

FILIPPE MITSUO DA CRUZ INONHE

DETECÇÃO E QUANTIFICAÇÃO DA NASALIDADE EM UM SINAL DE VOZ

Trabalho de Conclusão de
Curso apresentado à Escola de
Engenharia de São Carlos, da
Universidade de São Paulo

Curso de Engenharia Elétrica com Ênfase em Eletrônica

Orientador: Prof. Dr. José Carlos Pereira

São Carlos
2009

À minha família

AGRADECIMENTOS

A Deus, pela vida e por toda a sabedoria e entendimento.

A meus pais, que são minha escola prima, que me apoiaram durante toda minha jornada na graduação e que me proporcionam aprender todos os dias o que nenhuma escola pode me ensinar.

Aos meus irmãos Leonardo, Débora e Gabriel, que são minha vida e mais me auxiliam quando ninguém mais pode.

À madrinha Penha e toda a família pelo suporte emocional durante todo o processo de elaboração deste trabalho.

Ao Prof. Doutor José Carlos Pereira por se dispor a dar toda a orientação, dedicação, aprendizado e força que faz possível com que este trabalho se realize. Tenho uma eterna dívida de gratidão com este professor e “pai” de todos nós.

À Eugenia por toda paciência e força de vontade para ajudar este engenheiro. Não existem palavras suficientes para agradecê-la, mas fico com um: “Muito obrigado! De coração!”.

Ao Paulo que me auxiliou em tudo que precisei quando estava aqui.

Aos funcionários do Departamento de Engenharia Elétrica que sempre estavam disponíveis quando precisei e nunca se recusaram a dar esclarecimentos ou atender pedidos: Jussara, Shirley, Aura, José, Rosemberg, Petrócio, Zé, Alessandro e Denis

Aos vizinhos Kôji e Priscila, por todo o incentivo durante esta jornada. Todos os conselhos e todos os cafés, fossem dias, tardes ou noites.

Ao pessoal do Laboratório de Processamento de Sinais que me acolheram e sempre foram muito bondosos e gentis: Jamylle, Lianet, Allan, Regina, Edwin e Soledad. Muchas gracias a todos.

A todos que foram extremamente solícitos e se dispuseram a me auxiliar no momento em que eu mais precisava: Amandinha, Edson, Fer, Pancada, Indara, Jéssika, Natália, Moussa, Orieta, Camila, Cris, Beto, Amílcar e Kustela.

Aos amigos Aline e Eduardo que se mostraram presentes e solícitos com relação ao desenvolvimento deste trabalho.

A todo o Ministério Universidades Renovadas por todas as orações, amizades e apoio que sempre pude contar de cada um dos meus irmãos. Obrigado a todos por serem minha segunda família.

RESUMO

O objetivo deste estudo foi encontrar uma forma de detecção e quantificação de nasalidade para uma vogal pronunciada de forma oral e nasal através da análise do espectro densidade de potência de um sinal de voz e da modelagem de um sistema fonte-filtro para o trato vocálico. A vogal escolhida é a /a/ sustentada oral e nasal por oferecer menor carga acústica para o sistema glotal.

Foram feitas análises qualitativas e quantitativas para a observação das diferenças entre as duas formas de pronúncia da vogal. Entre as análises qualitativas estão a observação do espectro densidade de potências e da modelação fonte-filtro.

As análises quantitativas foram feitas através da relação da suavidade espectral entre altas e médias frequências e através da relação das potências em altas e médias frequências.

Dentre os quatro índices propostos, três deles se mostraram efetivos para a população de voluntárias femininas, porém não houve índice efetivo para a população de voluntários masculinos.

Este estudo não substitui as análises clínicas invasivas ou análises perceptivo-auditivas para detecção de nasalidade na voz, mas se mostra como um passo em direção a futuras análises não-invasivas com menor custo e que terão a mesma confiabilidade que as primeiras.

Palavras-chave: nasalidade, acoplamento nasal, quantificação, análise não-invasiva, espectro do sinal de voz, parâmetros acústicos.

ABSTRACT

The objective of this study was to detect and quantify the nasality on a sustained oral and nasal vowel /a/ by analyzing the Power Spectral Density of a voice signal and the source-filter modeling of the vocal tract. The elected vowel is the Brazilian /a/ due to its lower acoustic load to the glottal system.

Qualitative and quantitative analysis were performed to differentiate the oral and nasal sustained vowel. Power Spectral Density and source-filter modeling were the qualitative analysis methods.

And quantitative analysis was performed by spectral flatness high-to-mid frequencies ratio and by power high-to-mid frequencies ratio.

Among the four proposed indices, three were effective in the female volunteer population, but no effective index was found for male volunteer population.

This work cannot be used to substitute clinical invasive analysis or perceptive-auditive analysis, but it's a step forward in non-invasive low cost analysis, that will be as reliable as the first ones.

Keywords: nasality, nasal coupling, quantification, non-invasive analysis, voice signal spectrum, acoustic parameters

SUMÁRIO

1. INTRODUÇÃO.....	1
2. A VOZ HUMANA	3
2.1 PRODUÇÃO DO SINAL DE VOZ.....	3
2.2 MECANISMO VELOFARÍNGEO E NASALIDADE	5
3. MODELO ACÚSTICO.....	8
3.1 SISTEMA FONTE-FILTRO	8
3.1.1 MODELAGEM DE SISTEMAS LINEARES.....	9
3.1.1.1 MODELAGEM DE UM SISTEMA ARMA	11
3.2 PARÂMETROS ACÚSTICOS NO DOMÍNIO DA FREQUÊNCIA.....	12
3.3 EFEITOS DO ACOPLAMENTO NASAL NOS PARÂMETROS DE ANÁLISE.....	13
4. AVALIAÇÃO DA NASALIDADE NO SINAL DE VOZ.....	15
5. MATERIAL E MÉTODO	16
5.1 AMOSTRAS DE VOZ.....	16
5.2 ANÁLISE DO SINAL DE VOZ	17
5.3 CARACTERÍSTICAS ESPETRAIS.....	19
6. RESULTADOS	21
7. DISCUSSÃO E CONCLUSÕES.....	29
8. REFERÊNCIAS BIBLIOGRÁFICAS.....	31

LISTA DE FIGURAS

Figura 2.1: Subsistemas da produção do sinal de voz.	3
Figura 2.2: Mecanismo das pregas vocais para a produção do pulso glotal	4
Figura 2.3: Formação do sinal de voz	5
Figura 2.4: Exemplos de formantes das vogais /u/, /a/ e /i/	5
Figura 2.5: Padrões de fechamento velofaríngeo.....	6
Figura 3.1: Esquema do Sistema Fonte-Filtro na produção do sinal de voz.....	8
Figura 3.2: Modelos Lineares.....	9
Figura 5.1 Interface do software EXPStudio Áudio Editor	17
Figura 5.2: Espectro Densidade de Potência de uma vogal /a/ oral.	18
Figura 5.3: Ressonâncias de uma vogal /a/ oral.	19
Figura 6.1: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 14 – Feminino).....	25
Figura 6.2: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 12 – Feminino).....	26
Figura 6.3: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 08 – Masculino)	27
Figura 6.4: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 06 – Masculino)	28

LISTA DE TABELAS

Tabela 2.1: Descrição e função dos subsistemas de produção da voz.	3
Tabela 6.1: Características da vogal /a/ oral de voluntárias femininas.	21
Tabela 6.2: Características da vogal /a/ nasal de voluntárias femininas.	22
Tabela 6.3: Características da vogal /a/ oral de voluntários masculinos.....	23
Tabela 6.4: Características da vogal /a/ nasal de voluntários masculinos.	24

1. INTRODUÇÃO

A produção da voz humana envolve um complexo e preciso sistema que se subdivide em subsistemas de órgãos. O principal órgão do sistema fonatório é a laringe, por onde passa o ar que sai dos pulmões e lá é sonorizado através das vibrações das pregas vocais e articulados pela língua, palato, dentes e lábios. A articulação da voz é controlada pelo sistema nervoso central e periférico.

O mecanismo velofaríngeo controla o acoplamento do trato nasal ao trato vocal e, portanto, controla também a intensidade que as modificações da cavidade nasal causam no espectro densidade de potência do sinal de voz de uma determinada vogal que pode ser pronunciada de forma oral e nasal.

Todos os subsistemas articuladores da voz e os efeitos do acoplamento da cavidade nasal na fala serão descritos no capítulo 2.

A produção da voz pode ser comparada a um sistema fonte-filtro que, a partir do espectro densidade de potência da mesma, utilizando métodos de filtragem inversa, é possível modelar os parâmetros ARMA (auto-regressivos de médias móveis) desse filtro. Estes parâmetros fornecem ferramentas que podem auxiliar o estudo das modificações que ocorrem no espectro densidade de potência quando há o acoplamento do trato nasal na fala.

O capítulo 3 traz uma série de considerações ao redor do espectro densidade de potência e dos parâmetros acústicos que podem ser extraídos do mesmo. Também é abordada a forma de modelagem de filtro e comportamento de pólos e zeros do sistema. Insere-se por último o conceito de suavidade espectral que também será utilizado no estudo da quantificação de nasalidade em um sinal de voz da vogal /a/.

Atualmente as formas de avaliação de presença de nasalidade na fala são feitas a partir de análises subjetivas ou perceptivo-auditivas. Estas formas de análise clínica apresentam dois tipos de limitações: a primeira limitação diz respeito à dependência dos conceitos pessoais do avaliador a respeito de suas habilidades de percepção, de discriminação e de experiência. A segunda limitação está ligada à impossibilidade de uma quantificação exata ou, no mínimo, aproximada, do acoplamento nasal quando se faz este tipo de análise.

Existe também a possibilidade de utilizar métodos instrumentais como a nasometria e a técnica de fluxo-pressão para se fazer a análise desejada que oferecem

um resultado mais preciso de análise. Porém, são análises instrumentais e possuem um custo atrelado ao mesmo que pode ou não ser alto.

Como alternativa para as análises citadas anteriormente, tem surgido linhas de pesquisa que se utilizam de métodos computacionais que analisam apenas uma amostra de voz, proporcionando dados objetivos e mensuráveis para análise. Estes métodos possibilitam a quantificação na nasalidade e disponibilizam ao paciente um teste não-invasivo e de baixo custo.

A descrição dos atuais métodos utilizados na análise da nasalância e a proposição de alguns métodos computacionais para a análise da voz e quantificação de nasalidade serão discutidos no capítulo 4.

O objetivo deste estudo é desenvolver um método computacional que visa diferenciar e quantificar a nasalidade de um sinal de voz. Propõem-se análises da diferença do espectro densidade de potência entre a vogal /a/ ressoada de formas oral e da mesma vogal ressoada de forma nasal e também a criação de quatro diferentes índices que relacionam as altas frequências e as médias frequências. O material e método, os resultados e a conclusão do estudo serão apresentados respectivamente nos Capítulos 5, 6 e 7.

2. A VOZ HUMANA

2.1 PRODUÇÃO DO SINAL DE VOZ

Para produzir um sinal de voz, é necessária a sincronia de um complexo sistema, subdividido em subsistemas (Behlau, 2001 p.26), esquematizados na figura 2.1 e descritos na tabela 2.1:

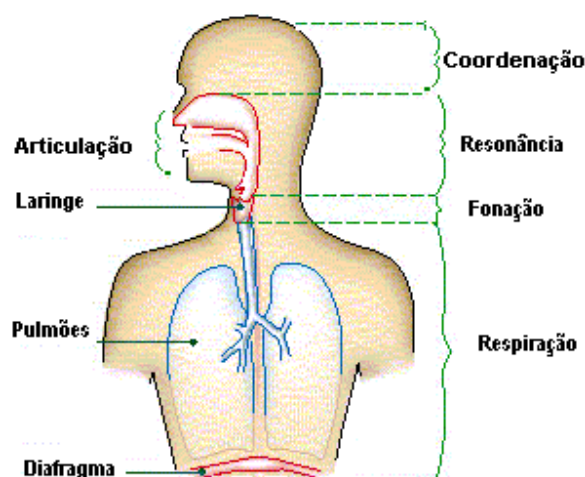


Figura 2.1: Subsistemas da produção do sinal de voz.

Subsistema	Função	Estruturas Associadas
Respiratório	Fonte de Energia	Pulmões
Fonatório	Fonte Vibratória	Pregas Vocais
Ressonância	Fonte de Ressonância	Cavidades Oral e Nasal
Articulatório	Articuladores	Língua, Lábios, Mandíbula, Palato e Dentes
Sistema Nervoso Central e Periférico	Coordenação	Córtex, Estruturas Sub-corticais e Nervos

Tabela 2.1: Descrição e função dos subsistemas de produção da voz.

A geração da voz se dá a partir de um som fundamental gerado pela vibração das pregas vocais que estão localizadas na laringe, principal órgão do sistema fonatório.

Este som básico é denominado *fonte sonora* ou *pulso glotal* e é responsável pela frequência fundamental (Dajér, M. E., 2006). A figura 2.2 mostra um esquema da formação do pulso glotal:

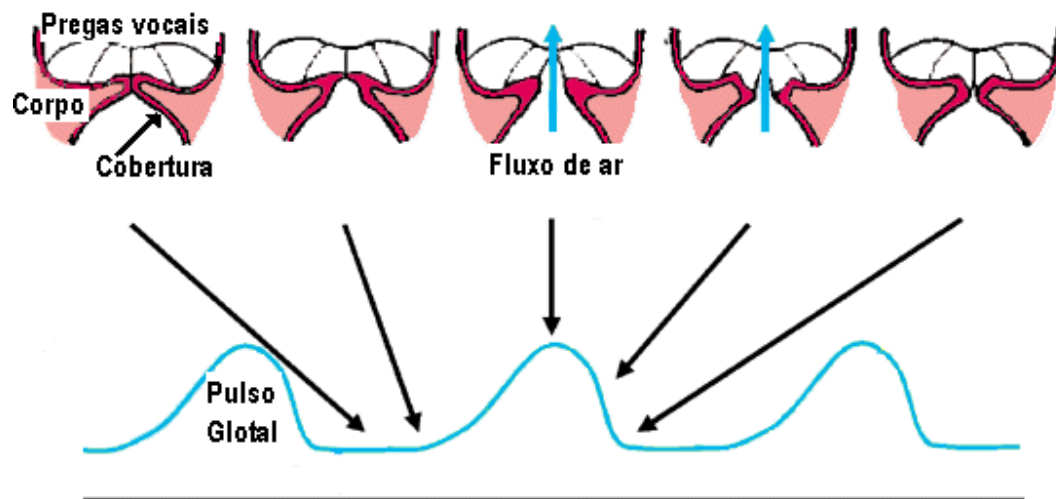


Figura 2.2: Mecanismo das pregas vocais para a produção do pulso glotal (adaptado de <http://biorobotics.harvard.edu/research/heather2.gif>).

A Frequência fundamental (F_0) também recebe o nome de primeiro harmônico. Para os homens a frequência fundamental varia em torno dos 113 Hz e para as mulheres esta varia em torno de 220 Hz (Souza, M. C. Q., 2003).

Entretanto, o som fundamental gerado pelo movimento das pregas vocais possui baixa intensidade e a frequência fundamental e as suas harmônicas, que são os múltiplos inteiros da frequência fundamental, não possuem amplificação. O espectro do pulso glotal apresenta uma intensidade que decresce 12 dB por oitava. A geração da voz humana se dá nas ressonâncias (amplificações, atenuações e modificações) da frequência fundamental e de suas harmônicas (Dajér, M. E., 2006).

O trato vocal atua como um ressonador que modula as frequências do espectro do pulso glotal. Esta ressonância atribui ao espectro densidade de potência um determinado formato para cada vogal pronunciada. Para se distinguir as vogais observando-se o espectro densidade de potência, introduz-se o termo formante, que é atribuído aos picos do espectro de uma vogal. Observa-se então os três primeiros formantes para se distinguir uma vogal da outra (Souza, M. C. Q., 2003).

A figura 2.3 mostra como se dá a formação, a partir do pulso glotal, dos formantes de uma vogal e a figura 2.4 trás alguns exemplos do espectro de algumas vogais:

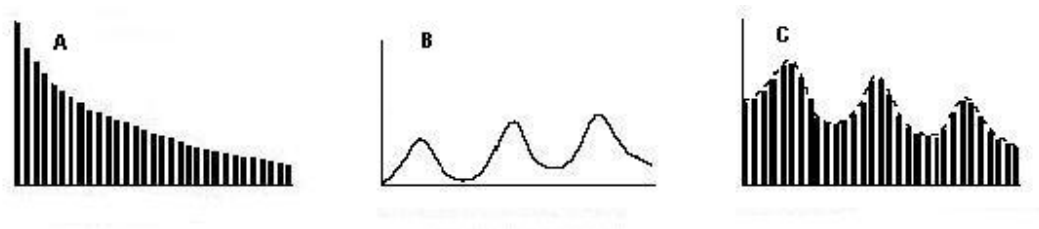


Figura 2.3: Formação do sinal de voz

(A) Pulso glotal sem amplificação. (B) Curvas de ressonância do trato vocal. (C) Espectro resultante da combinação do pulso glotal com a ressonância do trato vocal. (Dajer, M. E., 2006)

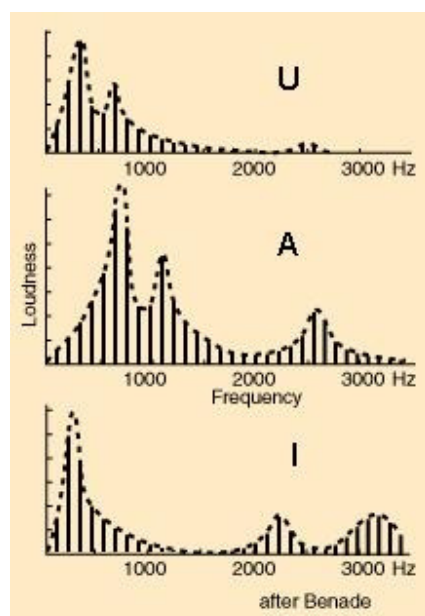


Figura 2.4: Exemplos de formantes das vogais /u/, /a/ e /i/

Adaptado de <http://hyperphysics.phy-astr.gsu.edu/hbase/music/vowel.html#c4>

2.2 MECANISMO VELOFARÍNGEO E NASALIDADE

As estruturas que compõem o Mecanismo Velofaríngeo (MVF) são as paredes laterais e posterior da faringe e o véu palatino (Souza, M. C. Q., 2003). O objetivo deste mecanismo é controlar a conexão entre a cavidade oral da cavidade nasal.

A separação das cavidades oral e nasal durante a produção de som caracteriza uma ressonância predominantemente oral. Neste caso, o MVF se mantém ativo e sela o acoplamento das cavidades. Por outro lado, o acoplamento das mesmas cavidades caracteriza uma ressonância predominantemente nasal, que se dá quando o MVF se mantém inativo (Kent, 1997).

São quatro os principais padrões de fechamento velofaríngeo: o coronal, em que uma linha orientada coronalmente é formada pela aproximação do palato mole na parede posterior da faringe; o sagital, onde há grande movimento das paredes laterais da faringe, resultando em uma linha movimentada no sentido sagital; o circular, produzido pela movimentação equivalente das paredes laterais da faringe e do véu palatino; e o circular com anel de Passavant, onde pode se observar a formação do anel de Passavant na parede posterior da faringe (Shprintzen, 1995). A figura 2.5 demonstra os tipos de fechamento velofaríngeo:

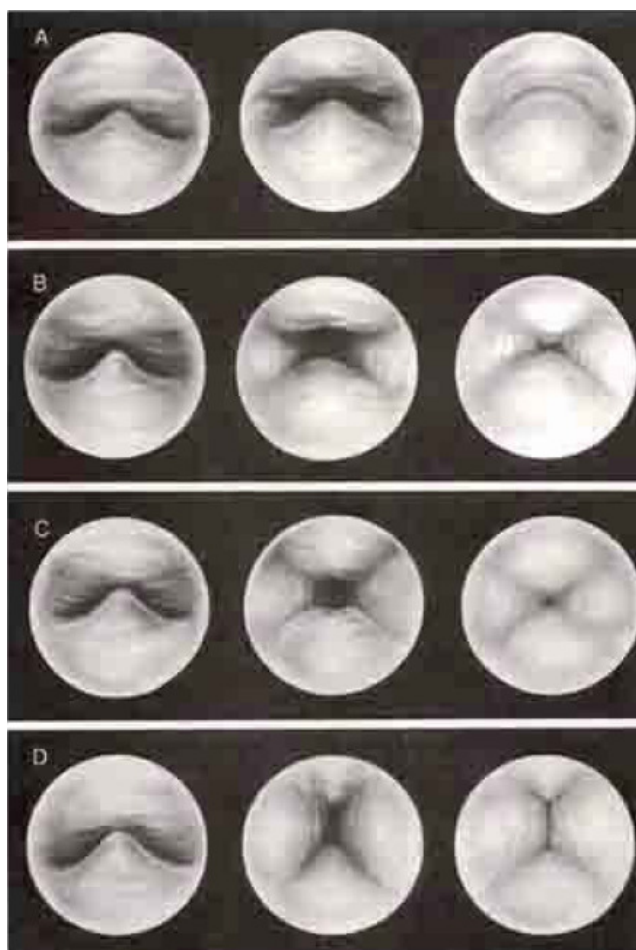


Figura 2.5: Padrões de fechamento velofaríngeo: (A) coronal, (B) circular, (C) circular com anel de Passavant e (D) sagital (Shprintzen, 1995)

A maior parte dos fonemas são ressoados pelas cavidades faríngea e oral, mas os fonemas nasais (tais como /m/, /n/ e /ã/ e outros) além de serem ressoados por estas cavidades também serão ressoados pela cavidade nasal (Souza, M. C. Q., 2003).

Se existir um acoplamento indesejável entre estas cavidades durante uma ressonância oral, a produção da fala pode ser comprometida, gerando alterações como escape de ar nasal audível, ronco nasal, articulação compensatória e hipernasalidade (Peterson-Falzone; Hardin-Jokes; Karnell, 2001). Esta deficiência de natureza estrutural é conhecida como Insuficiência Velofaríngea (IVF) (Fukushiro, A. P., 2007).

Exemplos de condições que podem gerar a IVF são a fissura palatina, as desproporções palato-faríngeas congênitas ou adquiridas e alterações anatômicas dos músculos do palato e do anel faríngeo, que podem, igualmente, gerar sintomas de fala, cuja eliminação depende de tratamento cirúrgico (Shprintzen, 2005; Rocha, 2007).

3. MODELO ACÚSTICO

3.1 SISTEMA FONTE-FILTRO

Conforme visto no capítulo anterior, o sistema de produção de voz pode ser esquematizado como um sistema fonte-filtro, onde o pulso glotal será considerado a fonte e o filtro será representado pelo trato vocal. A figura 3.1 esquematiza esta representação:

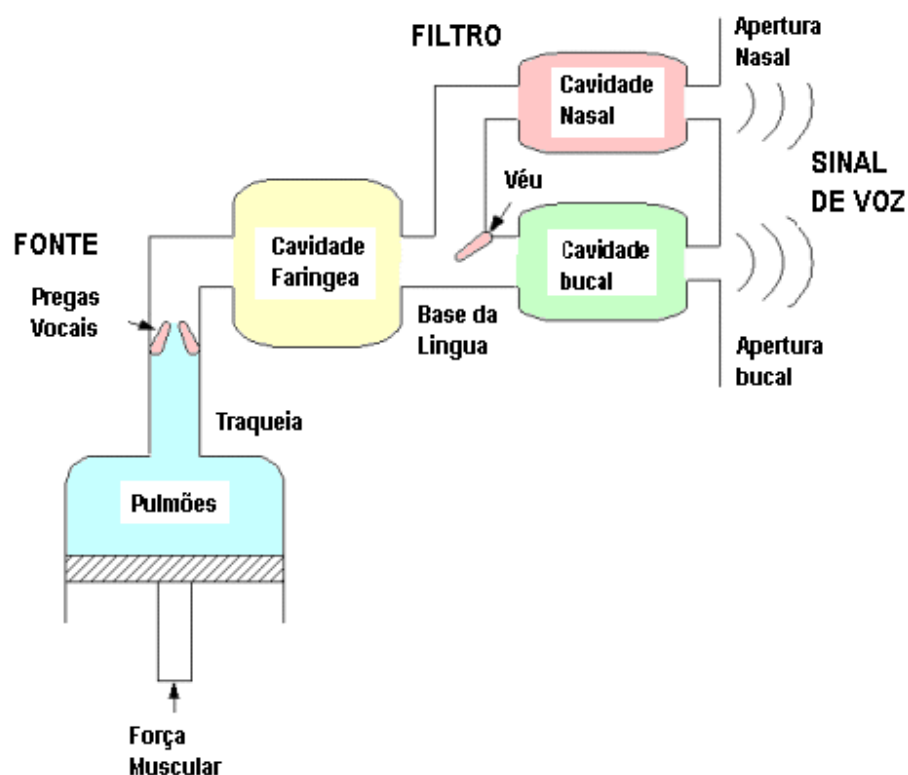


Figura 3.1: Esquema do Sistema Fonte-Filtro na produção do sinal de voz (adaptado de <http://ispl.korea.ac.kr/~wikim/research/speech.html>)

Desta forma, pode-se associar a anatomia vocal a um modelo matemático. Este fato facilita a análise qualitativa e quantitativa das características da fala de determinado indivíduo pois é possível modelar o sistema do trato vocal deste indivíduo através do sinal de voz obtido, utilizando-se o método de filtragem inversa,.

De posse do espectro densidade de potência do sinal de voz e do modelo matemático que rege este sinal, existe uma gama de análises que podem ser realizadas

sobre as propriedades vocais, como análise do formato do espectro e análises dos valores de potências presentes nas frequências deste espectro.

3.1.1 MODELAGEM DE SISTEMAS LINEARES

Assumindo a produção de voz como um sistema fonte-filtro, observa-se um modelo simplificado como o da figura 3.1.

Conhecendo-se o sinal de voz e assumindo o filtro do trato vocal como sendo linear, é possível utilizar técnicas de modelagem linear para se obter os parâmetros de filtro desejados.

Existem três tipos básicos de modelos de filtro com função transferência $H(z)$ sendo uma razão de polinômios. A figura 3.2 demonstra os três tipos de modelos:

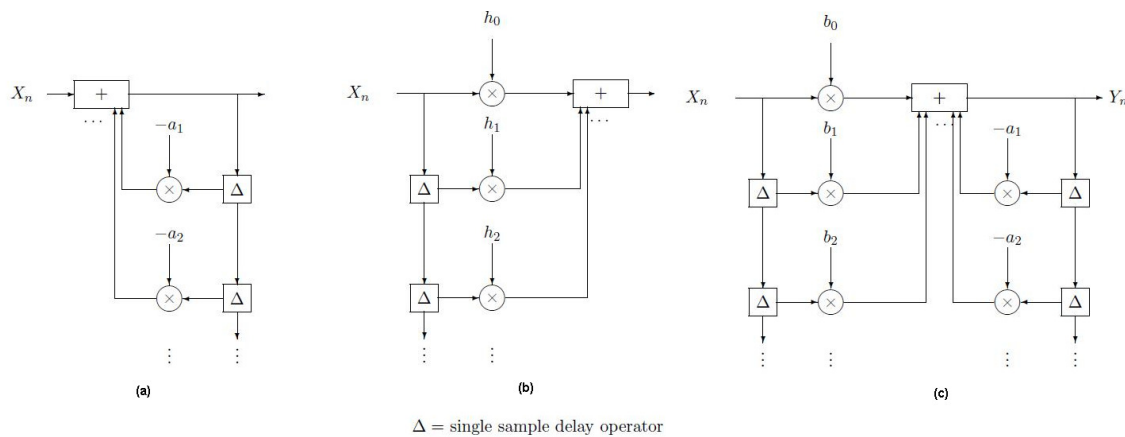


Figura 3.2: Modelos Lineares (a) AR, (b) MA, (c) ARMA
(Adaptado de Gray e Davisso, 2004)

O primeiro tipo de modelo é representado pela figura 3.3(a) e é chamado Modelo Autorregressivo (AR – *autorregressive* em inglês) ou Modelo Tudo-Pólo. É o mais simples e o mais usado entre os três tipos pois as equações que determinam seus parâmetros são lineares. Sua função transferência é dada por:

$$H(z) = \frac{b_0}{A(z)} \quad (3.1)$$

Onde $H(z)$ é o quociente entre ganho do sistema (b_0) e os parâmetros lineares autorregressivos (polinômio $A(z)$).

O segundo tipo é representado pela figura 3.3(b) e é denominado Modelo de Médias Móveis (MA – *moving average* em inglês) ou Modelo Tudo-Zero. É uma equação polinomial cujo denominador é unitário e, apesar de parecer de fácil desenvolvimento, a estimação de seus parâmetros é mais difícil que o modelo autorregressivo. A função transferência deste tipo de modelo é dada pela equação 3.2:

$$H(z) = B(z) \quad (3.2)$$

Onde $B(z)$ é o polinômio com coeficientes que representam os parâmetros lineares de médias móveis.

O último tipo de modelo é demonstrado na figura 3.3(c) envolve um filtro de numerador e denominador polinomiais não-triviais. É chamado Modelo Autorregressivo de Médias Móveis (ARMA – *autorregressive moving average* em inglês) ou Modelo Pólo-Zero. Este modelo é o mais flexível dos três e freqüentemente requer menos parâmetros que um modelo AR ou MA para alcançar um mesmo espectro densidade de potência. Entre os três modelos este é o de mais difícil desenvolvimento. A transferência de filtro deste modelo é dada por:

$$H(z) = \frac{B(z)}{A(z)} \quad (3.3)$$

Neste modelo $H(z)$ possui o seu numerador representado pelos parâmetros MA e o seu denominador representado pelos parâmetros AR.

Neste estudo, por se analisar a presença e quantificação da nasalidade, dar-se-á ênfase ao terceiro modelo, pois ao se acoplar o trato nasal ao trato oral, adicionam-se zeros ao filtro tudo-pólo que representa a cavidade oral (Feng, G. e Castelli, E., 1996) . Este modelo, então, provê maior confiabilidade às análises que serão realizadas.

3.1.1.1 MODELAGEM DE UM SISTEMA ARMA

A equação de diferenças que rege os parâmetros de um modelo geral ARMA pode ser representada da seguinte forma:

$$x[n] + a_1 x[n-1] + \dots + a_P x[n-P] = b_0 w[n] + b_1 w[n-1] + \dots + b_Q w[n-Q] \quad (3.4)$$

Sendo $w(n)$ a entrada (pulso glotal) e $x(n)$ a saída do filtro (sinal de voz), os coeficientes $\{a_k\}$ são os parâmetros AR e os coeficientes $\{b_k\}$ os parâmetros MA e a ordem destes parâmetros é dada por P e Q, respectivamente.

Existem diversos métodos para se fazer a modelagem de um sistema ARMA, porém o método que será descrito é o Método de Durbin (Durbin, 1959). É um método que consiste transformar a modelagem MA de um sistema ARMA na modelagem de dois sistemas AR, pela maior facilidade da modelagem AR.

Inicialmente, adota-se $y[n]$ como sendo a saída de um filtro com transformada z $A(z)$. O intuito é descrever um modelo MA de ordem Q para esta nova seqüência. Para isto, ao invés de se determinar diretamente um modelo MA, designa-se um modelo AR de ordem L maior que Q para esta mesma seqüência $y[n]$. A aproximação do modelo no domínio do tempo discreto será dada por:

$$Y(z) \approx \frac{1}{G^{(L)}(z)} \quad (3.5)$$

Agora, considerando a seqüência $g[0], g[1], \dots, g[L]$, que é a resposta ao impulso do filtro inverso $G^{(L)}(z)$, tem-se os coeficientes do polinômio $G^{(L)}(z)$. Então, se for escolhido um modelo AR de ordem Q para esta nova seqüência, tem-se a aproximação:

$$G^{(L)}(z) \approx \frac{1}{B(z)} \quad (3.6)$$

E então, a partir das equações (3.5) e (3.6), podemos concluir que $B(z)$ é uma aproximação de ordem Q um modelo MA para $Y(z)$ e então, com as aproximações AR e MA, podemos construir um modelo ARMA da forma:

$$H(z) = \frac{B(z)}{A(z)} \quad (3.7)$$

Existem diversos métodos computacionais que podem auxiliar na determinação dos coeficientes AR na modelagem de sistemas a partir de uma determinada sequência de dados.

3.2 PARÂMETROS ACÚSTICOS NO DOMÍNIO DA FREQUÊNCIA

De posse do sinal de voz é possível, através da Transformada Discreta de Fourier, utilizando métodos computacionais de processamento digital de sinais, obter-se o Espectro Densidade de Potência do sinal de voz, que é a representação da potência dos formantes deste determinado sinal em um gráfico. A figura 2.4 mostrou alguns exemplos de espectro densidade de potência de vogais:

A partir do espectro densidade de potência, é possível se obter diversos parâmetros que são descritos na literatura. Alguns exemplos de análise usados são os seguintes:

- LTAS (Espectro Médio de Tempo Longo): Frokjaer-Jensen (1976). É um parâmetro determinado pela média do espectro obtida de um segmento de voz de 80 ms de uma amostra de voz de 45 s.
- HPFR (Relação de Potência em Altas Freqüências): Shoji (1992). Utilizado para analisar o ruído em altas freqüências. Determina-se a partir do sinal de voz normalizando a média das curvas de densidade de potência espectral, em escala logarítmica, com deslocamento de 25,6 ms entre casa estimativa.
- Ra (Harmônicas Relativas): Kojima (1980) definiu este parâmetro para relacionar a energia entre as componentes harmônicas e ruidosas de um sinal construído a partir de trechos com período de freqüência fundamental. Este parâmetro é capaz de diferenciar vozes soprosas de ásperas.
- SNL (Nível Espectral de Ruído): Sansone e Emanuel (1970). É utilizado para diferenciar as vozes normais das ásperas.
- HNR (Relação Harmônicas/Ruído): Yumoto (1982). Quantifica elementos de voz do sinal patológico, tais como: componentes aleatórios na região da freqüência

formante, aparecimento de ruído de alta frequência e perda de magnitude nas frequências harmônicas.

- Hr (Intensidade Relativa de Harmônicas): Hiraoka (1984). Avalia a rouquidão na fala sustentada. Baseia-se na densidade de potência espectral para avaliar a relação entre as componentes harmônicas e as não-harmônicas.
- NNE (Energia de Ruído Normalizada): Kasuya (1986). É um parâmetro no domínio da frequência para estimar o ruído entre as componentes harmônicas, suportando alterações suaves na amplitude e no período.
- SF (Suavidade Espectral): Markel (1976). Faz uma avaliação através do espectro do sinal a quantidade de ruído contida neste. Este parâmetro quantifica a distância do espectro do sinal em relação ao espectro do ruído branco (sinal ideal contendo todas as frequências, considerado constante. Avaliado sobre o sinal residual da voz, obtém-se a Suavidade Espectral do Resíduo (SFR). Também pode-se obter uma avaliação do trato vocal quando realizado o cálculo da Suavidade Espectral do Filtro (SFF), sendo este um parâmetro apropriado para quantificação de nasalidade.
- VLHR (Relação de Baixas para Altas Frequências no Sinal de Voz): Lee, Yang, Kuo (2003). Utiliza análise espectral da voz para avaliar e quantificar obstrução nasal.

3.3 EFEITOS DO ACOPLAMENTO NASAL NOS PARÂMETROS DE ANÁLISE

Desde a década de 40 realizam-se análises espectrográficas com o intuito de se quantificar a nasalidade em um espectro de um sinal de voz. Para tal, utilizam-se sons vocálicos produzidos de maneiras diferentes. Por exemplo, vogais nasais produzidas por falantes normais (nasalidade), vogais produzidas por indivíduos funcionalmente nasais e vogais produzidas por falantes apresentando nasalidade associada com fissura palatal ou deformidade oral similar (Dickson, 1962).

House e Stevens (1956) estudando o espectro acústico de vogais nasais sintetizadas, concluíram que o efeito geral de incorporar o acoplamento nasal alargava e suavizava os picos no espectro vocálico. Mudanças espectrais evidentes ocorrem na região do primeiro formante, sendo razoável assumir que estas mudanças podem ser de grande importância para nasalidade. Exemplos desta diferença serão ilustradas mais adiante, na seção de Resultados. Outra observação feita pelos pesquisadores é que as

características específicas da nasalidade podem diferir para vogais diferentes, bem como para indivíduos diferentes. Relataram também que ao se introduzir o acoplamento nasal houve um aumento na largura de faixa dos formantes vocálicos, observando-se a presença de valores significantes em frequências mais altas.

As características apontadas por Dickson (1962) em seus relatos como relacionadas à nasalidade são: redução na intensidade do primeiro formante, redução da intensidade dos harmônicos de alta frequência, incluindo a região do primeiro formante, uma anti-ressonância e o aumento na largura de faixa dos formantes.

Analisando espectro da voz de alguns indivíduos funcionalmente nasais, Dickson (1962) encontrou as seguintes características: um aumento na largura de banda do formante, um aumento ou diminuição na intensidade dos harmônicos, um aumento ou diminuição na frequência formante, um aumento na frequência fundamental, diminuição na energia das frequências altas e intensidade reduzida do terceiro formante.

Quatro características importantes foram observadas por Schwartz (1968) ao observar o espectro de vogais nasais: redução na intensidade do primeiro formante, a presença de uma anti-ressonância a presença de harmônicos reforçados com frequências onde a energia não é normalmente esperada e a transferência das posições das frequências relativas dos formantes.

Em estudos com indivíduos de fala normal e fala hipernasal, Kataoka (1996) encontrou picos e vales entre o primeiro e segundo formante nas vogais nasais, bem como irregularidade na faixa superior contendo o terceiro formante.

Yang, Lee e Kuo (2003) em seus estudos, desenvolveram um índice que quantifica nasalidade em pacientes, através de uma relação de potências entre altas e baixas frequências.

4. AVALIAÇÃO DA NASALIDADE NO SINAL DE VOZ

A primeira e mais antiga forma de avaliação do sinal de voz é a avaliação perceptivo-auditiva. É uma forma subjetiva de avaliação e o julgamento é feito somente pela percepção do especialista, porém é freqüentemente aplicada na prática clínica devido às sua fácil execução (Genaro; Yamashita; Trindade, 2004; Laczi et al., 2005; Zraick et al., 2000). Mesmo sendo freqüentemente usada, esta forma admite observações sujeitas a erro, particularmente em pacientes com patologia vocal (Lablance, 1991).

Em específico para estudo de nasalidade em um sinal de voz, também encontramos diversos métodos instrumentados de avaliação, dentre eles: a nasofibrosopia, a videofluoroscopia, a nasometria e a técnica de fluxo-pressão, entre outras.

Introduzida por Fletcher em 1970, a nasometria se utiliza de um equipamento chamado nasômetro e permite estimar a ressonância da fala por meio da medida da nasalância, uma grandeza que reflete a quantidade de energia acústica nasal durante a fala (Fletcher et al., 1989; Dalston et al., 1992). A nasalância é calculada pela razão numérica entre a energia acústica total durante a fala (soma das energias acústicas oral e nasal), expressa em porcentagem. Esta análise é não-invasiva e oferece uma complementação à avaliação perceptivo-auditiva, para um diagnóstico mais preciso.

Outra técnica não-invasiva, descrita por Warren e Dubois (1964) é a técnica de fluxo-pressão, que permite calcular a área do orifício velofaríngeo com base no princípio hidrocínético de que a área de um orifício pode ser calculada conhecendo-se a diferença de pressão existente entre os dois lados do orifício e do fluxo aéreo que o atravessa. Permite, dessa forma, quantificar a função velofaríngea durante a fala (Warren, 1989; Trindade e Trindade Júnior, 1996; Trindade et al., 2007).

Baseados nas análises descritas anteriormente e buscando uma forma de minimizar os custos de análises, mantendo sua confiabilidade, especialistas têm buscado cada vez mais implementar análises acústicas de forma computacional utilizando apenas microfone e, através de programas computacionais, fazendo análises espectrais.

5. MATERIAL E MÉTODO

5.1 AMOSTRAS DE VOZ

Amostras de voz de 22 voluntários foram obtidas para este estudo. Destes, 14 eram mulheres e 8 eram homens. Todo o grupo era composto indivíduos com idades entre 18 e 35 anos (média 25,14 anos).

Todos os voluntários eram saudáveis, não exibiam anomalias craniofaciais, não apresentavam problemas auditivos e nem problemas neurológicos, não estavam gripados no dia da gravação e exibiam fonação e ressonância normais. Todos foram conscientizados da finalidade deste estudo.

As amostras de voz foram captadas por um microfone compatível com PC e foram gravadas com auxílio do programa *Análise Acústica 6.0* (Montagnoli, 1999), programa este desenvolvido pelo Departamento de Engenharia Elétrica da Universidade de São Paulo – campus São Carlos.

Os voluntários também foram instruídos a realizar estas pronúncias na posição sentada com os pés apoiados no chão.

Primeiramente, os voluntários foram instruídos a sustentar a vogal /a/ oral de maneira confortável por 5 segundos. Em seguida, pediu-se que sustentassem a vogal /a/ nasal também por 5 segundos.

A gravação das vozes foi feita no Laboratório de Processamento de Sinais do Departamento de engenharia Elétrica da USP. Em detrimento de este laboratório não possuir isolamento acústico, a gravação foi feita no período noturno, onde o movimento é menor e poder-se-ia diminuir a níveis aceitáveis a captação de ruídos indesejáveis.

As amostras foram gravadas com uma frequência de amostragem de 22050 Hz a 16 bits.

Os resultados foram submetidos à análise estatística com o teste t pareado realizado no programa *Origin*TM,

5.2 ANÁLISE DO SINAL DE VOZ

Para se analisar o espectro de voz, utilizou-se o software *MatLab* versão 6.5, que possui ferramentas, complementos e pacotes que possibilitam tal análise de maneira fácil e com alta confiabilidade.

Das amostras de som obtidas, retiraram-se os 2 segundos com maior estabilidade de amplitude do sinal com auxílio do software *EXPStudio Audio Editor* (figura 5.1) e este trecho foi normalizado para uma mesma intensidade (0 dB), dividindo-se todo o trecho pelo seu valor máximo.

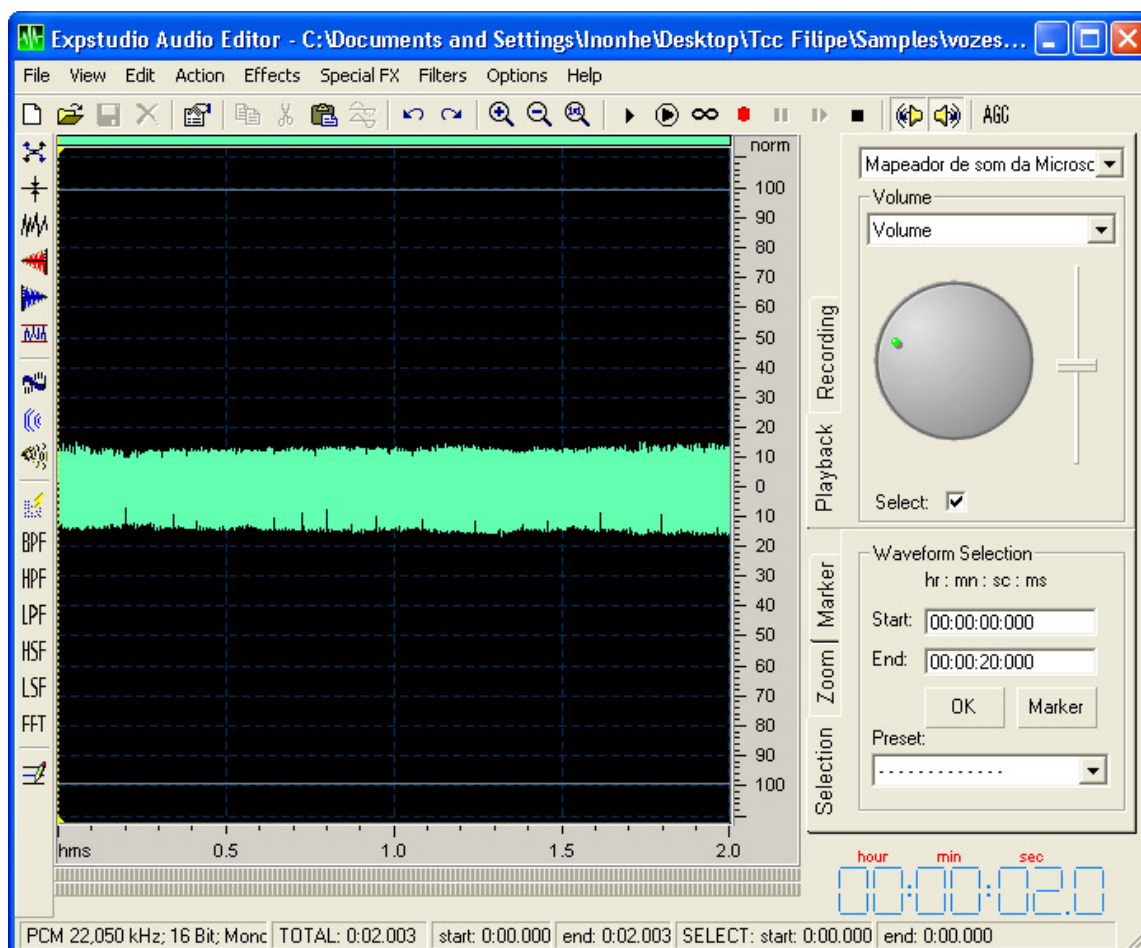


Figura 5.1 Interface do software EXPStudio Áudio Editor

O espectro densidade de potência do sinal resultante da operação anterior foi obtido através de um comando do *MatLab* que proporciona o periodograma médio a partir da estimação espectral de um sinal no tempo discreto usando o método modificado de

Welch. Utilizaram-se janelas retangulares de tamanho 50000. Este periodograma médio evidencia a frequência fundamental e seus harmônicos e suas devidas intensidades, conforme mostra a figura 5.2:

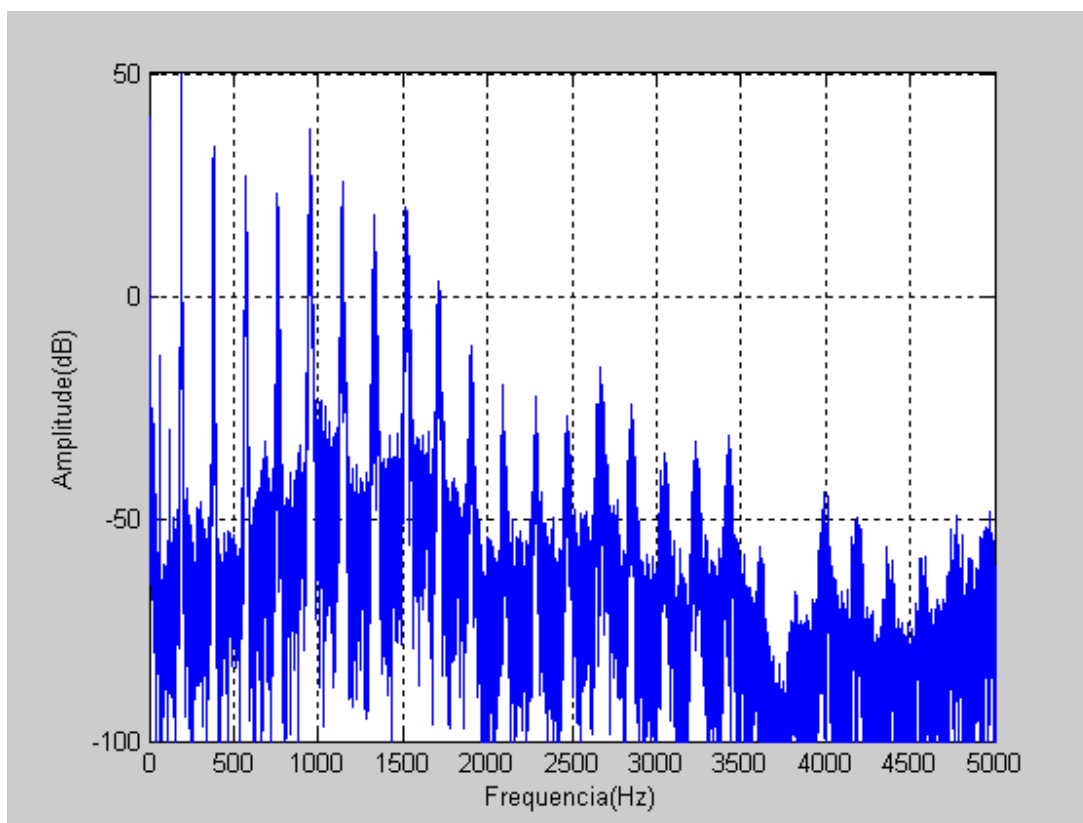


Figura 5.2: Espectro Densidade de Potência de uma vogal /a/ oral.

Através também do *MatLab* foi possível obter-se os parâmetros de filtro ARMA conforme procedimentos descritos no capítulo 3. Com estes parâmetros pôde-se evidenciar a resposta do filtro ao impulso, o que facilita a visualização dos formantes da vogal analisada. Utilizou-se 20 parâmetros MA e 30 parâmetros AR. Estes números de parâmetros foram obtidos por observação da melhor resposta com relação ao periodograma obtido anteriormente. A figura 5.3 demonstra um exemplo das ressonâncias de filtro evidenciando os formantes de uma vogal /a/ oral.

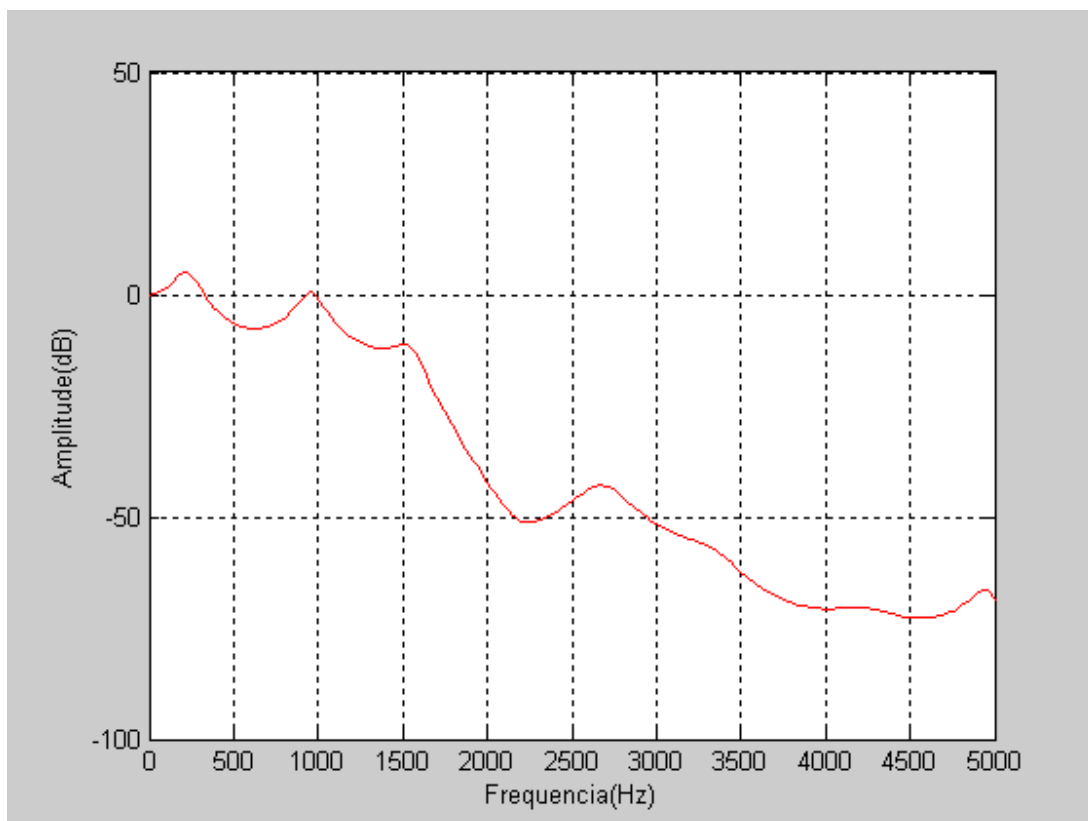


Figura 5.3: Ressonâncias de uma vogal /a/ oral.

De posse dos parâmetros ARMA do filtro do trato vocal, o gráfico no plano imaginário dos pólos e zeros do sistema e a frequência que representam e também se fazem importantes na análise de algumas características dos sinais de voz.

5.3 CARACTERÍSTICAS ESPETRAIS

Para os sinais de voz analisados, extraíram-se as seguintes características: SNL, SF (de todas as frequências, de baixas, de médias e de altas frequências), valores de pólos e zeros de filtro e relações de potências de faixas de frequências.

Todas foram obtidas computacionalmente com base na literatura e, através das mesmas, foi possível fazer algumas observações e associações que são o alvo de discussão deste estudo.

Feitas as devidas análises, propõe-se que exista, entre altas e médias frequências, alguma relação de suavidade espectral ou relação de potências, que

diferencie as pronúncias da vogal de forma nasal e de forma oral, pois observou-se que, na pronúncia nasal, existe um aumento de potência em altas frequências (3500 Hz – 4500 Hz) e uma diminuição da potência (aumento do vale) em médias frequências (1000 Hz – 2800 Hz), por causa da presença de uma ressonância na primeira faixa e uma anti-ressonância na segunda. Espera-se também um aumento do valor da potência na frequência fundamental nas vozes femininas. Tal aumento não é notado nas vozes masculinas, não se observando grande diferença nesta frequência.

Utilizou-se então os valores do periodograma e do espectro de filtro para se realizarem cálculos para se descobrir qual era mais efetivo e qual evidenciava mais a presença e quantidade de nasalidade.

Propuseram-se então 4 índices de relação diferentes: a Relação de Potências do Periodograma (RPP), a Relação de Suavidade Espectral do Periodograma (RSEP), a Relação de Potências de Filtro (RPF) e a Relação de Suavidade Espectral de Filtro (RSEF).

A RPP relaciona a soma das potências no periodograma das altas frequências e das médias frequências. Como se espera um resultado maior para resultados nasais nas vozes femininas, calcula-se também o valor da frequência fundamental e multiplica-se essa relação pelo valor da potência desta frequência.

A RSEP relaciona a suavidade espectral no periodograma das altas frequências e das médias frequências.

RPF relaciona a soma das potências dadas pelo espectro de filtro das altas para as médias frequências. Neste procedimento a relação é dada diretamente, sem utilizar o valor do primeiro formante.

Finalmente, a RSEF relaciona a Suavidade espectral do filtro da faixa de altas frequências e da faixa de médias frequências.

6. RESULTADOS

A partir do proposto no capítulo anterior, analisaram-se as amostras de voz e montaram-se as seguintes tabelas:

Na tabela 6.1 encontram-se os valores dos parâmetros avaliados para as vogais orais do grupo feminino. Para RSEP os valores variam de 0,1114 a 0,6745; para RPP os valores variam de 0,0004 a 5,6834; para RSEF os valores variam de 0,9211 a 272,7179 e para RPF os valores variam de 0,6493 a 1,1181. Os valores médios e o desvio padrão encontram-se na tabela.

Indivíduo	RSEP	RPP	RSEF	RPF
1	0,2881	0,3284	43,7398	0,8247
2	0,4614	0,9285	49,2307	0,7346
3	0,2802	0,1241	101,792	1,1181
4	0,2688	1,2075	47,9813	0,7941
5	0,1985	0,4949	146,9636	0,8539
6	0,4415	0,0483	25,4583	0,9865
7	0,4827	1,446	3,9224	0,5784
8	0,4042	5,6834	6,8353	0,6493
9	0,1114	4,86E-04	272,7179	1,0649
10	0,6745	0,9914	0,9595	0,6879
11	0,4913	1,0922	0,9211	0,653
12	0,4052	0,0299	47,7411	1,0252
13	0,3402	0,0188	45,1982	1,0798
14	0,3953	0,1956	63,7545	0,9098
Média	0,374521	0,899249	61,22969	0,8543
Desvio Padrão	0,141027	1,466991	73,16245	0,179688

Tabela 6.1: Características da vogal /a/ oral de voluntárias femininas.

Na tabela 6.2 encontram-se os valores dos parâmetros avaliados para as vogais nasais do grupo feminino. Para RSEP os valores variam de 0,2576 a 0,7611; para RPP os valores variam de 0,0059 a 35,6604; para RSEF os valores variam de 0,746 a 74,5279 e para RPF os valores variam de 0,4217 a 0,7615. Os valores médios e o desvio padrão encontram-se na tabela.

Indivíduo	RSEP	RPP	RSEF	RPF
1	0,4224	6,8718	18,2607	0,488
2	0,2848	0,0206	60,0464	0,7615
3	0,4574	35,6604	13,7231	0,6613
4	0,2576	1,9814	74,5279	0,7054
5	0,2809	0,8433	15,376	0,7417
6	0,6426	5,3679	3,0263	0,5024
7	0,3688	6,126	9,0837	0,4217
8	0,4553	1,0668	6,9647	0,5616
9	0,3021	0,0059	74,2979	0,6931
10	0,7611	5,022	0,5963	0,6282
11	0,6366	2,7732	0,746	0,5395
12	0,4352	0,9277	3,7676	0,6651
13	0,478	1,813	5,2137	0,6683
14	0,4934	13,4206	22,9639	0,6237
Média	0,4483	5,850043	22,04244	0,618679
Desvio Padrão	0,150323	9,319346	26,81273	0,101552

Tabela 6.2: Características da vogal /a/ nasal de voluntárias femininas.

Na tabela 6.3 encontram-se os valores dos parâmetros avaliados para as vogais orais do grupo masculino. Para RSEP os valores variam de ,2455 a 0,7179; para RPP os valores variam de 0,0294 a 1,1081; para RSEF os valores variam de 0,9736 a 19,3259 e para RPF os valores variam de 0,689 a 1,3504. Os valores médios e o desvio padrão encontram-se na tabela.

Indivíduo	RSEP	RPP	RSEF	RPF
1	0,6251	0,1376	3,048	0,8306
2	0,5912	0,0294	0,9736	0,9513
3	0,392	1,1081	6,4766	0,689
4	0,4469	0,4389	5,1005	0,6929
5	0,2455	0,0924	19,3259	0,7637
6	0,7179	0,0328	2,5243	1,052
7	0,3257	0,05	9,3226	0,8037
8	0,7645	0,0176	3,5914	1,3504
Média	0,5136	0,23835	6,295363	0,8917
Desvio Padrão	0,188817	0,377634	5,865376	0,223007

Tabela 6.3: Características da vogal /a/ oral de voluntários masculinos.

Na tabela 6.4 encontram-se os valores dos parâmetros avaliados para as vogais nasais do grupo masculino. Para RSEP os valores variam de 0,2282 a 0,6068; para RPP os valores variam de 0,06 a 0,4364; para RSEF os valores variam de 3,7341 a 40,3888 e para RPF os valores variam de 0,7169 a 1,0401. Os valores médios e o desvio padrão encontram-se na tabela.

Indivíduo	RSEP	RPP	RSEF	RPF
1	0,4151	0,1996	7,74	0,9365
2	0,3922	0,4364	3,7341	0,7169
3	0,3906	0,06	5,9002	0,8455
4	0,5315	0,2859	6,1216	0,8871
5	0,2282	0,2371	27,5856	0,8303
6	0,2826	0,0731	40,3888	0,8748
7	0,2523	0,0495	9,3043	1,0401
8	0,6068	0,2835	4,2903	0,7264
Média	0,387413	0,203138	13,13311	0,8572
Desvio Padrão	0,133246	0,136218	13,43506	0,105964

Tabela 6.4: Características da vogal /a/ nasal de voluntários masculinos.

Nas figuras seguintes, são demonstrados alguns espectros que demonstram bem a diferença entre a ressonância oral da ressonância nasal da vogal /a/:

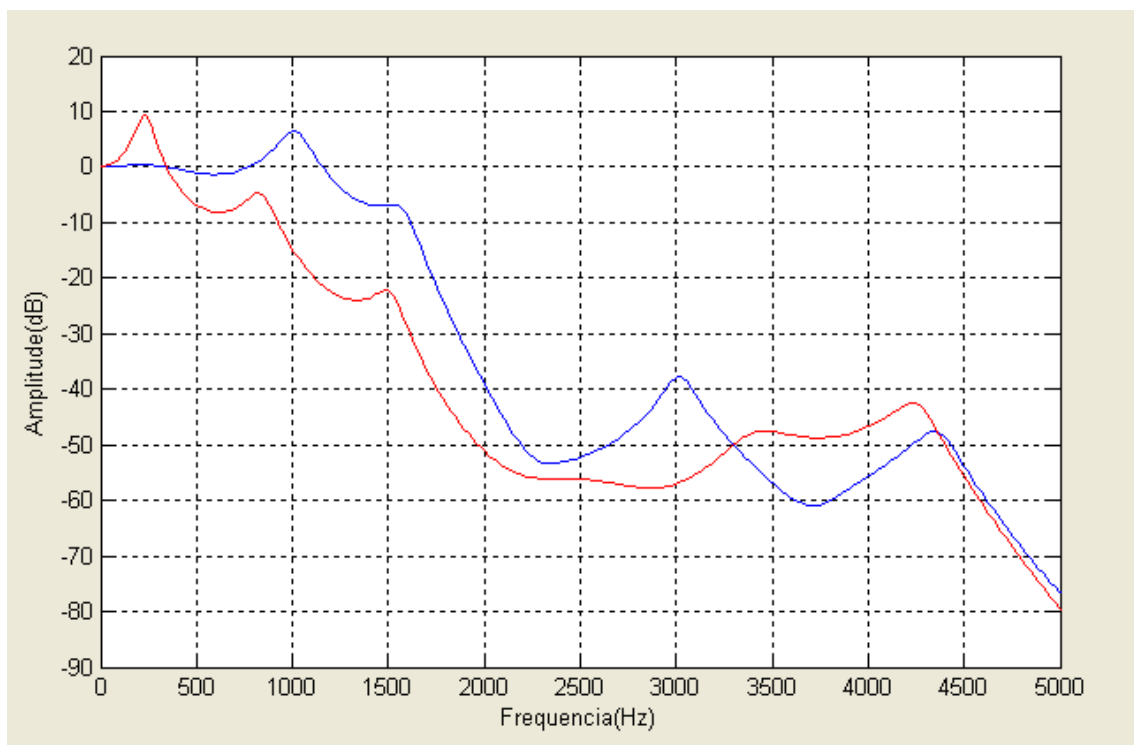


Figura 6.1: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/
(Indivíduo 14 – Feminino)



Figura 6.2: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 12 – Feminino)

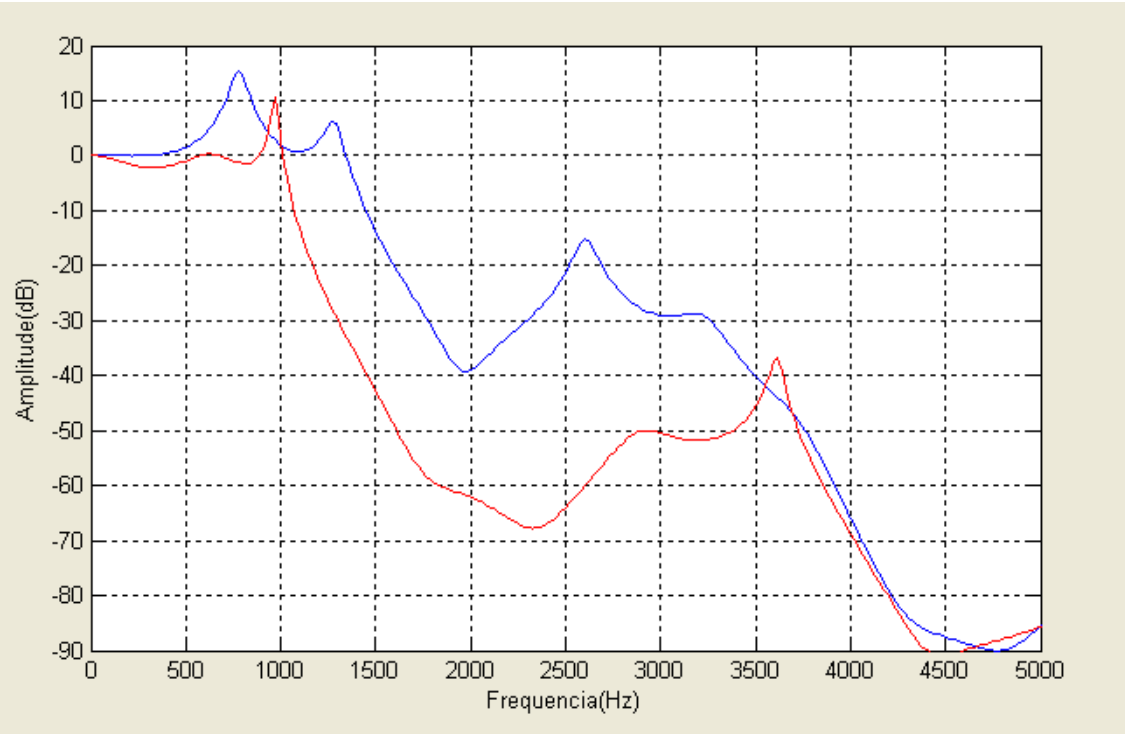


Figura 6.3: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 08 – Masculino)

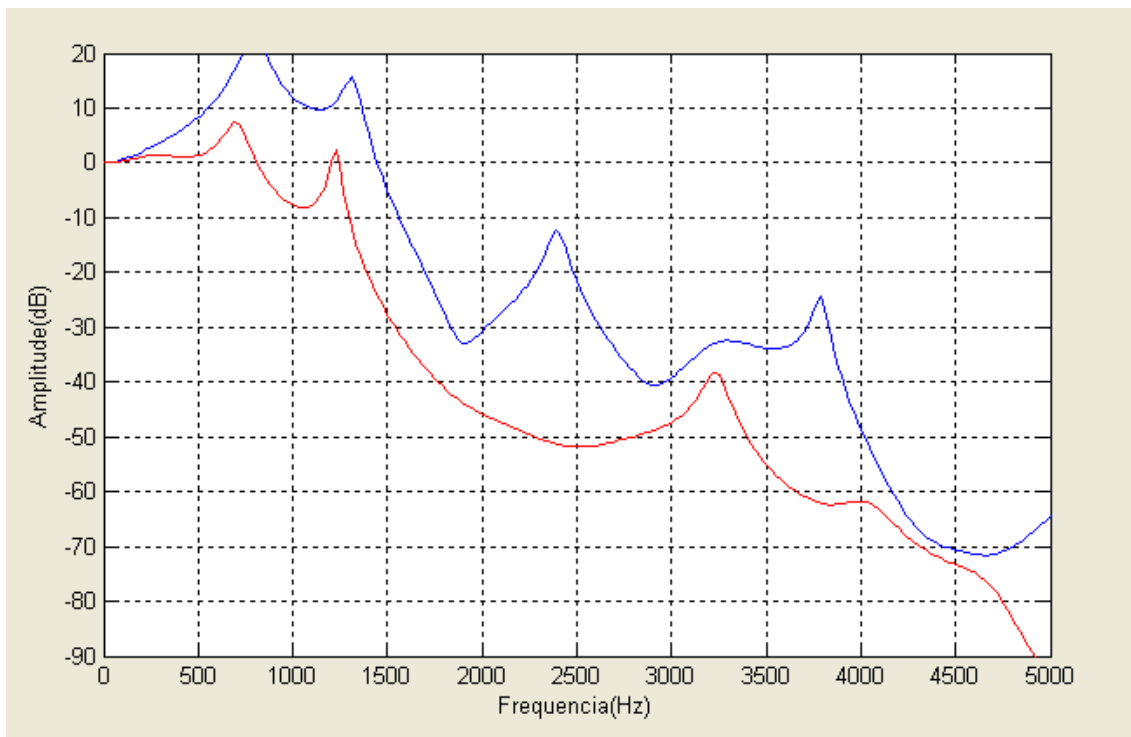


Figura 6.4: Comparação das Ressonâncias Oral (azul) e Nasal (vermelho) da Vogal /a/ (Indivíduo 06 – Masculino)

7. DISCUSSÃO E CONCLUSÕES

Notam-se, qualitativamente, as diferenças espectrais entre as ressonâncias oral e nasal da vogal /a/.

As figuras 6.1 a 6.4, escolhidas por evidenciarem bastante estas diferenças, levam a concluir, primeiramente, que com o acoplamento do trato nasal, a vogal /a/ nasalada possui uma anti-ressonância na faixa de frequência de 2200 Hz a 3000 Hz. Já a vogal /a/ pronunciada de forma oral possui uma ressonância nessa faixa. Esta característica foi observada em estudo semelhante realizado anteriormente (Souza, M. C. Q., 2003), com uma faixa de frequência maior do que a obtida neste estudo.

Característica significativa para as vozes femininas (figuras 6.1 e 6.2) é que se evidencia um aumento na amplitude da frequência fundamental, com relação à ressonância puramente oral.

Uma terceira característica é o aumento do “vale” que existe entre as frequências de 1000 Hz a 2800 Hz, quando existe acoplamento nasal na pronúncia da vogal.

Por último, concluiu-se ao se realizar a análise espectral que, apesar de haver acréscimo de zeros com o acoplamento nasal, também verifica-se que esta cavidade introduz um pólo em frequências mais altas (aproximadamente 3000Hz).

Com relação às características quantitativas demonstradas nas tabelas 6.1 a 6.4 e com o auxílio do programa computacional *Origin*TM, realizou-se o teste *t* pareado que mostra se existe diferença significativa entre duas amostras de um mesmo tomadas em momentos ou ocasiões diferentes (pronúncia vogal e nasal) para uma mesma população. O teste foi realizado com relevância de 5%, fato este que previne erros estatísticos do tipo 1 em 95% dos casos. De posse destes dados, obtiveram-se as seguintes conclusões:

- O teste *t* pareado realizado no programa *Origin*TM apontou que os parâmetros com mudanças significativas são o RSEP, o RSEF e o RPF das vozes femininas;
- O mesmo teste indicou que não há índices com mudança relevante para a população masculina. A causa deste fato pode ser o pequeno número de indivíduos masculinos sendo que algum dos índices pode ter relevância caso haja aumento na população de vozes masculinas analisadas.

- Dos índices significativos podemos concluir que:
 - Quanto maior o valor da RSEP, maior o valor da nasalidade. A partir de 0,45 pode-se suspeitar que exista a presença de nasalidade no sinal de voz;
 - Quanto menor o valor da RSEF, maior o valor da nasalidade. Se encontrado um valor menor ou igual a 20, existe grande possibilidade de que exista nasalidade no sinal de voz;
 - Quanto menor o valor da RPF, maior o valor de nasalidade. Um valor de RPF igual ou abaixo de 0,6 pode indicar presença de nasalidade.

A partir disto, conclui-se que, com uma cuidadosa análise do espectro densidade de potência e utilizando os parâmetros RSEP, RSEF e RPF em vozes femininas, obtém-se informações importantes sobre a possível existência de nasalidade e o grau que existe de acoplamento nasal ao trato vocal na pronúncia da vogal /a/.

Em estudo semelhante (Souza, M. C. Q., 2003), verificou-se que também existe uma variação do parâmetro Jitter, que é a perturbação da frequência em um sinal de voz, quando há o acoplamento do trato nasal, porém sem uma faixa definida de valores, mas que pode ser um parâmetro de grande valia em estudos futuros sobre a nasalidade na vogal /a/.

Ressalta-se, por último, que o custo destas análises é baixo e tais análises, em conjunto com análises perceptivo-auditivas realizadas por especialistas, podem gerar conclusões objetivas sobre a nasalidade. Estudos mais aprofundados embasados neste também podem fornecer informações mais precisas.

8. REFERÊNCIAS BIBLIOGRÁFICAS

BEHLAU, M. A Voz. O livro do especialista. Vol. I. Revinter, 2001.

COLEMAN, J. Introducing speech and language processing, Cambrige University Press, 2005.

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

DALSTON, R. M.; WARREN, D. W., DALSTON, E. T. A preliminary study of nasal airway patency and its potential effect on speech performance. Cleft Palate Craniofacial Journal, 1992.

DICKSON, D. R. An acoustic study of nasality. Journal of Speech and Hearing Research, June, 1962.

FENG, G.; CASTELLI, E. Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization. Journal of the Acoustic Society of America, June, 1996.

FLETCHER, S. G. Theory and instrumentation for quantitative measurement of nasality. Cleft Palate Craniofacial Journal, April, 1970.

FLETCHER, S. G.; ADAMS, L. E.; MCCUTCHEON, M. J. Cleft Palate speech assessment through oral-nasal acoustic measures. In: BZOCH, K.R., editor. Communicative disorders related to cleft lip and palate. III Ed. Boston: Little-Brown, 1989.

FORKJAER- JENSEN, B.; PRYTZ, S., Registration of Voice Quality. B&K Technical Review, 1976.

FUKUSHIRO, A. P. Análise perceptiva, nasométrica e aerodinâmica da fala de indivíduos submetidos à cirurgia de retalho faríngeo para correção da insuficiência velofaríngea. Tese (Doutorado) – Hospital de Reabilitação de Anomalias Craniofaciais, Universidade de São Paulo, Bauru, 2007.

GENARO, K. F.; YAMASHITA, R. P.; TRINDADE, I. E. K. Avaliação clínica e instrumental da fissura labiopalatina. In: FERREIRA, L. P.; BEFI-LOPES, D. M.; LIMONGI, S. C. O. (Edit.) Tratado de Fonoaudiologia. São Paulo: Rocca, 2004.

GRAY, R. M.; DAVISSON, L. D. Statistical signal processing, Cambridge University Press, 2004.

HIRAOKA, N.; KITAZOE, Y.; UETA, H.; TANAKA, S.; TANABE, M. Harmonic intensity analysis of normal and hoarse voices. Journal of the Acoustic Society of America, 1984.

HOUSE, A. S.; STEVENS, K. N. Analog studies of the nasalization of vowels. J. Speech Hear Disord, 1956.

KASUYA, H.; OGAWA, S.; MASHIMA, K.; EBIHARA, S. Normalized noise energy as a acoustic measure to evaluate pathologic voice. Journal of the Acoustic Society of America, 1986.

KATAOKA, R.; et al. Spectral properties and quantitative evaluation of hypernasality in vowels. Cleft Palate-Craniofacial Journal, Jan., 1996.

KATAOKA, R.; WARREN, D. W.; ZAJAC, D.J.; MAYO, R.; LUTZ, R. W. The relationship between spectral characteristics and perceived hypernasality in children. Journal of the Acoustic Society of America, May, 2001.

KAY, S. M. Modern spectral estimation: Theory and application, Prentice-Hall, 1988.

KENT, R. D. The speech sciences. San Diego: Singular, 1997.

KOJIMA H.; GOULD, W. J.; LAMBIASE, A.; ISSHIKI, N. Computer analysis of hoarseness. *Acta Otolaryngologica*, 1980.

LABLANCE, G. R. et al. Advances in non-invasive measures of vocal acoustic. *Ear, Nose and Throat Journal*, 1991.

LACZI, E.; SUSSMAN, J. E.; STATHOPOULOS, E. T.; HUBER, J. Perceptual evaluation of hypernasality compared to HONC measures: The role of experience. *Cleft Palate Craniofacial Journal*, 2005.

LEE, G.; YANG, C. C. H.; KUO, T. B. J. Voice low tone to high tone ratio – A new index for nasal airway assessment. *Chinese Journal of Physiology*, 2003.

LEE, G.; YANG, C. C. H.; WANG, C.; KUO, T. B. J. Effect of nasal descongestion on voice spectrum of a nasal consonant-vowel. *Journal of Voice*, 2005.

LINDQVIST-GAUFFIN, J.; SUNDBERG, J. Acoustic properties of the nasal tract. *Quarterly Progress and Status Report*, 1972.

MARKEL, J. D.; GRAY JR., A. H., *Linear Prediction of Speech*, Berlin: Springer Verlag, 1976

NARECE, I. L. Nasalância de crianças com fissura labiopalatina e nasalidade de fala normal: uma comparação dos dialetos mineiro e paulista. Dissertação (Mestrado) – Programa de Pós Graduação Interunidades em Bioengenharia (EESC/FMRP/IQSC), Universidade de São Paulo, São Carlos, 2007.

ROCHA, D.L., Tratamento cirúrgico da insuficiência velofaríngea. In: TRINDADE I. E. K., SILVA FILHO, O. G., organizadores. *Fissuras Labiopalatinas: uma abordagem interdisciplinar*. São Paulo: Santos, 2007.

SANSONE, F. E.; EMANUEL, F. W. Spectral noise levels and roughness severity ratings for normal and simulated rough vowels produced by adult males. *Journal of Speech and Hearing Research*, 1970.

SCHWARTZ , M. F., The acoustic os normal and nasal vowel production. Cleft Palate Journal, 1968.

SHPRINTZEN, R. J. Instrumental assessment of velopharyngeal valving. In: SHPRINTZEN, R. J.; BARDACH, J. Cleft palate speech management: A multidisciplinary approach. St. Louis: Mobsy, 1995.

SHPRINTZEN, R. J. The velopharyngeal mechanism. In: BERKOWITZ S., editor. Cleft lip and palate: diagnosis and management. New York: Springer, 2005.

SHOJI, K. , et al. High-frequency power ratio of breathy voice. Laryngoscope, 1992.

SOUZA, M. C. Q. Características espectrais de nasalidade. Dissertação (Mestrado – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2003.

THERRIEN, C. W., Discrete random signals and statistical signal processing. Califórnia: Prentice-Hall, 1992.

TRINDADE, I. E. K.; YAMASHITA, R. P.; GONÇALVES, C. G. A. B. Diagnóstico instrumental da disfunção velofaríngea. In: Trindade, I. E. K.; Silva Filho, O. G., organizadores. Fissuras labiopalatinas: uma abordagem multidisciplinar. São Paulo: Santos, 2007.

TRINDADE, I. E. K.; TINDADE JÚNIOR, A. S. Avaliação funcional da inadequação velofaríngea. In: CARREIRÃO, S.; LESSA, S.; ZANINI, A. S. Tratamento das fissuras labiopalatinas. II Ed. Rio de Janeiro: Revinter, 1996.

WARREN, D. W.; DUBOIS, A. B. A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech. Cleft Palate Journal, 1964.

WARREN, D. W. Aerodynamics assessment of velooaryngeal performance. In: BZOCH, K. R., editor. Communicative disorders related to cleft lip and palate. III Ed. Boston: Little-Brown, 1989.

YUMOTO, E.; GOULD, W. J.; BAER, T. Harmonics-to-noise ratio as a index of the degree of hoarseness. *Journal of the Acoustic Society of America*, 1982.

ZRAICK, R. I.; LISS, J. M.; DORMAN, M. F.; CASE, J. L.; LA POINTE; BEALS, S. P. Multidimensional scaling of nasal voice quality. *Journal of Speech, Language, and Hearing Research*, Aug., 2003.