ANDRÉ AUGUSTO SPADOTTO

ANÁLISE QUANTITATIVA DO SINAL DA DEGLUTIÇÃO

Tese apresentada à Escola de Engenharia
de São Carlos da Universidade de São
Paulo, para a obtenção do título de Doutor
em Engenharia Elétrica.

Área de concentração: Processamento de
Sinais e Instrumentação

Orientador: **Prof. Dr. José Carlos Pereira**

São Carlos
2009

FOLHA DE JULGAMENTO

Candidato: Bacharel **ANDRÉ AUGUSTO SPADOTTO**.

Tese defendida e julgada em 07/07/2009 perante a Comissão Julgadora:

Prof. Titular **JOSÉ CARLOS PEREIRA (Orientador)**
(Escola de Engenharia de São Carlos/USP)

APROVADO

Prof. Dr. **ARTHUR OSCAR SCHELP**
(Universidade Estadual Paulista "Julio de Mesquita Filho"/UNESP/campus de Botucatu)

APROVADO

Dr. **RODRIGO CAPOBIANCO GUIDO**
(Instituto de Física de São Carlos/USP/Pós-doutorando)

APROVADO

Prof. Associado **CARLOS DIAS MACIEL**
(Escola de Engenharia de São Carlos/USP)

APROVADO

Prof.^a. Dr.^a. **SUELY PEREIRA DE OLIVEIRA**
(University of Iowa)

APROVADO

Prof. Titular **GERALDO ROBERTO MARTINS DA COSTA**
Coordenador do Programa de Pós-Graduação em Engenharia Elétrica e
Presidente da Comissão de Pós-Graduação

Dedicatória

*Dedico esta tese aos meus pais
João Roberto Spadotto e Leonilda Pires Spadotto*

AGRADECIMENTOS

A DEUS, pois mesmo diante de meu cepticismo, creio que Ele me guia ao caminho certo que deve estar reservado para mim.

Aos meus pais Roberto e Leonilda pelo apoio incondicional, dando suporte familiar para que eu pudesse concluir este trabalho.

Aos meus irmãos Eduardo, Fernanda e Juliana pela amizade, compreensão e apoio em todos os momentos.

Ao meu avô Antônio Spadotto (*in memoriam*), que nos deixou quando a versão final desta tese estava para ser impressa.

À grande amiga Ana Rita Gatto, pois sem ela, esta tese simplesmente não existiria.

Ao amigo João Paulo Papa, pela ajuda com o classificador, mas também em muitos outros tópicos, principalmente nas parcerias em pesquisa.

Ao André Pagnin, Fernando Pagnin e Paula, meus cunhados, que sempre me apoiam no que for necessário.

À amiga Paula Cristina Cola, uma fonoaudióloga, que foi uma das idealizadoras, participou e sempre esteve disposta a ajudar em todo o decorrer deste trabalho.

Ao Danilo Spadoti, pelo companheirismo e por ter-me aberto as portas de sua casa.

Ao Prof. José Carlos Pereira, pela orientação necessária não só no que diz respeito à conclusão deste trabalho, mas em diversos assuntos acadêmicos e pessoais.

Ao médico Arthur Oscar Schelp, que se tornou um amigo, sempre incentivando e ajudando nas pesquisas científicas, além das orientações no âmbito profissional e pessoal.

Ao professor amigo Rodrigo Capobianco Guido, por sempre estar disposto a me receber quando sua ajuda era solicitada, além do fornecimento de muitos códigos-fonte.

A todos os parentes e amigos que sempre me incentivam a continuar os estudos, contribuindo direta ou indiretamente para que eu concluísse este trabalho.

Ao Arlindo Montagnoli Neto por sempre estar disposto a me ajudar.

A todos os funcionários do Departamento de Engenharia Elétrica da USP-São Carlos.

A Elena Palloni Gonçalves da Biblioteca da USP pela ajuda na padronização das referências.

A todos que aceitaram servir de voluntários para que pudéssemos coletar os exames e realizar testes.

Ao Prof. Roberto de Oliveira Dantas e a Danielle Ramos Domenis, ambos do Hospital das Clínicas da USP de Ribeirão Preto, que sempre me receberam e disponibilizaram o equipamento de RX, além de participar do processo de aprovação do Projeto pelo Comitê de Ética em Pesquisa.

Sou grato ainda ao Departamento de Neurologia, Psicologia e Psiquiatria da Faculdade de Medicina de Botucatu pela liberação das horas necessárias para que eu pudesse concluir este trabalho.

***“Mesmo desacreditado e ignorado por todos,
não posso desistir, pois para mim,
vencer é nunca desistir.”***

Albert Einstein

RESUMO

SPADOTTO, A.A. (2009). **Análise quantitativa do sinal da deglutição**. 105p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2009.

Neste trabalho, buscou-se compreender a morfologia e os componentes do sinal da deglutição. Na busca desse entendimento diversas técnicas foram empregadas. No intuito de fazer marcações fidedignas em trechos específicos do sinal, o qual foi analisado simultaneamente com a imagem da videofluoroscopia da deglutição, considerado o melhor método atual na avaliação da dinâmica da deglutição. Os parâmetros numéricos utilizados para análise também foram abrangentes e com base em técnicas atuais de processamento de sinais, como o emprego de transformada *wavelet*. Quanto à classificação dos sinais, foram utilizados classificadores modernos como Floresta de Caminhos Ótimos, Máquinas de Vetores de Suporte, Redes Neurais Artificiais e Classificador Bayesiano, dando maior ênfase ao primeiro, por possuir um custo computacional bem menor quando comparado aos outros 3, e conseqüentemente convergindo mais rapidamente ao resultado. Foram avaliados 84 sinais, divididos em 2 grupos separados pela consistência do bolo alimentar oferecido (líquido e pastoso). Na distinção e/ou caracterização desses tipos foi definido um subconjunto com 4 variáveis que proporcionou uma boa acurácia na separação das classes representantes de cada tipo de bolo alimentar.

Palavras-Chave: Floresta de Caminhos Ótimos. OPF. Sinal da deglutição. Transformada *Wavelet*.

ABSTRACT

SPADOTTO, A.A. (2009). **Quantitative analysis of the swallowing signal**. 105 p. Tese (Doutorado) - Escola de Engenharia de São Carlo, Universidade de São Paulo, São Carlos, 2009.

This work proposes to understand the morphology and the components of the swallowing signal. In pursuit of this understanding, a variety of techniques were employed. In order to make reliable markings on specific portions of the signal, the signal was examined simultaneously with videofluoroscopic swallowing, which is considered the best method in the evaluation of swallowing dynamics. The parameters used for numerical analysis were based on current signal processing techniques, such as: Wavelet Transform, Optimum Path Forest, Support Vector Machines, Artificial Neural Networks and Bayesian Classifier, emphasizing the first technique, due to a much lower computational cost when compared to the previous, and, consequently, the results converged much faster. Eighty four signals, divided into 2 groups separated by the consistency of food bolus offered (liquid and thickened), were evaluated. For distinction and / or characterization of such types, a subset with 4 variables was defined, providing a good accuracy in the separation of these classes representing each type of consistency of the food bolus.

Keywords: Optimum Path Forest. OPF. Swallowing Signal. Wavelet Transform.

LISTA DE FIGURAS

Figura 1 -	Estruturas anatômicas da deglutição	19
Figura 2 -	Fases da deglutição	21
Figura 3 -	Comparação da frequência de pico entre deglutições secas e molhadas	26
Figura 4 -	Translação (a) e dilatação (b), da <i>wavelet</i>	37
Figura 5 -	Decomposição multiresolução de um sinal <i>chirp</i> contendo um trecho ruidoso. Em (a) tem-se os detalhes do sinal $d_m(t)$ e em (b) os sinais de aproximação $x_m(t)$	42
Figura 6 -	Árvore logarítmica representando <i>wavelets</i>	43
Figura 7 -	(a) Função <i>scaling</i> de Haar e (b) função <i>wavelet</i> de Haar, ambas com versões de deslocamento e dilatação	45
Figura 8 -	Forma de onda das funções <i>scaling</i> e <i>wavelet</i> para <i>wavelets</i> de suporte compacto de Daubechies, para $N = 2, 3, 5, 7$ e 9	46
Figura 9 -	(a) Grafo completo ponderado nas arestas. (b) MST do grafo completo. (c) Protótipos escolhidos como sendo os elementos adjacentes de classes diferentes na MST (nós circulados). (d) Floresta de caminhos ótimos resultante para a função de valor de caminho $f_{\max}$ e dois protótipos. (e) Uma amostra de teste (triângulo) da classe 2 e suas conexões (linhas pontilhadas) com os nós do conjunto de treinamento. (f) O caminho ótimo do protótipo mais fortemente conexo	60
Figura 10 -	Área de trabalho do <i>software</i> de análise da deglutição	68
Figura 11 -	Área de trabalho do <i>software</i> de análise de vídeo e sinal simultâneos	70
Figura 12 -	Comparação da forma de onda dos sinais da deglutição entre indivíduos	71
Figura 13 -	Diferenças estatísticas ($p < 0,05$) entre momentos e grupos em relação aos parâmetros	75
Figura 14 -	Medidas estatísticas do tempo da deglutição	76
Figura 15 -	Intervalos entre eventos	76

Figura 16 - Medidas estatísticas dos 4 parâmetros que mais diferiram estatisticamente	77
Figura 17 - Histogramas da faixa de maior concentração de energia com 10 faixas	78
Figura 18 - Histogramas da faixa de maior concentração de energia com 19 faixas	79
Figura 19 - Média da energia por faixa de frequência	80
Figura 20 - Acurácia obtida com OPF, SVM, ANN e BC	81
Figura 21 - Tempo médio de execução dos classificadores	83
Figura 22 - Acurácia obtida com OPF	83
Figura 23 - Variáveis com acurácia acima da média obtida com OPF	84
Figura 24 - Medidas estatísticas da acurácia obtida com OPF	85

LISTA DE TABELAS

Tabela 1 -	Posição do bolo em relação a eventos	30
Tabela 2 -	Parâmetros numéricos para análise do sinal	50
Tabela 3 -	Intervalos de frequências para 10 intervalos	55
Tabela 4 -	Intervalos de frequências para 19 intervalos com sobreposição ...	56
Tabela 5 -	Momentos analisados	69
Tabela 6 -	Variáveis com diferença estatística entre os grupos	73
Tabela 7 -	Correlação com mais 70% entre os parâmetros para os grupos e momentos	74
Tabela 8 -	Duração dos tempos de deglutição	76
Tabela 9 -	Posição do bolo alimentar	80
Tabela 10 -	Tempo médio de execução dos classificadores	82
Tabela 11 -	Medidas estatísticas da acurácia obtida com OPF	85

LISTA DE SIGLAS

ANN -	<i>Artificial Neural Networks</i>
BC -	<i>Bayesian Classifier</i>
cm -	Centímetro
CO	Cavidade Oral
dB -	Decibel
DF -	Dimensão Fractal
DR -	Taxa de Dispersão
DWT -	<i>Discret Wavelet Transform</i>
EES -	Esfíncter esofágico superior
ES -	Entropia de Shannon
f0 -	Frequência do Pico de Energia
FIFO -	<i>First-in-first-out</i>
Hz -	Hertz
IQR -	Interquartil
KV -	Kilo-Volt
LPC -	<i>Linear Predictive Coding</i>
mA -	Mili-Ampère
MAD -	Média do desvio absoluto
mL -	Mililitro
ms -	Milissegundo
MST -	<i>Minimum Spanning Tree</i>
OPF -	<i>Optimum Path Forest</i>
P -	Potência
PLPC -	Potência do Envelope
SVM -	<i>Support Vector Machines</i>
ZA -	Estacionaridade
ZCR -	Taxa de Cruzamento por Zero

LISTA DE SÍMBOLOS

μ -	Média
$\psi(t)$ -	Função <i>wavelet</i>
$\phi(t)$ -	Função <i>scaling</i>
σ^2 -	Variância

SUMÁRIO

1 INTRODUÇÃO	15
2 FUNDAMENTOS DA DEGLUTIÇÃO	18
2.1 Anatomia e fisiologia da deglutição	18
2.2 Ausculta Cervical	24
2.3 Sinal Acústico da Deglutição	25
3 TEORIA <i>WAVELET</i>	35
3.1 A <i>Wavelet</i>	35
3.2 Transformada <i>Wavelet</i> Discreta	37
3.3 Transformada <i>Wavelet</i> Rápida	40
3.4 Transformada <i>Wavelet</i> por Multiresolução	42
3.5 Exemplos de <i>Wavelet</i>	44
3.5.1 <i>Wavelet</i> de Haar	44
3.5.2 <i>Wavelet</i> de Daubechies	45
4 CARACTERÍSTICAS NUMÉRICAS DO SINAL DA DEGLUTIÇÃO	47
4.1 Definição de parâmetros	49
4.2 Definição de faixas de frequência para base de energia	54
5 CLASSIFICADOR	57
5.1 Fundamentação teórica do OPF	59
5.2 Treinamento	62
5.3 Classificação	64
6 MÉTODOS	65
7 RESULTADOS	72
7.1 Estatística Convencional	74
7.2 Classificadores	81
8 DISCUSSÃO	86
9 CONCLUSÃO	94
REFERÊNCIAS	96
ANEXO 1.....	102
ANEXO 2.....	103
TRABALHOS CIENTÍFICOS	104

CAPÍTULO 1

INTRODUÇÃO

O estudo da deglutição e seus distúrbios (disfagia) é muito importante do ponto de vista clínico. É de grande valia a detecção precoce da disfagia, principalmente nos casos mais graves, nos quais o risco de aspiração, que se caracteriza pela entrada de alimento na via aérea, é elevado. Essa detecção pode evitar grandes riscos ao paciente, como, por exemplo, o desenvolvimento de uma pneumonia.

A disfagia orofaríngea é caracterizada por qualquer alteração na dinâmica da deglutição, podendo causar prejuízo aos aspectos nutricionais, hidratação, função pulmonar e integração social do indivíduo. O diagnóstico precoce é crucial para o prognóstico dos pacientes com disfagia. O melhor e mais objetivo método existente atualmente para avaliação da dinâmica da deglutição é a videofluoroscopia da deglutição, mas que impõe limitações pelo fato de expor os pacientes à radiação, não podendo ser realizado com frequência e nem ser muito prolongado.

A avaliação videofluoroscópica da deglutição depende de muito conhecimento e treinamento por parte do avaliador, em relação às estruturas anatômicas e processos neurofisiológicos envolvidos no ato de deglutir.

Ferramentas que permitem a detecção precoce de alterações no processo de deglutição são de suma importância, não só científica, mas também clínica. No campo da pesquisa, essas ferramentas também são importantes, pois fornecem

dados objetivos reduzindo muito a subjetividade por parte de quem avalia. Já na clínica, podem auxiliar na decisão das condutas terapêuticas, apontando quais manobras terapêuticas devem ser utilizadas ou qual a consistência alimentar mais segura para a alimentação. Este trabalho parte dessa idéia e estuda ferramentas não invasivas e inócuas baseadas na análise do sinal da deglutição, capturado por meio de um microfone, não causando risco ou desconforto ao avaliado.

Um dos maiores problemas na análise do sinal da deglutição decorre de sua irregularidade e do fato de não haver comprovação científica satisfatória da fonte de geração para cada evento específico. O sinal não possui métodos e parâmetros específicos para análise. Não há parâmetros numéricos bem descritos, que autorizem dizer que determinada deglutição está fora do padrão de normalidade, ou até mesmo em qual patologia tal deglutição se encaixa.

Outro tópico importante a se ressaltar é que a deglutição é um processo fisiológico que engloba muitos eventos, o que impossibilita a análise do sinal como um todo, embora este seja relativamente breve. Vale ressaltar que, pelo sinal, é muito difícil demarcar onde ocorrem tais eventos.

Diante do exposto, este trabalho pretende analisar o sinal da deglutição de pessoas saudáveis simultaneamente com a imagem videofluoroscópica. A análise do posicionamento das estruturas anatômicas pelas imagens auxilia a demarcar com mais fidedignidade onde ocorre cada evento. De posse das marcações em cada sinal, é possível extrair diversos parâmetros para análise desses sub-sinais, com o intuito de caracterizá-los.

Esta tese está organizada da seguinte maneira: o capítulo 2 descreve o funcionamento e as características da deglutição, nos capítulos 3, 4 e 5 é demonstrado o embasamento teórico-matemático das técnicas utilizadas para o

processamento e classificação do sinal, o capítulo 6 apresenta a metodologia aplicada ao estudo, no capítulo 7 os resultados obtidos são exibidos e, finalmente, a discussão e a conclusão são apresentadas no capítulo 8 e no capítulo 9, respectivamente.

CAPÍTULO 2

FUNDAMENTOS DA DEGLUTIÇÃO

Neste capítulo serão descritos, anatômica e fisiologicamente, o funcionamento da deglutição e as características sonoras do sinal da deglutição, além de se caracterizar a ausculta cervical.

2.1 Anatomia e fisiologia da deglutição

O ato de deglutir pode ser dividido em 4 fases: (1) preparatória oral, manipulação do alimento na boca e mastigação, se necessário; (2) oral, quando a língua impulsiona o alimento até ocorrer o desencadear da deglutição; (3) faríngea, condução do bolo alimentar pela faringe; (4) esofágica, na qual o bolo é conduzido do esôfago cervical e torácico até o estômago por meio de movimentos peristálticos esofagianos (LOGEMANN, 1983). Essas fases serão melhor explicadas mais adiante.

A figura 1 mostra as estruturas anatômicas envolvidas no ato de deglutir. Na figura são demonstrados sequencialmente a partir da boca: faringe, bolo alimentar, esfíncter esofágico superior, esôfago, esfíncter esofágico inferior, diafragma e finalmente o estômago.

Figura 1 - Estruturas anatômicas da deglutição
Fonte: Manual... (2009)

Fase preparatória oral: os movimentos padrões nessa fase da deglutição variam, dependendo do volume e da consistência da substância deglutida. Quando o alimento é colocado na boca, os lábios são mantidos fechados para que o alimento ou líquido não extravasem pela cavidade oral. Quando o bolo líquido é colocado na boca, a língua organiza o bolo alimentar formando um sulco, pressionada contra o palato, como se fosse uma concha contendo o líquido, preparando o alimento para posteriormente ser deglutido. Alguns sujeitos podem, primeiro, espalhar o líquido dentro da boca, manipulando-o para, depois, organizá-lo medialmente. A manipulação de substâncias com consistências pastosas depende da preferência do indivíduo, podendo ser manipuladas ou não dentro da cavidade oral de modo similar ao líquido. Já alimentos de consistência sólida requerem mastigação, que envolvem movimentos laterais rotatórios de mandíbula e língua para, assim, serem triturados, preparando-os para ser deglutidos (LOGEMANN, 1983).

Fase oral: quando a língua inicia o movimento ântero-posterior do bolo, com a elevação da língua progredindo sequencialmente mais adiante, empurrando o

bolo para trás; durante esse tempo uma ranhura central é formada, atuando como uma rampa para o alimento, guiando-o posteriormente. No momento em que o bolo passa pelo pilar das fauces, a fase oral é finalizada e a fase involuntária ou fase faríngea é desencadeada. A fase oral típica leva um tempo menor que um segundo para se completar e é considerada uma fase voluntária. Esta fase é de fundamental importância para desencadear a fase seguinte, pois sem esse ato voluntário de iniciar a deglutição, o ato de elicitar a fase seguinte pode mostrar-se atrasado ou reduzido.

Fase faríngea: fase involuntária. Inicia-se com o desencadear da deglutição que, em indivíduos normais, ocorre quando o bolo passa na região do pilar das fauces, no momento que é ejetado, região esta considerada a mais sensível para desencadear-lo. Há, sem dúvida, uma relação entre o ato voluntário de deglutir e o desencadear da deglutição. Está claro que simplesmente colocando alimento ou líquido na boca, a deglutição não será desencadeada, sendo necessária uma iniciação voluntária.

Durante essa fase as várias atividades fisiológicas que ocorrem simultaneamente, e que são críticas para o sucesso da deglutição, incluem: (1) elevação e retração do véu palatino e completo fechamento do esfíncter velo faríngeo impedindo a entrada de substância na cavidade nasal; (2) iniciação dos movimentos peristálticos faríngeos, carregando o bolo alimentar por ações sequenciais peristálticas dos constritores faríngeos pela faringe até passar pelo esfíncter cricofaríngeo; (3) elevação e fechamento da laringe nos três esfíncteres (epiglote / pregas arriepiglóticas, falsas pregas vocais e verdadeiras pregas vocais) para prevenir a entrada de alimento ou saliva na via aérea; (4) relaxamento do esfíncter cricofaríngeo para permitir que substâncias passem da faringe ao esôfago.

O tempo de trânsito faríngeo, que é o tempo durante o qual o bolo se move do ponto em que se inicia a fase faríngea, região do pilar das fauces até a junção do cricofaríngeo no esôfago, é normalmente de um segundo ou menos.

Fase esofágica: o tempo de trânsito esofágico pode ser medido do ponto em que o bolo entra no esôfago na junção do cricofaríngeo até passar ao estômago na junção gastro-esofágica. Esse tempo geralmente varia de oito a vinte segundos. As ondas peristálticas na faringe iniciam com o desencadear da deglutição e se propagam pelo esôfago.

Figura 2 - Fases da deglutição. Em (a) e (b) têm-se posições do alimento na fase oral já em (c) e (d) tem-se a localização do alimento na fase faríngea

Fonte: Manual... (2009)

Essas fases se inter-relacionam e configuram um processo dinâmico e complexo, com controle neuromotor. O sincronismo entre as fases permite que o alimento seja transportado da boca até o estômago sem que haja penetração e/ou aspiração laringotraqueal de alimento ou saliva (LOGEMANN, 1983).

Os distúrbios da deglutição são definidos como disfagia orofaríngea quando apresentam sinais e sintomas específicos, caracterizados por alterações em qualquer fase e/ou entre as etapas da dinâmica da deglutição, podendo ser congênitos ou adquiridos após comprometimento neurológico, com prejuízo dos aspectos nutricionais, hidratação, função pulmonar e integração social do indivíduo (FURKIM; SILVA, 1999). Os sintomas da disfagia orofaríngea neurogênica incluem dificuldade para iniciar a deglutição, regurgitação nasal, dificuldade no controle de saliva, episódios de tosse durante a alimentação e retenção do bolo alimentar no trato digestivo (BUCHHOLZ, 1994).

A abordagem ao indivíduo disfágico é composta de avaliação fonoaudiológica clínica e instrumental da deglutição, a qual depende do conhecimento do avaliador sobre as estruturas anatômicas e sobre os processos neurofisiológicos envolvidos na deglutição, importantes para compreender a inter-relação entre as fases e determinar o raciocínio clínico e terapêutico (SILVA; GATTO; COLA, 2003). A avaliação deve constar de informações de anamnese e procedimentos específicos que permitam verificar o funcionamento das fases da deglutição como auxílio ao raciocínio clínico.

Para avaliação objetiva da deglutição existem vários exames, dentre os quais a videofluoroscopia da deglutição. Esse exame auxilia no diagnóstico e na definição de condutas terapêuticas. A videofluoroscopia é o método pelo qual imagens dinâmicas de RX na tela podem ser documentadas em fitas de vídeo. O

processo permite a reavaliação dos exames, sem necessidade de repetição quando se deseja uma nova análise, reduzindo assim a exposição à radiação, tanto para o paciente como para o profissional que realiza o exame (COSTA et al., 1992).

O RX dinâmico examina detalhes da fisiologia oral, faríngea e esofagiana cervical durante a deglutição, mediante a deglutição de alimento com adição de bário (substância de contraste ao RX). As imagens fornecem informações desde as estruturas, bem como detalhes da fisiologia envolvidos no mecanismo de deglutição. Utilizam-se alimentos de consistências diferentes durante a avaliação. Nesse exame, o paciente permanece preferencialmente sentado, em posição lateral à do foco de geração das imagens, as quais também podem ser vistas no plano ântero-posterior (LOGEMANN, 1983).

Atualmente a videofluoroscopia da deglutição é considerada o método padrão ouro, para avaliar qualitativa e objetivamente a dinâmica da deglutição. É possível com esse RX dinâmico a visualização de todas as fases (LOGEMANN, 1983). No entanto, tal exame apresenta algumas limitações, como exposição do paciente a radiação, limitando a repetição periódica para monitoramento das terapias de deglutição. A análise dos parâmetros registrados pela videofluoroscopia é basicamente subjetiva e varia inter e intra-examinador, reduzindo consideravelmente a análise quantitativa dos dados.

Recentemente surgiram estudos analisando a videofluoroscopia da deglutição com o auxílio de *software*. O emprego de *softwares* permite uma análise quantitativa como, por exemplo, a medição do tempo do trânsito do bolo alimentar pelas fases da deglutição (KENDALL et al., 2000; SPADOTTO et al., 2006).

2.2 Ausculta Cervical

A ausculta cervical é um método não invasivo que se caracteriza por auxiliar na audição dos sons da deglutição mediante o uso de um instrumento de amplificação. Pode ser do modo convencional, com a utilização de um estetoscópio bi-auricular, ou por meio de transdutores, como microfone ou acelerômetro. A captura do sinal por transdutores permite a digitalização do sinal para um posterior processamento digital.

Embora o uso de ambos os tipos seja crescente, a ausculta cervical ainda não é aceita como ferramenta principal de avaliação de pacientes disfágicos, devido a sua subjetividade. Uma melhor comprovação científica deve ser fornecida para dar credibilidade e para a ferramenta ser aceita do mesmo modo que a videofluoroscopia, manometria faríngea, eletromiografia e endoscopia por fibra óptica. Um dos maiores problemas enfrentados para a comprovação é que a fonte de geração dos sons da deglutição e as diferenças sonoras entre sons de deglutições normais e disfágicas, não foram analisadas cientificamente, ou seja, com os estudos atuais existentes ainda não é possível estabelecer uma correlação entre o som obtido e o evento correspondente.

O uso da ausculta cervical convencional na avaliação da deglutição faz parte da rotina na avaliação de pacientes disfágicos. Esse método auxilia o clínico a distinguir entre deglutição normal e disfágica, podendo favorecer na detecção de sinais sugestivos de penetração laríngea e/ou aspiração antes ou após a deglutição. No entanto, é um método muito subjetivo que exige pessoas treinadas para executá-los. A análise acústica digital do sinal da deglutição visa trazer dados mais objetivos para a detecção de disfunções na deglutição e maior aceitação no meio científico.

2.3 Sinal Acústico da Deglutição

Lear, Flanagan e Morrees (1965), analisando o espectrograma do sinal da deglutição, reportaram em seu estudo, um pico súbito muito característico em 4600 Hz, o qual cai em 30 milissegundos (ms) para 1000 Hz. Depois se segue um elemento irregular em 3000 Hz com duração de 100 ms antes de um silêncio quase total com 200 ms de duração, seguido de outra banda irregular com frequência por volta de 1500 Hz e duração de 150 ms.

Logan, Kavanagh e Wornall (1967) documentam que a duração dos sons da deglutição varia de 250 a 800 milissegundos e que o número e a duração das explosões da energia sonora por deglutição são diferentes, mas o padrão é essencialmente inalterado.

Segundo Mackowiak, Brenman e Friedman (1967), três explosões acústicas temporalmente relacionadas são vistas em cada deglutição. O início da deglutição perdura por 50 ms (α - *alpha*), seguido por um intervalo de silêncio, geralmente variando de 100 a 150 ms (intervalo ab). O segundo componente acústico da complexa deglutição (β - *beta*) é geralmente o mais longo em duração e médias, com tempo entre 150 e 200 ms. As amplitudes dos componentes α e β são similares. Uma terceira explosão acústica (δ - *delta*) tem sido notada na maioria das deglutições utilizando água; esse componente normalmente aparece entre 300 e 400 ms após o final do sinal β e é de curta duração e baixa amplitude. A origem desse componente não é conhecida, mas a natureza da sequência temporal sugere duas possibilidades. A primeira seria a movimentação da epiglote no final da deglutição, e a outra, as atividades nas áreas inferiores do esôfago, com o líquido deglutido aproximando-se do esfíncter cárdio-esofágico.

Várias diferenças importantes são notadas entre deglutições secas (saliva) e molhadas (água). A primeira é que a intensidade absoluta da deglutição molhada é maior e tende a ter uma duração mais curta. Uma segunda é que as secas mostram um pico de energia entre 1000 e 1100 Hz para a duração completa do som da deglutição, enquanto as deglutições molhadas começam com uma baixa frequência (400 a 600 Hz) e aumentam a frequência de concentração de energia alcançando 1000 Hz no final da deglutição, conforme exhibe a figura 3. E finalmente uma terceira observação é que o componente δ do sinal está ausente em deglutições secas (MACKOWIAK; BRENNAN; FRIEDMAN, 1967).

Figura 3 - Comparação da frequência de pico entre deglutições secas e molhadas
Fonte: Modificado de Mackowiak, Brennan e Friedman (1967)

Hamlet, Nelson e Patterson (1988) documentam uma interessante teoria das causas fisiológicas dos sons da deglutição, relatando que a característica proeminente do sinal é um ruído de banda larga relativamente breve, de aproximadamente 200 ms, que corresponde à rápida passagem do bolo através da faringe e do esfíncter esofágico superior. Seu espectro contém maiores componentes de alta frequência para a consistência líquida do que para a pastosa.

Outras características de baixa amplitude do sinal correspondem a movimentos estruturais do hióide, laringe ou epiglote.

Hamlet, Nelson e Patterson (1990) relatam que a deglutição normal consiste, principalmente, de dois componentes temporais distintos, identificados por um duplo clique audível. Embora com uma variedade individual considerável, o primeiro componente com intervalo de 50 ms é seguido por uma longa concentração de energia com tempo entre 150 e 200 ms, chamado de explosão (*burst*). Na deglutição molhada um terceiro componente breve e de baixa amplitude pode ser ouvido 300 a 400 ms após o final do segundo. A concentração de energia acústica localiza-se abaixo dos 3000 Hz. Em uma detalhada investigação do fluxo de líquido pelo cricofaríngeo, os autores encontraram uma média do pico inicial no sinal da deglutição em 556 Hz e um segundo pico em 1384 Hz. Observam consistente correlação entre a rápida mudança no espectro de potência por volta de 1000 Hz e a passagem do bolo através do cricofaríngeo. Ressaltam ainda a importância da entrada do bolo pressurizado pela propulsão de língua, representada pelo movimento posterior de base de língua.

Robbins et al. (1992) descrevem em seu estudo que o ato da deglutição é gradualmente mais lento depois dos 45 anos de idade e com aproximadamente 70 anos é significativamente mais demorado que em indivíduos com menos de 45 anos. Os autores destacaram somente uma diferença significativa, registrando que a abertura do esfíncter esofágico superior das mulheres tem uma duração maior do que a dos homens. Atribuem isso ao fato de a faringe e o esfíncter esofágico superior serem menor anatomicamente nas mulheres, salientando que o esfíncter esofágico superior pode precisar de uma abertura maior para acomodar um bolo alimentar de tamanho equivalente.

Segundo Selley et al. (1994), o primeiro som da deglutição deve estar relacionado com eventos associados à elevação máxima da laringe e ao deslocamento da epiglote podendo, portanto, ser devido ao fato de o bolo, sob alta pressão, passar pelo esfíncter esofágico superior. O segundo som inicia quando a maioria do bolo alimentar está no esôfago, passando pelo esfíncter esofágico superior e a parte distal ainda está em nível da valécula.

Um estudo feito por Takahashi, Groher e Michi (1994) demonstrou, com a metodologia empregada, que o detector acústico mais apropriado para captura do som da deglutição era o acelerômetro, afixado com fita dupla face de papel, em consequência da ampla faixa de resposta em frequência e do baixo nível de atenuação. Um local ótimo indicado para fixação do acelerômetro é a área sobre a borda lateral da traqueia imediatamente inferior à cartilagem cricóide. Em sua pesquisa foram analisados como detectores o acelerômetro e o microfone de eletreto, juntamente com três tipos de fitas adesivas, e na avaliação do local do pescoço a ser afixado o detector, foram mapeadas 24 posições sob o pescoço e comparadas uma a uma.

Cichero e Murdoch (1998) sugeriram que os sons da deglutição são gerados de maneira similar aos sons do coração que, por sua vez, são gerados por complexos fechamentos de válvulas e pelas vibrações das estruturas musculares. Em seu estudo foi oferecida a hipótese de existirem três sons na deglutição. O primeiro corresponde à ejeção oral pela movimentação da língua contra a parede posterior da faringe e ao fechamento da válvula laríngea. Mais especificamente, a ação de fechamento das pregas vocais verdadeiras e movimento das pregas ariepiglóticas deslocando a epiglote para posição horizontal na faringe, favorecendo o fechamento laríngeo, e sua elevação e anteriorização, contribuindo para vibração

na faringe. Esse primeiro som está, segundo se afirma, na faixa de frequência entre 400 e 600 Hz, que é o primeiro formante da frequência ressonante natural da vibração do trato vocal. Dessa forma a elevação e anteriorização do complexo hio-laríngeo podem contribuir para formar o primeiro som da deglutição (MCCONNEL et al., 1988). A passagem pela faringe é relativamente silenciosa. Nesse estágio a língua sela a orofaringe, e o esfíncter velofaríngeo separa a nasofaringe. A abertura do esfíncter esofágico superior combinado com a onda de limpeza faríngea e mediada pelos constritores faríngeos produzem uma longa vibração na faringe, gerando assim o segundo som da deglutição. Esse segundo som tem aparentado ocorrer em aproximadamente 1000 Hz. Finalmente, diz-se que o terceiro som, ocorrendo de 300 a 400 ms após o segundo, é baixo em duração e amplitude. Sugerem que esse último som pode ser o resultado da abertura da via aérea, causando uma menor vibração da laringe (MACKOWIAK; BRENNAN; FRIEDMAN, 1967).

Boiron, Rouleau e Metman (1997) constataram diferenças na duração da apneia com mudanças no volume e consistência do bolo alimentar. Na deglutição de líquidos, quanto maior o volume maior era a duração da apneia, 600 ms para 5 mL, e 960 ms para 15 mL. Para consistência pastosa (iogurte) a duração era menor que para substâncias líquidas, 580 ms com 5 mL, e 920 ms com 15 mL.

Cichero e Murdoch (1998) relataram também que os sons gerados pela laringe durante a fonação, por exemplo, e o som modificado pelo trato vocal durante a fala são periódicos. Esses são os sons das partículas de ar movimentando-se em um caminho regular padronizado. Formas de onda acústica de sons periódicos mostram um padrão visual repetitivo. No entanto, sons aperiódicos são produzidos por movimentos randômicos de partículas de ar, e são reconhecidos como ruído.

Para sinais aperiódicos as componentes de frequência dos sons não estão relacionadas e não mostram um padrão visual repetitivo. Sons da deglutição são aperiódicos (HAMLET; NELSON; PATTERSON, 1988); por outro lado, sons fonatórios são periódicos. Pode-se assim concluir que sons fonatórios e sons da deglutição são gerados de maneiras diferentes. Qualquer similaridade nas frequências formantes de sons de fonação e sons de deglutição deveria ser atribuída à comunhão do trato vocal com o sistema definido na vibração.

Perlman, Ettema e Barkmeier (2000), verificaram a passagem do bolo durante a deglutição analisando simultaneamente: videofluoroscopia, sinal acústico, fluxo de ar e eletromiografia. A localização da parte proximal da maior concentração do bolo foi identificada para cada uma das 132 deglutições analisadas. Os dados são mostrados na tabela a seguir:

Tabela 1 - Posição do bolo em relação aos eventos

Evento	Videofluoroscopia	% de deglutições
Início da apneia	Cavidade oral	61
	Base de língua	14
	Valécula	20
	Seios piriformes	5
Início do Sinal 1	Cavidade oral	2
	Base de língua	8
	Valécula	27
	Seios piriformes	39
	Esfíncter esofágico superior	17
	Esôfago	7
Pico do Sinal 1	Seios piriformes	18
	Esfíncter esofágico superior	12
	Esôfago	70
Final apneia	Esôfago	48
	Parte distal já está no esôfago	52
Início sinal 2	Seios piriformes	1
	Esfíncter esofágico superior	1
	Esôfago	73
	Parte distal já está no esôfago	25
Pico sinal 2	Esôfago	66
	Parte distal já está no esôfago	34

Fonte: Modificado de Perlman, Ettema e Barkmeier (2000)

Pôde-se observar que o início do primeiro sinal acústico estava associado à presença da maior parte do bolo alimentar (66%) em valécula ou seios piriformes e seu pico em 70% das deglutições estava associado à parte proximal do bolo já no esôfago e em nenhuma deglutição se observava a parte proximal do bolo acima dos seios piriformes. Quando o segundo sinal era visualizado, a parte proximal do bolo estava no esfíncter esofágico superior (EES) ou abaixo dele; no momento do segundo pico o bolo estava dentro do esôfago. O início da apneia da deglutição estava associado em 95% dos casos analisados com a presença da parte proximal do bolo em cavidade oral ou valécula, e o final da apneia em 100% das vezes quando a parte proximal do bolo se encontrava em esôfago. Dessa forma, o período de apneia geralmente começa antes do alimento entrar na hipofaringe e termina depois do bolo entrar no esôfago (PERLMAN; ETTEMA; BARKMEIER, 2000).

Uma revisão metodológica para detecção de sons da deglutição foi feita por Cichero e Murdoch (2002a). Em sua pesquisa eles concordaram com Takahashi, Groher e Michi (1994) quanto à posição da fixação do detector acústico, porém o microfone apresentou melhores resultados quando comparado ao acelerômetro que foi indicado como sendo o melhor por Takahashi, Groher e Michi (1994). Metodologias diferentes foram empregadas nos estudos, mas o fator principal que contribui para a preferência pelo microfone é que Takahashi, Groher e Michi (1994) amplificaram somente o sinal do acelerômetro, causando assim uma rejeição prematura ao microfone.

Cichero e Murdoch (2002b) caracterizaram as mudanças que ocorriam nas características acústicas do som da deglutição de indivíduos normais, variando idade, sexo e volumes de bolo alimentar. Constataram que a duração do som da deglutição era sensível às mudanças de volume e também em relação à interação

volume-idade. Ou seja, a duração do som era significativamente mais curta para 15 mL do que para 5 ou 10 mL. O grupo mais jovem (18-35 anos) deglutia volumes menores (5 mL) com duração significativamente menor, comparado aos demais grupos (36-59 e >60 anos). Os parâmetros acústicos de duração, frequência e intensidade são caracterizados com duração média de 0,4 segundo, intensidade de 43 dB e uma faixa de frequência de 2200 Hz.

Stroud, Lawrie e Wiles (2002) objetivaram verificar a confiabilidade da ausculta cervical quando utilizada pelo especialista isoladamente de outros parâmetros clínicos. Avaliaram na pesquisa 16 pacientes, sendo 4 homens e 12 mulheres. Constataram ocorrência de falsos positivos com significância menor que 0,001, apresentando sensibilidade de 86% e especificidade de 56%.

Youmans e Stierwalt (2005), visando aumentar a objetividade da ausculta cervical, estudaram as faixas de normalidade para as características acústicas e temporais da deglutição em diferentes consistências e volumes do bolo alimentar, diferentes faixas etárias e sexo. Constataram aumento da duração e menor intensidade do sinal acústico com o aumento da idade. Não observaram mudanças em relação ao sexo. Em relação às características acústicas, a duração média mensurada foi 530 ms, intensidade aproximadamente de 60,8 dB e frequência 586,57 Hz.

Morinière, Beutter e Boiron (2006) identificaram 6 componentes de som, com variabilidade inter e intra-sujeitos, sendo que somente 3 componentes ocorriam com maior regularidade. Justificaram esses achados a mínimos intervalos entre os componentes de som, podendo levar à fusão desses quando o bolo passa rapidamente. Quantificaram os componentes principais no sinal da deglutição de indivíduos normais e suas durações. A duração média dos parâmetros acústicos era

710±28 ms, sendo o primeiro componente com duração média de 100±56 ms, o segundo 150±90 ms e 80±54 ms para o terceiro componente. O intervalo entre eles era de 100±66 ms e 190±120 ms respectivamente.

Segundo Leslie et al. (2007), a contribuição da ausculta cervical da deglutição não está clara. Na há uma evidência concreta de que a ausculta cervical dos sons da deglutição possa ser adotada na prática clínica, ou a inclusão dessa técnica a '*guidelines*' clínicos ou protocolos de gerenciamento. Os autores questionam o quanto outros parâmetros clínicos não influenciam os dados auscultados. Uma das principais críticas da ausculta cervical é que não tem correlação com a produção dos sons. Mais avaliações usando métodos concomitantes de áudio e imagem são necessárias antes dessa técnica subjetiva ter validade para o uso clínico. Em seu estudo, foi comparada a nasolaringoscopia simultaneamente com ausculta cervical, mas não se obteve dados conclusivos na identificação dos sons.

Nos últimos anos têm-se dado muitos créditos para a avaliação dos sons da deglutição usando ausculta cervical. O atrativo principal da ausculta cervical é obviamente o seu baixo custo. Manuais que descrevem técnicas relatam aos avaliadores que as características principais da ausculta são dois cliques seguidos de um sopro expiratório. Ainda hoje não pode se inferir qual movimento causa determinado som. Pode-se inferir que entre um número de deglutições, determinado som pode estar associado a algum movimento específico.

Borr, Hielscher-Fastabend e Lucking (2007), visando definir o perfil acústico da deglutição disfágica e as mudanças acústicas entre indivíduos jovens, disfágicos e idosos, encontram a duração da apneia da deglutição significativamente maior em grupo de idosos do que de jovens. Quando se correlacionavam jovens e

disfágicos além da duração da apneia, também se encontrava maior número de goles para completar a deglutição dos indivíduos disfágicos. Investigaram também a confiabilidade dos avaliadores utilizando a ausculta cervical com especificidade de 70% e sensibilidade de 94%.

Visando correlacionar o som ao evento fisiológico, Leslie et al. (2007) usaram a endoscopia da deglutição e o sinal de áudio. Foram avaliados 20 voluntários, sendo 8 homens e 12 mulheres. Nesse estudo encontraram seis componentes de som, sendo que nenhum componente pode ser claramente identificado em todos os indivíduos. Não se pode associar consistentemente nenhum desses componentes a um evento acústico. Os próprios autores levantam a hipótese de a endoscopia não ser o equipamento adequado para realizar esse estudo, apontando algumas desvantagens como a impossibilidade de se visualizar a fase oral e o momento de elevação de laringe.

Morinière et al. (2008) em seu estudo objetivaram entender a origem dos componentes de som da deglutição em sujeitos normais, para isso realizam a captura simultânea da imagem da videofluoroscopia e do sinal de áudio. Foram capazes de identificar três componentes, sendo o primeiro relacionado à ascensão da laringe (bolo em orofaringe ou hipofaringe), encontrado em 81% das deglutições e com duração média 106 ± 47 ms. O segundo relaciona-se à abertura do esfíncter esofágico superior (bolo passando pelo esfíncter esofágico superior), observado em 100% das deglutições e duração média de 185 ± 103 ms, e o terceiro, correspondendo ao retorno das estruturas (laringe descendo e alimento no esôfago), pode ser visualizado em 81% das deglutições, com duração média de 72 ± 38 ms. O intervalo entre estes componentes era 108 ± 44 ms e 236 ± 139 ms.

CAPÍTULO 3

TEORIA WAVELET

A análise por transformada *wavelet* utiliza pequenas funções ondulatórias conhecidas como *wavelets* para transformar o sinal analisado em uma forma que evidencie a informação contida no sinal, permitindo assim, uma melhor extração de características. Essa transformação do sinal é conhecida como transformada *wavelet*, que matematicamente é a convolução da função *wavelet* com o sinal.

3.1 A Wavelet

A transformada *wavelet* é um método de conversão de um sinal em outro sinal, no qual se quer tornar características do sinal original mais inteligíveis para estudos, ou de forma a permitir que o conjunto de dados originais seja descrito mais sucintamente. Para executar a transformada *wavelet* é necessário uma *wavelet*, que é uma função $\psi(t)$, a qual satisfaz alguns critérios matemáticos. Essas funções são manipuladas por um processo de translação (movimento junto ao eixo do tempo) e dilatação (espalhamento da *wavelet*) para transformar o sinal em outra forma que é variável em tempo e escala. A função *wavelet* $\psi(t)$ é conhecida como *wavelet* mãe ou *wavelet* de análise (ADDISON, 2002).

Para ser classificada como *wavelet*, uma função deve satisfazer certos critérios matemáticos, que são:

1 - A *wavelet* deve possuir uma energia finita:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (1)$$

2 - Se $\hat{\psi}(f)$ é a transformada de Fourier de $\psi(t)$

$$\hat{\psi}(f) = \int_{-\infty}^{\infty} \psi(t) e^{-i(2\pi f)t} dt \quad (2)$$

então a seguinte condição deve existir:

$$C_g = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty \quad (3)$$

Isso implica que a *wavelet* não tem componente de frequência zero, $\hat{\psi}(0) = 0$. A equação acima é conhecida como condição de admissibilidade e C_g é chamada de constante de admissibilidade.

3 - Um critério adicional que deve ser assegurado para *wavelets* complexas é que a transformada de Fourier deve ser real e nula para frequências negativas.

Escolhida a *wavelet* mãe, duas manipulações básicas podem ser executadas para permitir maior flexibilidade: alargamento ou estreitamento (dilatação) e movimentação (translação). Na figura 4 (a) é mostrada a translação e em (b) a dilatação. A dilatação e a contração da *wavelet* são definidas pelo parâmetro de dilatação ' a ' e o movimento da *wavelet*, pelo eixo do tempo é definido pelo parâmetro de translação ' b '. A função *wavelet* é definida como (ADDISON, 2002):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (4)$$

Figura 4 - Translação (a) e dilatação (b), da *wavelet*
Fonte: Modificada de Addison (2002)

3.2 Transformada *Wavelet* Discreta

Tradicionalmente os parâmetros a (escala) e b (deslocamento) assumem valores discretos o que, de um modo geral, é feito usando uma discretização logarítmica. Esse tipo de discretização de *wavelet* tem a seguinte forma:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \quad (5)$$

onde os inteiros m e n controlam a dilatação e translação da *wavelet* respectivamente, a_0 é um passo de dilatação fixo específico, com valor maior que 1, e b_0 é um parâmetro de localização devendo ser maior que zero. Os parâmetros de controle m e n estão contidos no conjunto de todos os inteiros, ambos positivos e negativos. Pode-se notar na equação (5) que o tamanho dos passos de translação, $\Delta b = b_0 a_0^m$ é diretamente proporcional a escala *wavelet* a_0^m .

A transformada *wavelet* de um sinal contínuo $x(t)$, usando *wavelets* da forma da equação (5) é então:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t) \frac{1}{a_0^{m/2}} \psi(a_0^{-m} t - nb_0) dt, \quad (6)$$

onde $T_{m,n}$ são valores da transformada *wavelet* discreta dados em localização e escala em grade com índice m,n . Para a transformada *wavelet* discreta, os valores $T_{m,n}$ são conhecidos como coeficientes *wavelet* ou coeficientes de detalhe. A teoria de '*wavelet frames*' fornece um embasamento para o estudo de *wavelets* discretas. *Wavelet frames* são construídos por amostragem discreta em parâmetro de tempo e escala de uma transformada *wavelet* contínua (DAUBECHIES, 1990).

Uma escolha comum para os parâmetros a_0 e b_0 da *wavelet* discreta são 2 e 1, respectivamente. Essa potência de 2 em escala logarítmica de ambos (dilatação e translação) é conhecida como grade diádica. A grade diádica talvez seja a mais simples e mais eficiente discretização para propostos práticos e permite por si própria a construção de uma base *wavelet* ortonormal.

Wavelets com grade diádica compacta geralmente são escolhidas por serem ortonormais (DAUBECHIES, 1988). Essas *wavelets* são ortogonais entre si e normalizadas para ter energia unitária, ou seja, o produto de cada *wavelet* com todos os outros no mesmo sistema diádico (aqueles que são transladados e/ou dilatados, ambas as versões) é zero (ADDISON, 2002). Isso significa que a informação armazenada no coeficiente *wavelet* $T_{m,n}$ não é repetida em outros locais e permite uma completa recuperação do sinal original sem redundância. Uma base ortonormal tem componentes vetoriais capazes de definir o sinal completamente e são perpendiculares entre si. A grade diádica possibilita um algoritmo computacional rápido (ADDISON, 2002; DAUBECHIES, 1988; DAUBECHIES, 1992).

Usando a grade diádica, a transformada *wavelet* discreta (DWT - *Discret Wavelet Transform*) pode ser escrita como:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t) \psi_{m,n}(t) dt \quad (7)$$

Escolhendo bases *wavelets* ortonormais, $\psi_{m,n}(t)$, pode-se reconstruir o sinal original em termos de coeficientes *wavelet*, $T_{m,n}$, usando a transformada *wavelet* discreta inversa como segue:

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} T_{m,n} \psi_{m,n}(t) \quad (8)$$

Wavelets discretas ortonormais diádicas são associadas com funções *scaling* e suas equações de dilatação. A função *scaling* é associada com o nivelamento do sinal e tem o mesmo formato da *wavelet*, dada por:

$$\phi_{m,n}(t) = 2^{-m/2} \phi(2^{-m}t - n) \quad (9)$$

E possui a seguinte propriedade:

$$\int_{-\infty}^{\infty} \phi_{0,0}(t) dt = 1, \quad (10)$$

onde $\phi_{0,0}(t) = \phi(t)$ é geralmente chamado de função *scaling* pai ou *wavelet* pai. A função *scaling* é ortogonal à translação de si própria, mas não para dilatação. A função *scaling* pode ser convoluída com o sinal para produzir os coeficientes de aproximação como segue:

$$S_{m,n} = \int_{-\infty}^{\infty} x(t) \phi_{m,n}(t) dt \quad (11)$$

Nas equações (9), (10) e (11) pode ser notado que os coeficientes de aproximação são simplesmente a média ponderada do sinal contínuo fatorado por $2^{m/2}$. Os coeficientes de aproximação em uma escala específica m são coletivamente conhecidos como aproximação discreta do sinal. Um sinal $x(t)$ pode ser representado usando uma série de expansão combinada usando os coeficientes de aproximação e coeficientes de detalhe. Se for adicionado o sinal detalhe em uma escala arbitrária (índice m) para a aproximação na escala, obtém-se um sinal aproximado de maior

resolução. Isso é chamado de representação multiresolução (ADDISON, 2002; DAUBECHIES, 1992; MALLAT, 1989).

A equação *scaling* descreve a função *scaling* $\phi(t)$ em termos de versões contraídas e deslocadas de si próprias, como descritas abaixo:

$$\phi(t) = \sum_k c_k \phi(2t - k), \quad (12)$$

onde $\phi(2t - k)$ é a versão contraída de $\phi(t)$, deslocada pelo eixo do tempo por um passo k e fatorado por um coeficiente *scaling* associado, c_k . Existe diferença na terminologia de equação *scaling* e função *scaling*. A equação basicamente diz como a função é construída. Integrando-se ambos os lados da equação acima, pode-se notar que os coeficientes *scaling* devem satisfazer a condição a seguir:

$$\sum_k c_k = 2 \quad (13)$$

Para criar um sistema ortogonal ainda é requerido que:

$$\sum_k c_k c_{k+2k'} = \begin{cases} 2 & \text{se } k' = 0; \\ 0 & \text{caso contrário} \end{cases} \quad (14)$$

Isso também mostra que a soma dos quadrados dos coeficientes *scaling* é igual a 2. Para coeficientes iguais são usados reversamente com sinais alternados para produzir a diferenciação da equação *wavelet* associada. Essa construção assegura que as *wavelets* e suas funções *scaling* correspondentes são ortogonais.

3.3 Transformada *Wavelet* Rápida

Os coeficientes de aproximação com escala de índice $m+1$ são dados por:

$$S_{m+1,n} = \frac{1}{\sqrt{2}} \sum_k c_k S_{m,2n+k} = \frac{1}{\sqrt{2}} \sum_k c_{k-2n} S_{m,k} \quad (15)$$

e usando essa equação podem ser gerados os coeficientes de aproximação de escala índice $m+1$ utilizando os coeficientes *scaling* da escala anterior. Similarmente, os coeficientes *wavelet* podem ser encontrados a partir dos coeficientes de aproximação de uma escala anterior usando os coeficientes *scaling* armazenados b_k como mostrado a seguir:

$$T_{m+1,n} = \frac{1}{\sqrt{2}} \sum_k b_k S_{m,2n+k} = \frac{1}{\sqrt{2}} \sum_k b_{k-2n} S_{m,k} \quad (16)$$

Pode-se notar que se os coeficientes de aproximação $S_{m_0,n}$ são conhecidos em uma escala específica m_0 , então, iterando a aplicação das equações (15) e (16) pode-se calcular os coeficientes *wavelet*, de aproximação e de detalhe, em todas as escalas maiores que m_0 . As duas equações acima representam o algoritmo de decomposição multiresolução (MALLAT, 1989). O algoritmo de decomposição é a primeira metade da transformada *wavelet* rápida, o qual permite computar os coeficientes *wavelet* em vez de computar a convolução da equação. A iteração das equações resulta respectivamente em uma filtragem passa-altas e passa-baixas de entrada ($S_{m,n+k}$) e saídas ($S_{m+1,n}$ e $T_{m+1,n}$). O vetor contendo as sequências $(1/\sqrt{2})c_k$ e $(1/\sqrt{2})b_k$ representa os filtros, onde $(1/\sqrt{2})c_k$ é o filtro passa-baixas, permitindo a passagem de sinais de baixa frequência e assim gerando uma versão nivelada do sinal, e $(1/\sqrt{2})b_k$ é o filtro passa-altas, permitindo a passagem de altas frequências correspondendo aos detalhes do sinal (STRANG; NGUYEN, 1996).

O algoritmo de reconstrução do sinal é dado por:

$$S_{m-1}(t) = \frac{1}{\sqrt{2}} \sum_k c_{n-2k} S_{m,k}(t) + \frac{1}{\sqrt{2}} \sum_k b_{n-2k} T_{m,k}, \quad (17)$$

onde se reutiliza k como índice de localização dos coeficientes transformados em escala de índice m para diferenciação de n , a localização de escala com índice $m-1$.

Por isso, na menor escala, $m-1$, os coeficientes de aproximação podem ser encontrados em termos da combinação dos coeficientes de aproximação e detalhe da próxima escala, m . O algoritmo de reconstrução é a segunda metade da transformada *wavelet* rápida. Na literatura, transformada *wavelet* rápida, transformada *wavelet* discreta, algoritmo de decomposição e reconstrução, transformada *wavelet* rápida ortogonal, algoritmo multiresolução, algoritmo pirâmide, algoritmo árvore, são todos usados com o mesmo intuito.

Figura 5 - Decomposição multiresolução de um sinal *chirp* contendo um trecho ruidoso. Em (a) tem-se os detalhes do sinal $d_m(t)$ e em (b) os sinais de aproximação $x_m(t)$

Fonte: Modificada de Addison (2002)

3.4 Transformada *Wavelet* por Multiresolução

A transformada *wavelet* opera em tempo contínuo (em funções) e em tempo discreto (vetores). A entrada é $f(t)$ ou $x(n)$. A saída é um conjunto de

coeficientes b_{jk} , o qual expressa a entrada na base *wavelet*. As bases *wavelet* e as bases de Fourier têm propriedades especiais. As escalas j e $j-1$ são intimamente relacionadas, assim como as frequências ω e $\omega/2$ também (ADDISON, 2002; MALLAT, 1989; STRANG; NGUYEN, 1996).

A natureza recursiva de *wavelets* pode ser demonstrada na forma de uma árvore de bancos de filtros, como na figura a seguir. Nesta representação os filtros passa-altas D computam diferenças da entrada. O passo de subamostragem simbolizado por $(\downarrow 2)$ mantém as diferenças $(x(2k) - x(2k-1))/\sqrt{2}$. Esses são resultados finais porque não são transformados novamente. Eles estão no fim de sua ramificação, em sua 'árvore logarítmica'.

Figura 6 - Árvore logarítmica representando wavelets
Fonte: Strang e Nguyen (1996)

O filtro passa-baixas C calcula as médias e realiza novamente a subamostragem $(\downarrow 2)$. Essas médias não estão na saída final, porque elas serão filtradas outra vez por C e D . O filtro C remove as baixas frequências ($\omega = 0$), enquanto o D remove as altas frequências ($\omega = \pi$). Juntos separam o sinal em bandas de frequências (STRANG; NGUYEN, 1996).

O banco de análise tem dois passos, filtragem e *downsampling*, e o banco síntese também tem dois: filtragem e *upsampling*. A operação de *downsampling* não é reversível. Abaixo um exemplo para melhor compreensão de $(\downarrow 2)$ e $(\uparrow 2)$.

$$y = \begin{bmatrix} \cdot \\ y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \\ \cdot \end{bmatrix} \quad (\downarrow 2)y = \begin{bmatrix} \cdot \\ y(0) \\ y(2) \\ y(4) \\ \cdot \end{bmatrix} \quad (\uparrow 2)(\downarrow 2)y = \begin{bmatrix} \cdot \\ y(0) \\ 0 \\ y(2) \\ 0 \\ y(4) \\ \cdot \end{bmatrix}$$

Existe uma transformada *wavelet* com características especiais, conhecida como *wavelet-packet*, cujo método de decomposição do sinal é similar ao do algoritmo de multiresolução. A diferença é que na decomposição *wavelet-packet* os coeficientes de aproximações e detalhes são decompostos em todos os níveis.

3.5 Exemplos de *Wavelet*

Existem muitos tipos de bases para cálculo da transformada *wavelet*. Serão citadas apenas duas: a mais simples de todas que é a de Haar, e uma das mais utilizadas que é a de Daubechies. A *wavelet* de Haar é demonstrada somente como contribuição teórica por possuir uma construção matemática mais simples que as demais, no entanto não foi utilizada neste trabalho.

3.5.1 *Wavelet* de Haar

A *wavelet* de Haar é o exemplo mais simples de uma *wavelet* ortonormal. Sua equação *scaling* contém somente dois coeficientes não-zeros e é dada por:

$$\phi(t) = \phi(2t) + \phi(2t - 1), \quad (18)$$

onde seus coeficientes *scaling* são $c_0 = c_1 = 1$. Os valores dos coeficientes são obtidos resolvendo as equações (13) e (14) simultaneamente. A equação correspondente da *wavelet* de Haar é:

$$\psi(t) = \psi(2t) + \psi(2t-1) \quad (19)$$

A *wavelet* de Haar é mostrada na figura abaixo.

Figura 7 - (a) Função *scaling* de Haar e (b) função *wavelet* de Haar, ambas com versões de deslocamento e dilatação

Fonte: Modificada de Addison (2002)

3.5.2 Wavelet de Daubechies

Os coeficientes são ordenados em duas sequências distintas; uma atua como um filtro de aproximação para os dados e a outra extrai os detalhes do sinal em cada escala. Daubechies é uma família de *wavelets* discretas na qual a *wavelet* de Haar é o membro mais simples. As funções *scaling* associadas com essas *wavelets* satisfazem as condições para serem ortogonais. Daubechies requer que suas *wavelets* tenham suporte compacto e sejam planas em algum grau. O nivelamento da *wavelet* é associado com a condição de momento, que pode ser expressa em termos de coeficientes *scaling*:

$$\sum_{k=0}^{N_k-1} (-1)^k c_k k^m = 0, \quad (20)$$

para inteiros $m = 0, 1, 2, \dots, N_k/2-1$. Essas *wavelets* têm $N_k/2$ momentos nulos. Alguns exemplos de *wavelets* de Daubechies podem ser vistos na figura 8. O comprimento do suporte das *wavelets* de Daubechies são $N_k/2-1$, por exemplo, a D2 (Haar), como é conhecida, tem suporte de comprimento igual a 1, a D4 tem suporte de comprimento 3, a D6 suporte 5 e assim por diante. O 'D' representa a família particular de *wavelets* de Daubechies e o número associado representa o número de coeficientes *scaling* não nulos, N_k . Pode-se ver na figura 8 que a função *scaling* permite a passagem de baixas frequências, atuando assim como um filtro passa-baixas, e a função *wavelet* associada permite a passagem de altas frequências, funcionando assim como um filtro passa-altas. A resposta em frequência dessa família de filtros é maximamente plana em $\omega = \pi$ e $\omega = 0$.

Figura 8 - Forma de onda das funções *scaling* e *wavelet* para *wavelets* de suporte compacto de Daubechies, para $N = 2, 3, 5, 7$ e 9 .

Fonte: Daubechies (1992)

CAPÍTULO 4

CARACTERÍSTICAS NUMÉRICAS DO SINAL DA DEGLUTIÇÃO

Existem muitos trabalhos com enfoques diferentes sobre análise do sinal da deglutição. Alguns trabalhos empregam diferentes técnicas para classificar entre deglutições normais e disfágicas, outros tentam identificar e localizar eventos no sinal acústico, e inúmeros outros tipos de estudos referentes ao sinal da deglutição são encontrados na literatura. Não existem parâmetros específicos definidos para analisar o sinal da deglutição. Serão mencionados a seguir alguns estudos recentes e relevantes sobre o tema em questão.

Cichero e Murdoch (2002a), em seu estudo, analisaram cinco variáveis: duração do som da deglutição, duração do pico presente no sinal da deglutição, intensidade do sinal, faixa de frequência do sinal e intensidade espectral da região da deglutição. Para isso variaram volume do bolo, idade e sexo e como resultado final encontraram a média do som da deglutição de aproximadamente 0,4 segundo, com intensidade de 43 dB e faixa de frequência de 2200 Hz. A intensidade de pico ocorre aproximadamente 0,19 segundo depois do início do registro da deglutição atingindo uma frequência de aproximadamente 550 Hz.

Lazareck e Moussavi (2004a, 2004b), pretenderam em ambos os estudos classificar deglutições entre normais e disfágicas, utilizando como parâmetro de classificação a dimensão de onda baseada em dimensão fractal. Em seus estudos foram analisados três tipos de alimentos oferecidos a 26 pessoas (15 controles e 11

pacientes). Os controles foram 12 crianças (3-16 anos) e 3 adultos (35, 38 e 54 anos). Os tipos de alimentos oferecidos foram: líquido, líquido engrossado e pastoso. O melhor resultado obtido na classificação foi com o líquido engrossado.

Também utilizando dimensão fractal, só que com o propósito de identificar e separar o som da deglutição e o som da respiração, Aboofazeli e Moussavi (2004), avaliaram 7 sujeitos saudáveis e, extraíndo as características do sinal da deglutição e aplicando-as em uma rede neural, obtiveram uma precisão de 91,7%.

Moussavi (2005), utilizou dimensão fractal e modelos ocultos de Markov para identificar os estágios da deglutição. Para o modelo foram assumidos 3 estados similares às fases oral, faríngea e esofágica. Foram utilizados sinais de deglutições de 4 crianças (5, 8, 10 e 12 anos). O mais interessante desse estudo foi que, de muitos estados assumidos inicialmente, o resultado final chegou a 3 estados para todos os sinais entre diferentes sujeitos. Outro resultado relevante é que o clique final nunca foi detectado como sendo um estado diferente.

Aboofazeli e Moussavi (2006) estudaram o emprego de filtro baseado em *wavelet* para separar o sinal da deglutição do sinal da respiração. Para esse fim, foi criado um algoritmo iterativo aplicando limiar para separar os sinais. O método proposto foi capaz de identificar corretamente 93% dos segmentos do sinal da deglutição.

Outro estudo usando novamente dimensão fractal, só que desta vez com o propósito de identificar fases no sinal. Foram definidas 3 partes do sinal da deglutição: som discreto inicial, som da transição do bolo e sinal da deglutição completa. Utilizaram para a pesquisa os sinais das deglutições de 6 crianças (3-16 anos) e 2 adultos (35 e 54 anos). Não conseguiram obter bons resultados na identificação do som inicial e na transição do bolo (YADOLLAHI; MOUSSAVI, 2007).

Spadotto et al. (2008a) utilizaram filtro baseado em *wavelet* para retirar ruídos e permitir uma melhor análise subjetiva do avaliador, quando feita pela audição do sinal. Foram empregados na pesquisa sinais de deglutições de 10 indivíduos saudáveis. Foi encontrado um aumento significativo da relação sinal-ruído.

Em outro estudo, Spadotto et al. (2009) extraíram características do sinal da deglutição com emprego de transformada *wavelet*; tais parâmetros representavam a energia do sinal contida em diversas faixas de frequência, previamente definidas. Depois foram aplicados os valores dos parâmetros obtidos em um classificador *Bayesiano* com o propósito de diferenciar entre deglutição normal e alterada. Foram avaliados 20 sujeitos controle e 20 pacientes, resultando em uma taxa de acerto de 76%.

Novamente utilizando transformada *wavelet* e a energia do sinal contida em faixas de frequências definidas, Spadotto et al. (2008b), selecionaram desta vez dois momentos distintos e específicos do sinal da deglutição, o primeiro sendo o momento do pico máximo e, outro, o momento de apneia da deglutição. Foram analisados sinais de 20 sujeitos controle e 20 pacientes. Os parâmetros extraídos serviram a um classificador chamado de *Optimum Path Forest*. A precisão obtida foi de 85,2% para o momento de pico e de 80,2% para o intervalo da apneia da deglutição.

4.1 Definição de parâmetros

Os parâmetros escolhidos para este estudo são listados na tabela 2, e logo após todos são explicados e definidos matematicamente. Foram escolhidos vários parâmetros que empregam técnicas diversas de processamento de sinais,

com o propósito de englobar diferentes características do sinal sobre aspectos variados.

Tabela 2 - Parâmetros numéricos para análise do sinal

Parâmetro	Símbolo
Média	μ
Variância	σ^2
Potência	P
Taxa de Cruzamento por Zero	ZCR
Entropia de Shannon	ES
Dimensão Fractal	DF
Frequência do Pico de Energia	f0
Potência do Envelope	PLPC
Estacionaridade	ZA
Taxa de Dispersão	DR

A média representada pela letra grega μ , equivale ao valor médio do sinal. É calculada somando todos os valores das amostras e dividindo pelo comprimento (SMITH, 1999). Sua fórmula matemática é a seguinte:

$$\mu = \frac{1}{N} \sum_{i=0}^{N-1} x_i \quad (21)$$

A variância mede a potência da flutuação do sinal em relação à média do mesmo (SMITH, 1999). É definida como:

$$\sigma^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} (x_i - \mu)^2 \quad (22)$$

A potência do sinal é um parâmetro dependente da energia por uma unidade de tempo (ADDISON, 2002). A energia por sua vez é definida como:

$$E = \sum_{i=0}^{N-1} (x_i)^2 \quad (23)$$

A potência de um sinal discreto é calculada pela equação a seguir:

$$P = \frac{\sum_{i=0}^{N-1} (x_i)^2}{N} \quad (24)$$

A taxa de cruzamento por zero, conta o número de vezes que o sinal passa pelo eixo de origem (inversão do sinal), geralmente indicando a presença de altas frequências, quando representada por um valor alto (CHEN, 1988). Definida como:

$$ZCR = \sum_{m=0}^N |(inversões)w(N-m)| \quad (25)$$

Definindo com um exemplo a entropia, pode-se dizer que é um parâmetro estatístico, o qual, de certo modo, mede a quantidade de informação produzida em média para cada letra de um texto na linguagem específica. Se a linguagem é traduzida com uma boa representação em dígitos binários (0 ou 1), a entropia S é o número médio de dígitos binários necessários por letra do idioma original (SHANNON, 1951). Um método utilizado para se calcular a entropia é usar séries de aproximação, como a mostrada a seguir:

$$S(p) = -\sum_i p_i \log(p_i) \quad (26)$$

A geometria fractal utiliza o conceito de auto-similaridade, no qual uma pequena parte de um objeto apresenta similaridades ao objeto como um todo. Recentemente o estudo fractal começou a ser desenvolvido matematicamente e aplicado em vários ramos da ciência e engenharia. A medida mais comum associada com a auto-similaridade é a dimensão fractal, que para sinais é um número compreendido ente 1 e 2 (AL-AKAIDI, 2004). A dimensão fractal é calculada da seguinte forma:

$$DF = 2 - |H| \quad (27)$$

Sendo que o H é conhecido como expoente de Hurst e obtido como:

$$H = \frac{\left(\frac{n \sum_{i=0}^{n-1} iP_i - \sum_{i=0}^{n-1} P_i \sum_{i=0}^{n-1} i}{n \sum_{i=0}^{n-1} i^2 - \left(\sum_{i=0}^{n-1} i \right)^2} \right) - 1}{2} \quad (28)$$

onde n representa a quantidade de amostras do sinal, sendo uma potência de 2 e P equivale ao logaritmo da energia dos elementos da transformada *wavelet-packet* em nível máximo.

Foi definido f_0 , como sendo a frequência que representa o maior valor de energia da transformada *wavelet-packet* em nível máximo. O primeiro passo para se obter este parâmetro é calcular a transformada *wavelet-packet* em nível máximo do sinal desejado, depois se localiza a folha que possui o valor de maior energia e se verifica qual faixa de frequência essa folha representa, e por último se calcula a média das frequências inicial e final. Essas frequências são calculadas da seguinte forma:

Frequência inicial:

$$\frac{k * TA}{2^{j+1}} \quad (29)$$

Frequência final:

$$\frac{(k + 1) * TA}{2^{j+1}} \quad (30)$$

onde TA equivale à taxa de amostragem do sinal, j representa o nível da transformada *wavelet-packet* e k , o número da folha da árvore no último nível da transformada, iniciando por zero.

Outro parâmetro estabelecido é o PLPC que representa a potência calculada sobre o envelope do espectro de potência do sinal. O envelope do espectro de potência é obtido pelo método do LPC (*Linear Predictive Coding*), o qual é calculado utilizando-se da combinação linear das amostras passadas de um sinal para prever as próximas amostras (DENG; O'SHAUGHNESSY, 2003). É representado da seguinte forma:

$$s[n] \approx \hat{s}[n] = -\sum_{i=1}^M a_i s[n-i], \quad (31)$$

onde $\hat{s}[n]$ é chamada de amostra estimada, a_i são chamados de coeficientes LPC e o M representa a ordem do LPC, que neste estudo foi de 27.

A estacionaridade implica que a média e a variância do sinal não se alteram ao longo do tempo. O *reverse arrangements test* é um teste simples não-paramétrico para a estacionaridade e tem sido bastante empregado em sinais fisiológicos. O valor do *reverse arrangements* (A) é obtido contando seu número da sequência $\{y_i, i = 1, \dots, M\}$, quando ocorre $y_i > y_j$ para $i < j$ (CHAU et al., 2005; LEE et al., 2006a; LEE et al., 2006b). Para o cálculo da estacionaridade os dados foram normalizados.

Pode se calcular a estacionaridade da seguinte maneira:

$$ZA = \frac{A - \mu A}{\sigma A}, \quad (32)$$

onde μA e σA , são respectivamente média e desvio padrão, que neste caso foram considerados sendo: $\mu A = N(N-1)/4$ e $\sigma A = \sqrt{N(N-1)(2N+5)/72}$.

A taxa de dispersão é uma taxa obtida entre a média do desvio absoluto (MAD - *mean absolute deviation*) e a faixa interquartil do sinal. Pode ser interpretada

por captar a diferença entre estimativa de propagação não-robusta (média de desvio absoluto) e robusta (intervalo interquartilico). A taxa de dispersão representa de uma maneira aproximada a natureza e a multiplicidade de observações atípicas dentro do sinal. Na ausência de tais observações atípicas, a taxa tenderia a esse valor (LEE et al., 2006a; LEE et al., 2006b). A média do desvio absoluto pode ser calculada por:

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_i - med(x)| \quad (33)$$

onde n é o número total de amostras do sinal, x_i é a amostra i do sinal e $med(x)$ é a mediana do sinal. O intervalo do interquartil denotado por IQR, é definido como:

$$IQR = q_{0.75} - q_{0.25} \quad (34)$$

onde $q_{0.75} - q_{0.25}$ representam, respectivamente, o primeiro e o terceiro quartil da distribuição da amplitude do sinal. Finalmente a taxa de dispersão é expressa como:

$$DR = \frac{MAD}{IQR} \quad (35)$$

4.2 Definição de faixas de frequência para base de energia

Além da análise dos parâmetros definidos na secção anterior foi definido neste estudo um segundo método para análise do sinal da deglutição, o qual é baseado no cálculo da energia presente no sinal em intervalos específicos de frequência. Outros trabalhos empregaram a mesma técnica (SPADOTTO et al., 2008b, 2009). As faixas de frequência analisadas são mostradas na tabela 3, para cada um dos 3 intervalos definidos.

Tabela 3 - Intervalos de frequências para 10 intervalos

Faixa	Frequência (Hz)		
	Intervalo 1	Intervalo 2	Intervalo 3
F-1	0-59	0-100	0-299
F-2	59-118	100-199	299-599
F-3	118-178	199-299	599-898
F-4	178-237	299-398	898-1198
F-5	237-296	398-498	1198-1497
F-6	296-355	498-598	1497-1797
F-7	355-415	598-697	1797-2096
F-8	415-474	697-797	2096-2396
F-9	474-533	797-896	2396-2695
F-10	533-592	896-996	2695-2994

Primeiramente delimita-se o momento do sinal a ser analisado, em seguida se extrai a transformada *wavelet-packet* em nível máximo, localizam-se as folhas da árvore de decomposição que representam as frequências especificadas na tabela 3. De posse da localização das folhas que representam os inícios e fins de cada faixa de frequência, delimita-se o intervalo de cada um dos 10 sub-sinais desejados, calculando-se então a energia de cada um desses trechos. A energia de cada sub-banda é obtida de seguinte forma:

$$E = \sum_k w_j^2(k) \quad (36)$$

onde $w_j(k)$ corresponde aos coeficientes de aproximação ou detalhe em nível j (CHEN; GUIDO; CHEN, 2007).

Foi definido analisar a energia contida em um intervalo de no máximo 3000 Hz, pelo fato de muitos trabalhos da literatura relatarem que a maior concentração de energia do sinal da deglutição está abaixo de 3000 Hz, conforme descrito na secção 2.3.

Somente para efeito de comparações, outro teste também foi utilizado empregando a mesma técnica, mas com sobreposição (*overlap*) de 50% das faixas

de frequência. Essas diferentes faixas de frequência analisadas são mostradas a seguir:

Tabela 4 - Intervalos de frequências para 19 intervalos com sobreposição

Faixa	Frequência (Hz)
F-1	0-59
F-2	32-92
F-3	65-124
F-4	97-156
F-5	129-188
F-6	161-221
F-7	194-253
F-8	226-285
F-9	258-318
F-10	291-350
F-11	323-382
F-12	355-415
F-13	388-447
F-14	420-479
F-15	452-511
F-16	484-544
F-17	517-576
F-18	549-608
F-19	581-641

Esses 4 intervalos diferentes de faixas de frequências demonstrados nas tabelas 3 e 4, serão utilizados nos cálculos da energia contido dentro de cada faixa e posteriormente empregados como conjuntos de dados para os classificadores.

CAPÍTULO 5

CLASSIFICADOR

A classificação de padrões visa encontrar a classe de cada amostra que pode ser particionada em c classes (rótulos), usando um vetor de atributos. Os métodos normalmente aprendem como as amostras se distribuem em diferentes classes no espaço de atributos, usando um subconjunto de treinamento e encontrando superfícies ou regras de decisão para o particionamento deste conjunto (DUDA; HART; STORK, 2000).

Os métodos de classificação podem ser divididos de acordo com seus algoritmos de aprendizado em 3 categorias: supervisionados, não-supervisionados, e semi-supervisionados (HUANG; KECMAN; KOPRIVA, 2006). Métodos não-supervisionados não possuem ou não exploram conhecimento algum sobre as classes das amostras de treinamento, enquanto, em métodos supervisionados, esta informação é conhecida e explorada. Métodos semi-supervisionados conhecem apenas parte ou fazem uso parcial dos rótulos das amostras de treinamento.

Entre as abordagens supervisionadas, Redes Neurais Artificiais (*Artificial Neural Networks* - ANN) (HAYKIN, 1998) e Máquinas de Vetores de Suporte (*Support Vector Machines* - SVM) (BOSER; GUYON; VAPNIK, 1992) são as mais populares. O uso de ANN tem seu desempenho melhorado quando se combinam várias redes, porém gerando um maior custo computacional. A SVM busca uma separabilidade linear ótima em um espaço de maior dimensão. Como desvantagens,

essa premissa pode não ser válida com dimensão finita e o custo computacional do SVM aumenta muito rápido com o número de vetores de suporte. Neste trabalho foi empregada a rede neural baseada em perceptron multicamadas (*ANN Multilayer Perceptron*).

Técnicas de reconhecimento estatístico de padrões tentam determinar superfícies de decisão baseadas em distribuições de probabilidade que melhor representam cada classe (JAIN; DUIN; MAO, 2000), as quais podem ser conhecidas a priori ou necessitam ser estimadas. Se os rótulos forem conhecidos, então um classificador Bayesiano (*Bayesian Classifier* - BC) pode ser utilizado para a tarefa de classificação das amostras. Em situações nas quais não se tem informação alguma sobre a densidade das classes, tem-se um problema de decisão não paramétrico, e tais densidades deverão ser estimadas.

Os classificadores supervisionados baseados em florestas de caminhos ótimos (*Optimum Path Forest* - OPF) preenchem a lacuna dos métodos existentes, ou seja, que sejam capazes de tratar classes não-separáveis sem assumir forma ou modelo paramétrico para essas classes. Classificadores OPF baseiam-se em um subconjunto de amostras protótipos (representantes de todas as classes do conjunto de treinamento) e em uma função de custo de caminho no grafo de treinamento, a qual busca agrupar amostras com atributos similares. Tais classificadores particionam o grafo de treinamento em uma floresta de caminhos de custo mínimo, onde cada árvore é enraizada em um protótipo e todas as amostras da árvore são rotuladas com o mesmo rótulo da raiz. A classificação de uma nova amostra se dá por encontrar o protótipo que lhe oferece o caminho ótimo entre os caminhos oferecidos por todos os protótipos, isto é, a classificação é baseada na força de conexão entre a amostra e o protótipo mais conexo, e não na simples distância

entre eles, como faz, por exemplo, os classificadores baseados em k-vizinhos mais próximos (FUKUNAGA; NARENDRA, 1975).

Neste trabalho foi utilizada a abordagem supervisionada modelando o problema como sendo um grafo completo, onde o peso da aresta entre duas amostras é dado pela distância entre elas. Para essa metodologia, utiliza-se a função $f_{\max}$, a qual associa como custo de um caminho o valor do arco de maior peso ao longo dele. Essa abordagem trata o problema de classes não-separáveis estimando protótipos nas regiões de intersecção entre elas. Os elementos mais representativos de cada classe do conjunto de treinamento, isto é, os protótipos, são escolhidos como sendo os elementos pertencentes às regiões de fronteira entre as classes. Os protótipos participam de um processo de competição disputando as outras amostras oferecendo-lhes caminhos de menor custo e seus respectivos rótulos. Ao final desse processo, obtemos um conjunto de treinamento particionado em árvores de caminhos ótimos, sendo que a união das mesmas nos remete a uma floresta de caminhos ótimos. Essa abordagem apresenta vários benefícios com relação a outros métodos de classificação de padrões supervisionados: é livre de parâmetros, possui tratamento nativo de problemas multiclases e não faz alusão a forma e/ou separabilidade das classes.

O algoritmo OPF com grafo completo foi primeiramente apresentado por (PAPA et al., 2007, 2008) e tem sido amplamente utilizado em diversas aplicações, tais como avaliação de descritores de textura, diagnóstico automático de patologias na laringe e classificação de impressões digitais.

5.1 Fundamentação teórica do OPF

Seja Z uma base de dados λ -rotulada e Z_1 e Z_2 os conjuntos de treinamento e teste, respectivamente, com $|Z_1|$ e $|Z_2|$ amostras, as quais podem

ser pixels/voxels/contornos tais que $Z = Z_1 \cup Z_2$. Seja $\lambda(s)$ uma função que associa o rótulo correto i , $i = 1, 2, \dots, c$ da classe i a qualquer amostra $s \in Z_1 \cup Z_2$.

Seja $S \in Z_1$ um conjunto de protótipos de todas as classes. Seja $\vec{v}$ um algoritmo que extrai n atributos de qualquer amostra $s \in Z_1 \cup Z_2$, e retorna um vetor de atributos $\vec{v}(s) \in \mathfrak{R}^n$. A distância $d(s, t)$ entre duas amostras, s e t , é dada pela distância entre seus vetores de atributos $\vec{v}(s)$ e $\vec{v}(t)$.

O classificador consiste em usar S , $(\vec{v}, d)$ e Z_1 criando uma partição discreta ótima, a qual é uma floresta de caminhos ótimos computada em $\mathfrak{R}^n$ pelo algoritmo da transformada imagem floresta (FALCÃO; STOLFI; LOTUFO, 2004).

Seja (Z_1, A) um grafo completo cujos nós são as amostras em Z_1 , onde qualquer par de amostras define um arco em A (isto é, $A = Z_1 \times Z_1$) (Figura 9a). Os arcos não precisam ser armazenados e o grafo não precisa ser explicitamente representado.

Figura 9: (a) Grafo completo ponderado nas arestas. (b) MST do grafo completo. (c) Protótipos escolhidos como sendo os elementos adjacentes de classes diferentes na MST (nós circutados). (d) Floresta de caminhos ótimos resultante para a função de valor de caminho f_{max} e dois protótipos. (e) Uma amostra de teste (triângulo) da classe 2 e suas conexões (linhas pontilhadas) com os nós do conjunto de treinamento. (f) O caminho ótimo do protótipo mais fortemente conexo, seu rótulo 2 e o custo de classificação 0.4 são associados a amostra de teste. Nota-se que, mesmo a amostra de teste estando mais próxima de um nó da classe 1, ela foi classificada como sendo da classe 2

O algoritmo baseado em OPF pode ser utilizado com qualquer função de custo suave que possa agrupar amostras com propriedades similares (FALCÃO; STOLFI; LOTUFO, 2004). Na versão OPF com grafo completo a função de custo abordada foi a f_{max} , que é dada como:

$$f_{max}(\langle s \rangle) = \begin{cases} 0 & \text{se } s \in S, \\ +\infty & \text{caso contrário} \end{cases} \quad (37)$$

$$f_{max}(\pi_s \cdot \langle s, t \rangle) = \max\{f_{max}(\pi_s), d(s, t)\},$$

O algoritmo baseado em OPF associa um caminho ótimo $P^*(s)$ de S a toda amostra $s \in Z_1$, formando uma floresta de caminhos ótimos P (uma função sem ciclos, a qual associa a todo $s \in Z_1$ seu predecessor $P(s)$ em $P^*(s)$, ou uma marca *nil* quando $s \in S$), como mostrado na figura (9d). Seja $R(s) \in S$ a raiz de $P^*(s)$ a qual pode ser alcançada por $P(s)$. O algoritmo computa, para cada $s \in Z_1$, o custo $V(s)$ de $P^*(s)$, o rótulo $L(s) = \lambda(R(s))$ e o seu predecessor $P(s)$, como segue.

Algoritmo 1 - Classificador supervisionado baseado em floresta de caminhos ótimos usando grafo completo

Entrada: Um conjunto de treinamento Z_1 λ -rotulado, protótipos $S \subset Z_1$ e o par (v, d) para vetor de atributos e cálculo das distâncias.

Saída: Floresta de caminhos ótimos P , mapa de valores de custo de caminhos V e mapa de rótulos L .

Auxiliares: Fila de prioridades Q , e variável cst .

1. **Para todo** $s \in Z_1$, **Faça** $P(s) \text{ nil}$ e $V(s) + \infty$.
2. **Para todo** $s \in S$, **Faça** $V(s) 0$, $P(s) \text{ nil}$, $L(s) = \lambda(s)$ e insira s em Q .
3. **Enquanto** Q é não vazia, **Faça**
 4. Remova de Q uma amostra s tal que $V(s)$ é mínimo.
 5. **Para cada** $t \in Z_1$ tal que $s \neq t$ e $V(t) > V(s)$, **Faça**
 6. Calcule $cst \max\{V(s), d(s, t)\}$.
 7. **Se** $cst < V(t)$, **Então**
 8. **Se** $V(t) \neq +\infty$, **Então** remova t de Q .
 9. $P(t)s$, $L(t)L(s)$ e $V(t)cst$.
 10. Insira t em Q .

As linhas 1-2 inicializam os mapas e inserem protótipos em Q . O laço principal calcula um caminho ótimo de S para cada amostra $s \in Z_1$ em uma ordem não decrescente de custos (linhas 3-10). A cada iteração um caminho de custo de ótimo $V(s)$ é obtido em P (linha 4). Empates são resolvidos em Q utilizando a política FIFO (*first-in-first-out*), ou seja, quando dois caminhos atingem uma determinada amostra s com o mesmo custo mínimo, s é associado ao primeiro caminho que o atingiu. O restante das linhas avalia se o caminho que atinge uma amostra adjacente t através de s é mais barato que o caminho que termina em t . Caso positivo, atualiza Q , $P(t)$, $L(t)$ e $V(t)$. No final do algoritmo, V armazena o valor do custo do caminho ótimo de S a cada amostra $s \in Z_1$ de acordo com f_{max} .

5.2 Treinamento

A fase de treinamento do classificador baseado em floresta de caminhos ótimos usando o grafo completo consiste, basicamente, em encontrar o conjunto S de protótipos, ou seja, os elementos mais representativos de cada classe. Várias heurísticas poderiam ser adotadas como, por exemplo, uma escolha aleatória de protótipos. Entretanto, tal escolha pode prejudicar o desempenho do classificador, tornando-o instável e com um alto grau de sensibilidade com relação aos protótipos escolhidos. Deseja-se, assim, estimar protótipos nas regiões de sobreposição de amostras e nas fronteiras entre as classes, visto que são regiões muito susceptíveis a erros de classificação.

Computando uma árvore geradora mínima (*Minimum Spanning Tree* - MST) no grafo completo (Z_1, A) , obtém-se um grafo conexo acíclico cujos nós são todas as amostras em Z_1 , e os arcos são não direcionados e ponderados (figura 9b). Seus

pesos são dados pela distância d entre os vetores de atributos de amostras adjacentes. Essa árvore de espalhamento é ótima no sentido em que a soma dos pesos de seus arcos é mínima se comparada a outras árvores de espalhamento no grafo completo. Os protótipos a serem escolhidos são os elementos conectados na MST com diferentes rótulos em Z_1 , isto é, elementos próximos de classes diferentes (figura 9c). Removendo-se os arcos entre classes diferentes, tais amostras adjacentes tornam-se protótipos em S e o Algoritmo 1 pode computar uma floresta de caminhos ótimos em Z_1 (figura 9d). Pode-se verificar que uma dada classe pode ser representada por múltiplos protótipos (isto é, árvores de caminhos ótimos) e deve existir pelo menos um protótipo por classe.

O Algoritmo 1 pode computar uma floresta de caminhos ótimos com erro zero de classificação em Z_1 , desde que a função f_{max} seja modificada. A ideia consiste, basicamente, em ponderar os arcos entre amostras de diferentes classes com um valor muito alto, impossibilitando assim que protótipos de uma classe conquistem elementos de outras classes. Desse modo, a função de valor de caminho f_{max} pode ser escrita da seguinte forma:

$$f_{max}(\langle t \rangle) = \begin{cases} 0 & \text{se } t \in S \\ +\infty & \text{caso contrário} \end{cases} \quad (38)$$

$$f_{max}(\pi_s \cdot \langle s, t \rangle) = \begin{cases} +\infty & \text{se } \lambda(t) \neq \lambda(s) \\ \max\{f_{max}(\pi_s), d(s, t)\} & \text{caso contrário.} \end{cases}$$

A equação 38 pondera todos os arcos $(s, t) \in A_k$ tais que $\lambda(t) \neq \lambda(s)$ com $d(s, t) = +\infty$, impedindo que tais arcos pertençam a algum caminho ótimo. Entretanto, vale destacar que, embora garanta erro zero de classificação no conjunto de treinamento, a equação 38 produz resultados similares aos obtidos usando a equação 37.

5.3 Classificação

Para qualquer amostra $t \in Z_2$, consideram-se todos os arcos conectando t com amostras $s \in Z_1$, tornando t como se fosse parte do grafo original (ver figura 9e, onde a amostra t é representada pelo triângulo no grafo). Considerando todos os possíveis caminhos entre S e t , desejamos encontrar o caminho ótimo $P^*(t)$ de S até t com a classe $\lambda(R(t))$ de seu protótipo $R(t) \in S$ mais fortemente conexo. Esse caminho pode ser identificado incrementalmente, avaliando o valor do custo ótimo $V(t)$ como:

$$V(t) = \min\{\max\{V(s), d(s, t)\}\}, \forall s \in Z_1. \quad (39)$$

Seja $s^* \in Z_1$ o nó que satisfaz a equação acima, isto é, o predecessor $P(t)$ no caminho ótimo $P^*(t)$. Dado que $L(s^*) = \lambda(R(t))$, a classificação simplesmente associa $L(s^*)$ como a classe de t (figura 9f). Um erro ocorre quando $L(s^*) \neq \lambda(t)$. Nesse caso, entretanto, deseja-se usar as amostras em Z_2 com o intuito de aprender a distribuição das classes no espaço de atributos e, consequentemente, melhorar o desempenho de classificadores baseados em OPF.

Vale ressaltar que, embora esteja mais próxima de um elemento da classe bola (figura 9e), a amostra a ser classificada é classificada como sendo da classe quadrado, o que demonstra que os classificadores baseados em OPF utilizam a força de conexidade entre as amostras para a classificação dos dados, ou seja, não são algoritmos baseados em conexidade local.

CAPÍTULO 6

MÉTODOS

Na fase de conhecimento atual, o sinal da deglutição tem sua aplicabilidade limitada, devido a vários fatores. Um deles é o fato de ser difícil identificar fases do sinal, correlacionando-as com eventos fisiológicos envolvidos na dinâmica da deglutição. Nesse estudo, antes de se fazer a análise do sinal, foi feita uma análise detalhada da videofluoroscopia da deglutição de cada indivíduo avaliado, como será detalhado um pouco mais a frente.

Foram avaliados nesta pesquisa 14 sujeitos saudáveis sem alteração do processo de deglutição, com idade entre 20 e 50 anos. Os sujeitos com queixas de dificuldades prévias ou atuais de deglutição foram excluídos de nossa coleta. Todos os sujeitos tiveram ciência do protocolo e assinaram o termo de consentimento livre e esclarecido.

A coleta de exames constou da captura da imagem de videofluoroscopia simultânea ao som da deglutição. Ambos os sinais foram digitalizados e armazenados em um gravador de DVD, marca Philips, modelo DVDR3455H/78. A qualidade de digitalização foi de 29,97 quadros por segundo para o vídeo, e de 44.100 amostras por segundo para o sinal, que na edição era convertida para 22.050 Hz. O sinal de vídeo provia de um instrumento radiológico Arcomax Phillips (Philips Medical Systems) modelo BV 300 (Veenpluis, The Netherlands).

O sinal de áudio era capturado por um microfone dinâmico colocado na região da cartilagem cricóide (borda lateral inferior) local considerado ótimo, segundo Takahashi, Groher e Michi (1994). O microfone era conectado a uma mesa de som, marca Wattsom, para pré-amplificação e controle de ganho. A saída de áudio da mesa de som era conectada diretamente ao gravador.

Muitos outros métodos de captura foram testados previamente, todos apresentando baixa qualidade no resultado final. Utilizou-se microfone de eletreto juntamente com um pré-amplificador desenvolvido para o propósito da pesquisa. Com esse tipo de microfone o sinal apresentava muita contaminação por sons do ambiente em volta do local onde foi realizada a captura. Na tentativa de coletar somente o sinal, também se tentou capturar o som com microfone de eletreto conectado diretamente à placa de som do micro-computador. O problema desse método é que os resultados variavam muito entre computadores devido aos vários modelos de placas de som e seus diferentes controles de ganho, e o sinal era facilmente saturado. Devido aos diversos testes efetuados, muitos exames foram desperdiçados no decorrer do trabalho, tanto exames simultâneos com a videofluoroscopia quanto somente os sinais de áudio capturados diretamente no computador. Esses testes com captura somente do sinal foram aprovados pelo Comitê de Ética e Pesquisa da Faculdade de Medicina de Botucatu, conforme o Anexo 1

Para se realizar a coleta foi firmada uma parceria com a Faculdade de Medicina de Botucatu e com o Hospital das Clínicas da USP de Ribeirão Preto para utilização do equipamento de RX, o qual fornecia o sinal de vídeo para gravação em equipamento externo. O projeto de pesquisa foi aprovado pelo Comitê de Ética e Pesquisa da USP de Ribeirão Preto, conforme o Anexo 2.

Cada indivíduo foi observado durante 6 deglutições sendo 3 de consistência pastosa, oferecidas em colher, com 7 mL cada, e 3 de substância líquida, oferecidas em copo descartável contendo 10 mL cada, totalizando 84 deglutições. Para posterior análise dividimos as 84 deglutições em dois grupos de 42, sendo um de deglutições de alimento pastoso e outro de líquido. Todos esses exames foram acompanhados e posteriormente avaliados por uma fonoaudióloga com habilitação em videofluoroscopia.

Na preparação das consistências foram utilizados os seguintes materiais: copo plástico descartável, seringa descartável de 20 mL, colher de plástico descartável, água na temperatura natural, espessante de alimento e bário. A consistência líquida consistia de bário líquido (sulfato de bário 100%, Bariogel, Laboratório Cristália, Itapira, SP, Brasil) e a consistência pastosa foi preparada com 50 mL de bário líquido acrescido de 4,5 g de espessante (*Thick & Easy, Hormel Health Labs, Austin, MN, USA*).

Para explicar melhor a metodologia empregada, será inicialmente discutida como é feita a análise digital da videofluoroscopia. Spadotto et al. (2008c) desenvolveram um *software* para análise quantitativa da videofluoroscopia. O *software* permite analisar o vídeo quadro-a-quadro e delimitar as fases, obtendo assim o tempo da fase em exame. A subjetividade ainda permanece, mas o emprego do *software* melhora e muito a eficiência e precisão na marcação de tempo.

A precisão torna-se importante quando o objetivo é verificar a eficiência de procedimentos terapêuticos dentro do processo de reabilitação do indivíduo disfágico. O acompanhamento da reabilitação é mais eficiente com a utilização do *software*. Quando se analisa a deglutição, as diferenças de tempo em cada fase são

de pequenas dimensões, impróprias para análise visual e mecânica. A figura 10 exibe a área de trabalho do *software*.

Figura 10 - Área de trabalho do *software* de análise da deglutição
Fonte: Spadotto et al. (2008c)

Com base nos moldes da pesquisa demonstrada acima, foi elaborada a ideia de analisar o sinal da deglutição, mas em momentos específicos. Tais momentos eram delimitados através de uma minuciosa avaliação da imagem da videofluoroscopia da deglutição de cada uma das 6 deglutições de cada voluntário da pesquisa.

Para tal fim foi desenvolvido um *software* que, além das ferramentas descritas anteriormente, possui também ferramentas para visualização da forma de onda do sinal. O *software* funciona de uma maneira integrada, sincronizadamente para o vídeo e áudio de cada deglutição. A cada avanço de quadro do vídeo é possível verificar o trecho do sinal e a forma de onda correspondente naquele momento específico.

Foram estabelecidos alguns momentos para uma posterior análise do sinal juntamente da videofluoroscopia e com o auxílio desse novo *software* foi possível delimitá-los. Os momentos analisados são mostrados na tabela a seguir:

Tabela 5 - Momentos analisados

Momentos	Variáveis
Fase Faríngea	Tempo
<i>Burst 1</i>	Tempo Posição do Bolo
<i>Burst 2</i>	Tempo Posição do Bolo
<i>Burst 3</i>	Posição do Bolo
Máxima elevação de Hióide	
Intervalo entre explosões 1 e 2	
Intervalo entre explosões 2 e 3	

Na análise do tempo de trânsito faríngeo, por meio desse *software*, utilizamos a definição proposta por Kendall et al. (2000) e Kendall; Leonard e Mckenzie (2001), considerando como início da fase faríngea da deglutição o momento em que o bolo alimentar atingia a região posterior da espinha nasal, localizada no final do palato duro e início do palato mole. Foi considerado como término da fase faríngea da deglutição o momento em que a parte distal do bolo alimentar passava pelo esfíncter superior do esôfago.

Na figura 11, tem-se uma amostra da visualização de uma imagem e de um sinal, aberta pelo programa desenvolvido para analisar ambos, simultaneamente. A imagem mostra o sinal do exame em seu tempo total e também o trecho específico do quadro em questão.

Ainda nesse programa, foi desenvolvida uma ferramenta que permite uma análise visual com o intuito de comparar a morfologia da onda de cada deglutição entre indivíduos. Delimitava-se o trecho que se desejava comparar, que nesse caso foi o tempo de duração da fase faríngea. Assim, era determinado o número de amostras da deglutição que teve uma maior fase faríngea e selecionava o mesmo número de amostras das outras deglutições, a partir do início especificado. Caso o sinal chegasse ao final era completado com zeros, apenas para se ter o mesmo número de amostras em todos os sinais, permitindo assim uma relação temporal entre os mesmos. Também se padronizou a amplitude fazendo-se a normalização pelo maior valor presente entre todos os sinais. Na figura 12 pode-se verificar que algumas deglutições se assemelham e outras se diferenciam bastante, relacionadas à forma de onda dos seus sinais.

Figura 12 - Comparação da forma de onda dos sinais da deglutição entre indivíduos

CAPÍTULO 7

RESULTADOS

Foram utilizados dois métodos diferentes para análise dos dados obtidos dos sinais. Inicialmente os dados foram analisados com ferramentas estatísticas convencionais, com o intuito de verificar diferenças entre os momentos e os grupos, com as 10 variáveis definidas no Capítulo 4 e também com as faixas de frequências determinadas. Depois foram utilizadas técnicas mais avançadas, como: redes neurais artificiais, máquina de vetores de suporte, classificador *Bayesiano* e o classificador OPF.

Antes de empregar os vetores de características relacionados na secção 7.2 como entrada de dados nos classificadores citados anteriormente, os vetores de características foram submetidos a um teste estatístico de comparação de médias (teste t) com intuito de se detectarem diferenças estatísticas significativas ($p < 0,05$) para cada característica isolada dentro de cada vetor, de características em relação aos grupos analisados (líquida e pastosa). Foi excluído o vetor de características que representa as 4 variáveis seleccionadas na secção 7.1, o qual levou em conta não só os grupos, mas também os momentos.

Tabela 6 - Variáveis com diferença estatística entre os grupos

Vetor	Momento	Variáveis									
10 Par	Burst 1	PLPC									
	Burst 2	μ	f0	PLPC	DR						
	FF	ES	PLPC	ZA							
Energia 600	Burst 1	F-1		F-6	F-10						
	Burst 2			F-6							
	FF		F-5	F-6							
Energia 600 over	Burst 1				F-10				F-17	F-19	
	Burst 2					F-11	F-12	F-13			
	FF	F-7	F-8	F-9	F-10	F-11	F-12				
Energia 1000	Burst 1			F-6	F-7	F-8	F-9	F-10			
	Burst 2		F-4								
	FF	F-3	F-4		F-7	F-8	F-9	F-10			
Energia 3000	Burst 1	F-3	F-4	F-5							
	Burst 2										
	FF	F-3	F-4	F-5	F-6						

A tabela 6 mostra quais variáveis que apresentaram diferença estatística ($p < 00,5$) entre os grupos (líquida e pastosa) para cada um dos 5 vetores de características em cada um dos momentos. O valor de cada faixa de frequência demonstrada na tabela pode ser obtido na seção 4.2. Embora nem todas as variáveis tenham apresentado significância entre os grupos, os vetores de características ainda assim foram aplicados como construídos originalmente, ou seja, sem a exclusão dessas variáveis isoladas. Por esse estudo ser pioneiro e empregar uma amostra pequena, não se pode aceitar ou rejeitar com convicção a análise estatística. Provavelmente, com uma grande amostra de dados e comparando com amostras de indivíduos patológicos, pode-se chegar a alguma definição por quais variáveis empregar na análise dos dados.

7.1 Estatística Convencional

Primeiramente foi verificada a correlação entre as 10 variáveis existentes. Foram calculadas as correlações nos 6 casos possíveis, ou seja, nos 3 momentos determinados para cada um dos dois grupos existentes. A tabela 7 mostra somente os valores da correlação considerando uma chance maior que 70%. Verifica-se que entre a variável P (potência) e a variável σ^2 (variância) uma alta correlação foi encontrada para os 3 momentos e na fase faríngea também houve correlação entre a variável DR (taxa de dispersão) com a ZCR (taxa de cruzamento por zero).

Tabela 7 - Correlação com mais 70% entre os parâmetros para os grupos e momentos

Tipo	Variáveis	Momento			
		<i>Burst 1</i>	<i>Burst 2</i>	Fase Faríngea	
		P- σ^2	P- σ^2	P- σ^2	DR-ZCR
	Líquida	0,976	0,961	0,939	
	Pastosa	0,938	0,963	0,977	0,742

Foi aplicado o teste de comparação de médias (teste de Tukey) para verificar se havia diferenças estatísticas significativas ($p < 0,05$) em relação aos 10 parâmetros, entre os 3 momentos definidos para estudo que eram: o *Burst 1*, o *Burst 2* e a Fase Faríngea e entre os dois grupos existentes que consistiam, o primeiro nas deglutições de substância líquida, e o segundo na substância pastosa. A figura 13 mostra onde houve níveis consideráveis de significância; os cones indicam que o nível de significância é igual ou menor ao estipulado, para as 10 variáveis em questão em relação aos grupos e a combinação dos 3 momentos, ou seja, *Burst 1* com *Burst 2*, *Burst 1* com fase faríngea e *Burst 2* com Fase Faríngea.

Figura 13 - Diferenças estatísticas ($p < 0,05$) entre momentos e grupos em relação aos parâmetros

Nota-se pela figura 13 que a variável ZA apresentou diferença estatística significativa para os 4 casos em questão, as variáveis PLPC, DF e P apresentaram em 3 casos, a variável σ^2 em 2 casos e as variáveis ES e ZCR em 1 caso. As variáveis DR, f_0 e μ não diferiram em nenhum caso.

Com base nesses resultados foram selecionados 4 parâmetros que mais vezes diferiram estatisticamente em relação às 4 comparações feitas, sendo 1 para os grupos (líquida e pastosa) e 3 para os momentos. Essas 4 variáveis selecionadas, que posteriormente serão utilizadas em um conjunto de dados para os classificadores, são: potência (P), dimensão fractal (DF), potência do envelope (PLPC) e estacionariedade (ZA).

Os tempos de duração computados para cada momento não apresentaram diferenças estatísticas entre os dois grupos. A tabela 8 e a figura 14 mostram valores estatísticos relacionados ao tempo para os 3 momentos. Pode-se notar que a deglutição de substância pastosa, na média, é um pouco mais demorada que a de substância líquida.

Tabela 8 - Duração dos tempos de deglutição

	Fase Faríngea		<i>Burst 1</i>		<i>Burst 2</i>	
	L	P	L	P	L	P
Média	659,63	693,40	87,27	78,20	112,93	85,50
Desvio Padrão	129,33	149,16	36,40	37,53	62,94	52,20
Mediana	633,97	633,97	66,73	66,73	100,10	66,73
Moda	533,87	600,60	66,73	66,73	66,73	33,37
Mínimo	500,50	467,13	33,37	33,37	33,37	33,37
Máximo	1034,37	1101,10	166,83	166,83	266,93	200,20

Figura 14 - Medidas estatísticas do tempo da deglutição

Foram computados também 3 intervalos de tempo entre o acontecimento de eventos fisiológicos e o do sinal, ocorridos durante a deglutição. Esses intervalos correspondem ao tempo em milissegundos entre: início de fase faríngea e máxima elevação do hióide (IFF-MH), *Burst 1* com *Burst 2* (1-2) e *Burst 2* com o início do *Burst 3* (2-3). Esses intervalos são demonstrados na figura 15.

Figura 15 - Intervalos entre eventos

A figura 16 mostra alguns valores estatísticos (média, mediana, mínimo e máximo) para as 4 variáveis que diferiram mais vezes considerando os grupos e momentos. Em (a) é exibido o valor do parâmetro P, em (b) o valor de DF, (c) o valor do PLPC e finalmente em (d) o valor de ZA.

Figura 16 - Medidas estatísticas dos 4 parâmetros que mais diferiram estatisticamente

Os próximos resultados referem-se à análise das energias contidas nas faixas de frequências pré-definidas na secção 4.2.

Na figura 17 são exibidos os histogramas da faixa de maior concentração de energia dos 3 momentos para cada um dos 3 intervalos de frequência (0-600 Hz, 0-1000 Hz e 0-3000 Hz) nos 2 grupos existentes. Para o cálculo dos valores do histograma foram analisados os valores de energias contidos em cada faixa de

frequência e demarcado em qual faixa havia uma maior concentração de energia, para todos os sinais de nosso estudo. Depois dessa marcação foram contados quantos casos havia em cada faixa de frequência. Em (a) têm-se os valores do *Burst* 1, em (b) representado o *Burst* 2 e em (c) exibindo os valores da fase faríngea (FF), em todos os casos são demonstrados os valores da deglutição de substância líquida e pastosa em cada um dos 3 intervalos.

Figura 17 - Histogramas da faixa de maior concentração de energia com 10 faixas

Nota-se pela figura 17 que o sinal não segue padrão específico em relação à frequência que possui maior concentração de energia. No entanto, pode-se notar que a faixa de frequência que exibe o pico de maior concentração, comparando com as demais faixas, situa-se sempre nas baixas frequências.

Figura 18 - Histogramas da faixa de maior concentração de energia com 19 faixas

Similarmente ao método empregado e demonstrado na figura 17 foi extraída também a energia em 19 faixas de frequência, em vez de 10 faixas, para se abranger intervalo similar, porém com uma sobreposição de faixas de 50%. Os histogramas da faixa de maior concentração de energia dos 6 casos existentes que são a combinação dos 3 momentos para cada um dos 2 grupos existentes estão mostrados na figura 18. Foi escolhido um agrupamento por momentos para uma melhor comparação entre os casos.

Para comparar os quatro casos diferentes aplicados com energias contidas em faixas de frequências, a figura 19 exibe a média entre todos os sinais utilizados, para cada momento e para cada tipo de deglutição. Em (a) tem-se as médias para o intervalo que contém de 0-600 Hz, em (b) a curva da média do intervalo entre 0-600 Hz com sobreposição, em (c) para a faixa de 0-1000 Hz e finalmente em (d) englobando o espectro de 0-3000 Hz.

Figura 19 - Média da energia por faixa de frequência

A tabela 9 mostra a posição do bolo alimentar, verificada com o auxílio da videofluoroscopia, em relação aos momentos, diferenciando entre substância líquida e pastosa. A abreviação CO representa cavidade oral.

Tabela 9 - Posição do bolo alimentar

Momento	Posição	Líquida (%)	Pastosa (%)
<i>Burst 1</i>	Orofaringe	71,8	87,5
	Faringe	20,5	9,4
	CO com escape posterior	7,7	0,0
	Não definido	0,0	3,1
<i>Burst 2</i>	CO + Faringe	25,6	50,0
	CO + Faringe + pequena parte em esôfago	53,9	25,0
	CO + Faringe + maior parte em esôfago	12,8	12,5
	Faringe + maior parte em esôfago	7,7	0,0
	Não definido	0,0	12,5
<i>Burst 3</i>	Parte distal do alimento em faringe	41,0	75,0
	Parte distal do alimento já em esôfago	46,2	21,9
	Não definido	12,8	3,1

7.2 Classificadores

Para classificação foram definidos 6 conjuntos de testes, sendo eles:

- Os 10 parâmetros definidos na secção 4.1
- Os 4 parâmetros mais significativos selecionados na secção 7.1
- A faixa de frequência de 0-600 Hz detalhada na secção 4.2
- A faixa de frequência de 0-600 Hz com sobreposição, detalhada na secção 4.2
- A faixa de frequência de 0-1000 Hz detalhada na secção 4.2
- A faixa de frequência de 0-3000 Hz detalhada na secção 4.2

A figura 20 exibe a acurácia dos 4 métodos de classificação utilizados (OPF, SVM, ANN e BC). Pode-se verificar que os classificadores se comportaram de maneira similar.

Figura 20 - Acurácia obtida com OPF, SVM, ANN e BC

A tabela 10 mostra o tempo médio de execução, em segundos, para cada classificador empregado em todos os conjuntos de testes estabelecidos e em cada um dos 3 momentos. O OPF tem um tempo muito menor de execução quando comparado com SVM e ANN e um tempo próximo, mas ainda menor, para o BC.

Tabela 10 - Tempo médio de execução dos classificadores

Momento	Variável	Classificador			
		OPF	SVM	ANN	BC
FF	10 Par	0,000964	1,925501	5,036465	0,002653
	4+	0,000679	2,283511	5,757901	0,001920
	Energia 600	0,000937	1,638272	0,852766	0,002147
	Energia 600 over	0,001046	1,716643	0,872152	0,002577
	Energia 1000	0,001007	1,788516	1,607049	0,001852
	Energia 3000	0,000869	1,917076	2,304783	0,004007
Burst 1	10 Par	0,000911	1,962406	3,639658	0,002649
	4+	0,000736	2,846782	6,138845	0,001972
	Energia 600	0,005320	1,679538	1,460969	0,002156
	Energia 600 over	0,001022	1,877829	0,066448	0,002035
	Energia 1000	0,000970	1,890635	1,558226	0,001939
	Energia 3000	0,000867	2,205440	0,075933	0,002092
Burst 2	10 Par	0,000862	2,020966	5,287574	0,002389
	4+	0,000692	2,693536	6,212895	0,001732
	Energia 600	0,001065	1,645393	1,622748	0,001770
	Energia 600 over	0,000928	1,640397	3,115564	0,002457
	Energia 1000	0,000966	1,637332	4,746125	0,004655
	Energia 3000	0,000820	2,032972	0,821940	0,001541

A figura 21 exibe o gráfico dos tempos descritos na tabela 10 em escala logarítmica, no qual se pode notar a grande diferença de tempo e o alto desempenho do classificador OPF. Vale ressaltar que o OPF foi, em média, 1713 vezes mais rápido que o classificador SVM, 2477 vezes mais rápido que a ANN e aproximadamente 2 vezes mais rápido que o BC.

Figura 21 - Tempo médio de execução dos classificadores

Como neste trabalho foi focado o emprego do classificador OPF, será demonstrado o resultado da acurácia obtida com o mesmo, juntamente com o desvio padrão para cada caso. Na figura 22 os resultados estão separados pelos momentos; para cada um dos 6 conjuntos de testes, as barras em azul representam a fase faríngea, as barras verdes, o *Burst 1*, e as vermelhas, o *Burst 2*.

Figura 22 - Acurácia obtida com OPF

A acurácia média entre todos os conjuntos de testes foi de aproximadamente 70% e nos 3 momentos analisados o conjunto que apresentou melhor resultado foi o que representa a variável que mais diferiu estatisticamente entre as 10, conforme explicado na secção 7.1.

A figura 23 mostra as variáveis que apresentaram uma acurácia acima da média; a presença do cone indica que a variável obteve uma acurácia acima da média, a qual foi calculada para cada momento (Fase faríngea, *Burst 1* e *Burst 2*).

Figura 23 - Variáveis com acurácia acima da média obtida com OPF

Pode-se notar que a variável 4+, a qual representa os 4 parâmetros mais significativos seleccionados na secção 7.1, foi a que apresentou acurácia acima da média para os 3 momentos em análise. A faixa de frequência de 0-600 Hz com sobreposição, e a faixa de frequência de 0-1000 Hz, ambas detalhadas na secção 4.2, apresentaram acurácia acima da média para 2 momentos. Finalmente a variável que representa os 10 parâmetros, definidos na secção 4.1, e a faixa de frequência de 0-600 Hz, detalhada na secção 4.2, apresentaram acurácia acima da média para 1 momento.

A tabela 11 mostra algumas medidas estatísticas da acurácia obtida com o classificador OPF e a figura 24 exibe um gráfico com essas medidas.

Tabela 11 - Medidas estatísticas da acurácia obtida com OPF

Momento	Medidas estatísticas			
	Mínimo	Média	Máximo	Desvio Padrão
FF	68,13%	71,98%	74,38%	2,69%
<i>Burst 1</i>	56,25%	68,13%	83,13%	10,55%
<i>Burst 2</i>	63,93%	67,56%	71,07%	3,25%

Figura 24 - Medidas estatísticas da acurácia obtida com OPF

Pode-se constatar que a acurácia não sofreu grande variação com relação às variáveis; apenas no *Burst 1*, o qual apresentou uma acurácia média de 68,13% e um desvio padrão de 10,55%, essa diferença ficou mais explícita.

CAPÍTULO 8

DISCUSSÃO

Neste trabalho fez-se a proposta de analisar o sinal da deglutição envolvendo técnicas adversas sob diferentes aspectos. Para análise desse tipo de sinal ainda não existe nenhuma técnica específica definida. Não existe uma padronização em termos de tempo de captura, frequência de amostragem ou mesmo de parâmetros de análise utilizados.

Existem muitos problemas tanto na captura quanto na análise do sinal da deglutição. No decorrer desta pesquisa foram feitos muitos testes antes de serem adotados os padrões definidos no capítulo 6 e empregados nesta tese. Estes padrões foram definidos com o intuito de minimizar os problemas, sempre com base em dados levantados na literatura recorrente.

A primeira dificuldade encontrada em nosso estudo foi na escolha do método de captura, já que segundo, Takahashi, Groher e Michi (1994), o detector acústico mais apropriado, para captura do som da deglutição, era o acelerômetro em consequência da ampla faixa de resposta em frequência e do baixo nível de atenuação. No entanto, em revisão desta metodologia em 2002, Cichero e Murdoch apresentaram melhores resultados para o microfone quando comparado ao acelerômetro (CICHERO; MURDOCH, 2002a). Apesar de metodologias diferentes terem sido utilizadas nestes estudos, os últimos justificam a diferença dos resultados pelo fato de que Takahashi, Groher e Michi (1994) amplificaram somente o sinal do

acelerômetro, causando assim uma rejeição prematura ao microfone. Dessa forma, com base nesses dados de literatura, iniciamos nossa coleta utilizando um microfone de eletreto.

Muitos problemas foram encontrados em nossa pesquisa com o emprego de microfone de eletreto, como por exemplo, a saturação do sinal, que pode ser justificada pelo fato de o sinal da deglutição conter, em um determinado momento, uma variação abrupta de amplitude. Muitos conversores analógico-digitais possuem controles de ganho para minimizar variações desse tipo, mas estes controles, em nosso caso, se tornam prejudiciais à qualidade do sinal. Por esse motivo optamos por utilizar microfone dinâmico, que por sua vez era conectado a um pré-amplificador de uso profissional e posteriormente ligado à entrada de linha do conversor (placa de som ou do gravador de DVD quando empregado juntamente com a videofluoroscopia), evitando dessa forma a saturação do sinal e conseguindo uma ampla faixa de resposta em frequência e baixo nível de atenuação.

Com relação à localização do detector acústico, este foi fixado na área próxima à borda lateral da traqueia imediatamente inferior a cartilagem cricóide, região esta considerada ótima por Takahashi, Groher e Michi (1994) e Cichero e Murdoch (2002a), pelo fato de ter apresentado a maior média de magnitude do sinal e o menor desvio padrão dos picos máximos do som da deglutição.

Sob a questão da frequência de amostragem foi feito um levantamento bibliográfico dos artigos mais relevantes (CICHERO; MURDOCH, 2002b; HAMLET; NELSON; PATTERSON; 1990; LEAR; FLANAGAN; MORREES, 1965; MACKOWIAK; BRENNAN; FRIEDMAN, 1967; YOUNG; STIERWALT, 2005), e definida uma taxa que permitisse certa tranquilidade com relação ao corte de frequências, o que resultou na frequência de amostragem de 22050 Hz.

Outro fator que dificultava na captura era o tempo definido de coleta, evidentemente levando em conta a não necessidade de edição destes sinais. Como nesta tese foi empregada a coleta simultânea de vídeo e som, esse tempo não foi importante, porém foram realizados primariamente outros tipos de coletas somente do sinal da deglutição com o intuito de auxiliar para uma melhor compreensão do sinal. Foi muito difícil estipular o tempo de captura, pois muitas pessoas demoravam a desencadear a deglutição. Esse fator ainda se agravava mais quando se tratava de pacientes com disfagia, que possuíam uma fase oral muito lenta.

Depois das dificuldades na coleta demonstradas acima, o maior problema de todos era como seria feita a análise desse sinal, o que deveria ser analisado ou quais momentos do sinal eram importantes nesta análise. O problema de delimitação dos intervalos e momentos foi minimizado e muito com o auxílio da imagem da videofluoroscopia adquirida sincronizadamente com o sinal. Certamente contando com o auxílio de uma fonoaudióloga com experiência na área de análise de videofluoroscopia, a qual realizava uma minuciosa análise dos exames coletados e fazia a marcação dos pontos previamente definidos.

Resolvido os problemas mencionados e de posse das marcações que forneciam pontos para delimitação de trechos importantes do sinal, surgia a questão do que seria analisado e quais parâmetros seriam utilizados para tal fim. Considerando esse problema, foi feito um extenso levantamento bibliográfico e com base nessa pesquisa escolhidos alguns métodos numéricos, os quais foram empregados nesta análise; também foram criados pelo autor outros métodos, todos já demonstrados no capítulo 4.

Na maior parte das deglutições avaliadas, que resultou em um total de 71, sendo descartados 15,48% por fatores adversos, foi possível localizar 3

componentes de som, mas em 5 deglutições (7,04%), só foi possível detectar 2 componentes de som, mesmo realizando a análise simultânea com a videofluoroscopia da deglutição, o que permitia uma análise muito mais minuciosa. Estes dados corroboram os estudos de Hamlet, Nelson e Patterson (1990); Mackowiak, Brenman e Friedman (1967), e Morinière, Beutter e Boiron (2008). Acredita-se que nos pacientes nos quais não foi possível detectar todos os componentes de som, este achado possa ser justificado devido a mínimos intervalos entre os componentes de som levando a fusão destes quando o bolo passa rapidamente.

Em relação aos tempos de duração de cada parâmetro analisado neste estudo encontram-se: duração total da fase faríngea $659,63 \pm 129,33$ ms para a consistência pastosa, $693,40 \pm 149,16$ ms para a consistência líquida, um tempo ligeiramente maior que a pastosa. Esses dados corroboram o estudo de Boiron, Rouleau e Metman (1997). No trabalho desses autores citados anteriormente é citada a hipótese de que o tempo de abertura do esfíncter esofágico superior depende tanto do volume quanto da consistência deglutida, e que quanto maior o volume deglutido, maior o tempo de abertura deste. No entanto, em relação à consistência, não foi constatada diferença estatisticamente significativa para as diferentes viscosidades, dados estes que corroboram o estudo de Kim et al. (1994), concluindo que o tempo de esvaziamento faríngeo não é influenciado pela consistência deglutida.

Concordando com nossos achados, Dantas et al. (1990) consideraram que ao aumentar a viscosidade do bolo alimentar, também aumenta a resistência para passagem, resultando na diminuição da velocidade de deslocamento do bolo e consequentemente maior tempo para a consistência com maior viscosidade. Desta

forma deve-se ressaltar que neste estudo, apesar da ligeira diferença entre as consistências testadas, não foi encontrada diferença estatisticamente significativa.

No entanto, quando se trata dos componentes de som os resultados são contrários, encontrando valores menores para a consistência pastosa. Primeiro componente de som (*Burst 1*) com duração de $87,27 \pm 36,40$ ms na consistência líquida e $78,20 \pm 37,53$ ms na pastosa, e o segundo componente de som (*Burst 2*) com duração média de $112,93 \pm 62,94$ ms, para a líquida e $85,50 \pm 52,20$ ms para pastosa. Nesse achado não se pode utilizar a mesma hipótese utilizada acima (DANTAS et al., 1990), que quanto mais viscoso, mais lento seria a passagem e maior seria o tempo do componente de som, e o mesmo raciocínio seria utilizado para a consistência líquida, que passaria mais rapidamente pela faringe e esfíncter esofágico superior e desse modo provoca reflexos mais rápidos. Dessa forma, para confirmar esses achados seria necessário uma amostra maior, para que se pudesse confirmar essa diferença estatisticamente, já que neste estudo essa diferença não é significativa.

Apesar de neste estudo, na maioria das deglutições serem detectados 3 componentes de som (3 *Bursts*), somente foi possível quantificar 2 desses componentes, pela dificuldade de demarcação do final do último componente. Os tempos desses dois componentes foram exibidos anteriormente. Foi possível também quantificar o intervalo entre esses componentes: Intervalo (1-2) 82 ± 66 ms (água contrastada com bário), 95 ± 48 ms (pastosa); Intervalo (2-3) 339 ± 127 ms (água contrastada com bário), 322 ± 145 ms (pastosa).

Estudos como de Mackowiak, Brenman e Friedman (1967), também determinam 3 explosões acústicas, mas só quantificam as duas iniciais, com um intervalo mais curto que o deste estudo entre as duas primeiras explosões de 100 a

150 ms, e um intervalo mais longo entre a segunda e terceira explosão de 300 a 400 ms. No entanto, com relação à duração dos componentes, não se observam exatamente os mesmos achados, mas pode-se dizer que é seguido o padrão com um primeiro componente mais curto e um segundo mais longo (HAMLET; NELSON; PATTERSON, 1990).

Apesar de Morinière, Beutter e Boiron (2006) e Morinière et al. (2008) terem identificado 6 componentes de som, enfatizaram que somente 3 ocorriam com maior sistematicidade; os outros ocorriam em casos esporádicos. Da mesma forma não se encontram valores semelhantes, com variabilidade considerável dos achados em cada estudo. Mas, novamente pode-se encontrar uma relação temporal entre os estudos, como citado acima, ou seja, apesar dos números não serem iguais, seguem um padrão com a primeira explosão mais curta, juntamente com o primeiro intervalo, e o segundo componente mais longo, semelhante ao segundo intervalo.

Com todos esses dados, pode-se perceber a grande variabilidade das deglutições, mesmo intra-sujeito. Dessa forma objetivou-se nesta pesquisa definir parâmetros para que fosse possível uma análise numérica e mais automática desses parâmetros, minimizando assim a subjetividade da análise. Os parâmetros foram usados a posteriori nos classificadores.

Das dez variáveis estudadas, somente 2 apresentaram correlação e dessa forma somente uma variável, se necessário, poderia ser descartada da análise. Essas variáveis se referem à potência e variância. Destas variáveis estudadas, somente duas estão presentes em trabalhos sobre disfagia, que são estacionaridade e taxa de dispersão (LEE et al., 2006a; CHAU et al., 2005).

A análise estatística entre os 10 parâmetros, os momentos definidos no estudo (*Burst 1*, *Burst 2* e Fase Faríngea) e os grupos estudados (pastosa e líquida),

auxiliaram a definir as 4 variáveis que posteriormente foram utilizadas em um subconjunto de treinamento para os classificadores. São elas: potência (P), dimensão fractal (DF), potência do envelope (PLPC) e estacionaridade (ZA). Essas variáveis foram as que mais apresentavam diferenças estatísticas ($p < 0,05$). Deve-se relatar aqui ser possível que em novos estudos com um número maior de amostras essas variáveis possam ser alteradas.

Neste estudo foi buscado entender se o sinal e seus momentos estavam relacionados a faixas de frequências específicas, e se existia sistematicidade nesses achados. O sinal é muito variável (figuras 17, 18 e 19) não podendo estabelecer relação direta entre qualquer faixa de frequência e os componentes de som. Dessa forma, mais uma vez esses achados direcionam a análise dos dados para o uso de técnicas mais avançadas de classificação. Spadotto et al. (2008b, 2009) usaram faixas de frequências com o intuito de classificar a deglutição de sujeitos normais e com disfagia orofaríngea, encontrando uma especificidade em torno de 76% a 86%.

Lazareck e Moussavi (2004a, 2004b), que pretenderam em ambos os estudos classificar deglutições entre normais e disfágicas, utilizaram como parâmetro de classificação a dimensão de onda baseada em dimensão fractal, e obtiveram uma classificação maior que 90%. Lee et al. (2006a, 2006b) utilizaram um classificador de base radial para detectar aspiração em crianças obtendo uma acurácia entre 70,94 e 78,46% variando os parâmetros utilizados na pesquisa. Recrutaram para a pesquisa 117 crianças com risco de aspiração e, embora este seja um problema ético na coleta, devido ao risco oferecido às crianças, os pesquisadores obtiveram a aprovação do comitê de ética. Aboofazeli e Moussavi (2004) empregaram dimensão fractal e rede neural para classificar sinais da deglutição e sons da respiração, obtendo uma acurácia média de 91,7%.

Os trabalhos citados acima possuem enfoques diferentes do utilizado neste estudo. Aqui o intuito foi de tentar compreender os eventos que ocorrem em deglutições de pessoas saudáveis. Depois, em um trabalho futuro, os achados poderiam servir de base para comparações com deglutições de pessoas com disfagia. Como foram empregadas duas consistências diferentes, buscou-se exibir as diferenças e/ou semelhanças entre a deglutição de ambas.

Os resultados entre os classificadores apresentados foram equivalentes, considerando média e desvio padrão dos 4 classificadores empregados. Porém, em relação ao tempo de execução, o total foi muito menor para o OPF. Cabe destacar que os tempos de execução apresentados correspondem às etapas de treinamento e teste. Embora o classificador SVM seja um problema de otimização quadrático, o mesmo torna-se exponencial quando é necessário escolher os parâmetros ótimos, ou seja, aqueles que minimizam o erro no conjunto de treinamento. Tais parâmetros são escolhidos de maneira exaustiva pelo pacote de implementação utilizado para o classificador SVM (PAPA et al., 2008). As redes neurais tiveram sua eficiência comprometida devido à conhecida complexidade de execução e demora na convergência de algumas situações do algoritmo de treinamento por retro-propagação utilizado nas redes neurais com perceptron multicamadas (ANN).

Com base na figura 22, que exibe os resultados do OPF em cada um dos 3 momentos (*burst 1*, *burst 2* e fase faríngea) para cada um dos 6 conjuntos de testes, definidos na seção 7.2, pode ser verificado que nos 3 momentos o conjunto que possibilitou uma melhor classificação foi o conjunto de dados que representava a seleção das 4 variáveis que mais se diferenciaram estatisticamente ($p < 0,05$).

CAPÍTULO 9

CONCLUSÃO

Com base nos resultados apresentados nas figuras 22 e 23, pode-se constatar que o melhor vetor de características para representar os momentos, em um conjunto de dados para classificadores, seria o conjunto representado pela variável 4+. Empregando também a energia contida em faixas de frequência em novos conjuntos de dados representativos, seria interessante o emprego de duas faixas, ou seja, dois conjuntos de dados: o primeiro conjunto seria o que contém a faixa de 0-600 Hz com sobreposição, e o segundo conjunto, o que engloba a faixa de 0-1000 Hz.

A variabilidade entre as variáveis calculadas, assim como a quantidade de energia contida em determinadas faixas de frequência em cada amostra, foi muito grande, não conseguindo assim apresentar nenhum comportamento padrão que caracterizasse o sinal.

Conclui-se neste estudo que muitas outras pesquisas ainda serão necessárias para um melhor entendimento sobre a produção e a análise do som da deglutição. A ideia de analisar o sinal concomitante com a videofluoroscopia é muito interessante, porém requer uma análise muito mais demorada e dificulta bastante a coleta, consequentemente diminuindo muito a capacidade de se adquirir amostras. Com as pesquisas atuais, só se pode dizer que, entre um número específico de deglutições, determinado som pode estar associado a algum movimento em

particular. Um fator que poderia vir a minimizar essa constatação seria realizar uma análise com um grande número de casos, variando idade e sexo. A princípio ainda com o auxílio da videofluoroscopia, podendo chegar a uma melhor compreensão dos eventos da deglutição, para posteriormente ser possível obter resultados somente com a gravação do sinal.

Um trabalho futuro poderia ser o de aumentar o número de amostras de sinais de pessoas saudáveis bem como começar a coletar de pessoas com disfagia. Talvez, com base em novos resultados, essa nova proposta de análise, a qual se propõe avaliar a deglutição somente pelo sinal sonoro, poderia começar a ser empregada na prática clínica.

REFERÊNCIAS

ABOOFAZELI, M.; MOUSSAVI, Z.M.K. (2004). Automated classification of swallowing and breath sounds. In: ANNUAL INTERNATIONAL CONFERENCE OF THE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 26., 2004, San Francisco. **Proceedings...** New York: IEEE.

_____. (2006). Automated extraction of swallowing sounds using a wavelet-based filter. In: ANNUAL INTERNATIONAL CONFERENCE OF THE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 28., 2004, New York. **Proceedings...** New York: IEEE. p.5607-5610.

ADDISON, P.S. (2002). **The Illustrated wavelet transform handbook**: introductory theory and applications in science, engineering, medicine and finance. Boca Raton: CRC Press.

AL-AKAIDI, M. (2004). **Fractal speech processing**. Cambridges: Cambridge University Press.

BOIRON, M.; ROULEAU, P.; METMAN, E.H. (1997). Exploration of pharyngeal swallowing by audiosignal recording. **Dysphagia**, Berlin, v.12, n.2, p.86-92, Apr.

BORR, C.; HIELSCHER-FASTABEND, M.; LÜCKING, A. (2007). Reliability and validity of cervical auscultation. **Dysphagia**, Berlin, v.22, n.3, p.225-234, July.

BOSER, B.E.; GUYON, I.M.; VAPNIK, V.N. (2007). A Training algorithm for optimal margin classifiers. In: WORKSHOP ON COMPUTATIONAL LEARNING THEORY, 5., 1992, MorganKauffmann. **Proceedings...** New York: ACM. p.144-152.

BUCHHLOZ, D.W. (1994). Dysphagia associated with neurological disorders. **Acta Otorhino-laryngologica Belgica**, Belgium, v.48, n.2, p.143-155.

CHAU, T. et al. (2005). Investigating the stationarity of paediatric aspiration signals. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, New York, v.13, n.1, p.99-105, Mar.

CHEN, C.H. (1988). **Signal processing handbook**. New York: Dekker.

CHEN, S.H.; GUIDO, R.C.; CHEN, S.H. (2007). Voice activity detection in car environment using support vector machine and wavelet transform. In: IEEE INTERNATIONAL SYMPOSIUM ON MULTIMEDIA WORKSHOPS, 9., 2007, New York. **Proceedings...** New York: IEEE. p.252-255.

CICHERO, J.A.Y.; MURDOCH, B.E. (1998). The Physiologic cause of swallowing sounds: answers from heart sounds and vocal tract acoustics. **Dysphagia**, Berlin, v.13, n.1, p.39-52, Jan.

_____. (2002a). Acoustic signature of the normal swallow: characterization by age, gender, and bolus volume. **Annals of Otology, Rhinology and Laryngology**, St.Louis, v.111, p.623-632, July.

_____. (2002b). Detection of swallowing sounds: methodology revisited. **Dysphagia**, Berlin, v.17, n.1, p.40-49, Jan.

COSTA, M.M.B. et al. (1992). Videofluoroscopia: um novo método. **Radiologia Brasileira**, São Paulo, v.25, p.11-18.

DANTAS, R.O. et al. (1990). Effect of swallowed bolus variables on oral pharyngeal phases of swallowing. **American Journal of Physiology**, Bethesda, v.258, p.G675-G681.

DAUBECHIES, I. (1988). Orthonormal bases of compactly supported wavelets. **Communications on Pure and applied Mathematics**, New York, v.41, n.7, p.909-996, Oct.

_____. (1990). The Wavelet transform, time-frequency localization and signal analysis. **IEEE Transactions on Information Theory**, New York, v.36, n.5, p.961-1005, Sept.

_____. (1992). **Ten lectures on wavelets**. Philadelphia: Society for Industrial and Applied Mathematics. (CBMS-NSF Regional Conference Series in Applied Mathematics, 61).

DENG, L.; O'SHAUGHNESSY, D. (2003). **Speech processing**: a dynamic and optimization-oriented approach. New York: Marcel Dekker.

DUDA, R.O.; HART, P.E.; STORK, D.G. (2000). **Pattern classification**. 2nded. New York: Wiley-Interscience.

FALCÃO, A.X.; STOLFI, J.; LOTUFO, R.A. (2004). The Image foresting transform: theory, algorithms, and applications. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v.26, n.1, p.19-29, Jan.

FUKUNAGA, K.; NARENDRA, P.M. (1975). A Branch and bound algorithms for computing k-nearest neighbors. **IEEE Transactions on Computers**, New York, v.24, n.7, p.750-753, July.

FURKIM, A.M.; SILVA, R.G. (1999). **Programas de reabilitação em disfagia neurogênica**. São Paulo: Frôntis.

HAMLET, S.; NELSON, R.; PATTERSON, R. (1988). Sounds of swallowing. **The Journal of the Acoustical Society of America**, New York, v.83, Suppl.1, p.S23, May.

HAMLET, S.; NELSON, R.; PATTERSON, R. (1990). Interpreting the sounds of swallowing: fluid flow through the cricopharyngeus. **Annals Otology, Rhinology and Laryngology**, St.Louis, v.99, p.749-752.

HAYKIN, S. (1998). **Neural networks**: a comprehensive foundation. Upper Saddle River: Prentice-Hall.

HUANG, T.M.; KECMAN, V.; KOPRIVA, I. (2006). **Kernel based algorithms for mining huge data sets**: supervised, semi-supervised, and unsupervised learning. New York: Springer-Verlag. (Studies in Computational Intelligence, 17).

JAIN, A.K.; DUIN, R.P.W.; MAO, J. (2000). Statistical pattern recognition: a review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v.22, n.1, p.4-37, Jan.

KENDALL, K.A.; LEONARD, R.J.; MCKENZIE, S.W. (2001). Accomodation changes in bolus viscosity in normal deglutition: a videofluoroscopic study. **Annals of Otology, Rhinology and Laryngology**, St.Louis, v.110, p.1059-1065.

KENDALL, K. A. et al. (2000). Timing of events in normal swallowing: a videofluoroscopic study. **Dysphagia**, Berlim, v.15, n.2, p.74-83, Mar.

KIM, C.H. et al. (1994). Effect of viscosity on oropharyngeal and esophageal emptying in man. **Digestive Diseases and Sciences**, New York, v.39, n.1, p.189-192, Jan.

LAZARECK, L.J.; MOUSSAVI, Z.M.K. (2004a). Classification of normal and dysphagic swallows by acoustical means. **IEEE Transactions on Biomedical Engineering**, New York, v.51, n.12, p.2103-2112, Sept.

LAZARECK, L.J.; MOUSSAVI, Z.M.K. (2004b). Swallowing sound characteristics in healthy and dysphagic individuals. In: ANNUAL INTERNATIONAL CONFERENCE OF THE ENGINNERING IN MEDICINE AND BIOLOGY SOCIETY, 26., 2004, San Francisco. **Proceedings...** New York: IEEE. v.2, p.3820-3823.

LEAR, C.S.C.; FLANAGAN, J.B.; MORREES, C.F.A. (1965). The Frequency of deglutition in man. **Archives of Oral Biology**, New York, v.10, n.1, p.83-99, Jan./Feb.

LEE, J. et al. (2006a). A Radial basis function classifier for pediatric aspiration detection. In: ANNUAL INTERNATIONAL CONFERENCE OF THE ENGINNERING IN MEDICINE AND BIOLOGY SOCIETY, 28., 2006, New York. **Proceedings...** New York: IEEE. p.3553-3556.

LEE, J. et al. (2006b). A Radial basis classifier for the automatic detection of aspiration in children with dysphagia. **Journal of NeuroEngineering and Rehabilitation**, London, v.3, n.14, p.1-17, July.

LESLIE, P. et al. (2007). Cervical auscultation synchronized with images from endoscopy swallow evaluations. **Dysphagia**, Berlim, v.22, n.4, p.290-298, Oct.

LOGAN, W.J.; KAVANAGH, J.F.; WORNALL, A.W. (1967). Sonic correlates of human deglutition. **Journal of Applied Physiology**, Washington, v.23, n.2, p.279-284, Aug.

LOGEMANN, J.A. (1983). **Evaluation and treatment of swallowing disorders**. Austin: Pro-ed.

MACKOWIAK, R.C.; BRENMAN, H.S.; FRIEDMAN, M.H.F. (1967). Acoustic profile of deglutition. **Proceedings of the Society for Experimental Biology and Medicine**, Malden, v.125, n.4, p.1149-1152, Aug./Sept.

MALLAT, S.G. (1989). A Theory for multiresolution signal decomposition: the wavelet representation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v.11, n.7, p.674-693, July.

MANUAL Merck. (2009). **Espasmo difuso do estômago**. Seção 9 - distúrbios digestivos. Capítulo 101 - distúrbios do esôfago. Disponível em: <www.msd-brazil.com/msdbrazil/patients/manual_Merck/mm_sec9_101.html>. Acesso em: 10 jun. 2009.

MCCONNEL, F.M.S. et al. (1988). Timing of major events of pharyngeal swallowing. **Archives of Otolaryngology Head & Neck Surgery**, Chicago, v.114, n.12, p.1413-1418, Dec.

MORINIÈRE, S.; BEUTTER, P.; BOIRON, M. (2006). Sound component duration of healthy human pharyngoesophageal swallowing: a gender comparison study. **Dysphagia**, Berlim, v.21, n.3, p.175-182, July.

MORINIÈRE, S. et al. (2008). Origin of the sound components during pharyngeal swallowing in normal subjects. **Dysphagia**, Berlim, v.23, n.3, p.267-273, Sept.

MOUSSAVI, Z.M.K. (2005). Assessment of swallowing sounds' stages with hidden Markov model. In: ANNUAL INTERNATIONAL CONFERENCE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 27., 2005, Shanghai. **Proceedings...** New York: IEEE. p.6989-6992.

PAPA, J.P. et al. (2007). Design of robust pattern classifiers based on optimum-path forests. In: INTERNATIONAL SYMPOSIUM ON MATHEMATICAL MORPHOLOGY, 8., Rio de Janeiro, 2007. **Proceedings: mathematical morphology and its applications to signal and image processing**. São José dos campos: MCT/INPE. p.337-348.

PAPA, J.P. et al. (2008). A Discrete approach for supervised pattern recognition. In: INTERNATIONAL WORKSHOP ON COMBINATORIAL IMAGE ANALYSIS, 12., 2008, Buffalo. **Proceedings...** Berlim: Springer. v.4958, 136-147.

PERLMAN, A.L.; ETTEMA, S.L.; BARKMEIER, J. (2000). Respiratory and acoustic signals associated with bolus passage during swallowing. **Dysphagia**, v.15, n. 2, p.89-94, Mar.

ROBBINS, J. et al. (1992). Oropharyngeal swallowing in normal adults of different ages. **Gastroenterology**, New York, v.103, n.3, p.823-829, Sept.

SELLEY, W.G. et al. (1994). The Synchronization of respiration and swallow sounds with videofluoroscopy during swallowing. **Dysphagia**, Berlim, v.9, n.3, p.162-167, June.

SHANNON, C.E. (1951). Prediction and entropy of printed english. **Bell System Technical Journal**, New York, v.30, p.50-64.

SILVA, R.G.; GATTO, A.R.; COLA, P.C. (2003). Disfagia orofaríngea neurôgenica em adultos - avaliação fonoaudiológica em leito hospitalar. In: JACOBI, J.S.; LEVY, D.S.; SILVA, L.M.C. **Disfagia-avaliação e tratamento**. Rio de Janeiro: Revinter. p.181-96.

SMITH, S.W. (1999). **The Scientist and engineer's guide to digital signal processing**. 2nded. San Diego: California Technical.

SPADOTTO, A.A. et al. (2006). Método quantitativo para análise da disfagia orofaríngea em pacientes com doença cérebro vascular. **O Dendrito**, Curitiba, v.12, p.29.

_____. (2008a). Denoising swallowing sound to improve the evaluator's qualitative analysis. **Computers and Electrical Engineering**, Oxford, v.34, n.2, p.148-153, Mar.

_____. (2008b). Oropharyngeal dysphagia identification using wavelets and optimum path forest. In: INTERNATIONAL SYMPOSIUM ON COMMUNICATIONS, CONTROL AND SIGNAL PROCESSING, 3., 2008, Malta. **Proceedings...** New York: IEEE. p.735-740.

_____. (2008c). Software para análise quantitativa da deglutição. **Radiologia Brasileira**, São Paulo, v.41, n.1, p.25-28, jan./fev.

_____. (2009). Classification of normal swallowing and oropharyngeal dysphagia using wavelet. **Applied Mathematics and Computation**, New York, v.207, n.1, p.75-82, Jan.

STRANG, G.; NGUYEN, T. (1996). **Wavelets and filter banks**. Wellesley: Wellesley-Cambridge.

STROUD, A.E.; LAWRIE, B.W.; WILES, C.M. (2002). Inter and intra-rater reliability of cervical auscultation to detect aspiration in patients with dysphagia. **Clinical Rehabilitation**, London, v.16, n.6, p.640-645, Sept.

TAKAHASHI, K.; GROHER, M.E.; MICH, K. (1994). Methodology for detecting swallowing sounds. **Dysphagia**, Berlim, v.9, n.1, p.54-62, June.

YADOLLAHI, A.; MOUSSAVI, Z.M.K. (2007). Feature selection for swallowing sounds classification. In: ANNUAL INTERNATIONAL CONFERENCE OF THE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 29., 2007, Lyon. **Proceedings...** New York: IEEE. p.3172-3175.

YOUMANS, S.R.; STIERWALT, J.A.G. (2005). An Acoustic profile of normal swallowing. **Dysphagia**, Berlim, v.20, n.3, p.195-209, Sept.

ANEXO 1

Universidade Estadual Paulista
Faculdade de Medicina de Botucatu

Distrito Rubião Junior, s/nº - Botucatu - S.P.
CEP: 18.618-970
Fone/Fax: (0xx14) 3811-6143
e-mail secretaria: capellup@fmb.unesp.br

Registrado no Ministério da Saúde em 30 de
abril de 1997

Botucatu, 05 de maio de 2008

Of. 147/08-CEP

Ilustríssimo Senhor
Prof. Dr. Arthur Oscar Schelp
Departamento de Neurologia/Psicologia e Psiquiatria da
Faculdade de Medicina de Botucatu.

Prezado Dr. Schelp,

De ordem da Senhora Coordenadora deste CEP, informo que o Projeto de Pesquisa "**Análise do som na deglutição**", a ser conduzido por Vossa Senhoria, com a participação de Ana Rita Gatto e André Augusto Spadotto, recebeu do relator parecer favorável, aprovado em reunião de 05 de maio de 2.008.

Situação do Projeto: **APROVADO**. Ao final da execução deste Projeto, apresentar ao CEP "**Relatório Final de Atividades**".

Atenciosamente,

Alberto Santos Capelluppi
Secretário do CEP.

ANEXO 2

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

www.hcrp.fmrp.usp.br

Ribeirão Preto, 12 de novembro de 2008

Ofício nº 4102/2008
CEP/SPC

Prezados Senhores,

O trabalho intitulado **“ANÁLISE DO SOM DA DEGLUTIÇÃO”**, foi analisado pelo Comitê de Ética em Pesquisa, em sua 276ª Reunião Ordinária realizada em 10/11/2008 e enquadrado na categoria: **APROVADO, bem como o Termo de Consentimento Livre e Esclarecido**, de acordo com o Processo HCRP nº 10073/2008.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (IGH-GCP), bem como a Resolução nº 196/96 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente,

PROF. DR. SÉRGIO PEREIRA DA CUNHA
Coordenador do Comitê de Ética em
Pesquisa do HCRP e da FMRP-USP

Ilustríssimos Senhores
DANIELLE RAMOS DOMENIS
PROF. DR. ROBERTO OLIVEIRA DANTAS (Orientador)
Depto. de Clínica Médica

TRABALHOS CIENTÍFICOS

Periódicos

- PEREIRA, J.C. et al. (2006). Residual signal auto-correlation to evaluate speech in parkinson's disease patients. **Arquivos de Neuro-Psiquiatria**, São Paulo, v.64, p.912-915.
- SPADOTTO, A.A. et al. (2008). Denoising swallowing sound to improve the evaluator's qualitative analysis. **Computers and Electrical Engineering**, Oxford, v.34, n.2, p.148-153, Mar.
- _____. (2008). Software para análise quantitativa da deglutição. **Radiologia Brasileira**, São Paulo, v.41, n.1, p.25-28, jan./fev.
- _____. (2009). Classification of normal swallowing and oropharyngeal dysphagia using wavelet. **Applied Mathematics and Computation**, New York, v.207, n.1, p.75-82, Jan.

Conferências (completo)

- PAPA, J.P. et al. (2008). Optimum path forest classifier applied to laryngeal pathology detection. In: INTERNATIONAL CONFERENCE ON SYSTEMS, SIGNALS AND IMAGE PROCESSING, 15., 2008, Bratislava. **Proceedings...** Elosváquia: House STU. p.249-252.
- SPADOTTO, A.A. et al. (2008). Oropharyngeal dysphagia identification using wavelets and optimum path forest. In: INTERNATIONAL SYMPOSIUM ON COMMUNICATIONS, CONTROL AND SIGNAL PROCESSING, 3., 2008, Malta. **Proceedings...** New York: IEEE. p.735-740.

Conferências (resumo)

- BARBOZA, R.; SPADOTTO, A.A. (2007). On a new FHN dynamics for excitable media. In: INTERNATIONAL SYMPOSIUM ON SYNCHRONIZATION IN COMPLEX NETWORKS, 2007, Leuven. p.67-68.
- COLA, P.C. et al. (2006). Medição do tempo faríngeo da deglutição em indivíduos após acidente vascular encefálico. **GED - Gastroenterologia e Endoscopia Digestiva**, v.25, p.111. Apresentado a Semana Brasileira do Aparelho Digestivo, São Paulo, 2006.
- _____. (2008). Sour taste and cold temperature effects on pharyngeal swallowing transit time after ischemic hemisphere stroke. In: ANNUAL DYSPHAGIA RESOURCE SOCIETY MEETING, 16., 2008, Charleston.

- _____. (2009). The Influence of sour taste and cold temperature on pharyngeal swallowing transit. In: ANNUAL DYSPHAGIA RESOURCE SOCIETY MEETING, 17., 2009, New Orleans.
- GATTO, A.R. et al. (2008). Quantitative analysis of the oral and pharyngeal transit on the control of therapeutic efficacy on neurogenic oropharyngeal dysphagia: a case report. In: WORLD CONGRESS FOR NEUROREHABILITATION AND NEURAL REPAIR, 5., 2008, Brasília. v.22, p.592.
- _____. (2009). Effects of sour taste and cold temperature on oral transit time of swallowing. In: ANNUAL DYSPHAGIA RESOURCE SOCIETY MEETING, 17., 2009, New Orleans.
- SILVA, R.G. et al. (2008). Protocolo para controle de eficácia em disfagia orofaríngea neurogênica (PROCEDON): uma proposta. In: CONGRESSO BRASILEIRO DE FONOAUDIOLOGIA, 16., 2008, Campos do Jordão.
- _____. (2009). Efficacy protocol in oropharyngeal dysphagia: a proposal. In: Dysphagia Resource Society, 2009, New Orleans. In: ANNUAL DYSPHAGIA RESOURCE SOCIETY MEETING, 17., 2009, New Orleans.
- SPADOTTO, A.A. et al. (2006). Método quantitativo para análise da disfagia orofaríngea em paciente com doença cérebro-vascular. **O Dendrito**, Curitiba, v.12, p.29. Apresentado ao 6º Encontro da Sociedade Brasileira de Investigação Neurológica, Ribeirão Preto, Ribeirão Preto.
- _____. (2008). Digital swallowing sounds: a useful method for dysphagic patients assessment. In: ANNUAL DYSPHAGIA RESOURCE SOCIETY MEETING, 16., 2008, Charleston.
- _____. (2009). Recognizing of specific signal's bursts during swallowing pharyngeal phase. In: ANNUAL DYSPHAGIA RESOURCE SOCIETY MEETING, 17., 2009, New Orleans.

Revisor de periódico

- IEEE International Conference on Multimedia & Expo, 2008
- IEEE International Conference on Multimedia & Expo, 2009
- IEEE Transactions on Biomedical Engineering, 2009