega vocal. O limite entre a camada profunda da lamina própria (CPLP) e o músculo vocal é difuso, sendo que algumas fibras da camada profunda inserem-se no músculo como se pode observar na figura 2.1.8. (Gray, Hirano, Sato, 1993).

Figura 2.8: Imagem de um corte histológico das pregas vocais. Pode-se visualizar a concentração das diferentes fibras que diferenciam cada uma das camadas.
www3.umdj.edu/.../lab10/images/ftruecords.jpg

Segundo o modelo de corpo e cobertura as cinco camadas descritas podem ser agrupadas segundo sua estrutura e propriedades mecânicas em três seções: cobertura (epitélio e CSPL), transição (CIPL) e corpo (músculo vocal).

Hirano observou em estudos com luz estroboscópica que durante a fonação a cobertura da prega vocal apresenta um deslocamento intenso e fluido em direção contrária a gravidade, enquanto o corpo oferece estabilidade e tono muscular de base (figura 2.9). Por sua vez a camada de transição tem a

função de acoplamento entre a camada superficial da lamina própria e o músculo vocal. (Hirano; Bless, 1993)

Figura 2.9: Imagem eletrostroboscópica das pregas vocais normais em fonação. Podem-se visualizar o deslocamento da mucosa (onda mucosa), produzida pela passagem do fluxo aéreo das pregas em vibração. www.laryngograph.com/img/digivid.png

Portanto, segundo este modelo, a passagem do fluxo aéreo proveniente dos pulmões e a diferença de pressões da cavidade (efeito Bernoulli) associadas às propriedades mioelásticas e mecânicas das pregas vocais produzem a vibração da cobertura, gerando o som vocal ou pulso glotal. (figura 2.10)

Figura 2.10: Desenho esquemático do mecanismo das pregas vocais para a produção do pulso glotal.
biorobotics.harvard.edu/research/heather2.gif

2.1.2.2 Funções da Laringe

A laringe tem três funções principais; proteção de vias aéreas, respiração e fonação. As funções de proteção e respiração são consideradas como as mais primitivas, sendo a fonação, do ponto de vista filogenético, uma função desenvolvida posteriormente. (Zemlin, 2000)

2.1.2.2.1 Função de Proteção

A laringe apresenta dois mecanismos de proteção das vias aéreas inferiores. Um deles é reflexo da tosse, este tem três fases: a inspiratória, na qual laringe se abre amplamente para permitir uma inspiração rápida e profunda; a fase compressiva, onde a glote se fecha bruscamente e se ativam os músculos expiratórios; e finalmente a fase expulsiva, na qual as pregas vocais verdadeiras oferecem mais resistência aos fluxos inspiratórios que expiratórios. Nesta fase a laringe subitamente se abre de forma abrupta e ampla, criando um fluxo expiratório rápido e explosivo. A saída de ar forçada elimina o muco ou outro material que tenha iniciado o reflexo na via aérea. (Sasaki, Weaver, 1997)

O outro mecanismo de proteção acontece durante a deglutição. (Lopes, Campos, 1994). Esta ação é reflexa e involuntária; é desencadeada por receptores nervosos da mucosa da faringe e laringe, que produz o fechamento da glote e a inibição da respiração. Este fechamento inicia-se com a aproximação das pregas vocais é seguida pela aproximação das bandas ventriculares entre si e contra a base da epiglote. A ação da musculatura extrínseca contribui no fechamento esfinteriano de estruturas da região supraglótica. A elevação da laringe provoca um aumento da pressão intralaríngea que promove a tração da epiglote contra as bandas ventriculares, auxiliando no fechamento da laringe no processo de deglutição.

2.1.2.2.2 Função respiratória

Desde o ponto de vista filogenético a respiração é a função mais antiga e a que ocupa maior parte de tempo (Belhau, 2001.p. 25). A glote abre durante uma fração de segundo antes que o fluxo de ar passe por seu interior pela contração do diafragma; esta abertura é realizada pela contração dos músculos cricoaritenóideos posteriores e é coordenado pelo centro respiratório. Durante a respiração, as pregas vocais ficam afastadas uma da outra, o que permite que o fluxo aéreo dos pulmões ingresse e saia sem dificuldades. (Zemlin, 2000). A figura 2.11 mostra a imagem das pregas vocais de um sujeito adulto, sem queixa vocal, em posição respiratória.

A descida da laringe durante a inspiração provoca o estiramento das pregas vocais e das bandas ventriculares, as cartilagens aritenóideas deslizam lateralmente; e a glote se abre. Como resultado das variações de tamanho da glote durante a respiração, a laringe possivelmente desempenha um papel importante na regulação da resistência intrínseca da via aérea durante a respiração. (Zemlin, 2000).

Figura 2.11: Imagem das pregas vocais em posição respiratória. Abertura glotal máxima.
<http://www.texasvoicecenter.com/diseases.html>

2.1.2.2.3 Função fonatória

A função fonatória é provavelmente a mais complexa das três funções da laringe. Segundo alguns autores é uma função neurofisiológica inata, mas que se forma ao longo da vida, de acordo com as características anatomofuncionais do indivíduo e os aspectos emocionais de sua história pessoal. (Perkins, Kent, 1986)

Para produzir voz, as pregas vocais são aproximadas através do posicionamento das cartilagens aritenóides e o fluxo de ar que sai dos pulmões provoca a sua vibração, produzindo um som ou frequência fundamental (figura 2.12).

Figura 2.12: Imagem das pregas vocais em posição fonatória. Fechamento glotal máximo.
<http://www.texasvoicecenter.com/diseases.html>

Pelas características particulares da fonte produtora do som, pode-se observar uma certa variabilidade de ciclo a ciclo vibratório. Em outras palavras os ciclos vibratórios podem ser considerados similares, mas não idênticos o que faz com que o sinal de voz humano seja considerado um sinal acústico quase-periódico (Tosi,1979b). Essas variações próprias do som fundamental podem ocorrer como perturbações na frequência, denominada *jitter*, ou perturbações na amplitude, *shimmer*. (Baken, Orlikoff, 2000).

O som fundamental gerado pelos ciclos vibratórios da mucosa das pregas vocais tem uma intensidade muito baixa, e frequência fundamental e seus harmônicos sem amplificação alguma. O espectro do pulso glotal apresenta uma intensidade que decresce 12 dB por oitava, como demonstra a figura 2.13A.

A voz humana como resultado final é muito mais complexo que esse sinal gerado nas pregas vocais. O pulso glotal é amplificado e modificado ao longo do trato vocal. O trato vocal atua como um verdadeiro filtro (equalizador), realçando ou atenuando as componentes de frequências contidas no pulso glotal. O som básico da vibração da mucosa das pregas vocais será modificado pelo fenômeno de ressonância provocado pelo trato vocal que consiste na amplificação de certas faixas de frequência e no amortecimento de outras dependendo das características de cada ressonador ou frequências naturais de ressonância (figura 2.13B). (Titze, 2000).

Pelas características do sistema de ressonância do trato vocal existem inúmeras possibilidades na estrutura tridimensional, revestimento das paredes que geram diversas curvas de resposta traduzida como infinitas qualidades vocais, ou seja, cada voz é única e individual. (Figura 2.13C).

Fig. 2.13A: Desenho esquemático do espectro de frequências do som fundamental sem amplificação. Fig 2.13B: Curvas de ressonância do trato vocal. Fig 2.13C Espectro resultante da frequência fundamental mais as ressonâncias do trato vocal.

O trato vocal também é responsável pela articulação dos sons da fala; palato, língua, dentes e lábios modificam o som laríngeo formando as vogais e as consoantes, elementos básicos da fala. Finalmente esses elementos ao serem processados e coordenados pelo sistema nervoso se transformam em linguagem. (Hammarberg, Gauffin, 1995)

A qualidade de uma voz não depende simplesmente do desempenho e coordenação de diferentes subsistemas, depende também da atividade de todos os músculos envolvidos na produção da voz, e da integridade de todos os tecidos do aparelho fonador e do sistema nervoso central e periférico. (Behlau, 2001).

Resumindo, os espectros de som e as curvas de ressonância de uma voz dependem das características anatômicas e funcionais das estruturas, e da conjunção das dinâmicas dos diferentes sistemas envolvidos no processo de produção de voz. Cabe mencionar que determinadas opções de amplificação e ressonância dependem das escolhas, de base psico-emocional, do falante que se evidenciam no padrão articulatorio e de ressonância. (Rosen; Sataloff, 1997).

2.1.3 Avaliação acústica e perceptiva auditiva dos sinais de voz humana:

A compreensão deste complexo mecanismo é fundamental para os profissionais da voz que dependem da sua performance vocal, assim como para diversas áreas da ciência. Pesquisas na área da saúde, como é o restabelecimento da voz em pacientes com câncer de laringe, na área das telecomunicações como a transmissão e recepção de informações, a indústria e o desenvolvimento tecnológico tem como foco central à voz e seu complexo sistema de produção.

Atualmente, a avaliação tanto perceptivo-auditiva quanto acústica da voz humana é uma prática comum para os profissionais fonoaudiólogos e

otorrinolaringologistas, e consiste em uma série de procedimentos e técnicas que tem como finalidade explorar o comportamento e a dinâmica do sistema vocal de um sujeito.

2.1.3.1 Avaliação perceptiva auditiva da voz.

A avaliação perceptiva auditiva se baseia no julgamento subjetivo que um ouvinte faz à respeito das características de uma voz. Na clínica pode-se considerar como uma estratégia a comparação de uma voz com um sistema de referência interno do avaliador, que depende de diversos fatores e varia de sujeito a sujeito. Este tipo de análise faz parte da prática dos profissionais da voz e é objeto de estudo e aprimoramento, já que por sua característica de subjetividade o grau de confiabilidade tem sido questionado por diversos pesquisadores. (Carding; Carlson; Epstein; Mathieson; Shewell. 2000.p 133-138), (Gerratt; Kreiman. 2001 p. 2560–2566).

Várias propostas de avaliação vocal perceptiva têm sido apresentadas na literatura mundial. As escalas de avaliação vocal perceptiva mais difundidas são: “Modelo Fonético de Descrição da Qualidade Vocal – Fonte e Filtro” (Laver, 1980), para avaliação global da voz. E a escala “GRBAS”, criada pelo Comitê para Testes de Função Fonatória da Sociedade Japonesa de Logopedia e Foniatria (SJLF) em 1969, para avaliação da voz em nível glótico. Esta última foi recentemente adaptada como Escala de Avaliação Perceptiva da Fonte Glótica: RASAT. (Pinho; Pontes. 2002).

Na literatura encontramos que Isshiki (1980) define a rouquidão como uma irregularidade vibratória da mucosa das pregas vocais durante a fonação, causada por uma fenda glótica maior ou igual a $0,5 \text{ mm}^2$. Em patologias como nódulos vocais, hiperemias e edemas pode-se observar uma rouquidão característica. Segundo diversos autores a rouquidão provoca ruídos em baixa frequência (Hirano; Yoshida; Tanaka; Hibi. 1990 p.679-83), (Hammarberg; Gauffin. 1995 p. 283-303), (Hammarberg, 2000 p. 93-102), que mascaram o traçado espectrográfico dos harmônicos inferiores. Na avaliação perceptiva

auditiva a voz apresenta-se irregular, produzindo ruído similar ao som de uma bandeira tremulando ao vento forte.

O termo aspereza é empregado para descrever uma característica vocal associada à rigidez de mucosa, provoca por certa irregularidade vibratória dependendo da presença ou não de fenda glótica e de associações com outras alterações laringeas. Se a característica de rigidez predomina, a qualidade vocal parece uma voz seca e sem projeção.

Segundo alguns autores (Hirano; Yoshida; Tanaka; Hibi. 1990 p.679-83), (Hammarberg; Gauffin. 1995 p. 283-303), (Hammarberg, 2000 p. 93-102) na análise acústica a presença de ruídos em alta frequência é característica. Em termos funcionais, isso se deve ao grande desperdício de ar durante a emissão vocal necessário para desencadear vibração da mucosa enrijecida, favorecendo o aparecimento de harmônicos de baixa frequência. O exemplo clássico de voz *áspera* por rigidez de mucosa é encontrado nos casos de sulco vocal. Cistos, pontes mucosas e bolsas geralmente apresentam voz *áspera* associada à rouquidão mais intensa, devido à presença de edema que freqüentemente os acompanha.

Isshiki (1980 p. 119-54), alerta para o fato que, devido à complexa estrutura de funcionamento do sistema produtor da fala, patologias com características soprosas podem ser confundidas com as percentualmente definidas como roucas.

O adequado estado das pregas vocais, sua densidade muscular, o estado da mucosa que a reveste permitem que ambas sejam aduzidas otimamente, possibilitando a criação de uma condição ideal de pressão e velocidade volumétrica do fluxo de ar, gerando assim o sinal quase periódico que excitará a estrutura supraglotal.

Um funcionamento inadequado de qualquer músculo, principalmente intrínseco, acarretará uma vocalização soprosa (Shrivastav, 2005). Ao ocorrer uma superadução prolongada das pregas vocais, estas ficam passíveis de originarem uma fenda na região posterior, também causando escape de ar.

Qualquer interferência nas bordas de aproximação das pregas vocais resultará em disfonias e ruído turbulento, ou seja, aparecimento de soprosidade na avaliação perceptual.(Eadie; Baylor, 2005)

2.1.3.2 Avaliação acústica dos sinais de voz

Desde o primeiro espectrógrafo de som desenvolvido na Segunda Guerra Mundial por Potter, Kopp e Green (1947 p.8-15), até nossos dias, cientistas e estudiosos de diferentes campos têm criado e desenvolvido diversos métodos para a pesquisa e diagnóstico laríngeo como laringoscopia, eletroestroboscopia, glotografia, análise acústica entre outras. Com o desenvolvimento tecnológico e a introdução do processamento digital de sinais das últimas décadas foi possível obter parâmetros acústicos que oferecem medidas objetivas do sinal de voz. Portanto a análise acústica se apresenta como um método não invasivo capaz de prover informações quantitativas sobre o aparato vocal.

A avaliação acústica da voz oferece algumas informações relativas ao bloco patológico que originou tal modificação acústica, embora possam ocorrer situações em que se encontre um determinado padrão acústico e nenhuma patologia é verificada em exames mais precisos (videolaringoscopia e videoestroboscopia). A implementação da análise acústica no processamento de sinais de voz permitiu definir os parâmetros e medidas acústicas usadas na atualidade na clínica de voz.

Para uso clínico, os parâmetros considerados como mais robustos são frequência fundamental e seus índices de perturbação, as medidas de ruído, e a espectrografia acústica (Belhau, 2001 p.26). Embora este tipo de análise seja considerado objetivo e confiável, devem porém, ser acompanhados do correspondente análise perceptivo auditivo. (Kent; Weismer; Kent; Vorperian; Duffy, 1999 p.146-86).

2.1.3.2.1 Parâmetros acústicos no domínio do tempo

A frequência fundamental ou também chamada frequência glótica é o número de ciclos vibratórios completos da mucosa das pregas vocais por segundo, sendo que quanto mais ciclos por segundo, mais alta será a frequência e mais agudo o sinal de voz, e de forma inversa quanto menos ciclos por segundo mais grave será o sinal acústico vocal. A frequência fundamental de um indivíduo e sua variação estão relacionadas com o comprimento natural das pregas vocais, o alongamento, a massa vibrante e a tensão envolvida em integração com a pressão subglótica. Pode-se observar que a frequência fundamental está relacionada fortemente com o gênero e a idade, porém considerada um dos parâmetros acústicos mais robustos.

A vibração das pregas vocais produz um pulso de ar glotal quase periódico. Esta quase periodicidade, embora normal no funcionamento da laringe, está limitada a certos valores. A utilização de medidas de perturbação do período do sinal acústico pode discriminar casos patológicos dos normais. Baseado nesse fato, várias medidas na perturbação de frequência e amplitude são propostas na literatura. A primeira foi o Fator de Perturbação de Frequência (PFF), (Lieberman, 1961 p. 597-603). Sendo definido como a frequência relativa das perturbações de período maiores que 0,5 ms ocorrendo em uma vogal sustentada. É definida, portanto, como a diferença de tempo entre as durações de períodos sucessivos no sinal de voz.

Utilizando a frequência para substituir o período e variando a janela de dados significativos, pesquisadores propuseram diversas medidas de perturbação. Dessa forma, tem-se na literatura o Quociente de Perturbação de Frequência (FPQ), o Quociente de Perturbação de Período de “*pitch*” (PPQ), Razão de *Jitter* (JR) que fornece a relação entre a média das perturbações pelo período médio, proposto para compensar a redução do *Jitter* em frequências fundamentais relativamente altas. Além da Razão de *Jitter*, existe também o Fator de *Jitter*, análogo ao JR, porém definido no domínio da frequência. A nomenclatura atual engloba todos esses fatores de perturbação de período (frequência) numa única denominação: *Jitter*. (figura 2.14)

Figura 2.14: Exemplo de onda complexa de voz de uma vogal sustentada /e/. As linhas vermelhas assinalam os períodos ciclo a ciclo para cálculo do *jitter*.

Por outro lado, quando a perturbação é calculada sobre a amplitude do sinal, a maioria dos parâmetros descritos acima tem seu correspondente e a denominação é *Shimmer* (figura 2.15). Uma outra característica do sinal a ser considerada além das perturbações de período e amplitude é a forma de onda do sinal. A relação sinal ruído e o espectro do sinal fornecem informações para a discriminação de vozes patológicas e normais.

Figura 2.15: Exemplo de onda complexa de voz de uma vogal sustentada /e/. As linhas vermelhas assinalam as sucessivas amplitudes ciclo a ciclo para cálculo do *shimmer*.

2.1.3.2.2 Parâmetros acústicos no domínio da frequência

A análise no domínio da frequência possibilita um outro conjunto de parâmetros acústicos da voz. O espectrógrafo é um instrumento utilizado na análise de frequência da voz, o qual verifica a distribuição espectral de energia

de um determinado trecho da voz. Como resultado dessa análise tem-se um gráfico tempo x frequência, sendo a densidade de energia indicada pela intensidade. Além disso, a frequência fundamental e os formantes das vogais sustentadas são visualizados no espectrograma. (figura 2.16)

Figura 2.16: Exemplo de um espectrograma de uma vogal sustentada. O tempo está no eixo horizontal, a frequência no eixo vertical e a intensidade do sinal no grau de contraste do traçado.

Diversos estudos comprovam que existem diferenças significativas entre espectros de vozes normais e patológicas. As componentes espectrais do ruído em vozes patológicas são originadas pelo fluxo de ar turbulento gerado pelo fechamento incompleto da glote ou por vibração assíncrona das pregas vocais. (Hammarberg; Gauffin, 1995).

Essas componentes são distribuídas ao longo do espectro de forma variada e a extensão da distribuição depende da gravidade da patologia. Além disso, a presença do ruído contribui para a rouquidão, que é o primeiro sintoma perceptivo para inúmeras patologias.

Figura 2.17: Exemplo de um espectro de frequência de um sinal de voz, da vogal sustentada /a/ de um sujeito de gênero masculino sem queixa vocal.

Proposto por Shoji et al. (1992 p.337-345), a Relação de Potências em Alta Frequência (HFPR) calcula a média do logaritmo da densidade espectral de potência em intervalos de frequência bem definidos.

Alguns outros parâmetros definidos no domínio da frequência são: Relação Harmônico /Ruído (HNR) ; Intensidade Relativa de Harmônicas (Hr) determinado somando-se todas as intensidades das harmônicas (da segunda harmônica em diante) normalizadas pela intensidade total do espectro de voz (fundamental, harmônicas e ruído), sendo que o limite estabelecido para pessoas normais é de 67,5%; Harmônicas Relativas (Ra) que é a relação entre as componentes harmônicas e não harmônicas obtidas em um segmento de três períodos; Relação de Soprosidade (Br) que estabelece a razão entre a energia em altas frequências e a energia total do sinal; Nível Espectral de Ruído (SNL) que analisa as intensidades espectrais em intervalos de frequência e por último a energia de ruído normalizada (NNE) que considera o sinal composto de componentes periódicas mais ruído aditivo.

Neste trabalho estamos propondo uma técnica alternativa de análise quantitativa de voz, visando caracterizar efeitos da voz não contemplados pelas técnicas acima descritas.

2.2 CAOS DETERMINISTICO E ANÁLISE NÃO LINEAR.

2.2.1 Introdução:

No começo da década de 1960, o meteorologista Edward Lorenz, quando tentava resolver um problema de predição meteorológica usando um sistema de equações simples enunciou um fenômeno característico da teoria do Caos, a “sensibilidade às condições iniciais”. Dessa forma ele vislumbrou as bases dos primeiros atratores do caos, que posteriormente conformaram a Teoria do Caos. (Lorentz, 1963).

Os cientistas consideram a Teoria do Caos como uma das grandes descobertas do século vinte. Até antes dela, a visão do mundo e dos fenômenos físicos, era uma dicotomia simplista, fenômenos determinísticos vs fenômenos randômicos, onde os sistemas com padrão regular, previsível e estável eram considerados determinísticos, enquanto os sistemas que apresentam padrões irregulares, erráticos e instáveis eram considerados aleatórios.

Quando se observam comportamentos irregulares em sistemas determinísticos, estes são considerados geralmente como ruído. Estes sistemas determinísticos que apresentam um comportamento irregular na verdade são melhor interpretados sob o enfoque moderno do Caos Determinístico. Sob este enfoque pode-se observar que sistemas governados por leis matemáticas simples podem ter comportamentos altamente imprevisíveis. O resultado desses sistemas é determinado pelas propriedades internas e as condições iniciais.

Os métodos e técnicas de dinâmica não linear, como mapa de Pointcaré, dimensão fractal, dimensão de correlação, entropia de Kolmogorov e Expoente de Lyapunov, são empregados para à análise desse tipo de sistemas (Kantz; Schreiber, 1997), (Ott ; Sauer; Yorke, 1994).

Os métodos de dinâmica não linear e seus conceitos são amplamente conhecidos e aplicados na literatura de estatísticas e resultam triviais na comunidade de processamento de sinais, mas são relativamente novos e desconhecidos para a comunidade da área da saúde. Por esse motivo propomos uma breve revisão de alguns conceitos diretamente relacionados com o Caos Determinístico e análise não linear.

2.2.2 Sistemas não lineares e Dinâmica não linear

Um sistema não linear, por exemplo, um pêndulo pode ser descrito usando a posição e a velocidade angulares, sendo esses dados empregados para construir o *espaço de estado* ou *espaço de fase*, do comportamento do sistema. (figura 2.18)

Figura 2.18: Desenho esquemático da reconstrução do espaço de estado ou espaço de fase, do comportamento de um pendulo, empregando os vetores de posição e velocidade angular.

Uma serie temporal é uma coleção de observações feitas seqüencialmente ao longo do tempo. Através das séries temporais pode-se obter uma observação da dinâmica subjacente de um sistema, e com essa observação escalar realizar a reconstrução do espaço de estado do

comportamento do sistema. Packerd et al. (1980 p. 712) demonstrou como isto pode ser realizado numericamente, e Takens (1981 p. 366-381), subseqüentemente formalizou a prova de como isto pode ser alcançado. Outros autores posteriormente estenderam o trabalho consideravelmente (Sauer; Yorke; Casdagli, 1991 p. 579–616).

Para aplicar a dinâmica não linear como ferramenta de análise é necessário que o sistema possua um número relativamente baixo de variáveis ; $X_i(t)$ ($i = 1, 2, \dots, m$) para que as séries temporais possam ser projetadas no espaço de fase, que é medida nas m coordenadas. Tendo como referência parâmetros fixos (condições externas são constantes), o $X(t)$ vetor m -dimensional é quem traça a trajetória e estabelece um atrator depois de alguns comportamentos transitórios, portanto o atrator é a representação dinâmica de um sistema no espaço de fase.

Existem três tipos de atratores: os pontos atratores (figura 2.19), que são pontos fixos a partir dos quais o sistema evolui; os atratores periódicos (figura 2.20) que são círculos repetitivos de estados; e os *atratores estranhos* que são atratores não periódicos, ou em outras palavras são as representações dinâmicas características dos sistemas caóticos, (Ruelle;Takens,1971 p.167), (Ruelle;Takens, 1971 p.343), e apresentam auto-similaridade de escala (caráter fractal) e dimensão fractal associada. O primeiro atrator estanho foi descrito por Edward Lorentz (1963), como um sistema dinâmico determinístico não linear de três dimensões derivado de equações simples e foi chamado de *atrator de Lorentz*. (figura 2.21).

Figura 2.19: Desenho esquemático de um ponto atrator.

Figura 2.20: Desenho esquemático de um atrator periódico (circular).

Figura 2.21: Desenho esquemático de um atrator estranho ou atrator de Lorenz.
<http://www.nationmaster.com/encyclopedia/Chaos-theory>

Para aplicar a dinâmica do caos num sistema também deve considerar-se que o sistema seja determinístico e não linear, em outras palavras, o sistema deve ser governado interiormente por leis matemáticas simples, característica determinística; e a entrada e a saída do sistema não seguem a proporcionalidade em nenhuma forma, característica não linear. Outro critério para ser considerado é a mudança qualitativa repentina na saída do sistema, ou propriedade de bifurcação. Por último a representação do sistema no espaço de fase deve apresentar propriedade fractal e auto-similaridade. (Kantz; Schreiber, 1997).

O espaço de fase é definido por variáveis dinâmicas múltiplas compostas de posições e velocidades. A trajetória de um atrator no espaço de fase mostra qualitativamente a dinâmica do sistema. A técnica de reconstrução do espaço de fase resulta apropriada quando as variáveis dinâmicas múltiplas não podem ser determinadas. Por exemplo, os ciclos vibratórios das pregas vocais representam um sistema dinâmico, e podem ser descritos como uma trajetória no espaço de fase com evolução no tempo. (Goldberger, 1996).

2.2.2.1 Propriedade de bifurcação

Num sistema dinâmico chama-se de bifurcação à duplicação de um período, uma mudança de um ponto N do atrator à um ponto $2N$ do atrator, que acompanha o início de um comportamento caótico. Representa a manifestação repentina de uma solução qualitativamente diferente para um sistema não linear quando algum parâmetro é variado. Existem quatro variedades de bifurcações: “bifurcação flip”, bifurcação fold”, “bifurcação Pinchfork” e “bifurcação transcritica”. (Gallez; Babloyantz, 1991 p.381-91). Um diagrama de bifurcação é a representação visual da sucessão de duplicação de períodos produzidos quando aumenta r .

A figura 2.22 mostra um diagrama de bifurcação, onde r corre longitudinalmente no eixo x . Para cada valor de r o sistema primeiro se estabelece e, a continuação são marcados valores sucessivos de x para centos de interações.

Figura 2.22: Desenho esquemático de um diagrama de bifurcação.
www.belisarius.com/.../beckerman/non_linear.htm

2.2.2.2 Propriedade Fractal

O termo "fractal" (Mandelbrot, 1975), foi criado em 1975 pelo pesquisador Benoît Mandelbrot, este conceito está geralmente relacionado a formas geométricas com estruturas auto-similares, também é um método para caracterizar semelhanças de objetos complexos e irregulares. Alguns trabalhos demonstraram que sinais fisiológicos apresentam dimensão fractal relacionada à sua auto similaridade, e permite quantificar a irregularidade. (Courtemanche; Winfree, 1991 p.431-44), (Gallez; Babloyantz, 1991 p.381-91).

Uma das características é a auto-semelhança, isto é, têm sempre o mesmo aspecto visual independentemente da escala a que se observa, porque o todo é exatamente igual à ampliação de uma das partes. Esta propriedade resulta do fato de serem construídos pela iteração 'ad infinitum' da mesma regra de construção. (Mandelbrot, 1983).

2.2.2.3. Reconstrução do espaço de fase

Para a reconstrução do espaço de fase de um sistema pode-se usar a técnica de tempo de atraso (Packard; Crutchfield; Farmer; Shaw, 1980 p.712). Por suas características dinâmicas, um sinal de voz humana, pode ser representado como uma série temporal simples de comprimento N , e registrado como $x(t_1), x(t_2), x(t_3), \dots$, onde $x(t_i) \in \mathbf{R}$, $t_i = t_0 + i\tau$ ($i=1,2,\dots,N$), e τ o período de amostragem. Com o vetor de tempo de atraso pode-se reconstruir o espaço de fase como:

$$X(t) = \{x(t), x(t - \tau), \dots, x(t - (m-1)\tau)\}$$

onde m é a dimensão de imersão e τ é o tempo de atraso. Quando $m > 2D+1$, (onde D é a dimensão de Hausdorff), a reconstrução do espaço de fase preserva as propriedades topológicas e a invariância geométrica do atrator original (Takens, 1981 p. 366–381). Os atratores obtidos desta forma se denominam *atratores reconstruídos*. Portanto a dinâmica de um sistema pode ser analisada por meio do método de reconstrução do espaço de fase.

As series temporais se caracterizam por ter comprimento e precisão finitos, portanto, a seleção do τ é importante na reconstrução do espaço de fase. Se τ for muito baixo, as amostras atrasadas estarão fortemente correlacionadas, e as trajetórias dos atratores estendem-se ao redor de uma linha diagonal no espaço de fase reconstruído. No caso contrario, se τ é muito grande, a trajetória dos atratores reconstruídos no espaço de fase caracterizam-se por uma auto-interseção.

Para estimar o tempo de atraso um dos métodos propostos na literatura é o Método de Informação Mutua (Fraser; Swinney, 1986 p.1134-1140), o qual, geralmente depende de duas variáveis. Fraser e Swinney (1986), propõem como critério efetivo para a seleção de um tempo de atraso τ , o primeiro valor mínimo da curva de informação mutua por tempo de atraso τ , e afirmam que esta medida garante que as variáveis defasadas sejam o mais independentes.

2.2.2.4 Dimensão de correlação

O método de dimensão de correlação descreve a complexidade dos comportamentos dinâmicos. Proposto por Grassberger e Procaccia (1983 p. 189-208) é uma medida geométrica de uma trajetória no espaço de fase que descreve o quanto dois pontos de uma trajetória estão correlacionados. Se um sistema for muito complexo terá uma maior dimensão de correlação, o que requer mais variáveis de estado para descrever os comportamentos dinâmicos.

No caso dos atratores reconstruídos o número de variáveis não é conhecido. Assim para reconstruir os atratores é necessário arbitrar-se a dimensão do espaço de fase, conhecida também como dimensão de imersão (*embedding dimension*, E_D).

A dimensão de correlação pode se classificada em: dimensão zero de um ponto fixo (estado estático); limite circular de dimensão um (oscilações periódicas); “torus” quase periódico de duas dimensões (superposição de duas ou mais oscilações com frequências dependentes) e trajetória caótica de dimensão fractal (oscilações aperiódicas). Com este método pode-se distinguir caos do ruído branco, a dimensão de correlação (D_2) do ruído branco não converge quando é aumentada a dimensão de imersão m ; ao contrario a dimensão de correlação (D_2) de um sistema caótico converge a um valor finito quando a dimensão de imersão m é aumentada.

Segundo (Grassberger; Procaccia, 1983), construindo a função $C(r)$, que é a probabilidade de que dois pontos qualquer de uma órbita estejam mais

próximos um do outro que r . Isto é realizado calculando a separação entre cada par de N pontos. Numa série temporal de extensão N a dimensão de correlação pode ser calculada por :

$$D_2 = \lim_{r \rightarrow 0} \lim_{N \rightarrow \infty} \frac{\log C(N, r)}{\log r}$$

onde C é a constante da largura escolhida e r é a separação euclidiana entre dois pontos. ($r = [(x_2 - x_1)^2 + (y_2 - y_1)^2]^{1/2}$).

2.2.2.5 Expoente de Lyapunov

O método de estimação do expoente de Lyapunov foi desenvolvido para definir um número característico para um sistema dinâmico com tempo de atraso. Estes números devem ser calculados na convergência ou divergência exponencial de trajetórias que começam próximas umas de outras (figura 2.23). O número do expoente de Lyapunov caracteriza o comportamento de um sistema dinâmico e é equivalente à dimensão do sistema. (Rosenstein, M.T.; Collins, J.J.; De Luca, C. J. 1993), (Stefanski; Dabrowski; Kapitaniak, 2005 p. 1651-1659).

Figura 2.23: Esquema das trajetórias de duas órbitas de um sistema dinâmico geradas por os pontos X_0 e $X_0 + \Delta x_0$ usando uma equação ou sistema de equações.

Considerando dois pontos no espaço, X_0 e $X_0 + \Delta x_0$, onde cada um deles gera uma órbita no espaço usando uma equação ou um sistema de equações. Essas órbitas podem ser imaginadas como funções paramétricas de uma variável, que pode ser o tempo. Se uma de essas órbitas é usada como órbita de referencia, então o afastamento entre duas órbitas será também uma função do tempo.

Como a sensível dependência pode aparecer somente em algumas porções do sistema, este afastamento é também uma função da localização do valor inicial e tem a forma de $\Delta x(X_0, t)$. Num sistema com pontos atratores fixos ou pontos atratores periódicos, $\Delta x(X_0, t)$ diminui assintoticamente com o tempo (Stefanski; Dabrowski; Kapitaniak, 2005). Se um sistema é instável as órbitas divergem exponencialmente por um período de tempo, mas finalmente se estabilizam.

Portanto o cálculo do expoente de Lyapunov é adequado para estudar a taxa media exponencial de divergência de duas órbitas próximas inicialmente usando a fórmula:

$$\lambda = \lim_{t \rightarrow \infty, |\Delta x_0| \rightarrow 0} \frac{1}{t} \frac{|\Delta x(x_0, t)|}{|\Delta x_0|}$$

O expoente de Lyapunov, “ λ ”, serve para distinguir os diferentes tipos de órbitas tanto em sistemas discretos quanto em sistemas contínuos. Um sistema pode ter mais de um expoente de Lyapunov, sendo $\lambda_{\max}$ usado para detecção da presença de caos. (Wolf; Swift; Swinney; Vastano, 1985 p. 285–317).

Se $\lambda_{\max} < 0$, a trajetória da órbita alterada é atraída à um ponto fixo (figuras 2.24A e 2.24B). Um expoente de Lyapunov negativo é característico de sistemas dissipados ou não-conservativos. Estes sistemas apresentam um comportamento estável assintótico, quanto mais negativo o expoente, maior é a estabilidade. Pontos fixos superestáveis ou pontos periódicos superestáveis apresentam um $\lambda_{\max} = -\infty$. (Hilborn, R. 1994).

Se $\lambda_{\max} = 0$ a órbita é um ponto fixo neutro, indica que o sistema é periódico ou quase periódico. Um sistema físico com $\lambda_{\max} = 0$, é considerado conservativo.

Figura 2.24A: Exemplo de órbitas de um sistema dissipado com um ponto fixo atrator.
Figura 2.24B: Exemplo de órbitas de um sistema dissipado com uma órbita atratora.

Por exemplo, pelo fato da frequência ser independente da amplitude, o espaço de fase de dois osciladores harmônicos simples com diferentes amplitudes resultaria em dois círculos concêntricos, onde as órbitas permanecem com uma separação constante. (figura 2.25). (Bryant; Brown; Abarbanel, 1990 p. 1523-26).

Figura 2.25: Exemplo de duas órbitas de um sistema conservativo com um ponto fixo neutro e órbitas fixas neutrais.

No caso de $\lambda_{\max} > 0$, a órbita é instável e caótica. Pontos vizinhos, mesmo muito próximos, divergem de forma arbitrária, e são considerados instáveis. Num sistema discreto, as órbitas se apresentam como um conglomerado de múltiplos pontos espalhados, nos quais às vezes, é possível visualizar um padrão. Nos sistemas contínuos o espaço de fase pode-se apresentar como um conjunto confuso de linhas. (figura 2.26).

Figura 2.26: Exemplo de órbitas num sistema discreto com $\lambda_{\max} >$

2.2.2.6 Entropia de Kolmogorov

Um sistema pode apresentar um grau elevado e finito de precisão no seu estado inicial, embora a capacidade de prever estados posteriores diminua conforme a trajetória começa a divergir. Deste modo, se perde parte da informação, ou resulta necessário mais informação para descrever o sistema com a precisão inicial, ou em outras palavras é um aumento da entropia do sistema. (Grassberger; Procaccia, 1983).

A entropia de Kolmogorov K ou taxa de variação da informação, se define como taxa média da perda de informação de um sistema, na medida que evolui no tempo. (Grassberger; Procaccia, 1983 p.2591-93). A formula proposta pelo algoritmo de Grassberger e Procaccia (1983 p. 198-208) é:

$$K_2 = \frac{1}{\tau} \ln \frac{C_m(r)^2}{C_{m+1}(r)^2}$$

Se o sistema apresenta um comportamento regular (estado estático, oscilação periódica, ou quase periódica), $K = 0$. Para sistemas caóticos com graus finitos de liberdade $0 < K < \infty$. Em sistemas estocásticos, não determinísticos com comportamentos randômicos, $K = \infty$. (Men; Liang; Zhao, 2004 p.74-8).

2.2.3 Caos na produção dos sinais de voz humana.

A aplicação do caos determinístico e a dinâmica não linear como uma ferramenta útil na análise de sistemas dinâmicos vem acontecendo em diversas áreas como, por exemplo, na bolsa de valores, análise cognitiva em psicologia, meteorologia, etc.

Na área da medicina nos últimos anos os estudo tem focalizado na análise das propriedades caóticas em EEG, sistema respiratório, ritmos cardíacos (Richter; Schreiber, 1998), (Sataloff; Hawkshaw, 2001), (Poon; Merrill, 1997 p. 492-95), (Hornero; Alonso; Jimeno; Jimeno; Lopez, 1999 p.84-90) e vibração das pregas vocais (Jiang; Zhang, 2001), (Awrejcewicz, 1990 p.151-156), (Herzel, 1993 p. 399-413), (Berry; Herzel; Titze; Krischer, 1994 p.3595-3604), (Jiang; Zhang, 2002 p.2127-33). Esses sistemas têm uma característica em comum, não podem ser categorizados como aleatórios, já se seguem determinados padrões próprios, mesmo que o resultado seja muitas vezes altamente imprevisível.

Os sinais de voz humana possuem manifestações de dinâmica não linear, alguns exemplos são voz rouca, voz crepitante, ou as quebras de voz. Podemos mencionar como trabalhos pioneiros em demonstrar a não linearidade na vibração das pregas vocais, o modelo matemático desenvolvido por Awrejcewicz (1990 p. 151-56) que desvia a dificuldade de conduzir experimentos fisiológicos numa laringe em fonação, e demonstra a função

laríngea através de um modelo matemático, o realizado por Herzel, Mende e Wermke, (1990 p.418-424) que evidenciou que o choro de recém nascidos é uma fonte rica de características raras e desconhecidas na voz do adulto, que apresenta “quebra nos sub-harmônicos” e turbulência, que envolvem duplicação de período e aperiodicidade respectivamente. E o estudo de Baken (1990 p.185-197), que comparou as vibrações das pregas vocais entre pessoas normais e pessoas com patologia laríngea comprovada.

A partir destes trabalhos, considerados pioneiros na demonstração da presença da dinâmica não linear, foram desenvolvidos outros trabalhos tanto na área de processamento de sinais, quanto nas áreas da fisiologia e medicina, evidenciando que a teoria de dinâmica não linear oferece uma ferramenta para avaliar e quantificar características específicas de sistemas dinâmicos que produzem saídas caóticas, independentemente do campo da ciência que o empregue.

Este método oferece uma potencial precisão no diagnóstico diferencial baseado em medições acústicas e fisiológicas. (Zhang; Jiang, 2004 p.1266–1269).(Hertrich; Lutzenberger; Spieker; Ackermann, 1997 p.652–654). (Giovanni; Ouaknine; Triglia, 1998; p.341–354), (Matassini; Hegger; Kantz; Manfredi, 2000; 413–418), (Jiang; Zhang, 2002;294–295).(Zhang; Jiang; Biazzo; Jorgensen; Berman, In press).

A dinâmica do caos oferece à pesquisa na área da produção vocal um futuro promissor, mas existem ainda muitas dificuldades pela frente. Dados de um sistema fisiológico real são difíceis de analisar. Os sistemas fisiológicos em geral incluem comportamento aleatório, e sofrem às influências de variáveis contaminantes.

Os avanços da tecnologia em pesquisa, os sistemas de informática e os desenvolvimentos científicos alcançados atualmente nos laboratórios de pesquisa ao redor do mundo permitem que a busca de conhecimento e compreensão dos sistemas, através da dinâmica não linear, nos aproxime do ideal de que uma vez que o caos possa ser controlado, será possível

proporcionar sistemas de avaliação e tratamentos alternativos para uma reabilitação mais efetiva para os diferentes distúrbios vocais.

3- MATERIAIS E MÉTODOS

3.1 Materiais:

3.1.1 Base de dados:

Foram utilizados 66 sinais de voz humana de 22 sujeitos na faixa etária de 19 a 39 anos, 12 de gênero feminino e 10 de gênero masculino, das vogais sustentadas /a/, /e/ e /i/. As vozes pertencentes ao banco de vozes da Bioengenharia foram classificadas como vozes normais, sem alterações patológicas, através de exame médico.

Os sinais de voz foram gravados com uma taxa de amostragem de 22.050 Hz, e digitalizados diretamente num microcomputador com sistema operacional *Microsoft Windows XP Professional- Versão 2002. AMD Athlon™ XP 1700+1.46 GHz, 512 MB RAM*, através do procedimento descrito a continuação:

Com o individuo em posição sentada, com tronco reto, costas encostadas no respaldar da cadeira, braços e mãos relaxadas sobre as pernas; o microfone “de cabeça” é ajustado à cabeça do sujeito e colocado cuidadosamente a uma distancia de 5 cm da boca com uma inclinação de 45 graus.

Foi solicitada a emissão de uma vogal sustentada /a/, numa frequência e intensidade confortáveis, estável e tão prolongada quanto possível. O mesmo procedimento se realiza como as vogais /e/ e /i/. Cada uma das vogais é gravada 3 vezes, para escolher o sinal com características mais estáveis para a análise. Os sinais de voz foram armazenados numa pasta pertencente a cada sujeito voluntário e acompanhados por outros dados referentes fornecidos pelo sujeito.

3.2 Métodos:

Os dados selecionados para este trabalho foram avaliados através de três diferentes metodologias de análise: análise acústica tradicional, análise perceptivo-auditiva e análise de dinâmica não linear.

3.2.1 Análise tradicional de parâmetros de perturbação acústica

Para a coleta dos sinais, e análise tradicional de medidas de perturbação acústicas foi utilizado o programa *Análise de Voz 5.0*. (Montagnoli, 1997). As medidas de perturbação acústica selecionadas para análise foram: “*jitter*” ou perturbação da frequência fundamental, “*shimmer*” ou perturbação da amplitude e o *Coeficiente de Excesso (EX)*, correlato estatístico da relação harmônico / ruído.

3.2.2 Avaliação perceptivo auditiva

Para avaliação perceptivo auditiva das vozes foi utilizada a escala *RASAT* “Escala de avaliação perceptiva da fonte glótica” proposta por Pinho e Pontes (Pinho; Pontes, 2001), criada a partir da escala “*GRBAS*”, criada pelo Comitê para Testes de Função Fonatória da Sociedade Japonesa de Logopedia e Foniatria (SJLF, 1969).

A Escala *RASAT* avalia a qualidade da voz, considerando variações em 5 diferentes aspectos da emissão: Rouquidão; Aspreza; Soprosidade; Astenia e Tensão. Emprega uma escala de graduação que varia de 0 a 3 para cada aspecto avaliado.

3.2.2.1 Procedimento

A avaliação foi realizada por uma fonoaudióloga treinada na área de voz, ouvindo 3 vezes a gravação da vogal sustentada (trecho completo), e anotando numa planilha de dupla entrada, a avaliação estabelecida segundo a escala proposta, para cada aspecto contemplado.

Para a avaliação de: Rouquidão; Aspereza; Soprosidade; Astenia e Tensão, se considera: Normal (0) quando nenhuma alteração vocal é percebida pelo ouvinte; Leve (1) para alterações vocais discretas ou, em caso de dúvida, se a alteração está presente ou não; Moderado (2) quando a alteração é evidente e; Intenso (3) para alterações vocais extremas.

Para notação na avaliação das vozes se utilizou a letra maiúscula como indicador da característica avaliada, ou seja, para Rouquidão (R); para Aspereza (A); para Soprosidade (S); para Astenia (S) e Tensão (T); e o grau de alteração como sub-índice do lado direito da letra maiúscula. Por exemplo, um sinal de voz avaliado com escala *RASAT* que apresente rouquidão grau 1, aspereza 0, Soprosidade 0, Astenia 0 e Tensão 1 será notada como: **R₁ A₀ S₀ A₀ T₁**.

As vozes consideradas para formar parte da amostra deste trabalho foram as que apresentaram na avaliação auditiva perceptiva com escala *RASAT* grau 0 e/ ou 1 nos diferentes aspectos avaliados, ou seja, vozes sem alteração ou com alteração leve na qualidade vocal.

3.2.3 Análise de dinâmica não linear

Para análise da dinâmica não linear dos sinais de voz foi empregado o programa “*ANL*” *Análise Não Linear*, desenvolvido a partir do pacote *Tisean*. (Hegger, Kantz, e Schreiber, 1999)

Foram empregados o método de reconstrução de espaço de fase para descrever as características da dinâmica não linear dos sinais de voz, com a

técnica de tempo de atraso (Fraser, A.M.; Swinney, H.L. , 1986), o cálculo da divergência das órbitas, e o Máximo Expoente de Lyapunov ($\lambda_{\max.}$), (Rosenstein; Collins; De Luca, 1993).

A reconstrução do espaço de fase, o cálculo da divergência das órbitas e o Máximo Expoente de Lyapunov foram implementadas no *MATLAB 6.1*.

3.2.3.1 Procedimento:

Para análise com *ANL*, o programa abre o arquivo do sinal de voz completo como mostra a figura 3.1. Nessa figura pode-se selecionar um trecho qualquer do sinal. No exemplo da figura 3.1 a parte superior mostra o sinal completo (em relação à amplitude e o tempo) da vogal sustentada /e/ de um sujeito de gênero feminino; e o trecho escolhido encontra-se delimitado por duas linhas paralelas de cor vermelha; na parte inferior se observa o gráfico da relação frequência- tempo ou o espectrograma desse mesmo sinal.

Para este trabalho foram escolhidos os trechos que apresentaram maior estabilidade na amplitude de onda, cada um deles com 10 ciclos completos para cada uma das vogais sustentadas.

Figura 3.1: Exemplo de um sinal de voz completo da vogal sustentada /e/. Na parte superior sinal em amplitude por tempo. Na parte inferior representação do sinal no espectrograma.

No exemplo da figura 3.2 pode-se observar em detalhe, delimitado por duas linhas paralelas vermelhas, os 10 ciclos seleccionados do sinal completo da figura 3.1 e seu respectivo espectrograma, visando estabelecer uma referência sobre a estabilidade do sinal.

Figura 3.2: Exemplo do gráfico do trecho selecionado do sinal de voz da vogal sustentada /e/. Na parte superior sinal em amplitude por tempo, 10 ciclos completos. Na parte inferior representação do sinal no espectrograma.

Uma versão mais detalhada dos componentes de frequência do sinal de voz são representados num gráfico de espectro de frequências como mostra o exemplo da figura 3.3.

Figura 3.3: Exemplo do gráfico do espectro de frequências do sinal de voz.

A continuação o ANL apresenta um gráfico com curva de informação mutua por tempo de atraso da amostra. Este gráfico indica em que grau parte da série temporal está relacionada com outras partes da série temporal. Ela mede a dependência geral das duas variáveis diferentes e oferece uma melhor estimativa na eleição de tempo de atraso.

Para a seleção do tempo de atraso da amostra se empregou o método apresentado por Fraser e Swinney, que emprega como critério o tempo dado pelo primeiro mínimo local da função de informação mutua. (Fraser; Swinney, 1986).

Como exemplo a figura 3.4 mostra a curva de informação mutua por atraso de amostras do trecho de sinal da figura 3.2.

Para a seleção do tempo de atraso o aplicativo proporciona um indicador que marca um ponto na curva de informação mutua correspondente ao valor do tempo de atraso escolhido. No exemplo da figura 3.4 a intersecção das linhas

transversais vermelhas marcam a escolha de 7 como primeiro valor mínimo da curva de informação mutua por atrasos de amostras.

Figura 3.4: Exemplo do gráfico da curva de informação mutua por atrasos de amostras. Onde 7 é o valor mínimo selecionado sinalado pela intersecção das linhas vermelhas.

Após da seleção do tempo de atraso, exemplificado na figura 3.4, o *ANL*, apresenta um gráfico bidimensional da reconstrução do espaço de fase do trecho de sinal, neste caso os 10 ciclos escolhidos na figura 3.2.

Este gráfico proporciona um padrão visual traçado pelas órbitas dos atratores ou em outras palavras, uma representação da dinâmica do sistema. Na figura 3.5 pode-se observar o gráfico da reconstrução do espaço de fase dos 10 ciclos escolhidos para análise.

Figura 3.5: Exemplo do gráfico da reconstrução do espaço de fase do trecho selecionado do sinal de voz da vogal sustentada /e/.

O ANL oferece outros dois gráficos ou padrões visuais da dinâmica do sistema. Um deles mostra a reconstrução do espaço de fase em função do tempo, exemplificado na figura 3.6. Nesta figura pode-se observar a evolução no tempo do traçado das órbitas dos atratores do trecho de sinal de voz analisado.

Figura 3.6: Exemplo do gráfico da reconstrução do espaço de fase do trecho selecionado do sinal de voz da vogal sustentada /e/ em função do tempo.

Por último, o *ANL* apresenta o padrão visual da reconstrução do espaço de fase num gráfico tridimensional. A figura 3.7 mostra o gráfico em três dimensões da dinâmica do sistema representada no traçado das órbitas dos atratores.

Figura 3.7: Exemplo do gráfico da reconstrução do espaço de fase do trecho selecionado do sinal de voz da vogal sustentada /e/ num plano tridimensional.

Para medir o grau de divergência ou convergência das trajetórias dos atratores no espaço de fase foram calculados o logaritmo da divergência das órbitas, e o Máximo Expoente de Lyapunov ($\lambda_{\max.}$), empregando os algoritmos propostos por Rosenstein, M.T.; Collins, J.J.; De Luca, C. J. (1993) e exemplificados na figura 3.8.

Figura 3.8: (superior) Curva dos valores do logaritmo de divergência das órbitas dos atratores no espaço de fase. (Inferior) Curva do cálculo do Máximo expoente de Lyapunov do sinal de voz.

4. RESULTADOS E DISCUSSÃO

4.1 Introdução:

Os resultados foram obtidos considerando-se três métodos de análise de forma de compará-los. As análises foram realizadas através do “*Análise de Voz 5.0*” para a obtenção dos parâmetros acústicos da voz, através da escala de avaliação da fonte glótica “*RASAT*” para avaliação perceptivo auditiva dos sinais de voz e do *ANL* “Análise não Linear”, visando à reconstrução do espaço de fase do sinal e o grau de convergência e divergência das órbitas dos atratores.

4.2 Análise tradicional de parâmetros de perturbação acústica. “*Análise de Voz 5.0*”.

Os sinais de voz analisados com “*Análise de Voz 5.0*” mostraram como resultado porcentagens de perturbação da frequência, *Jitter*, com valores na faixa de 0.32% até 2.2%. Foi incluído um sinal com 3.15% de *jitter*, a fim de evidenciar o efeito deste parâmetro quando é realizada a análise com métodos não lineares. Os valores achados para a perturbação da amplitude, *shimmer*, foram entre 0.85% e 7.5%. Os valores para o Coeficiente de excesso foram entre 16.5dB e 145.2 dB.

4.3 Avaliação Perceptiva Auditiva. “*RASAT*”.

Os 66 sinais de voz considerados para este trabalho que foram analisados com escala *RASAT*, apresentaram grau 0 ou 1 para os diferentes aspectos

avaliados. Sendo que 39 sinais de voz apresentaram grau 0 para todos os aspectos avaliados ($R_0 A_0 S_0 A_0 T_0$); 17 amostras com sopro de grau 1 e o resto das características avaliadas com grau 0 ($R_0 A_0 \mathbf{S}_1 A_0 T_0$); com rouquidão grau 1 e 0 nas restantes características ($\mathbf{R}_1 A_0 S_0 A_0 T_0$) foram encontrados 5 sinais de voz; 3 vozes com tensão grau 1 e 0 no resto das características ($R_0 A_0 S_0 A_0 \mathbf{T}_1$) e 2 vozes com astenia grau 1 ($R_0 A_0 S_0 \mathbf{A}_1 T_0$). As 27 amostras de sinais de voz que apresentaram grau 1 em alguma de suas características, mantiveram o resto dos aspectos com grau 0.

4.4 Análise Não Linear “ANL”

4.4.1 Validação da ferramenta.

4.4.1.1 Reconstrução do espaço de fase e parâmetros acústicos em sinais periódicos.

Para avaliar o aplicativo de análise não linear *ANL*, foram realizados testes com um sinal periódico que permite que o operador modifique suas características de modo a simular os parâmetros da voz. Assim é possível alterar a frequência fundamental F_0 , *Jitter*, *shimmer*, além das amplitudes dos harmônicos. Através da validação da ferramenta é possível estabelecer como se manifestam as mudanças e os efeitos destes parâmetros no traçado das órbitas dos atratores no espaço de fase, para posterior identificação de anomalias nos sinais de voz.

No exemplo da figura 4.1 se observa um gráfico da reconstrução do espaço de fase de um sinal artificial com uma frequência única e amplitude constante. Note-se o traçado das órbitas dos atratores é único, independentemente do tamanho do trecho selecionado, já que por ser um sinal periódico o traçado das órbitas dos atratores se sobrepõem no espaço. A figura 4.2 mostra a reconstrução do espaço de fase do sinal na evolução do tempo.

Figura 4.1: Exemplo da reconstrução do espaço de fase de um sinal artificial com uma frequência pura, com 0% de *shimmer* e 0% de *jitter*.

Figura 4.2: Exemplo da reconstrução do espaço de fase de um sinal artificial no tempo de uma frequência pura, com 0% de *shimmer* e 0% de *jitter*.

Nesses testes pode-se observar que o traçado das órbitas dos atratores quando as porcentagens de *shimmer* são alteradas tem um comportamento de espaçamento homogêneo no percurso da órbita no espaço de fase.

Na figura 4.3 mostra o comportamento das órbitas dos atratores no espaço de fase para o mesmo sinal quando sua amplitude é variada aleatoriamente em 15%. O trecho escolhido do sinal tem 10 ms de duração, e o atraso estabelecido pela função mutual é de 5 unidades. Na reconstrução do espaço de fase deste sinal pode-se observar que o traçado dos atratores se apresenta espalhado de forma proporcional à amplitude do sinal. Este comportamento é característico da presença de *shimmer* num sinal.

Figura 4.3: Exemplo da reconstrução do espaço de fase de um sinal artificial de frequência pura com 15% de *shimmer*. O círculo vermelho sinala a porção do espaço de fase que será aumentado na figura 4.4..

Na figura 4.4 pode-se observar a ampliação de uma parte do traçado onde fica visível o espaçamento proporcional no traçado dos atratores. Na reconstrução do espaço de fase o sinal apresenta órbitas com trajetórias próximas, mas que não passam exatamente pela mesma região do espaço de fase.

Figura 4.4: Imagem aumentada de uma parte do traçado das órbitas da reconstrução do espaço de fase de um sinal artificial de frequência pura com 15 por cento de *shimmer* (figura 4.3).

Visando estabelecer o comportamento do sinal no espaço de fase, foi variado, também aleatoriamente, o período do sinal. Na reconstrução do espaço de fase deste sinal o traçado dos atratores apresenta um comportamento diferenciado do anterior onde além do espaçamento, existe uma contorção das órbitas sobre si mesmas.

Figura 4.5: Exemplo do espaço de fase reconstruído de um sinal artificial de frequência pura com 3 de *jitter*. O círculo vermelho sinala a porção do espaço de fase que será aumentado na figura 4. 6..

Nas figuras 4. 5 se observa o espaço de fase reconstruído do sinal artificial com alteração randômica de 3 por cento na frequência, e o exemplo da figura 4.6 é uma imagem aumentada da trajetória das órbitas dentro do círculo vermelho.

Nesta figura fica evidente que os atratores convergem em algumas regiões do espaço de fase e divergem e se espalham em outras, característica mencionada anteriormente.

Figura 4.6: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase de um sinal artificial de frequência pura com 3 de *jitter*.

Figura 4.7: Exemplo da reconstrução do espaço de fase em 3 D de um sinal artificial de frequência pura com 3 de *jitter*. Pode-se observar os pontos de convergência das órbitas e as zonas onde os traçados se espalham.

4.4.1.2 Efeitos da complexidade do sinal de voz na reconstrução do espaço de fase.

O espaço de fase também é eficiente no estabelecimento de padrões visuais de um sinal complexo. Nas figuras 4.8; 4.9; 4.10 e 4.11 são mostrados os padrões de sinais periódicos com uma frequência única (figura 4.8), com duas componentes (figura 4.9), com três harmônicos (figura 4.10) e com quatro componentes (figura 4.11). Note-se que a complexidade da figura aumenta conforme à presença de harmônicos de maior ordem.

Figura 4.8: (superior) Exemplo de um sinal periódico com uma frequência fundamental. (inferior) reconstrução do espaço de fase do sinal.

Figura 4.9: (superior) Exemplo de um sinal periódico de uma frequência fundamental e uma harmônica. (inferior) reconstrução do espaço de fase do sinal.

Figura 4.10: (superior) Exemplo de um sinal periódico com uma frequência fundamental e dois harmônicos. (inferior) reconstrução do espaço de fase do sinal.

Figura 4.11: (superior) Exemplo de um sinal periódico com e uma frequência fundamental e três harmônicos. (inferior) reconstrução do espaço de fase do sinal.

4.4.2 Implementação da ferramenta “ANL” para análise de sinais de voz.

4.4.2.1 Estabelecimento de padrões das vogais nos sinais de voz humana:

A produção das vogais se origina com a vibração das pregas vocais que por ser um sistema oscilador baseado nas características aerodinâmicas e mio-elásticas produz o mesmo pulso para as diferentes vogais. A caracterização das diferentes vogais é determinada pelas estruturas de conformação supra-glóticas, amplificando ou atenuando faixas específicas de frequências contidas nos pulsos glotais.

O gesto vocal para a emissão da vogal /a/ introduz amplificação em várias faixas de frequência produzindo um sinal complexo e rico em harmônicas.

A figura 4.12 mostra os padrões da vogal /a/ onde a presença de harmônicos de ordem superior acarretam vários “laços” em suas órbitas produzindo padrões complexos no espaço de fase.

Figura 4.12: Exemplos da reconstrução do espaço de fase de sinais de voz humana da vogal sustentada /a/.

A figura 4.13 apresenta a reconstrução do espaço de fase para a vogal /e/. Nota-se que as órbitas apresentam menor complexidade devido ao reforço do trato vocal estar localizado em poucas faixas de frequências.

Figura 4.13: Exemplos da reconstrução do espaço de fase de sinais de voz humana da vogal sustentada /e/.

Finalmente a figura 4.14 mostra padrões da vogal /I/. Esta vogal tem reforço em uma faixa de frequências prevalente. Esta característica está retratada como um laço único em suas órbitas. Devido à relação alta entre o formante principal e os outros formantes, a conformação de laços fica prejudicada, isto é, o diâmetro dos laços é desprezível face ao diâmetro da órbita principal.

Figura 4.14: Exemplos da reconstrução do espaço de fase de sinais de voz humana da vogal sustentada /I/.

Observe-se que o ruído presente nas vogais provoca dispersão no traçado da órbita, ou seja, quanto mais rugosa a voz mais dispersa a órbita.

4.4.2.2 Reconstrução do espaço de fase, cálculo da convergência das órbitas, Máximo Expoente de Lyapunov e parâmetros acústicos em sinais de voz.

Através da validação e implementação do programa *ANL* como ferramenta de análise foi possível visualizar os efeitos das mudanças dos parâmetros no traçado das órbitas dos atratores no espaço de fase de sinais de voz humana consideradas normais, ou sem alterações na qualidade vocal, segundo avaliação perceptiva auditiva e acústica.

Antes de descrever os resultados obtidos na análise dos sinais de voz com *ANL*, deve-se recordar que na validação da ferramenta com um sinal periódico foi observado que quando as porcentagens de *shimmer* são alteradas os atratores apresentam um comportamento de espaçamento homogêneo no percurso das órbitas no espaço de fase, sendo que quanto maior a porcentagem de *shimmer* num sinal, mais espalhadas serão as trajetórias dos atratores, e de modo inverso, um sinal com baixa porcentagem de *shimmer* apresentará um espaço de fase com trajetórias muito próximas umas de outras.

No caso de alteração na porcentagem de *jitter*, os testes de validação mostraram que os atratores traçam trajetórias, que no decorrer da dinâmica se contorcem sobre si mesmas fornecendo convergências e divergências alternadas na projeção das órbitas.

Quanto maior for a porcentagem de perturbação do período, maior o número de alternâncias e maior a dispersão das órbitas nos lóbulos de divergência. Em sinais com valores baixos de *jitter* as órbitas tendem a passar

próximas umas de outras, atenuando a dispersão nos lóbulos e diminuindo o número de contorções das órbitas.

Num gráfico de reconstrução de espaço de fase em duas dimensões esta característica se apresenta como se as órbitas se contorcessem sobre si mesmas, ou em outras palavras, as trajetórias se apresentam espalhadas heterogeneamente no seu percurso no espaço de fase.

4.4.2.2.1 Reconstrução do espaço de fase e Coeficiente de Excesso (Ex) em sinais de voz.

Diferentemente dos testes realizados com sinais periódicos, para análise das vozes foi considerado o Coeficiente de Excesso (Ex), além da perturbação do período e da amplitude (*jitter* e *shimmer*).

O sistema de produção de voz é considerado um sistema de alta complexidade que envolve a dinâmica de diversos sub-sistemas de órgãos, portanto propenso à presença de ruído. O fluxo de ar turbulento proveniente dos pulmões e as condições da mucosa laríngea são alguns dos fatores considerados como possíveis fontes de ruído na dinâmica deste sistema. Devido às características de alta complexidade do sistema de produção de voz foi calculado o Coeficiente de Excesso (Ex), que basicamente é o correlato estatístico da relação harmônico /ruído (HNR).

Figura 4.15A: reconstrução do espaço de fase do sinal de voz da vogal sustentada /e/ com 3.13 % de *jitter*, 6.96 % de *shimmer* e 136.05 de Ex, com avaliação perceptiva auditiva $R_0 A_0 S_1 A_0 T_0$.

Figura 4.15B: reconstrução do espaço de fase do sinal de voz da vogal sustentada /e/ com 0.4 % de *jitter*, 2.15 % de *shimmer* e 16.5 de Ex, com avaliação perceptiva auditiva $R_0 A_0 S_1 A_0 T_0$.

A figura 4.15 mostra de forma comparativa na reconstrução do espaço de fase os parâmetros acústicos *Jitter*, *shimmer* e Ex (Coeficiente de Excesso) de dois sinais de voz de sujeitos sem queixa vocal e semelhante qualidade de voz segundo avaliação perceptivo auditiva.

Note-se que para uma pequena relação entre o sinal e o ruído (figura 4.15B) as órbitas apresentam um traçado irregular, e as trajetórias dos atratores sobrepostas dão a impressão de dispersão. Este efeito difere da característica do *shimmer*, que embora apresente dispersão, suas órbitas têm traçado mais uniformes.

Para o sinal mostrado na figura 4.15A os parâmetros acústicos tem os seguintes valores: *jitter* de 3.13%; *shimmer* de 6.96 % e Ex de 136.05dB e avaliação perceptiva auditiva $R_0 A_0 S_1 A_0 T_0$, enquanto os valores dos parâmetros para o sinal mostrado na figura 4.15B são 0.4% para *jitter*, 2.15% para *shimmer* e

16.5dB para Ex, e a avaliação perceptiva auditiva apresenta a mesma valoração $R_0 A_0 S_1 A_0 T_0$.

Se considerarmos apenas *jitter* e *shimmer* o espaço de fase da figura 4.15A deveria ser mais dispersa que a figura 4.15 B. A alta relação entre os valores de Ex dos dois sinais se sobrepõe aos valores de *jitter* e *shimmer* realçando o ruído contido no sinal da figura 4.15 B.

Neste caso, a avaliação perceptiva auditiva de ambos sinais é semelhante, embora existam diferenças na dinâmica de cada sistema que ficam em evidencia com a utilização da reconstrução do espaço de fase como ferramenta de análise. Uma importante característica desta ferramenta é o aumento da resolução, isto é, diferenças imperceptíveis para o ouvido humano são realçadas com a técnica não linear.

Desse modo, num espaço de fase reconstruído a relação harmônico/ ruído de um sinal de voz, calculada através do Ex, pode ser observada no traçado de cada órbita; isto é, cada trajetória apresenta regiões oscilantes e irregulares seguidas de outras com traçado liso e uniforme.

O efeito de um valor baixo do Coeficiente de Excesso ($Ex = 67.84\text{dB}$) no comportamento dos atratores está exemplificado em detalhe no gráfico de reconstrução de espaço de fase da figura 4.16, onde as regiões irregulares encontram-se demarcadas em vermelho e amplificadas na figura 4.18.

Figura 4.16: Reconstrução do espaço de fase do sinal de voz da vogal sustentada /a/ com 0.36 % de *jitter*, 2.92 % de *shimmer* e 67.84dB de Ex, com avaliação perceptiva auditiva $R_0 A_0 S_1 A_0 T_0$.

Em contraste a figura 4.17 mostra um padrão visual diferente da dinâmica dos atratores de um sinal de voz que, embora com características semelhantes na qualidade vocal, grau 1 de soproidade, ($R_0 A_0 S_1 A_0 T_0$) e índices similares de *jitter* e *shimmer*, apresenta um valor elevado no Coeficiente de Excesso (Ex= 197.31dB).

Nesta figura encontra-se demarcada uma região do espaço de fase onde se observa que embora as órbitas apresentem trajetórias dispersas, estas encontram-se limitadas a uma região mais restrita do espaço de fase.

Figura 4.17: Reconstrução do espaço de fase do sinal de voz da vogal sustentada /a/ com 0.35 % de *jitter*, 1.96% de *shimmer* e 197.31dB de Ex, com avaliação perceptiva auditiva $R_0 A_0 S_0 A_0 T_0$.

Na avaliação preceptiva auditiva os sinais de voz das figuras 4.16 e 4.17 diferem no aspecto de soproidade, que do ponto de vista funcional está associado ao escape de ar no momento da coaptação glótica; este excesso de ar na fonação pode ter origens anátomo- fisiológicos diferentes, mas para o ouvido humano se traduz como soproidade. O correlato acústico desta característica é o aumento na porcentagem de *shimmer* ou um valor de Coeficiente de Excesso (proporção harmônico/ ruído) baixo.

Observe-se nas figuras 4.18 e 4.19 as imagens aumentadas das áreas selecionadas dos padrões visuais das figuras 4.17 e 4.18 respectivamente. Nelas pode-se apreciar em detalhe e de forma comparativa, a regularidade no traçado de cada órbita e a distribuição das trajetórias num espaço de tempo, além de possibilitar uma melhor resolução na hora de avaliar estas diferenças na dinâmica vocal.

Figura 4.18: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase do sinal de voz da figura 4.16

Figura 4.19: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase do sinal de voz da figura 4.17.

Nestes sinais foram calculados os valores do logaritmo da divergência das órbitas, e o Máximo Expoente de Lyapunov ($\lambda_{\max}$). Estas medidas permitem distinguir os diferentes tipos de órbitas tanto em sistemas discretos quanto em sistemas contínuos, e possibilitam o cálculo da divergência ou convergência exponencial das trajetórias dos atratores no espaço de fase.

A estimativa destes valores foi calculada usando os algoritmos propostos no trabalho de Rosenstein, M.T.; Collins, J.J.; De Luca, C. J. (1993).

Figura 4.20: Gráfico do cálculo do Máximo Expoente de Lyapunov ($\lambda_{\max}$) do sinal de voz da vogal sustentada /a/ da figura 4.16.

Figura 4.21: Gráfico do cálculo do Máximo Expoente de Lyapunov do sinal de voz da vogal sustentada /a/ da figura 4.17.

Nas figuras 4.20 e 4.21 pode-se observar na parte superior a curva do cálculo do logaritmo da divergência das órbitas e na parte inferior a curva do cálculo do máximo Expoente de Lyapunov ($\lambda_{\max}$) para as dimensões da 2 à 7 dos sinais reconstruídos nas figuras 4.16 e 4.17 respectivamente.

Segundo (Wolf; Swift; Swinney; Vastano, 1985), (Rosenstein; Collins; De Luca. 1993) o Máximo Expoente de Lyapunov ($\lambda_{\max}$) pode ser: $\lambda_{\max} < 0$, quando a trajetória é uma órbita periódica estável; $\lambda_{\max} = 0$, quando a órbita é um ponto fixo neutro ou se encontra dentro de um ciclo limite, isto indica que o sistema é periódico ou quase periódico; ou $\lambda_{\max} > 0$, quando pontos vizinhos divergem de forma arbitrária, portanto a órbita é instável e caótica.

4.4.2.2 Reconstrução do espaço de fase e *Shimmer* em sinais de voz.

Nos exemplos das figuras 4.22 e 4.23 pode-se observar o comportamento de sinais de voz analisados com ANL, que apresentam diferentes porcentagens de perturbação da amplitude (*shimmer* = 0.98% na figura 4.22 e *shimmer* = 2.91% na figura 4.23). O sinal da figura 4.23 foi avaliado com grau 1 no aspecto da soproiedade na análise perceptiva auditivo ($R_0 A_0 S_1 A_0 T_0$), entanto o sinal da figura 4.22 não apresentou nenhum grau de desvio na escala ($R_0 A_0 S_0 A_0 T_0$). As medidas acústicas de *jitter* e EX mostraram valores semelhantes, sendo diferentes somente nas porcentagens de perturbação da amplitude. O *shimmer* é a perturbação ou variabilidade da amplitude ciclo a ciclo. (Horii, 1982). Este parâmetro acústico esta geralmente correlacionado com a presença de ruído na emissão vocal, descrito em termos perceptivos como rouquidão ou soproiedade.

Figura 4.22: Exemplo do espaço de fase reconstruído do sinal de voz da vogal sustentada /e/ com 0.4 % de *jitter*, 0.98 % de *shimmer* e 139.13dB de EX, com avaliação perceptiva auditiva $R_0 A_0 S_0 A_0 T_0$.

Figura 4.23: Exemplo do espaço de fase reconstruído do sinal de voz da vogal sustentada /e/ com 0.46 % de *jitter*, 2.91 % de *shimmer* e 129.73dB de E_x , com avaliação perceptiva auditiva $R_0 A_0 S_1 A_0 T_0$.

Esta diferença resulta evidente ao se observar que as órbitas dos atratores, embora com traçados muito próximos, são afastadas de forma homogênea no percurso das trajetórias no espaço de fase mostrado na figura 4.22; em comparação com a configuração mais dispersas das órbitas da figura 4.23.

Para estas vozes, a análise perceptiva auditiva, embora seja uma avaliação subjetiva, detecta uma mínima mudança na qualidade vocal, do mesmo modo que a avaliação acústica acusa um aumento na porcentagem do *shimmer*, e o *ANL* possibilita que uma pequena diferença na sensação acústica e /ou no cálculo de valor médio da perturbação da amplitude (*shimmer*), sejam visíveis no espaço de fase.

Figura 4.24: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase do sinal de voz da figura 4.22.

Figura 4.25: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase do sinal de voz da figura 4.23.

Nas imagens aumentadas das figuras 4.24 e 4.25 observam-se em detalhe as características dos traçados dos atratores das áreas selecionadas das figuras 4.22 e 4.23 respectivamente

De igual modo, foi calculado o logaritmo da divergência das órbitas e o máximo Expoente de Lyapunov ($\lambda_{\max}$), para as dimensões da 2 à 7, para os sinais reconstruídos nas figuras 4.22 e 4.23.

As figuras 4.26 e 4.27 mostram os gráficos das curvas desses valores de divergência para as vozes das figuras 4.22 e 4.23 respectivamente.

Figura 4.26: Gráfico do cálculo do Máximo Expoente de Lyapunov do sinal de voz da vogal sustentada /e/ da figura 4.22.

Figura 4.27: Gráfico do cálculo do Máximo Expoente de Lyapunov do sinal de voz da vogal sustentada /e/ da figura 4.23.

Os exemplos apresentados anteriormente mostraram de forma comparativa, nas figuras 4.16 e 4.17, o comportamento de sinais de voz com diferentes valores da relação harmônico/ ruído, através de cálculo do Coeficiente de Excesso (Ex); e o comportamento na reconstrução do espaço de fase de vozes de diferentes porcentagens de *shimmer*, nas figuras 4.22 e 4.24. Também foram calculadas medidas de divergência das órbitas dos atratores reconstruídos e o Máximo Expoente de Lyapunov para estes sinais.

4.4.2.2.3 Reconstrução do espaço de fase e *Jitter* em sinais de voz.

Seguem a continuação, exemplos da representação dinâmica do comportamento de sinais de voz com diferentes porcentagens de perturbação do período (*jitter*), analisados com ANL. (Figuras 4.28 e 4.29).

Figura 4.28: Exemplo do espaço de fase reconstruído de um sinal de voz da vogal sustentada /i/ com 0.63 % de *jitter*, 2.4 % de *shimmer* e 110.64 de Ex , com avaliação perceptiva auditiva $R_0 A_0 S_0 A_0 T_0$.

A voz analisada na figura 4.28 mostrou na avaliação com escala RASAT que perceptivamente pode ser considerada sem alteração na qualidade vocal ($R_0 A_0 S_0 A_0 T_0$). Já o sinal vocal da figura 4.29 foi avaliada com grau 1 no aspecto de aspereza ($R_0 A_1 S_0 A_0 T_0$).

Os valores médios da análise acústica de ambos sinais foram considerados similares para o Ex e o *shimmer*, apresentando uma diferença significativa

somente no valor médio da perturbação do período (*jitter* = 0.63% sinal da figura 4.28 e *jitter* = 2.11% para o sinal da figura 4.29). Baixas porcentagens de *jitter* em vozes sem alterações patológicas podem ser causadas por uma pequena variação na massa ou na tensão das pregas vocais, ou ainda a presença de muco sobre estas estruturas. (Baken, 1987).

Ao observar os padrões visuais obtidos com *ANL*, a variabilidade na frequência ou período fundamental resulta mais nítida e com melhor definição como pode-se observar nas figuras de reconstrução de espaço de fase.

Figura 4.29: Exemplo do espaço de fase reconstruído de um sinal de voz da vogal sustentada /i/ com 2.11% de *jitter*, 3.38 % de *shimmer* e 103.7 de Ex , com avaliação perceptiva auditiva R_0 A_1 S_0 A_0 T_0 .

Observe-se que a disposição das projeções no espaço de fase dos sinais das figuras 4.28 e 4.29 apresentam regiões com cruzamento de órbitas e outras onde há paralelismo entre as mesmas, caracterizando o correlato visual de *jitter*. Embora a convergência e divergência no percurso das trajetórias seja comum a

ambos os sinais, este comportamento característico da presença de *jitter*, encontra-se mais acentuado em toda a trajetória dos atratores do sinal da figura 4.29.

As figuras 4.30 e 4.31 mostram imagens aumentadas que confirmam as características mencionadas anteriormente dos traçados dos atratores das áreas selecionadas das figuras 4.28 e 4.29 respectivamente.

Figura 4.30: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase do sinal de voz da figura 4.28.

Figura 4.31: Imagem aumentada de uma parte do traçado das órbitas dos atratores na reconstrução do espaço de fase do sinal de voz da figura 4.29.

As curvas do cálculo do logaritmo da divergência das órbitas e do Máximo Expoente de Lyapunov ($\lambda_{\max.}$) dos sinais das figuras 4.28 e 4.29 encontram-se apresentados nas figuras 4.32 e 4.33 respectivamente.

Figura 4.32: Gráfico do cálculo do Máximo Expoente de Lyapunov do sinal de voz da vogal sustentada /i/ da figura 4.28.

Figura 4.33: Gráfico do cálculo do Máximo Expoente de Lyapunov do sinal de voz da vogal sustentada /i/ da figura 4.29.

O cálculo do logaritmo da divergência das órbitas e dos valores do Máximo Expoente de Lyapunov ($\lambda_{\max}$) apresentaram curvas e valores com resultados semelhantes para todos os sinais de voz analisados.

Em todos os casos as curvas do logaritmo da divergência teve o mesmo comportamento de oscilação para os primeiros valores até ficar estável num só valor.

Para todos os sinais de voz analisados o cálculo do Máximo Expoente de Lyapunov ($\lambda_{\max}$) tendeu a zero ($\lambda_{\max} = 0$), resultado característico de um sistema periódico ou quase periódico, como é o caso dos sinais de voz; e característico da dinâmica não linear ou caos determinístico dentro de um ciclo limite.

Nos resultados apresentados neste capítulo pode-se observar que os métodos de avaliação acústica tradicional e /ou avaliação auditiva perceptiva são em alguns casos insuficientes para refletir à dinâmica do sistema de produção de voz. A reconstrução do espaço de fase como ferramenta de análise sinais vocais proporciona padrões visuais que permitem “enxergar” dados que ficam encobertos nos cálculos de outras ferramentas.

5. CONCLUSÃO:

Pode-se concluir depois dos ensaios e os testes realizados com os sinais periódicos e com as vozes humanas, que a técnica de análise não linear de reconstrução do espaço de fase se apresenta como uma ferramenta útil para discriminar a dinâmica dos sinais de vozes normais através da avaliação da heterogeneidade das diferentes regiões das trajetórias dos atratores no espaço de fase.

Os resultados das análise dos sinais de voz evidenciam a eficiência da ferramenta, já que diferenças entre ciclos são imperceptíveis para o ouvido humano e muitas vezes para os métodos de análise linear e podem mascarar alterações fisiológicas nos mecanismos de produção vocal, que ficam evidentes na análise não-linear.

Com a técnica de reconstrução do espaço de fase é possível obter padrões visuais da dinâmica do sistema de produção vocal que independe da frequência fundamental do sinal de voz, já que o espaço de fase em duas dimensões não apresenta modificações com a mudança de frequência. Além disso permite visualizar a amplitude dos formantes nos traçados das órbitas secundárias dos atratores, estabelecendo uma relação entre o formante principal e os outros formantes, pelas características da conformação de laços. Para componentes em frequências com amplitudes similares as órbitas formam laços por cruzamento, mas se a relação da amplitude da frequência fundamental para a amplitude das harmônicas é muito elevada, o traçado das órbitas fica prejudicado, ou seja, o diâmetro dos laços é desprezível face ao diâmetro da órbita principal. Quando esta relação se torna muito elevada a presença dos laços no espaço de fase cedem lugar às flutuações na órbita principal.

Através desta técnica pode-se visualizar os parâmetros acústicos convencionais como jitter, shimmer e Ex (coeficiente de Excesso).

No espaço de fase estes parâmetros estão contidos nas características do traçado dos atratores bem como no comportamento de uma única órbita desses atratores. Em sinais de voz com perturbações da amplitude do período as órbitas dos atratores são deslocadas de maneira quase uniforme apresentando traçados com espaçamento proporcional. Em outras palavras a presença de shimmer num sinal de voz apresenta órbitas com trajetórias próximas, e aproximadamente paralelas.

Para os sinais de voz com perturbações no período, jitter, os padrões visuais da reconstrução do espaço de fase mostram que o comportamento do traçado dos atratores sofre uma contorção das órbitas sobre si mesmas, diferenciando-se do traçado característico com presença de shimmer.

A característica visual do coeficiente de excesso Ex no espaço de fase se constitui num movimento errático no traçado da órbita.

Em suma, a avaliação perceptivo auditiva é subjetiva, a avaliação acústica baseia-se em valores médios dos parâmetros enquanto a análise dinâmica não linear fornece valores instantâneos e uma maior nitidez das características da voz, aumentando desta forma a resolução das avaliações.

Sugestões para trabalhos futuros

Como sugestão para trabalhos futuros podemos citar :

- a) Estabelecimento de protocolos adequados à técnica de reconstrução de espaço de fase.
- b) Estabelecimento de padrões visuais para alterações vocais
- c) Estabelecimento parâmetros de quantificação da dinâmica das órbitas do sinal de voz no espaço de fase.

REFERENCIAS BIBLIOGRAFICAS

AWREJCEWICZ, J. *Bifurcation Portrait of the Human Vocal Cord Oscillations*. Journal of Sound and Vibration, 136 (1) pp.151-156. (1990)

BANBROOK, M; MCLAUGHLIN, S; MANN, I. *Speech Characterization and Synthesis by Nonlinear Methods*. IEEE Transactions On: Speech And Audio Processing, 7. (1999)

BAKEN, R.J. *Clinical Measurements of Speech and Voice*. Boston: College Hill. Pp. 542-547. (1987)

BAKEN, R.J. *Irregularity of vocal period and amplitude: A first approach to the fractal analysis of voice*. J Voice ;4:185–197. (1990)

BAKEN RJ, ORLIKOFF RF. *Clinical Measurement of Speech and Voice*. San Diego, CA: Singular Publishing Group;(2000).

BÉLAIR, J. J.; GLASS, L.U.; HEIDEN, L.U.; MILTON, J. *Dynamical Disease: Mathematical Analysis of Human Illness*. American Institute of Physics, New York. (1995).

BELHAU, M. *A Voz. O livro do especialista*. Vol. I. p 25. Revinter. (2001).

BELHAU, M. *A Voz. O livro do especialista*. Vol. I. p 26. Revinter. (2001).

BERRY, D.A.; HERZEL, H.; TITZE, I.R.; KRISCHER, K. *Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions*. J AcoustSoc Am. 95:3595–3604. (1994).

BRYANT, P.; BROWN, R.; ABARBANEL, H. *Lyapunov exponents from observed time series*. Phys. Rev. Lett. 65 (13), 1523-1526 (1990).

CARDING, P.; CARLSON, E.; EPSTEIN, R.; MATHIESON, L.; SHEWELL, C. *Formal perceptual evaluation of voice quality in the United Kingdom*. Logoped Phoniat Vocol.;25:133–138. (2000).

COMITÉ PARA TESTES DE FUNÇÃO FONATÓRIA DA SOCIEDADE JAPONESA DE LOGOPEDIA E FONIATRIA - Fita de Vídeo ilustrativo da “GRBAS”. (1969)

COURTEMANCHE, M.; WINFREE, A.T. *Re-entrant rotating waves in a Beeler-Reuter based model of two-dimensional cardiac electrical activity*. Int J Bifurcation Chaos 1: 431–44. (1991)

- EADIE, T.L.; BAYLOR, C.R. *The Effect of Perceptual Training on Inexperienced Listeners. Judgments of Dysphonic Voice. Journal of Voice, ARTICLE IN PRESS, (2005).*
- FRASER, A.M.; SWINNEY, H.L. *Independent coordinates for strange attractors from mutual information. Phys Rev Lett. ;33:1134-1140. (1986)*
- GALLEZ, D.; BABLOYANTZ, J. *Predictability of the human EEG: a dynamical approach. Biol Cybern; 64:381–91. (1991)*
- GERRATT, B.R.; KREIMAN, J. *Measuring vocal quality with speech synthesis. J Acoust Soc Am. ;110:2560–2566. (2001)*
- GIOVANNI, A.; OUAKNINE, M.; TRIGLIA, J.M. *Determination of largest Lyapunov exponents of vocal signal: application to unilateral laryngeal paralysis. J Voice.;13:341–354. (1998).*
- GOLDBERGER, A. L. *Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. The Lancet. pp. 1312-14. (1996)*
- GRASSBERGER, P.; PROCACCIA, I. *Measuring the strangeness of strange attractors. Physica D. 9:189–208. (1983)*
- GRASSBERGER, P.; PROCACCIA, I. *Estimation of the Kolmogorov entropy from a chaotic signal. Phys Rev A.; 45: 2591–2593 (1983)*
- GRAY, S.; HIRANO, M.; SATO, K. *Molecular and cellular structure of vocal fold tissue. In TITZE IR: Vocal Folds Physiology. San Diego: Singular. P. 1-35. (1993).*
- HAMMARBERG, B.; GAUFFIN, J. *Perceptual and acoustics characteristics of quality differences in pathological voices as related to physiological aspects. In: Fujimura O, HIRANO, M – Vocal fold physiology. San Diego, Singular Publishing Group Inc., p. 283-303. (1995).*
- HAMMARBERG, B. *Voice research and clinical needs. Folia Phoniat. Loop., 52: 93-102.(2000)*
- HEGGER, R.; KANTZ, H.; SCHREIBER, T. *Practical implementation of nonlinear time series methods: The TISEAN package. Chaos. Volume 9, Issue 2, pp. 413-435. (1999).*
- HERTRICH, I.; LUTZENBERGER, W.; SPIEKER, S.; ACKERMANN, H. *Fractal dimension of sustained vowel productions in neurological dysphonias: an acoustic and electroglottographic analysis. J Acoust Soc Am.;102:652–654. (1997)*
- HERZEL, H. *Bifurcations and chaos in voice signals. Appl. Mech. Rev.;46:399–413. (1993).*

HERZEL, H., BERRY, D., TITZE, I.R., SALEH, M. *Analysis of vocal disorders with methods from nonlinear dynamics*. Journal of Speech and Hearing Research, 37, pp 1001-1007. (1994).

HILBORN, R. *Chaos and Nonlinear Dynamics*. Oxford: Oxford University Press, (1994).

HIRANO, M.; YOSHIDA, T; TANAKA, S.; HIBI, S. *Sulcus vocalis: functional aspects*. Ann. Otol. Rhinol. Laryngol.; 99:679-83. (1990).

HIRANO, M.; KAKITA, Y. *Cover-body theory of vocal fold vibration*. Speech Science. ed. R.G. Daniloff, San Diego, California: College-Hill Press, pp. 1-46. (1985)

HIRANO, M. *Objective evaluation of the human voice: Clinical aspects*. Folia Phoniatica, Vol. 41, , pp. 89-144. (1989)

HIRANO, M. *Structure of the vocal fold in normal and diseased states: Anatomical and physical study*. Proceedings of the Conference on the Assessment of Vocal Pathology. American Speech and Hearing Association Report, Vol. 11, p. 69. (1981)

HIRANO, M.; BLESS, D.M. *Videostroboscopic Examination of the Larynx*. San Diego: Singular Publishing Group Inc., (1993).

HIRANO, M. *Histological Color atlas of the Human Larynx*. San Diego: Singular, (1993).

HIRANO, M. *Laryngeal Histopathology*. In COLTON, R; CASPER, J. Understanding voice problems. A Physiological Perspective Of The Diagnosis And Treatment. 2nd.ed. Baltimor: Williams & Wilkins, (1996).

HORII, Y. *Jitter and Shimmer differences among sustained vowel phonations*. Journal of Speech and Hearing Research; 25:12-14. (1982).

HORNERO, R.; ALONSO, A.; JIMENO, N.; JIMENO, A.; LOPEZ, M. *Nonlinear analysis of time series generated by schizophrenic patients*. IEEE Eng Med Biol Mag. ;3:84-90. (1999).

IMAMURA, R.; TSUJI, D.H.; SENNES, L.U. *Fisiologia da laringe*. In: CAMPOS, C.A.H.; COSTA, H.O. Tratado de Otorrinolaringologia. São Paulo: Rocca; p. 743-50. (2002).

ISSHIKI N. *Recent advances in phonosurgery*. Folia Phoniatr. (Basel), 32:119-54. (1980)

JIANG, J. J.; ZHANG, Y. *Modeling of chaotic vibrations in symmetric vocal folds*. J. Acoust. Soc. Am. 110 (4). (2001).

JIANG, J.J.; ZHANG, Y. *Chaotic vibration induced by turbulent noise in a two-mass model of vocal folds*. J Acoust Soc.Am.;112:2127–2133. (2002).

JIANG, J.J.; ZHANG, Y. *Nonlinear dynamic analysis of speech from pathologic subjects*. Electron Lett. 38:294–295. (2002)

KANTZ, H.; SCHREIBER, T. *Nonlinear Time Series Analysis*. Cambridge, UK: Cambridge University Press; (1997).

KENT, RD; WEISMER, G.; KENT, FJ.; VORPERIAN, HK.; DUFFY, JR. *Acoustics studies of dysarthric speech: methods, progress and potential*. J of Commun Disord 32:146-86. (1999)

KUMAR, A.; MULLICK, K. *Nonlinear dynamical aspects of speech*. Journal of the Acoustical Society of America, 100 pp. 737 – 793. (1996).

KRUEL, T. M.; EISWIRTHB, M.; SCHNEIDER, F.W. *Computation of Lyapunov spectra: Effect of interactive noise and application to a chemical oscillator*. Physica D 63 117-137. (1993).

LAVER, J. *The phonetic description of voice quality*. Cambridge University Press, Cambridge. (1980).

LE HUCHE, F.; ALLALI, A. *La voz*. Anatomía y fisiología de los órganos de la voz y el habla. 2da ed. Barcelona, Masson, (1993).

LIEBERMAN, P. *Perturbation in vocal pitch*. Journal of Acoustics Society of America., 33, 597-603. (1961).

LOPES, O.; CAMPOS, C.A. *Tratado de Otorrinolaringologia*. São Paulo: Rocca. (1994).

LORENZ, E. N. *Deterministic no periodic flow*. J. Atmos.Sci. 20 : pp.130-141.(1963).

MANDELBROT, B. *"The Fractal Geometry of Nature"*.(1983)

MATASSINI, L.; HEGGER, R.; KANTZ, H.; MANFREDI, C. *Analysis of vocal disorders in a feature space*. Med Eng Phys.;22:413–418. (2000)

MEN, B.; LIANG, C.; ZHAO, X. *Chaotic Analysis on Precipitation Time Series of Sichuan Middle Part in Upper Region of Yangtze*. Nature and Science, 2(1): 74-78. (2004)

MENDE, W., HERZEL, H., WERMKE, K. *Bifurcations and chaos in newborn infant cries*. Phys Lett A;145: 418–424. (1990)

OTT, E.; SAUER, T.; YORKE, J. A. *Coping with Chaos*. New York: Wiley eds. (1994)

PACKARD, N.H.; CRUTCHFIELD, J.P.; FARMER, J.D., SHAW, R.S. *Geometry from a time series*. Phys Rev Lett.;45:712. (1980).

PERKINS, W.; KENT, R. *Functional Anatomy of Speech Language and Hearing*. 2nd ed. Boston: College Hill. Pp65-117. (1986)

PINHO, S.; PONTES, P. *Escala de avaliação perceptiva da fonte glótica: RASAT*. Vox Brasilis - Ano 8 - nº 3. (2002).

POON, C.S.; MERRILL, C.K. *Decrease of cardiac chaos in congestive heart failure*. Nature. ;389:492–495. (1997).

POTTER, R.K.; KOOP, GA; GREEN, H. “*The soundspectrograph*” In: Visible speech. New York, Dover. p.8-15. (1966)

RICHTER, M.; SCHREIBER, T. *Phase space embedding of electrocardiograms*. Chaos-Dyn. 9807035 V1. Germany. (1998)

ROSEN, D.C.; SATALOFF, R.T. *Psychology of Voice Disorders*. San Diego, CA: Singular Publishing Group; (1997).

ROSENSTEIN, M.T.; COLLINS, J.J.; DE LUCA, C. J. *A practical method for calculating largest Lyapunov exponents from small data sets*. Physica D 65 117-134 (1993)

RUELLE, D.; TAKENS, F. *Nature of Turbulence*. Commun. Math Phys.20 (3), 167 (1971)

RUELLE, D.; TAKENS, F. *Nature of Turbulence*. Commun. Math Phys.23 (4), 343 (1971)

SASAKI, C.T.; WEAVER, E.M. *Physiology of the larynx*. Am J Med;103(5A):9s-18s. (1997)

SATALOFF, R.T.; HAWKSHAW, M. *Chaos in Medicine: Source Readings*. San Diego, CA: Singular Publishing Group. (2001).

SAUER, T.; YORKE, J. A.; CASDAGLI, M. *Embedology*. J. Stat. Phys., vol. 65, pp. 579–616. (1991)

SHOJI, K.; MITSUI, Y.; MAKOTO, I.; TETSURO, N.. *Parallel Hardware Algorithms with Redundant Number Representations for Multiple-Valued Arithmetic VLSI*.: 337-345. (1992)

SHRIVASTAV, R. *Multidimensional Scaling of Breathy Voice Quality: Individual Differences in Perception*. Journal of Voice, ARTICLE IN PRESS, (2005).

STEFANSKI, A.; DABROWSKI, A; KAPITANIAK, T. *Evaluation of the largest Lyapunov exponent in dynamical systems with time delay*. Chaos, Solitons & Fractals, Volume 23, Issue 5, Pages 1651-1659 (2005).

TAKENS, F. "*Dynamical systems and turbulence*". Lecture Notes in Mathematics. Berlin, Germany: Springer-Verlag, vol. 898, pp. 366–381. (1981)

TAKENS, F. "*Detecting strange attractors in turbulence*," in Lecture Notes in Mathematics, ed. Rand and Young ~Springer-Verlag, Berlin. Vol. 898, pp 366–381.(1981)

TITZE IR. *Principles of Voice Production*. 2nd ed. Iowa City, IA: NCVS; (2000).

TOSI, O. "*Voice identification: theory and legal applications*". University Park, Baltimor. (1979).

TUCKER, H. *Anatomy of the larynx*. In TUCKER, H: The Larynx. New York: Thieme, (1993)

WOLF, A.; SWIFT, J.B.; SWINNEY, H.L.; VASTANO, J.A. *Determining Lyapunov exponents from a time series*. Physica D ;16:285–317. (1985)

ZEMPLIN, W.L. *Princípios de Anatomia e Fisiologia em Fonoaudiologia*. In Fonoaudiologia. 4th ed. Porto Alegre: Artes Médicas. (2000)

ZHANG, Y.; JIANG, J.J. *Chaotic vibrations of a vocal-fold model with a unilateral polyp*. J Acoust Soc Am. 115:1266–1269. (2004).

ZHANG, Y.; JIANG, J.J.; BIAZZO, L.; JORGENSEN, M.; BERMAN, M. *Perturbation and nonlinear dynamic analyses of voices from patients with unilateral laryngeal paralysis*. J Voice. In press.(2005).