

# **Determinação dos Limiares de Normalidade dos Parâmetros Acústicos da Voz**

Luciana Mara de Oliveira Andrade

Dissertação apresentada ao Programa de Pós-Graduação Interunidades em Bioengenharia / Escola de Engenharia de São Carlos / Faculdade de Medicina de Ribeirão Preto / Instituto de Química de São Carlos, da Universidade de São Paulo para obtenção do título de Mestre em Bioengenharia.

ORIENTADOR: Prof. Dr. Marcos Grellet

São Carlos – SP  
2003

*Dedico este trabalho a DEUS que em  
sua infinita Bondade permitiu que mais  
esta etapa da minha vida fosse  
concluída.*

## **AGRADECIMENTOS**

*Ao Prof. Dr. Marcos Grellet pela dedicação e orientação, sem as quais não seria possível realizar este sonho.*

*Ao Prof. Dr. José Carlos Pereira pela imensa colaboração, paciência, confiança e ensinamentos não só para a realização deste trabalho, mas também como exemplo de vida.*

*Ao Dr. Marcelo de Oliveira Rosa por suas sugestões e esclarecimentos.*

*Ao Dr. Orivaldo Lopes da Silva pelo incentivo e conhecimentos transmitidos.*

*Aos amigos Parê, Granato, Maira e Mori pelo companherismo e acolhida no laboratório LIM.*

*À Daniel Espanhol Razera pela amizade, ajuda e sugestões que tanto contribuíram para a concretização deste trabalho.*

*À Jussara Melo Vieira pela inestimável amizade conquistada.*

*À Paula Carvalho Viana pelo apoio e incentivo.*

*Aos meus pais Aulus e Ana, minha irmã Rose e meu futuro sobrinho Enzo pelo incentivo e confiança.*

*Ao Departamento de Otorrinolaringologia da Faculdade de Medicina de Ribeirão Preto/ Universidade de São Paulo.*

*À Fapesp pela concessão de auxílio financeiro para o desenvolvimento deste trabalho.*

*À todos que direta ou indiretamente contribuíram para a realização deste feito.*

## **SUMÁRIO**

<b>LISTA DE FIGURAS</b>	<b>I</b>
<b>LISTA DE TABELAS</b>	<b>II</b>
<b>LISTA DE ABREVIATURAS</b>	<b>III</b>
<b>LISTA DE SÍMBOLOS</b>	<b>V</b>
<b>RESUMO</b>	<b>VI</b>
<b>ABSTRACT</b>	<b>VII</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
<b>2 OBJETIVOS</b>	<b>3</b>
<b>3 REVISÃO DE LITERATURA</b>	<b>4</b>
<b>3.1 VOZ</b>	<b>4</b>
<b>3.2 PRODUÇÃO VOCAL</b>	<b>7</b>
<b>3.3 ANATOMIA VOCAL</b>	<b>9</b>
<b>3.4 FUNÇÕES DA LARINGE</b>	<b>12</b>
3.4.1 Função Respiratória	12
3.4.2 Função Deglutitória	13
3.4.3 Função Fonatória	13
<b>3.5 ANATOMIA DAS PREGAS VOCAIS</b>	<b>13</b>
<b>3.6 VIBRAÇÃO DAS PREGAS VOCAIS</b>	<b>14</b>
3.6.1 Freqüência Fundamental	17
3.6.2 Intensidade Vocal	18
3.6.3 Qualidade Vocal	18
<b>3.7 PARÂMETROS ACÚSTICOS</b>	<b>20</b>
3.7.1 Parâmetros Acústicos no Domínio do Tempo	21
3.7.2 Parâmetros Acústicos no Domínio da Freqüência	24

<b>4 METODOLOGIA</b>	31
<b>5 RESULTADOS E DISCUSSÃO</b>	34
<b>6 CONCLUSÕES</b>	45
<b>7 REFERÊNCIAS BIBLIOGRÁFICAS</b>	46

**LISTA DE FIGURAS**

Figura 1: Posicionamento das pregas vocais	8
Figura 2: Esqueleto da Laringe	10
Figura 3: Sinal vocal no domínio do tempo	20
Figura 4: Sinal vocal no domínio da frequência	21
Figura 5: Histograma da frequência fundamental do sexo masculino	35
Figura 6: Histograma da frequência fundamental do sexo feminino	36
Figura 7: Histograma do jitter	37
Figura 8: Histograma do shimmer	38
Figura 9: Histograma do SFR	39
Figura 10: Histograma do SFF	40
Figura 11: Histograma do P.A	41
Figura 12: Resíduo de voz normal para o PA	42
Figura 13: Resíduo de voz patológica para o PA	42
Figura 14: Histograma do EX	43

**LISTA DE TABELAS**

Tabela 1: Tabela com análise estatística	44
Tabela 2: Tabela comparativa de resultados	44

**LISTA DE ABREVIATURAS**

APQ - Quociente de perturbação de amplitude

Br - Razão de soproidade

BRI - Índice soproidade

CPP - Proeminência de pulso cepstral

Et al - Colaboradores

EX - Coeficiente de excesso

$F_0$  - Frequência fundamental

$F_c$  - Limiar de alta frequência

FFT - Espectro por Transformação de Fourier

FPQ - Quociente de perturbação de frequência

H/L - Razão da média da energia espectral menor que 4 kHz e maior/igual que 4 kHz

H1A - Relação entre o 1º. harmônico e o 2º. harmônico

HFPR - Relação de potência em altas frequências

HNR - Razão harmônico/ruído

Hr - Intensidade harmônica relativa

LTAS - Espectro médio de longo termo

NNE - Energia do ruído normalizado

P/A - Relação entre a amplitude de pico pela amplitude média do valor absoluto do sinal

PA - Amplitude da autocorrelação do sinal residual

PPF - Fator de perturbação de pitch

PPQ - Quociente de perturbação do período de frequência

ppvv - Pregas vocais

Ra - Razão da energia dos harmônicos

RAP - Perturbação média relativa da frequência



RPK - Pearson  $r$  no pico de autocorrelação

S/N - Razão sinal/ruído

SFF - Suavidade espectral do filtro inverso

SFR - Suavidade espectral do sinal residual

$T_p$  - Período fundamental médio

## LISTA DE SÍMBOLOS

/a/ - Vogal a

/e/ - Vogal e

/i/ - Vogal i

dB - Decibéis

Gb - Gigabyte

Hz - Hertz

Mb - Megabyte

p – Nível de Significância do Teste Estatístico

## **RESUMO**

Existe uma gama de pesquisas destinada à análise digital da voz, cujo intento é o de propiciar ferramentas que possam auxiliar no processo diagnóstico e terapêutico. O intercâmbio entre engenharia e áreas da saúde, como medicina e fonoaudiologia, tem apresentado expansiva importância para um melhor entendimento e aprimoramento de um “saber” mais comum, onde a aplicabilidade de dados quantitativos associados a dados qualitativos acresce conhecimentos nestas duas vertentes.

O intuito deste trabalho é o de utilizar dados quantitativos através de uma ferramenta de análise acústica da voz e avaliar sua acuracidade. Para tanto, fez-se exames clínicos e avaliações perceptivas em 130 voluntários.

Os resultados obtidos conjoinam com a proposição deste estudo, sendo expressivamente promissores com a exatidão da inter-relação dos achados, com o qual pode-se concluir a veracidade da ferramenta utilizada.

**ABSTRACT**

The diagnosis and therapeutic process of the voice has incorporated digital voice analysis as a useful tool for the specialists.

The interchange between engineering and health sciences has improved the understanding of this interface area and the association between quantitative data and qualitative evaluation enhanced both.

The purpose of this work is to use the acoustical analysis of voice in order to evaluate how accurate is that tool compared with clinical and perceptual diagnostics. A set of 130 subjects was used to evaluate this comparison.

The results are consistent and match favorably the clinic and perceptual achievements.

## 1 INTRODUÇÃO

Para poder discorrer sobre a voz, faz-se necessário situá-la enquanto elemento de relevada importância no processo de comunicação humana, visto que, é um componente da linguagem oral e é, também, um dos principais vínculos do relacionamento entre as pessoas.

A voz tem origem na laringe, ou seja, na vibração das pregas vocais, decorrente da passagem do fluxo aéreo originado nos pulmões. Esta voz, então produzida, possui características próprias que variam de acordo com o sexo, a pessoa e com a faixa etária, além de refletir o estado e o comportamento laríngeos, caracterizando o que se chama de qualidade vocal. Assim, vê-se que pela qualidade vocal, pode-se conhecer e/ou inferir sobre as condições anátomo-fisiológicas das estruturas laríngeas.

Devido à complexidade da voz, as avaliações subjetiva e objetiva são importantes para descrevê-la. Essas avaliações podem ser utilizadas como meio de auxílio no diagnóstico e no tratamento das doenças que acometem a laringe.

Além de ser a primeira ferramenta de um clínico, a avaliação perceptivo-auditiva vocal, ou seja, aquela realizada subjetivamente, é fundamental e insubstituível por ser a mais acessível e não necessitar de aparatos. O que não descarta a importância das avaliações objetivas, pois estas se complementam com as subjetivas dando mais consistência às constatações clínicas. Para isto, o confronto de achados perceptivos e acústicos faz-se necessário, estabelecendo-se, portanto, em casos em que ambos não

direcionam para o mesmo diagnóstico, um falso positivo, quando o referencial subjetivo não detecta problemas vocais e a análise acústica o acusa. O contrário também pode acontecer, ou seja, auditivamente perceber-se alteração vocal e a análise acústica não detectar problemas vocais, gerando-se um falso negativo.

Diante disto, estudos para aprimoramento dos programas de análise acústica assumem grande destaque para o clínico, principalmente por induzirem a um aprofundamento no conhecimento das características objetivas da voz.

## 2 OBJETIVOS

Determinar os valores médios dos parâmetros acústicos da voz: *jitter*, *shimmer*, *suavidades espectrais do filtro e do sinal residual*, o *índice de vocalização* e o *coeficiente de excesso*.

Avaliar as taxas de falso positivo e falso negativo da ferramenta de análise acústica da voz, através da inspeção videoestroboscópica e avaliação perceptivo-auditiva.

### **3 REVISÃO DE LITERATURA**

#### **3.1 VOZ**

A avaliação subjetiva da voz iniciou-se no século XIX e é essencial ao fonoaudiólogo, o qual, utiliza-se deste instrumento para avaliação e posterior reabilitação do sujeito. Nesta avaliação, o treinamento e a sensibilidade do especialista se faz mister, pois o mesmo é capaz de detectar e reconhecer alterações vocais devido à habilidade desenvolvida através do treinamento de sua audição, e a sensibilidade para perceber se tais alterações não estão intrinsecamente ligadas ao seu estado emocional momentâneo.

A voz é o resultado de fenômenos fisiológicos que, juntos, determinam a emissão acústica, o que subsidia a objetividade no que tange à voz. Análises, são realizadas em diferentes domínios Nieto (1995), para discriminar as diferentes particularidades da voz. A avaliação objetiva dos vários aspectos da voz como por exemplo medidas de perturbação em amplitude e frequência permite uma quantificação da mesma propiciando, desta forma, a determinação de alterações na voz que podem ocorrer ao longo do tempo e mudanças resultantes de intervenções. Assim como o audiograma auxilia o especialista nos diagnósticos e tratamento de possíveis disacusias, um teste objetivo da voz pode auxiliar no diagnóstico de disfonias (ROSA et al., 1999). Devido à natureza numérica, são permitidas a documentação e a comparação entre



resultados. Nota-se que a evolução da tecnologia propicia uma apuração bem mais minuciosa das alterações na voz.

Para melhor diagnosticar alterações vocais, existe uma miríade de ferramentas e pesquisas disponíveis, que devem ser utilizadas por especialistas que tratam da voz. Mais que interesse científico, o foco do clínico está no aprimoramento dos resultados do diagnóstico e do tratamento. Pelo uso seletivo de medidas relevantes, estas podem ser usadas como realimentação na melhoria de técnicas de tratamento.

Dentre as ferramentas, a análise acústica da voz, devido a sua objetividade, tem expressiva utilização. A evolução histórica da mesma, descrita em décadas, facilita a compreensão dos fatos marcantes (BOONE E MACFARLANE, 1988). Seu desenvolvimento se deve, em grande parte, à evolução da tecnologia através dos estudos no campo da Engenharia.

Na década de 20, a análise acústica da voz era realizada utilizando oscilogramas por meio da visualização de ondas ou gráficos de amplitude pelo tempo.

Na década de 40, iniciou-se a utilização da Análise de Fourier com Análise de Henrici para obtenção da análise espectral.

Na década de 50, o Espectrógrafo surgiu revolucionando a metodologia de análise da voz através do espectrograma que é uma gravação visual tridimensional da distribuição de energia em frequência, tempo e intensidade.

Na década de 70, o processamento digital de sinais revolucionou o estudo da voz. Os programas permitiram análises combinadas no domínio do tempo e da frequência. Foram produzidos os Espectrogramas Digitais e Programas de Análise Computadorizados da Voz.

Davis (1979), coloca que a avaliação perceptivo-auditiva da voz fornece informações sobre a qualidade vocal, visto que, patologias laringeas são acompanhadas

freqüentemente por mudanças na mesma. No entanto, esta avaliação pode receber críticas devido sua subjetividade.

A análise acústica do sinal vocal é promissora e vem sendo mais utilizada com a evolução da tecnologia digital, acrescentando dados à avaliação e ao diagnóstico. Sua associação com métodos de imagem como laringoscopias e estroboscopias aumenta o grau de precisão do diagnóstico.

A maioria dos trabalhos realizados com amostras vocais utiliza equipamentos analógicos para gravação da voz em fita magnética, para posterior utilização em estudos no computador. Holmes (1975), analisou a gravação analógica em meio magnético e verificou que mesmo gravadores analógicos de boa qualidade provocam distorções de fase nas baixas freqüências do sinal gravado. Assim é mais interessante a utilização de gravação vocal direta em equipamento digital.

Segundo Rosa (1998), a voz do paciente pode ser captada com uso de microfones tipo cardióide e tipo acelerômetro. O microfone cardióide ou convencional, é mais difundido e para melhor definição deve ser unidirecional, com ganho dinâmico e larga faixa de resposta em freqüência. O microfone acelerômetro ou de contato, capta através da pele a vibração da glote, não é influenciado pela supra-glote, mas pelas camadas do pescoço que atenuam o sinal captado.

Krom (1994), cita que nas técnicas de análise acústica da voz, o sinal vocal pode ser analisado com estabilidade usando vogais sustentadas. A preferência por vogais sustentadas é devida a maior facilidade de controle e padronização da produção da fonação e menor atuação da articulação do que na fala contínua.

Os sinais de voz apresentam diferenças de energia sonora entre as porções inicial e final dos trechos vocálicos devido ao ataque do fluxo de ar pela glote e perda de capacidade volumétrica de ar, respectivamente, sendo necessário definir o trecho

vocálico do sinal digitalizado e eliminar os trechos não vocálicos. Krom (1995), mostrou que estes intervalos no sinal vocal produzem maior perturbação em amplitude e frequência, o que pode caracterizar apenas uma condição de energia acústica do fluxo de ar e não uma evidência patológica. Em decorrência de tal fato, devem ser desconsiderados esses dois momentos (inicial e final) padronizando o tamanho da amostra vocal.

Para melhor compreensão dos programas existentes, destinados à análise de voz, se faz necessário à distinção entre os domínios temporal e frequencial (NIETO et al., 1995).

A análise de voz no domínio do tempo é relacionada à representação gráfica da onda sonora em relação às variações de pressão sonora ao longo do tempo. Neste domínio foram obtidas as formas da onda, análise de Frequência Fundamental, do *jitter* e do *shimmer*.

A análise de voz no domínio da frequência é relacionada à representação gráfica das variações da amplitude das componentes espectrais distintas que compõem a voz. Neste domínio foram obtidos o Espectrograma, o Espectro por Transformação de Fourier (FFT), traçado de formantes, análise de componentes harmônicas e análise de frequência fundamental.

### **3.2 PRODUÇÃO VOCAL**

Por não existir um único órgão destinado à produção vocal, faz-se necessário a ação integrada e harmoniosa de uma série de sistemas, que participam ativamente no processo de vocalização. Dentre os vários sistemas envolvidos, pode-se destacar o

sistema respiratório, adaptado a partir da respiração vital, tornando-se passiva na inspiração e ativo na expiração.

A voz é produzida pelo fluxo de ar expiratório que advém dos pulmões e passa pelas pregas vocais (Figura 1), as quais são encontradas aproximadas colocando-as em vibração. A frequência de vibração é determinada pela elasticidade ou tensão que as pregas vão apresentar durante a produção do som, o que é indicado pelo alongamento e espessura das mesmas. Assim sendo, quanto mais rápido elas vibrarem (frequência alta), mais aguda é a altura percebida da voz. A intensidade da vocalização ou a intensidade percebida da voz é determinada pelo vigor em que as pregas vocais vibram.

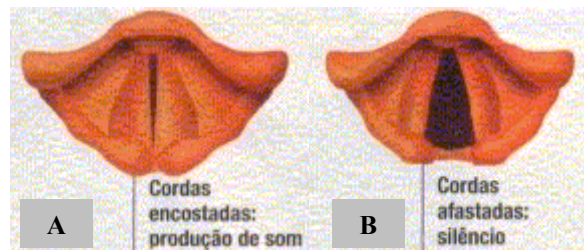


Figura 1: Posicionamento das pregas vocais durante a produção vocal (A) e durante a respiração (B) (adaptação de <http://www.gastroweb.com.br/imgend/laringe.htm>, acessado em 12/04/03)

Os sons vocálicos são produzidos por pulsos de ar quase periódicos gerados na laringe pelo movimento cíclico das pregas vocais que excitam o trato vocal (/a/, /e/, /i/ /o/, /u/). Há, ainda, os sons fricativos produto da passagem turbulenta de ar através de alguma constrição formada no trato vocal (/f/, /v/) e os sons plosivos, que são aqueles produzidos pelo fechamento completo do trato vocal com conseqüente aumento da pressão anterior à obstrução e liberação abrupta desta (/d/, /t/, /p/, /b/).

Assim, vê-se que a voz é uma forma inteligível de som, a qual é produto de uma excitação (gerador) e dos efeitos ressonantes do trato vocal (filtro).

Dos três modos supra citados, os sons vocálicos são os que apresentam, durante a vibração das pregas vocais, características ressonantis completa do trato vocal.

Portanto, do ponto de vista de análise do aparato vocal, este modo de geração de voz contém toda a informação fisiológica sobre o sistema.

### 3.3 ANATOMIA VOCAL

Há oito músculos extrínsecos na laringe, sendo quatro abaixo do osso hióide e quatro acima dele. Por isso são divididos em dois grupos: supra-hióides (digástrico, gênio-hióideo, milo-hióideo e estilo-hióideo) e infra-hióides (tiro-hióideo, esterno-hióideo, omo-hióideo e esternotireóideo). Todos os músculos extrínsecos têm suas fixações laríngeas a partir do osso hióideo. As fixações fora da laringe incluem várias estruturas, tais como a mandíbula, mastóide e estruturas do tórax.

No grupo supra-hióideo, os músculos digástrico (ventre anterior), o gênio-hióideo e o milo-hióideo, quando contraídos, puxam o osso hióideo, e conseqüentemente a laringe, para frente. O ventre posterior do digástrico e do estilo-hióideo quando contraídos puxam o hióideo posteriormente.

No grupo infra-hióideo os músculos esternotireóideo, esterno-hióideo e o omo-hióideo, quando contraídos, puxam a laringe para baixo. O músculo tiro-hióideo contribui para determinar o ângulo da cartilagem tireóidea com relação à cartilagem cricóidea.

Segundo Dângelo e Fattini (1995), a ação dos músculos supra-hióideos com os infra-hióideos mantêm o osso hióideo, propiciando uma base firme para os movimentos da língua.

O “esqueleto” da laringe (Figura 2) é composto de cartilagem hialina, sendo as mais importantes às cartilagens tireóideas, cricóidea e aritenóideas.

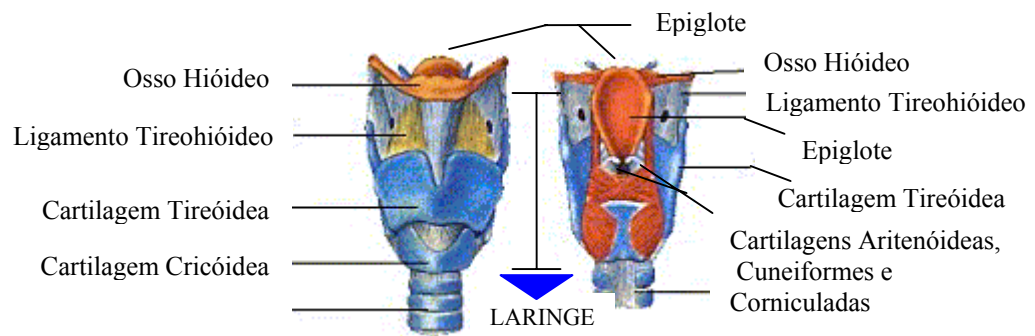


Figura 2: Esqueleto da Laringe  
(adaptação de <http://www.corpohumano.hpg.ig.com.br/respiracao/laringe/laringe.html>, acessado em 12/04/03)

A tireóidea é a maior cartilagem da laringe e é composta por duas lâminas unidas ao meio. Zemlin (1998), coloca que esse ângulo nos homens é de aproximadamente 80° e nas mulheres é de 90°, aproximadamente.

A cricóidea é a segunda maior cartilagem laríngea. A cartilagem tireóidea articula-se com a cricóidea sobre sua superfície póstero-lateral, permitindo assim, movimento de rotação e um leve movimento ântero-posterior.

Há duas cartilagens aritenóideas, cada uma posicionada sobre um lado da linha média, sobre a superfície supraposterior da cartilagem cricóidea. A base horizontal da cartilagem aritenóidea, que se articula com a cricóidea, é côncava.

Segundo Vonlenden e Moore (1961), movimento de rotação ocorre devido à forma dessas superfícies.

Na base de cada cartilagem aritenóidea encontramos três ângulos. O mais anterior projeta-se para dentro da laringe e é chamado de processo vocal, sendo o ponto de fixação posterior da prega vocal; o ângulo póstero-lateral projeta-se para fora da laringe e recebe o nome de processo muscular por ser o local de fixação de vários músculos, tais como o cricoaritenóideo posterior, abductor da laringe, e o cricoaritenóideo lateral, adutor da região anterior das ppvv; o ângulo póstero-lateral não recebe nenhum nome.

A epiglote é uma cartilagem em forma de folha fixada sobre a superfície mesial da cartilagem tireóidea na junção das duas lâminas tireóideas (Figura 2). Essa cartilagem auxilia no direcionamento de líquidos e alimentos ao esôfago durante a deglutição. Na fonação, ela se movimenta durante a produção de diferentes vogais e consoantes, podendo dificultar a visualização das pregas vocais. Entretanto, durante a produção das vogais /i/ e /u/, onde a posição lingual resulta no movimento da epiglote para frente, a visualização das pregas vocais é permitida.

Há cartilagens remanescentes como as corniculadas e as cuneiformes (Figura 2), as primeiras formam o ápice da aritenóidea e as segundas são encontradas dentro da prega ariepiglótica.

São cinco músculos intrínsecos da laringe, cujas fixações encontram-se associadas a estruturas dentro da laringe.

O músculo aritenóideo tem duas partes e posiciona-se entre as duas cartilagens aritenóideas. Uma parte consiste de fibras musculares que percorrem uma direção horizontal. A outra parte consiste de fibras que correm a partir da base de uma aritenóidea até o ápice da outra aritenóidea. O efeito conjunto promove a adução das aritenóideas e fecham a via aérea posterior extrema.

O músculo cricoaritenóideo lateral é pareado, percorrendo os lados e a superfície superior da cartilagem cricóidea para o processo muscular da aritenóidea. Esse músculo puxa o processo muscular anteriormente, balançando a cartilagem aritenóidea medialmente e aduzindo as próprias pregas vocais. Os músculos aritenóideo e o cricoaritenóideo lateral funcionam juntos para aduzir as pregas vocais.

O cricoaritenóideo posterior é o único que abduz as pregas vocais, pois um movimento de curta duração de seus movimentos rapidamente abduz as pregas vocais o suficiente para cessar sua vibração.

O músculo cricótireóideo possui formato de leque e localiza-se entre as cartilagens cricóidea e tireóidea. Sua função consiste em: diminuir o espaço entre a tireóidea e a cricóidea; aumentar a distância entre as cartilagens tireóidea e aritenóidea; aumentar o comprimento das pregas vocais; diminuir sua massa; e aumentar sua tensão, bem como a frequência vocal.

O músculo tiroaritenóideo compõe o corpo das pregas vocais e é par, cada um deles com a forma de um grosso feixe, originam-se no ângulo da cartilagem tireóidea e inserção principalmente no processo vocal. A ação deste músculo é encurtar e aduzir as ppvv, diminuindo a distância entre as cartilagens aritenóideas e tireóidea, tornando-se um feixe mais largo e reduzindo a frequência da voz gerada.

### **3.4 FUNÇÕES DA LARINGE**

A laringe possui várias funções, das quais as mais importantes são: a respiratória, a deglutitória e a fonatória.

#### **3.4.1 Função Respiratória**

Esta função ocupa a maior parte do tempo desse órgão. A abertura da laringe permite a entrada e a saída do ar dos pulmões. Durante o ato inspiratório a laringe é ligeiramente rebaixada, passando a ser tracionada ligeiramente para cima durante a expiração; esse deslocamento é proporcional à intensidade da respiração. A traquéia e o diafragma participam da descida da laringe, contribuindo, junto com o músculo cricoaritenóideo posterior, para o afastamento das aritenóideas e das pregas vocais. Os músculos infra-hióideos contribuem para a descida da cartilagem tireóidea e, conseqüentemente, para a abertura da laringe.



### 3.4.2 Função Deglutitória

Primeiramente a deglutição ocorre na boca e é voluntária, sendo seguida de uma fase involuntária, onde ocorre o fechamento máximo da laringe. Esta é a fase faríngea. Durante a deglutição, na fase faríngea, a laringe se fecha de baixo para cima protegendo as vias aéreas, sendo essa proteção uma função vital.

### 3.4.3 Função Fonatória

A fonação é uma habilidade adquirida pela espécie humana (BEHLAU E REHDER, 1997). Com a evolução, o ser humano aprendeu a coordenar órgãos que, fisiológica e anatomicamente apresentavam como funções primordiais a respiração, alimentação e proteção das vias aéreas inferiores. Deste modo, a voz apesar de sua beleza e aprimoramento, é uma função adaptada e, como tal, depende da integridade e equilíbrio das estruturas do trato vocal.

Ao emitir a voz, as pregas vocais se aproximam da linha média, controlando a saída do ar que advém dos pulmões. O ar, por sua vez, ao passar pela laringe, coloca em vibração as pregas vocais, que estão próximas entre si. As pregas vocais, neste instante, fecham-se e abrem-se numa seqüência rápida, realizando a função fonatória.

## **3.5 ANATOMIA DAS PREGAS VOCAIS**

Considerando-se o funcionamento vocal, o conceito mais importante sobre a estrutura da prega vocal é que esta é um vibrador com várias camadas epiteliais. De um modo geral, pode-se dizer que a prega vocal é composta de mucosa e músculo. A mucosa divide-se em epitélio e lamina própria. A função do epitélio é cobrir e manter a forma da prega vocal. O epitélio é formado por tecido escamoso estratificado.

A lâmina própria, por sua vez, pode ser dividida em três camadas: superficial, intermediária e profunda (HIRANO, 1977). A camada superficial, freqüentemente referida como espaço de Reinke, é frouxa e flexível, sendo essa camada que vibra mais intensamente durante a fonação. A camada intermediária consiste principalmente de fibras elásticas e a profunda é composta de fibras de colágeno podendo ser comparadas às linhas de algodão. Por não haver um limite definido entre as camadas intermediária e profunda, dá-se o nome de ligamento vocal à junção destas, sendo que o número de fibras elásticas diminuem e as de colágeno aumentam em direção ao corpo da prega vocal.

O corpo da prega vocal é composto pelo músculo vocal e quando esse se contrai, funciona como um feixe de elásticos muito rijo. Não há uma delimitação entre as fibras colágenas da camada profunda e as fibras mais superficiais do músculo vocal.

As estruturas descritas podem ser agrupadas em três seções: a cobertura, (consistindo do epitélio e da camada superficial da lâmina própria); a transição (compostas pelas camadas intermediária e profunda da lâmina própria); e o corpo (composto do músculo vocal).

Durante a fonação, o deslocamento da cobertura apresenta uma componente vertical contrária a gravidade produzindo um movimento elipsoidal, enquanto que o corpo oferece estabilidade e o tono muscular de base. A transição, por sua vez, serve de acoplamento entre a camada superficial e o músculo vocal (BEHLAU, 2001).

### **3.6 VIBRAÇÃO DAS PREGAS VOCAIS**

A vibração das pregas vocais pode ser alterada ou gerar novos modos de vibração em decorrência de mudanças assimétricas na massa, elasticidade e tensão nas próprias pregas vocais (ppvv).

Outras alterações, como paralisia dos músculos respiratórios, podem causar pressão subglótica insuficiente, alterando as forças aerodinâmicas que agem sobre as pregas vocais o que, também, alterará as vibrações.

A conceituação de vozes normais e disfônicas é negociável, visto que, implica uma certa subjetividade. A dicotomia existente entre normalidade versus disfonia é um dos temas mais controvertidos na área fonoaudiológica. A palavra normal, que significa segundo a norma quando aplicada ao comportamento humano, defronta-se a questões influenciáveis, como, por exemplo, gosto pessoal, modismo, fatores raciais, culturais e sexuais. Desta maneira, pode-se formular apenas padrões de vozes ditas normais.

Tabith (1995), coloca que normalidade vocal implicam a dependência de vários fatores, como normalidade anatômica e funcional, boa coordenação entre respiração e fonação, boa saúde geral, ausência de problemas psicológicos relevantes e emissão agradável.

São sugeridas por Aronson (1990), três questões para o julgamento de vozes normais: 1ª) se a voz é adequada para oferecer ao ouvinte inteligibilidade de fala; 2ª) se suas propriedades acústicas são aceitáveis; 3ª) se a voz preenche as demandas profissionais e sociais do falante. O autor acresce seu discurso colocando que a variedade vocal é ilimitada e os padrões de adequação vocal são amplos.

Para Behlau (2001), a voz deve ser produzida sem esforço e com conforto, identificando o sexo e a faixa etária a que pertence. A autora propõe a utilização do termo voz adaptada, ao invés de voz normal, pois, a produção vocal deve ser aceitável socialmente e apropriada às características sócio- histórico-culturais do falante.

Encontra-se na literatura uma variedade de classificações destinadas a conceituar as disfonias, dentre as quais, explanar-se-á algumas destas.

Aronson (1980), classifica as disfonias em três grupos principais: disfonias orgânicas, (como tumores, inflamações, problemas endócrinos, traumas e distúrbios neurológicos), distúrbios psicogênicos (como estresse emocional, nódulos, quadros de psicose, falsete mutacional, voz infantil no adulto, iatrogênia e mutismo) e disfonias de origem indeterminada (como a disfonia espasmódica).

As disfonias são classificadas por Brandi (1990), em dois grupos: disfonias comportamentais (como desvios do padrão vocal, conduta vocal hiperfuncional ou hipofuncional) e disfonias estruturais (decorrentes de anomalias congênitas, ou por laringopatias).

Garcia-Tabia e Cobeta (1996), dividem as disfonias em quatro grupos: disfonias de causa orgânica, (podendo ser congênitas ou adquiridas como as traumáticas, inflamatórias, neoplásicas, endocrinológicas, neurológicas e pós-cirurgia laríngea), disfonias de causa funcional, (podendo ser habituais ou psicogênicas), disfonias por lesões associadas (como nódulos, pólipos, edema de Reinke e granuloma de contato) e disfonias de causa psiquiátricas (esquizofrenia, distúrbios afetivos, transtorno bipolar e outras).

Consideradas por Pinho (1998), três formas de classificar as disfonias, a saber: disfonias funcionais, em que há presença de distúrbio vocal com ausência de alterações orgânicas significativas; disfonias orgânicas secundárias, que são disfonias decorrentes do uso inadequado da voz e geram alterações orgânicas consequenciais; e disfonias orgânicas primárias, disfonias que surgem independente do uso da voz.

Behlau e Pontes (1995), classificam as disfonias em três categorias: disfonias funcionais, que são disfonias provenientes de desvios do processo básico de produção vocal, apresentando como consequência mais comum, a fadiga vocal; disfonias orgânico-funcionais, que são disfonias funcionais diagnosticadas tardiamente, gerando

uma lesão secundária, como por exemplo, o nódulo vocal; e disfonias orgânicas, cuja etiologia independe do uso da voz e advém de quadros inflamatórios ou infecciosos, gripes e laringites, alterações vocais por carcinoma da laringe e ou doenças neurológicas como doença de Parkinson e Esclerose Lateral Amiotrófica.

Por haver vertiginosa aceitação no meio fonoaudiológico adotar-se-á, neste trabalho, a classificação ditada por (BEHLAU E PONTES, 1995).

De qualquer forma, os sintomas acústicos percebidos durante o processo patológico da laringe são mudanças na frequência fundamental, na intensidade da voz ou em sua qualidade. Esses sintomas são indicativos de alterações orgânicas e/ou funcionais e a natureza desses sintomas varia para cada paciente e também com o estágio da patologia.

### 3.6.1 Frequência Fundamental

A frequência fundamental de um som vocálico é função da massa, da elasticidade e do comprimento das pregas vocais. Também depende da pressão subglótica e da configuração do trato vocal (carga acústica).

Segundo Nieto et al. (1995), a frequência fundamental é diretamente proporcional à tensão das pregas vocais ou a pressão subglótica e inversamente proporcional à massa e a largura das pregas vocais .

A frequência fundamental é definida como o número de vibrações por segundo produzidas pelas pregas vocais. Em mulheres adultas, a frequência média situa-se ao redor de 200 Hz e em homens adultos, cerca de 100 Hz (BEHLAU E PONTES, 1995). O pitch é considerado a forma perceptiva da frequência fundamental (BEHLAU, 1997).

A constatação de que um paciente tem uma regulação de frequência adequada envolve um julgamento, por um especialista treinado, em comparação com vozes de

peessoas que apresentam tamanho, sexo e idade similar. Nos distúrbios funcionais que envolvem alta frequência, as pregas vocais são muito exigidas no ponto de máximo deslocamento. Este abuso vocal pode produzir laringite, causando o desenvolvimento de nódulos, ou agravando uma patologia existente.

### 3.6.2 Intensidade Vocal

A intensidade depende diretamente da resistência que a glote oferece à passagem de ar, ou seja, uma glote eficiente gera aumento da pressão subglótica, responsável pela variação da intensidade (BEHLAU, 2001).

Fatores como velocidade da emissão do ar e quantidade de ar emitido também interferem na intensidade vocal, e são diretamente relacionados à pressão subglótica, ou seja, quanto maior for a pressão subglótica, maior será a velocidade e a quantidade do ar emitida.

A etiologia de vozes com intensidade fraca pode ser atribuída à coaptação glótica deficiente, gerando assim uma pressão subglótica insuficiente (por exemplo, em casos de paralisia dos músculos respiratórios, ou por paralisia de pregas vocais cuja boa coaptação está afetada).

A forma perceptiva da intensidade é denominada *loudness*.

### 3.6.3 Qualidade Vocal

O termo qualidade vocal, antigamente conhecido como timbre é o que caracteriza e identifica a impressão vocal. Uma degradação na qualidade vocal, geralmente definida como rouquidão, é frequentemente o primeiro, e algumas vezes o único sintoma de disfunção laringeal.

Vários são os termos utilizados para conceituar a voz, levando assim a divergências no que concerne a nomenclatura. A imprecisão da terminologia é tão acentuada que se apropria de vocábulos inerentes a outros sentidos, como por exemplo, ao se dizer que a voz é clara, utiliza-se assim do sentido da visão, ou quando se infere que a voz é rude, sendo esse um atributo do sentido tátil. No entanto, tais terminologias são corriqueiramente usadas.

Sonnimem e Hurne (1992 apud BELHAU e PONTES\*, 1995), nomeiam as vozes em mais de cinquenta termos.

Foi adotado por Behlau e Pontes (1995), vinte e um tipos de vozes. São eles: voz rouca, soprosa, áspera, sussurrada, fluida, gutural, comprimida, tensa-estrangulada, bitonal, diplofônica, polifônica, monótona, trêmula, pastosa, crepitante, infantilizada, virilizada, presbifônica, hipernasal, hiponasal e com nasalidade mista.

Revisando-se a literatura, nota-se o estudo minucioso de diversos autores quanto à notificação dos diferentes tipos de vozes.

Neste trabalho, três tipos de alterações na qualidade vocal são sugeridos fundamentais, pois acredita-se que as vozes rouca, soprosa e áspera podem suprir e englobar toda e qualquer alteração vocal. O intuito de tal proposta está fundamentado em facilitar uma classificação para os profissionais atuantes, para que estes se expressem dentro de um linguajar mais comum, não mais subdividindo-se, mas interagindo entre si, dentro de uma abordagem mais simplista.

Propõe-se estabelecer assim, uma “neo-conceituação” apropriando-se de termos já existentes e consagrados, destacando a concordância deste intento com os correlatos fisiológicos estabelecidos por Pinho (1998), nos três tipos de vozes citados, a saber: *voz rouca*, correspondente à presença de irregularidade vibratória da mucosa das pregas

---

\* BEHLAU, M. & PONTES, P. Avaliação e Tratamento das Disfonias. São Paulo: Lovise, 1995.

vocais durante a produção vocal; *voz áspera*, que ocorre quando há redução ou ausência da onda de mucosa que reveste as pregas vocais; *voz soprosa*, correspondente à presença de ruído de fundo, audível.

Esta visão de uma abordagem simplista somente tem o intuito de propiciar aos profissionais da área e áreas afins uma unificação de termos, não desmerecendo os pesquisadores, nem se atendo à idéia reducionista.

### 3.7 PARÂMETROS ACÚSTICOS

A análise dos parâmetros acústicos evidencia, segundo vários autores, diferenças entre as vozes consideradas normais e patológicas. Diversos pesquisadores propuseram-se a quantificar os parâmetros acústicos vocais, embasados na frequência e na amplitude do sinal vocal.

Para tanto, estudaram o sinal vocal, como citado por Nieto (1995), nos domínios do tempo e da frequência.

No domínio do tempo, vê-se o sinal de voz bidimensionalmente (Figura 3): em um eixo cartesiano tem-se o tempo como variável independente e, no outro, a amplitude em função do tempo.

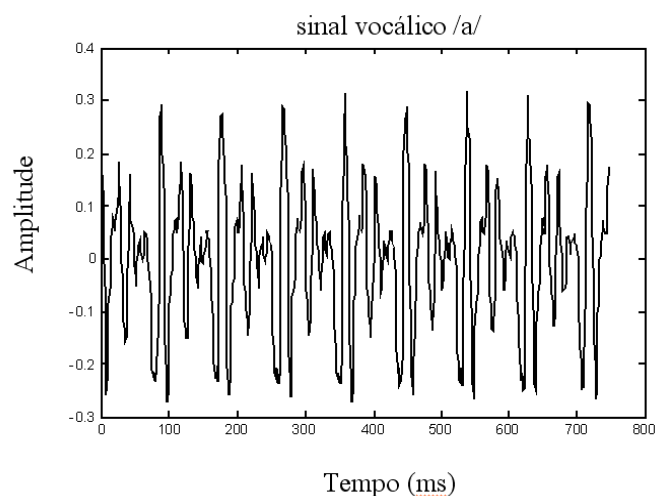


Figura 3: Sinal vocal no domínio do tempo.



No domínio da frequência, o sinal vocal é analisado em suas componentes de frequência, bidimensionalmente, em relação à amplitude, ou seja, é o espectro vocal (Figura 4).

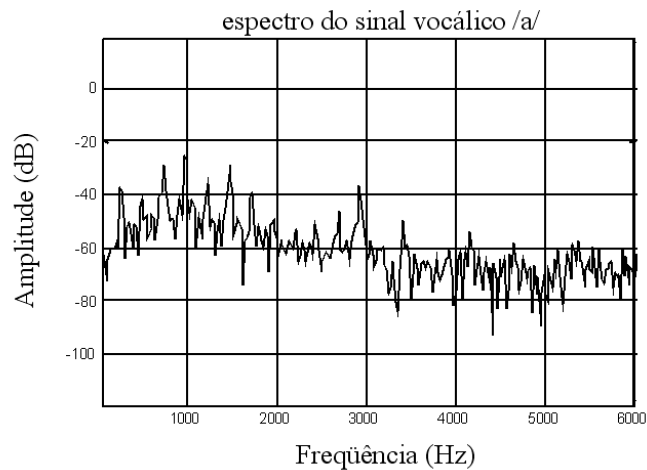


Figura 4: Sinal vocal no domínio da frequência.

### 3.7.1 Parâmetros Acústicos no Domínio do Tempo

Lieberman (1963), verificou através da distribuição das diferenças entre períodos sucessivos, que os casos patológicos apresentavam-se na região desta curva, acima de 0,5 ms. Com isso definiu a medida PFF (fator de perturbação de pitch ou fator de perturbação de frequência), que é o número de vezes que o módulo da diferença entre o período a ser avaliado e o período médio excede 0,5 ms, normalizado pelo período médio. Perturbações de frequência, derivadas a partir das formas das ondas acústicas, podem refletir variações na periodicidade glotal, variações no envoltório da onda ou mudanças na configuração do trato vocal. Perturbações com magnitudes maiores ou iguais a 0,5 ms podem ser induzidas por um decaimento da pressão do ar através da glote e indica variações no padrão vibratório das pregas vocais. As perturbações de frequência que foram medidas para vozes normais são menores do que aproximadamente 0,2 ms. Sujeitos com algum tipo de patologia na laringe podem ter

maiores perturbações de frequência que sujeitos com vozes normais, com o mesmo período médio fundamental. O fator de perturbação é sensível ao tamanho e à localização da alteração tecidual patológica na laringe do sujeito. Para Lieberman, lesões na lâmina própria causam perturbações menores que lesões no músculo.

Koike (1973), desenvolveu a medida RAP (perturbação média relativa da frequência), que considera as estruturas supraglotais, pois conforme o autor, o contato do microfone com a garganta é o melhor indicador da aperiodicidade laríngea. Em última análise é uma dessensibilização do período, através de uma filtragem passa-baixas, usando o valor médio e o janelamento de três períodos da frequência.

Em Takahashi e Koike (1975), as vozes ditas patológicas são percebidas como rouca, áspera e soprosa, porém não são pertinentes para um diagnóstico mais preciso. Para elucidar um diagnóstico por meio da voz, alguns estudos perceptivos são necessários. E um dos procedimentos propostos por Osgood\* (1957), citado pelos autores, é a técnica diferencial semântica. Esta técnica tem sido utilizada por muitos pesquisadores para estudar a impressão auditiva da voz e da fala humana. São considerados quatro fatores perceptivos: fluidez, rouquidão, magnitude e extensão vocal, sendo que culturas de línguas diferentes não interferem no resultado, ou seja, a percepção é a mesma. Após terem feito uma análise perceptiva, fizeram uma comparação por meio de técnicas acústicas envolvendo frequência fundamental ( $F_0$ ), quociente de perturbação de frequência (FPQ) e quociente de perturbação de amplitude (APQ) dos trechos mais estáveis da vogal /a/, sustentada durante, aproximadamente, 1,5 segundos e captadas pelo microfone de contato. O parâmetro FPQ foi definido como a perturbação média normalizada do período da frequência pico-a-pico e o APQ como a

---

\* TAKAHASHI, H. E KOIKE, Y. Some Perceptual Dimensions and Acoustical Correlates of Pathologic Voices. *Acta Oto-laryngologica*, Suppl.338, 1-24, 1975.

perturbação média normalizada do período de amplitude pico-a-pico. Definiram estes parâmetros com base em um janelamento de 11 períodos.

Davis (1979), usando características e sinais obtidos por meio do filtro inverso residual, desenvolveu um perfil vocal para demonstrar informações acústicas sobre a voz e servir como um histórico médico do paciente. Por meio das informações acústicas, Davis demonstrou que as medições poderiam discriminar falantes normais de patológicos. São seis parâmetros acústicos determinados para definir este perfil, a saber: PPQ (quociente de perturbação do período de frequência), APQ (quociente de perturbação da amplitude), EX (coeficiente de excesso), PA (amplitude da autocorrelação do sinal residual), SFF (suavidade espectral do filtro inverso) e SFR (suavidade espectral do sinal residual). Os parâmetros SFF e SFR são estudados no domínio espectral. O PPQ e o APQ são semelhantes ao RAP de Koike e ao APQ de Koike e Takahashi. Visando esclarecer as diferenças entre esses parâmetros acústicos, Davis estudou a influência do tamanho da janela na precisão do resultado. Investigou os benefícios das alterações no tamanho das janelas, encontrando que a janela com cinco períodos produz resultados melhores nas medidas de perturbação APQ e PPQ na discriminação de vozes patológicas ou normais. O parâmetro EX, definido por Cramer\* (1958, *apud* DAVIS 1979), permite quantificar a relação sinal-ruído do sinal residual. O PA pode ser considerado como um índice de vocalização e é definido como a máxima amplitude na função autocorrelação do sinal. Dessa forma, o PA é alto para sons vocálicos, pequeno para sons fricativos sonoros e zero para sons fricativos surdos.

---

\* DAVIS, S.B. Acoustic Characteristics of Normal and Pathological Voices. In Lass, N.J. (ED) Speech and language: Advances in Basic Research and Practice, vol.1, New York Academic Press, 271-335, 1979

### 3.7.2 Parâmetros Acústicos no Domínio da Frequência

Frøkjær-Jensen e Prytz (1976), registraram 22 vozes saudáveis e 50 patológicas (pacientes que apresentavam redução na energia espectral ou seu aumento na região do 1º. formante durante a recuperação) de indivíduos de ambos os sexos em um sistema desenvolvido no próprio laboratório dos autores. Analisaram-nas por meio do espectro médio de longo termo (*Long-time-average-spectra - LTAS*) e experimentaram comparar suas qualidades vocais propondo um novo parâmetro, chamado  $\alpha$ , que é uma medida das relações de intensidade das frequências mais altas e mais baixas do espectro da fala. Ou seja,  $\alpha$  é o resultado da soma das intensidades acima de 1 kHz dividida pela soma das intensidades abaixo de 1 kHz. O parâmetro  $\alpha$  é independente da distância do microfone, do nível de amplificação, porque o espectro acima de 1 kHz está normalizado em relação ao espectro abaixo de 1 kHz. Esses pesquisadores tentaram encontrar novos métodos para LTAS no balanço entre as partes altas e baixas do espectro da fala, afirmando que esse balanço depende da fonte vocal e parece estar correlacionado com o termo “qualidade vocal”, o qual definem como propriedade auditiva, isto é, um aspecto da percepção da voz humana ainda mal elaborado. Demonstraram que o parâmetro  $\alpha$ , bem como o desenvolvimento de instrumentação e métodos, podem ser usados com finalidade diagnóstica e, ainda, como avaliação da terapia vocal, não só pela representação do que ocorreu durante o processo terapêutico mas também pela comparação da qualidade vocal antes e após a terapia de voz.

Kojima et al. (1980), separaram os componentes harmônicos da voz usando um computador com a finalidade de obterem um cálculo objetivo para avaliar quantitativamente a rouquidão. Determinaram o parâmetro  $R_a$ , que é a razão da energia dos harmônicos múltiplos de três pela energia das componentes harmônicas entre estes, consideradas pelos autores como ruído, em uma janela com três períodos de frequência,

pela Transformada de Fourier. Compararam os resultados desse parâmetro com a impressão auditiva de cinco pessoas sem conhecimento prévio sobre voz. A obtenção das amostras (58, no total) foi feita por meio da produção da vogal sustentada /a/. Os resultados calculados mostraram uma forte correlação com a impressão auditiva e este estudo sugeriu que este método de separação dos componentes harmônicos da voz rouca pode ser útil no cálculo quantitativo da rouquidão, permitindo sua graduação em vários níveis.

Yumoto et al. (1982), desenvolveram a razão harmônico-ruído (HNR), que é computacionalmente menos complexa e consome menos tempo para avaliar quantitativamente o grau de rouquidão. Basearam-se na suposição que a onda acústica de uma vogal sustentada consiste de um componente periódico e de um componente ruidoso (média zero na distribuição da amplitude). Em uma cabina acústica registrou-se a vogal /a/ sustentada, emitida em um nível de sonoridade confortável para um grupo de pessoas com vozes normais na faixa etária de 19 a 60 anos de idade (22 homens e 20 mulheres) e um grupo de sujeitos de 21 a 68 anos de idade com várias patologias laringeas pré e pós operatório (12 homens e 8 mulheres). Na tela do monitor do computador foram exibidas duas formas de onda: uma original na qual foi determinado o período mais estável para análise (600 ms) e uma onda filtrada (passa-baixa de 500 Hz) para determinação da frequência pelo método semi-automático de cruzamento por zero. Por meio de um filtro banda-estreita obteve-se espectrogramas da fonação original e a extração do componente ruidoso. Foram usados 50 períodos para calcular a HNR. Os resultados mostraram correlação entre componentes ruidosos da fonação, mesmo quando o indivíduo apresentou rouquidão severa. Consideraram a HNR menor que 7,4 dB como patológica, baseados na distribuição da HNR para vozes normais (5% - taxa de falso alarme). Os autores compararam a razão harmônico-ruído

(HNR) com a razão sinal/ruído (S/N) proposta por Kojima et al. (1980) e verificaram que os valores encontrados em dB para (S/N) foram maiores que os encontrados para (HNR) e justificaram esta discrepância devido a diferença metodológica entre os dois estudos. O coeficiente de correlação de Spearman foi 0,849 e significativo com  $p=0,001$ , o que confirmou que a HNR foi um índice útil para avaliar quantitativamente o componente ruidoso relativo ao componente harmônico da vogal, mas afirmaram a necessidade de estudos futuros para confirmar a relação entre a HNR e as medidas psicofísicas do grau de rouquidão.

No estudo de Hiraoka et al. (1984), o tempo de intervalo para análise foi fixado independentemente, sendo a frequência fundamental e a frequência harmônica determinadas automaticamente pela Transformada de Fourier. A avaliação objetiva foi realizada considerando as características de vozes roucas, que mostraram a intensidade da frequência fundamental mais proeminente, comparada com os harmônicos do espectro vocal, conforme afirmação dos autores. A intensidade harmônica relativa (Hr), obtida por meio de uma porção estável da vogal /a/ sustentada, denota a porcentagem das intensidades das componentes harmônicas relativas a intensidade vocal total. Os autores estabeleceram um limite de 67,2% para Hr como discriminador entre vozes normais e patológicas. Das vozes normais examinadas, 95% tiveram a intensidade harmônica maior que o valor crítico de 67,2%, enquanto que 90% das vozes roucas tiveram a intensidade harmônica menor que o valor crítico.

Kasuya et al. (1986), desenvolveram uma medida de energia do ruído normalizado (NNE) que estimou os componentes ruidosos da fonação de uma vogal sustentada e propuseram uma medida acústica para a detecção de patologias laríngeas. O sinal vocal foi filtrado através de um filtro pente adaptativo operando no domínio da frequência. Basearam-se em um sinal de voz com janelamento de 7 períodos de uma

vogal que continha um componente periódico e um componente ruidoso aditivo. Em um ambiente silene, um grupo de 64 sujeitos com vozes normais e outro grupo com 186 sujeitos com patologias laringeas emitiram, por 5 vezes, a vogal japonesa /e/ sustentada em uma frequência confortável e em um nível de intensidade vocal aumentado. Determinou-se o período fundamental médio ( $T_p$ ) pela autocorrelação em janelas de 40 ms (KASUYA et al., 1984). O comprimento da janela de dados a ser analisado foi determinado em  $7 T_p$  e usado para análise espectral sobre todo intervalo. Segundo os autores, o nível de ruído pode ser melhor avaliado na faixa de altas-frequências para a discriminação de vozes normais e patológicas. Os experimentos mostraram que o câncer glótico avançado T2-T4 pode ser perfeitamente detectado como patológico com um erro de 9,4% para sujeitos normais. Para o caso de paralisia do nervo recorrente e nódulos das pregas vocais, os erros são 9,1% e 20%, respectivamente. Em comparação com essas doenças, as razões dos erros foram um pouco altas para pólipos nas pregas vocais e laringites. Câncer avançado supra e subglótico e câncer faringeal também foram discriminados de amostras normais porque eles exibiram invasão direta sobre as pregas vocais. A medida NNE comparada com os parâmetros HNR (YUMOTO et al. 1982) e o Hr (HIRAOKA et al., 1984) mostrou-se mais eficiente para discriminação de vozes patológicas e normais, além de também ser útil para discriminar câncer glótico T2-T4, paralisia do nervo recorrente, nódulo nas pregas vocais de casos normais, apesar de ter encontrado um falso positivo para seis sujeitos com vozes normais, os quais considerou com câncer.

Fukazawa et al. (1988), desenvolveram o índice Br que indica o ruído turbulento em vozes soprosas e testaram sua validade. Basearam-se na razão entre a energia do sinal filtrado por um passa-alta de segunda ordem pela energia do sinal original. Um grupo de 24 sujeitos com vozes normais e um grupo de 31 indivíduos com rouquidão

(10 casos de pólipos na prega vocal, 15 de paralisias do nervo recorrente e 6 de câncer glótico), emitiram a vogal /a/ sustentada diante de um microfone a uma distância de 15 cm da boca. Por meio de um microcomputador a voz foi processada e utilizando uma amostra de comprimento 25,6 ms calculou-se o índice Br. A impressão auditiva para avaliar a soproidade e a rouquidão foi feita por dois experientes médicos de acordo com o método de Isshiki (1964). A correlação entre o índice Br e a percepção de soproidade foi significativa ao mesmo nível que a correlação perceptiva entre os avaliadores. Por outro lado, a correlação entre o índice e a percepção de rouquidão não foi significativa. No grupo de sujeitos com vozes normais realmente não houve presença de soproidade nem de rouquidão, com uma média de 27,0 variando de 8,3 a 75,7. Nos casos de vozes com patologias, por meio da caracterização psicoacústica, a correlação entre soproidade e o índice Br foi 0,734 e a correlação entre rouquidão e o índice Br foi somente 0,211. As vozes patológicas tiveram um índice Br maior do que as normais. Estatisticamente, a média do índice Br dos 31 sujeitos roucos foi maior que dos 24 sujeitos com vozes normais ( $p=0,01$ ). Já a média dos casos com câncer laríngeo foi maior que dos casos com paralisia do nervo recorrente e dos casos com pólipos ( $p=0,05$  e  $p=0,02$ , respectivamente).

Shoji et al. (1992), investigaram um novo método para discriminar vozes normais de soprosas através da relação de potência em altas-freqüências pela potência total do sinal vocal (HFPR), cujo limiar de alta-freqüência ( $F_c$ ) foi estimado em torno de 5 kHz, devido à maior concentração de energia da voz normal estar situada abaixo desta freqüência. Para isso, gravaram a vogal sustentada /a/ de 24 sujeitos com voz soprosa (3 deles em pré e pós cirúrgico) e 16 indivíduos formaram o grupo controle (9 homens e 7 mulheres). Encontraram que a razão de duas variâncias em ANOVA demonstrou que os dois grupos, sujeitos com vozes alteradas e controles, começam a



apresentar faixas de valores diferentes para o índice HPFR a partir de  $F_c$  igual a 4 kHz e a 6 kHz, regiões em que não ocorreram sobreposições. Concluíram que para  $F_c$  (6 kHz) o limite de separabilidade entre vozes normais e soprosas é  $-30$  dB. Além disso, o índice HFPR é, também, uma ferramenta útil para a avaliação de fonocirurgia.

Hillenbrand et al. (1994), mediram a relação entre as avaliações de sopro e uma gama de medidas acústicas. Estas medidas enquadraram-se em 3 categorias: a) medidas da periodicidade do sinal; b) medidas da amplitude do primeiro harmônico e c) medidas de inclinação espectral. A amostra foi constituída por 8 homens e 7 mulheres na faixa etária de 22 a 37 anos de idade sem relato de problemas vocais, de fala ou de audição. Cada falante foi treinado a produzir 3 variações vocais (normal, moderadamente soprosa e muito soprosa), das quais se estimou a média da frequência fundamental. Esta frequência foi calculada a partir da leitura de um trecho do texto “Rainbow Passage” (FAIRBANKS, 1940), processado pelo equipamento Visi-Pitch da Kay Elemetrics/interface IBM PC (HORII, 1983). Incluiu-se, ainda, as vogais [a], [I], [æ] e [o], sustentadas por, aproximadamente, 3 segundos, perfazendo um total de 12 amostras vocálicas (4 vogais x 3 variações vocais). Vinte ouvintes, estudantes de graduação de Patologias da Fala-Linguagem (Universidade de Michigan), foram recrutados para avaliar as amostras vocálicas, após terem realizado audiometrias (frequências 0,5; 1; 2; 4 e 6 kHz). Na análise acústica utilizou-se a proeminência de pulso cepstral (CPP), que é a distância do pico à reta de regressão linear normalizado para todas as amplitudes; o Pearson  $r$  no pico de autocorrelação ( RPK); a relação entre a amplitude de pico pela amplitude média do valor absoluto do sinal (P/A); o índice de sopro (BRI), que consiste na média do Br em janelas de 25,6 ms com passo de 10 ms do sinal; a razão da média da energia espectral na faixa menor que 4 kHz e maior/igual que 4 kHz (H/L), que é uma modificação do HFPR, com o limiar de

freqüência em 4 kHz, apresentando a relação de energias superior e inferior a este limite; e a relação entre o 1º. harmônico e o 2º. harmônico (H1A), com unidades em dB.

Os resultados encontrados pelos autores demonstraram que os parâmetros CPP e RPK são os mais consistentes para quantificar soproidade. Os índices BRI, H/L e H1A apresentaram fraca correlação com as medidas de soproidade, para confirmação dos achados os autores fizeram modificações nas operações matemáticas não encontrando resultados confiáveis para esses índices.

#### 4 METODOLOGIA

Cento e trinta indivíduos, de ambos os sexos, na faixa etária de 21 a 45 anos de idade colaboraram para este estudo com suas vozes. Do total de voluntários, 77 indivíduos eram do sexo feminino, sendo os 53 restantes do sexo masculino.

A seleção destes indivíduos ocorreu aleatoriamente, à exceção da faixa etária, que teve a idade mínima estabelecida pelo estatuto do Comitê de Ética do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (HC/FMRP/USP) e a idade máxima para se evitar vozes em fase de senilidade.

Após assinatura do termo livre e esclarecido (ANEXO A) para participação neste estudo, devidamente aprovado pelo comitê referido acima (ofício nº 2259/2002 ), os indivíduos foram conduzidos a uma entrevista com a pesquisadora para preenchimento de uma ficha de identificação contendo dados pessoais como: nome, sexo, idade, data de nascimento e endereço.

A partir daí, o indivíduo foi orientado sobre o processo de registro da voz, realizando um treinamento prévio, evitando assim, possíveis influências psicológicas e, conseqüentemente, alterações vocais não existentes. Neste treinamento, foi solicitado que cada voluntário emitisse a vogal sustentada /a/ amostra vocal a ser coletada para este estudo em níveis de altura e intensidade confortáveis, a fim de que não ocorresse tensão excessiva sobre pregas vocais .

Depois do treinamento, a gravação das amostras vocais dos voluntários foi realizada diretamente em computador, em sala tratada acusticamente – Laboratório de Voz do HC/FMRP/USP – onde o voluntário foi posicionado sentado, com o microfone a 10 cm à frente de sua boca e solicitado a emitir a vogal sustentada /a/ por 5 segundos. O controle da duração da emissão estava previamente padronizado pelo programa de captação/análise do sinal de voz: *Análise de Voz*, versão 2.0, da DSP Instrumentos Ltda.

O computador foi um Pentium de 100 MHz, 16 Mb de memória e 1 Gb de winchester. A conversão analógico-digital foi realizada através da placa de aquisição de dados convencional, marca *Soundblaster* SB16, sendo a taxa de amostragem utilizada de 22050 Hz com resolução de 16 bits.

O microfone utilizado foi do tipo cardióide, marca LeSon, modelo SM58, unidirecional e com ganho dinâmico, segundo estudos de Rosa (1998).

A mesma amostra vocálica serviu simultaneamente para análise acústica computadorizada e para avaliação perceptivo-auditiva da qualidade vocal, realizada pela pesquisadora.

A avaliação perceptivo-auditiva da qualidade vocal consistiu na seguinte classificação: voz neutra para vozes sem alterações na qualidade vocal enquanto que para vozes com qualidade vocal alterada seguir-se Behlau e Pontes (1995).

As amostras vocais armazenadas no computador foram transferidas para um cd e levadas para o Laboratório de Instrumentação e Microeletrônica do Departamento de Engenharia Elétrica da Escola de Engenharia de São Carlos da Universidade de São Paulo (LIM/SEL/EESC/USP). Neste local, realizou-se a análise acústica computadorizada, por meio do programa *Análise de Voz* já referido anteriormente.

Na análise, cada amostra vocal do /a/ sustentado teve seus trechos inicial e final desprezados, conforme estudos de Krom (1995), considerando-se apenas o terço médio da emissão. Este teve a duração de 30 janelas de 512 pontos, o qual corresponde a 696,6 ms. Essa duração de tempo é decorrente da frequência de amostragem usada (22.050).

De acordo com este programa de análise vocal, determinou-se os seguintes parâmetros sobre o sinal vocal: *jitter (em %)*, *shimmer (em %)*, *suavidade espectral do filtro (SFF, em dB)*, *suavidade espectral do sinal residual (SFR, em dB)*, *índice de vocalização (PA em valor absoluto)* e *coeficiente de excesso (EX, em valor absoluto)*.

Todos os voluntários passaram, ainda, por uma inspeção de videoestroboscopia no HC/FMRP/USP realizada por profissional habilitado. Durante esta inspeção, solicitou-se que cada um dos voluntários emitisse a vogal sustentada /i/ durante o tempo que o profissional julgasse necessário para esta avaliação. Ainda este profissional, no final desta avaliação, fornecia o diagnóstico das ppvv avaliadas para o voluntário ao mesmo tempo que confrontava seus achados clínicos com a avaliação perceptivo-auditiva, realizada anteriormente.

A pesquisadora anotou, então, junto à sua avaliação perceptivo-auditiva, o resultado clínico da inspeção videoestroboscópica para comparações posteriores.

## 5 RESULTADOS E DISCUSSÃO

Na avaliação perceptivo-auditiva nenhuma das amostras vocais apresentou qualidade vocal alterada, sendo todas, então, classificadas como neutras.

Pela clínica, através da videoestroboscopia, houve em todas as amostras vocais a confirmação diagnóstica de ausência de alterações vocais. Ou seja, todas as pregas vocais apresentaram adequada mobilidade, com ciclo glótico satisfatório. Pôde-se, ainda, observar que nos aspectos físicos, as pregas vocais tinham coloração também adequada, sem alterações de massa.

Na avaliação objetiva dos parâmetros contidos no programa utilizado obteve-se os seguintes resultados:

Frequência fundamental ( $F_0$ )

O histograma da figura 5 mostra a frequência relativa de ocorrências para a  $F_0$  dos voluntários masculinos. A curva parece com a normal (gaussiana) o que corrobora com o tamanho da amostra mostrando-a suficiente para este estudo. A análise estatística dos parâmetros acústicos obtidos estão na tabela 1 apresentada posteriormente.

Aproximadamente 70% do total das amostras do sexo masculino concentrou-se entre 110 e 146,7 Hz, sendo o menor valor encontrado neste estudo foi em 95,6 Hz e o maior valor em 175,9 Hz.

Nota-se neste estudo que houve uma diferença aproximadamente de 28% referente ao valor médio citado por Bhelau e Pontes (1995), o valor médio resultante encontrado foi de 128,65 Hz com desvio padrão de 20,75 Hz.

Algo que chama a atenção é a distribuição mais dispersa do pitch masculino em relação ao feminino (figura 6). Não se encontra na literatura estudos a respeito de tal ocorrência. Considerando-se esta amostra representativa da população, este fato pode requerer maiores investigações.

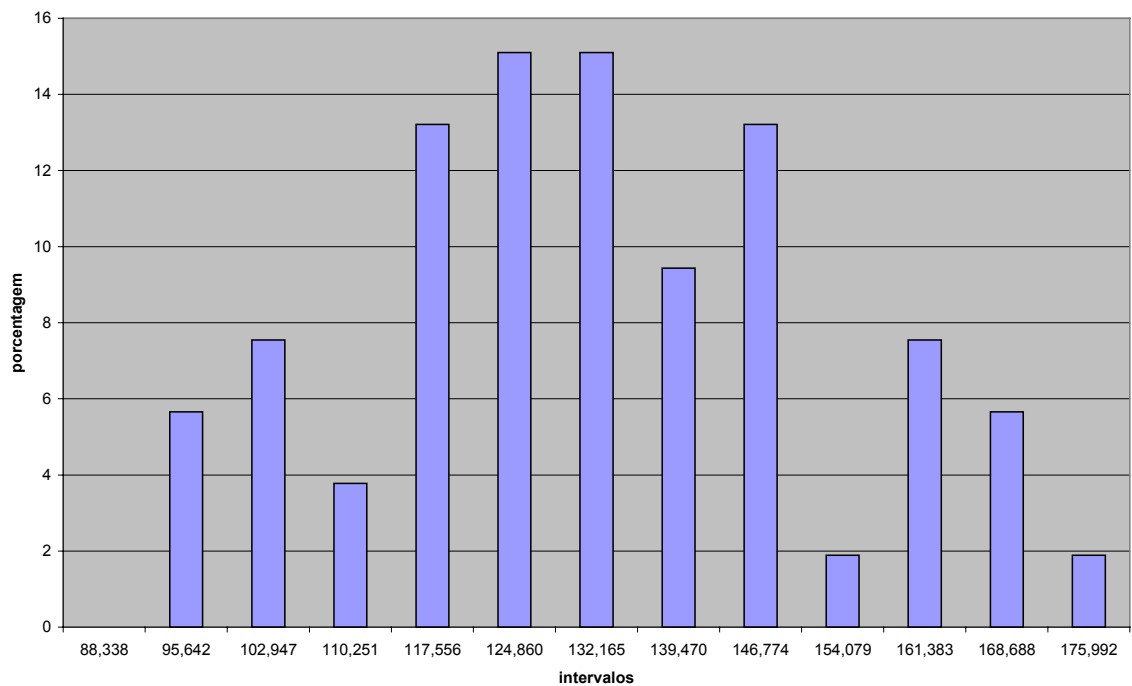


Figura 5: Histograma da frequência fundamental do sexo masculino

O valor médio encontrado para a frequência fundamental dos voluntários femininos é de 209,61 Hz ( figura 6). Vê-se a forte concentração em torno da média em contraste para o grupo masculino.

Aproximadamente 46% do total das amostras concentrou-se em 203,5 Hz, tal dado aproxima-se dos achados de Behlau e Pontes (1995), que discorrem que a frequência fundamental do sexo feminino situa-se ao redor de 200 Hz.

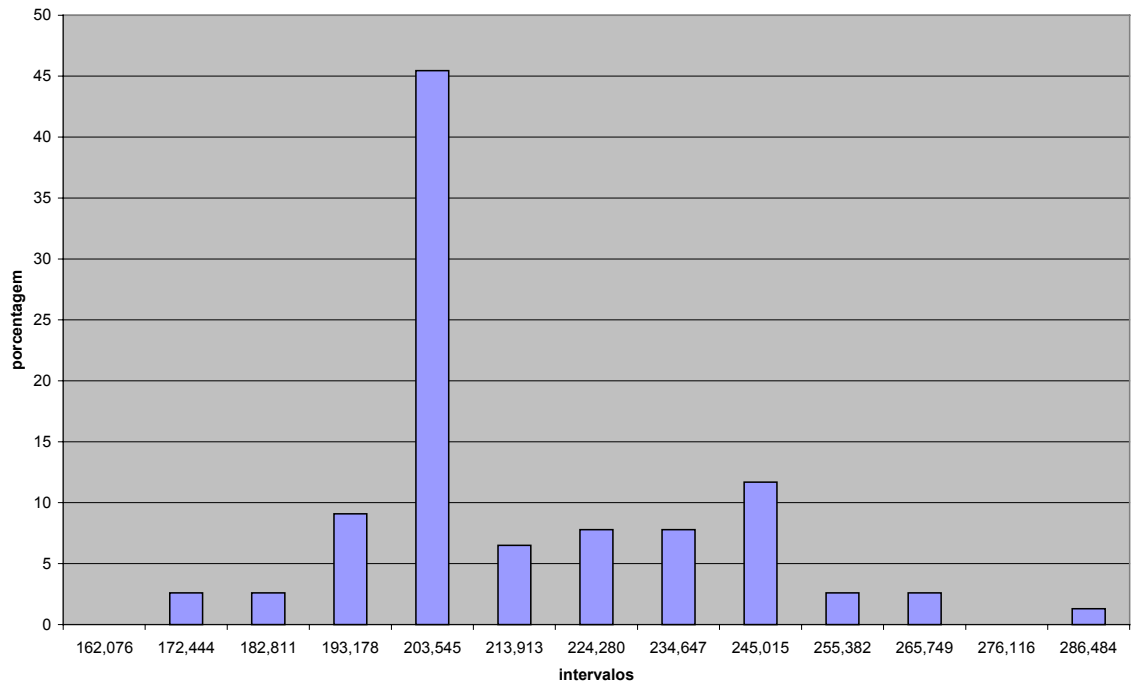


Figura 6: Histograma da frequência fundamental do sexo feminino

#### Perturbação de período de pitch (*Jitter*)

Outro parâmetro avaliado foi a perturbação de período de pitch. A figura 7 mostra a distribuição deste parâmetro. Percebe-se através deste histograma a forte concentração para valores pequenos, aproximadamente 63% dos voluntários encontram-se abaixo de 3%. O valor médio obtido para o jitter, considerando-se toda amostra vocal, é 1,99% com desvio padrão 2,16%.

Apenas para exercício de comparação e reportando-se aos achados de Lieberman para PFF, que considera voz normal para variações de período até 0,5 ms, neste estudo, para vozes masculinas, a frequência média é de 128,65 Hz fornecendo um período de 7,77 ms. Portanto, 0,5 ms proposto por Lieberman representa 6,43% de variação do período (jitter). Para as vozes femininas os valores são 209,61 Hz e 4,77 ms para pitch e período, respectivamente. O jitter teórico resultante é 10,48%.

Pelos achados neste estudo verifica-se que o limite do parâmetro proposto por Lieberman é conservador.



Para o Jitter, Davis encontrou em seus estudos os valores de 0,99% para a média e desvio padrão 2,28%. Independentemente das amostras vocais selecionadas, os resultados estão bastante próximos daquele estudo.

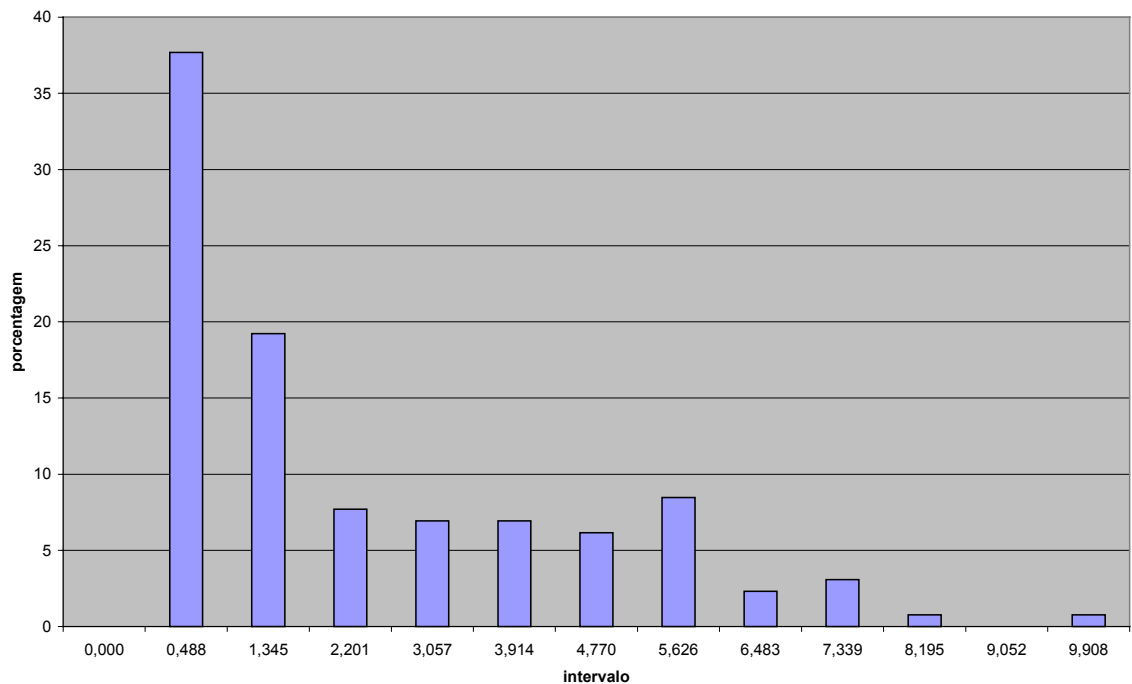


Figura 7: Histograma do jitter

#### Perturbação de amplitude (*Shimmer*)

O histograma relativo a este parâmetro está mostrado na figura 8. Nota-se a concentração em torno do valor médio 9,17%, sendo seu desvio padrão 3,17%. Estes valores estão em consonância com Davis, que encontrou os valores 7,22% e 4,47% respectivamente.

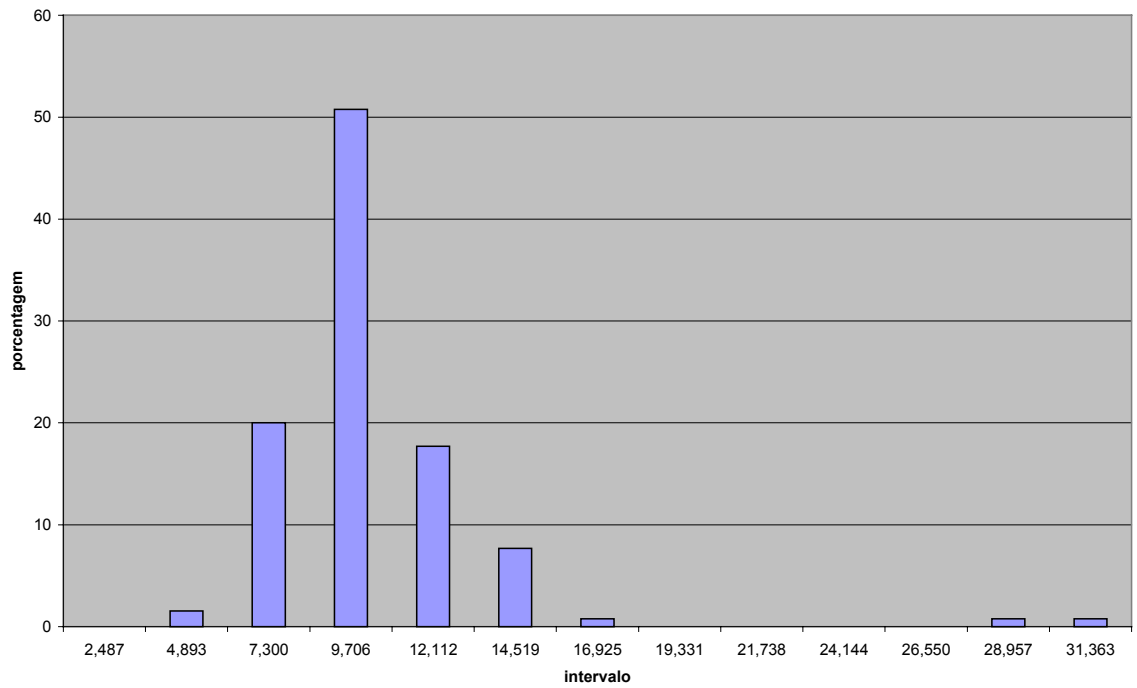


Figura 8: Histograma do shimmer

#### Suavidade espectral do resíduo (*SFR*)

Em relação ao parâmetro de suavidade espectral do resíduo, que em última análise avalia o espectro do sinal residual, os valores obtidos assemelham-se favoravelmente com os dados de Davis. Os valores encontrados neste trabalho são - 10,91 dB e 2,13 de desvio padrão, enquanto os dados de Davis são: - 10,50 dB e 2,50. A distribuição dos valores pode ser visualizada na figura 9.

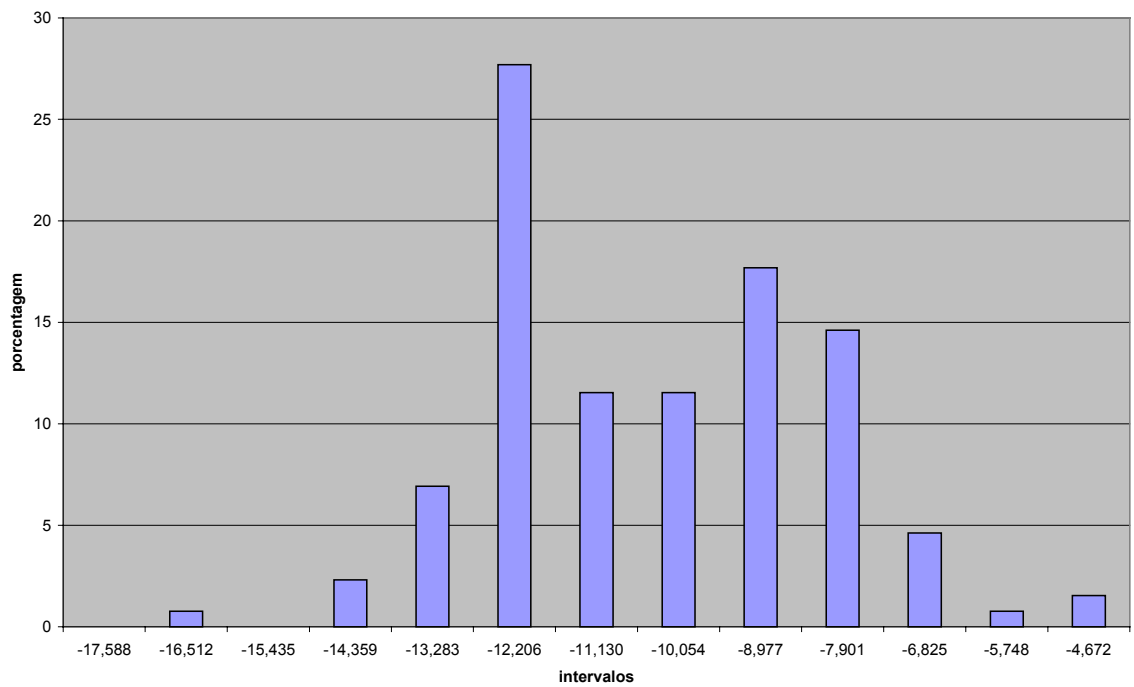


Figura 9: Histograma do SFR

#### Suavidade espectral do filtro (*SFF*)

A suavidade espectral do filtro está intrinsecamente relacionada ao trato vocal. Sua avaliação estabelece numericamente as características formantes. Neste estudo os valores encontrados foram -12,73 dB para o valor médio com desvio padrão de 3,41. Para termos de comparação os valores obtidos por Davis são -11,85 dB e 1,84 respectivamente.

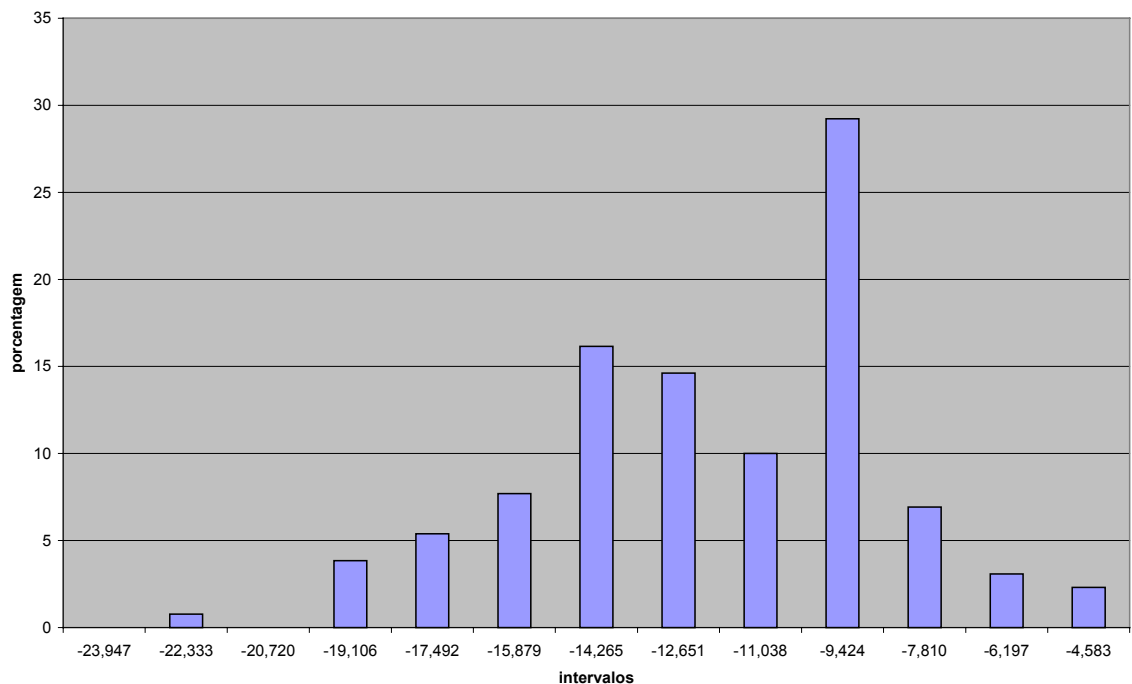


Figura 10: Histograma do SFF

### Índice de Vocalização ( $P. A$ )

Este parâmetro avalia a qualidade vocal realçando a “nitidez” da vocalização. Para um som fricativo surdo, por exemplo, o valor do  $P. A$  é zero. Já para uma vogal perfeita, o valor do  $P. A$  é a unidade. Os valores intermediários, portanto, estabelecem uma escala de qualidade vocal. A distribuição deste parâmetro pode ser visto na figura 11. Nota-se deste histograma que 96% dos voluntários apresentam  $P.A$  maior que 0,4. O valor unitário é atribuído a voz perfeita.

Há que se ressaltar, ainda, os erros de processamento tais como, quantização do dado (decorre do processamento de digitalização do sinal analógico, podendo ser considerado uma fonte de ruído), indução eletromagnética (captação de campos eletromagnéticos pela instrumentação utilizada), ruído ambiente (todas as fontes sonoras, exceto o sinal a ser captado, presentes no ambiente), não robustez do algoritmo detetor do pitch.

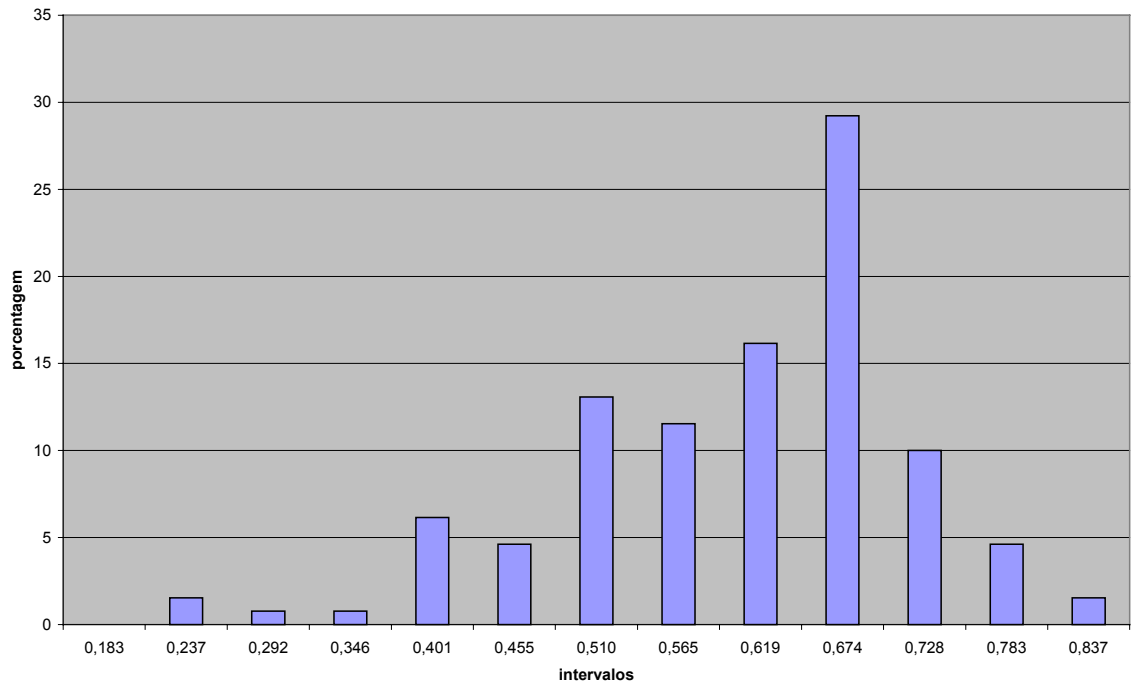


Figura 11: Histograma do P.A

Para efeito de ilustração as figuras 12 e 13 apresentam o cálculo da correlação do resíduo para vozes normal e patológica respectivamente.

Note a maior regularidade dos picos para a voz normal, embora haja ainda grande componente ruidosa resultante de erros de processamento, como os já citados. Para a voz patológica a estrutura regular é mascarada pela maior quantidade de ruído fornecendo, deste modo, um menor valor de P.A.

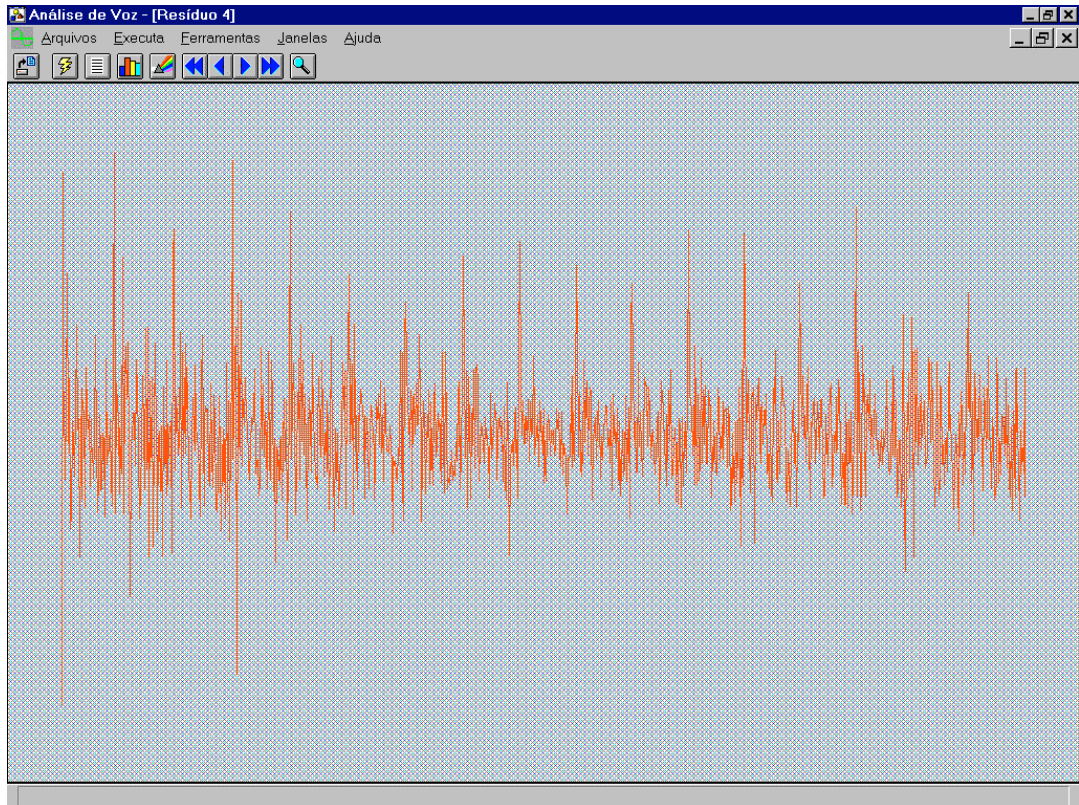


Figura 12: Resíduo de voz normal para o PA

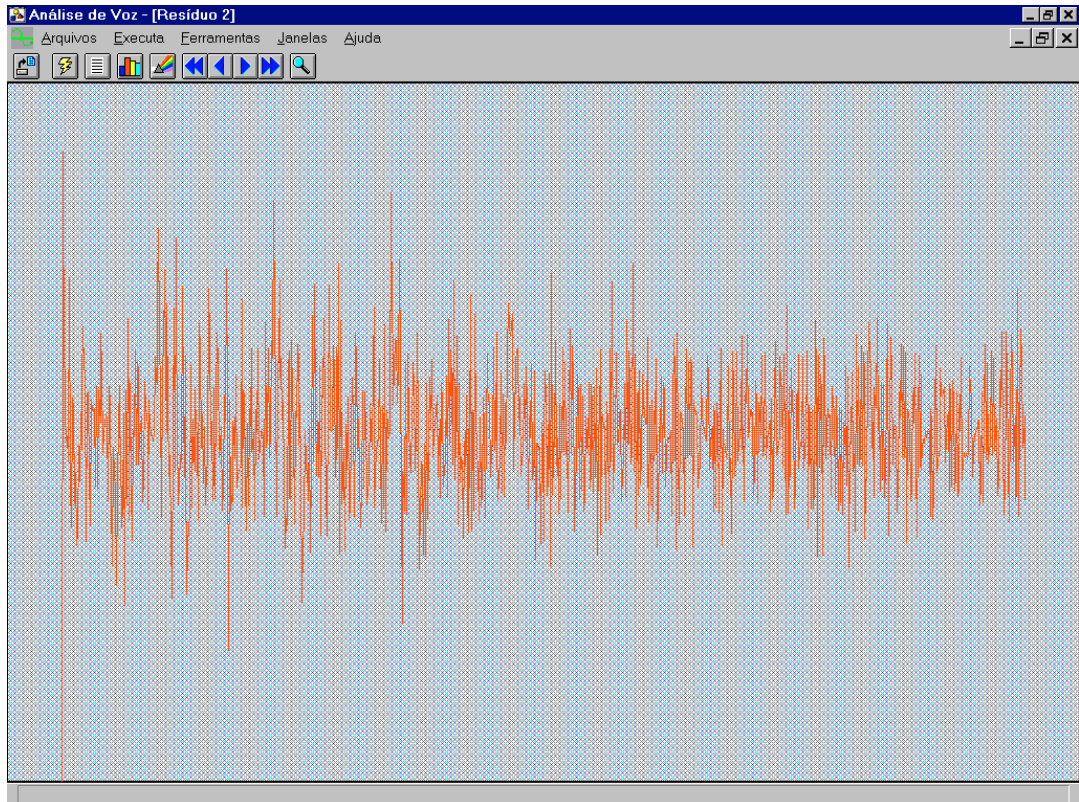


Figura 13: Resíduo de voz patológica para o PA

### Coeficiente de Excesso ( $EX$ )

Este parâmetro é um correlato da relação sinal-ruído. Portanto, quanto maior seu valor melhor será a qualidade vocal. Os valores obtidos por Davis são 5,17 para a média e 4,29 de desvio padrão. Do gráfico apresentado na figura 14 vemos que todos os voluntários apresentam valores de  $EX$  maiores que 5, podendo ser este o limite entre normalidade e patológico.

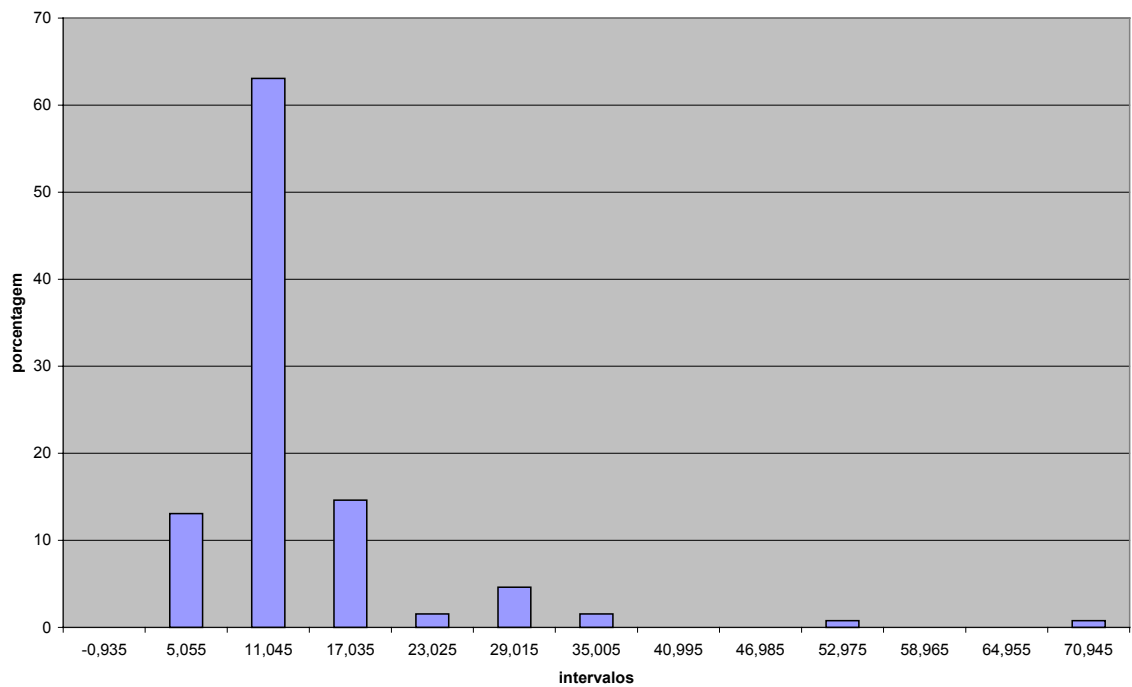


Figura 14: Histograma do EX

Sumarizando, elaborou-se uma análise estatística dos parâmetros acústicos avaliando a simetria e a curtose, cujos valores são apresentados na tabela 1. Apresenta-se também um quadro comparativo entre os valores obtidos por Davis e os encontrados neste estudo (tabela 2).

	SIMETRIA	CURTOSE
Jitter	11,5541	72,5008
Shimmer	113,1123	2276,9625
SFR	1,7130	55,9840
SFF	-14,1819	380,5350
PA	-0,0012	0,0006
EX	350,5444	7499,6247

Tabela 1: tabela com análise estatística

	DAVIS	ANDRADE
Jitter	0,99 ± 2,28	1,99 ± 2,16
Shimmer	7,22 ± 4,47	9,17 ± 3,17
SFR	-10,50 ± 2,50	-10,91 ± 2,13
SFF	-11,85 ± 1,84	-12,73 ± 3,41
PA	0,725 ± 0,105	0,579 ± 0,116
EX	5,17 ± 4,29	9,50 ± 5,60

Tabela 2: tabela comparativa de resultados

Considerando a especificidade e as peculiaridades da população envolvida neste estudo, vê-se que os achados, apesar de certas diferenças quantificáveis dos encontrados por Davis, são condizentes com seu estudo.



## 6 CONCLUSÕES

O comportamento gaussiano das distribuições encontradas para todos os parâmetros avaliados justifica a amostragem realizada. Isto é, uma amostra de 130 voluntários é representativa da população regional.

O valor da média encontrado neste estudo relativo a frequência fundamental do sexo feminino encontra-se em concordância com os achados de Behlau e Pontes, no entanto o mesmo não ocorreu com os achados relativos à frequência fundamental do sexo masculino.

O critério adotado para estabelecimento dos limiares foi o valor que abrange pelo menos 70% da amostra. Assim, os limiares propostos como padronizadores de normalidade e as porcentagens associadas são: jitter 3% (70,8%), shimmer 10% (75,4%), SFR – 9 dB (76,9%) SFF – 10,28 dB (72,3%), P.A 0,52 (73%) e EX 6,5 (73,8%).

O programa utilizado neste trabalho comportou-se de maneira promissora, visto que não apresentou nenhum falso positivo, fato que corroborou com os achados clínicos e perceptivos.

Não houve investigação relacionada às vozes patológicas devido ao não aparecimento de tais amostras durante este estudo, apesar da aleatoriedade dos voluntários. Tal fato descartou a averiguação de um provável falso negativo. Fica proposto para trabalho futuro, o estudo comportamental da distribuição destes dados.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, L. M. O., VIEIRA, J. M., RAZERA, D. E., GUERRA, A. C., PEREIRA, J. C. **Medidas de Perturbação da voz: um novo enfoque.** Revista Fonoaudiologia-Brasil, Brasília-DF, v.2, n.2, p.39-46, 2002.

ARONSON, A.E. **Clinical Voice Disorders.** New York, Decker, 1980.

ARONSON, A.E. **Clinical Voice Disorders.** 3<sup>rd</sup> ed. New York: Thieme, 1990.

BEHLAU, M.& PONTES, P. **Avaliação e Tratamento das Disfonias.** São Paulo: Lovise, 1995.

BEHLAU, M. & REHDER, M.I. **Higiene Vocal para o Canto Coral.** Rio de Janeiro: Revinter, 1997.

BEHLAU, M. (org). **Voz “O Livro do Especialista”.** Rio de Janeiro: Revinter,2001.

BOONE, D. & MCFARLANE, S. **The Voice and the Voice Therapy.** 4 ed. Englewood Cliffs, Prentice-Hall, 1988.

BRANDI, E. **Voz Falada: Estudo, Avaliação e Tratamento.** Rio de Janeiro: Atheneu, 1990.

DANGELO, J.G. & FATTINI, C.A. **Anatomia Humana Sistêmica e Segmentar.** 2ed, Atheneu, 1995.

DAVIS, S.B. **Acoustic Characteristics of Normal and Pathological Voices.** In Lass, N.J. (ED) *Speech and language: Advances in Basic Research and Practice*, vol.1, New York Academic Press, 271-335, 1979.

FROKJAER-JENSEN, B. & PRYTZ, S. **Registration of Voice Quality,** B & K Technical Review, 3-17, 1976.

FUKAZAWA, T. EL-ASUOOTY, A E HONJO, I. **A New Index for Evaluation of the Turbulent Noise in Pathological Voice.** Journal of the Acoustical Society of America, 83, 1189-1193, 1988.

GARCIA-TABIA, R., COBETA, I. **Clasificación de Las Disfonías. In los Transtornos de la Voz.** Madrid: Garsi, 1996.

HILLENBRAND, J., CLEVELAND, R. A & ERICKSON, R. L. **Acoustic Correlates of Breathy Vocal Quality**. Journal of Speech and Hearing Research, 37, 769-778, 1994.

HIRAOKA, N., KITAZOE, Y., UETA, H., TANAKA, S. E TANABE, M. **Harmonic-Intensity analysis of Normal and Hoarse Voices**. Journal of the Acoustical Society of America, 76, 1648-1651, 1984.

HIRANO, M. **Structure and Vibratory Behaviour of the Vocal Folds**. In M. Saswashima and F.S. Cooper (Eds.), Dynamic Aspects of Speech Production. Tokio. University of Tokio Press. 1977.

HOLMES, J. N. **Low Frequency phase Distortion of Speech Recording**. The Journal of the Acoustical Society of America, 747-749, 1975.

KASUYA, H., OGAWA, S., MASHIMA, K. EBIHARA, S. **Normalized Noise Energy as an Acoustic Measure to Evaluate Pathologic Voice**. The journal of the Acoustical Society of America, 80, 1329-1334, 1986.

KOIKE, Y. **Applications of Some Acoustic Measures for the Evaluation of Laryngeal Dysfunction**. Studia Phonologica, 7, 17-23, 1973.

KOJIMA, H., GOULD, W.J., LAMBIASE, A. E ISSHIKI, N. **Computer Analysis of Hoarseness**, Acta Otolaryngologica, 89, 547-554, 1980.

KROM, G. **Consistency and Reability of Voice Quality Ratings for Different Types of Speech Fragments**. Journal of Speech and Hearing Research, v.37, 985-1000, 1994.

KROM, G. **Some Spectral Correlates of Pathological Breathy and Rough Voice Quality for Different Types of Vowel Fragments**. Journal of Speech and Hearing Research, v.38, 794-811, 1995.

LIEBERMAN, P. **Some acoustics Measures of the Fundamental Periodicity of Normal and Pathologic Larynges**, Journal of the Acoustics Society of America, 35, 344-353, 1963.

NIETO, A., DEL PALACIO, A. J., LORENZO, A. J., VEGAS, A., COBETA, I. **Los Ordenadores en el Análisis de la Voz: Aplicaciones Clínicas**. Acta Otorrinolaringologica Española. V.46, n.4, 241-245, 1955.

PINHO, S. **Avaliação e Tratamento da Voz**. In Pinho (org): **Fundamentos em Fonoaudiologia. Tratando os Distúrbios da Voz**. Rio de Janeiro: Guanabara-Koogan, 2-48, 1998.

ROSA, M.O. **Análise Acústica da Voz para Pré-Diagnóstico de Patologias da Laringe**. Dissertação (mestrado) – Escola de Engenharia-Universidade de São Paulo. São Carlos, 1998.

ROSA, M.O., PEREIRA, J.C., GRELLET, M. **Signal Processing and Statistical Procedures Identify Laryngeal Pathologies**. 6<sup>th</sup> IEEE International Conference on

Electronics, Circuits and Systems (ICECS'99) – Advanced Program. Paphos, Chipre, Set. 1999.

SHOJI, K., REGENBOGEN, E. YU, J.D., BLAUGRUND, S.M. **High-frequency Power Ratio of Breathly Voice**. Laryngoscope, 102, 267-271, 1992.

TABITH, A.JR. **Foniatria: Disfonias, Fissuras, Paralisia Cerebral**. 7ed. São Paulo, Cortez, 1995.

TAKAHASHI, H. E KOIKE, Y. **Some Perceptual Dimensions and Acoustical Correlates of Pathologic Voices**. Acta Oto-laryngologica, Suppl.338, 1-24, 1975.

VONLEDEN, H., MOORE, G. P. **The Mechanics of the Cricoarytenoid Joing**. Arch Otolaryngol. 73: 541-550. 1961.

ZEMPLIN, W. R. **Speech and Hering Science: Anatomy and Physiology**. 3ed. Englewood Cliffs, WJ: Prentice Hall, 1998.

YUMOTO, E., GOULD, W. J. E BAER, T. **Harmonics-to-Noise Ratio as na Index of the Degree of Hoarseness**. Journal of the Acoustical Society of America, 71, 1544-1549, 1982.

## **ANEXO**

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

Campus Universitário Monte Alegre – Fone: 602.1000  
CEP 14048-900 – Ribeirão Preto – São Paulo

**TERMO DE CONSENTIMENTO PÓS-INFORMAÇÃO**

**NOME DA PESQUISA:**

“Determinação dos Limiares de Normalidade dos Parâmetros Acústicos da Voz”

**PESQUISADOR RESPONSÁVEL:**

Luciana Mara de Oliveira Andrade CRFa: 9959

**ORIENTADOR RESPONSÁVEL:**

Marcos Grellet CRM:10.069

**INFORMAÇÕES AOS PACIENTES:**

1. Os sujeitos que possuem problemas de laringe, podem apresentar alterações na voz, durante a gravação da mesma e posterior análise no computador, fato que necessita de averiguação para um melhor conhecimento.
2. A pesquisa a ser realizada, tem a finalidade de investigar a eficácia de um programa de computador para voz, através da gravação da mesma, via microfone/computador e exame otorrinolaringológico denominado videoestroboscopia, que consiste em filmar a laringe/garganta. A filmagem será realizada, utilizando-se um instrumento metálico, que será introduzido dentro da boca, não provocando nenhuma dor.
3. Os procedimentos não envolvem qualquer risco e, com o conhecimento adquirido através da pesquisa, poder-se-á verificar se existe possibilidade de pré-diagnosticar problemas de voz com o uso de programa de voz computadorizado.

**EU, \_\_\_\_\_**  
**RG no : \_\_\_\_\_, abaixo assinado, tendo recebido**  
**as informações acima, e ciente dos meus direitos abaixo relacionados, concordo**  
**em participar.**

1. A garantia de receber a resposta a qualquer pergunta ou esclarecimento a qualquer dúvida a cerca dos procedimentos, riscos, benefícios e outros relacionados com a pesquisa;
2. A liberdade de retirar meu consentimento a qualquer momento e deixar de participar no estudo sem que isso traga prejuízo à minha pessoa;
3. A segurança de que não serei identificado e que será mantido o caráter confidencial da informação relacionada com a minha privacidade;
4. O compromisso de me proporcionar informação atualizada durante o estudo, ainda que esta possa afetar minha vontade de continuar participando;
5. O compromisso de que serei devidamente acompanhado e assistido durante todo o período de minha participação na pesquisa;
6. Que se existirem gastos adicionais estes serão absorvidos pelo orçamento da pesquisa,

**Tenho ciência do exposto acima e desejo participar desta pesquisa sob orientação da fonoaudióloga que subscreve este documento.**

Ribeirão Preto, \_\_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

---

Assinatura do paciente